UNIVERSITÉ DE MONTRÉAL

ON-LINE STEADY-STATE DATA RECONCILIATION FOR ADVANCED COST ANALYSIS IN THE PULP AND PAPER INDUSTRY

MILAN KORBEL

DÉPARTEMENT DE GÉNIE CHIMIQUE

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION

DU DIPLÔME DE PHILOSOPHIAE DOCTOR

(GÉNIE CHIMIQUE)

DÉCEMBRE 2011

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

ON-LINE STEADY-STATE DATA RECONCILIATION FOR ADVANCED COST ANALYSIS IN THE PULP AND PAPER INDUSTRY

présentée par : KORBEL Milan

en vue de l'obtention du diplôme de : Philosophiae Doctor

a été dûment accepté par le jury d'examen constitué de :

M. PERRIER Michel, Ph.D., président

M. STUART Paul, Ph.D., membre et directeur de recherche

M. FRADETTE Louis, Ph.D., membre

M. GUNN Eldon, Ph.D., membre

DEDICATION

To whom I owe everything

"Since all measurements and observations are nothing more than approximations to the truth, the same must be true of all calculations resting upon them, and the highest aim of all computations made concerning concrete phenomena must be to approximate, as nearly as practicable, to the truth. But this can be accomplished in no other way than by a suitable combination of more observations than the number absolutely requisite for the determination of the unknown quantities."

Gauss, K. G. *Theory of Motion of Heavenly Bodies*, New York, Dover, 1963, p. 249.

ACKNOWLEDGEMENTS

First of all, I would like to thank Professor Paul Stuart for all his valuable advices, guidance and tireless support during these years. Jogging together and speaking about science and hypothesis will always make me remember and smile. I am very grateful for all your endless support in many difficult times.

I would like to thank Taiwen Jiang and Larry Wasik for their tremendous help with understanding many important and difficult issues concerning my project. A very big thank you goes out to Bernard Lupien and Jean Bigelow for their invaluable assistance and advices.

I am indebted to my many of my fellow students and colleagues to support me and to endlessly discuss our various project topics. This was special and enriching life experience learning the wide spectrum of our research domains and giving me with many great friendships. Special thank goes to Louis Patrick, Jean-Christophe, Jose, Matty, Virginie Martin and Jean-Martin for their kind friendships and our close collaboration. And of course, I would like to specially thank Agnès who solve many of my administrative problems during these years. There are so important and so easily forgotten by me.

A very special thank you goes out to my "mama Québécois", who always supported me during all these happy years but also when things were not going as they should. Thank you Catherine, you are a real friend.

A samozrejme, že ďakujem mojím rodičom Ĺudmile a Milanovi, a taktiež mojej milovaanej sestričke Monika doma na Slovensku, že mi uľahčili možnosť isť do Kanady a spraviť si doktorát tisícky kilometrov vzdialený od nich. Ďakujem za celé tieto roky podpory a pochopenia.

And last but not least big thank you goes to Agathe, whose love, trust, honesty and support were always the deepest source of my inspirations.

RÉSUMÉ

L'industrie nord-américaine des pâtes et papiers fait présentement face à plusieurs défis pour survivre dans le contexte actuel. Dans cette optique, le fait de pouvoir comprendre les marges de chacun des produits devient indispensable afin de déterminer un prix de vente optimal et de montrer la vraie rentabilité de la production. Cependant, les systèmes et les pratiques de comptabilité actuels basés sur des recettes prédéterminées ne fournissent qu'une estimation *ad-hoc* de ces valeurs, et ne peuvent alors seulement servir que de point de repère pour l'évaluation de la performance. Par ailleurs, l'implantation des systèmes de gestion de l'information dans les entreprises papetières a permis une meilleure compréhension de la dynamique des procédés et des affaires. Ceci a d'ailleurs permis le développement de méthodes avancées qui les assistent dans le contrôle des coûts et donc dans l'amélioration de la rentabilité. Ces systèmes sont d'une importance capitale pour les producteurs de produits de commodité tels que les usines de papier journal, où de faibles coûts de production sont essentiels pour la survie des entreprises.

Cette thèse a pour objectif de développer une méthodologie permettant une analyse en ligne des coûts manufacturiers pour l'évaluation des coûts marginaux réels, et d'utiliser cette information pour la prise de décision au niveau tactique et stratégique. Cette méthode utilise des données de procédés en temps réel et de coûts provenant du système de gestion de l'information de l'usine. De plus, l'information obtenue peut être exploitée au niveau stratégique de prise de décision afin d'évaluer les impacts des coûts de procédé de diverses alternatives de projets de rétro-installation. Cette méthodologie comprend trois étapes principales. Lors de la première étape, une technique de traitement de signaux, basée sur la transformation multiéchelle d'ondelettes et leur filtrage, est appliquée afin d'analyser simultanément chaque segment du réseau d'instrumentation de l'usine. Cette étape nettoie les données de mesures du bruit à haute fréquence et des anomalies, et cherche à identifier les périodes où le procédé s'approche d'un régime permanent. La seconde étape améliore davantage la qualité des données de procédés en comparant l'ensemble des variables de l'usine à un modèle fondamental de procédé. Cette information sur les procédés est mise à jour par coaptation et correction des mesures biaisées. Troisièmement, cette information opérationnelle est intégrée aux données financières dans un modèle de coûts axé sur les opérations afin de calculer et d'analyser les coûts de production de différents régimes

d'opération. Cette méthodologie a été appliquée à une étude de cas considérant une usine de papier journal et une implantation potentielle du bioraffinage en rétro-installation.

Il a été constaté que le fait d'utiliser une combinaison de la technique des ondelettes avec les techniques de réconciliation de données classiques apportait plusieurs avantages au niveau de l'usine. D'abord, une implantation en ligne de cette technique a été en mesure de fournir un nombre important d'ensembles de données extraites du système de gestion de l'information de l'entreprise. Ces ensembles ont pu ensuite être utilisés pour représenter les opérations en régime permanent. Ils ont aussi fourni une base statistique suffisante pour caractériser la production selon différents régimes d'opérations. Ce faisant, la méthodologie a permis d'améliorer la qualité des données et d'identifier les endroits où l'incertitude des mesures était importantes. Pour le cas relativement simple de l'usine de papier journal, la technique a été en mesure d'obtenir plusieurs ensembles de données représentant les opérations avec un niveau de précision acceptable. Par ailleurs, la combinaison de cette implantation en ligne avec la méthode de réconciliation de données a permis d'améliorer la performance du système d'instrumentation en identifiant les appareils présentant des erreurs systématiques. De plus, il a été montré que si cette technique était implantée à l'usine pour une utilisation quotidienne, elle permettrait d'identifier les mesures erronées en temps réel, améliorant significativement l'instrumentation et le diagnostic des anomalies de procédés. Par ailleurs, l'évaluation des coûts manufacturiers des différents régimes d'opération a fourni de nouvelles façons de visualiser la structure de coûts de l'usine, permettant ainsi d'interpréter de façon transparente les coûts de procédé. À titre d'exemple, dans l'application de ce cadre méthodologique à l'étude de cas de l'usine papetière, les régimes d'opération les plus rentables et les plus coûteux pour la production d'un même grade de papier ont pu être identifiés. La caractérisation et l'interprétation des variances de coûts entre les différents régimes d'opération ont aussi permis d'identifier divers problèmes de production. L'application de cette méthodologie pour l'évaluation des coûts manufacturiers de scenarios de rétro-installation a montré la capacité de cette méthodologie pour utiliser les informations opérationnelles basées sur les régimes afin d'améliorer la prise de décision au niveau stratégique de l'usine. Par exemple, il a été montré que la rentabilité opérationnelle de nouvelles lignes de production intégrées dépend fortement de chacun des régimes d'opérations actuels de l'usine papetière. Les différences entre chacun de ces régimes peuvent ainsi être visible d'une perspective de procédé et permettent l'évaluation des marges des futurs produits et combinaisons de produits. Entre autres, ces informations sur les différents régimes d'opérations permettraient d'améliorer la rentabilité de futures bioraffineries en fournissant l'information nécessaire pour utiliser de façon optimale la flexibilité des procédés selon les conditions de marché.

Les travaux futurs comprennent l'élargissement de ce travail dans un cadre méthodologique pour l'aide à la prise de décision pour d'investissements stratégiques au niveau corporatif. De plus, une analyse des coûts marginaux basée sur les données réelles et sur une analyse de la performance des opérations pourrait être ajoutée à cette méthodologie afin d'analyser différentes options de procédés de bioraffinage forestier implantés en rétro-installation et différents niveaux de flexibilité.

ABSTRACT

The North American pulp and paper industry currently faces many challenges. Due to its commodity-focused and capital intensive nature, the industry experiences increasing difficulty to survive in the current global market place. Knowing individual product margins becomes essential to determine the optimal unit prices, thus uncovering the real operating profitability of manufacturing. However, current cost accounting systems that are based on predetermined resource spending can provide only ad-hoc assessment of these values, thus serve only as a mill benchmark for cost performance evaluation. The implementation of information management systems in pulp and paper companies has enabled a better understanding of both business and production processes. This allow for the development of advanced methodologies that assist the forestry sector in better controlling costs and improving profits. These systems are of especially critical importance for commodity producers such as newsprint mills, where low production costs are essential for business survivability.

The objective of this thesis was to develop a methodology for online manufacturing cost analysis using real-time process and cost data available from the information management systems that would be capable to assess actual product margin costs, and use this information for operational and tactical decision-making. Furthermore, the knowledge from applying this methodology can be explored in the strategic decision-making level for addressing the process-cost impact of retrofit design alternatives. This methodology consists of three major steps. First, a signal processing technique, based on multiscale wavelet transformation and filtering, is applied to simultaneously analyze every segment of the plant-wide instrumentation network. A second step further improves process data quality by confronting the plant-wide set of variables to the underlying fundamental process model using data reconciliation techniques. The plant-wide manufacturing information is updated by coaptation and correction of biased measurements. Third, this operational knowledge is integrated with the financial data in an operations-driven cost model to calculate and analyze production costs of operating regimes for short and long term benefits of the company. The methodology is applied to a case study considering current newsprint mill characteristics and potential retrofit biorefinery implementation.

It was found that using a combination of the wavelet technique with classical data reconciliation technique provides multiple facility-level benefits. An online implementation of this technique

was able to provide a significant number of data sets that were extracted from the information management systems as potential candidates to represent plant-wide and near steady-state operation. These data sets have provided sufficient statistical basis for characterising manufacturing operation per different operating regime. By doing this automatically, the methodology was able to enhance the quality of data and highlight the manufacturing region where the uncertainty in measurements is significant. The number of near steady-state candidates that can be detected was increased, when state identification parameters were being relaxed. However, it was shown that the uncertainty in the resulting data sets is increasing with relaxing the steady state assumption. In the analyzed rather simple newsprint operation, the technique was able to acquire multiple near steady-state data sets representing plant-wide operation with satisfactory level of accuracy. Moreover, the on-line implementation in combination with data reconciliation method, helped to improve measurement sensor performances by identifying sensors with systematic errors. This valuable information was then used to tune individual instruments further, and hence improve the overall performance of the methodology. Furthermore, it was shown that if this technique is implemented at the facility level for everyday use, it would help identify biased measurements in real-time and thus improve instrumentation and process troubleshooting significantly.

The manufacturing cost assessment based on these data sets that represent individual operating regime, has provided a new insights into the cost structure of the facility with transparent and visible process-cost interpretation capabilities. The application of the overall methodological framework, in the context of real production processes, has proved the ability to identify favourable and costly operating regimes when producing the same product grade. The characterisation and interpretation of the cost variances between individual regimes as well as within the same operating regime helped to identify process problems. Further application of the methodology for evaluating manufacturing costs of retrofit design scenarios have shown the ability to exploit the current operations-driven manufacturing knowledge based on regimes to enhance strategic decision-making at the facility. The results from this application showed that the operational profitability of new integrated production lines strongly depends on the operational differences in current manufacturing regimes of core business products. These differences in manufacturing costs can be visible from a process perspective and enable assessment of individual future product and mix-product margins. This information is essential for margin-centric supply chain planning of the enterprise and for exploring process flexibility to achieve an optimal production profile according to market conditions.

Future work includes the expansion of this work into strategic investment decision-making at the corporate level in order to enhance tactical and strategic planning. Furthermore, marginal cost analysis based on real-data and operations-performance analysis could be included in the methodological framework in order to obtain more flexible forest biorefinery retrofit designs with good strategic fit.

CONDENSÉ EN FRANÇAIS

L'implantation des systèmes de gestion de l'information dans les entreprises papetières a permis une meilleure compréhension de la dynamique des procédés et des affaires. Bien que les ingénieurs et les comptables incorporent maintenant les données en temps réel dans leurs pratiques quotidiennes, l'utilisation de celles-ci est souvent limitée pour la résolution *ad*-hoc de problèmes. L'information critique comprise dans ces données n'est souvent pas visible pour les gestionnaires. Les tendances des données sont étudiées, mais l'information est rarement extraite à partir des variables mesurées. Si les systèmes de gestion de l'information pouvaient être exploités à leur plein potentiel, les activités de prise de décisions stratégiques, tactiques et opérationnelles, seraient grandement améliorées par cet accès à une information nouvelle et pertinente.

Les entreprises de pâtes et papiers fabriquent plusieurs produits de commodités selon des spécifications individuelles à chaque client pour satisfaire ces derniers. Le fait de pouvoir comprendre les marges de chacun des produits devient alors indispensable afin de déterminer un prix de vente optimal et de montrer la vraie rentabilité de la production. Les systèmes et les pratiques de comptabilité actuels ne fournissent qu'une estimation *ad-hoc* de ces valeurs. En fait, on émet généralement une hypothèse d'homogénéité des produits, se traduisant alors par une distorsion au niveau des coûts lors de l'analyse. Les méthodes habituelles de comptabilité des coûts basées sur les recettes ne peuvent alors seulement servir que de point de repère pour l'évaluation de la performance. D'autre part, les calculs de coûts réels obtenus par les méthodes traditionnelles ne fournissent que des coûts agrégés qui sont évaluées de façon top-down. La séparation de ces coûts agrégés en groupes de coûts correspondant à chacun des produits est généralement basée sur le volume, incorporant alors plusieurs variations du procédé en raison de perturbations au niveau des matières premières ou bien de la dynamique du procédé. Cette façon de faire est souvent loin de ce qui se produit en réalité, ce qui rend l'estimation des coûts peu fiable pour la détermination de la vraie rentabilité d'un produit. Les comptables et les ingénieurs de l'usine reconnaissent généralement que le taux auquel chaque usine génère des coûts peut varier considérablement, et ce, même lorsque ces usines fabriquent le même produit. Ainsi, pouvoir déterminer la véritable marge de contribution d'un produit représente un réel défi pour les comptables de l'industrie des procédés, étant donnée que les données de procédé et de coûts sont toutes deux biaisées.

Cette thèse a pour objectif de développer une méthodologie permettant une analyse en ligne des coûts manufacturiers pour l'évaluation des coûts marginaux réels, et d'utiliser cette information pour la prise de décision au niveau tactique et stratégique. Cette méthode utilise des données de procédés en temps réel et de coûts provenant du système de gestion de l'information de l'usine. De plus, l'information obtenue peut être exploitée au niveau stratégique de prise de décision afin d'évaluer les impacts coûts-procédé de diverses alternatives de projets de rétro-installation. Cette méthodologie a été appliquée à une étude de cas considérant une usine de papier journal et une implantation potentielle du bioraffinage en rétro-installation

Une étape essentielle de cette méthodologie est la considération de la flexibilité et des différentes options de production. Les différents régimes d'opération ou façons d'opérer afin de fabriquer un produit doivent être clairement identifiées. Par exemple, pour le cas étudié, il était intéressant d'analyser la variabilité potentielle des coûts de diverses « recettes » utilisées pour produire un même grade de produit, mais utilisant différentes configurations de procédé caractérisées par 1) les différents points de consignes correspondant à une certaine qualité de pâte, 2) le type et l'âge des plaques de raffineurs, et 3) le débit de production. Ensuite, les trois étapes principales de la méthodologie ont été utilisées. Lors de la première étape, une technique de traitement de signaux, basée sur la transformation multiéchelle d'ondelettes et leur filtrage, est appliquée afin d'analyser simultanément chaque segment du réseau d'instrumentation de l'usine. Cette étape nettoie les données de mesures du bruit à haute fréquence et des anomalies, et cherche à identifier les périodes où le procédé s'approche d'un régime permanent. La seconde étape améliore davantage la qualité des données de procédés en comparant l'ensemble des variables de l'usine à un modèle fondamental de procédé. Cette information sur les procédés est mise à jour par coaptation et correction des mesures biaisées. Troisièmement, cette information opérationnelle est intégrée aux données financières dans un modèle de coûts axé sur les opérations afin de calculer et d'analyser les coûts de production de différents régimes d'opération.

Lors de la *première étape*, l'espace temps-fréquence de chaque point de mesure du système d'instrumentation de l'usine est analysé selon la méthode de transformation et de filtrage des ondelettes proposée. Les données en temps réel sont purifiées du bruit à haute fréquence et des anomalies de procédés. De plus, elles sont utilisées afin d'identifier diverses occurrences de régime permanent. Deux étapes essentielles sont nécessaires pour initialiser la technique des ondelettes :

- Recueillir de l'information pour chacun des points de mesure et analyser les données historiques dans l'optique d'identifier la coupe optimale de transformée d'ondelette (WT) pour chaque variable, et
- Analyser les données historiques pour initialiser les paramètres optimaux de détection des régimes permanents.

Après avoir appliqué la transformée d'ondelette, le bruit gaussien et les anomalies sont extraites ou éliminées des tendances de procédé. Le signal sans bruit est ensuite analysé afin d'identifier l'occurrence de régimes permanents suivant une méthodologie en trois étapes :

1. Le point de départ de la période de régime permanent est détectée en utilisant les caractéristiques de la WT et de sa première dérivée (valeurs prédéterminées du paramètre alpha)
2. Les composantes à haute fréquence du signal mesuré qui n'ont pas été préalablement éliminées, sont extraites par filtrage. La durée du régime permanent peut ensuite être estimée.
3. La fin du régime permanent est détectée en utilisant les caractéristiques de l'analyse de la WT.

Cette analyse a d'abord été effectuée pour un petit sous-système de l'usine. Les résultats montrent que cette technique est robuste et peut améliorer significativement la précision des variables mesurées, tout en fournissant des candidats de régime permanent variables et multi-variables. Ensuite, pour plusieurs variables présentant des comportements dynamiques importants, les hypothèses de pseudo régime permanent ont été relaxées en relaxant les paramètres de seuil pour la détection des conditions optimales de régime permanent. Les impacts de ces hypothèses vis-à-vis la précision des coûts ont ensuite été abordés afin de calculer les incertitudes reliées aux coûts. Les valeurs de seuil retenues sont directement reliées aux dynamiques du système impliquées dans l'hypothèse de régime permanent. Or, cette information est déjà calculée pour chaque variable, chaque sous-système et l'opération du procédé global. Cette dernière peut alors être utilisée pour améliorer les résultats de l'analyse des coûts en fournissant une valeur de confiance sur les marges d'un produit.

Lors de la *seconde étape*, l'exactitude et la validité des ensembles de données obtenus lors de la première étape sont améliorées par l'utilisation de la réconciliation des données basée sur une simulation. Cette étape est nécessaire pour fournir des données adéquates au modèle fondamental du procédé correspondant. Lors de l'étude de cas (usine de papier journal), plusieurs instruments mal calibrés ont ainsi pu être identifiés, et leurs valeurs erronées ont pu être ré-estimées et corrigées. La performance de ces instruments de mesure a donc été améliorée grâce à ces informations, qui permirent de mieux les calibrer. De plus, il a été montré que si cette technique était implantée à l'usine pour une utilisation quotidienne, elle permettrait d'identifier les mesures erronées en temps réel, améliorant significativement l'instrumentation et le diagnostic des anomalies de procédés. Une estimation des erreurs de mesure par la matrice de variance-covariance est ensuite effectuée en se basant sur l'analyse de performance de l'instrumentation de l'étape de prétraitement. Les poids de confiance pour chacune des variables sont attribués avec le jugement technique de l'ingénieur. Cette étape de la méthodologie garantit que la représentation des régimes d'opération de l'usine soit en accord avec la modélisation de procédé. Ainsi, le réseau du système a pu être vérifié simultanément pour d'éventuelles erreurs systématiques, aidant les employés de l'usine dans leurs activités reliés à l'instrumentation et à la résolution de problèmes. Il a pu être démontré que l'échec de certains instruments de mesure, ou simplement une mauvaise calibration de ces derniers, menait à des activités de production plus coûteuses.

Lors de la *troisième étape*, l'information caractérisant les différents régimes de production est utilisée conjointement avec les données financières de l'entreprise afin de développer un modèle innovateur de coûts basé sur les opérations donnant des informations pertinentes à propos des coûts manufacturiers, permettant ainsi d'évaluer les marges réelles de chacun de produit. Dans ce modèle, l'évaluation des coûts manufacturiers de chaque régime d'opération est basée sur les principes de la comptabilité par activités (comptabilité ABC). Le modèle de coût est développé avec le détail nécessaire pour extraire l'information liée aux changements lors de l'opération et aux modifications àla conception, et est utilisé pour évaluer les coûts de production par tonne pour chacun des différents grades de papier journal. Ainsi, les coûts manufacturiers directs et indirects de l'usine de papier journal à l'étude ont pu être caractérisés et interprétés afin d'identifier les régimes de production les plus rentables. Plusieurs implications ont pu être tirés des résultats de cette étude de cas pour générer des économies potentielles et procurer un avantage à court terme. D'abord, une amélioration de la compréhension et de la visibilité de la structure de coûts de l'usine a permis d'effectuer une interprétation des variances de coûts entre l'hiver et l'été, et a aidé à identifier les coûts associés aux transitions. À titre d'exemples, il a été constaté que l'arrêt des unités de récupération impacte de façon considérable les coûts de production en été. Aussi, les modifications quant à l'alimentation en matières premières ont pu être traduites en termes de variance de consommation de vapeur, d'électricité et de produits chimiques. Par ailleurs, il a été constaté qu'une variance significative au niveau des coûts survient quelques temps avant une transition planifiée entre deux produits. Cette variance est principalement due au fait que certains paramètres du second régime d'opération sont utilisés avant même d'avoir terminé d'opérer selon le premier régime.

Lors de l'application de cette méthode pour l'analyse stratégique de futurs projets de rétro-installation, il a été montré qu'un modèle de coûts basé sur les opérations utilisant les principes de la comptabilité ABC permettait d'améliorer le processus de prise de décision en fournissant des informations additionnelles habituellement non considérées. En effet, le modèle de coûts présenté calcule les coûts reliés aux procédés de production et n'effectue pas seulement une évaluation des coûts des produits. Il est donc possible d'analyser les implications de l'utilisation de divers régimes d'opérations pour chaque future option de procédé de bioraffinage implanté en rétro-installation. Entre autres, ces informations sur les différents régimes d'opérations permettraient d'améliorer la rentabilité de ces futures bioraffineries en fournissant l'information nécessaire pour utiliser de façon optimale la flexibilité des procédés selon les conditions de marché.

Les contributions les plus importantes de ce travail sont les suivantes :

- Une méthode en ligne basée sur la transformation d'ondelettes et de leur filtrage qui est en mesure de fournir une représentation en régime permanent précise de petits sous-systèmes. Cette méthode peut aussi représenter des systèmes à l'échelle de l'usine pour des applications industrielles où les opérations sont plutôt stables et simples.

- Un cadre systématique pour l'analyse et la gestion des incertitudes lors de la représentation d'opérations près du régime permanent. Cette approche pragmatique peut être utilisée pour attribuer automatiquement un niveau d'infidélité à un ensemble de données près du régime permanent représentant un régime de production.

- Une approche pratique mais valable pour la validation par modèle de données de procédés basées sur des mesures en temps réel. Cette approche combine la technique de prétraitement par ondelettes et la réconciliation de données basée sur la simulation en un seul système capable de fournir des ensembles de données réconciliés pour des systèmes à faible redondance.
- L'évaluation des implications économiques de procédés de fabrication par une approche reliant les données en temps réel et les données réconciliées de l'usine à unemodélisation des coûts basée sur les principes de la comptabilité par activité. Cette approche unique aide à caractériser et interpréter certaines relations de coûts de procédé complexes, et aide à la résolution de problèmes liés aux procédés.
- L'utilisation d'une technique unique d'épluchage de données de procédés qui caractérise les régimes d'opération à l'intérieur d'un environnement de production complexe. Cette technique permet d'analyser les coûts réels d'un procédé de production. L'information obtenue permet aussi de caractériser et d'interpréter les différences en termes de profit de différents produits et régimes d'opération, et fournit de nouveaux conseils pour l'amélioration continue.
- L'évaluation des impacts des coûts de procédé et des implications de projets stratégiques de rétro-installation par la combinaison systématique de données en ligne de procédés et d'un modèle de coûts avancé basé sur les principes de la comptabilité par activité. L'utilisation d'informations réelles de procédés et donc d'une meilleure compréhension de la structure de coûts de l'usine permet ainsi une meilleure prévision de la performance future de l'activité principale de l'usine, et met en évidence les combinaisons de produits les plus rentables.

Les aspects suivants constituent quelques possibilités de recherches futures :

- La méthode présentée pour le calcul des marges de produit pourrait être utilisée afin d'effectuer une gestion de la chaine logistique, une planification et un ordonnancement centrés sur les marges.
- Afin d'analyser différentes options de procédés de bioraffinage forestier implantés en rétro-installation et différents niveaux de flexibilité, une analyse des coûts marginaux

basée sur les données réelles et sur une analyse de la performance des opérations pourrait être ajoutée à cette méthodologie.

- La méthodologie développée dans cette thèse a été appliquée pour caractériser, interpréter et guider les activités de réduction de coûts pour une seule usine. Ce cadre pourrait être utilisé pour analyser tous les sites de production d’une entreprise dans l’optique d’améliorer la planification stratégique de l’entreprise.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

Variables and Parameters

A	Distance between measured value and steady state average
B	Distance between successive measurements
C_j	Coefficient of the smoothed signal at scale j
D_j	Coefficient of the detailed signal at scale j
J	Selected wavelet scale for on-line data treatment
G_1	Filtered distance between measured value and steady state average
G_2	Filtered distance between successive measurements
R	Ratio use to detect steady state duration
S	Sampling time
WT	Wavelet transforms
$\frac{dWT}{dt}$	First derivative of wavelet transform
$\bar{x}$	Filtered average
$\alpha_1, \alpha_2, \alpha_3$	Threshold used for steady state starting and ending point identification
$\beta_1,\ \beta_2, \beta_3$	Filtering parameters
$\sum_{i \varepsilon I_j} C_{j,i} \varphi_{j,i}$	Smoothed or approcimated signal
$\sum_{k \varepsilon K_L} D_{j,k} \psi_{j,k}$	Detailed signal
$\varphi_{j,i}$	Discretized scaling function
$\psi_{j,k}$	Discretized wavelet function
σ	Standard deviation
τ_i	Response time associated with variable i
n	Duration of true steady state
FC	fixed costs
Q	mass flow
R	revenue
VC	variable cost
x	non-correlated variables of Taylor expansion function
Σ	Summation
y	Measurement set

W	Variance-covariance matrix (Weighted matrix)
v_s,	Simulated variable
v_m	Measured variable

Abbreviations

ABC	Activity Based Costing
ABCEM	Activity-based cost and environmental management
ADMT	Air-Dry Metric Ton
BPI	Biopulping International
CVP	Cost-volume-profit analysis
DCS	Digital control systems
EtOH	Ethanol
FBC	Functional-Based Costing
FBR	Forest Biorefinery
FCF	Free Cash Flow
FSC	Full Standard Costing
GAAP	Generally Accepted Accounting Practices
GPK	Grenzplankostenrechnung (cost management system)
HAC	Acetic Acid
HP	High Pressure Steam
IMS	Information management systems
IRR	Internal Rate of Return
LP	Low Pressure Steam
M&E	Mass and Energy
MSE	Mean square error

MTE	Measurement trend error
NG	Natural Gas
O&M	Operation and Maintenance
ODCA	Operations-driven costing approach
PLA	Polylactic Acid
PSE	Process Systems Engineering
PWC	Process work center
RCA	Resource consumption accounting
RER	Relative Error Reduction
ROI	Return on Investment
SDDR	Simulation driven data rectification
t	Tons
TMP	Thermomechanical pulping process
VPP	Value Prior-to-Pulping

LIST OF APPENDICES

CHAPTER 1 INTRODUCTION

1.1 Problem statement

Canadian pulp and paper products are mature and standardized commodity goods, many of them beyond their product lifetimes. Every day, paper manufacturers are being challenged by diminishing market conditions and by increasing resource and energy prices. Different corporate business strategies are being investigated to remain viable in these tough market conditions and to achieve short- and long-term competitive advantage. Today's focus on managing at the strategic level by choosing the number of facilities, their locations, and their capacities is helping pulp and paper companies to reduce their enterprise cost curves temporarily. However, managing at the tactical level, in the facility itself, will not only tighten control of manufacturing costs at each individual mill, but will also provide critical information for long-term strategic planning and decision-making. To sustain a successful business in North America, however, these commodity products must first be manufactured at the lowest possible mill-level cost. Today, tactical or operating decisions are based mainly on information derived from mill benchmarking, home-grown cost accounting systems, or a combination of both. It is rare that this information is based on actual process measurement data from information management systems.

In recent years, developments in information technology and management systems and their use in the pulp and paper industry have expanded significantly. The use of these systems is, however, still limited, and they are often used only for troubleshooting. To enhance production profitability and product quality, the information captured in real-time data could be used to promote the development of decision-making tools which would enable mill personnel to react promptly to process and market changes. One of the major opportunities for exploiting the information extracted from the data available from information management systems (IMS) in pulp and paper companies lies in the field of manufacturing cost accounting, where operating-cost-related efficiency improvements remain to be achieved. To grasp effectively the operating knowledge that resides within cost-accounting data, the underlying operating characteristics must be integrated with the cost data. Because traditional cost-accounting practices use top-down cost allocation per volume of production using weekly or monthly statements, some other approach must be used to account for process operations. Activity-based costing (ABC), which has been developed in the last couple of decades, and its process-based characteristics could enhance production-cost modelling in the continuous manufacturing industries. Many companies, mostly in the discrete manufacturing industry, are saving millions of dollars through well-informed decisions based on results from ABC and its granular view of resource consumption. In the continuous manufacturing industries, however, only a few implementations have been done, with the majority occurring within oil and petrochemical companies. These home-grown enterprise-specific practices are often kept confidential as company know-how and are not available to the public.

Recently, a few applications of ABC-like cost accounting have been demonstrated in several case studies aimed at improving the visibility of manufacturing costs and using this information for higher-level analysis such as supply-chain management and retrofit design decision-making activities. The use of lower-level process data together with financial data in this "bottom-up" cost accounting approach has yielded a better understanding of complex continuous production environments such as those found in pulp and paper mills. In these high-level applications, the relatively long time scale used for cost modelling (weeks to years) is adequate. Further decreasing the time scale (to hours) for production cost assessment could enable tracking of actual product margins and their variations due to changes in operating practices. To do this, the biased process measurement data from information management systems must be corrected; otherwise, wrong cost information could distort the process-cost characterization and interpretation activities. Usually, data reconciliation techniques that enable validation of measured data against a process model are used for this task. This is a common practice in the petrochemical industry, where the fact that production remains fairly stationary enables the use of averaged process data (usually every hour) for data reconciliation. In a continuous manufacturing process which is characterized by faster process dynamics, several holding tanks, and process loops, data reconciliation results may result in an incorrect process characterization. In dynamic papermaking processes, regular use of data reconciliation with averaged measurements would provide an inaccurate process representation if not done carefully. Engineering judgment must be used to evaluate measurement trends to identify a time frame where data reconciliation could be used. This necessary step prohibits the use of classical data reconciliation in on-line applications for validating data from pulp and paper manufacturing processes. Furthermore, because of their age and characteristics, pulp and paper facilities do not offer sufficient system redundancy in their instrumentation networks for classical data reconciliation techniques to be used.

An on-line method that would be able to extract a measurement data set corresponding to a time frame when the manufacturing operation was near a steady-state process condition would be necessary to enable and facilitate the use of on-line data reconciliation in the pulp and paper industry. Generally, various types of filters are used to analyze the process status around a process unit or a small process subsystem. Very often, such an approach to process state identification fails because of the sudden occurrence of measurement spikes (outliers/abnormalities), resulting on false rejections of near-steady-state occurrence. Furthermore, for large systems or plant-wide applications, near-steady-state occurrence is not obvious because of the natural process dynamics involved between and within the individual sub-systems. The combination of these difficulties prohibits on-line application of data reconciliation in pulp and paper facilities. If a method existed that would provide reconciled data across the entire plant-wide operation and that hence would represent accurately pseudo-steady-state operating regimes (characterized by operating practices), the appropriate advanced cost models could then provide a unique operating profitability assessment of these production processes. This approach would provide, for the first time in the pulp and paper industry, access to the actual operating margins for each product, enhancing significantly mill cost-control strategies and providing an opportunity to use this information for far-reaching applications.

The aim of this work was therefore to develop a systematic methodology which would be able to provide valuable information on product operating margins for on-line industrial applications and to facilitate the use of this information to generate valuable decision-making knowledge. The near-steady-state detection technique should be able to detect and eliminate abnormalities, clean high-frequency noise from process data, and identify when plant-wide operation is near steady-state conditions. The data reconciliation technique should be able to compensate for the lack of redundancy in pulp and paper processes and to incorporate steady-state detection tools to enable an automatic flow of information among tools. The cost-accounting framework should be able to capture cost-process relationships eloquently in the form of visible and transparent resource-consumption insights. The cost analysis must further be able to analyze process cost impacts to generate a better understanding of the integration of new production facilities in retrofit design analysis.

1.2 Objectives

The discussion in the problem statement leads to the formulation of the main hypothesis of this work, which is entitled, "On-line steady-state data reconciliation for advanced cost analysis in the pulp and paper industry":

> *Given the availability of data and the emergence of information management systems, significant improvements in the profitability of pulp and paper mills can be achieved by coupling real-time process information with product and cost information, and interpreting both on the basis of plant-wide reconciled pseudo-steady-state data.*

This main research assumption can be divided into three sub-hypotheses:

- *Near steady-state conditions of a plant-wide operation can be identified from real-time data by using signal processing techniques based on wavelet transform.*
- *It is possible to use reconciled process data representing near steady-state operating regimes for instrumentation and process troubleshooting in pulp and paper mills.*
- *Product-based cost information can be assembled for continuous processes by resolving the product cost structure for different process operating regimes, thus providing information on the product operating margin in each case.*

Given the problem statement and the hypothesis as formulated, a systematic methodology was developed to demonstrate the value of the proposed framework in advanced production cost analysis for short- and long-term benefits. The methodology is presented through a case study of an integrated thermo-mechanical newsprint mill, which has the following objectives:

<u>Main objective</u>

- *To develop a practical methodology for making available on-line plant-wide reconciled process and business data in a form suitable for advanced decision-making and to demonstrate the value of this approach using a case study.*

<u>Sub-objectives</u>

- *To develop an on-line technique for analyzing process status and to identify when the manufacturing operation is near steady-state conditions.*
- *To develop a process model for data reconciliation technique in order to acquire reconciled near steady-state data sets that represent plant-wide operating regime and that can be simultaneously used for detecting biased instruments.*
- *To develop an operation-driven cost methodology that would systematically assess the manufacturing costs of an operating regime, with the ability to interpret cost variability and to use this information for the short- and long-term benefit of the company.*

1.3 Thesis organization

This thesis is organized as follows: In Chapter 2 the relevant literature is reviewed in order to identify the gaps in the body of knowledge. The next chapter presents the methodology and the case study to which the methodology is applied. Chapter 4 synthesizes the results obtained in the process of demonstrating the methodology. In Chapter 5 general conclusions are given, followed by the contributions to knowledge and recommendations for future work presented in Chapter 6.

CHAPTER 2 LITERATURE REVIEW

In this chapter, a critical review of pertinent literature was carried out. The thread of thoughts starts with section 2.1 covering the general principles of manufacturing costs assessment. Then, a critical review of classical cost analysis based on standard recipes and monthly spending, followed by the recent cost control improvements over the last decades.

The assessment and analysis of process-based manufacturing costs requires using lower time-scale than the traditional practices (hours). Essentially, this requirement creates the necessity to improve the accuracy of process characterisation. The representation of manufacturing processes must be based on reconciled process measurements (Section 2.2) in order to validate the measured variables with a process model, and thus make the overall fundamental manufacturing cost balances justified. Usually, production cost analyses are steady-state applications and hence section 2.3 deals with the pertinent literature done in the field of signal analysis for process state identification.

2.1 Manufacturing cost analysis

2.1.1 Introduction

Every business environment must exploit some level of cost-control strategies in order to analyze its variability in performance. For this purpose, cost and financial accounting measures are exploited. Cost accounting is the pillar of the accounting framework that provides valuable financial insight to decision makers. Commonly, the information provided is confidential and is used only internally to help managers find the optimal way to maximize the company's profits. The environment and the outcome of decision-making activities is the cost accounting system. Many business companies and production facilities use several various cost accounting systems for problem-solving. Considering that the limits of practice are entirely within the company's control, the prepared cost reports can be based on whatever rules, standards, or rational bases. Generally, the cost accounting knowledge is exploited in the second pillar of an accounting framework, financial accounting. This branch of accounting deals with public corporate information used solely for a company's financial statements, and its preparation must follow generally accepted accounting principles (GAAP).

The general control elements of cost accounting can be divided into three pools: material, labor, and overhead costs. Direct material and labor costs are generally variable costs and are a function of the number of units manufactured or sold. Overhead costs, on the other hand, are fixed costs that do not change with the level of production. For instance, management salaries, rent, or depreciation expenses do not vary from month to month, even though the rate of production is never the same. The ability to track these various cost elements accurately determines the value of the accounting system to final decision-making activities. In the early 20th century, this task was not difficult because overhead costs were negligible compared to material and labor. However, it became more complex to account correctly for indirect and overhead costs once the face of manufacturing had shifted from a labor-intensive to a machine-intensive environment.

Clearly, the ultimate goal in every organization should be to control these different types of costs. Often a company chooses to use only one costing system, even though there are several approaches available. The most commonly used in today‘s industrial practice are summarized in the table below:

Table 2-1: Cost accounting systems

Traditional costing framework	Funcional or activity driven
Cost-volume-profit analysis	Throughput accounting
Standard cost accounting	Lean accounting
	Resource consumption accounting
	Activity-based costing

The first two systems, cost-volume-profit and standard costing, are often referred to as *traditional* or *normal* costing and are used extensively in the pulp and paper sector. This traditional approach was created for the needs of the early industrial era when the total costs were dominated by variable elements (Enotes, 2011). The overhead and other indirect costs are accounted for based on simple volume-based measures such as labor or machine hours. Therefore, a product with a low level of labor hours is allocated less overhead cost. However, the actual costs may be very different if this product requires special attention or testing. The resulting unit production cost becomes even more distorted when overhead and other indirect costs begin to dominate overall manufacturing costs. Then it is strongly recommended that other supplementary costing systems be used. The following four costing principles are rather new in management accounting. These enhanced costing methods belong to the group of functional-based costing (FBC). The common factor of all four is the introduction of more levels of detail to the company's cost structure. This is done by simply assigning and allocating the individual costs to the unit-level of operation. This approach then aggregates the granular representation of costs in to the mill level for cost reporting and/or cost control purposes. The allocation bases and drivers are expressed by the use of production throughput or direct labour hours and machine hours. Thus, individual processing units, subsystems and activities (departments) are characterised as the consumer of resources. The end products are then consuming these individual activities, instead of the direct assignment of all costs to end product.

- Throughput accounting was developed for the enterprise-wide level, to help identify factors that limit the enterprise in achieving its established goals (Eliyahu 1992).
- In lean accounting, the essential philosophy is to preserve value with less work. This approach was developed for the car industry which was aiming to eliminate waste while simultaneously minimizing production costs and time, using techniques such as poka-yoke (Robinson 1997) or value-stream mapping (Rother 1999).
- Resource consumption accounting (RCA) is a fully integrated and complex managerial approach that uses available state-of-the-art methods. The combination of the German *Grenzplankostenrechnung* (GPK) cost management system and activity-based costing principles create a system that can be used and interpreted by non-accountants. The next two sections explain more in details the cost structures and cost practices that are incorporated within these different types of cost accounting system.

2.1.2 Traditional cost accounting practices

Generally, continuous industry such as mining, pulp and paper, in some cases petrochemical and chemical industries use accounting practices that are dominated by traditional costing, because of its simplicity and the wide understanding of this approach among accountants. An important part of standard costing is a variance analysis. By breaking down the overall variance into the three pools listed below, this analysis helps managers identify where the difference between actual and budgeted costs has occurred (Table 2.2):

Table 2-2: Example of some of the elements of traditional cost variance analysis

Cost-volume variances
Material-cost variation
Volume variation
Labor-cost variation

This valuable information assists managers to identify the source of the overall cost variance, but not the cause of it. For instance, if the variance is largely due to material-cost variation, accountants with the help of process engineers need to drill down into historical process data to interpret the variance and take appropriate action.

The problem is that traditional costing considers all costs as variable with regard to production volume. This often creates inaccuracy in fixed costs whenever the volume of production changes. Furthermore, arbitrary rather than cause-and-effect overhead allocation makes the traditional approach highly inappropriate in a multiproduct environment. Another problem in the current general accounting profession, not only in the forestry sector, is the emphasis on financial accounting. Most of the time, decision-makers must create their own cost analysis based on financial accounting reports. However, these statements contain aggregated and distorted costs with no activity data incorporated, leading to poorly informed decisions. There are a few existing advanced systems at the academic level (operations-driven cost modeling developed by Janssen, 2011; cost-function modeling by Fogelholm, 2000) or already being used by advanced processing industries such as the petrochemical sector. The pillar of these approaches is the principles of activity-based costing, which is briefly discussed in the following section.

2.1.3 Activity based costing accounting

Activity-based costing (ABC) is a new philosophy that emerged in the 1980s in response to overhead allocation discrepancies (Kaplan 1989). By simply adding an activity as a link between resource consumption and a cost object, the knowledge of costs incurred in the organization is improved significantly. The activity becomes a fundamental cost item whose value is directly assigned to the final cost objects such as products and customers. In other words, the rate of resource spending is traced to an activity, and the activity is then traced to the product, as shown in Figure 2.1 (Korbel and Stuart (c))

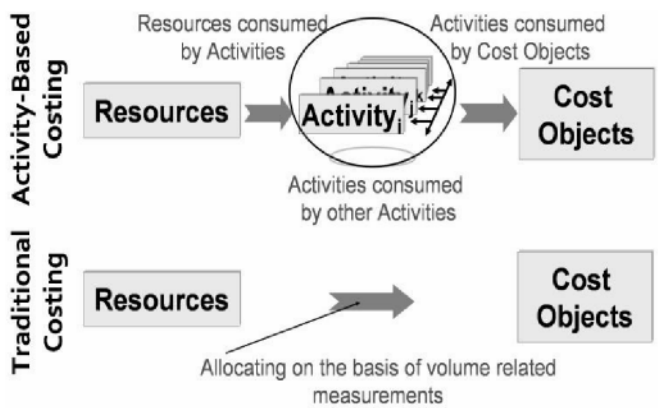


Figure 2.1: Activity-based costing and traditional costing

The significant advantage of using ABC is that it attempts to assign all costs to final cost objects, including marketing, engineering, and administrative costs. This added ability to trace indirect costs directly enables accountants to track overheads rationally and as closely as they track direct costs. This is done by making use of so-called *drivers*. As shown in Figure 2.1, resources are linked to activities by resource drivers, and similarly activities are linked to cost objects by cost drivers. According to this definition, resource drivers determine the amount of a resource consumed by each activity, while activity drivers specify how different cost objects (products, customers) consume these activity costs. Labor hours, kWh, and number of shipments are examples of resource drivers, whereas number of customers and number of products are examples of the second stage, the activity driver. The difference between these drivers is that the former focuses on why things happen and the latter on what happens (Emblemsvåg and Bras

2001). The implementation of an ABC system may be a complex and expensive task, and therefore it is important to determine the minimum number of appropriate drivers that will meet accounting objectives.

As shown in Figure 2.1, the process-oriented character of ABC means that it is implemented in two simple and logical stages, while structure-oriented traditional costing is implemented in one. This fundamental principle is the basis for increasing the accuracy of the cost data (Drucker, 1996). Traditional costing cannot encompass this critical linkage between actual causes and associated costs. Furthermore, advanced ABC has recently evolved into multistage systems where individual activities can be used by other activities before being used by final cost objects, thus enhancing even more the accuracy of cost modeling (Emblemsvåg and Bras, 2001).

In a continuous manufacturing context, the process-oriented character of an ABC system and the causal relationships between cost drivers and activities make the method highly suitable for modeling and analyzing costs. The availability of real-time cost and process data from information management systems (IMS) makes ABC easier to implement. It must be made clear that ABC is a cost accounting system that can help managers understand their actual costs and improve their profits efficiently. Traditional methods are complementary to the financial reporting prepared according to GAAP.

2.1.4 ABC-like cost accounting

A cost accounting system that is used by a wide spectrum of industries is the *resource consumption accounting* (RCA), whose development has been strongly influenced by German cost accounting and ABC principles. The structure is very close to variable costing, a well-documented method discussed in cost accounting textbooks, but rarely used by industry. RCA and its variations are extensively used by advanced processing industries such as mining, petrochemicals, and chemicals. Often their costing methods are confidential and inaccessible to the public or to researchers. In general, RCA is based on three fundamental pillars (for further details, the reader should refer to Friedl and Kupper (2005)):

- View of resources: The use of a high volume of cost pools establishes a clear linkage between resource spending and a company's costs and revenues;

- Quantity-based model: The value of the costing system is created in this pillar by the use of operations data and models. Traditional costing uses the output of variance analysis with dollar values, which create overhead-costs bias due to their fixed nature. By contrast, RCA exploits causal operational relationships;
- Cost behavior: Understanding the nature of costs is a very important aspect of the third pillar of RCA. The clear distinction between direct, indirect, variable, and fixed costs is based on aggregating pools.

There have been significant changes in recent years, although not well documented; some forestry companies are approaching now ABC-like costing for improved decision-making activities. For example, Fogelholm (2000) has discussed the difficulties of product costing in the paper making industry and its potential industrial application. This approach is now a pillar of Metso Automation's MetsoDNA - Dynamic Network of Applications (Metso, 2011) that some companies are presently using for product-customer decision making as well as it helps their budgeting activities. The application seek to anticipate and determine the resource requirements for the next individual customer orders based on current raw material content, dimensions and quantities of the paper product (Fogelholm, 2004).

Several applications of process-driven accounting frameworks have been developed based on ABC philosophy with potential industrial applications. For example, an approach that integrates ABC principles with environmental metrics to perform analytical economic and environmental assessment for decision-making activities was developed by Emblemsvåg and Bras (2001). Their activity-based cost and environmental management (ABCEM) system is extensively discussed in their 2001 book. The use of an uncertainty variable introduces extra complexity and versatility into the system. The ABCEM has been applied to a wide range of industries including furniture, carpets, and supply vessels, where it has provided insights and highlighted potential areas for improvement.

Lastly, a sophisticated ABC-like approach that integrates process and cost information into one system is called operations-driven costing (Janssen and Laflamme-Mayer 2011) method. The basis of this approach is in making a link between costs and process operations data using principles similar to those of activity-based costing. This approach is similar to RCA in some aspects, but is more versatile because it includes an in-depth engineering understanding of the process operation. The following section discusses this approach in more detail.

2.1.4.1 Operations-driven costing approach

The *operations-driven costing approach* (ODCA) is an interdisciplinary approach developed by accountants and process engineers in the pulp and paper industry. Understanding the cost of process performance is a critical success factor for paper mills. Chemical and process engineers are concerned with developing systematic tools and methodologies for both optimal design and optimal process operation. These procedures range from nano to industrial scales (Puigjaner, 2006). The concept can be understood from the supply chain point of view, where on the one hand, product quality is determined on the nano or micro scales, and on the other hand, the desired product properties are determined by its functionality and structure (Figure 2.2 Grossmann, 2005).

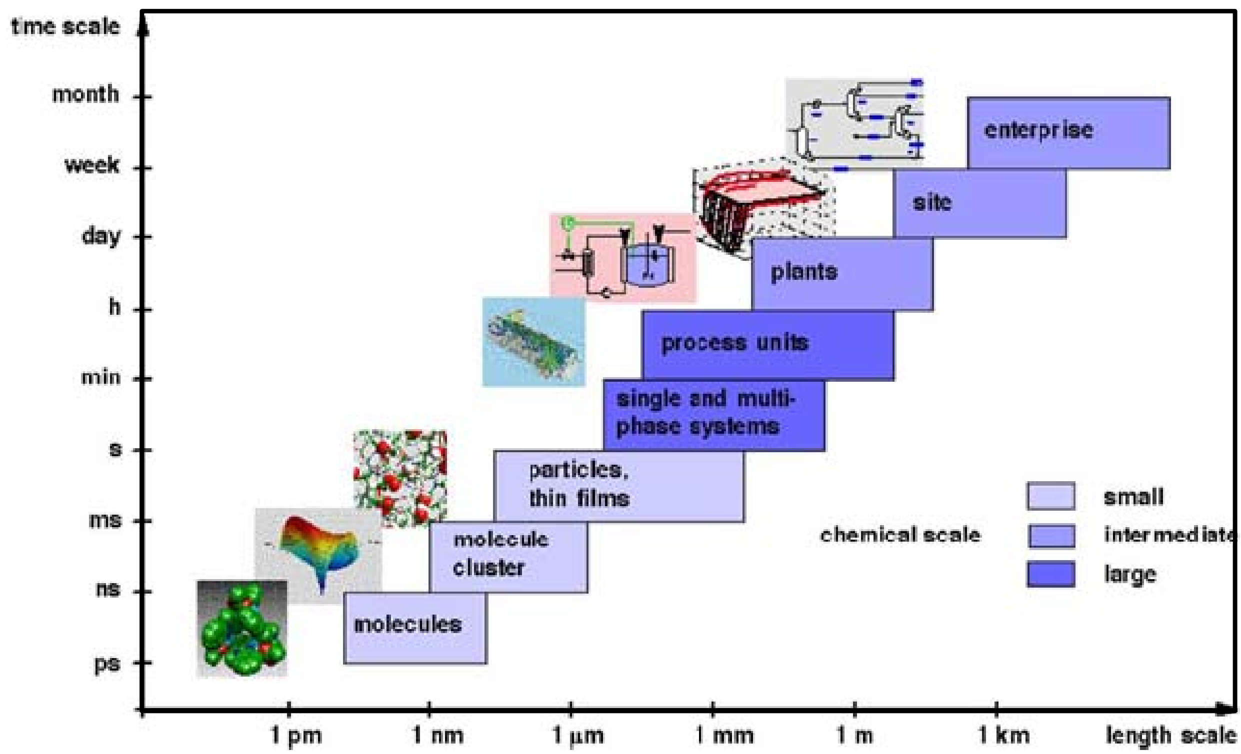


Figure 2.2: From micro scale to macro scale complexity or the “chemical supply” chain

In the pulp and paper industry, fiber micro properties influence the quality of pulp and paper products. On the macro industrial scale, as an analogy to the principles presented on the figure

2.2., a reflection of the micro complexity of the fiber structure can be brought to light using information extracted from real-time data through IMS. The most practical way of doing this in the pulp and paper industry is to develop tools and methodologies for macro or mega scale applications that are based on real-time data and that reflect the meso and micro scales according to the general chemical-engineering definition of complexity levels.

Janssen and Laflamme-Mayer (2011) developed an operations-driven cost modeling framework to provide in-depth understanding of resource consumption by integrating process and cost data. The bottom-up structure (Figure 2.3, Laflamme-Mayer, 2011) provides mutual communication between different business levels. The resulting generic framework can be used to enhance the understanding of manufacturing processes both for design and for operational decision support.

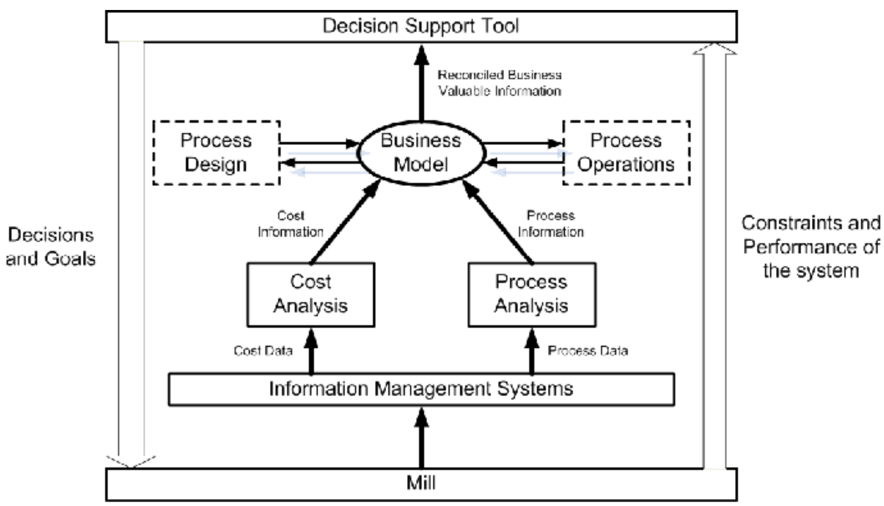


Figure 2.3: Overview of the bottom-up process-based approach

Later, Laflamme-Mayer (2011) presented in work an application of operations-driven cost modeling to assess the production costs for different product campaigns. This information was then used for planning and scheduling and optimization of high-level supply-chain analysis. The understanding and differentiation of product margins for each campaign can be used to enhance the current *ad hoc* representations of product margins. This versatile view of manufacturing costs in the paper industry is revolutionary and has tremendous value for reducing production costs.

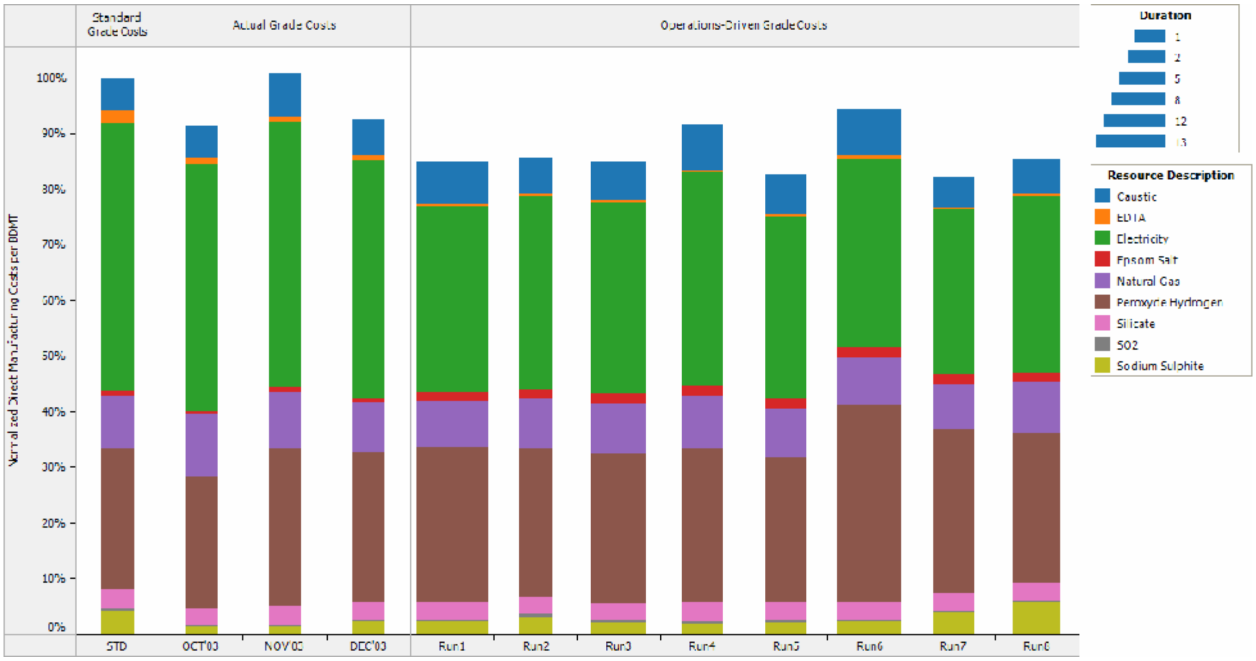


Figure 2.4: Comparison of standard cost, actual grade cost, and operations-driven grade costing

Figure 2.4 (Laflamme-Mayer, 2011) compares standard costing with actual and operations-driven costing information. Standard costs represent how the resources should have been used to manufacture a particular grade; actual grade costs are the true resource consumptions at the end of the three-month period. The operations-driven grade-cost assessment breaks up the three-month period into segments corresponding to campaign runs. From these results, it is clear that manufacturing the same product varies significantly from one campaign to another.

2.1.5 Product profitability assessment

In the continuous industries, such as pulp and paper producers, the multiple-product environment is characterized by numerous customers with different needs and corresponding product specifications and prices. It is therefore crucial to understand individual product margins; however, as discussed above, the current cost practices and systems provide only approximate values that are based on time-framed (usually monthly) spending information.

These traditional accounting systems use practices that aggregate costs over the manufacturing period and use a standard recipe, e.g. it is based on experience from the process operation. This overall cost information incorporates various changes in process operation due to mechanical (process) or raw material disturbances. By experience and understanding of production, engineers and accountants can recognize that within the manufacturing period, the generation of cost differs from one product to another as well as within the production process for the same product. However, it is not a simple task to determine these cost variances. First, current accounting practices cannot accommodate cost information from a process perspective, and second, the discrepancies in the current cost data are significant. In 2003, a survey by Ernst & Young and the Institute of Management Accountants (Ernst and Young, 2003) indicate that 98% of respondents claim that cost reporting is distorted, with indirect costs and overhead allocation being the main biases reported, and almost 40% believe that the cost data they receive are significantly inaccurate. The following section summarizes several pertinent applications to determine product margins in the continuous industries:

- Operational Performance Simulator (from Acorn Systems©) - the engine of this application is the integration of operational data with financial planning software. It allows decision makers to create proactive plans and analyze different operational scenarios. The ABC character of the application is providing enhanced cost-control via benchmarking key performance indicators. The applications are based on standard costs and use several real-time data points for benchmarking purposes.

- Fogelholm (2000) has developed a methodology for product costing in the paper making industry and for its potential industrial application (Metso Automation's Metso DNA ©). The use of standard drivers (instead of real-time data) for profitability assessment enhanced the insights and understanding of the cost variances between different productions recipes.

- Quesada-Pineda (2004), the use of ABC was explored for forestry and wood sector. The methodology is based on classical ABC principles. The division of different operating activities into batch tasks helped indenting optimal sequence of production.

- SAS® Profitability Management – the profitability assessment provides increased analytic power and faster results based on ABC principles. This power allows ABC analysis of complex business models. Several petrochemical facilities use SAS® Profitability Management for real-time profitability analysis. Different tools of data mining and statistical analysis are used for results transparency uncovering the hidden insights into various drivers of profitability. This system is based on reconciled data using Simagine© form OSISoft®. However, very often biased values in business data were reported due to process data inaccuracy.

- Laflamme-Mayer (2011), application of ABC-like cost accounting to address the costs of production in different product campaigns, and use this information for supply chain optimization. The ABC-like character of cost accounting system used, have provided improved insights into production costs, and helped to enhance the overall corporate profitability.

2.1.6 Manufacturing costs assessment for retrofit design evaluation

Generally the main financial indicator in retrofit design assessment is the profitability of the project with the emphasis on profitability of the future operation. The profitability of the projects are expressed by different types of metrics, e.g. return on investment (ROI), turnover ratio and payback time), and more recently characterised measures, such as discounted cash flow rate of return, net present value (NPV), or internal rate of return (IRR) (Dimian 2003). For profitability of the future operation, different types of manufacturing costs such as variable costs (materials, energy, and chemicals) are generally estimated using monthly statements and different purchasing measures. The level of variable costs must be calculated based on material and energy balances. On the other hand, the fixed and indirect manufacturing costs (maintenance, administration, labour, operating supplies, insurances, rents, other overheads, etc.) are estimated based on the predetermined rules for the operating requirements. For this purpose very often ratios and factors of capital investment costs are exploited (Dimian 2003).

Financial analysis for process retrofit projects usually seeks to evaluate capital investment cost and ignores the necessity to account for evaluating operating cost. However, recently some work have been done using advanced accounting methods for continuous process industries

- Janssen (2007) have introduced the use of operations-driven cost modelling in a single-product and multi-feedstock system retrofit design problems. The industrial case study was looking at the increased deinked pulp production and cogeneration at an integrated newsprint mill. The drivers characterising resource and activity consumptions were defined by the use of data from the information management systems. The processes of new retrofit options were simulated by suing simple first principle analysis. Fixing by-product cost flows was used to approximate the main product cost flow between activities and overhead costs. This cost information have allowed for further used in net present value analysis of the design projects and compare different alternatives.
- Janssen et al, (2011), presented a generic retrofit process design methodology uniting process and supply chain level assessment steps. The particularity of this sustainable design methodology is that: 1) generation of design options is based on process data, techno-economic study and environmental impact analysis; 2) operations driven cost modelling; 3) Life Cycle Analysis and Supply Chain –level profitability assessment; 4) multi-criteria decision making (MCDM) process for selecting the preferred design option using the whole set of criteria covering economic, environmental and supply chain profitability.
- Sadhukhan (2008) presented a new methodology for retrofit process design based on value analysis. A design superstructure was formed summarising all retrofit alternatives and then linear mathematical programming was exploited to identify the optimal operating conditions in each of the subsystems of the whole master problem. Next, for manufacturing cost evaluation, these favourable operating states are selected. Cost association between multiple results of operational segments are not considered, e.g. all products of the segment are assigned the very same manufacturing costs. The methodology was applied to a case study being oil refinery.
- Janssen et al, 2011, explored the use of marginal costs of energy (steam and electricity) and the impact to production rate change on project profitability using the ABC-like costing method. The results indicate that system constraints govern in that specific

context and that design capacity was the most suitable with regards to all marginal performance measures. (Janssen, Naliwajka et al.).

- Hytönen and Stuart, (2010), presented a sophisticated methodology for identifying promising retrofit integrated forest biorefinery strategies - design decision making under uncertainty. In his work, ABC-like accounting was used to improve transparency into uncertainty of raw material prices on production costs.
- Process integration investment decision making was studied by Berntsson et al. (2009) under uncertainty, focusing on improving energy efficiency. Several process integration tools were used consecutive steps.: pinch analysis, process simulation and optimisation and scenario planning

In summary, if production facilities are able to determine their actual values of individual product margins, the implications are enormous. In order to do this actual operating knowledge must be involved. The use of IMS becomes pertinent to extracting process measurement data that provide knowledge about the underlying process. However, the lack of reliability in certain measurements as well as the lack of instrumentation on site makes this task very challenging.

2.2 Process data reconciliation

2.2.1 Introduction

The data reconciliation problem is an old industrial application which was proposed first by Kuehn and Davidson (1961) to minimize the error between measured data and the underlying process model. Since they first published their pilot solution to the linear steady-state data reconciliation problem, further studies have led to progress in this area. Crowe (1996) proposed to solve the non-linear data reconciliation problem by successive linearization. Liebman and Edgar (1988) demonstrated that reconciliation results can be improved by non-linear programming instead of successive linearization when solving the non-linear data reconciliation problem. Tjoa and Biegler (1991) showed that non-linear programming together with a contaminated normal (Gaussian) objective function other than the least-squares objective function can improve the results further. Many other developments in data reconciliation and gross error detection have been proposed in numerous papers (Arora & Biegler, 2001; Johnston & Kramer, 1995).

The most usual estimator exploited for data reconciliation is the weighted least-squares estimator, which is very sensitive to the potential presence of systematic errors, often referred to as gross errors. If gross errors exist in the measurement data, the weighted least-squares estimator will yield incorrect estimates which will then significantly deflect reconciliation of other measurements. The critical task of identifying the presence of gross errors and estimating their values remains a challenge in practical industrial applications. Several methods to solve this problem have been proposed, for instance, the measurement test gross-error detection method presented by Tamhane and Mah (1985) and the modified iterative measurement test gross-error detection algorithm presented by Serth and Heenan (1986). Other statistical approaches have also been used, such as the generalized likelihood ratio (Narasimhan & Mah, 1987), the maximum power test (Crowe, 1992), and the principal component test (Tong & Crowe, 1995). The method proposed here exploits the statistics of the historical measured process data. The analysis of each measured variable is compared to its historical values. If a change is detected, then the current systematic error is estimated, and the biased measurement value is corrected. Data reconciliation is then repeated with the new corrected value of the measurement.

2.2.2 Formulation of data reconciliation problem

Measured data quality affects not only the quality of high-level tasks such as optimization and cost accounting, but also the quality of any estimated process model. Therefore, reliable and consistent measurement data play an important role in process plants. Random and gross errors can result in poor-quality measured data. Data processing and data reconciliation can be beneficial in minimizing measurement errors. The general form of data reconciliation is the minimization of measurement errors subject to the constraints of the physical process. Random errors are minimized by the use of data processing techniques and refined further in a reconciliation step. However, systematic errors must be identified and estimated before reconciliation. For a steady-state application, inconsistencies between instrument values and a steady-state process model are also caused by process dynamics. In this case, data reconciliation helps to distribute the errors caused by the steady-state assumption systematically onto the whole set of variables while still satisfying the underlying process model. For purposes of illustration,

Figure 2.5 presents a simple data reconciliation problem around a single unit, a splitter. The illustration shows the minimisation problem on the variables (for simplicity in illustration the third variable is equal to 100 and is considered to be as "perfect" measurement). The physical constraint is a steady-state process model. The measurements are an average values within the steady-state duration (the length of steady-state must correspond to at least the time lag of the splitter processing unit)

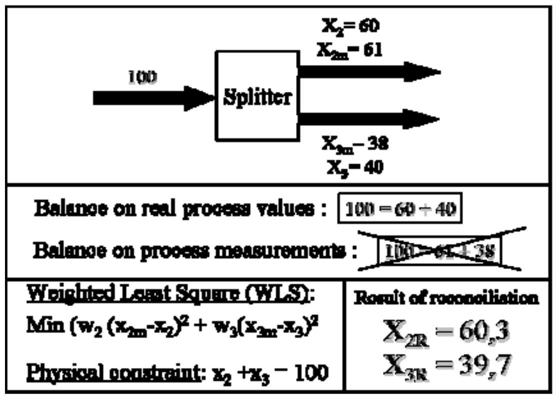


Figure 2.5: Example of the data reconciliation concept around a splitter

A simplified mathematical formulation of the data reconciliation approach may be written as the weighted least-squares minimization problem of the difference between measured/unmeasured and reconciled values of variables with regard to instrument and process characteristics:

Equation 2.1: General formulation of data reconciliation problem

$$Min\sum_{i=1}^{K}\left(\frac{(measured/unmeasured)_i - estimated_i}{weight_i}\right)^2$$

Subject to: mass, energy, component balances

With an assumption of normally distributed random errors with no systematic errors present, this constrained minimization procedure was first introduced by Kuehn and Davidson (1961). It is important that analytical or hardware redundancy[1] of measured variables be present if data reconciliation is to be performed. The character of the problem represented by Equation 2.1 depends on the formulation of the constraints, e.g., linear, non-linear, or dynamic. Furthermore, data reconciliation not only validates and estimates measured values, but also provides estimates for variables that are not being measured (often referred to as a coaptation process).

Many methods in the literature provide simplifications and solutions of the problem stated in Equation 2.1 by eliminating the unmeasured process variables from the problem statement. In linear data reconciliation, Crowe (1983) used a projection matrix method to decouple the unmeasured variables from the constraints. Other methods have also been used, such as a Gauss-Jordan elimination procedure (Madron, 1992) and QR decomposition (Sanchez, 1996). For non-linear data reconciliation, the procedure is based on successive linearization of the constraints, and the resulting simplified problem is then solved using Crowe's (1983) projection matrix (Pai and Fisher, 1988, Liebman 1988, Veverka, 1997). Crowe (1986) extended his previous technique to non-linear (bilinear) processes by a two-step projection matrix technique which significantly reduced the computational effort for bilinear systems. Many other authors have addressed the computational challenges of data reconciliation for particular cases. However, there is a lack of practical applications to the pulp and paper industry, which is to some extent due to the low system redundancy of the papermaking operation. Monitoring of a sufficient number of variables to ensure redundancy is limited by installation and maintenance costs (Jacob, 2003), inaccurate measurement techniques, and the current unavailability of instrumentation. Hence, reconciling process data in the pulp and paper mills becomes a challenging and often impossible task.

2.2.2.1 Weighting matrix estimation

The success of data reconciliation methods generally depends on the hypothesis that the errors are normally distributed, and hence on the quality of estimation of the variance-covariance matrix. This symmetric and positive-definite matrix quantifies the uncertainties in each

---

[1] It is said that a measured variable has hardware redundancy if two instruments are used to measure its value. On the other hand, analytical (software or spatial) redundancy of a variable is ensured when its value can be estimated in two independent ways, e.g., by a measurement and by a value from a process model.

instrument value (Benqlilou, 2004). If the process is truly at steady state, the covariance matrix can be estimated by the direct method (Bagajewicz, 2000), which is simply a sum of standard deviations within the time of the true SS. The mean value can be calculated as:

Equation 2.2: Weighting matrix estimation

$$\bar{y}_i = \frac{1}{n}\sum_{k=1}^{n} y_{i,k}$$

and the covariance matrix can be estimated as:

Equation 2.3: Covariance matrix can be estimation

$$cov(y_i, y_j) = \frac{1}{n-1}\sum_{k=1}^{n}(y_{i,k} - \bar{y}_i)(y_{j,k}\ \ \bar{y}_j)$$

where *n* is the duration of true steady state and *y* is the measurement set. This direct estimation of the covariance matrix helps to correct instrument values on an optimal statistical basis. Because it is known that the process is never at true steady state, the process of estimating the variance-covariance matrix becomes more complex (Gedeon, 1984; Crowe, 1996; Chen, 1997).

A highly simplified and common practice in industrial applications is to use engineering judgment for matrix estimation by allocating uncertainty weights to each instrument (Narasimhan and Jordache, 2000). This pragmatic approach, which is also used in the current study, takes into account knowledge of the process dynamics around each particular instrumentation sub-network as well as information about each instrument's accuracy, precision, and reliability.

2.2.2.2 Gross error handling

Because data reconciliation is limited to the elimination of random errors, systematic errors must be eliminated *a priori*. There are generally two principal issues linked to gross error (GE) handling, e.g. the gross error detection and its value estimation. Furthermore in any proper GE detection technique, the four following points should always be addressed:

- GE detection – possibility to identify the existence of one and/or multiple GEs in the measurements
- GE location – possibility to locate one or multiple gross errors

- GE identification – possibility to determine the GE type,
- GE estimation – possibility to estimate the magnitude of GE.

Several methods are available to do this, ranging from pure statistics through neural networks to time-series screening. The efficiency and practical usefulness of the statistical approaches seem to be superior to the others. In the present work, historical knowledge about the potential locations and relative sizes of biases is used. This approach can be situated in the framework of measurement adjustment using statistical methods such as the measurement test (Mah, 1982 or Crowe, 1983). In this type of method, the data are first reconciled, and then each measurement point is tested for possible bias. If gross errors are present, then their values are estimated by solving a simple non-linear problem (McBrayer, 1995):

Equation 2.4: Gross error estimation

$$Min\ B(y,b)$$

$$\text{s.t.}\quad f(y) = 0$$

$$y_{min} < y_i < y_{max}$$

$$b_{min} < b_i < b_{max}$$

$$B(y,b) = (-)_{i-1} \ldots + [(y_i - (y_{mi}\text{-}b_i))/s_i\ ]^2 + \ldots(-)_{i+1}$$

where y_i is the i^{th} measured variable, i y is the i^{th} estimate, s_i is the measurement noise standard deviation of the i^{th} measured variable and b_i is the estimate of bias on the i^{th} measured variable. It must be noted here that the bias, bi is also included in the inequality constraints which allows for setting limits on the range of biases that are admissible. Once the values of the gross errors are known, the biased measurement value is corrected, and the data reconciliation procedure can be run again. The process is repeated until convergence to a minimum value of squared error is achieved.

2.2.3 Plant-wide data reconciliation

In large-scale applications such as plant-wide optimization or cost modelling, the need for reliable[2] and consistent[3] plant-wide data sets is critical. To reconcile process data plant-wide, it is common practice to use a data set that consists of averaged variables over a specific, fairly stationary time period. However, the discrepancies in the measurements are due not only to random errors (assumed to be normally distributed), but to process dynamics as well. Clearly, this type of error does not follow the assumption of a normal distribution. Bagajewicz and Jiang (2000) has shown that in systems with no significant hold-ups, this error can be neglected. However, a papermaking operation is a collection of unsteady-state manufacturing processes with many hold-ups and recirculation loops; hence, significant error may be created if averaged process data are used with no systematic approach.

A different approach to plant-wide data validation is dynamic data reconciliation, which has received significant attention within the last decade, although large-scale industrial applications still remain to be developed. Clearly, the inhibiting factor is the high computational requirements of these procedures. However, with today's advances in information technology, and considering that processes are actually never at steady state, it may be better to consider using dynamic data reconciliation techniques even for near-steady-state processes (Narasimhan and Jordache 2000). On the other hand, from a practical perspective, for on-line applications, it would be wise to extend steady-state reconciliation to deal with dynamic situations (Benqlilou, 2000). Process optimization and process control would undoubtedly benefit from dynamic data reconciliation; however, for cost modelling, steady-state, not dynamic, data sets are required. In fact, as mentioned by Bagajewicz and Jiang (2000), for the time being, there are more pressing problems to resolve, for example gross error detection, which is closely related with the problem of data reconciliation.

---

[2] "Reliable process data" refers here to estimated variable values that are close to the true values of the process variables.

[3] "Consistent process data" refers here to data sets that are consistent with the underlying process model.

Several publications are devoted to gross error handling, and many methodologies and techniques have been proposed. However, our capabilities to detect and correct gross errors are still limited. As for current commercial software available for gross error handling, the main technique used today is the serial elimination strategy. As mentioned by Bagajewiczand Jiang (2000), vendors need to improve their strategies, for example by implementing methodologies to handle uncertainty and to enhance processing of gross errors in an industrial context.

2.2.4 Industrial applications of data reconciliation

Many commercial software packages for process analysis and simulation today provide integrated functionality for data reconciliation. All these applications are for linear steady-state data reconciliation. Bagajewicz and Rollins (2002) discussed the functionality of one academic and eight commercial packages and concluded that most of them deal with material and component balances. Only a few provide the advanced possibility to connect directly to DCS systems for on-line applications. Generally, all the packages were developed for the petrochemical industry, with embedded features such as phase equilibrium constraints or model libraries for some petrochemical process units. The most popular packages in the industry are Sigmafine (OsiSoft), Datacon (Invensys), and Vali (Belsim).

Industrial applications of any commercial software in the papermaking industry are scarce because of the dynamic nature of the process and its lack of measurement redundancy. Summary of some industrial applications are listed in the following section:

- The Sigmafine package has been used for off-line data conditioning (Jiang 2003a) and has been assessed for possible on-line application in a recausticizing plant at a kraft paper mill. The application was limited to material balances because the package cannot accommodate non-linear constraints such as energy balances.

- A further commercial implementation of this software involves a dynamic application to track pulp stock from batch digesters (Rankin 2009).

- Another package used in papermaking is the CADSIM® Plus simulation software from Aurel Systems. It has been applied to on-line energy accounting for the steam utility system in a kraft paper mill (Wasik 2007). This practical approach uses a parallel process simulator module to perform a process data validation procedure, and performs data reconciliation of dynamic data.

- In the work by Osisoft (2003), the use of a process information system at the Alabama River mill in Perdue Hill is described. The new system focuses on plant coordination in order to better synchronize processes and holding tanks, thus avoiding slowdowns, speedups, and shutdowns. Personnel can identify and examine a process problem in order to solve it rapidly. The applications are used also for monitoring performance, troubleshooting and reporting.

- Another application of Osisoft (2002), at the Georgia Pacific mill in Plattsburgh, NY. The data management system monitors and collects process data from the paper machine and from the winders. The output of the implemented system is used for maintenance, process troubleshooting, and monitoring grade specifications.

2.2.4.1 The use of optimisation and simulation modules for data reconciliation

In the case of CADSim software that is very adaptive to papermaking environment, the simulation solver is sequential, but the data rectification solver is simultaneous. The principle is to select several key and independent variables (by the plant operator or analyst) and use these to model the process operation using simulator modules (Figure 2.6, Korbel et al (b)) in order to estimate the whole set of process variables. The pillar of the data rectification module is the modified version of the simplex optimization technique. In the first iteration, the algorithm compares the changes in the free simulated variables to their measured values. The simulation and iteration process repeats until the minimum least-squares error between the simulated variables and the measured values is obtained. The output of the rectification process is the set of simulated variables, including rectified measured values and other calculated variables not available from measurements.

The mathematical description of the minimization problem is identical to that of a classical data reconciliation procedure, with the difference that the constraint of the minimization problem is not the underlying process model, but rather is user-defined:

Equation 2.5: Formulation of simulation driven data reconciliation problem

In the frame of the method implementation, a minimization of the weighted square error between the simulated values, which are necessarily balanced, and the corresponding measured variables, which are inherently unbalanced:

$$min_{x_i} \sum_{i=1}^{n} w_i \left(\frac{y_i - x_i}{span_{x_i}} \right)^2$$

$$\text{s.t.} \quad \mathbf{f}(\mathbf{x}, \mathbf{z}) = 0$$

$$\mathbf{g}(\mathbf{x}, \mathbf{z}) \geq 0$$

where:

x_i = reconciled value;

y_i = measurement – free variables (FV);

w_i = weight;

z_i = non-measurement variables - computed variables (CV)

$span_{xi}$ = normal operating span for variable x_i ;

The vector x is subject to constraint equations, i.e. mass and energy conservation laws and specified inequality constraints. The objective function is added to the simulation to be reconciled using mathematical functions native to the simulator used. The iterative search for values of x is performed using optimization module based on a simulated annealing version of the well known simplex optimization algorithm. Furthermore, normalized values are used to calculate the objective function due to the variety of physical units met in pulp and paper operation, for instance volumetric flows which can be in the thousands and mass fractions which are between 0 and 1.

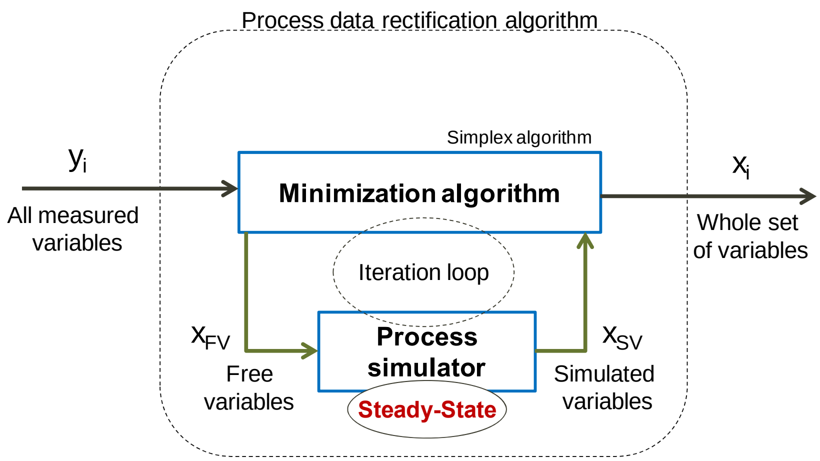


Figure 2.6: Iterative algorithm in simulation-driven process data rectification

In summary, the extensive work done in the literature is focusing on the improvements of solving efficiently the optimization problems related to all types of data reconciliation: linear, bilinear, non-linear or dynamic data reconciliation problems. These applications are suited to high redundant and only small industrial sub-systems, such as in chemical and petrochemical, pharmaceutical and in some cases mining industries. However, plant-wide applications are challenging in industrial concept due to the process dynamics involved (Bagajewitz, 2001). Dynamic data reconciliation is very computationally expensive for on-line industrial applications (Benqlilou, 2004) and plant-wide steady-state data reconciliation creates errors due to process dynamics. Papermaking industry is a special manufacturing environment characterized by both of these types of challenges: low redundancy (if only small amount of measurements are available not allowing performing reconciliation on the overall system) in measurements and highly dynamic processes. To the authors' knowledge, a methodology that would provide plant-wide and steady-state data sets in low redundant systems has not been addressed.

2.3 Signal processing

2.3.1 Introduction

Signal processing is an area of research in many fields, such as of electrical engineering, control and systems engineering, and applied mathematics that deals with operations for both, discrete or continuous time. Signals can represent various elements, such as images, sounds, continuous measurement data and sensor values, for instance astronomical data representing series of solar flares, control system signals, transmission signals, and many others. Signals can be analog and digital electrical representations of time-, spatial- or time-spatial varying physical quantities. The first signs for principles in signal processing date to the development of classical numerical analysis techniques of the 17th century (Schafer, 1974). Further digital refinement of the first ides of time series analysis is associated with the digital control systems of the 1940s and 1950s.

In the continuous industries signal processing helps to analyze time-frequency signals that are being measured by sensors and captured by the information management systems at the facility. The main goal of their application in everyday practice is to provide more accurate measured data for plant control. There are several types of errors that are affecting the usability of process signals (Figure 2.7, Bellec, 2004):

- Random errors (measurement fluctuations, process noise)
- Outliers or abnormalities
- Systematic errors (gross errors)

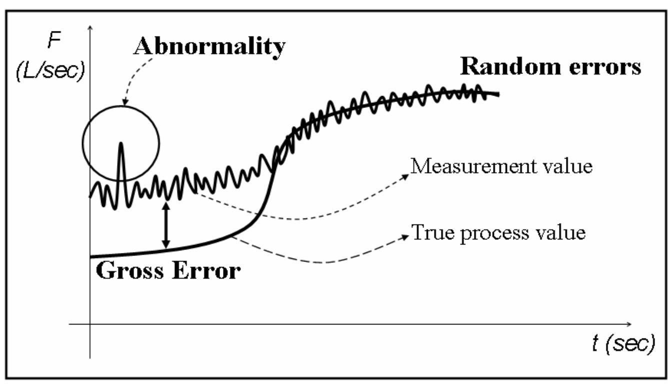


Figure 2.7: Errors in measurements

The random errors are assumed to follow normal distribution and are usually caused by the external disturbances such as equipment vibration or ambient conditions, but can be also caused by the measurement process itself. Typical outliers are caused by electricity fluctuation or wiring problems. Systematic errors, on the other hand, are associated with bias and are caused by for example drift in electricity supply, instrument miscalibration, incorrect instrument installation or process leaks. The presence of random or systematic errors if not processed will certainly lead to poor decisions, which will adversely affect all information in other business layers. Data processing, data reconciliation and gross error detection and estimation techniques deal with the errors elimination/minimization from process measurements.

2.3.1.1 Information technology and management systems in P & P mills

Pulp and paper facilities are more and more integrating process information systems (Enterprise Resource Planning (ERP), Enterprise Asset Management (EAM) and Process Information Systems (PIS)) that are helping them to improve managing their activities such as purchasing and asset management, plant monitoring and maintenance, production planning and scheduling, inventory management, shipping and customer service, and companywide financials. The advantage of using these systems has been recognized in 90s by pulp and paper facilities based on the successes reported by chemical companies (Scharpf, 1999). However, the gaps between the functionality of these systems and the papermaking needs prevent their application. This is mostly due to the particular nature of the pulp and paper manufacturing processes. Couple of surveys has been carried out in order to provide information on how the growth in information technology and management systems is exploited in the pulp and paper industry. A review by Fadum (1996) points out on the limited use of process information systems only for troubleshooting and that this system should be exploited more to enhance production profitability and product quality. Shaw's (1999) survey points out on the indeed increased availability and usage of acquired data, however no advanced use of these data was reported by surveyed users (engineers, IT and mill managers). Yeager (2000) has shown that today's real-time data availability promotes possible development of decision-making tools, which would allow mill personnel to react promptly. A recent detailed survey (Janssen et al, 2003, 2004) concludes that the interpretation of data from data management systems is on its ascent in pulp and paper mills. Different types of errors, that are present in the real-time process data, are the main reason why these systems are not being exploited to their full extent. For this purpose various pre-processing techniques are used. In the pulp and paper facilities it is done using various types of filters. Use of the analog or digital filters that are usually incorporated in distributed control systems is not however, sufficient to eliminate the effect of abnormal measurements completely. Some larger outliers can be eliminated using the permissible lower and upper bounds of process variables;

however, many outliers pass unprocessed and remain within these bounds. The presence of abnormalities decreases the performance of any process-state identification system (Shankar, 2000). In order to eliminate this problem, data pre-processing must be applied. This can be termed as a process of converting real-time measurements from manufacturing operation into useful operational knowledge. In order to obtain a decent steady-state data set, measurements need to be cleaned before steady-state identification can be performed. There are various techniques presented in the literature that deal with data processing and steady-state detection. Since this field is vast, only pertinent techniques are being presented in this thesis. Wavelet transform for data processing have been proved by many authors as the most robust technique for extracting the true process trend and omitting the random and abnormal errors (Jiang, 2000, Cao and Rhinehart, 1995; Benqliou, 2004; Flehmig et al, 1998; Nounou and Bakshi, 1999).

2.3.2 Wavelet transform

The recent development in continuous wavelet transform (WT) into time series analysis has provided its benefit also for process trend estimation in industry The measurement process signal *f(t)* (defined by measurement values) consist of two unobservable components, the desired process trend T(t) and process noise N(t) (stochastic component).

Equation 2.6: Process trend representation

$$f(t) = T(t) + N(t)$$

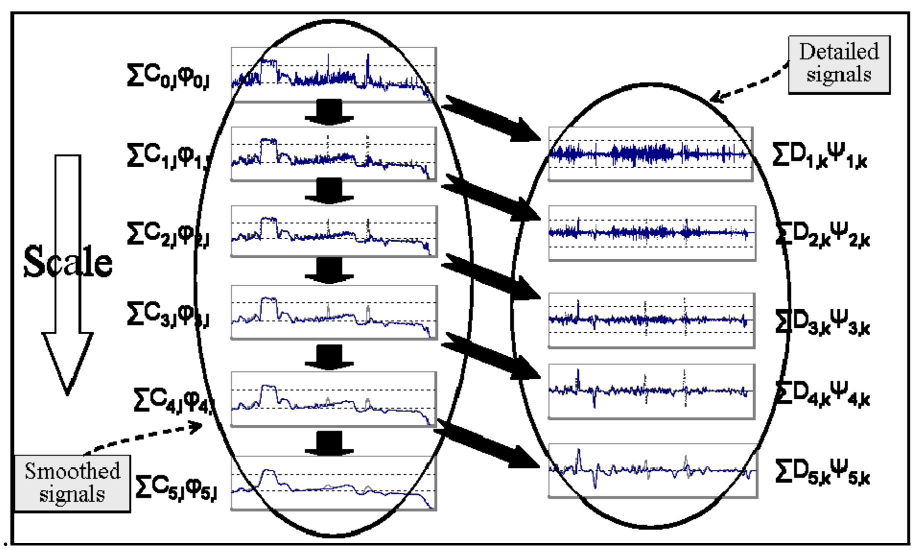


Figure 2.8: Multi-scale wavelet representation of real time measurements.

Wavelet transform technique is used to decompose this one-dimensional time series into two-dimensional time-frequency space. This feature is of use when a desired signal is corrupted by multiple events occurring at different locations in time and space (N(t) and T(t)). Jiang (2000) used this basic idea of process trend decomposition in a very efficient methodology for abnormality detection and denoising of the real-time process data. This process signal is decomposed into diverse frequency components at different scales (see Figure 2.8, Korbel et al, (a)). It uses the fundamental idea to represent the series of measurements as limit successive approaches at different frequencies:

Equation 2.7: Wavelet transformation of a signal

$$f(\mathrm{t}) = f(0) = \sum_{i\in I_0} c_{0,i}\varphi_{0,i} = \sum_{i\in I_1} c_{1,i}\varphi_{1,i} + \sum_{i\in k_1} c_{1,i}\psi_{1,i} = \dots = \sum_{i\in I_j} c_{j,i}\varphi_{j,i} + \sum_{j=1}^{J}\sum_{k\in k_1} d_{j,k}\psi_{j,k},$$

Where $\sum_{i\in I_j} c_{j,i}\varphi_{j,i}$ is the smoothed signal representing the low frequency part of the original signal with so-called mother coefficients $c_{j,i}$ and $\sum_{j=1}^{J}\sum_{k\in k_1} d_{j,k}\psi_{j,k}$ is the detail signal representing the high frequency components with father coefficients $d_{j,i}$. Individual isolated components are then analyzed and modified by altering its coefficients $c_{j,i}$, $d_{j,i}$ to $c'_{j,i}$, $d'_{j,i}$. This way the wavelet transform has the features to denoising signal from different frequencies by thresholding the coefficients below given threshold value. The validated ones then serve for process trend reconstruction using the inverse wavelet transformation. The thresholding is the process of discarding values that are below a threshold and keeping values over the threshold, which is often used in data compression or image processing. This process of tuning is done by applying the wavelet algorithm to each measurement separately.

2.3.2.1 Abnormality detection

Abnormalities are errors represented as large changes at high frequency or can be defined as high amplitude peaks of short duration. Such changes in real time can be detected using the first order WT, which is proportional to the first derivative of the smoothed signal (Equations 2 and 3). Since the extrema of the first derivative indicate fast changes in the function under study, one can detect such changes in a set of measurements using the first order WT (Jiang et al 2003a) and remove them from the process measurements.

For the first order WT:

Equation 2.8: Derivative wavelet signal

$$\psi_j(t) = \frac{d\varphi_j(t)}{dt}$$

Equation 2.9: Smoothed wavelet signal

$$WT_j f(t) = f * \varphi_j(t) = f * (2^j \frac{d\varphi_j}{dt})(t)$$

Abnormality detection is very important task in data processing due to the fact that if pre-processing via filters is applied the spikes will distort process trend (Narasihman and Jorache, 2000). Bakshi and Stephanopoulos (1993) proposed a wavelet based approach for multiscale extraction of trends, which is capable of characterizing different process features according to the corresponding information varying with successive scales. The method proposed by Jiang (2003a) and used in this study is particular for identification of abnormalities at a single scale, usually at the finest scale (time sampling scale).

2.3.2.2 Data pre-processing using wavelets

Pre-processing raw measured data by means of trend analysis involves a de-noising of data and elimination of abnormal data in measurements which in turn leads to better estimation accuracy. Wavelet de-noising utilizes the temporally redundant information of measurements. These trends are theorized to be more accurate than their measurements though they are usually inconsistent with underlying process model; therefore reconciliation has to be employed to resolve this conflict. In a way, the wavelet noise elimination creates measurements obtained by more accurate instruments (Benqliou, 2003). It can be also argued that wavelet based trend de-noising equalizes the uncertainty in process measurements with different standard deviations

Complete removal of the unsuitable high frequency features will be achieved if the correct cutting scale is employed. According to Jiang and al (2003a), the optimal choice of scale for signal denoising is based on the process dynamic, which is relatively easy to approximate off-line but cannot be predicted in real time and thus is not useful for a real time application. The technique proposed in the present paper bases the scale choice on historical data, and assumes that the scale choice is time invariant. The inconsistencies related to that assumption will be removed by using a low pass filter in subsequent steps of the applied methodology. As discussed previously, choosing a WT scale that is too high creates a distortion of process measurements, and leads to an inaccurate reflection of process trends. On the other hand, choosing a scale that is too low will leave the smoothed signal dominated by noise and unsuitable temporal features. This second possibility does not affect the true process trend. At an under-evaluated scale, the process trend is still available, but it is corrupted with high frequency measurements. Therefore, in a subsequent step, the corrupted smoothed signal can be refiltered to isolate the process trend from higher frequency perturbations. This opportunity is not possible in the case where the scale is over-estimated due to the distortion created in the signal.

In order to keep the true signal properties intact, one should choose a scale that does not affect the process trend. By studying historical process measurements, one can investigate the optimal cutting criterion for different process operations. The scale employed for on-line implementation can be selected in such a way that high frequency features are mostly deleted, and the true process trend is not affected by signal distortion. To do so, one should test the performance of the optimum cutting scale (for off-line data treatment) proposed by Jiang et al (2000a) on historical data and compare it to the previous scale (filtered data at lower frequencies).

The theoretical formalization of threshold in the context of removing noise via thresholding wavelet coefficients was presented by Donoho (1995). This method estimates threshold by

Equation 2.10: Wavelet threshold estimation

$$j = \sigma(2\log N)^{1/2}$$

where N is the size of the wavelet coefficient arrays and σ is the noise standard deviation. The rationale for this choice is that the matched filter is theoretically the optimal detection filter. This condition is best suited only for stationary white noise. Recently, some new methods have been presented, which estimates threshold according to wavelet coefficients at different scales (Jiang 2003a).

2.3.3 Process steady-state identification

In order to generate near steady-state data set candidates from real-time measurements, signal processing technique must be employed. Several methods for on-line process status identification have been presented in the literature based on statistics or filtering (Cao and Rhinehart, 1995, Bakshi and Stephanopoulos, 1993), which creates data distortion when abnormalities are present. With the advent of wavelet transform theory, the signal processing field has evolved into more multidimensional analysis of trends allowing for multiscale and accurate representations of functions. Flehmig et al (1998) explored the wavelet transform features to approximate process measurements. Nounou and Bakshi (1999) used wavelet features to identify and to remove random and gross errors. Recently, Jiang et al (2003a) proposed a wavelet method for the detection of near steady state periods. These methods are used for offline signal representation. The wavelet data processing can be used efficiently to eliminate random noise and abnormalities, and simultaneously analyze the trend for steady-state occurrences. False detection of the process steady state can lead to misinterpretation of true process features, especially if the incorrect steady state data are subsequently reconciled. Under-estimating the true process steady state periods can lead to only partial correction of gross errors (Figure 2.9a and Figure 2.10), while over-estimating steady state periods can result in false input to data reconciliation (Figure 2.9b).

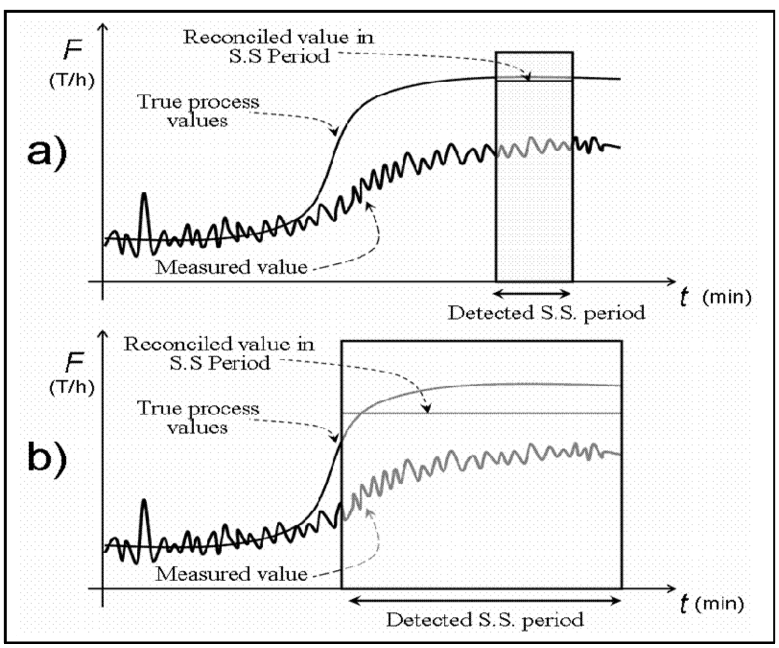


Figure 2.9: Inaccurate estimation of steady state periods (Korbel et al, (a))

a) Under-estimated period b) Over-estimated period

A variety of techniques for on-line process status identification have been proposed in the literature. Bakshi and Stephanopoulos (1993) developed a geometric approach for the description of process trends. Cao and Rhinehart (1995) proposed a steady state identification technique based on the comparison of data variances calculated in different ways. In this method, a weighted moving average is used to filter the sample mean. Then, the filtered mean square deviation from the new mean is compared with the filtered squared difference of successive data. This method uses a low pass filter to estimate the mean value. On the one hand, the computational requirements and storage are significantly reduced. On the other hand, low pass filters are less sensitive to the presence of abnormal measurements. Furthermore, using a weighted average to filter the calculated variances creates a delay in the characterization of process measurement frequency. These delays can cause detection problems in periods where the signal properties vary in real time.

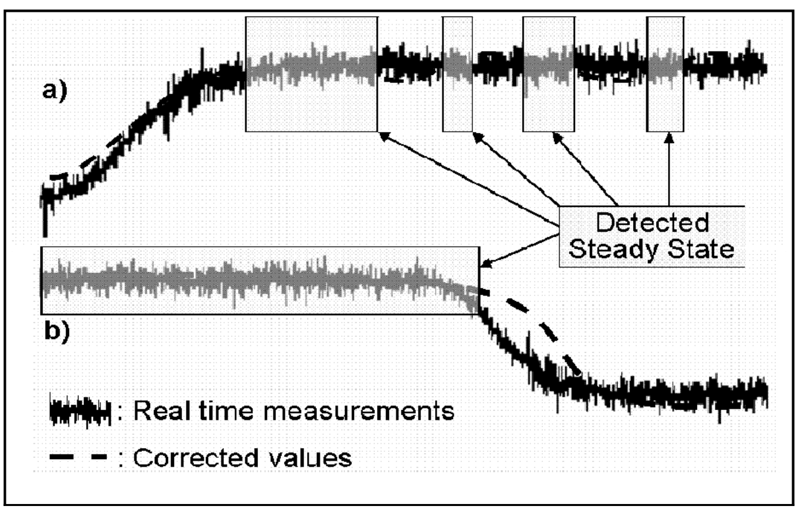


Figure 2.10: Multi-scale decomposition of real-time measurements

a) Impact of under-scaling on steady state identification (wavelet transform is corrupted by high frequency noise)

b) Impact of over-scaling on steady state identification (extensive smoothing of process signal, creating distortion of process state representation)

Flehmig et al (1998) used wavelet transform features to approximate process measurements by a polynomial of limited degree and to identify process trends. Nounou and Bakshi (1999) used wavelet features to identify and to remove random and gross errors. More recently, Jiang et al (2003a) proposed a wavelet transform (WT) based method for the detection of near steady state periods. The wavelet based multi-scale data processing technique was used to eliminate random noise and abnormalities. Then, the process status was analyzed according to the modulus of the first and second order wavelet transforms. This method can accurately analyze high frequency components and abnormalities. When applying the multi-scale method, the accurate choice of scale is critical. If the scale selected is too low, the WT will be corrupted by high frequency noise, i.e., process status identification is corrupted by temporal features. If the scale selected is too high, then process measurements are excessively smoothed, which creates distortion in the process signal. This creates a deviation from the true process trend and leads to an incorrect reflection of process status.

Jiang et al (2003a) proposed selecting the optimal scale (the scale at which the most of the high frequency noise is removed without distorting the actual process trend) by taking into consideration the response time constants and sampling intervals. This criterion is adequate for off-line purposes, but is not practical for on-line treatment of real time data because on-line measurements can be corrupted with different high frequency features over time. Therefore, the scale choice must be known a priori for on-line wavelet-based treatment of real time data. Furthermore, this method uses the second order WT of the signal to distinguish zero-crossing points from steady state periods (steady state is detected when WT is near zero however fast changes in the process variable correspond to a zero value of WT). The second WT is directly proportional to the second derivative of the smoothed signal at the sample cutting scale. It is adequate to represent process trends but requires great computational speed and storage. Finally, in the so-called direct approach, linear regression of the measured values is calculated over a data window, and a t-test is performed on the regression slope. This approach is executed over a specified time period, which is not ideal when dealing with real time measurements.

2.4 Gaps in the body of knowledge

Based on the pertinent literature review the following holes in the body of knowledge were identified:

Signal processing

On-line signal processing method for detection of steady state operating conditions based wavelet transform is missing. Furthermore, there is no industrial application of on-line signal processing technique that would provide automatically process data sets representing near steady-state operating conditions of a small subsystem and a plant-wide manufacturing in the pulp and paper facilities and simultaneously pre-process real-time data from random errors and abnormalities.

Plant-wide data reconciliation

For higher level analysis it is critical to ensure plant-wide data reliability and consistency. There is no systematic methodology in the literature that would address plant-wide and steady-state data reconciliation in the low redundant industries such as pulp and paper sector. The information gained from having multiple plant-wide data sets would provide enhanced representation of manufacturing operation for higher level applications, such as cost analysis. Furthermore, there is no evidence in the literature about an industrial application for plant-wide and steady-state data reconciliation coupled with on-line wavelet technique for process state identification and process data cleansing in the pulp and paper industry.

Advanced manufacturing cost analysis

A critical element of a facility business model is the cost modelling methodology and how it uses process knowledge. The product-based cost analyses resolving operating regimes in the pulp and paper industry have not yet taken place.

There is no manufacturing cost methodology capable of providing information on actual product margins in the pulp and paper industry that exploits real-time reconciled process data in combination with mill cost data. Furthermore there is no activity-based costing –like methodology that systematically assesses the impact of retrofit implementation of new production facility on the current core business manufacturing costs that are based on real-time data measurements.

CHAPTER 3 OVERALL METHODOLOGICAL APPROACH

3.1 Overall methodology

Generally, manufacturing facilities such as pulp and paper mills are creating goods according to production recipes. Within individual recipes, different operating regimes are chosen by operators to produce goods in optimal and safe manner or different operating regimes are selected as a respond to external perturbations (change in raw material quality). These operating regimes are characterized by control set points and/or the type of production assets utilized. These different production set-ups result in potential resource consumption rate variances. These changes can be assessed and interpreted by the proposed methodology (Figure 3.2) and thus validate the main hypothesis.

The methodology that will be able to capture production cost variances due to changes in operating conditions consists of three main steps:

1. Pseudo steady-state detection of a process operation
2. Plant-wide steady-state data reconciliation
3. Operations-driven cost analysis

The scope of the process and cost analysis must be first defined by clear characterization of operating regimes. The time boundaries are selected (starting and ending points) for process data analysis that correspond to duration of production campaign runs and operating regimes (Figure 3.1). Since the operator action to external perturbations or desired changes in operating set-up is aimed to a steady-state condition, the operating regime must be represented by steady-state data sets. The transient periods that are inherently occurring between individual steady state conditions cannot be accounted for cost assessment of regime because their duration and variation is never occurring the same way for the same operating regime. Cost analysis based on such data would not represent operating regime correctly. The aim of the method is to characterize periods of operation that we may take an action upon to enhance cost profitability. However, as stated earlier, some of the regimes occur as a respond to external perturbations and cannot be avoided. On the other hand an action to these perturbations is answered by different operating set-ups and hence the cost representation of them would be essential to understand.

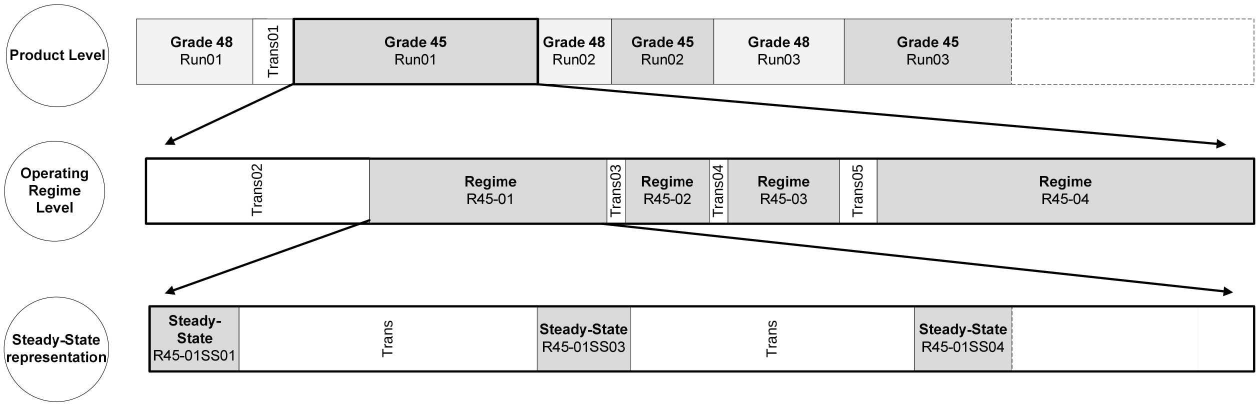


Figure 3.1: Different levels of analysis

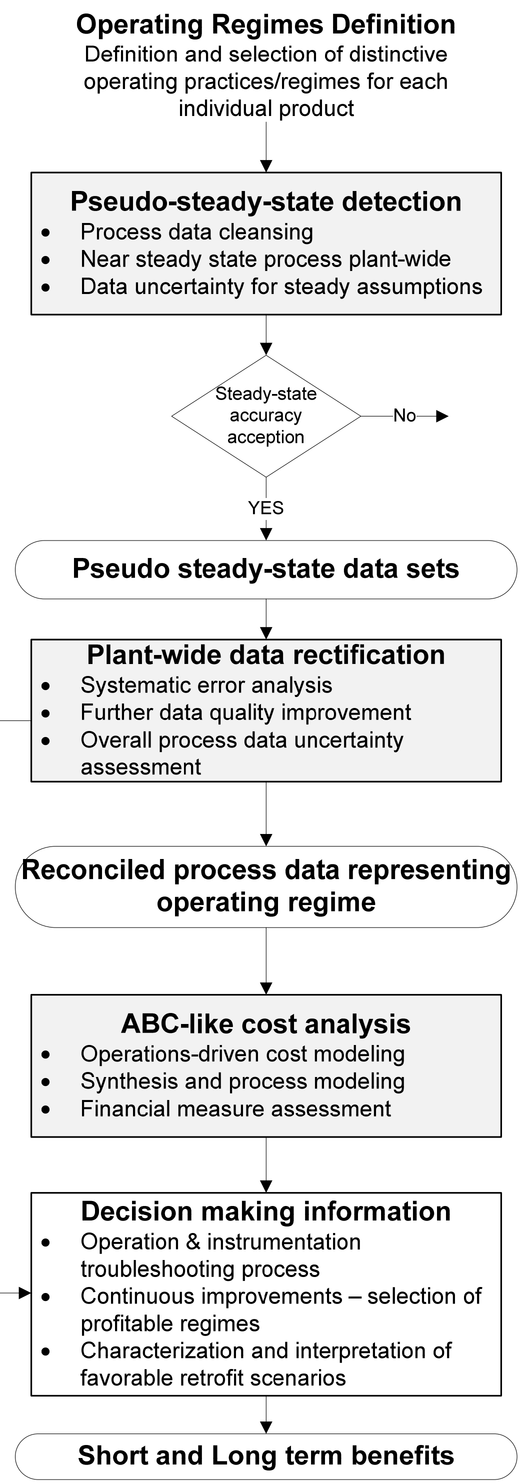


Figure 3.2: Overall methodological approach for production cost assessment of operating regimes

Following sections describe each step of the overall methodology in more details as a block of activities:

3.1.1 Pseudo steady-state detection

Generally, if near steady-state conditions of an operation need to be detected in paper making facilities, the measurement process data must be analyzed manually. Often, only the average of the relevant variables within the analyzed time framework are calculated, and used for the needs of the analyst. This off-line and conventional process analysis is very laborious and often does not provide accurate data for optimization or advanced cost accounting (Narasimhan and Jordache, 2000). Conventional on-line techniques are based on filters and have difficulties with abnormal measurements. The use of the wavelet transform of signals and data filtering is tested for the purpose of an on-line industrial application for plant-wide steady-state detection.

First, the proposed technique is used to pre-process and analyze the time-frequency space of each measurement (within each individual operating regime). This step in the proposed methodology is done to get rid off high-frequency noise and most process abnormalities in the real-time process data in order to enhance the detection of near steady-state operation. The performance of steady-state identification is tested by comparing the outputs of different conventional methods (Korbel et al, a) on a small sub-system of a newsprint operation.

For each measurement variable, the knowledge from historical data (from information management systems at the mill) is used to tune several parameters of the method (wavelet transform and filtering). This is done by applying the wavelet algorithm to each measurement separately. The set of optimal parameters (the wavelet parameters of each measured variable that provide the most robust performance of processing step possible) is then stored in the database for testing the on-line applicability of the method for the small sub-system and then plant-wide operation. The detection of the optimal parameters is described in Appendix A (Korbel et al, (a)). It can be summarized in four critical steps:

- Select independent and comprehensive variables for determining near-steady states;
- Establish near steady-state criteria for each variable and for the system as a whole;

- Determine the minimum length of the steady-state periods according to process system delays;
- Determine criteria for threshold values for steady-state periods (for each variable and for the system as a whole).

This step requires profound engineering analysis of the underlying process dynamics to represent the true process trend correctly by the wavelet transformation.

On the one hand, the ability of detecting near steady-state operation of larger systems (plant-wide operation) strongly depends on the process dynamics involved. On the other hand, the characterization of the plant-wide data sets with the steady-state assumption depends on the judgment of the analyst. The plant-wide operation is divided into smaller sub-systems in order to analyze process dynamics for each. This division of the mill operation into small sub-systems corresponds to the third and forth steps of the methodology, where these subsystems are defined as cost centers of production activities in order to enhance tracking of changes in resource consumption. Sensitivity analysis of the process variables for each part of the mill is performed. The variables that have a major influence on the process sub-system dynamics are selected and only these key variables are used for steady-state detection thereafter. The parameters of the method that allow for detection of a steady state (Appendix A) are tuned, to enhance the frequency of steady-state identification within each subsystem. This way representation of near-steady state condition for plant-wide operation can be selected more frequently. The validity of these data sets is then analyzed using steady-state data reconciliation (next step described in details in the next section). If the assumption of steady-state is wrong, the discrepancies in the data reconciliation results will manifest as a high magnitude of least square error, indicating that the given plant-wide data set is not valid (i.e. the process can not be assumed to be at steady state or a large bias measurement is present). Such data sets are omitted from further use.

The results of the on-line signal processing technique were multiple data sets representing plant-wide, near steady-state operation. These are further used in data reconciliation for further enhancement of the manufacturing knowledge they represent.

3.1.2 Plant-wide steady-state reconciliation

Papermaking facilities are known of having a very poor instrumentation network. Only the necessary instruments for control and safety are installed. In order to validate the plant-wide data sets and identify if certain measurements exhibit biased values, data reconciliation must be employed. Conventionally in the pulp and paper facilities, several lengthy manual tests are necessary for reconciling process data due to a lack of instrumentation redundancy. Some small parts of the modern mill operations can be reconciled using data from information management systems using averaged data, however with the need of the analyst (chemical engineer) to search for good periods (periods with relatively steady operation) that would provide good estimates.

The use of a simulation model using CADSIM software (section 2.2.4.1), coupled with the wavelet steady-state detection technique is tested for on-line, plant-wide application of steady-state data reconciliation. Parallel to the simulation-driven approach (presented and discussed in literature review: section 2.2.4.1), conventional data reconciliation is used (where applicable) to compare the outcomes and thus validity of the method (Figure 3.3). This strategy of comparing two different techniques is assumed to validate the approach and highlight the variance in outcomes. For instance the difference in using only average values within operating regime duration instead of characterizing regime by steady state is essential to manifest. This way the use of proposed method will manifest its advantages and contributions to engineering field.

First, the process simulation model is built using information from flow sheets of the actual production and close interaction with mill personal (operators and process engineers). Several empirical equations (for instance back steam generation from operation of high consistency refiners as a function of electricity demand) based on years of practice were added to the simulation model, thus decreasing the degrees of freedom (increasing system's redundancy).

In order to establish a weighting matrix for the optimization module (optimization engine of the simulation-driven data reconciliation, presented in the section 2.2.4.1), the historical data for each instrument is used. The knowledge gained from the analysis from the previous data processing step is used. The standard deviation for each measurement is calculated. This knowledge coupled with years of experience (close collaboration with mill personal) in the given manufacturing environment is used to express the trust (the selection of weighting procedure is described in the Appendix I, the sensitivity testing of this choices is discussed in Appendix J) in each individual instrument. The system of data reconciliation coupled with wavelet pre-processing is tested iteratively, which allows for better tuning of the weights in time.

The outcomes of the data reconciliation are updated and stored in the information management system for each operating regime. The least square error and the variance between individual measurements are compared to these values in situations when biased measurements are detected. This allows for systematic detection and estimation of the gross error (biased measurement) and can assist in calibrating the faulty instrument.

The results of plant-wide steady state data reconciliation represent the process characterization of each individual operating regime. These are further used in the operations-driven cost analysis step that is described in more details next.

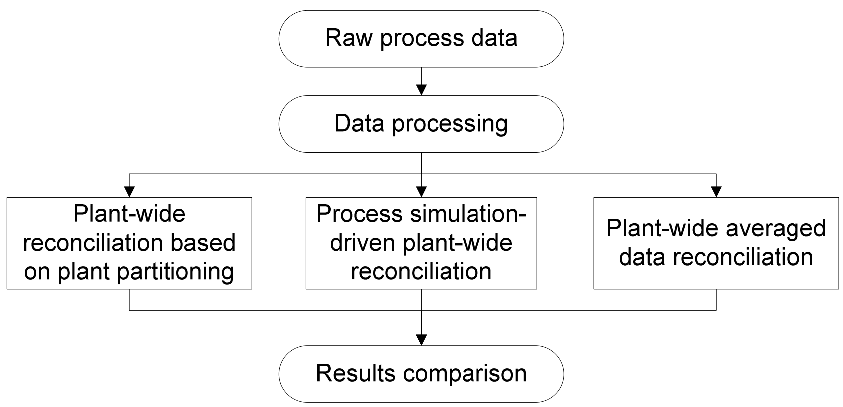


Figure 3.3: Methodological steps in plant-wide data reconciliation

3.1.3 Operations-driven cost analysis

The papermaking production processes can be dissected into different levels, according to the representation in figure 3.1. First, traditional accounting techniques (standard and actual costing) are used to estimate the manufacturing costs of the current production. The same practices are also applied to estimate the manufacturing costs of the analyzed future retrofit scenarios (presented in the following section – case study). Three different retrofit scenarios (described more in detail in the following section: Case study) have been chosen for analysis based on the case study mill's preferences. The information on production costs based on traditional techniques is used as a benchmark to compare the outcomes of the proposed operations-driven cost methodology.

The steady-state process simulation model and real-time data are used to define resource and activity drivers of the cost model for current business and for the forest biorefinery options.

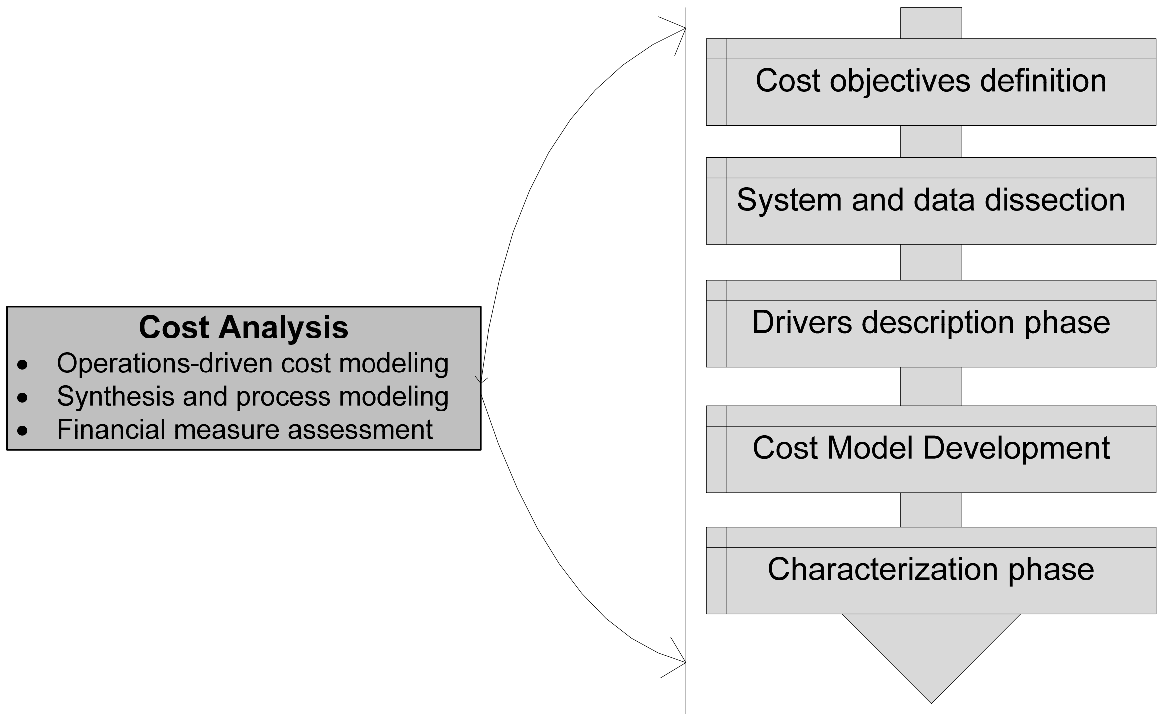


Figure 3.4: Methodological steps for building operations-driven cost modeling framework

The focus of cost-model development is first to characterize the direct and indirect manufacturing costs of the studied facility (for both, current core business and integrated forest biorefinery strategies – Figure 3.5) in order to identify the most profitable and most costly operating regimes. This information can be interpreted to define cost variances. The five methodological steps for cost model development and analysis are (Figure 3.4):

1. **Cost objectives definition** - The first important step is to clearly define the scope and the objectives of the cost analysis procedure. It is necessary to develop guidelines for identifying relevant cost items and for characterizing the desired cost behavior for use in decision-making.

2. **System and data dissection** - The production system is divided into smaller subsystems to enable the qualification and quantification of different types of cost drivers. This step increases the in-depth cost analysis capabilities of the system. The separate subsystems are called Process Work Centers (PWCs) and are further divided into individual Processing Units for increased cost-tracking transparency. The system division step is based on rules derived from the first and second step of the overall methodology (Korbel et al. (b)). Then, individual production runs of a given grade can be dissected into operating regimes at this stage, or additional production regimes can be identified and added to the cost analysis.

3. **Driver description phase** - This step involves intensive discussion with mill personnel to identify the cost drivers. This phase is of critical importance because it structures the shape of the cost model and the characterization and interpretation of the results.

    a. *Resource drivers:* The characterization and measurement of the resource consumption rates of processing units and process activities are based on process data. For instance, flow measurement is a resource driver for a given flow medium.

    b. *Process activity drivers*: These drivers characterize the linkage between operating conditions and the consumption of a resource driver. This phase identifies what information is necessary to characterize the intensity of a process activity within a process work center. For instance, the pressure in a vessel will characterize the required rate of steam flow to be input.

    c. *Process work center drivers*: The boundaries of each PWC are defined in the second phase of the methodology. The interpretation must be intuitive to capture the cost-insight capabilities of the method clearly in a graphic user interface. The aim is to explain cost generation better at a mill-wide level. One of the important cost centers is the overhead work center, where the drivers must be clearly defined to achieve indirect-cost transparency throughout all mill departments. For instance, the work center driver for maintenance is the head count for a given subsystem of the operation.

4. **Cost model development** - The model development follows the operations-driven cost modeling framework presented in Laflamme-Mayer et al. (2011). The supporting pillar is the integration of process and financial information based on ABC-like principles. The systematic consideration and cost aggregation of individual production processes and their operating condition into the plant-wide manufacturing operation are essential principles of this stage.

5. **Characterization phase -** The last phase involves the characterization of the costs incurred and the interpretation of the results based on the objectives defined in phase one. Process understanding is the key element at this stage. Therefore, interaction with mill personnel is necessary to interpret the results. Sometimes, steps 2-4 will need to be repeated to arrive at a satisfactory level of in-depth cost understanding. The results of this stage, when validated, can be clearly visualized and used for decision-making support.

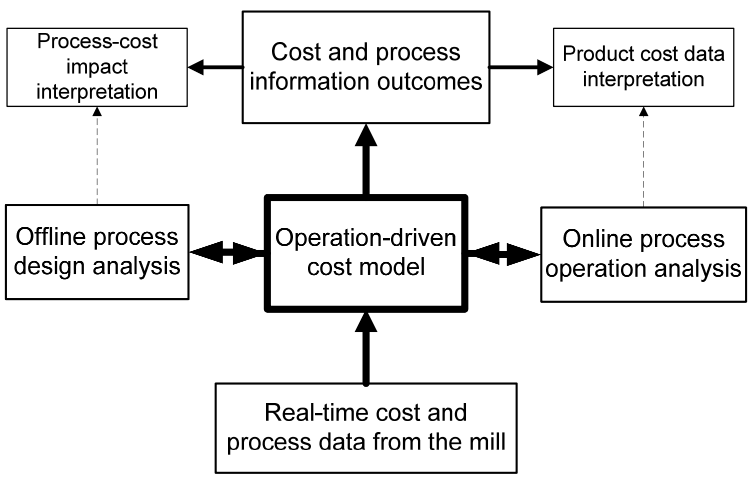


Figure 3.5: Dataflow between the operation-driven cost model and process operation and retrofit design analysis

3.2 Case study – Advanced manufacturing cost analysis of current and retrofit design process operations

3.2.1 Background

3.2.2 Product cost assessment of a core business - base case

The base-case mill is an existing integrated newsprint mill. The thermo-mechanical facility produces different pulp qualities based on paper mill demand and specifications, with the throughput matched to that of the paper mill. Two newsprint products with different basis weights, 48 $g.m^{-2}$ and 45$g.m^{-2}$, are produced (Figure 3.6, Korbel et al (b)).

<u>Existing mill configuration</u>

The following manufacturing steps are involved in the base-case mill:

- One paper machine that is 8.4 metres wide built by Beloit in 1985 with a production rate of 680 tonnes/day of newsprint,
- One thermomechanical pulping line consists of a single series of Sunds CD-82 two-stage
- refiners and related equipment with an average production of 680 tonnes/day of pulp.

The following supporting processes are also part of the base-case mill configuration:

- A wastewater treatment plant processing 30,000 m^3/day
- A boiler plant producing 2500 GJ/day of steam
- A steam recovery unit in the TMP line, producing 3000 GJ/day of steam.

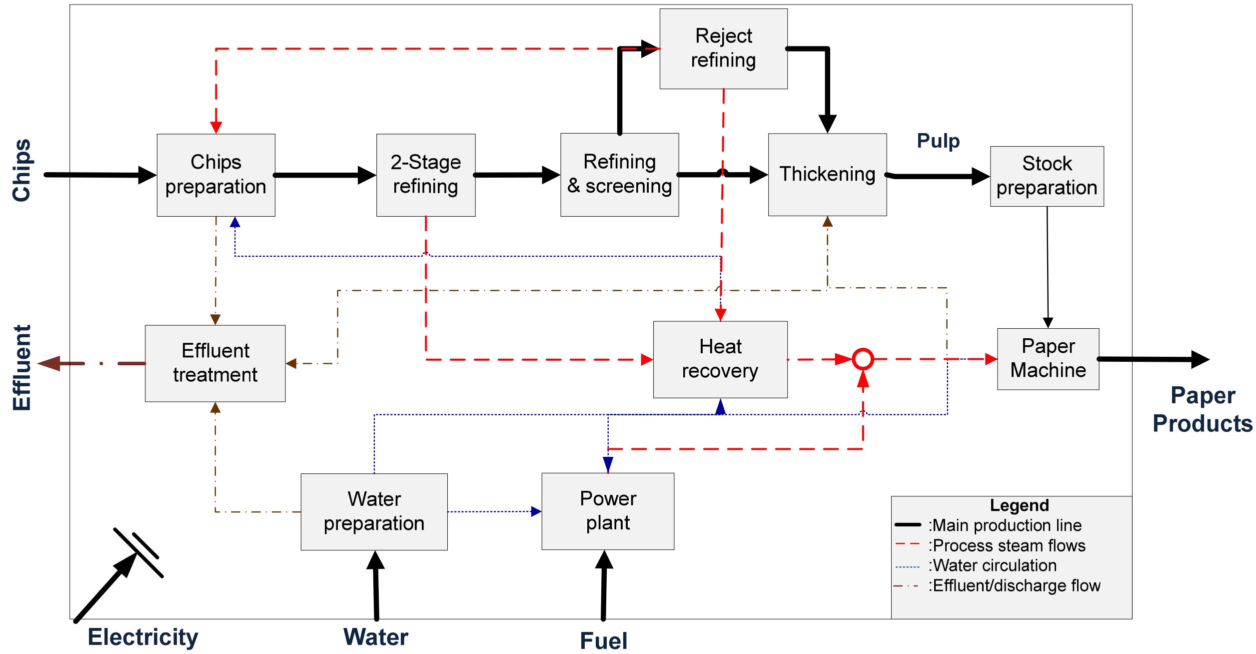


Figure 3.6: Simplified block diagram of a complex thermo-mechanical pulping process with paper mill (for detailed mill model presentation see Appendix I)

3.2.3 Product cost assessment of selected biorefinery scenarios

The case mill under analysis is a highly competitive newsprint mill (in the first quartile of manufacturers) with limited access to biomass. Hence, they have chosen to investigate a biorefinery strategy that could be integrated into their existing operations. Three major forest biorefinery retrofit options at an integrated newsprint mill were selected for production cost analysis:

- Cellulosic ethanol production: ~3000 gallons per day ethanol production from hemicelluloses extracted before pulping
- PLA production: 11.5 tons per day of polylactic acid (PLA) production from lactic acid extracted from hemicelluloses before pulping
- Biocomposite production: 80 tons per day of biocomposite pellets produced from the blending of TMP fibres and polypropylene.

The first two retrofit options are based on the sugar platform, i.e., sugars are the feedstock for production of these biochemicals. Ethanol and PLA products share the same process design up to the fermentation unit (Figures 3.7 and 3.8, Korbel et al (d)). The third forest biorefinery alternative is based on mechanical blending of TMP fibres with a plastic matrix to create a biocomposite material (Figure 3.9, Korbel et al (d)).

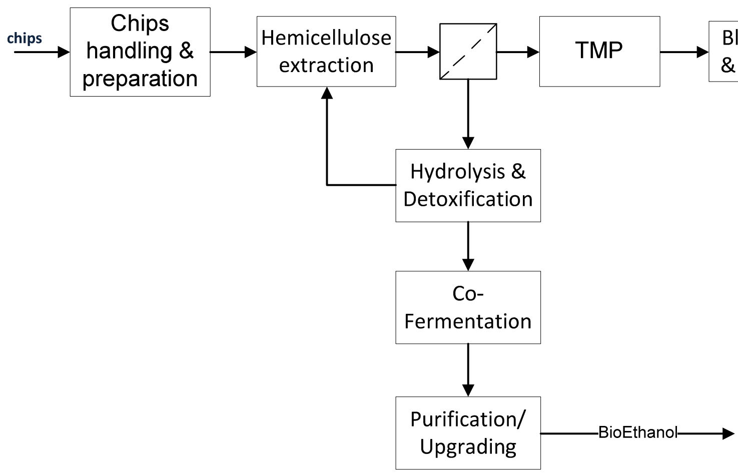


Figure 3.7: Simplified flowsheet of simultaneous ethanol production.

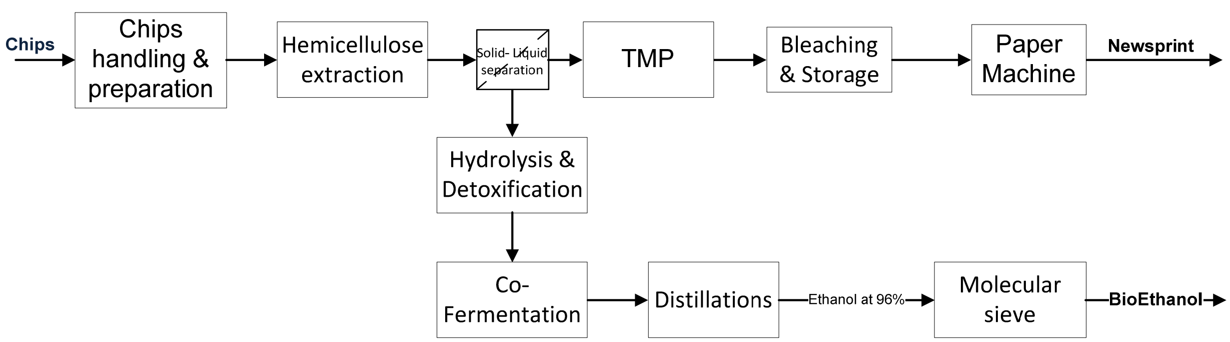


Figure 3.8: Simplified flowsheet of simultaneous PLA production

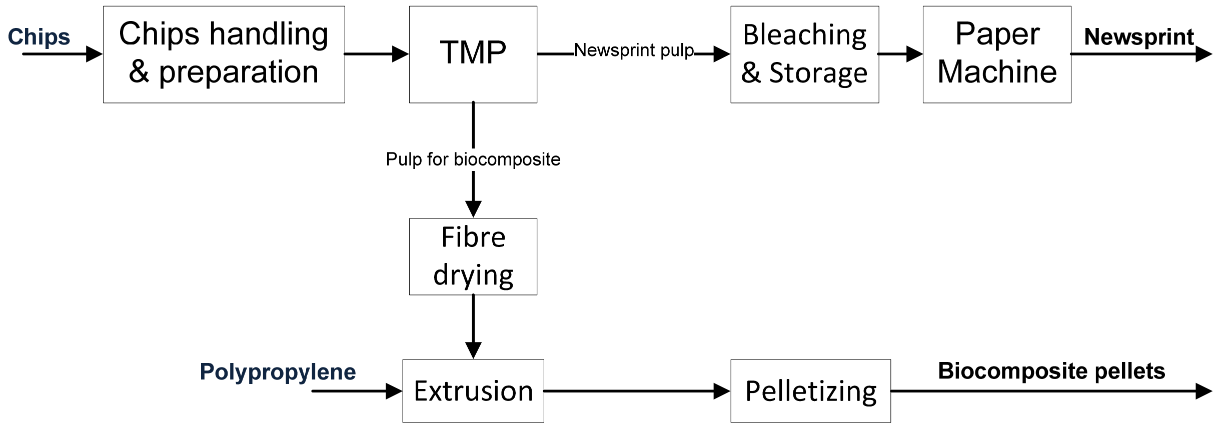


Figure 3.9: Simplified flow-sheet of simultaneous biocomposite production

CHAPTER 4 PUBLICATION EXECUTIVE SUMMARY

4.1 Presentation of publications

The following articles that are published in, in review or submitted to peer-reviewed scientific journals are presented in Appendices A to E of this thesis (Figure 4.1 illustratively presents links between publications).

- Korbel, M., Jiang, T., Stuart, P.R., “Steady state identification for on-line data reconciliation based on wavelet transform and filtering”, Submitted to Journal of Computers and Chemical Engineering

- Korbel, M., Wasik, L., Stuart, P.R., “Practical methodology for plant-wide process data rectification in the pulp and paper industry” Submitted to J-FOR journal

- Korbel, M. Stuart, P.R., "Assessment and interpretation of advanced cost data for process improvement in an integrated newsprint mill”, Submitted to International Journal of Production Economics

- Korbel, M. Stuart, P.R., “Product margin assessment for process cost-impact analysis”, Submitted to Industrial & Engineering Chemistry Research

- Korbel, M. Stuart, P.R., “On-line methodology for operations-driven cost assessment of operating regimes using real-time process data”, Submitted to TAPPI journal

Other complementary publications listed below are included in Appendices F and G.

- Korbel, M. Stuart, P.R., “Cost integration methodology and the forest biorefinery” Chapter of a book: El-Halwagi, Mahmoud and Paul Stuart, Integrated Biorefineries: Design, Analysis, and Optimization, CRC Press/Taylor & Francis (2012)

- Korbel, M., Garrigues, L., Stuart, P.R., 2007, “Process and Business Data Reconciliation in the Pulp & Paper Industry”, TAPPI Conference, Innovations in Engineering, Pulping & Environmental, Jacksonville, FL.

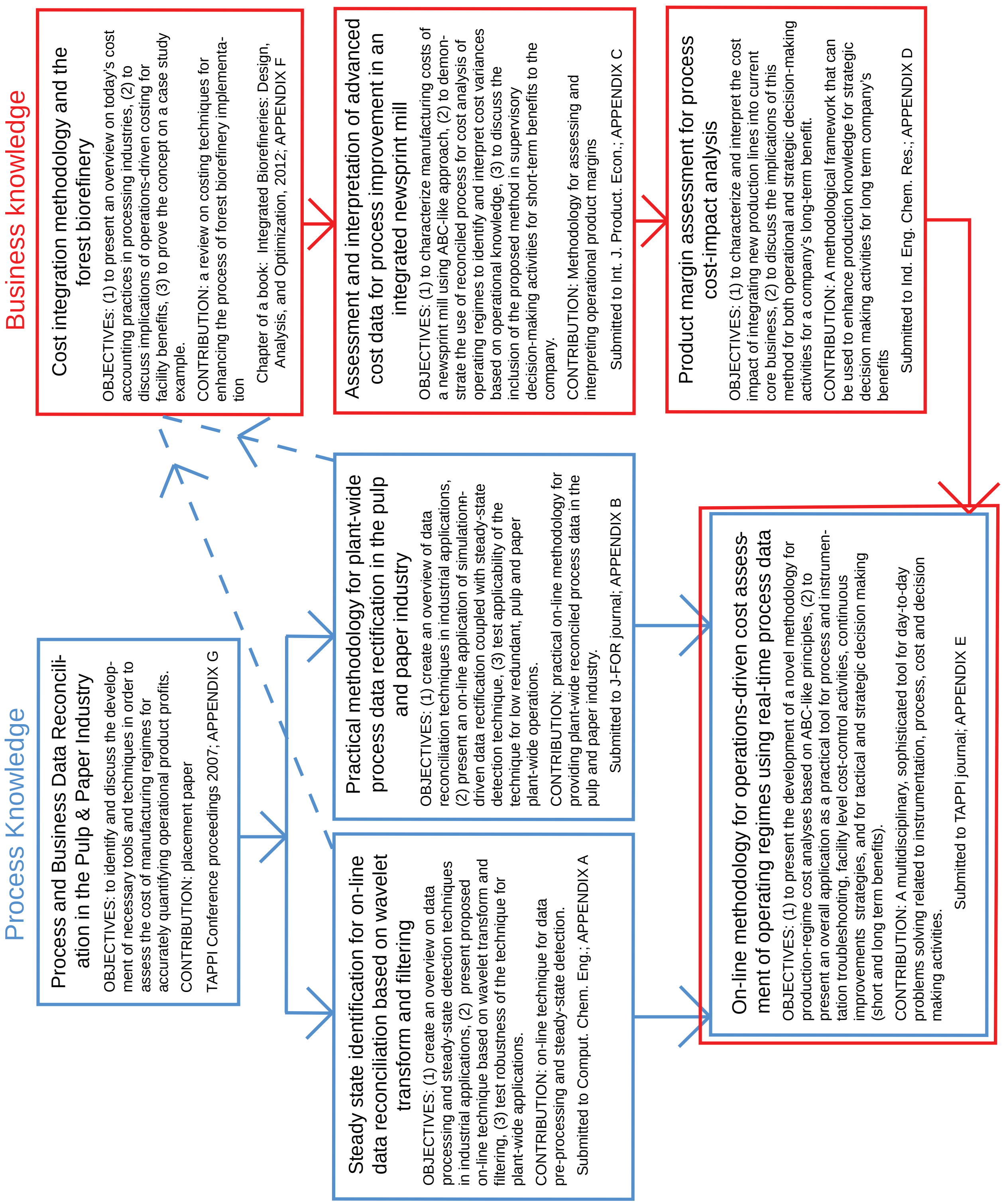


Figure 4.1: The line of thoughts between individual publications.

4.2 Links between publications

The following discussion links different segments of this Ph.D. work that were presented in individual publications. Figure 4.1 demonstrates the links discussed.

Theoretical background on data cleansing focusing on data processing and reconciliation, in the industrial context of this Ph.D. work, was reviewed in order to gather necessary information for making real-time data consistent with underlying process operations and available for higher level cost analyses. This review was presented during the TAPPI Conference - Innovations in Engineering, Pulping & Environmental in Jacksonville, Florida (2007) (Appendix G). A more detailed report on the practical value that can be obtained from the smart dissection of these process data by integrating them with financial information is reported in a chapter of a biorefinery design book (Appendix F). In the former publication, in addition to the theoretical background, a case study example on an existing newsprint mill was presented to concretize the value of smart data dissection. These two pilot papers are not part of the body of thesis but serve only as a reference.

The methodology that is necessary to pre-process the raw real-time data and identify process conditions on-line, is described in detail in Appendix A. Two case studies are presented in this article. The first case study, being a sub-system of the newsprint process operation, proves the capability of the method for on-line use. The comparison to conventional offline methods was carried out in order to validate the method's robustness. The results indicate that the method is capable to provide multivariable near steady-state data sets of high quality, in some cases with better identification capabilities than the conventional methods. The second case study work, being a large system (plant-wide operation), was done in order to 1) analyze the near steady-state probability of occurrence for different mill sub-systems and of the whole mill as a function of adjusting the method's parameters, and 2) address the accuracy of plant-wide pseudo steady-state assumptions and its influence on the final higher-level applications (advanced cost analysis in the context of this Ph.D. work). This work was submitted to the International Journal of Computers and Chemical Engineering and it is under investigation for potential application and implementation in the CADSIM simulation software for improved data accuracy.

The on-line identification of near steady-state operation has created the critical basis for plant-wide and steady-state data reconciliation in the pulp and paper industry. This manufacturing sector presents many challenges from making traditional data reconciliation possible. Therefore, a very practical and simulation-based data rectification method was explored for plant-wide data validation. This work is described in detail in Appendix B. It was found that the presented method overcomes redundancy issues at paper making mills and makes reconciled process data for large systems available for further use.

This new process information acquired by careful process data analysis was explored for advanced cost analysis. An operations-driven cost accounting model was developed, that uses for the first time, reconciled real-time process data for operational and retrofit design decisions. These two publications are described in detail in Appendices C and D. The results from the proposed cost modeling framework provide superior insight into production costs compared to conventional cost accounting techniques. Two case studies have been done to validate the above statement. The first case study focusing on the newsprint business presented the characterization and interpretation of the current production costs while making visible actual product margins. The second case study (Appendix D) takes advantage of knowing this valuable information for analyzing future company's opportunity in exploring/adding new business. Particularly, the focus was to analyze the cost impact of integrating new production lines into the core business facility.

Finally, the overall on-line operations-driven cost accounting methodology for short and long term company's benefits, which were discussed in detail in above mentioned articles, is united into one publication (Appendix E). This publication assembles all methods into one methodology, discussing every major step with case study results. The overall methodology is proposed to be used as a multidisciplinary tool for day-to-day instrumentation & process troubleshooting, cost assessment of operating regime (thus providing a framework for continuous improvements by selection the profitable ones), and use this information for tactical and strategic decision making activities at the facility.

4.3 Synthesis

The synthesis part of this thesis presents an overview of the most pertinent results of the work done in this Ph.D. research project to highlight the main values and deliverables from the proposed methodology. The overall methodology was proposed at the outset of the project and then systematically addressed as summarized in this section. The focus was on three essential aspects of the method after operating-regime scope definition: 1) on-line data pre-processing and near-steady-state identification, 2) plant-wide process data rectification, and 3) advanced cost analysis for true product margin assessment with short- and long-term benefits to the company. For each of these aspects, a comparison with traditional process data or cost analysis is performed.

Because the case study is based on an existing newsprint mill, the majority of the results are presented as normalized values due to the confidentiality requirements of this highly competitive commodity business environment.

4.3.1 Operating-regime scope definition

The first essential step in application of the proposed methodology is to analyze the process operation under consideration to determine its manufacturing options and flexibility. The various operating regimes or operating practices that create the final products must be clearly identified. In the base-case operation studied here, the focus was on analyzing the potential cost variances of the same grade "recipe," but with different operating set-ups characterized by 1) groups of process control setpoints corresponding to a given pulp quality, 2) the type and age of the refiner plates used, and 3) the volume of production throughput. Changes in control setpoint strategy were generally observed in response to two manufacturing situations to which process operations need to adjust:

- Mechanical causes: a change in raw material (wood chips) quality (humidity, dimensions, or variability) or a change in the ratio of different types of chips used (high or low ratio of softwood or addition of hardwood). The pulp and paper mill under analysis uses the same "recipe" to specify the range of fibre types and chemical requirements needed for producing a given grade.

- Process causes: a planned change in the manufacturing operation, for instance, a change of grade or preparation for a partial or plant-wide shut-down.

An illustrative example of a change in process regime is presented in Figure 4.2. This figure illustrates a scenario in which the thermo-mechanical pulping process, which is producing a given pulp quality (characterized by its freeness[4] number (CFN)) must respond to the need of the paper machine for a change in pulp specifications (due to frequent paper breaks). The quality is improved by lowering the pulp freeness value. The manipulated control parameters STP1–STP6 are adjusted through a transient period to arrive at the final required freeness value.

[4] CFN stands for Canadian freeness number, which is simply a measure of the drainage capacity of different pulp types. The drainage rate is related to fibre swelling and to surface conditions.

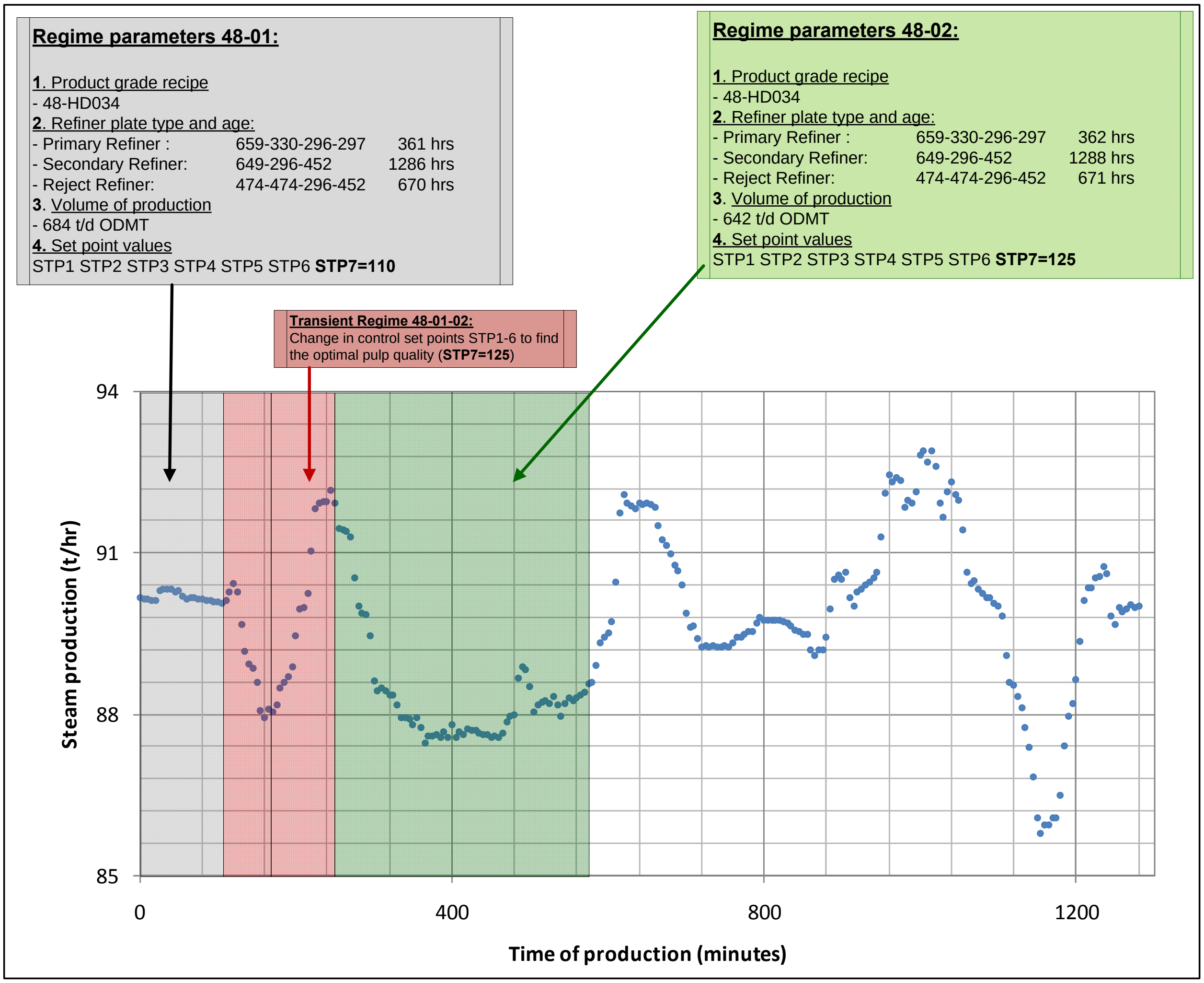


Figure 4.2. An illustration of a operating regime change (production settings) due to required change in pulp quality (presented as a change in one monitored variable - steam production from boilers and recovery units)

When the time boundaries for process analysis are chosen, the raw measurement data can be extracted for process analysis and then cost analysis.

4.3.2 Signal processing using multiscale wavelet trend decomposition

The first important activity phase in the proposed methodology is careful cleansing of real-time process data from high-frequency noise and process abnormalities and identification of when the process is near steady-state operation. First, a test case study, which was a small subsystem of the overall operation, was chosen to analyze the suitability of the method for industrial application. The technique used involves multiscale wavelet decomposition of measurement trends (Figure 2.8). Two essential steps are required to initialize the wavelet technique:

- Gathering of information on each individual measurement point and analysis of its historical values to identify the optimum wavelet transform (WT) cutting scale for each variable, and
- Analysis of historical data to determine the optimal steady-state values of the detection parameters (alpha parameters).

According to the two steps described above, the sensor network at the mill was analyzed, and each measurement point was characterized by its accuracy and precision values. Multiple decomposition trials and tests of each variable trend were carried out to identify the two essential scaling and parameter values. This knowledge is maintained as a matrix representation of the sensor network, whose values will potentially need to be changed over longer time periods. Implementation of adaptive techniques would improve the long-term degree of automation of the system.

The two simultaneous tasks of data pre-processing and pseudo-steady-state detection are discussed further in the following section.

4.3.2.1 Tuning the on-line steady-state detection technique

Figure 4.3 presents the principle of the method by a schematic representation of the overall algorithm. The measurement data were extracted from the data management system (DMS) as a noisy signal. The optimal WT scale was analyzed using an iterative procedure to find the optimal values. After applying a wavelet transform of the chosen scale (data pre-processing), Gaussian noise and other abnormalities were extracted or discarded from the process trend.

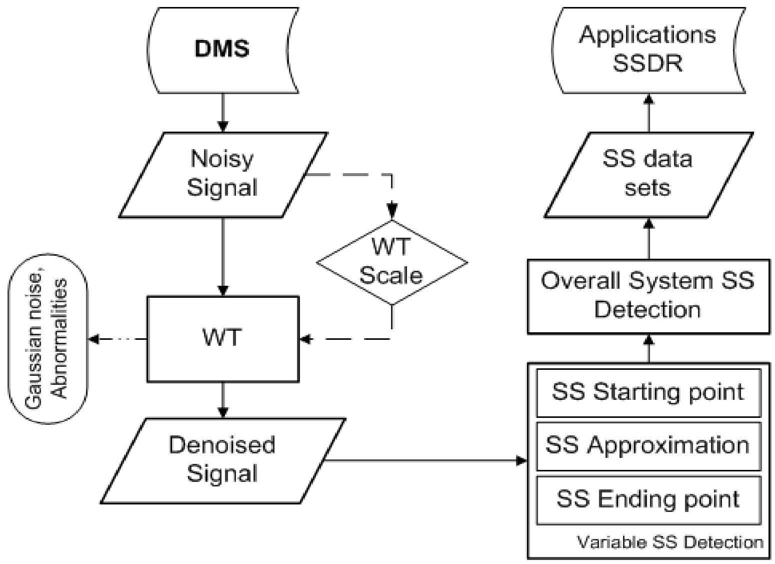


Figure 4.3: Schematic representation of an algorithm for on-line pseudo-steady-state detection

The de-noised signal was then analyzed for potential steady-state occurrence using a three step simultaneous methodology (see Appendix A for a detailed mathematical description of each step):

1. The starting point of the steady-state period is detected using the WT characteristics and its first derivative (values of the pre-determined alpha parameter),
2. High-frequency features of the measured signal, which were not eliminated in the first step, are removed by filtering, and the steady-state duration is approximated by the use of filters (for more details see Appendix B on how the method is performing), and
3. Finally, the steady-state end point is detected through WT feature analysis.

The results from the application of the proposed technique to two different case studies indicated that the method is robust and can provide significant improvements to the accuracy of measured variables. The first case study on a small sub-system was intended to prove the method's robustness by comparing its results to those of conventional steady-state detection techniques (Figure 4.4). To determine the impact on overall system accuracy, the Sigmafine software package from OsiSoft Inc. was used for linear steady-state data reconciliation. The results further indicated that the proposed methodology was able to reduce the probability of not detecting a multivariable operating pseudo-steady state by at least 46% compared to other methods. On the other hand, false detection was only marginally reduced. In critical situations, such as when the system was at a true steady state, but systematic error (bias value) was present, the proposed method was still able to detect that the operation was at a pseudo-steady state and hence to reduce the overall system error (and actually to identify the steady-state period). In fact, for several variables, the use of the proposed methodology in combination with a steady-state data reconciliation technique resulted in higher accuracy.

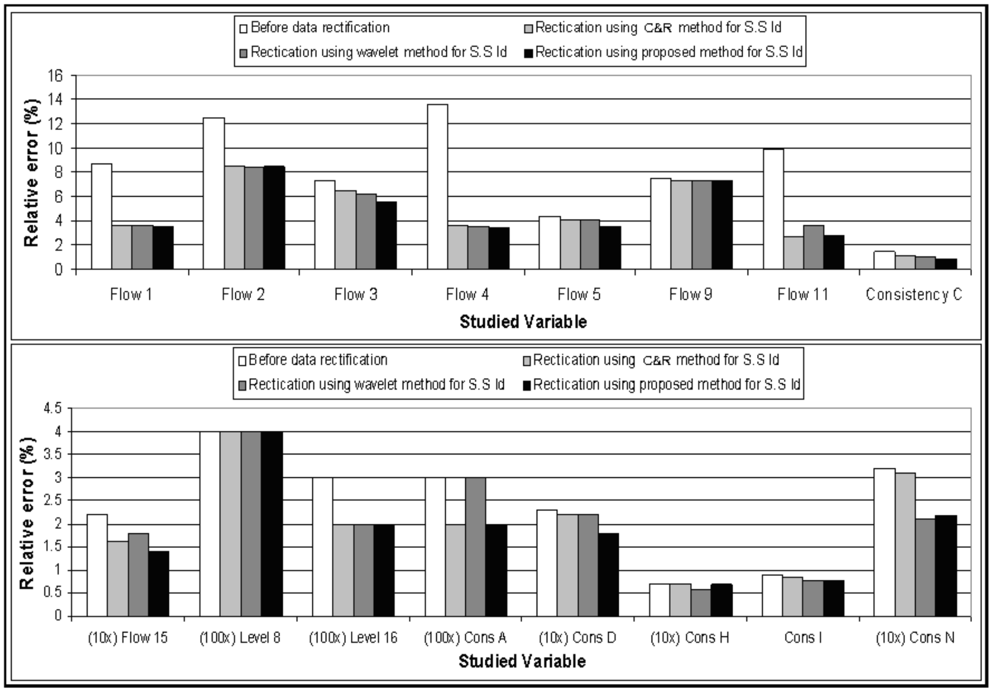


Figure 4.4: Improvement in accuracy achieved by use of various on-line steady-state detection techniques.

These results, which were obtained from application to a small sub-system of the process operation, proved the capability of the method to be used on-line in a real industrial context. However, for higher-level analysis such as optimization or advanced production cost analysis, plant-wide near-steady-state operation is required. This restriction is due to the steady-state nature of these applications and because only steady-state models can provide the process knowledge of various operational regimes which must be used in the context of this Ph.D. research. As already discussed there are few reasons of why steady-state detection is important. The changes in operating control set points are aimed to arrive to safe and steady-state operating conditions. Transitions between them will always occur but cannot be taken as a representation of a given operating conditions. Furthermore, the engine of advanced operations-driven cost analysis is steady-state mass and energy balances coupled with cost information.

The expression for ‘‘steady-state’’ is a misnomer in a rigorous sense because no system parameter is ever steady, and its measurements will always vary to some degree with time. Thus, detection of near steady-state periods requires first to establish an assumption of what near steady-state is, and then evaluate whether the operation satisfies this criteria. The assumptions that were taken into account were based on the analysis of process dynamics. Different trends of measurement variables and their representation with wavelet decomposition present different levels of variations due to process dynamics. Therefore a second case study was carried out to analyze the frequency of multivariable steady state (MSS) occurrence in each section of the mill and in the whole plant-wide operation. Furthermore, the pseudo-steady-state assumption was relaxed by using predetermined alpha parameters to define a optimal steady state. The impact of these assumptions on cost accuracy was then addressed to determine cost uncertainties. The accepted value of alpha is directly linked to the level of system dynamics involved in the steady-state assumption, which was expressed for each variable, each subsystem, and the whole process operation. This information can be used to enhance the results of cost analysis as a confidence value for product margins.

For the purpose of the test case, a relatively unstable (very dynamic) period of operation was chosen that lasted for 600 min. For each subsystem, several key variables were chosen; for instance, seven variables were defined for PWC1 (chips treatment). These variables were chosen to determine whether the subsystem could be assumed to be in a pseudo-steady state or whether it was undergoing a transition (dynamic) period. For each of the key variables, the degree of fluctuation in the WT and its first derivative was analyzed to determine the optimal alpha values. This was achieved by selecting successive measurements at steady state, performing the first-order WT, and then computing the standard deviation of the WT modulus (σ_{WT}). The method was then used to detect the overall MSS for only these selected variables.

Figure 4.6 shows the increase in the ability to detect pseudo-steady states with increasing values of the wavelet threshold of the alpha parameter for steady-state assumption (however, the probability for false steady-state detection is also increased). It was noticed that only certain variables exert fast dynamic behaviour and hence are responsible for the subsystem’s being characterized as not being at steady-state. Figure 4.5 (Korbel et al, (b)) summarizes three aspects of the problem:

- the evolution of the wavelet transform for each key variable
- the corresponding state identification using a binary representation (1=steady state identified, 0=transient period), and
- the number of variables in steady state at each period. If the number of variables in steady state equals the number of key variables (e.g., 7 in the case of the given subsystem) then a MSS of PWC1 was assumed.

The value of the wavelet-transform thresholding parameter has been successfully increased from its minimum value to a value where the dynamics of the process do not permit a MSS to be detected. Note that relaxing the alpha values increases the frequency of single-variable steady-state occurrence (assumption of steady-state) and consequently the number of MSS identified. This systematic approach to MSS detection has been shown to be very practical; however, during the analysis, special attention was paid to identifying the impact of overestimating the steady-state assumption for high-transient dynamic periods or hold-ups. To quantify this impact, the quality (the error associated with the near-stationary assumption) of the pseudo-steady-state data sets obtained was investigated.

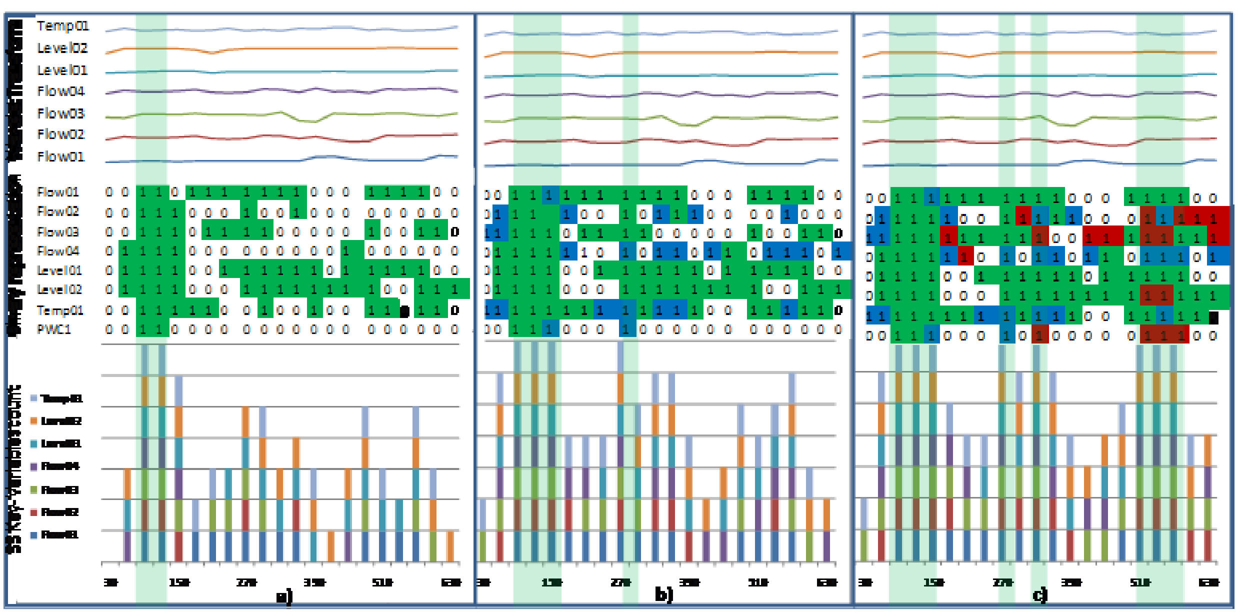


Figure 4.5: Plant-wide steady state analysis

Increase in the number of steady states obtained with relaxation of the value of the wavelet threshold (alpha). A: the value of the threshold is set to 0.1, leading to two multivariable steady states (MSS) identified, B: the value of the threshold is set to 1.2, leading to four pseudo-steady states detected (the two new MSS are represented in blue), C: a threshold value of 5 has increased the number of steady states to eight (the four new MSS are represented in red).

4.3.2.2 Quality of pseudo-steady-state data sets

Clearly, relaxation of the steady-state assumption (steady state assumption includes some degree of process dynamics, relaxing this assumption means here to incorporate more and more process dynamics into near steady-state representation) by increasing the alpha value enables more MSS to be detected. However, this may result in biased information, for instance in estimating production cost, possibly leading to wrong decisions. Therefore, a careful sensitivity analysis of each key variable was carried out to highlight systematically the variables that were causing the overestimation of MSS. For this purpose, two measures were selected

1. A multivariable measure describing the offset of a sub-system or a whole system from some kind of trend (steady-state). In this case, MTE (measurement trend error) was selected as the measure:

$$MTE = \sum_{i}^{n} \frac{1}{\Delta t} \int_{t_0}^{t_0+\Delta t} \left\| y'_{M} \right\|_2^2 dt$$

   where y'_M refers to the first derivative of the variable y_M, which was determined by an algorithm presented in Appendix A.

2. The absolute error of the end application (production cost analysis) corresponding to the absolute error in the product margin estimation. The cost results were compared to the base-case values for the very same operating regime, which were quantified before this analysis using several near-steady-state data sets, thereby statistically ensuring the validity of the results.

Figure 4.7 presents the relation between the probabilities of MSS occurrence (how often near steady-state assumption occurs) in each sub-system of the plant-wide operation and the MTE values. From the figure, it is evident that the PWC3 sub-system (pulp screening) is characterized by relatively little dynamic variation of variables, which enables a larger number of detections of potentially near steady-state condition. For the time frame analyzed, a plant-wide MSS could not be detected without relaxing the steady-state parameters. The alpha values correspond to the maximum threshold value that at least one of the key variables (in most cases, steam flow) had to attain to reach a required number of steady states. This increase in the threshold value for some variables was necessary to obtain at least one plant-wide near steady-state data set in all cases (Figure 4.6.). The impact of this assumption was then analyzed by looking at the MTE of the system and the absolute error in product margin estimation. For comparison, Figure 4.6(b) also shows the value obtained by the *ad-hoc* approach that analysts are using (in petrochemical or other processing industries) an average of variables within the time frame. The higher cost error is due to the fact that the average values incorporate inconsistencies due to presence of tanks ("hold-ups") and due to the nonlinearity nature of process model. This approach can be used with relatively accurate results for linear processes operating frequently near steady-state conditions (such as petrochemical and chemical industry).

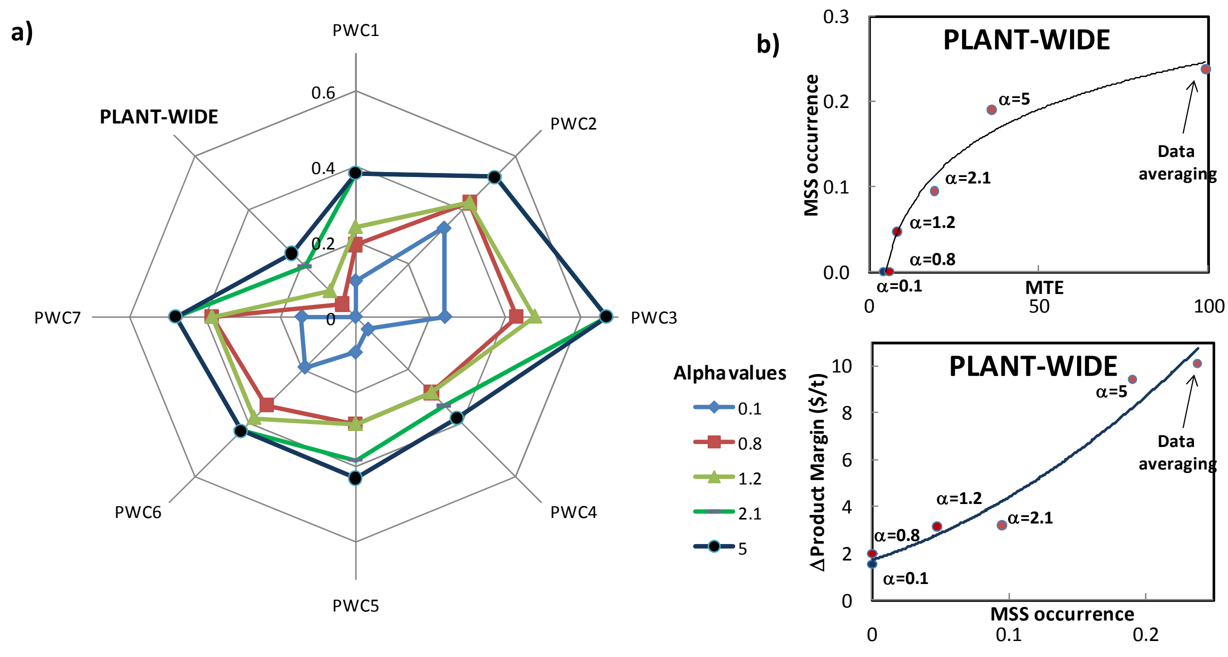


Figure 4.6: Steady-state probability of occurrence

A: MSS occurrence as a function of the threshold values for different sub-systems and for the overall plant.

B: the relation between the MTE and MSS occurrence, as well as the relation between absolute changes in the product margin and the absolute errors in the data set.

In summary, within each operating regime, the method presented here has provided accurate and reliable steady-state candidates for production cost analysis. Furthermore, after iterative tuning of the parameters of the wavelet algorithm (a simple program written in the C language), this data pre-processing and pseudo-steady-state detection task was performed automatically for several production regimes. In this way, the method proved its applicability for on-line use with a low need for user interaction. Note that the thermo-mechanical pulping operation under study is a very stable manufacturing environment, and hence the method was able to identify several pseudo-steady-state operations without the necessity to relax steady-state assumptions using alpha values.

This portion of the study has demonstrated that the method is robust for on-line industrial application and can provide plant-wide pseudo-steady-state data sets within the operating regimes analyzed. Furthermore, a careful analysis of steady-state assumptions can be done automatically while providing an increase in detection rate. This feature increases the visibility and transparency of errors for each PWC and each individual variable. This knowledge can be used to provide guidance in situations in which the method is applied in more dynamic manufacturing environments (such as a Kraft operation) to quantify the impact of steady-state assumptions.

4.3.3 Plant-wide process data rectification

The first block of tasks, including signal processing for process data cleansing and steady-state identification, provided a very good basis for characterizing the manufacturing processes involved in the production of different newsprint grades. However, when these pseudo-steady-state operation data sets were used for product cost analysis, unclear and biased results were produced. Even though the pre-processed data sets had been cleansed from random errors, systematic errors and errors due to the process dynamics themselves were still present. These discrepancies must be identified and eliminated before the data are used in end applications.

Several miscalibrated instruments and highly inaccurate[5] individual sensors were the main causes of the overall discrepancies in results.

The variance-covariance matrix estimation of the measurement errors was performed based on the instrumentation performance analysis from the pre-processing step. Trust weights were allocated to all measured variables based on engineering judgment.

A data reconciliation technique using the Sigmafine software from OSISoft (c) was then used to identify potential systematic errors while providing a complete set of process variables (by a process of coaptation). However, the validation of process measurements showed that only certain variables could be reconciled, depending on the level of redundancy involved. Classical steady-state data reconciliation methods are not applicable to newsprint production because of the low redundancy of the instrumentation network. The number of currently installed sensors is sufficient for production control and safety, but not enough to create the level of software redundancy required for classical data reconciliation. A certain level of redundancy was created by the use of process simulation together with a separate optimization module.

.

4.3.3.1 Development of the simulation-driven data rectification model

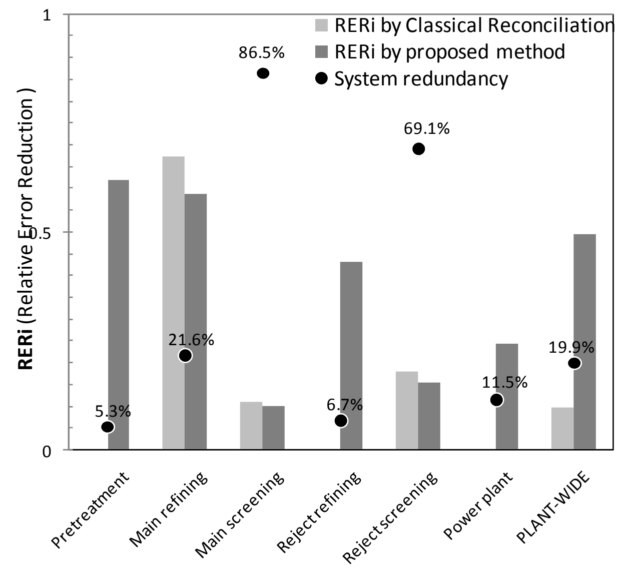
The next step was to choose a model-based cleansing method that could identify and estimate biased measurements and provide complete plant-wide process data sets. The heart of the data-reconciliation step of the methodology is the process simulation using the CADSim (c) software from Aurel Inc. The reason for this choice was that this particular software is well adapted to the papermaking industry. Various process parameters and equations describing processing units were able to increase the number of degrees of freedom for the problem at hand. Furthermore, the minimization of least-squares error, which is the engine of data reconciliation, could be performed in various ways, enabling the use of system redundancy corresponding to zero degrees of freedom.

A process model of the integrated newsprint mill was constructed by a classical flowsheet definition using standard building blocks which describe fundamental operations such as mixing

[5] Accuracy is the ability of a sensor to measure the correct or “true” value (Bagajewicz, 2001)

or separating, but also papermaking-specific units such as chip refiners and paper machines. The steady-state simulation was performed using several key variables from real-time steady-state process data sets to ensure zero degrees of freedom. The sequential solver for steady-state nonlinear systems calculated each process module in turn with its output/input streams for the next module, an approach which provides additional software redundancy. The results of the simulation were then compared to the whole set of measurements, and the least-squares error was calculated. The advantage of this technique is that the measured and simulated variables can then be iteratively reconciled by minimizing a weighted least-squares error while satisfying the simulation model and other user-defined modelling constraints. The optimization module uses a simplex search algorithm to arrive at the optimal solution.

When comparing the performance of the proposed method in low-redundancy systems to classical data reconciliation, it was shown that the relative error reduction (RER of pulp volumetric flow) was similar in several process sub-systems with some degree of redundancy (e.g., Main and Rejects, Main Refining) (Figure 4.7). In other sub-systems of the mill where too few measurement sensors are available, classical data reconciliation will at best result in some level of data coaptation process (input = output model). However, with the use of a simulation-driven approach, it is sufficient to maintain zero degrees of freedom to obtain a more significant relative error reduction (close to 62% error reduction in the chip pre-treatment section of the operation) than classical reconciliation could provide (no error reduction). It is important to mention that because the iterative process of error minimization is taking place between simulated and measured variables, process model quality is of critical importance. In the context of this Ph.D. research, the model is assumed to represent precisely the underlying process operation.



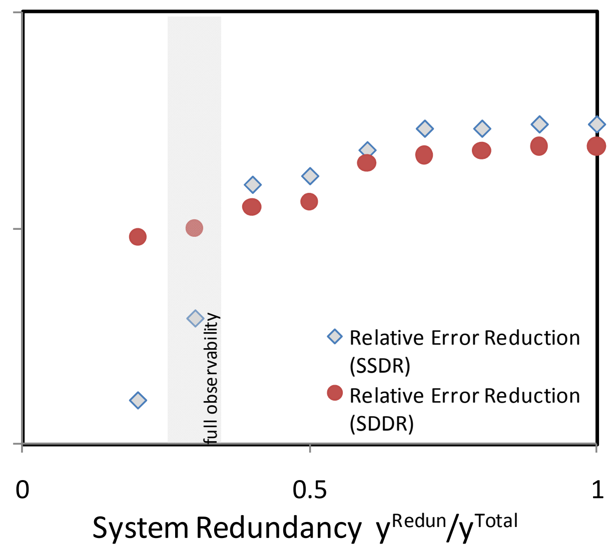


Figure 4.7: Relative error reduction of production rate in different sub-systems of the process operation

Figure 4.8: Relative error reduction of pulp flow measurement and its dependence on overall system redundancy

For plant-wide process data reconciliation, the classical method reduced the error in production rate by approximately 10% (only redundant variables were corrected). On the other hand, the simulation-driven approach could reduce the error by approximately 50% (Figure 4.8, from Korbel and Stuart, d). Assuming that simlation model is correct, it is apparent that the estimators produced by the proposed methodology are superior to those produced by the classical approach. The difference in the outcomes of the two methods reflects the fact that the simulation module includes more system redundancy because the specific papermaking modules are described with more equations (including empirical relationships). Furthermore, the nature of the optimization module enables more practical equality and inequality constraints to be implemented. Historically, similar constraints have been based on the engineering judgment and experience of mill personnel involved with particular processing units and can be in many cases difficult to use in the optimization formulation of classical data reconciliation.

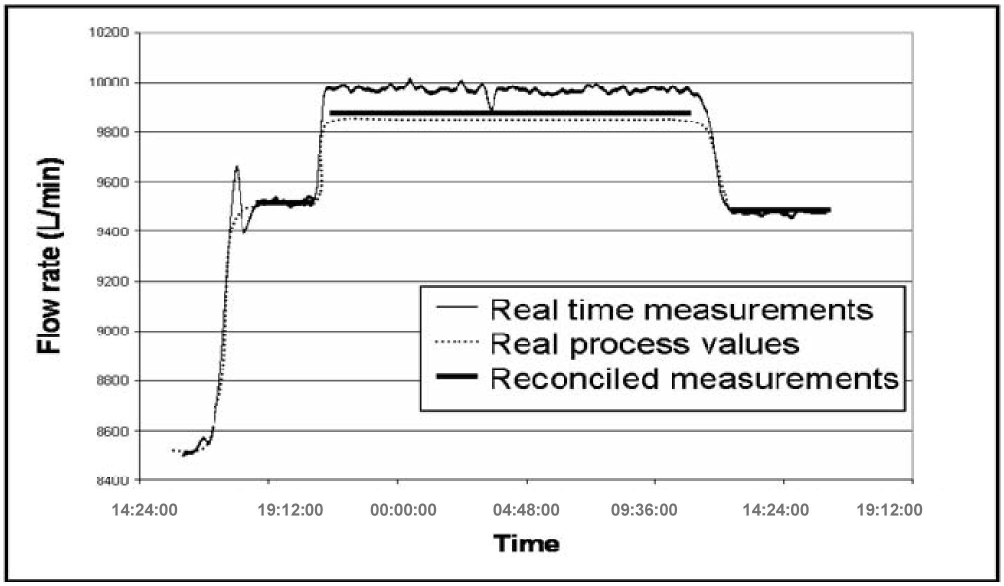


Figure 4.9: Example of a corrected biased measurement.

In a more general approach, the method was applied to various hypothetical production situations where the level of redundancy was being increased from its current value of 19.9% by adding new instruments. The process of adding new measurement points was performed systematically based on a methodology for instrumentation design and upgrade (Appendix B). The following assumptions and steps were followed:

- A hypothetical steady-state data set representing the manufacturing process was created for each level of system redundancy (the weighting matrix was updated based on vendor-supplied accuracy and precision values).
- For each level of redundancy, reconciliation with both techniques was carried out, and a relative error reduction was calculated to analyze the impact of system redundancy on the performance of each method.

As can be deduced from the results presented in Figure 4.8, the advantage of using this method is the possibility of validating process data for low-redundancy systems, where at least zero degree of freedom are present. This is of crucial importance in the papermaking industry, where low-redundancy systems are common because of poor instrumentation-network accuracy. With increasing redundancy values, the simulation-driven method will provide an additional possible improvement of only 20% (from approximately 50% to approximately 70%). Clearly, this fact illustrates that the quality of the simulation model is of critical importance.

The presence of miss-calibrated or malfunctioning instrumentation is determined by the gross error handling process. This task is performed in a very practical way. Historical knowledge about each instrument is used for systematic error testing by analyzing the squared error. If the values are very high compared to the historical data, the most commonly occurring biased measures are checked for potential gross error presence. The bias estimate can then be assessed according to historical values corresponding to similar situations (an example of flow measure correction is shown on Figure 4.9.)

In summary, the simulation-driven method has proved to be practical and its performance to be comparable to that of classical data reconciliation for systems with high redundancy. The risks of the method lie with the trust and accuracy of the simulation model. The method's main practical advantages are:

- More of the equations that characterize the process operation are already implemented in individual processing units as a fundamental basis of the simulation engine.
- The use of the CADSim simulation for reconciling the papermaking operation is very convenient for the user (no need to develop an additional reconciliation model for model-based data validation).

This portion of the study has demonstrated that the presented simulation-driven data rectification (reconciliation) is capable of making pre-processed real-time data consistent with the underlying process model in a low-redundancy manufacturing environment.

Both methods, wavelet signals processing and data rectification, are then combined into a single unique methodology that was then applied to analyze manufacturing operations and provide plant-wide consistent data sets automatically in real time. Figure 4.10 presents an analysis performed on one of the months studied. This period of time was found to be a very stable operation in which the mill operated mainly in three different operating regimes for each grade. For each operating regime, multiple steady states had been identified and subsequently reconciled. Many other operating regimes have occurred but only for short time period. The selected ones that are presented were being periodically repeated according to the needs of paper machine. It must be noted that the relatively stable production was attributed to exceptionally homogenous feedstock quality (humidity and size of chips) and to some extent an ambient conditions (temperature, pressure and ambient humidity). Some external perturbations (change in chips quality) have occurred however, these transient regimes were not analysed due to their very short time periods of occurrence. After the passage of several perturbations, the production was always stabilised to one of the tree operating regimes based on paper machine requirements. These cyclic operating situations serve as very good basis for analysing the variance between chosen manufacturing conditions.

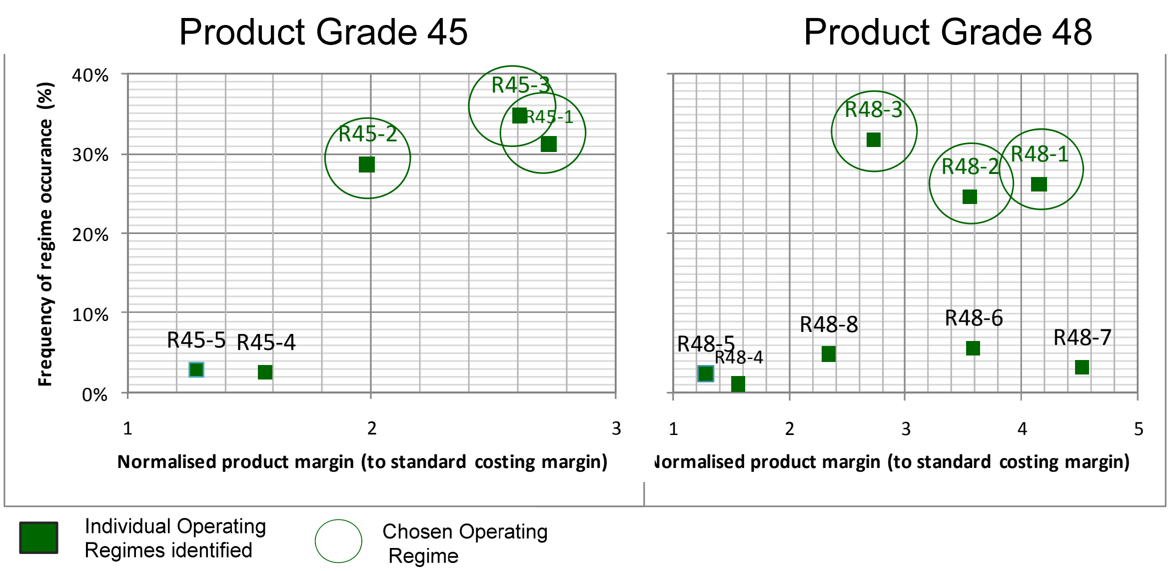


Figure 4.10: Operating-regime selection based on frequency of occurrence

(three regimes for each product were selected based on their frequency of occurrence).

4.3.4 Advanced cost analysis for true product margin assessment

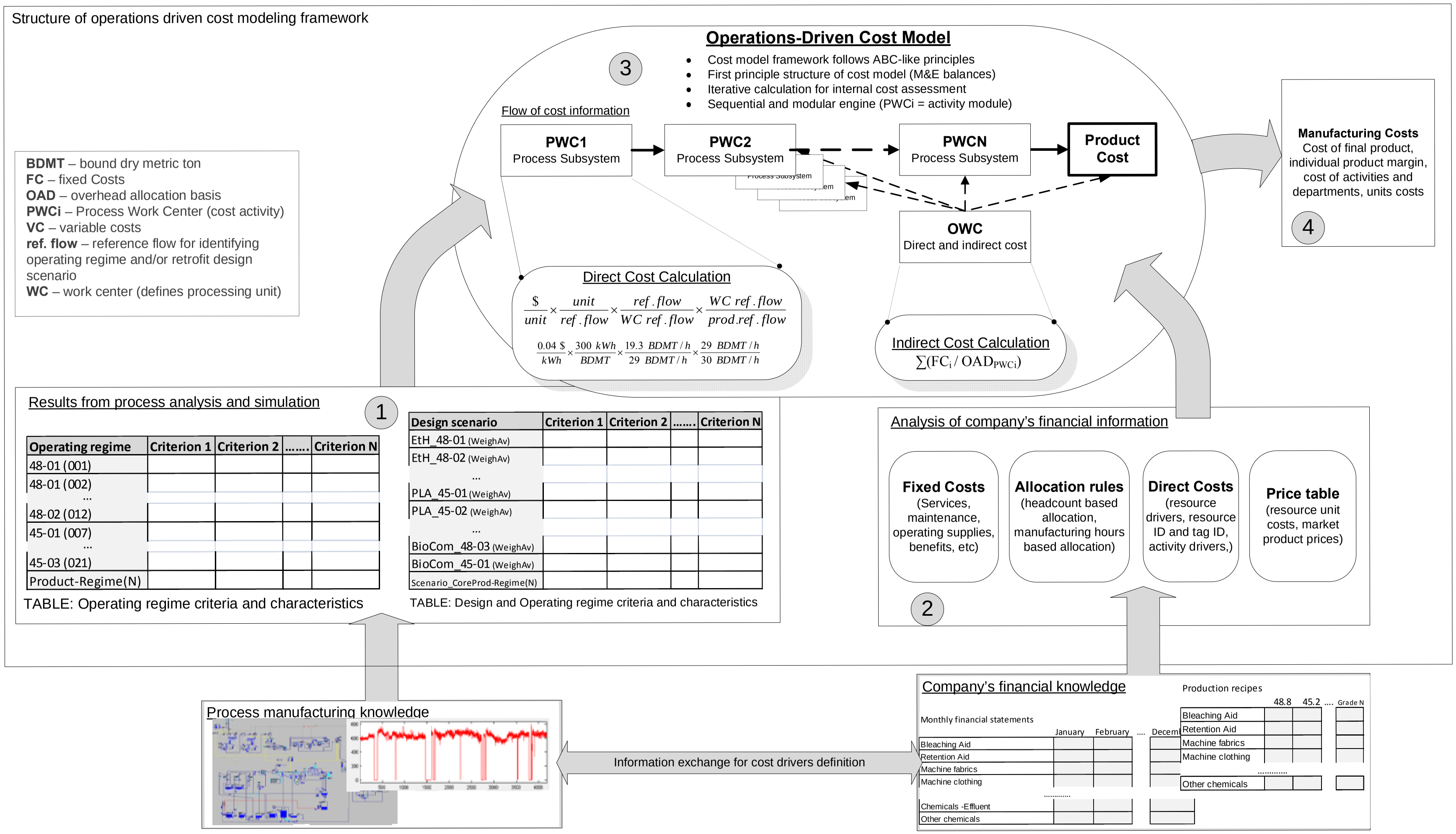
4.3.4.1 Development of the cost modelling framework

The goal of the initial work was to develop an on-line data processing methodology that would generate accurate real-time process data capable of characterizing the plant-wide operating regime for cost modelling and analysis. The third pillar of the methodology—the operations-driven cost model—was then developed to explore these novel and unique insights into papermaking production for the short- and long-term benefit of the company.

The operations-driven cost modelling approach consists of four steps (Figure 4.11) and was implemented using the Impact: EDCTM software:

1. Characterization of the process operation based on real-time process data. The data are dissected to describe multiple operating regimes for manufacturing products in the core business and in the various biorefinery retrofit scenarios.
2. Defining and organizing cost data and cost drivers into matrices that correspond to underlying fundamental (mass and energy) equations.
3. Modeling and calculation of manufacturing costs for operating regimes and biorefinery design alternatives.
4. Analysis, interpretation, and evaluation of cost model outcomes.

The core of the methodology for manufacturing cost assessment of operating regimes and hence for product cost distribution is the ABC-like philosophy. The model was developed at the level of detail necessary to extract complex cost information on operating and design changes and was used to assess the production costs per tonne of each newsprint grade (as a cost object). The individual cost activities, referred to as process or overhead work centres (PWC and OWC respectively),



$$\frac{\$}{unit} \times \frac{unit}{ref.flow} \times \frac{ref.flow}{WC\ ref.flow} \times \frac{WC\ ref.flow}{prod.ref.flow}$$

$$\frac{0.04\ \$}{kWh} \times \frac{300\ kWh}{BDMT} \times \frac{19.3\ BDMT/h}{29\ BDMT/h} \times \frac{29\ BDMT/h}{30\ BDMT/h}$$

$$\sum(FC_i / OAD_{PWCi})$$

Figure: 4.11: Implementation of the operations-driven cost modeling framework for manufacturing costs assessment of operating regimes and product margins: 1 – Process operation knowledge matrix representation, 2 – definition of cost items 3 – Cost modeling engine, 4 – Process operation cost evaluation

Figure 4.11: Implementation of operations-driven cost modeling framework

were defined to capture and represent the chain of production as it moves through manufacturing sub-systems (in some cases mill departments). The direct cost (raw material resources) was linked to these activities based on the process model, whereas the indirect or overhead costs were linked based on predetermined allocation rules and drivers. The cost items in both categories are stored in a matrix representation for convenient access and manipulation using database cost scripts incorporated in the software. Each PWC consists of the following essential elements:

- The process operation criteria and characteristics describing the process regime or retrofit design alternative,
- The integration of cost with mass and energy flow along the manufacturing operation,

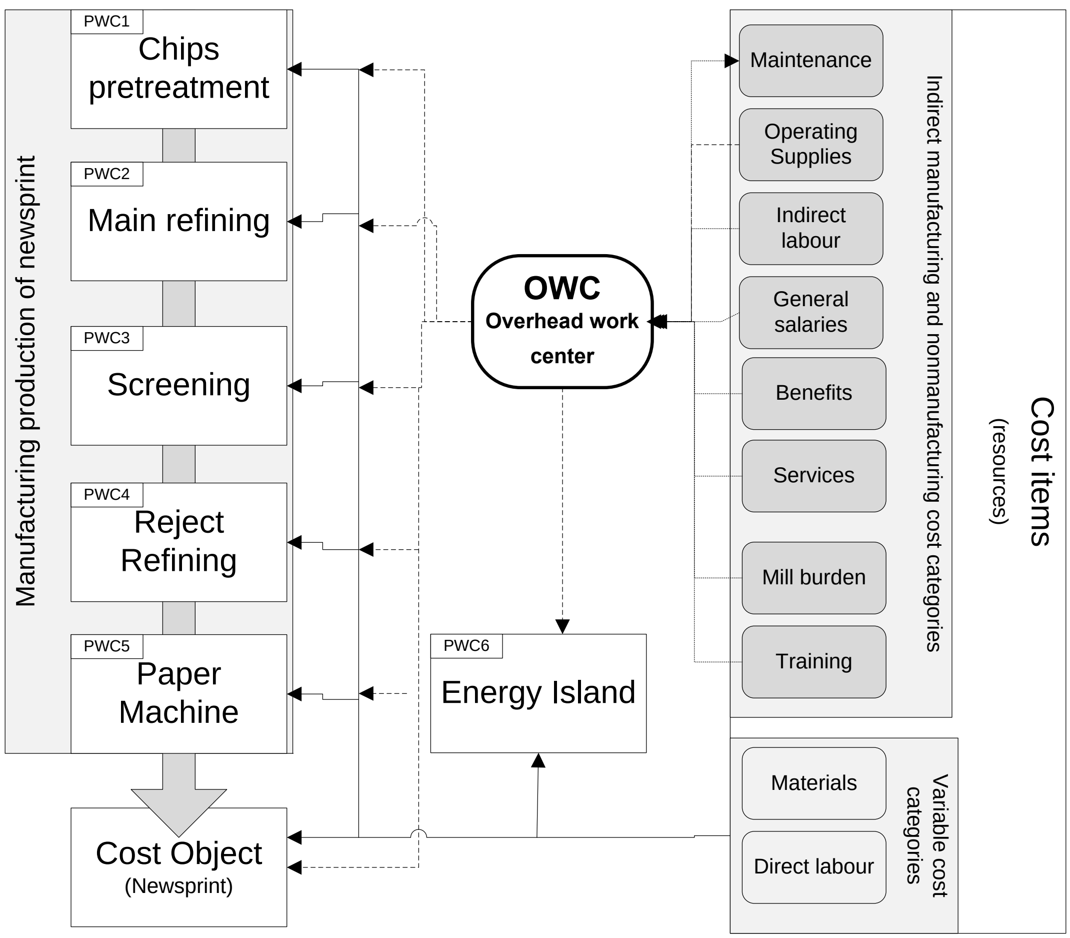


Figure 4.12: Definition of the process and overhead work centers which capture various cost categories within the current business base-case mill.

- Specific calculations (cost-related or simply unit conversion-related) for individual PWCs,
- The allocation (or assignment) of indirect and overhead costs, and
- The core ABC-like engine (operations-driven cost calculation).

Figure 4.12 shows how different types of costs were traced and allocated within the costing framework. Direct and overhead production costs were addressed for each PWC, for each process regime (Figure 4.13) or retrofit alternative (Figure 4.14). The second type of costs that cannot be traced (but can be allocated) by traditional accounting were directly associated with each individual PWC based on allocation information. This association makes the indirect costs behave similarly to direct costs by introducing the link between overhead cost pools and cost objects (newsprint or future FBR products) using PWCs. The costs used were obtained by averaging the cost output of several steady-state data sets representing individual operating regimes. The error associated with each operating regime was relatively small (maximum of 4.2% relative error) when compared to the cost variations among the individual operating regimes.

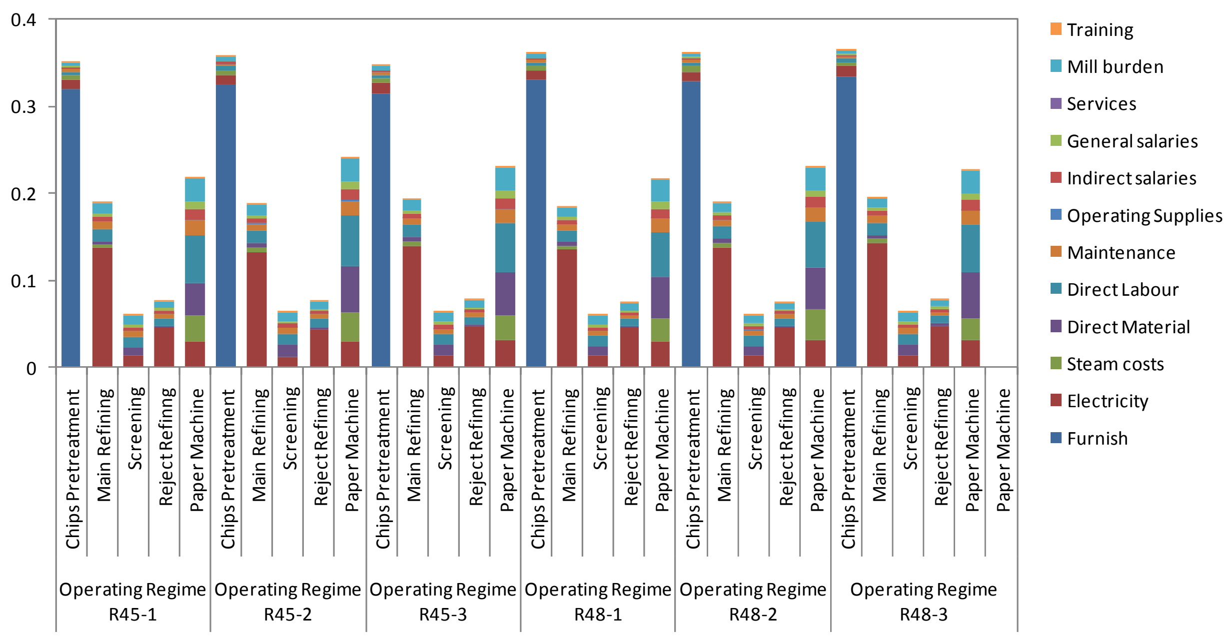


Figure 4.13: Production costs of different operating regimes divided into process work centres and by types of cost

Overheads allocation was performed per ton of newsprint produced, and therefore the variances between regimes are due mainly to production volume. The cost variations of the paper-mill department are associated mainly with variations of steam price, whereas the cost variations in the TMP mill are associated mainly with electricity consumption. The three operating regimes presented here are characterized by different pulp qualities and production volumes, and hence the “free” steam production from refining high-consistency pulp varies. This variance as well as the production changes in the paper mill must be compensated for by producing more high-pressure steam from the boilers (using natural gas, electricity, or oil fuels), which changes the internal unit price of steam. The unit price of steam is calculated iteratively as a ratio of high-pressure steam production price and recovered-steam production price. The iteration step is necessary because the recovered steam price depends on multiple interactions between PWCs. The electricity cost variation is due mainly to the specific energy difference between operating regimes, the production rate, the ratio of reject volumetric rate to mainline production rate, and potential changes in refiner plate characteristics. However, the mill generally uses the same types of plates for a long period of time; therefore, the impact of this parameter was not addressed. The electricity variance due to the change in the reject-line ratio is attributable to the change in the specific electricity consumption in reject pulp refining. The cost of fibre remains essentially constant, with slight variations due to small yield fluctuations, which becomes essential when comparing different grade recipes (e.g., when comparing the productions of 48.8 $g.m^{-2}$ and 45.1 $g.m^{-2}$ grades). The increase in costs due to steam price and the dependence of costs on the interrelation of process activities, as well as the increase in electricity costs, can be captured and interpreted only because of the operations-driven nature of the cost model. The cost model has integrated the resources consumed and their related costs with the process activities in each PWC and has brought process and financial knowledge closer together. This type of analysis demonstrates that the operations-driven cost model proposed here is able to unify the flows of cost and process information to increase the transparency of production costs. Furthermore, the characterization of production costs and the interpretation of variances using lower-level real-time process data have never been done before in the pulp and paper industry. Hence, the opportunity to use this approach for continuous mill improvements will minimize manufacturing costs and increase the cash flow of the company, thus providing competitive advantage.

A similar characterization of mill manufacturing costs for a different one-month period was tried to validate the operating-cost assessment procedure and to compare the outcomes from the proposed method to the values acquired from traditional cost-accounting techniques. Figure 4.14 presents the results of an analysis of several campaign runs at the beginning of the month, producing a 48.8 $g.m^{-2}$ grade. The results are dissected into different operating regimes and cost items. Between individual runs, another product grade is being produced, or a long transient period occurred. For simpler representation, the overhead cost items were united into one cost pool—the overhead costs.

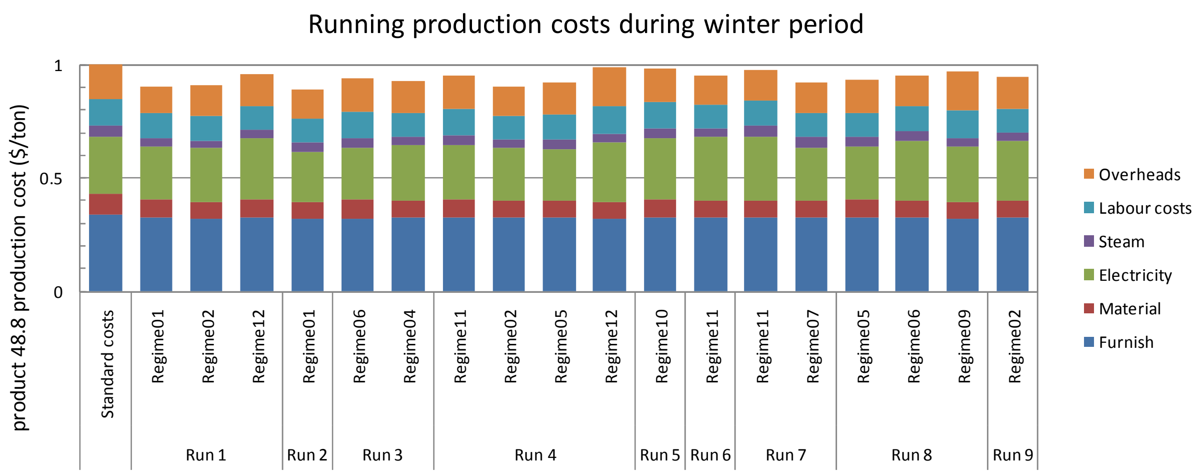


Figure 4.14: Manufacturing costs of newsprint in different operating regimes

Manufacturing costs of a 48.8 $g.m^{-2}$ newsprint product within the time frame of campaign costs analyzed and its corresponding operating regimes (normalized to standard costs).

A few intriguing cases were identified:

- During several runs, steam costs, electricity consumption, use of direct materials, and overheads were significantly different. After closer analysis and discussions with mill personnel, it was determined that a different furnish ratio and a different chip pile were used for these runs, creating a peak in specific electricity consumption for different operating regimes while producing the same grade (regime 01-12)
- Manufacturing costs for the same operating regime but within different campaign runs can vary significantly (e.g., operating regime 09, campaign runs 1, 2, and 9):

  - This variance was determined to be caused by an increase in the use of bleaching chemicals as well as an increase in electricity consumption in the primary refiner. This disturbance was caused by a change in raw material characteristics. After further analysis, it was determined that the reason was an instrument which was miscalibrated and that therefore measured slightly higher production throughput than was actually being produced (a difference in chip moisture). This biased measurement caused the specific energy to move outside the optimal range, creating higher energy consumption.
  - The steam unit decrease was caused by an increase in steam recovery, minimizing the need to produce high-pressure steam. The steam consumption in the paper mill is was identical.

- Operating regime 11 cost variance between runs 6 and 7 was mainly due to the unit steam price. This variance was attributed to excess production of high-pressure steam because of low efficiency of the steam recovery unit due to a mechanical cause: the microfibers present in the dirty steam coming from the high-consistency refiners had clogged the recovery unit. This problem had not been identified for a couple of days, resulting in significant profit losses. Note that the unit steam price fell back to its optimal value (regime09 and regime02 of runs 8 and 9) after the problem had been corrected.
- Comparing the outcomes of the proposed methodology to the values generated by classical cost accounting, it can be concluded that standard costing is an *ad-hoc* method that does not provide the essential process perspective on costs incurred.

This clear process operation visibility enables a better understanding of the mill's cost structure. Furthermore, because process operating regimes are defined based on process characteristics and conditions, the manufacturing costs of the same product can be addressed in multiple ways, making the actual cost distribution of a given product grade available for the first time in pulp and paper mills. Figure 4.15 shows the manufacturing information covering the whole set of operating regimes that were identified and analyzed during the whole month (for the 48.8 $g.m^{-2}$ grade). Each regime is labelled by its corresponding total production cost (normalized to the average value) and its probability of occurrence. The width of the bar corresponds to the cost range of the regime because of the use of multiple steady-state data sets for regime costing (the abnormalities caused by process or raw material were omitted). The thick line inside each of the bars represents the weighted average of near-steady-state costs. The colors towards the red spectrum indicate the more costly operating regimes, whereas the bars with green colors represent the more profitable regimes. The grey bars represent average values. The cost variance in producing the same grade under different operating regimes is significant (~22$ per ton of paper).

This section of the study has demonstrated that the proposed methodology is able to provide cost distributions for different products. This essential analysis enables identification of less profitable operating regimes with the capability to interpret the causes from a process perspective.

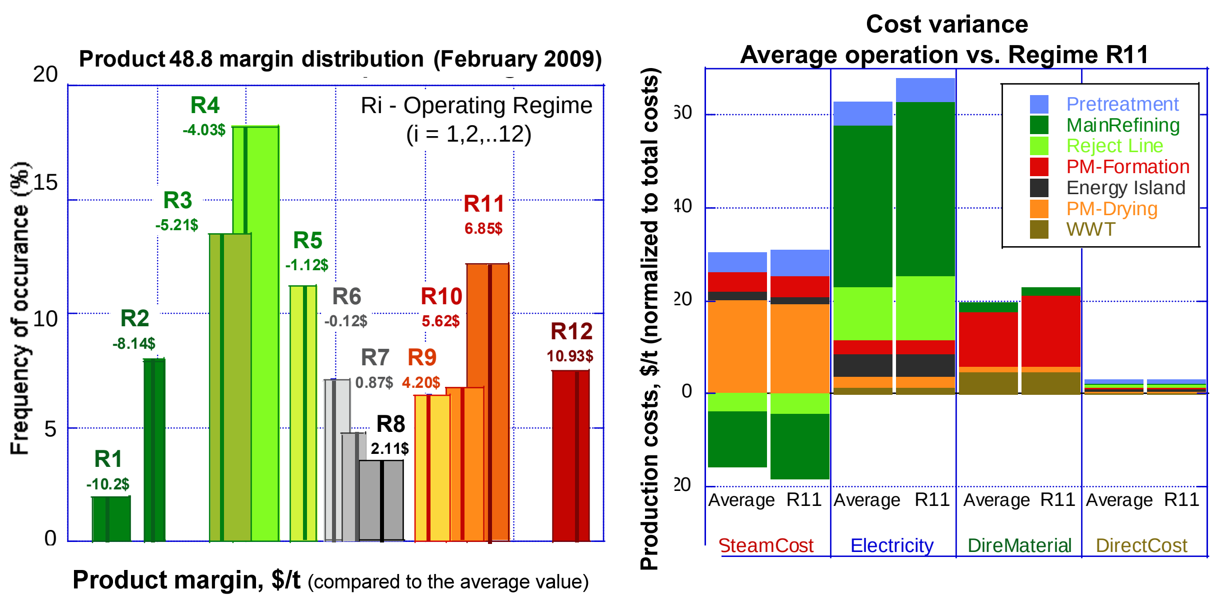


Figure 4.15: Product operating profitability distribution

(Wide range of product profit margins for a 48.8 g.m^{-2} product for one month of operation.)

The knowledge acquired from characterization of the current mill was used to explore the future production costs of three integrated forest biorefinery (FBR) options. The profitability analysis expressed with EBIDTA values for different FBR scenarios is presented in Figure 4.6. From the company's production profit margin, it was concluded that the option of the integrated biocomposite production line would increase the company's cash flow significantly. The manufacturing costs were calculated based on current mill information and simulated data for the new process. The simulation was integrated into the existing base-case simulation. All three retrofit options are tightly integrated with the current core business manufacturing processes by multiple flows. These retrofit scenarios were represented in the operations-driven cost model as different cost objects. Various driver types are used in the model to describe resource consumption by activities or cost objects. Because many of the activity drivers in chemical operations, which characterize direct costs, are based on continuous material flows, the ABC parameters and definitions are reorganized to capture these new activities: for instance, various by-product streams, new material flows, and process steam.

The impact on manufacturing costs of current core business products is presented in Figure 4.17. To analyze the true impact of process integration, the first set of cost analyses within the stable production month was used (only three main operating regimes per product grade were considered (Figure 4.10) and the corresponding manufacturing-cost analysis presented in Figure 4.13). This choice is justified by the high accuracy of the results as a consequence of using many steady-state data sets to quantify the manufacturing costs of a single operating regime.

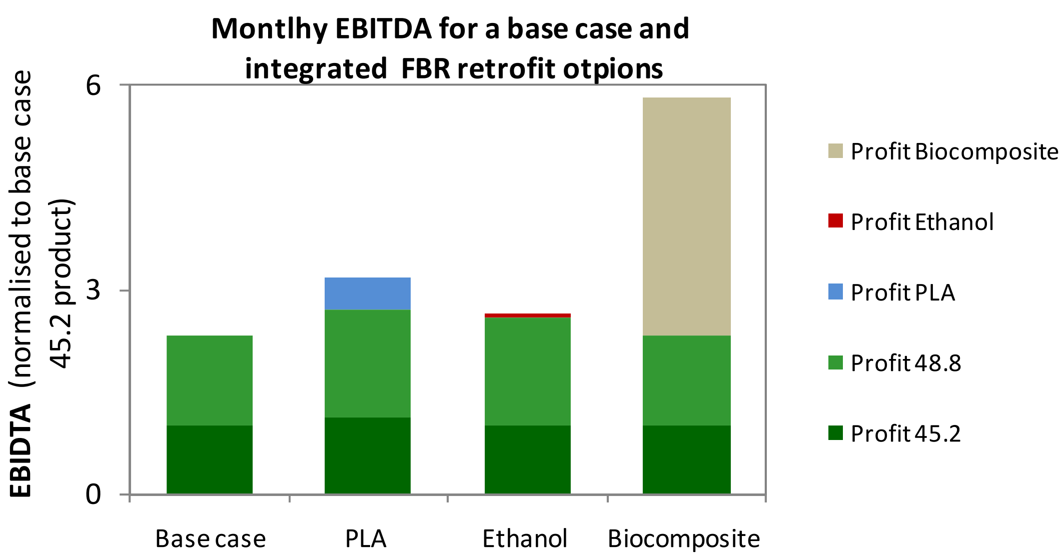


Figure 4.16: Overall manufacturing profitability for each biorefinery scenario

(Overall manufacturing profitability expressed as EBIDTA (broken down by product) of each scenario under consideration.)

When looking at the integrated biocomposite scenario, the cost impact on the core business, newsprint products (both 45 and 48 grades), is minimal. Generally, the variance in manufacturing costs among all three scenarios is due to steam price, electricity costs, and overhead costs. The cost-impact characterization and interpretation can be summarized in several findings:

- Variance in steam unit price between different FBR scenarios is due to the increase in the overall operation demand for producing high-pressure steam in both VPP options (bio-ethanol and PLA scenario). Both PLA and bio-ethanol processes require steam for the final purification step; however the production volume is so small that the cost impact on core business products is marginal.
- The increase in pulp throughput, from the pre-treatment step to the secondary high-consistency refiners (to account for producing 80 tons per day of biocomposite pellets), increases the production of dirty (low-pressure) steam from the primary refiner. This impact was manifested as a decrease in steam price; however, the new processing line requires steam for biocomposite pellets drying, and therefore the overall net impact increases the steam unit price marginally.
- This change in pulp flow rate reduces the specific energy of the primary refiner (from 980 kWh/ODMT to 920 kWh/ODMT), which has only a small impact on electricity consumption. These process changes and cost impacts on core business products are minimal in the case of the biocomposite scenario.
- Pre-treatment of chips using acetic acid, which is common to both VPP scenarios, provides significant cost savings for the core products. The change in specific energy is almost 25%.
- The difference in overhead costs is due to the sharing of indirect costs with the new process operation. However, due to the increase in labor and maintenance, the overall mill burden increased, which manifested itself as a small change in the net impact on core products.
- Direct material and furnish were not affected by any retrofit option. However, the increase due to chip unit price was not considered because of the small production volumes that

are required by the FBR options. However if larger volumes were considered, this impact might become significant.

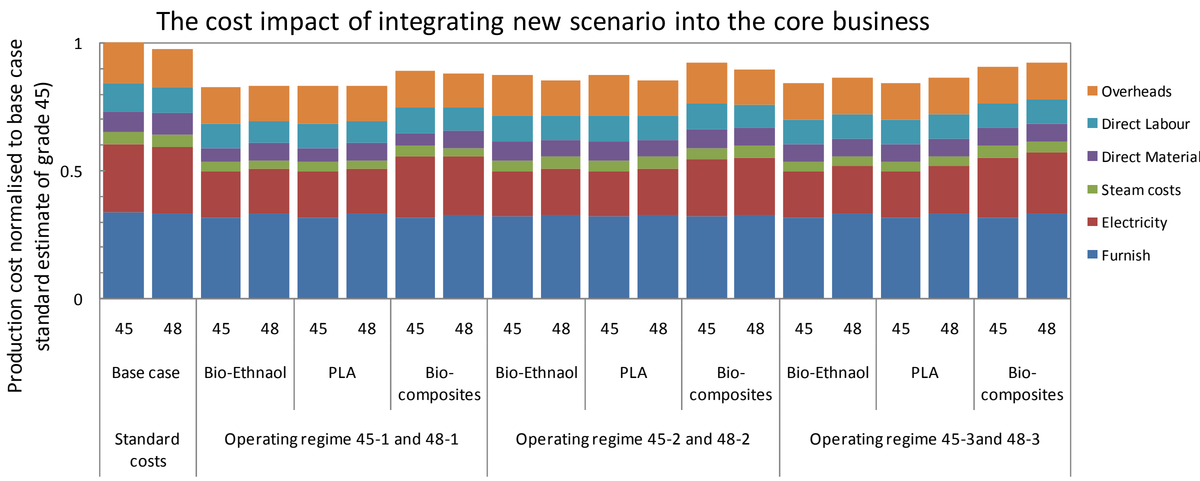


Figure 4.17: Cost-process impact assessment of different retrofit design alternatives

(Comparison of the cost impact on core products for each scenario (grades 45 and 48))

From the findings presented above, it could be concluded that the biorefinery retrofit options involving ethanol and PLA are more attractive when considering the cost savings for paper products. One of these retrofit options might be prioritized before the biocomposite scenario if the newsprint market were to have a positive price forecast which would ensure an increase in the company's long-term cash flow. However, the overall cash-flow increase in the biocomposite scenario is significantly larger compared to the two VPP options. This will provide business security for the company over the long term, assuming that biocomposite market prices follow the predicted trend.

When considering specific retrofit options and their impact on core paper products, the proposed methodology provides granular cost performance analysis of manufacturing, as represented by different regimes (Figure 4.18). The PLA production costs are increased when producing 45 newsprint grades. This difference is mainly due to steam and overhead costs. The steam unit price was determined to be lower because of the increase in the demand of high-pressure steam when producing grade 48. The specific steam usage per ton of PLA changed only marginally; however, the increase in the internal steam price caused the overall steam costs to be increased. The R45-3 operating regime is the most profitable operating scenario for simultaneous PLA and 45-grade productions (the absolute cost difference is more visible from Figure 4.19). The favourable position of the R45-3 operating regime is mainly due to the higher production rate of PLA, which is enabled by the highest rate of parallel production of paper grade 45. Most of the variance can be observed to be due to overhead costs because the allocation base (tons of PLA produced) increased PLA profitability. The specific use of steam and electricity changed only marginally due to the corresponding increase in steam demand and hence unit steam price. Similarly, the R48-1 operating regime was identified as the most profitable for parallel PLA and grade 48 production. The manufacturing costs for this scenario are significantly reduced compared to any option with parallel 45-grade production. The difference is due to the increase in PLA production, which increases the allocation base (tons of PLA produced). The allocation basis for overhead reallocation to departments is based on head count per these departments (for more details see Appendix I for cost classification and calculation). The overall mills overhead costs are spread across the new integrated facility. Since the tons of PLA production is significantly smaller than of paper products, the OH contribution is large. However the specific steam cost is higher because of the significant increase in steam unit price resulting from the increase in overall steam demand.

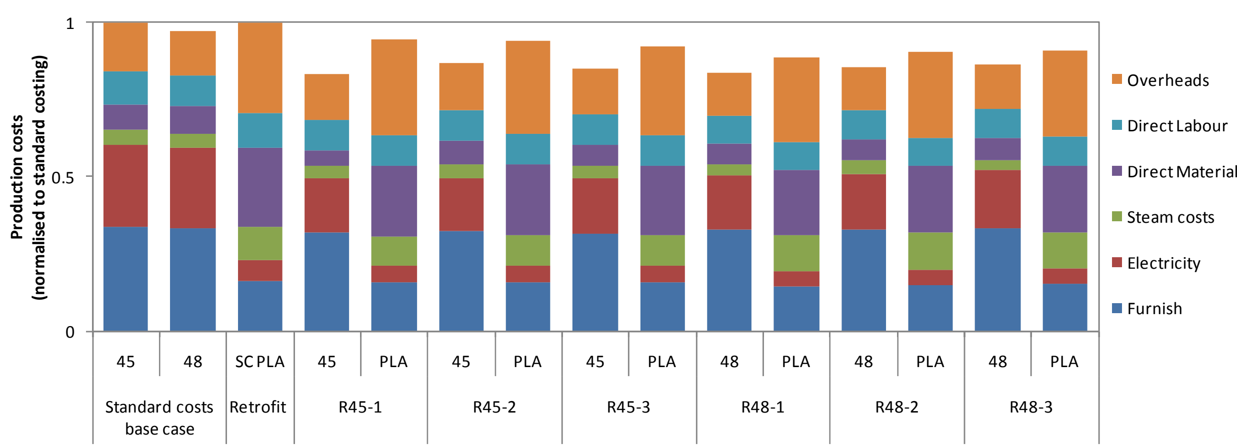


Figure 4.18: Process-cost impact of PLA scenario to core business production

(Production costs of PLA option within the biorefinery (paper grades and PLA) scenario for each operating regime (paper grade costs are normalized to standard costing of grade 45, whereas PLA costs are normalized to PLA standard costs to improve visibility))

The variations in the different scenarios stand out even more when looking at the contribution margin (operating profitability) of each product for different scenarios characterized by operating regime (Figure 4.19). The values of the changes in actual or true product margin (normalized to the base case calculated by standard costing) provide information which can be explored in strategic decision-making. The near-zero or negative effect (simultaneous biocomposite production in regime R45-3) on the margin relative to the base case indicates that the biocomposite option may appear to be unattractive. PLA production appears to be very attractive when looking at only this potential decision-making parameter: in some cases, the margin increases by over 200% (PLA and grade 45 productions in R45-1 and R45-3 regimes). However, different conclusions are obtained when looking at the gross profit margin of the production in the retrofit option with biocomposite production (Figure 4.16).

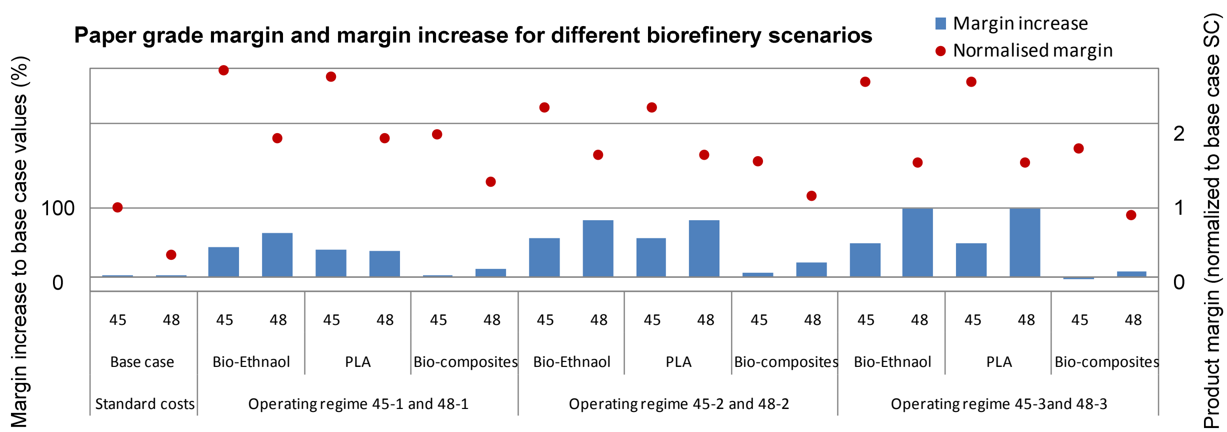


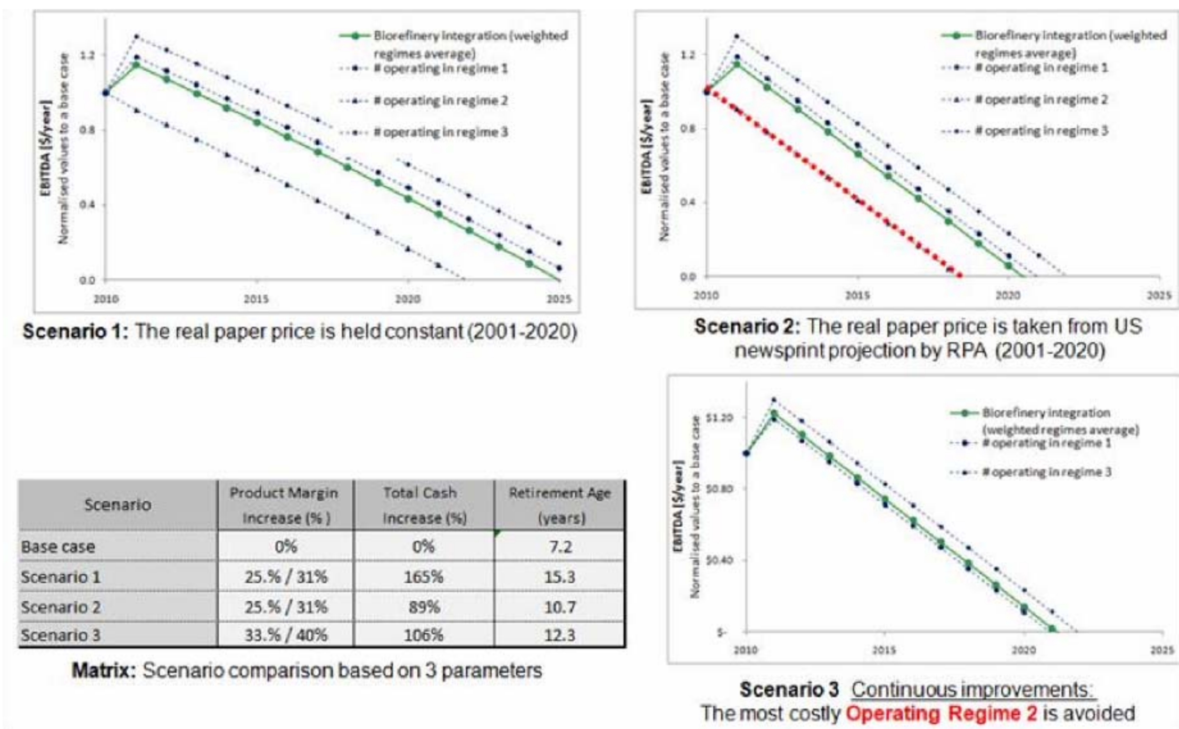
Figure 4.19: Operating profitability of each core product under biorefinery scenarios

Further long-term company benefits from using operations-driven cost modeling based on ABC-like integration of real-time data with cost data can be seen from a simple analysis of the future company's cash flow. EBITDA (earnings before interest, taxes, depreciation, and amortization), also called operating profit, which serves as an indicator of the cash-generation potential of the retrofit scenario, was taken as a measure of long-term mill benefit expression. Figure 4.20 presents a tree of possible future scenarios and their interpretations in the matrix. Each of these scenarios has been analyzed under the following assumptions:

- Constant quantity of paper products sold to customers (only grade 48 is assumed),

- Yearly increase in production efficiency due to operating improvements,
- Yearly labor and raw material cost increases,
- Selling prices for the base case and for scenarios 2 and 3 are taken from predictions by RPA (2001–2020); the selling price is held constant for scenario 1.

The assumption of constant newsprint price after PLA scenario implementation shows a business cash-flow increase of more than 165% compared to that of the current business. This will make the new business model break even more than 15.3 years from the present (8 years more than the base case). By assuming simultaneous continuous improvements by simply avoiding mill operation in regime R48-3, and assuming a change in future paper price, an increase of 106% in total cash flow within the analyzed business period may be achieved. Furthermore, the product margin of the 48.8 g.m^{-2} grade is increased by nearly 40%, enabling newsprint production to continue more than 22 months longer than in the base case.



Scenario	Product Margin Increase (%)	Total Cash Increase (%)	Retirement Age (years)
Base case	0%	0%	7.2
Scenario 1	25.% / 31%	165%	15.3
Scenario 2	25.% / 31%	89%	10.7
Scenario 3	33.% / 40%	106%	12.3

Matrix: Scenario comparison based on 3 parameters

Figure 4.20: EBIDTA forecast and the potential impact of regime costing on product margins, and retirement age of a business.

In summary, this analysis of potential retrofit design changes has benefitted from having an operations-driven cost model based on ABC-like principles in place which systematically integrates process and cost data. The model focuses on process costs instead of only on product cost assessment and is able to analyze the implications of operating regimes in future retrofit design options. This enables the identification of profitable operating conditions in current and future manufacturing environments and provides an increase in the profitability of retrofit design alternatives.

4.4 Conclusions

This chapter outlined the implementation of the on-line methodology for assessing operating product margins, and its results related to the case study involving characterisation and interpretation of current core business and to the case study involving the implementation of three integrated forest biorefinery scenarios. First, signal processing technique, based on wavelet transform and filtering, was applied to analyze the time-frequency space of individual sensor of the whole plant-wide instrumentation network. This process was done for several operating regimes that were chosen for production cost analysis. This step provided mill personnel with multiple steady-state data sets to use in different application. The online ability of the method, to detect steady-state has permitted process engineers to gather several data sets without an effort. This is due to the wavelet processing features which allow identifying abnormal measurements and excluding them from the analysis. The application of the method to case study has shown that the accuracy of production costs was improved when compared to the use of average data within the analyzed operating regime. In the second step, simulation-driven data reconciliation was applied to analyse the sensor network for the presence of potentially biased instruments. The estimation and correction of their values was carried on if presence was confirmed. The method was able to identify multiple measurements with biased values. The sizes of their errors were estimated and correction was applied. This is due to the ability of the method to be applied in the low redundant system as the case study was. When classical reconciliation techniques were used, the system has failed to crosscheck measurements and detect biased measures. In the third step, the use of process knowledge is exploited in the ABC-like cost accounting framework. The advantages of this approach have been demonstrated in two case studies for short and long term newsprint facility benefits. The improved visibility of the mill cost structure permitted process-based interpretation of cost variances between several operating regimes, which creates potential framework for continuous improvements. Several process-related problems were identified when using the ABC-like character of the method. For instance, it was found that the shutdown of recovery unit have a large cost impact on production costs in summer periods. The impact of the change in feedstock properties has manifested as the variance in steam, electricity and chemicals consumptions. Furthermore, it was found that, significant cost variance occurred, when operating in regime that corresponds to a different product grade. This type of interpretations is possible due to the ABC-like structure of the model which emphasizes the process costs, whereas the traditional cost approach focuses on the product cost. In the last step, the process-based production knowledge of the current business was used to analyze future facility manufacturing costs for three retrofit design scenarios. The transparency of cost results due to ABC-like method have allowed for improved view on the potential cost-process impacts based on actual and real-time process data.

CHAPTER 5 GENERAL DISCUSSION

The implementation of information management systems in pulp and paper companies has enabled a better understanding of both business and production processes. Even though mill engineers and accountants have incorporated the use of real-time data into their daily practices, they are often limited to using these data only for *ad-hoc* problem solving. The critical information captured in the data has not yet been made visible to decision-makers. Data trends are studied, but information is seldom extracted from the actual measured variables. If information management systems at the mill can be exploited to their full potential, decision-making activities will be enhanced significantly by access to new and insightful manufacturing information for operational, tactical, or strategic decision-making.

Understanding individual product margins becomes essential to determine the optimal unit prices and to reveal the true profitability of production. Current cost accounting systems at pulp and paper mills provide only an *ad-hoc* assessment of these values. The common simplification made by all accountants, assuming product homogeneity, creates distortions to the real costs incurred in the time frame under analysis. Standard costing methods based on standard recipes can serve only as a mill benchmark for performance evaluation. On the other hand, actual cost calculations using traditional methods provide only aggregated costs that are assessed in a top-down manner. The division of such aggregated costs into cost pools corresponding to individual products is usually volume-based and therefore incorporates various changes in process operation due to process dynamics, raw material disturbances, or both. For this reason, this assumption is often far from reality, making the estimated costs unreliable for determining the true product profitability that is critically important for decision-making. Mill accountants and engineers recognize that the rate at which each mill generates costs may vary significantly, even when the mills are manufacturing the very same product. Determining the true operating margins of individual products is clearly a challenging task for accountants in the processing industries because both process and cost data are biased.

The aim of this work was therefore to develop a methodology for on-line manufacturing cost analysis, using real-time process and cost data available from information management systems, which would be capable of assessing actual product margin costs and of using this information for operational and tactical decision making (Figure 3.1). Furthermore, the knowledge gained from applying this methodology can be explored at the strategic decision-making level for addressing the process-cost impact of retrofit design alternatives. The methodology is applied in a case study which considers both current newsprint mill characteristics and potential retrofit biorefinery implementations.

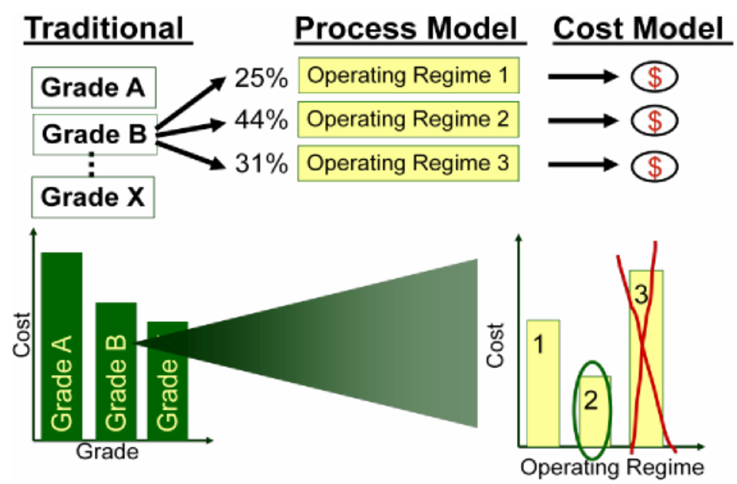
5.1 Manufacturing cost assessment

With the use of advanced data-processing tools and methodologies, the operating costs of manufacturing can be expressed by assessing the costs of different operating strategies (operating regimes). Every paper product is manufactured according to a particular production recipe. However, within this recipe, different operating strategies can be followed by operator choice or as a result of natural process-material interactions. These operating regimes are driven by process design characteristics and operating practices. For instance, the use of different chip-refining plates, changes in the control setpoint strategy for freeness control, and the open or closed nature of process loops and units are examples of operating regimes. In an environment more complex than the case study examined here, operating regimes could be defined by the level of process flexibility or the use of different production lines to arrive at the same product specifications; different product recipes could be as well used for cost analysis.

In response to these needs, an operations-driven cost approach was developed to capture the necessary cost characterization from a process perspective and interpretation. The overall structure of this cost-modeling vision, as discussed throughout the thesis, can be understood from Figure 5.1 (Korbel and Stuart, 2012). Traditional cost-accounting procedures permit *ad-hoc* profitability analysis of different products (grades). To go further and to understand the actual costs incurred from a chemical engineering perspective, a process model should be used to assess the profitability of individual operating regimes using the probability of occurrence of each regime. At this stage, the information can be used by operational decision-makers to choose the most profitable operating regimes and to eliminate costly ones and by tactical decision makers to enhance planning and scheduling and to explore process flexibility for the short-term benefit of the company.

This clear visibility of the process operation provides a better understanding of the mill's cost structure. Furthermore, because process operating regimes are defined based on process conditions, the manufacturing costs of the same product can be addressed in multiple ways, making the cost distribution of a given product grade available for the first time in pulp and paper mills.

Several operating regimes were recommended to the mill personnel based on the results of the case study. It was also concluded that some of the costly operating regimes could not be avoided due to process or raw material constraints or because they occur as a natural response to ensure operating safety.



5.1: Smart data dissection for an operations-driven cost modeling approach

Because pulp and paper facilities operate in an item-based or order-driven manufacturing environment, the use of a regime costing system creates a sustainable option for the corporation. Not only will short-term savings in manufacturing costs be generated, but also high-value supply-chain modeling and the analysis of a potential retrofit or transformation of the business to a forest biorefinery will benefit from these valuable insights into production knowledge. This was shown in a case study application of the method to strategic analysis of the process-cost impacts of possible future retrofit design and manufacturing alternatives. This analysis proved that the benefits from an operations-driven cost model based on ABC-like principles will enhance strategic decision-making knowledge. These benefits arise because the cost model focuses on process costs instead of only on product-cost assessment and is able to analyze the implications of operating regimes in future retrofit design options. Interpretation of different core operating strategies and their actual impact on producing a parallel mix of products in the future were discussed. The essential knowledge gained from these granular cost results can be exploited in the company's strategic planning activities to enhance decision-making information and to identify the optimal option for a complex multi-product manufacturing environment.

5.2 Data reconciliation

To provide operations-driven cost analysis with process knowledge, data reconciliation must be performed, first, to validate the process measurements and second, to help estimate the extent of bias in measurement signals. As discussed briefly in the literature (Section 2.2.1), two major types of applications can be distinguished based on an analysis of major industrial types of data-reconciliation applications (Narasimhan and Jordache (2000):

- *Process unit reconciliation* (fundamental first-principles balance models for processing units) applications for different type of reactors (especially pyrolysis reactors or catalytic and cracker-reformer units in the petrochemical industry), including distillation and separation columns in the petrochemical industry
- *Plant-wide reconciliation* of production and utilities accounting for refineries. Many refineries are already saving on production costs by using data-reconciliation techniques.

However, industrial applications for plant-wide data reconciliation are generally home-grown techniques for petrochemical facilities (Romagnoli and Sanchez, 2000). In pulp and paper facilities, the lack of measurements has restricted data reconciliation to be performed only as an off-line method with very scarce access to data measurements. Usually, manual tests are performed over several days (Jacob, 2003) to gather a single snapshot that represents the plant-wide (or processing unit) operation. The application of this laborious procedure is challenging and very impractical for the needs of process-based applications such as the proposed methods for operations-driven cost accounting and operating-regime cost analysis that have been developed in this work. Several trials of classical data reconciliation have failed in application because of lack of redundancy in measurements.

In response to this lack of industrial applications of data reconciliation in the papermaking sector, the focus of this work was to elaborate a framework for steady-state data reconciliation using coupled simulation and optimization models for this purpose. The choice of the CADSIM software from Aurel Systems Inc. was justified on the basis of its functionality and degree of adaptation to pulp and paper facilities. Close collaboration with the software developers has created an opportunity to access the source code and to define customized empirical equations. In this way, analytical system redundancy was increased by increasing the degrees of freedom of the problem. This approach facilitated the establishment of a system with zero degrees of freedom using several necessary key variables.

This approach has proved to be very practical for on-line industrial application in the pulp and paper industry. The approach provides robust estimators of process variables with relatively low least-squares error. The sequential nature of the software creates a delay when too many iterations are necessary to arrive at the optimal solution. When data pre-processing and steady-state detection using wavelets were used, the approach provided robust estimators corresponding to near-steady-state operation. It must be emphasized that model accuracy is the key factor in the approach to provide good measurement estimators corresponding to real operation.

5.3 Signal processing

The use of data pre-processing methods for industrial applications was investigated for the purpose of providing accurate representation of the underlying production processes. Data pre-processing in the pulp and paper facilities is done using filters. Use of the analog or digital filters that are usually incorporated in distributed control systems is not sufficient to eliminate completely the effect of abnormal measurements. Some larger outliers can be eliminated using the permissible lower and upper bounds of process variables; however, many outliers pass unprocessed and remain within these bounds. The presence of abnormalities decreases the performance of any process-state identification system (Shankar, 2000). The method proposed here uses wavelet-transform features to decompose the signal in the time-frequency domain. Each part of the signal in the frequency space is analyzed by the wavelet transform. Comparison of the WT module extrema (as discussed in Section 2.3.2.1 in the critical literature review) makes it possible to identify and discard outliers.

As mentioned in the critical literature review, several robust methods for steady-state identification exist and are used for process state analysis. However, only a few applications (using filters) have been reported that focus on on-line applications of state identification. Flehmig (1998) built a mathematical framework for state identification using wavelets for potential on-line applications. Parallel work by Jiang (2000) introduced an actual algorithm for off-line industrial applications. Therefore, in response to the lack of on-line industrial applications of data pre-processing and steady-state detection, especially in the pulp and paper industry, the focus of this work was to elaborate on Jiang's efficient and robust method by introducing a new three-step methodology and thereby to offer on-line industrial application of near-steady-state process detection.

The method shows robust performance and the ability to detect near-steady state from real-time measurements. This is done using key variables representing the state of individual sub-systems, from which the multivariable near-steady state of the plant-wide operation can be identified. A very practical analysis includes a framework for assessing the uncertainty in the steady-state assumption due to operating dynamics. Indeed, this may be a very challenging task because of the complex manufacturing environment of a newsprint operation, which consists of a number of tanks and process loops. The case study was, however, a rather simple and stable operation. A modern, computerized single-line thermo-mechanical pulping operation with a single paper machine exhibits near-steady-state operation fairly often. It was therefore possible to detect several steady-state data sets to characterize the same operating regime and thus to validate each step of the methodology by statistically comparing the offsets of the various data sets from each other.

The method has several advantages that have simplified the cost assessment of the integrated thermo-mechanical newsprint mill case study. On-line use of the method facilitated the extraction of multivariable steady-state candidates. Note that in the thermo-mechanical mill under study, the method was implemented as a trial for data pre-processing. The control personnel and operators at the mill have agreed that the results are superior to the current practice of using the filters in the DCS system.

CHAPTER 6 CONCLUSIONS AND RECOMMENDATIONS

Canadian P&P facilities have tremendous opportunities to increase their level of competitive advantage by developing tools and methodologies to exploit their information management systems. These systems have gathered vast amounts of process and cost data which are not now fully exploited. The use of the activity-based costing philosophy and its variations could help managers improve forest-company profits based on these valuable data. This thesis presented ABC-like methodology for on-line manufacturing cost assessment that is based on using this real-time process and cost data from information management systems. The supporting pillar of the method is an on-line technique that is able to detect near-steady-state operation and to establish that the steady-state data sets are relatively accurate. The accuracy and validity of the operating-regime representation with the use of near-steady-state data increases with the number of near-steady states identified. When wavelet signal processing is combined with a data reconciliation method, the analysis can provide a complete set of plant-wide reconciled data representing operating regimes.

It was found that this methodology provides granular cost results thus provide transparency to manufacturing-cost knowledge of a current business environment, enhancing multiple opportunities for short and long term company's benefits. This methodology can be used as a tool in day-to-day operations that would assist mill personal on multiple aspects of the organisation. The strategy of continuous mill improvement can benefit from this information, process flexibility can be explored when tracking paper prices, a margin-centric supply chain will benefit from providing actual product margins for each operating regime, process-driven explanation of cost variances on a daily basis will enhance cost-control practices, and outcomes from signal and data processing will provide enhanced instrumentation and process troubleshooting. These are but a few essential examples of the potential implications of the proposed method for short-term facility's benefits. Furthermore, the analysis of potential retrofit design scenarios would benefit from having an operations-driven cost model based on ABC-like principles in place, providing systematic integration of process and cost data. The model focuses on process costs instead of only on product cost assessment and is able to analyze the implications of operating regimes for future retrofit design options. This leads to the identification of profitable operating conditions in both current and future manufacturing environments and makes possible increased profitability of retrofit design alternatives.

6.1 Contributions to the body of knowledge

A development and industrial application of multiscale signal processing technique that is used for efficient on-line near steady-state detection based on wavelet transform and filtering. The individual methodological contributions can be listed as follows:

- An on-line methodology based on wavelet transforms and filtering which is able to provide highly accurate steady-state representation of small subsystems. This ability can be elaborated to plant-wide systems for industrial applications in cases where the operation under study is fairly stable and simple.

A development and industrial application of plant-wide and steady-state data reconciliation in low redundant systems that is based on process simulation and optimisation

- A practical approach of model-based measurement data validation, that is based on linking wavelet steady-state detection technique with simulation-driven data-reconciliation method into one methodological framework that is capable of providing plant-wide reconciled data sets that representing operating regimes.
- A practical approach for identifying presence of biased measurements in the pulp and paper manufacturing processes.

Development and application of costing method based on real-time data and ABC-like cost accounting principles for assessing actual product operating margins for short and long term company's benefits. The operations-driven cost modeling approach enables the analyst to analyze, characterise and interpret different operating regime alternatives based on real-time process data. More specifically:

- A way to address cost implications of manufacturing processes by linking real-time and plant-wide reconciled process data with advanced cost modelling on the basis of activity-based costing principles. This unique approach helps to characterize and interpret complex cost-process relationships and enables process troubleshooting

- Use of a unique dissection of process data to characterize operating regimes within the complex manufacturing environment. This approach makes it possible to analyze the actual cost of production processes. The resulting information also enables the characterization and interpretation of differences in the operating profit of individual products and regimes and provides a facility which can give new guidance on continuous improvements.
- A systematic approach to address the process-cost impacts and implications of strategic retrofit projects systematically by combining real-time plant-wide process knowledge and advanced cost modelling based on ABC-like principles. The use of real process knowledge and the actual understanding it provides of the facility cost structure enables better forecasting of the future performance of the core business and highlights the most profitable product-mix options.

On-line methodology for assessing operating profits in P&P facility. The methodology is a combination of PSE tools, data processing tools and existing cost accounting methods

- Combination of several techniques, such as data processing, process state identification, data reconciliation and operations-driven costing are used in day-to-day process operation analysis, manufacturing scheduling/planning and facility strategic planning, result in a multidisciplinary tool that provide multiple benefits:
    - With better understanding of the process operation, equipment efficiency can be estimated more precisely. As a result, plant benefits such as improvement in maintenance both for instruments (calibration) and for equipment (cleaning) will emerge.
    - On-line identification of the causes of process problems will make it possible, for example, to locate process leaks and product losses and to detect instrument faults, with the possibility of tracking the origin of the problem back in time.
    - Enhanced process control will be achieved when on-line reconciliation results are used to update process status and overall balances. Moreover, de-noising of process data using wavelets can help process and control engineers to maintain the process closer to the optimum.

- Furthermore, using accurate process data in combination with business, quality, and environmental data will form a knowledge base for continuous improvement in. An operations-driven cost model will be used to assess and evaluate different operating regimes to select the most profitable ones.

6.2 Future work

Overall methodology

The proposed methodology can be used for providing reliable data for further process integration applications, particularly multivariable analysis, real-time optimization, and bridge methods to enhance the identification of energy projects.

- *Supply chain:* The use of actual product margins assessment can be exploited in margin centric supply chain management and planning/scheduling to improve facility and corporate tactical planning

- *Marginal cost of energy:* marginal cost analysis based on real-data and operations-performance analysis can be added to the framework of the developed methodology in order to analyse more flexible forest biorefinery retrofit designs with good strategic fit

- *Corporate strategy:* the developed methodology was applied to characterise, interpret and guide cost savings strategies on one manufacturing facility. The framework of this methodology could be used to analyse all productions sites of a company, to enhance the corporate strategic planning

- *Process flexibility:* the use of actual operating profits for individual products and their operating regimes in multiproduct environment could help to enhance company's profitability when combined with the market price monitoring.

- *Sustainable and knowledge-based manufacturing*: Identification of sustainable production regimes by maximizing expected profit simultaneously with minimizing environmental impact and still meeting the quality requirements could become possible through the development of real-time optimization techniques based on operations-driven representation of regimes. The results could lead to find improved overall process efficiency, including improved production efficiency, lower costs, better quality, better environmental performance, and improved safety.

- *Instrumentation network:* The process driven character of the proposed methodology could be used to study the cost associated with the lack of instrumentation. Another study associated with instrumentation to analyse the cost savings that could be achieved if certain instruments are installed (for instance fibre quality measurement)
- *Wavelet processing for process control:* The study to analyze costs associated with the use of multiscale wavelet processing technique for process control. It is documented that data pre-processing using wavelet improves the accuracy of measurement of controlled variables

REFERENCES

3C Software, *3C Software*, consulted from 2006 to 2011, from www.3Csoftware.com

E-notes, http://www.enotes.com, consulted from January to July, 2011, from http://www.enotes.com/accounting-historical-perspectives-reference/accounting-historical-perspectives

Abramovich, F. and Silverman, B.W. (1998). Wavelet decomposition approaches to statistical inverse problems, *Biometrika*, 85 (1), 115–129.

Alekman, S.L. (1994). *Control for the process industries* (volume 7, No 11, p 62), Chicago: Putman publications.

Arora, N. and Biegler, L.T. (2001). Redescending estimators for data reconciliation and parameter estimation. *Computers and Chemical Engineering*, 25, 1585.

Bagajewicz, M.J. and Sanchez, M.C. (1999). Design and upgrade of nonredundant and redundant linear sensor networks. *AIChE Journal*, 45(9), 1927-1938.

Bagajewicz, M.J. and Jiang, Q. (2000). Comparison of steady state and integral dynamic data reconciliation. *Computers and Chemical Engineering.* 24(11), 2367-2383.

Bagajewicz, M.J. (2001). *Process Plant Instrumentation and Upgrade*. Lancaster, Pennsylvania: Technomic Publishing.

Bagajewicz M.J. and Rollins, D.K. (2002). Data Reconciliation. In Liptak B.G. (ed.) *Instrument Engineers' Handbook* (3rd Edition, volume 3: Process Software and Digital Networks). Bose Roca: CRC Press.

Bahri, P.A., Bandoni, A., Romagnoli, J. (1996). Operability assessment in chemical plants. *Computers and Chemical Engineering*, 20 (Suppl. B), S787.

Bakshi, B.R. and Stephanopoulos, G. (1993). Representation of Process Trends III. Multiscal Exraction of Trends from Process Data. *Computers and Chemicals Engineering*, 18, 267-302.

Bartko, J.J. (1991). Measurement and reliability: statistical thinking considerations. Schizophrenia Bulletin 17(3), 483–489.

Bellec, S., Taiwen J., Kerr, B., Diamond, M., Stuart, P. (2005). On-line processing and steady state reconciliation of pulp and paper mill process data. *91st Annual Meeting Preprints* (Book B, 105-110).

Benqlilou, C. (2004). *Data Reconciliation as a Framework for Chemical Processes Optimization and Control*. PhD thesis, Universitat Politecnica de Catalunya, Barcelona, Spain.

Bethea, R.M. and Rhinehart, R.R. (1991). *Applied Engineering Statistics*. New York: Marcel Dekker.

Bhat A. and Saraf N. (2004). Steady-State Identification, Gross Error Detection and Data Reconciliation for Industrial Process Units. *Industrial and Engineering Chemistry*, 43, 4323-4336.

Bras, B. and Emblemsvåg, J. (1995). Activity-Based Costing and Uncertainty in Designing for the Life-Cycle. In: Huang G.Q., (Ed.), *Design for X: Concurrent Engineering Imparatives*, London: Chapman & Hall.

Brown, C.D. (1999). *Accounting and Reporting Practices of Nonprofit Organizations: Choices and Applications*. New York: American Institute of Certified Public Accountants.

Burrows, G.H., Brown, C.A., Thom, T.W., King, J.M.C., Frearson, J., (2001). Real-time Cost Management of Aircraft Operations. *Management Accounting Research*, 12(3), 281-298.

Cao, S. and Rhinehart, R.R.(1995). An Efficient Method for On-line Identification of Steady State. *Journal of Process Control*, 5(6), 363-374.

Chakraborty, A. and Linninger, A.A. (2002). Plant-wide waste management. 1. Synthesis and multiobjective design. *Industrial and Engineering Chemistry Research,* 41 (18), 4591-4604.

Chen, J., Bandoni, A., Romagnoli, J.A. (1997). Robust estimation of measurement error variance/covariance from process sampling data. *Computers and Chemical Engineering*, 21, 593-600.

Cooper, R., (2000). Cost management: from Frederick Taylor to the present. *Journal of Cost Management*, September/October, 4–9.

Crama, Y., Pochet, Y., Wera, Y. (2001). A discussion of production planning approaches in the process industry. *Core Discussion Paper*, 42, University of Liege.

Crowe, C.M., Campos, Y.A.G., Hrymak, A. (1983). Reconciliation of process flow rates by matrix projection. Part I: linear case, *AIChE Journal*, 29(6), 881-888.

Crowe, C. M. (1992). The Maximum-Power Test for Gross Errors in the Original Constraints in Data Reconciliation. *Can. J. Chem. Eng.*, 70, 1030.

Crowe, C.M. (1996). Data reconciliation - progress and challenges. *Journal of Process Control*, 6 (2/3), 89-98.

Dabros, M., Perrier, M., Forbes, F., Fairbank, M., Stuart, P.R. (2005). Model-Based Direct Search Optimization of the Broke Recirculation System in a Newsprint Mill. *Journal of Cleaner Production*, 13(15), 1416-1423.

Dansereau, L.P., El-Halwagi, M.M., Stuart, P. (2009). Sustainable supply chain planning for the forest biorefinery. *Design for Energy and the Environment: Seventh International Conference on the Foundation of Computer-Aided Process Design* (p. 1101), Breckenridge, Colorado, USA.

Donoho, D. L., Johnstone, I. M., Kerkyacharian, G., and Picard, D., (1995) "Wavelet Shrinkage: Asymptotia?", *J. Royal Stat. Soc*. B, 57, 2, 301-369.

Drucker, P.F., Collins, J., Kotler, P., Kouzes, J., Rodin, J., Rangan, V.K., Hesselbein, F. (2008). *The Five Most Important Questions You Will Ever Ask About your Organization*, San Francisco, CA, USA: Leader to Leader Institute.

Emblemsvåg, J. and Bras, B. (2000). *Activity based cost and environmental management - A different approach to the ISO 14000 compliance*. Norwell, MA, USA: Kluwer Academic Publishers.

Ernst & Young (2003). *2003 survey of Cost Management Accounting*. IMA's Cost management updates, issues 28, 62.

Fadum, O. (1996). Use of process information system provides economic benefit for mills. *Pulp & Paper,* 70 (4), 59-61.

Farquhar, P.H. (1984). Utility assessment methods. *Management Science*, 30 (11), 1283.

Flemig F., Watzdorf, R.V., Marquardt, W. (1998). Identification of Trends in Process Measurements using the Wavelet Transform. *Computers and Chemicals Engineering,* 22, 491-496.

Fogelholm, J. (2004). Reporting and simulation of economic steering information. *27th Annual Congress of the European Accounting Association*. Prague, Czech Republic.

Fogelholm, J., (2000). *Cost Function Modelling in the Paper Industries*. Ph.D. thesis, Helsinki University of Technology, Department of Industrial Engineering and Management, Espoo, Finland.

Friedl, G. and Kupper, H.-U. (2005). Relevance added: ABC with German cost accounting. *Strategic Finance*, June, 56–61.

Gale, R. (2006). Environmental costs at a Canadian paper mill: a case study of Environmental Management Accounting (EMA). *Journal of Cleaner Production*, 14(14), 1237–1251.

Garg, A., Ghosh, D., Hudick, J., Nowacki, C. (2003). Roles and practices in management accounting today: results from the 2003 IMA – E&Y survey. *Strategic Finance* 1–6.

Goldratt, E.M. and Cox, J. (1992). *The Goal: A Process of Ongoing Improvement* (2nd revised edition). New Berlin, Wisconsin, USA: North River Press.

Greenwood, T.G. and Reeve J.M. (1992). Activity based cost management for continuous improvement: a process design framework. *Journal of Cost Management for the Manufacturing Industry* 5(4), 22-40.

Grossmann, I.E. (2005). Enterprise-wide optimization: A new frontier in process systems engineering. *AIChE Journal* 51(7), 1846-1857.

Grossmann, I.E. (2005). Challenges in the New Millennium: Product Discovery and Design, Enterprise and Supply Chain Optimization, Global Life Cycle Assessment. *Computers and Chemical Engineering,* 29, 29-39.

Gruber, P., O'Brien, M. (2002). Polylactides "Natureworks® PLA". In: Doi, Y., Steinbüchel, A. (eds) Biopolymers- polyesters III applications and commercial products. (Vol 4., 235-249) Weinheim: Wiley-VCH.

Harisson, R., Leroux, R., Stuart, P.R. (2003). Multivariate Analysis of Refiner Operating Data From a Thermo-Mechanical Pulp Newsprint Mill. *Pulp & Paper Canada*, 105(4), 24-27.

Helliwell, J.F., Cox, A.J. (1979). Simulation analysis of energy production in the B.C. pulp and paper industry. MIT Energy Laboratory Working paper no. MIT-EL-79-009WP. Lab. for Product and Process Design, Department of Chemical Engineering, University of Illinois, Chicago, IL, USA.

Horngren, C.T., Foster, G., Datar, S.M., Teall, H.D., Gowing, M. (2006). *Cost Accounting: A Managerial Emphasis* (4th Canadian edition). Scarborough, ON, Canada: Prentice Hall Canada Inc.

Horngren, C.T., Sundem, G.L., Stratton, W.O. and Teall, H.D. (1999). *Management Accounting* (3rd edition). Scarborough, ON, Canada: Prentice Hall Canada Inc.

Houtman, C. and Horn, E. (2011). Pilot trials of hemicelluloses extraction prior to thermomechanical pulp production: Part I. Submitted to *TAPPI Journal*.

Huang, H.J., Ramaswamy, S., Tschirner, U.W., and Ramarao, B.V. (2008). A review of separation technologies in current and future biorefineries. *Separation and Purification Technology* 62(1), 1–21.

Hunt, C., Kenealy, W., Horn, E., Houtman, C. (2004). A biopulping mechanism: creation of acid groups on fiber. *Holzforschung* 58, 434-439.

Hytönen, E. and P. R. Stuart (2010). "Biofuel Production in an Integrated Forest Biorefinery Technology Identification Under Uncertainty." *Journal of Biobased Materials and Bioenergy*, 4(1): 58-67.

Jackson, C. and Kennedy, J.P. (2000). Paper industry's `E' challenge is leveraging its information. *Pulp and Paper*, 74(6), 5.

Jacob, J. and Paris, J. (2003a). Data Sampling and Reconciliation – Application to Pulp and Paper Mills. Part I: Methodology and Implementation, *Appita Journal*, 56(1), 25-52.

Jacob, J. and Paris, J. (2003b), Data Sampling and Reconciliation – Application to Pulp and Paper Mills. Part II: Case Studies, *Appita Journal*, 56(2), 116-120.

Janssen, M., (2007). Retrofit design methodology based on process and product modeling. PhD thesis, École polytechnique de Montreal, Montreal, Canada.

Janssen, M., Laflamme-Mayer, M., Stuart, P.R. (2003). Survey of data management systems used in the pulp and paper industry, Proceedings, Foundations of Computer-Aided Process Operations (FOCAPO 2003). *Coral Springs FL*, 551–554.

Janssen, M., Laflamme-Mayer, M., Zeinou, M.-H., Stuart, P.R. (2004). Survey indicates mills' need to exploit it systems with new business model. *Pulp and Paper,* 78(6), 46-51.

Janssen, M., Laflamme-Mayer, (2011). Development of an operations-driven cost model for continuous processes. Part I: Framework for design and operations decision making. *Journal J-FOR*, in print

Janssen, M. P., Naliwajka, P., and Stuart, P.R. (2011). Multi-criteria decision making for retrofit process design using process- and supply chain-level criteria. Submitted to *J-4 journal*

Janssen, M., Stuart, P.R. (2004). Business Model Framework for Using Real-Time Process Data for Design Decision Making. *Proceedings of the C-DEN Design Conference*. Montreal, Canada.

Jiang, T., Chen, B., He, X. (2000). Industrial Application of Wavelet Transform to the On-line Prediction of Side Draw Qualities of Crude Unit. *Computers and Chemical Engineering*, 24, 507-512.

Jiang, T., Chen, B., He, X., Stuart, P.R. (2003a). Application of a steady-state detection method based on wavelet transform. *Computers and Chemical Engineering*, 27, 569–578.

Jiang, T., Chen, B., Jasmin, K., Stuart, P.R. (2003b). Strategy for Improving Data Quality for a Kraft Pulp Mill Recausticizing Plant. *Proceedings of the Computer-Aided Process Operations (FOCAPO) Conference* (467-470). Coral Springs, Florida, USA.

Johnson, L.P.M. and Kramer, M.A. (1995). Maximum Likelihood Data Rectification: Steady-State Systems. *AIChE Journal*, 41, 2415.

Jubien, G. and Bihary, G. (1994). *Control for the process industries* (volume 7, No 11, p 64). Chicago: Putman publications.

Juran, J.M. (2004). *Architect of Quality: The Autobiography of Dr. Joseph M. Juran* (1st ed.). New York City, USA: McGraw-Hill.

Kaplan, R.S. and Atkinson, A.A. (1989). *Advanced Management Accounting*. Englewood Cliffs, NJ, USA: Prentice-Hall.

Kaplan, R.S. and Cooper, R. (1998). *Cost & Effect : Using Integrated Cost Systems to Drive Profitability and Performance*. Boston: Harvard Business School Press.

Kaplan, R.S. and Anderson, S.R. (2004). Time-driven activity-based costing. *Harvard Business Review*, 82(11), 131.

Kaplan, R.S. (1988). One cost system isn't enough. *Harvard Business Review*, 66(1), 61.

Kenealy, W., Horn, E., Houtman, C. (2007). Vapor-phase diethyl oxalate pretreatment of wood chips: Part 1. Energy savings and improved pulps. *Holzforschung*, 61(3), 223–229.

Keys, D.E. and Van Der Merwe, A. (1999). Cost management: what insights does German cost management have for U.S. companies? *Management Accounting Quarterly*, 1 (Fall), 1–8.

Knepper, J.C. and Gorman, J.W. (1980). Statistical Analysis of Constrained Data Sets. *AIChE Journal*, 26(2), 260-264.

Korbel, M., Jiang, T., Stuart R.P. (2011a). Steady-state identification for on-line data reconciliation based on wavelet transform and filtering. Submitted to *Computers and Chemical Engineering*. (APPENDIX A)

Korbel, M., Wasik, L., Stuart, P.R. (2011b). Practical methodology for plant-wide process data rectification in the pulp and paper industry. Submitted to *J-FOR Journal*. (APPENDIX B)

Korbel, M. and Stuart, P.R. (2011c). Assessment and interpretation of advanced cost data for process improvement in an integrated newsprint mill, submitted to *International Journal of Production Economics*. (APPENDIX C)

Korbel, M. and Stuart, P.R. (2011d) Product margin assessment for process cost-impact analysis, submitted to *Industrial and Engineering*. (APPENDIX D)

Korbel, M. and Stuart, P.R. (2011e). Online methodology for operations-driven cost assessment of operating regimes using real-time process data, submitted to *TAPPI journal*. (APPENDIX E)

Korbel, M. and Stuart, P.R., (2012). Cost Integration Methodology and the Forest Biorefinery. In El-Halwagi, M, Stuart, P.R. (Eds). *Integrated Biorefineries: Design, Analysis, and Optimization.* CRC Press/Taylor and Francis.

Kuehn D.R., Davidson H. (1961). Computer Control II. Mathematics of Control, *Chem. Eng. Process* 57: 44-47.

Laflamme-Mayer, M. (2008). Process-based modeling and analysis methods for supply chain management in the pulp and paper industry. PhD thesis, École Polytechnique de Montréal, Montréal, Canada.

Laflamme-Mayer, M., Janssen, M., Stuart, P.R. (2011). Development of an operations-driven cost model for continuous processes, Part I: Framework for design and operations decision making. *Journal of Science and Technology for Forest Products and Processes* (J-FOR), 1(1).

Laflamme-Mayer, M., Linkewich, J., Stuart, P. R. (2005). Methodology for using realtime process data for supply chain optimization in the pulp and paper industry. *91st Annual Meeting Preprints* (Book C, 279-286).

Lafourcade, S., Labidi, J., Koteles, R., Gélinas, C. Stuart, P.R. (2003). Thermal pinch analysis with process streams mixing at a TMP-newsprint mill. *Pulp and Paper Canada*, 104(12), 74-77.

Lail, P.W. (2000). ERP and the paper industry: Where is the fit? *PIMA's Papermaker*, 74 (4), 32-38.

Levin, H.M. and McEwan, P.J. (2000). *Cost-Effectiveness Analysis: Methods and Applications* (2nd editon). Thousand Oaks, CA, USA: Sage.

Liebman, M. and Edgar, T. (1988). Data Reconciliation for Nonlinear Processes. *AIChE Annual Meeting,* Washington, DC, USA.

Liebman M.J., Edgar T.F., Lasdon L.S. (1992). Efficient Data Reconciliation and Estimation for Dynamic Processes Using Nonlinear Programming Techniques. *Computers Chem. Eng.*, 16, 963-986.

Madron, F., and Veverka, V. (1992). Optimal Selection of Measuring Points in Complex Plants by Linear Models. *AIChE Journal*, 38(2), 227.

Mah, R.H.S., Tamhane, A.C. (1982). Detection of gross errors in process data. *AIChE Journal*, 28(5), 828–830.

Malik, S.A. and Sullivan, W.G. (1995). Impact of ABC information on product mix and costing decisions. *IEEE Transactions on Engineering Management*, 42, 171-176.

Mallat, S., Hwang, W.L. (1992). Singularity Detection and Processing with Wavelets. *IEE Transactions on Information Theory*, 38, 617-641.

Mansoornejad, B., Chambost, V., Stuart, P. (2010). Integrating Product Portfolio Design and Supply Chain Design for the Forest Biorefinery. *Computers and Chemical Engineering,* 34(9), 1497-1506.

Laflamme-Mayer, M, Janssen, M., Stuart, P. (2011). Development of an operations-driven cost model for continuous processes, Part I: Framework for design and operations decision making. *Journal of Science and Technology for Forest Products and Processes (J-FOR)*, 1(1).

McBrayer, K.F., Edgar, T.F. (1995). Bias detection and estimation in dynamic data reconciliation. *J. Process Control*, 5(4), 285–289.

McDermott, T. (2002a). Getting value from IT investments. *Solutions!,* 85, 40-41.

McDermott, T. (2002b). Liberate your business with IT. *Solutions!*, 85, 47-48.

Meyer-Pinson, V., Ruel, K., Gaudard, F., Valtat, G., Petit-Conil, M., Kurek, B. (2004). Oxalic acid: a microbial metabolite of interest for the pulping industry. *Comptes Rendus Biologiques,* 327(9–10), 917–925.

Metso (2010), *METSO, inc.*, consulted from 2010-2011, http://www.metso.com

Mueller, D.C. (1990). *The Dynamics of Company Profits: An International Comparison*. New York, USA: Cambridge University Press.

Narasimhan, S. and Mah, R.S.H. (1987). Generalized likelihood ratio method for gross error identification. *AIChE Journal*, 33, 1514.

Narasimhan, S., and Jordache, C. (2000). *Data reconciliation and gross error detection: An intelligent use of process data*. Houston, TX, USA: Gulf Publishing Company.

National Association of Accountants (1985). *Statements on Management Accounting (4B): Allocation of Service and Administrative Costs*. Montvale, NJ, USA: National Association of Accountants.

Nounou, M. and Bakshi, B.R. (1999). On-line Multiscale Filtering of Random and Gross Errors without Process Models. *AIChE Journal*, 45(5), 1041-1058.

Quesada-Pineda, H., (2004). The ABCs of Cost Allocation in the Wood Products Industry: Applications in the Furniture Industry. *Virginia cooperative extension*, publication 420-147.

Robinson, H. (1997). Using poka-yoke techniques for early defect detection. *Proceeding of the Sixth Annual Conference on Software Testing, Analysis and Review*, 119-142.

Rother, M. and Shook, J. (1999). *Learning to See: Value-Stream Mapping to Create Value and Eliminate Muda.* Lean Enterprise Institute.

Roussel, G. (2003). Real-time process data optimized electricity purchasing. *Pulp and Paper Canada*, 104 (12), 16-19.

Sadhukhan, J., M. A. Mustafa, et al. (2008). "Value analysis tool for feasibility studies of biorefineries integrated with value added production." *Chemical Engineering Science,* 63(2008): 503-519.

Sammons Jr. N.E., Yuan W., Eden M.R., Aksoy B., Cullinan H.T. (2008). Optimal Biorefinery Product Allocation by Combining Process and Economic Modeling. *Chemical Engineering Research and Design*, 86(7), 800-808.

Sánchez M. and Romagnoli, J. (1996). Use of Orthogonal Transformations and Classification/Data Reconciliation. *Comp. Chem. Engng*, 20, 483- 493.

Scharpf, S., (1999). ERP & MES: Information revolution sparks an IS evolution. *PIMA's papermaker*, 81(3), 42-48.

Serth, R.W. Serth and Heenan, W.A. (1986). Gross error detection and data reconciliation in steam metering systems. *AIChE Journal*, 32, 733.

Steen, M. and Steensland, T., (1994). Real-time profit monitoring and activity-based cost management. *TAPPI Journal*. 77(2), 105–111.

Stettler, M., Marechal, F. (1999). *Efficient energy conversion in the pulp and paper industry: application to a sulfite wood pulping process*.

Sundholm, Jan. (1999). Mechanicul Pulping. In: Sundholm, J. (Ed). *Volume 5 of the Papermaking Science and Technology Series*. Printed Jyväskylä, Finland: Gummerus Printing.

Tamhane, A.C. and Mah R.S.H. (1985). Data reconciliation and gross error detection in chemical process networks. *Computers chem. Engng*, 17, 409.

Tatsiopoulos, I.P. and Panayiotou, N. (2000). The integration of activity based costing and enterprise modeling for reengineering purposes. *International Journal of Production Economics,* 66 (1), 33-44.

Tjoa, I.B., Biegler, L.T. (1991). Simultaneous strategies for data reconciliation and gross error detection of nonlinear systems. *Comput. Chem. Eng.*, 15(10), 679–690.

Tong, H., (1995). Studies in Data Reconciliation and the Detection of Gross Errors. PhD Thesis, McMaster University, Hamilton, ON, Canada.

Tong H. and C. Crowe, 1996, Detecting Persistent Gross Errors by Sequential Analysis of Principal Components. *Comp Chern. Engng.,* S20, S733-S738.

Turney, P.B.B. (2010). Activity-based costing: An emerging foundation for performance management. *Cost Management* (July/August), 33-42.

U.S. Census Bureau (1998). *Quarterly Financial Report for Manufacturing, Mining, and Trade Corporations*. Series QFR-98-3. Washington: GPO. Available from www.census.gov.

Van Heiningen, A.R.P., (2010). Integrated Forest Products Refinery (IFPR). Final report. University of Maine, USA. Consulted on May 29, 2011 from http://www.osti.gov/bridge/servlets/purl/979929-FuFfqu/

Veverka, V.V. and Madron, F. (1997). Material and energy balancing in the process industries: From microscopic balances to large plants. In *Computer-aided chemical engineering series* (Vol. 7). Amsterdam, The Netherlands: Elsevier.

Willson, J.D. and Colford, J.P. (1990). *Controllership: The Work of the Managerial Accountant*. New York, USA: Wiley.

Yeager, R. (1998). Information system coordinates operation at Alabama River market pulp complex. *Pulp and Paper,* 72, 97-98.

Yeager, R., (1999). Costing software helps Sappi mills make the grade more efficiently. *Pulp and Paper*, 73(12), 63–66.

APPENDIX A –

STEADY STATE IDENTIFICATION FOR ON-LINE DATA RECONCILIATION BASED ON WAVELET TRANSFORM AND FILTERING

Submitted to Computers and Chemical Engineering

Steady state identification for on-line data reconciliation based on wavelet transform and filtering

Milan Korbel, Steve Bellec, Taiwen Jiang, Paul Stuart[6]

NSERC Design Engineering Chair in Process Integration, Department of Chemical Engineering, École Polytechnique de Montréal

PO Box 6079, Succ. Centre-Ville, Montreal Que., Canada H3C 3A7

Abstract

In order to derive higher value operational knowledge from raw process measurements, advanced techniques and methodologies needs to be exploited. In this paper a methodology for online steady-state detection in continuous processes is presented. It is based on a wavelet multiscale decomposition of the temporal signal of a measured process variable, which simultaneously allows for two important pre-processing tasks: filtering-out the high frequency noise via soft-thresholding and correcting abnormalities by analysing the maximums of wavelet transform modulus. Wavelet features involved in the pre-processing task are simultaneously exploited in analysing a process trend of measured variable. The near steady state starting and ending points are identified by using the first and the second order of wavelet transform. Simultaneously a low filter with a probability density function is employed to approximate the duration of a near stationary condition. The method provides an improvement in the quality of steady-state data sets, which will directly improve the outcomes of data reconciliation and manufacturing costs. A comparison with other steady-state detection methods on an example of case study indicates that the proposed methodology is robust and suitable for online implementation.

[6] Telephone +1 514 340 4711 extension 4384; fax +1 514 340 5150

Email address: paul.stuart@polymtl.ca

Nomenclature

A	Distance between measured value and steady state average
B	Distance between successive measurements
C_j	Coefficient of the smoothed signal at scale j
D_j	Coefficient of the detailed signal at scale j
J	Selected wavelet scale for online data treatment
G_1	Filtered distance between measured value and steady state average
G_2	Filtered distance between successive measurements
R	Ratio use to detect steady state duration
S	Sampling time
WT	Wavelet transform
$\frac{dWT}{dt}$	First derivative of wavelet transform
$\bar{x}$	Filtered average
$\alpha_1, \alpha_2, \alpha_3$	Threshold used for steady state starting and ending point identification
$\beta_1,\ \beta_2, \beta_3$	Filtering parameters
$\sum_{i \varepsilon I_j} C_{j,i} \varphi_{j,i}$	Smoothed or approcimated signal
$\sum_{k \varepsilon K_L} D_{j,k} \psi_{j,k}$	Detailed signal
$\varphi_{j,i}$	Discretized scaling function
$\psi_{j,k}$	Discretized wavelet function
σ	Standard deviation
τ_i	Response time associated with variable i

Introduction

With the increasing use of model based techniques in continuous processes, such as process data reconciliation (Jiang et al, 2003a, Korbel et al (a)), plant-wide optimization (Dabros et al, 2005) or advanced operations-driven cost modeling (Korbel et al, 2012), identification of pseudo steady state operating conditions is critical. The efficiency of these applications relies on a near steady state quality as well as on the ability to identify the near steady state process operations. Unfortunately, process measurements are inherently corrupted with various sources of error (instrument miscalibration or malfunction, power supply fluctuation, as well as wiring and process noise), which can lead to misidentification of near steady state process operations. These problems result in process measurements not being used to their full potential. In this paper, a method for online steady-state detection is proposed and its robustness is compared to two known methods taken from literature. Once the steady state data sets are identified and extracted, data reconciliation is applied in order to improve the quality of data used for plant-wide applications.

Pre processing raw process data involves cleansing high frequency noise and elimination of abnormalities in measurements. This process creates operational data with better estimation accuracy. Wavelet de-noising utilizes the temporally redundant information of measurements so that random errors are reduced and denoised trends are extracted. Although these trends are considered to be more accurate than raw measurements they might be inconsistent with process model constraints, therefore reconciliation has to be employed to resolve this conflict. Since it can be argued that the denoised trends obtained by wavelet transform can be considered as data obtained by more accurate instruments (Benqliou, 2001), the inconsistency in data are due to process dynamics itself. Hence the weighting matrix in data reconciliation step can be quantified by systematically defined engineering rules while avoiding the complications with variance/covariance matrix calculation (Korbel et al. a).

The second step after data pre-processing is steady-state detection. False detection of the process steady state can lead to misinterpretation of true process /features, especially if the incorrect steady state data are subsequently reconciled. Under-estimating the true process steady state periods can lead to only partial correction of gross errors (Figure 1a), while over-estimating steady state periods can result in false input to data reconciliation (Figure 1b).

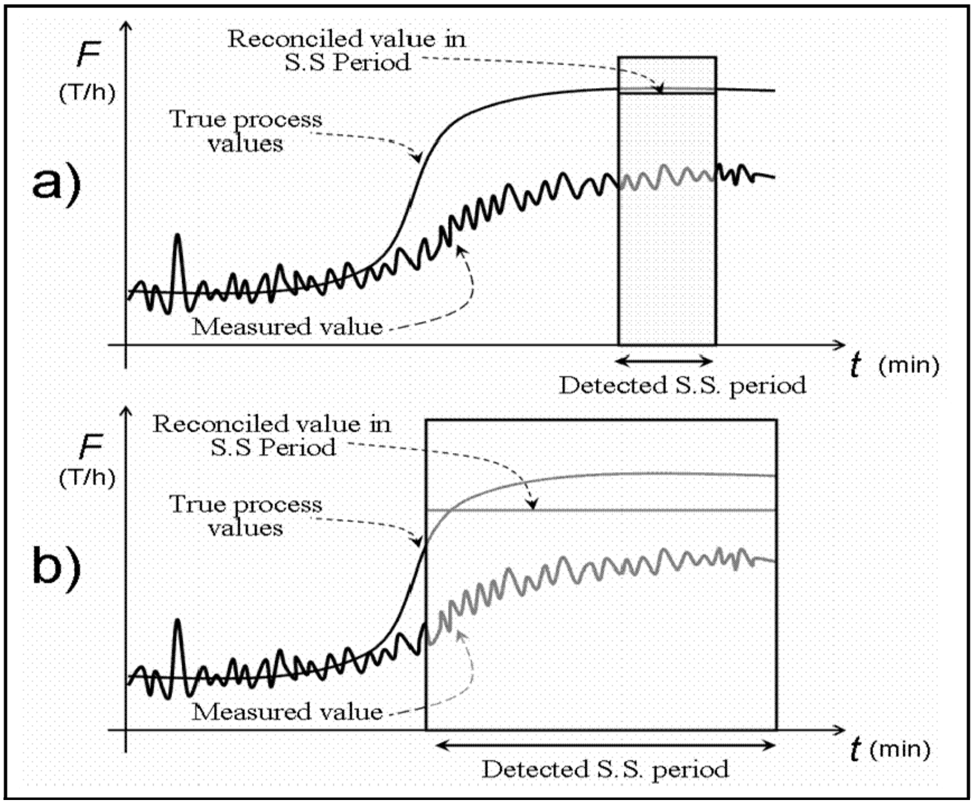


Figure 1: Inaccurate estimation of steady state periods

a) Under-estimated period b) Over-estimated period

A variety of techniques for on-line process status identification have been proposed in the literature. Bakshi and Stephanopoulos (1993) developed a geometric approach for the description of process trends. Cao and Rhinehart (1995) proposed a steady state identification technique based on the comparison of data variances calculated in different ways. In this method, a weighted moving average is used to filter the sample mean. Then, the filtered mean square deviation from the new mean is compared with the filtered squared difference of successive data. This method uses a low pass filter to estimate the mean value. On the one hand, the computational requirements and storage are significantly reduced. On the other hand, low pass filters are less sensitive to the presence of abnormal measurements. Furthermore, using a weighted average to filter the calculated variances creates a delay in the characterization of process measurement frequency. These delays can cause detection problems in periods where the signal properties vary in real time.

Flehmig et al (1998) used wavelet transform features to approximate process measurements by a polynomial of limited degree and to identify process trends. Nounou and Bakshi (1999) used wavelet features to identify and to remove random and gross errors. More recently, Jiang et al (2003a) proposed a wavelet transform (WT) based method for the detection of near steady state periods. The wavelet based multi-scale data processing technique was used to eliminate random noise and abnormalities. Then, the process status was analyzed according to the modulus of the first and second order wavelet transforms. This method can accurately analyze high frequency components and abnormalities. When applying the multi-scale method, the accurate choice of scale is critical. If the scale selected is too low, the WT will be corrupted by high frequency noise, i.e., process status identification is corrupted by temporal features. If the scale selected is too high, then process measurements are excessively smoothed, which creates distortion in the process signal. This creates a deviation from the true process trend and leads to an incorrect reflection of process status.

Jiang et al (2003a) proposed selecting the optimal scale by taking into consideration the response time constants and sampling intervals. This criterion is adequate for off-line purposes, but is not practical for on-line treatment of real time data because on-line measurements can be corrupted with different high frequency features over time. Therefore, the scale choice must be known a priori for on-line wavelet-based treatment of real time data. Furthermore, this method uses the second order WT of the signal to distinguish zero-crossing points from steady state periods. The second WT is directly proportional to the second derivative of the smoothed signal at the sample cutting scale. It is adequate to represent process trends but requires great computational speed and storage. Finally, in the so-called direct approach, linear regression of the measured values is calculated over a data window, and a t-test is performed on the regression slope. This approach is executed over a specified time period, which is not ideal when dealing with real time measurements.

In this paper, a steady state detection technique for on-line estimation of process status is proposed. First, by using multi-scale wavelet data processing (coupled with historical process data analysis in order to select the appropriate wavelet cutting scale), random noise and abnormalities are eliminated. Then, the process status is evaluated with a 3-step method based on wavelet transform and statistical theory. The steady state period starting point is identified using wavelet transform and its first derivative. Then the steady state duration is approximated by coupling a hypothesis test with filtration. Finally, the end point of the period is identified by using wavelet transform features. The robustness of proposed method is addressed in two case studies. In the first case study, the occurrence of pseudo steady-state operation of a small scale process (Stock preparation for paper mill) is being investigated by different methods and compared. The second case study extends the pseudo steady-state assumption to different levels of process variability in order to capture a large scale steady-state operation. Further use of different pseudo steady-state data sets in production cost modeling, points out the magnitude of potential cost inaccuracies.

Multi-scale process data analysis using wavelets

High frequency process measurement features corrupt the process trends, and have a direct impact on process status identification. Jiang et al (2000) proposed a multi-scale wavelet method for processing measurements that is effective for removing noise and detecting abnormalities in real time. The approach for steady state detection presented in this study builds on this multi-scale wavelet methodology.

Generally, the multiscale processing techniques exploit a succession of approximation functions of increasing scale U = {Ui} (Figure 2). At each level of approximation, the smoothed function (or a signal) U_i has its corresponding detail function Yj. This detailed signal is the orthogonal complement of approximation function U_i in its higher resolution space U_{i+1}, i.e.

$$U_{i+1} = U_i \oplus Y_i, \qquad (1)$$

And therefore the multiscale methods can approximate any function by decomposing the high resolution function into a low resolution approximation function Vjo and a succession of detail functions of increasing resolution (Flehming, 1998).

.

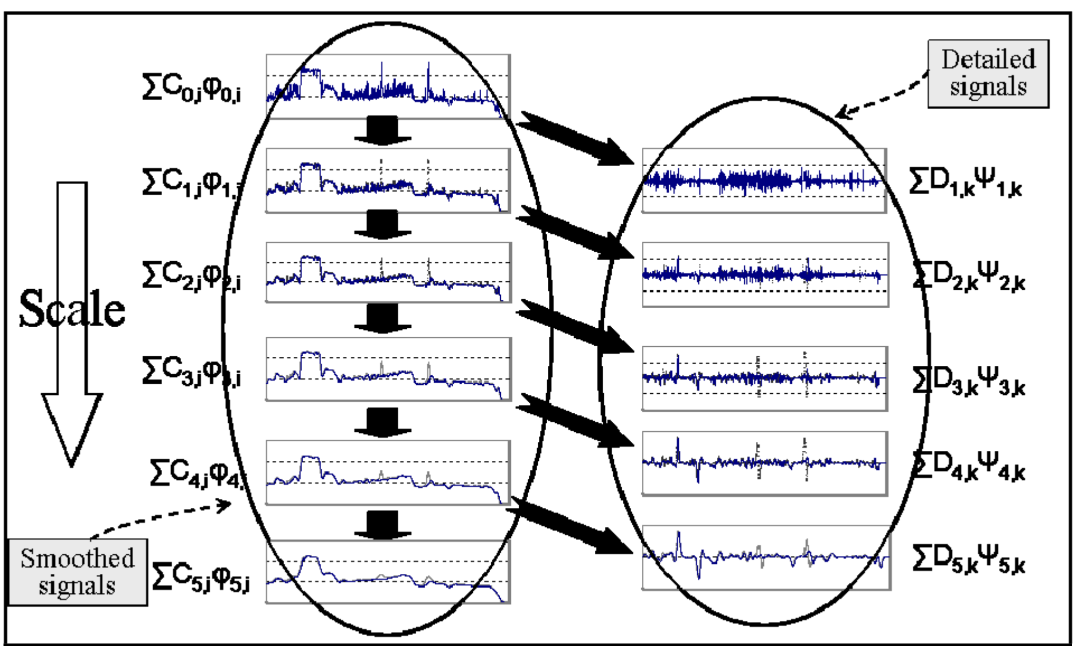


Figure 2: Multi-scale wavelet representation of real time measurements

The theory of wavelet transformation and its use in data processing is well documented and reader should refer to Flehming, 1998 or Jiang, 2002 for more details.

Abnormality detection

Abnormalities are defined as high amplitude peaks of short duration. In other words, they are errors represented as large changes at high frequency. Such changes in real time can be detected using the first order WT, which is proportional to the first derivative of the smoothed signal (Equations 2 and 3). Since the extrema of the first derivative indicate fast changes in the function under study, one can detect such changes in a set of measurements using the first order WT (Jiang et al 2003a) and remove them from the process measurements.

For the first order WT:

$$\psi_j(t) = \frac{d\varphi_j(t)}{dt} \qquad (2)$$

$$WT_j f(t) = f * \varphi_j(t) = f * (2^j \frac{d\varphi_j}{dt})(t) \qquad (3)$$

Abnormality detection is very important task in data processing due to the fact that if pre-processing via filters is applied the spikes will distort process trend (Shankar, 2000). Bakshi and Stephanopoulos (1994) proposed a wavelet based approach for multiscale extraction of trends, which is capable of characterizing different process features according to the corresponding information varying with successive scales. The method proposed by Jiang (2002) and used in this study is particular for identification of abnormalities at a single scale, usually at the finest scale.

Process data de-noising via thresholding

Real time measurements contain noise at a higher frequency than the searched process trend. Pre-processing raw measured data by means of trend analysis involves a de-noising of data and elimination of abnormal data in measurements which in turn leads to better estimation accuracy. Wavelet de-noising utilizes the temporally redundant information of measurements. These trends are theorised to be more accurate than their measurements though they are usually inconsistent with underlying process model, therefore reconciliation has to be employed to resolve this conflict. In a way, the wavelet noise elimination creates measurements obtained by more accurate instruments (Benqliou, 2003). It can be also argued that wavelet based trend de-noising equalises the uncertainty in process measurements with different standard deviations. Hence the weighting matrix in data reconciliation step can be quantified by systematically defined engineering rules instead of the complex variance/covariance expression in practical applications.

Complete removal of the unsuitable high frequency features will be achieved if the correct cutting scale is employed. According to Jiang and al (2003a), the optimal choice of scale for signal denoising is based on the process dynamic, which is relatively easy to approximate off-line but cannot be predicted in real time and thus is not useful for a real time application. The technique proposed in the present paper bases the scale choice on historical data, and assumes that the scale choice is time invariant. The inconsistencies related to that assumption will be removed by using a low pass filter in subsequent steps of the applied methodology. As discussed previously, choosing a WT scale that is too high creates a distortion of process measurements, and leads to an inaccurate reflection of process trends. On the other hand, choosing a scale that is too low will leave the smoothed signal dominated by noise and unsuitable temporal features. This second possibility does not affect the true process trend. At an under-evaluated scale, the process trend is still available, but it is corrupted with high frequency measurements. Therefore, in a subsequent step, the corrupted smoothed signal can be refiltered to isolate the process trend from higher frequency perturbations. This opportunity is not possible in the case where the scale is over-estimated due to the distortion created in the signal.

In order to keep the true signal properties intact, one should choose a scale that does not affect the process trend. By studying historical process measurements, one can investigate the optimal cutting criterion for different process operations. The scale employed for on-line implementation can be selected in such a way that high frequency features are mostly deleted, and the true process trend is not affected by signal distortion. To do so, one should test the performance of the optimum cutting scale (for off-line data treatment) proposed by Jiang et al (2000a) on historical data and compare it to the previous scale (filtered data at lower frequencies).

The theoretical formalization of threshold in the context of removing noise via thresholding wavelet coefficients was presented by Donoho (1995). This method estimates threshold by

$$j = \sigma(2\log N)^{1/2} \tag{4}$$

where N is the size of the wavelet coefficient arrays and σ is the noise standard deviation. The rationale for this choice is that the matched filter is theoretically the optimal detection filter. This condition is best suited only for stationary white noise. Recently, some new methods have been presented, which estimates threshold according to wavelet coefficients at different scales (Xu Qiong 2000, Jiang 2002)

Steady state detection based on WT features, filtering, and statistical theory

Three-step methodology to perform steady state detection

Once abnormal measurements and high frequency features are removed, on-line real time steady state detection can be performed. To do so, a 3-step methodology based on wavelet transform and statistical theory is proposed, as follows:

1. The starting point of the steady state period is detected using WT characteristics and its first derivative,
2. High frequency features of the measured signal, which were not eliminated in the first step, are removed by filtering and steady state duration is approximated,
3. Finally, the steady state end point is detected through WT feature analysis.

Since a multi-scale WT has already been performed on the measurement signal, the WT features needed in step 1 and 3 are already known. As a result of this, along with the fact that a low pass filter is used in the second step of the methodology, no significant calculation is necessary to perform the proposed detection test.

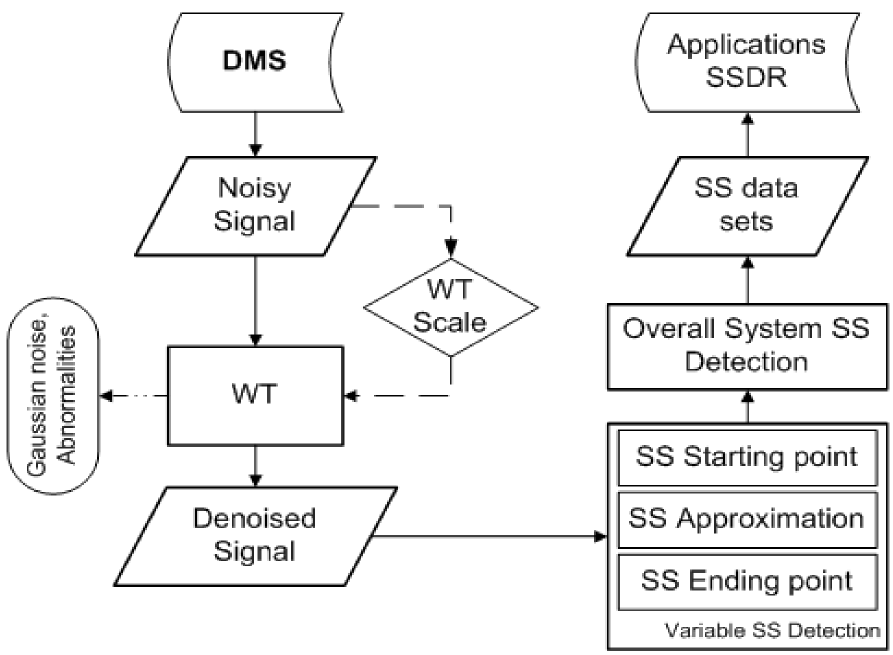


Fig 3: Schematic methodology representation for online near steady-state detection of operating condition

<u>Step 1</u>: Detecting the steady state starting point with WT and the first derivative

The first order WT is proportional to the first derivative of the smoothed signal. Hence the WT measures variation in the smoothed signal and can be used to represent process variations (Jiang *et al,* 2003a). WT extrema indicate fast changes in the data set, while near zero values indicate the presence of a slow change or a zero crossing point of the WT, which corresponds to a local extremum of the measurement signal.

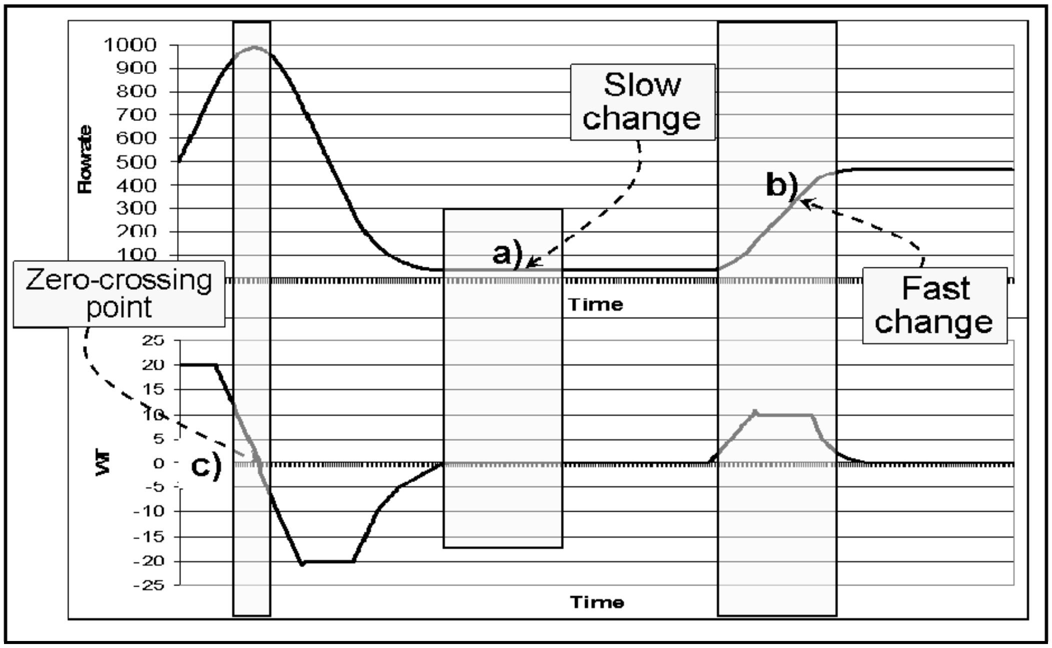


Figure 4: Using wavelet transform features to represent process trends

a) Slow change in process variation: 1st WT and its slope are near zero
b) Fast change in process variation: 1st WT value is not near zero
c) Zero-crossing point: 1st wavelet value is near zero but not the slope value

Since near-zero WT values are associated with both slow change and local extrema, one can rely on the first derivative (slope) of the WT modulus with respect to time to determine the signal properties. In the first case, the WT first derivative is a near-zero value, while a non-zero value is associated with an extremum value of the measurement signal. Based on the above information, the steady state starting point can be detected when the following equations are verified for the first time following a transient period:

$$\begin{cases} \left|WT(f)\right| < \alpha_1 \\ \left|\dfrac{dWT(f)}{dt}\right| < \alpha_2 \end{cases} \quad (4)$$

Figure 6 shows the methodology used to detect the steady state starting point by taking advantage of wavelet transform features. The starting point is identified when near steady state is detected in the measurement signals (equation 4 is verified when the first order WT and its first derivative are simultaneously at near-zero values). One can notice that zero-crossing points are discarded by using this approach.

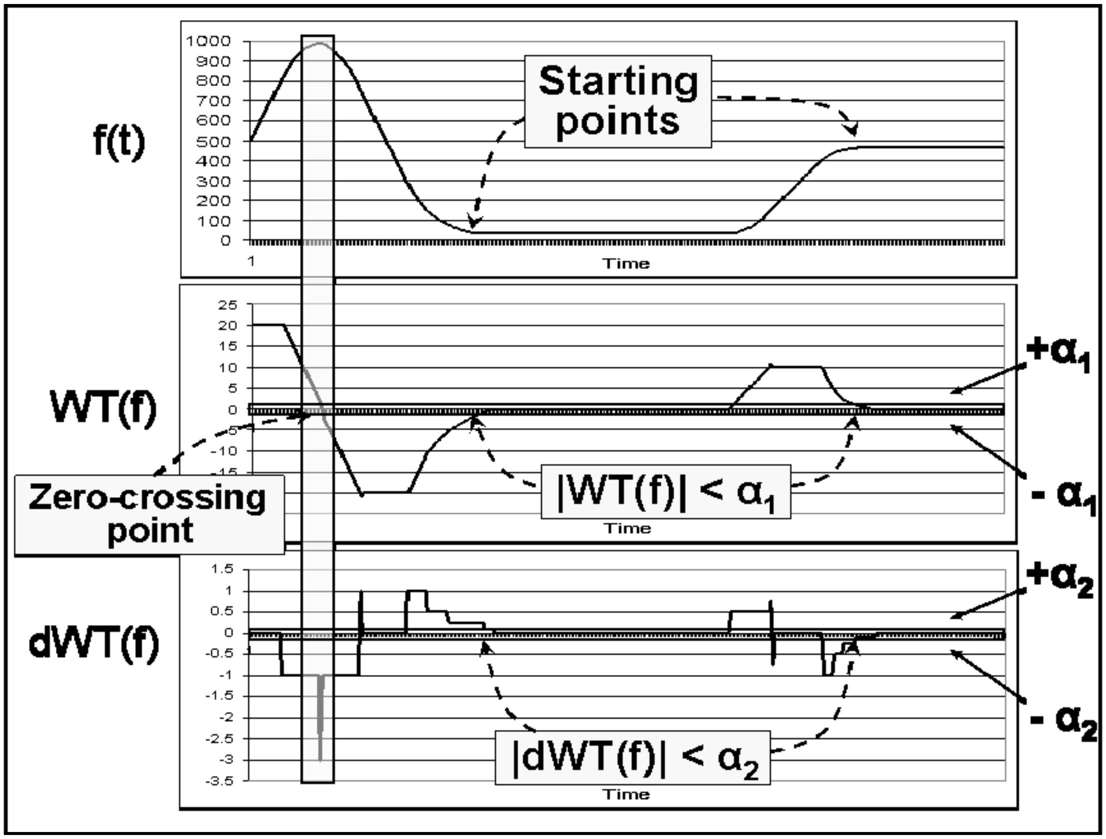


Figure 5: Detection of steady state starting point by using WT and first derivative

<u>Step 2</u>: Filtering steady state data

As discussed previously, over a long time period, the choice of WT scale for removal of high frequency features must be done carefully. Since the on-line scaling choice should be executed in order to insure that distortion is not incorporated into the signal, at certain points in time, some high frequency components are not removed from process measurements. The oscillation created by this phenomenon can lead to false detection by the WT method. Figure 7 shows that when high frequency residuals exceed the minimal threshold acceptance, steady state identification is partial. Therefore, using the method described in the previous section is not suitable for on-line applications over a long time period when selecting the scale as proposed.

To overcome this problem, the steady state period can be approximated through a hypothesis test. Cao and Rhinehart (1995) have proven that, when the process is at steady state and measurements are stationary and independent, there is a probability density function representing the ratio (R) between the filtered squared deviation from the mean and the filtered squared difference of successive measurements. According to Cao and Rhinehart (1995), if the process data mean varies, the R value will be greater than 1 for a period of time. In other words, divergence from the previously identified steady state starting point will be detected if $R > 1$.

The wavelet scales have previously been selected in order to minimize signal distortion, although not every high frequency component is eliminated in some parts of the measurement signal. Thus, an average value of the near steady state variable is approximated at each point in time using a conventional moving average filter (Equation 5). Because they can create distortion in the corrected signal, such filters are not suitable for correction during transient periods, but are useful over steady state analysis. Therefore, a third step will be needed to precisely identify the end point of the steady state.

$$\bar{x}_i = \beta_1 x_i + (1-\beta_1)\bar{x}_{i-1} \quad (5)$$

Figure 8 shows the basic concept behind the hypothesis test proposed to estimate the steady state duration. "*A*" represents the deviation from the mean, while "*B*" represents the difference between two successive data. To avoid corruption by high frequency features that remain in the signal, the two estimators are filtered (Equations 6 and 7) and their ratio is calculated and compared to the threshold value of 1 (Equation 8). At this step of the analysis, steady state is already detected and real time measurements are treated through a wavelet processing method. Therefore, all features with frequencies higher than a certain threshold have been eliminated, and utilization of a low pass filter on the estimators does not affect the steady state approximation of the average value.

$$G_{1,i} = \beta_2 A_i + (1-\beta_2)G_{1,i-1} \quad \text{where } A_i = x_i - \bar{x}_{i-1} \quad (6)$$

$$G_{2,i} = \beta_3 B_i + (1-\beta_3)G_{2,i-1} \quad \text{where } B_i = x_i - x_{i-1} \quad (7)$$

if $R_i = \dfrac{G_{1,i}}{G_{2,i}} \leq 1$, then the process is at near steady state (8)

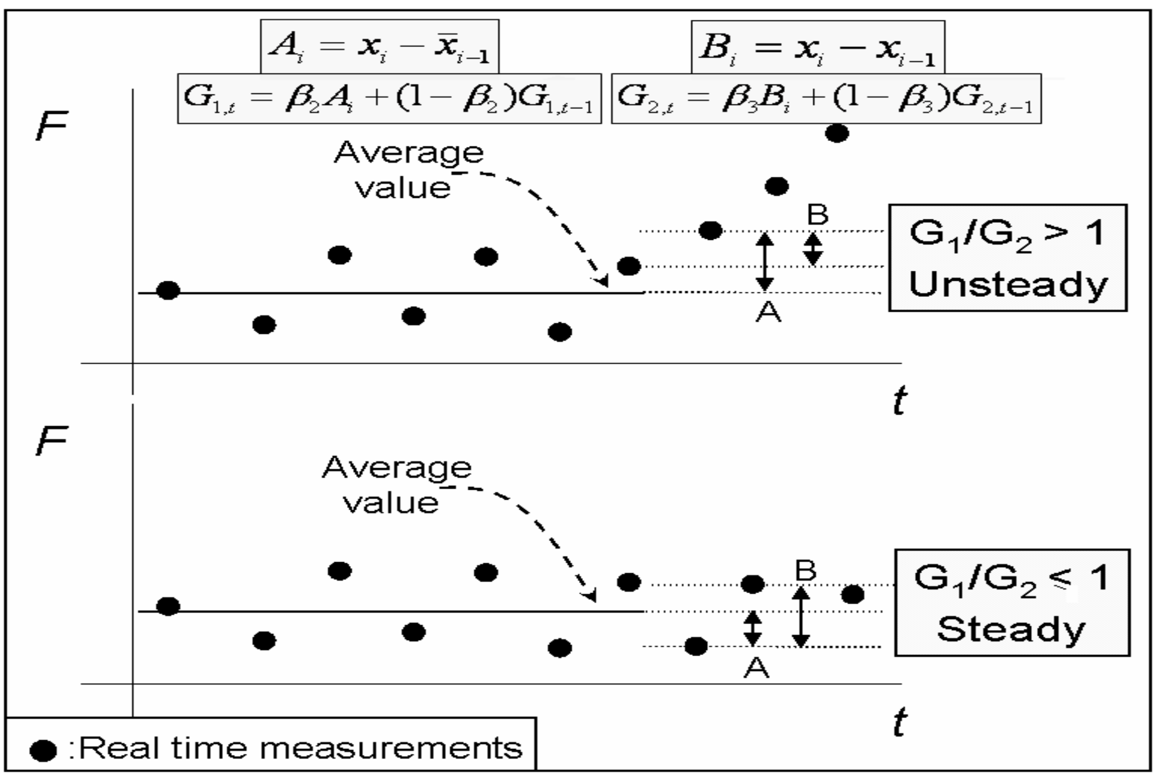


Figure 6: Representation of deviation from the mean and differences between successive data points

Figure 7 shows that as the transient period evolves, the distance between the measured value and filtered average (A_i) increases considerably. A low value of β_2 must be selected in order to over-estimate the steady state period. In this way, the variation in G_1 and R will be minimized. Consequently, the beginning of the transient period is assigned a low weight in the steady state mean approximation, therefore ensuring that the average value of the steady state will be affected only if transient periods occur over a certain time. The counterpart of this approach is that it creates a small over-estimation of the steady state period. In the third step of this methodology, the over-estimation of the steady state period will be corrected.

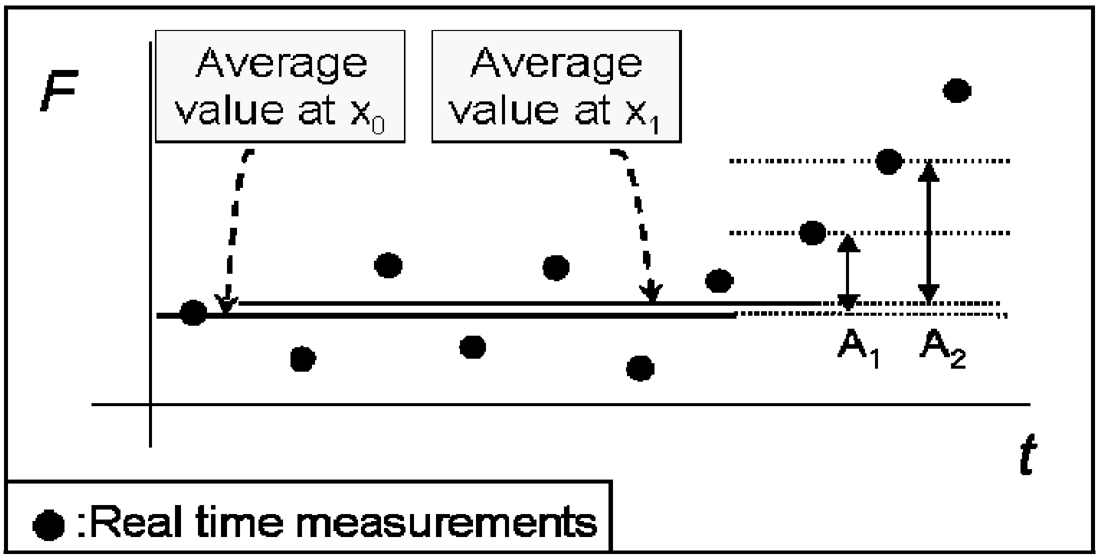


Figure 7: Evolution of deviation from average value (A) during transient periods

3.1.3 <u>Step 3</u>: Detecting the steady state period end point with WT

The final step of the methodology consists in identifying the end point of the steady state period. To do so, the last portion of the period identified in the previous step is analyzed. The WT is used to detect the last moment before the process status becomes transient. Since the WT value is proportional to the rate of change in the measurements, one can detect, by analyzing the data backwards in time, the end point of the steady state period by selecting the first value of the WT under a certain threshold:

$$|WT(f)| > \alpha_3 \quad (9)$$

Since the selected set of data is already at steady state, no analysis related to the zero-crossing is needed here.

3.2 Selecting threshold values and filtering parameters for steady state detection

In order to take advantage of the proposed methodology, an accurate selection of appropriate threshold values and filtering parameters is critical. Since the selection of these parameters is accomplished based on the degree of fluctuation of process measurements, threshold values and filtering parameters can be selected *a priori* based on historic measurements.

3.2.1 Selection of threshold values (α_1, α_2 and α_3)

The threshold values α_1 and α_2 are used to identify the steady state starting point while α_3 is used in the detection of the period's end point. More specifically, α_1 and α_2 are respectively used to ensure that the WT modulus and its first derivative are at near-zero values, while α_3 is helpful to characterize process features in the WT modulus at the end of the steady state period.

The first two threshold values are used to detect the steady state starting point. These are selected according to the degree of fluctuation in the WT and its first derivative. To choose α_1 effectively, it is necessary to select successive measurements under steady state, perform the first order WT, and compute the standard deviation of the WT modulus (σ_{WT}) (Jiang *et al*, 2003a). The selection of α_2 is made to differentiate zero-crossing points from true steady state values. When comparing the true steady state WT modulus to the zero crossing point, one can notice that the steady state values are associated with a slow change in the WT modulus (Figure 5). Therefore, one can perform the first derivative on the WT and compute its standard deviation ($\sigma_{\frac{dWT}{dt}}$). Then the threshold values are selected as follows:

$$\alpha_1 = \sigma_{WT} \quad (10)$$

$$\alpha_2 = \sigma_{\frac{dWT}{dt}} \quad (11)$$

Selection of α_3 is based on the WT features and is used to identify significant changes in the process trends. To choose α_3 effectively, multiple sets of different historic steady state process measurements are selected. Then, the first order WT and its standard deviation are computed. Starting with a λ value of 1, the following threshold value is calculated and the efficiency of the method for the selected data sets is determined:

$$\alpha_3 = \lambda\sigma_{WT} \quad (12)$$

If the end point of the steady state period is detected too late, then one should decrease the λ value (around 0.1) and repeat. If it is detected too early, one should increase the λ value. This process is repeated until a suitable solution for all steady state sets is achieved based on historical data.

3.2.2. Selection of filtering parameters (β_1, β_2 and β_3)

Filtering parameters have a significant impact on the performance of the second step of steady state detection:

- β_1 is useful to eliminate the presence of residual random noise in the treated measurement signal.
- β_2 is useful to help approximate the length of the steady state period.
- β_3 is useful to insure that all steady state periods are detected.

To avoid, as much as possible, undesired results due to unexpected on-line events, the investigation of the optimal filtering parameter must be performed simultaneously on different historic steady state periods. The tuning of optimal filtering parameters should be influenced by the sampling period associated with the instrument used, the response time of the selected variables, and the preliminary data treatment carried out on the raw measurements. In order to select the appropriate parameters, one should follow these steps (inspired by Bhat and Saraf, 2004) :

1. Select a starting value for β_3 (around $\frac{S}{\tau_{mean}}$)
2. Select a starting value for β_2 (around $\frac{S}{\tau_{long}}$)
3. Select a starting value for β_1 (around $0.05(J+1)$)
4. Increment β_1 slowly while the filtered average is affected by the high frequency component
5. Increment β_2 slowly while every steady state period is over-estimated
6. Increment β_3 slowly while all steady state periods are detected. Return to step 5, until convergence is achieved for β_2 and β_3

4. Case studies

4.1 Case study I: Application of the Methodology to a Paper Machine Process

The robustness of the method was investigated on a simulated data set inspired by a pulp stock preparation system at an integrated Kraft paper mill. The plant is equipped with a process data management system (PI system from OsiSoft inc.) that stores the data approximately every 10 seconds. The paper machine under study has multiple process operations associated with different characteristics of the process measurements and high frequency features of the signal.

Hypothetical true process data were created in order to systematically analyze method's performance. Based on historical data, a compilation of characteristics (mean and noise features) associated with the different process operating regimes was performed. Once those characteristics were known, hypothetical data sets with known errors were built according to the real process flowsheet (Figure 10) and operations. Since no down sampling of process measurements is needed to use multi-scale WT for real time process measurements, the generated data are created in order to match the same sampling rate as the real process measurements (each 10 seconds).

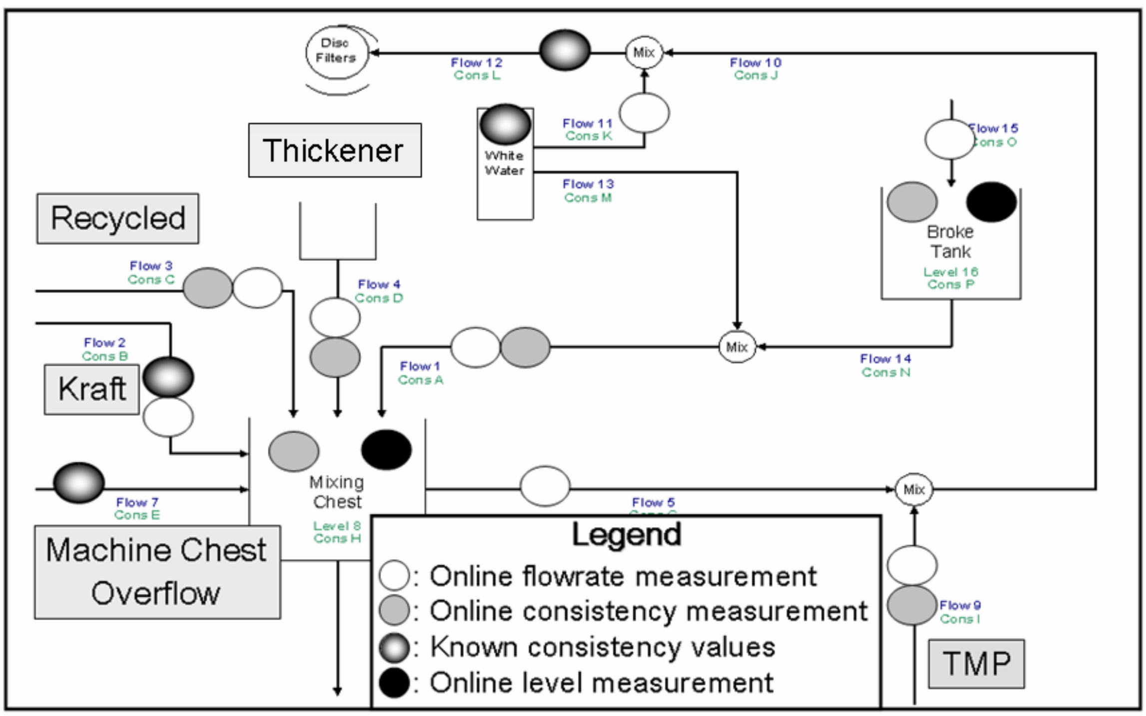


Figure 8: Paper machine stock approach system used to evaluate the methodology

To demonstrate the representativeness of the created data sets, the system operating regimes were analyzed and compared to the created data over a one-week period. Identified regimes were reproduced and errors were added to create a hypothetical corrupted set of data (generated data). Figure 11 shows the results of the investigation for a particular flowrate measurement over time. Although the general characteristics of the two signals are the same, the hypothetical data can contain more or less noise in particular regions, and some parts of the signal can be corrected to fit the overall mass balances around the system.

Performance indices

In order to investigate the efficiency of the proposed methodology, it was compared to two other methods used to detect steady state. The comparison was carried out by calculating, for each method, the type I (steady state is not detected) and type II (false detection) errors related to steady state detection. The hypothesis used to perform this test is:

H_0: Process is at steady state

H_1: Process is not at steady state

The accuracy and precision of the steady state detection technique on on-line steady state data reconciliation was also evaluated. Accuracy was calculated for each steady state period as the relative error between the true process value and the reconciled measurement. Precision is defined as the standard deviation of such errors. Therefore, low values for those two indices indicate high accuracy and precision.

Efficiency of the on-line steady state detection method

The proposed methodology for on-line steady state detection was compared to 1) the method based on first and second order WT (Jiang *et al,* 2003a), and 2) the method of Cao and Rhinehart (1995). For the first method, WT based parameters were selected according to the description in Jiang *et al* (2003a), while optimum parameters for the second method were selected based on the approach proposed by Bhat and Saraf (2004). The investigation was performed on the simulated data sets, which represent 80 hours (sampling at 10-second intervals) of steady state operation divided into 72 different scenarios over a one-week period.

Figure 9 shows the quantity of type I errors related to the identification of true steady state periods for each sampling point in time. One can notice that, for the system under study, the application of the proposed method reduces considerably the incidence of type I error. This can be explained by the presence of a large amount of different high frequency features in the measurements.

1. Filtering the raw measurement variance with a low pass filter can lead to misidentification of process trends when high frequency features undergo sudden variation
2. Utilization of the wavelet features in a conventional way, using a single cutting scale per variable, can lead to under-scaling or over-scaling problems when dealing with signals associated with multiple frequency features

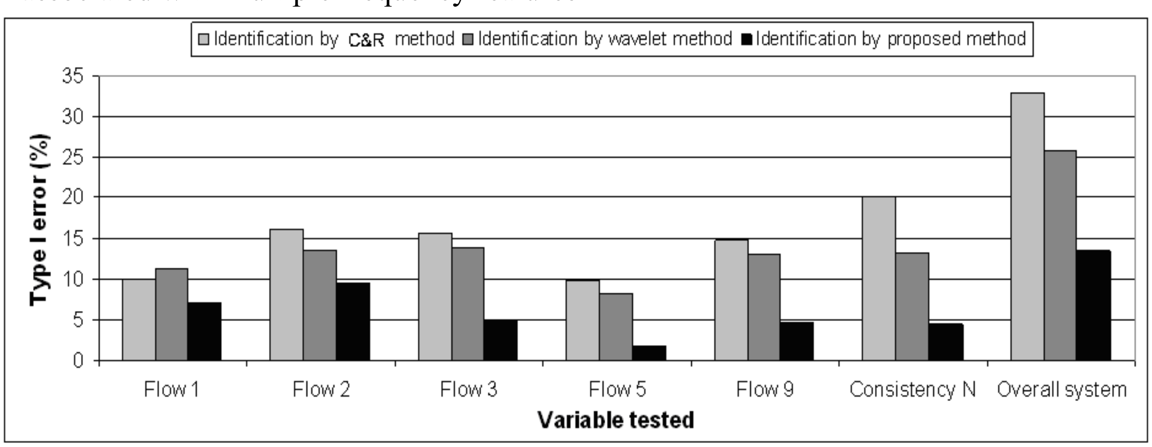


Figure 9: Type I errors related to on-line steady state detection

In order to reconcile the steady state data, one must first identify the overall system steady state periods. Figure 10 shows the quantity of type II errors related to the identification of true steady state periods. One can notice that the effect of the chosen steady state identification method does not have a high impact on the type II errors related to the overall system status recognition. Since every variable under study must be near steady state to record the overall system steadiness, type II errors related to each component are diluted in the process of identifying the overall system status.

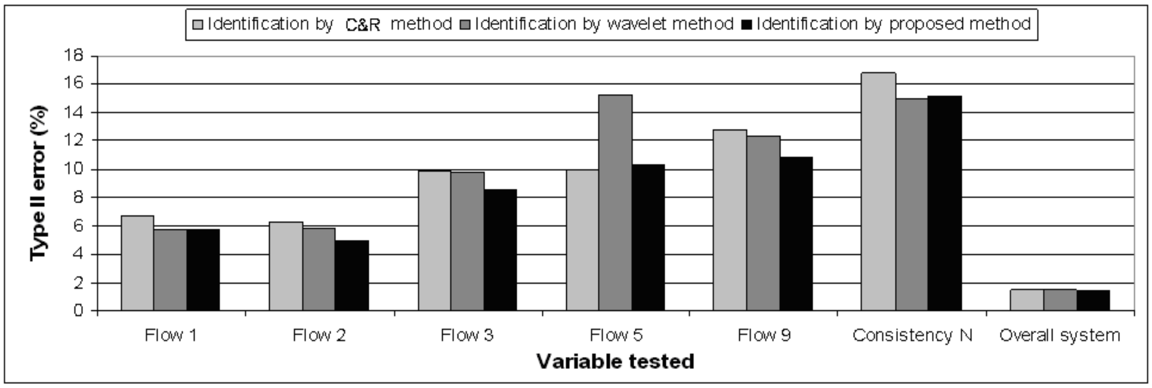


Figure 10: Type II errors related to on-line steady state detection

Impact on on-line steady state data reconciliation

To analyze the impact of steady state detection on on-line reconciliation efficiency, the results of the steady state analysis by the three different methods were used to reconcile the data. Experimental data were reconciled by applying the exact same reconciliation method to each data set.

Overall methodology for steady state data reconciliation

To reconcile on-line steady state data, the overall methodology proposed by Bellec *et al* (2006) was used. First, to correct and filter abnormal measurements and high frequency features, data are analyzed based on wavelet features. Then, overall system steady state detection is performed. This step ensures the selection of near steady state data sets for steady state reconciliation. Finally, the data are reconciled using the Sigmafine software from OsiSoft Inc.

Steady state data reconciliation results

Figure 11 summarizes the improvement of the average accuracy of the steady state data before and after the application of the different methodologies on the simulated data sets (72 steady state periods). Here, the average relative error is used to represent accuracy in the data sets. Therefore, low values on the figure indicate a high accuracy. Equation 13 shows how the average relative error was calculated.

$$\%\mathrm{RE} = average\left(100\% * \frac{|\mathrm{Targeted.value} - \mathrm{measurement}|}{\mathrm{Targeted.value}}\right) \quad (13)$$

One should note that the results obtained from this equation and shown in Figure 11 represent an average improvement for each variable over the 72 steady state periods. Therefore, even a small improvement shown on the figure can represent a large improvement over a particular period of time. For example, Flow 9 was significantly corrected for 1 period containing a gross error (Figure 11) and was marginally corrected over the 71 other periods. Correction for that specific period resulted in a 87% reduction of the error, but the average improvement presented in Figure 14 is small due to the other period for which no steady state correction was performed.

The choice of steady state identification method has a direct impact on the quality of the steady state data reconciliation results. Generally, better status identification leads to improvement in accuracy at the reconciliation step (see Figure 1 for details). To capture the full potential of the results presented in Figure 11, one should simultaneously use the results in Figures 10 and 9. By taking into consideration the fact that the reconciliation is performed only on periods detected as being near steady state, one should notice that there are more periods available for correction by applying the proposed methodology. However, only a small improvement is observed in Figure 11 since the corrections obtained for the additional periods are small due the absence of gross errors (irregularity in the conservation of mass around the system).

An analysis of Figure 11 indicates that for some specific variables (cons N, level 8, and level 16), the relative error does not seem to be reduced by the utilization of the new method. Results on the graph for those variables are multiplied by factors of 10 and 100. These are small errors in the range of 0.05%, which are more difficult to correct using data reconciliation techniques.

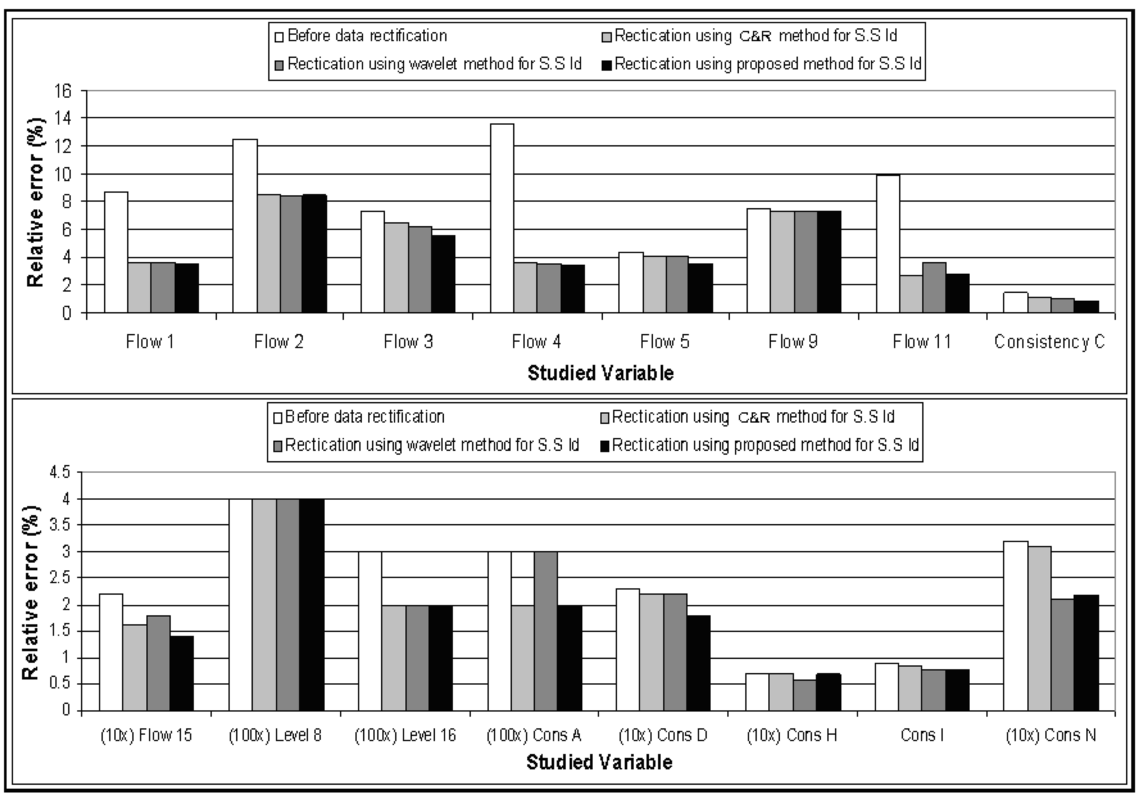


Figure 11: Improvement in accuracy achieved by the application of different on-line steady state detection techniques

Conclusions to case study I

When compared to other on-line steady state methods in a case study based on simulated data, it was shown that the proposed methodology reduced type I errors by at least 46 % for overall steady state status identification, while type II errors were also improved, but in a marginal way. In cases where the steady state system was characterized by an inconsistency in the conservation of mass, the reduction of type I errors lead to the improvement of steady state data in the problematic set of measurements. In fact, for several cases, the impact of using the proposed methodology in combination with steady state data reconciliation techniques resulted in a higher accuracy for real time measurements. Using this approach for steady state detection in a general online data reconciliation methodology could potentially lead to multiple benefits for on-line process operations. The advantages include efficient on-line identification of out-of-calibration instruments, improved process control due to the improvement of measurement accuracy, and more accurate process operation planning and optimization.

4.2. Case Study II: Steady-state quality in large scale processes

The proposed method, as shown on the case study I, shows promising results for the online applications applied to small scale processes. However, many plant-wide applications require steady state data sets across the whole plant. Due to the fact that the large scale true stationary state almost never occurs, practical assumptions need to be taken into account. Measurement trend error (MTE) can be used to quantitatively analyze the distance from the steady-state trend (Flehmig, 1998).

The measured trend error (MTE) reflects the deviation of the de-noised measured variable y_M from the notion of a steady-state trend:

$$MTE = \sum_{i}^{N} \frac{1}{\Delta t} \int_{t_0}^{t_0+\Delta t} \|y'_{Mt}\|_2^2 dt$$

Where y'_M refers to the first derivative of the measured variable y_M which is determined by an algorithm presented by Abramovich, 1998. MTE is multivariable measure with respect to each measurement process variable since the time starting point and its length is the same for each.

The large scale steady-state data set quality versus small scale is addressed. The case study process operation used for this purpose is thermomechanical pulping (Figure 12). The plant is divided into several smaller subsystems in which a multivariable steady-state of each is analyzed alongside the overall plant wide dynamics. If and only all subsystems are simultaneously identified to operate at near steady-state, the plant-wide operation is assumed to be at pseudo steady-state as well. However, this condition is very uncommon. The values of wavelet tuning parameters were being relaxed in order to allow for more pseudo steady-states being identified. The quality of each assumption is then carefully addressed by looking at MTE as well as analyzing the potential impact of the steady-state assumption to the production cost analysis.

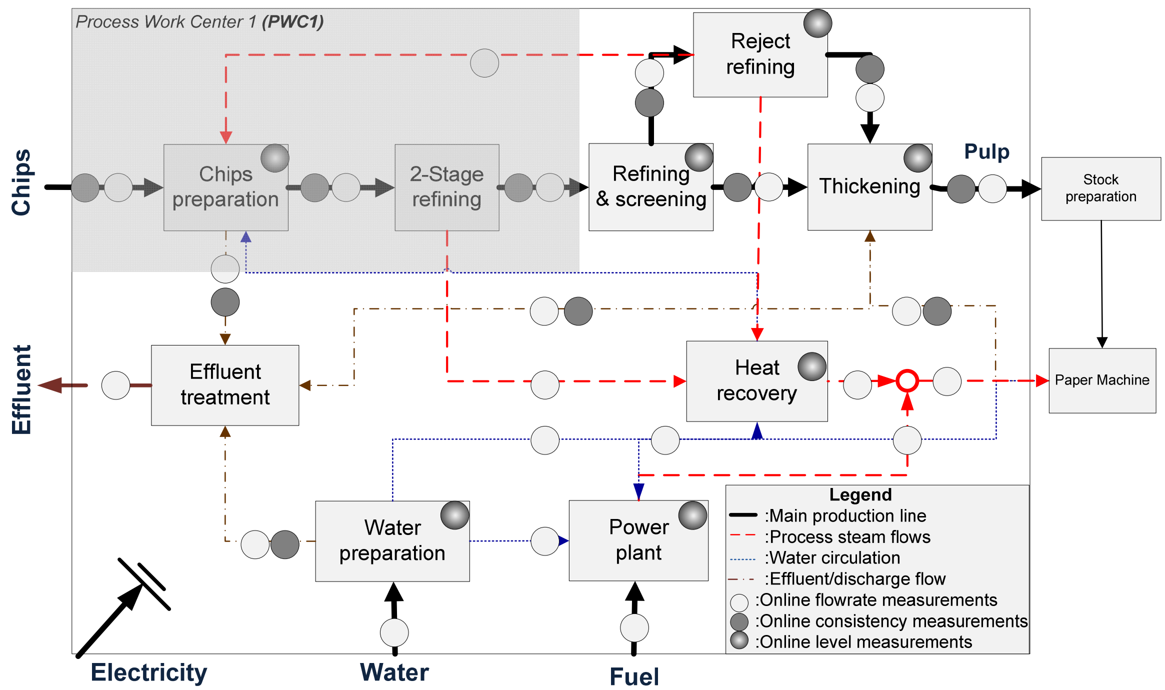


Figure 12: Simplified block diagram of the thermomechanical newsprint process (showing here an example of plant division into seven parts, PWC1represents the firs subsystem)

The analysis is presented on the first process subsystem PWC1 (chips pretreatment), from where seven measured quantities have been chosen to identify when the subsystem is assumed to be stationary. The time snapshot of the analysis was 630 minutes which is the time corresponding to the production of one product within identical process conditions (operating regime).

From figure 13 one can understand that each variable is of different sensitivity to the steady-state occurrence. The Flow01 (warm process water) is at steady state 62% of the time and will increase only by 5% when the threshold value is changed from 0.1 to of 1.2. Further increase in alpha values have no influence on the frequency of steady-state identified in Flow01. On the other hand, variable Flow02 (low pressure process steam) and Flow03 (washed chips flow) stationary state detection strongly depend on the alpha values. Clearly, highest single variable steady-state occurrence creates the potential for multivariable steady-state detection. And therefore the goal is to understand the link between detecting a multivariable steady-state (MSS) and the corresponding level of relaxing steady-state assumption by manipulating the threshold value alpha.

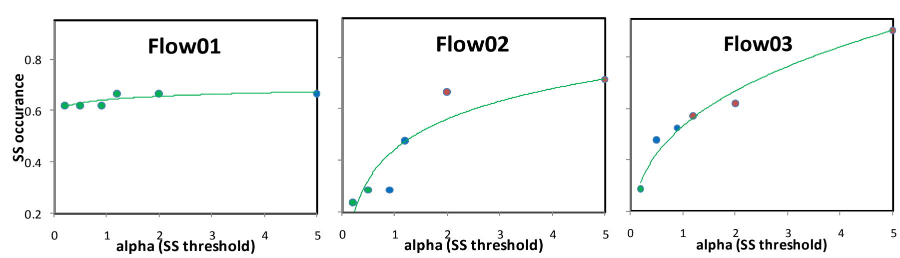


Figure 13: The difference in the rate of steady-state detection relative to the increase of threshold values in key-variables

The threshold values for the measured variables whose steady state occurrence does not influence the overall stationary state detection is tuned to the lowest value. For example alpha values for Flow01 trend detection can be kept at the value α=0.12 as soon as analyst decides using the first tree MSS for higher level analysis. If more are required, then the value is set to α=0.8 for four MSS or increased to α=2.1 for five. On the other hand, for variables that drive the state of the system, one needs higher values of thresholding to achieve relatively small amount of steady-state.

The figure 14 illustrates the evolution of wavelet transform for each key-variables, their corresponding state identification with the use of binary representation (1=steady-state, 0-transient period) and the count of stationary variables per each period. If the count equals the number of variables (e.g. 7 in the case of chips washing process) then the multivariable steady-state (MSS) of a subsystem (PWC1) is assumed. The value of wavelet transform thresholding parameter has been successfully increased from starting value α=0.1 to α=5. The analysis corresponding to values 0.1, 1.2 and 5 are presented on the figure 14. The alpha values were chosen based on standard deviation of wavelet transfer modulus of the key variables that behave dynamically. Clearly, relaxing the alpha values is increasing the single variable steady-state occurrence (as it is presented in figure 13) and inherently an increased number of MSS is achieved. It must be said that the values of alpha that correspond to 5 and above, represents the case where all the dynamics of the process are involved in steady-state assumption. This type of data is used in the classical steady-state data reconciliation procedures, e.g. averaging variables over a certain time period.

This systematic approach for multivariable steady-state detection is very practical; however the analyst must be cautious for certain steady-state assumptions can include transient periods. In order to analyze the errors or uncertainties with associated steady-state assumptions, the quality of pseudo steady-state data sets is investigated next.

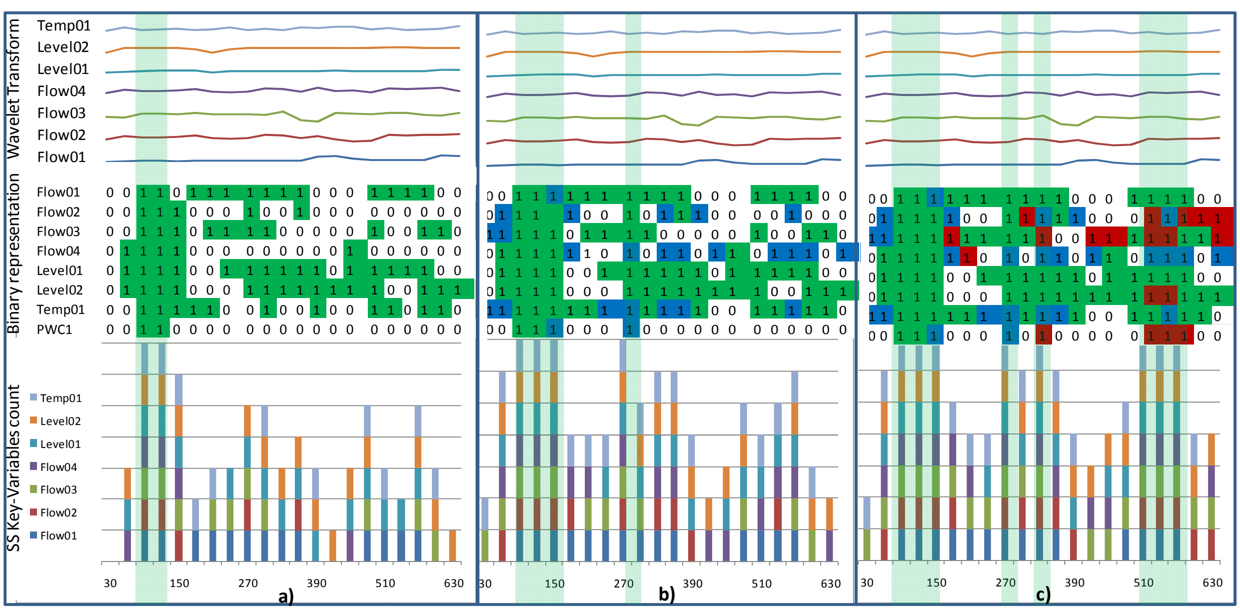


Figure 14: Number of steady state increases with relaxing the values of wavelets threshold (alpha). A: the value of threshold is set to 0.1 leading to two Multivariable Steady States (MSS) identified, B: the value of threshold is set to 1.2 assuming four pseudo steady-states detected (2 new MSS are

represented blue), C: the value of 5 for thresholding has increased stationary statutes to eight (4 new MSS are represented red).

The value of measurement trend error (MTE) is calculated for each MSS data set and its relation to the occurrence of multivariable steady-state is analyzed. This information is critical in deciding the sufficient level of data set quality for the end application such as product cost modeling.

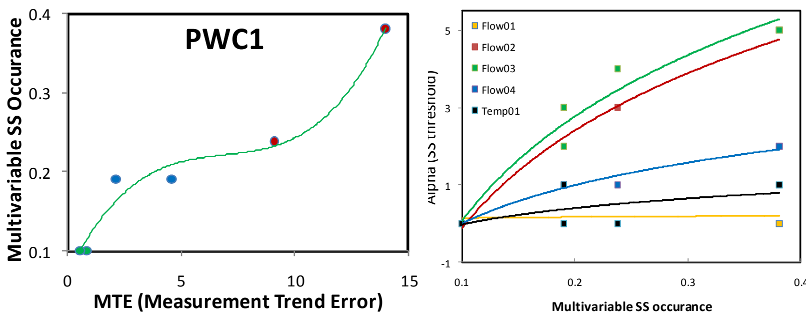


Figure 15: A: MSS occurrence as a function of MTE, B: Sensitivity analysis – the change in MSS occurrence as a function of threshold change

The sensitivity analysis presented on the figure 15B helps to understand the dynamics of each key variable and hence their influence on the overall MSS detection. As one can see from the figure, the variables flow02 and flow03 are the most sensitive to the change in alpha.

Since the pseudo steady-state data sets are used for cost analysis of the production, a direct link between the data quality and the error in cost information can be critical for decision makers. The figure 16 presents an example of cost sensitivity analysis of different measurement entities. The accuracy of calculating the product margin strongly depends on the quality of steady-state of some variables. The change in the values for steady-state thresholding in the case of Flow03 variable, affects the final results for product margin values significantly.

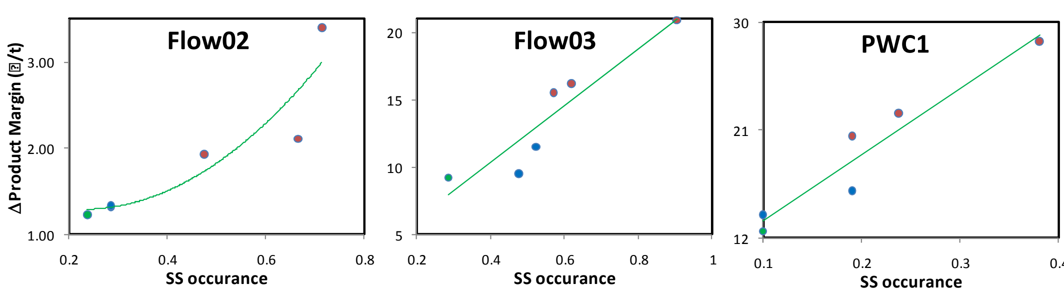


Figure 16: A, B, C – direct link between change in the product margin and the steady-state occurrence for variables flow02, flow03 and the whole multivariable subsystem – PWC1

The above analysis has been done on a few key-variables that are assumed to control the process subsystem (PWC1) state of operation. However, in order to address plant-wide steady state, the remaining six PWCs must be analyzed. The figure 17 presents the outcome of the calculation analyzing the influence of the threshold value on the plant-wide multivariable steady-state (PWMSS).

As one can understand from the figure 17, the plant-wide steady state does not occur when the low values of alpha are used (green values). The first stationary period is detected when alpha is increased to the value α=1.12 (blue values). Further increase produces more SS detected; however, it is clear that many transient periods are included. The decision to use such assumptions for pseudo steady-state is questionable and the decision makers should consult sensitivity analysis for each variable and subsystem, such as the ones presented below on the figures 18-20.

PWC1	0	0	1	1	1	0	0	0	1	0	1	0	0	0	0	0	1	1	1	0	0
PWC2	0	0	1	1	1	0	0	0	1	0	1	0	0	0	0	1	1	1	1	1	1
PWC3	1	0	1	1	1	1	1	1	1	0	1	1	1	0	0	0	1	1	1	1	1
PWC4	1	0	1	1	1	0	0	0	1	0	1	0	0	1	0	0	1	1	1	0	0
PWC5	0	0	1	1	1	0	0	0	1	0	1	0	0	1	1	1	1	1	1	0	0
PWC6	0	0	1	1	1	0	0	0	1	0	1	0	0	1	0	0	1	1	1	0	0
PWC7	0	0	1	1	1	0	0	0	1	0	1	0	0	1	0	0	1	1	1	0	1
PLANT					**1**				**1**		**1**						**1**	**1**			

Figure 17: the binary representation of the plant-wide operation state (1=pseudo stationary, 0=transient period)

The increase in alpha values increases the probability to detect a plant-wide pseudo steady-state; however it may lead to inaccurate cost information and bad-informed decisions. Figure 18A presents the relation between the MSS occurrence in each section of the operation with relation to the plant-wide level. From the figure we can understand that the subsystem PWC3 is relatively the least dynamic part of the mill allowing for higher detection of steady-state. The plant-wide MSS is not detected at the highest accuracy of the method, the threshold must be relaxed at least to a=0.8 in order to identify fist MSS. If more data sets are required then the threshold must be increased to 2.1 or 5 to reach near 20% of MSS occurrence within the analyzed time frame.

As mentioned above, the MTE is a multivariable measure describing the offset of a subsystem or a whole system from the notion of a trend – a steady-state in this case. Therefore one can be interested in analyzing the values of MTE in relation to the increase in plant-wide steady-state occurrence (figure 18B). The alpha values correspond to the maximum threshold value that at least one of the key variables had to employ in order to reach a required number of steady-state. The increase in the thresholding value for some variables is necessary in order to reach at least one plant-wide steady-state data set (figure 18A). The impact of this assumption is then analyzed by looking at the MTE of the system and the error of the end application that is using the assumed steady-state data set. For comparison, the figure 18B shows also the value obtained by the ad-hoc approach that many analysts are using – an average of variables within the time frame.

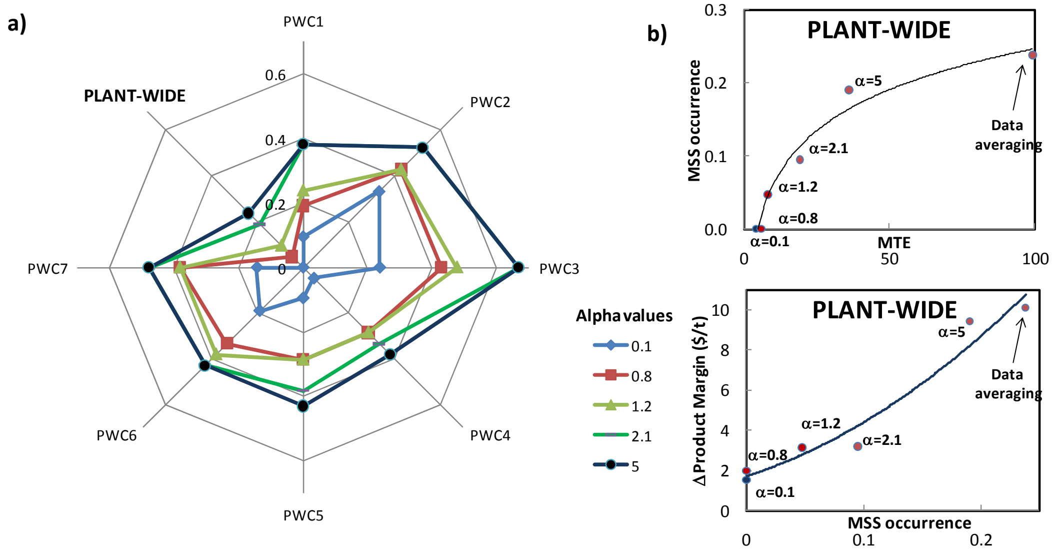


Figure 18: A- the MSS occurrence as a function of threshold values for different subsystems and plant-wide level. B – the relation between the MTE and MSS occurrence as well as the relation between the absolute change in the product margin due to error in the data set the MSS occurrence

Conclusions and final remarks

A robust methodology that is efficiently detecting steady state operating conditions in on-line applications has been proposed. The method identified when variable is at steady state using three signal processing steps, which reduces type I and type II errors related to the identification of steady state periods. The optimum WT cutting scale and parameters were pre-determined based on historical data. In practical applications, these properties will vary over long term periods and the implementation of adaptive techniques would improve the automation of the system.

The method was used in a pragmatic analysis (case study 2) of the data sets quality when different parameters were relaxed. As expected, the system over and under estimate the steady-state periods to some extent. However if plant-wide applications are to be used, this is a very practical approach to extracting near or pseudo steady-state data sets. The extraction process must be done carefully and analyst must take into account the accuracy of his/hers assumptions. Finally, the comparison to the usual steady-state assumptions, via averaging data within the analyzed time period, shows that even the highest threshold values provide more accurate outputs from the end application, in our case, production cost analysis.

The online application of this methodology can be used to automatically extract potential multivariable and plant-wide steady-state periods for higher level applications, for instance real-time optimization or production cost analysis of different operating regimes. The system can be implemented as a part of the information management system at the mill and hence automatically update a steady-state database. By doing this, the cost accounting practices could alter more towards process or operations driven approach which would significantly improve the daily process troubleshooting from a cost-process perspective.

Acknowledgements

This work was completed with the financial assistance of the Natural Sciences and Engineering Research Council (NSERC) of Canada.

References

F. Abramovich, B.W. Silverman, Wavelet decomposition approaches to statistical inverse problems, Biometrika 85 (1) (1998) 115–129.

BAKSHI, B.R., STEPHANOPOULOS, G. (1993). Representation of Process Trends III. Multiscal Exraction of Trends from Process Data. *Computers and Chemicals Engineering*, Vol.18, pp 267-302.

BELLEC, S., JIANG, T., KERR, B., DIAMOND, M., STUART, P.R. (Corrected version submitted in 2006). On-line Processing and Reconciliation of Pulp and Paper Mill Data, *Pulp & Paper Canada*.

BENQLILOU, C.; Tona, R. V., Espuña, A., Puigjaner, L. (2001), On-line application of dynamic data reconciliation. Proceedings of the Conference on Process Integration, Modeling, and Optimization for Energy Saving and Pollution Reduction PRESS 2001, AIDIC: Milano, Italy, May 2001; p 403.

BHAT A.**,** SARAF N.,(2004). Steady-State Identification, Gross Error Detection and Data Reconciliation for Industrial Process Units. *Industrial and Engineering Chemistry*, Vol 43, pp 4323-4336.

CAO, S., RHINEHART, R.R.(1995). An Efficient Method for On-line Identification of Steady State. *Journal of Process Control*, Vol.5, No.6, pp 363-374.

DABROS, M., PERRIER, M., FORBES, F., FAIRBANK, M., STUART, P.R. (2005). Model-Based Direct Search Optimization of the Broke Recirculation System in a Newsprint Mill, accepted to *Journal of Cleaner Production*.

FLEHMIG F., WATZDORF, R.V., MARQUARDT, W. (1998). Identification of Trends in Process Measurements using the Wavelet Transform, *Computers and Chemicals Engineering*, Vol.22, pp 491-496.

HARISSON, R., LEROUX, R., STUART, P.R. (2003). Multivariate Analysis of Refiner Operating Data From a Thermo-Mechanical Pulp Newsprint Mill, *Pulp & Paper Canada*, 105:4, pp 24-27.

JANSSEN, M., STUART, P.R. (2004). Business Model Framework for Using Real-Time Process Data for Design Decision Making". Proceedings - *C-DEN Design Conference*, Montreal, QC. Paper 07B5.

JIANG, T., CHEN, B., He, X. (2000). Industrial Application of Wavelet Transform to the On-line Prediction of Side Draw Qualities of Crude Unit. *Computers and Chemical Engineering*, Vol.24, pp 507-512.

JIANG, T., CHEN, B., HE, X., STUART, P.R. (2003a). Application of a Steady State Detection Method Based on Wavelet Transform, *Computers and Chemical Engineering*, Vol. 27, pp 569-578.

JIANG, T., CHEN, B., JASMIN, K., STUART, P.R. (2003b), Strategy for Improving Data Quality for a Kraft Pulp Mill Recausticizing Plant, Proceedings - *FOCAPO Conference*, pp 467-470.

KORBEL, M, WASIK, L, STUART, P.R. (A), Practical methodology for plant-wide process data rectification in the pulp and paper industry, submitted to International Journal of Production Economics

KORBEL, M, STUART, P.R. (2012), Cost integration methodology and the forest biorefinery, *Integrated biorefineries: design, analysis and optimization, Taylor and Francis, 2012*

MALLAT, S., HWANG, W.L. (1992). Singularity Detection and Processing with Wavelets. *IEE Transactions on Information Theory*, Vol.38, pp 617-641.

NOUNOU, M., BAKSHI, B.R. (1999). On-line Multiscale Filtering of Random and Gross Errors without Process Models. *AIChE Journal*, 45:5, pp.1041-1058.

APPENDIX B –

PRACTICAL METHODOLOGY FOR PLANT-WIDE PROCESS DATA RECTIFICATION IN THE PULP AND PAPER INDUSTRY

Submitted to the Journal of Science and Technology for Forest Products and Processes

(J-FOR)

Practical methodology for plant-wide process data rectification in the pulp and paper industry

Milan Korbel, Larry Wasik[7], Paul Stuart[8]

NSERC Design Engineering Chair in Process Integration, Department of Chemical Engineering, École Polytechnique de Montréal

P.O. Box 6079, Succ. Centre-Ville, Montreal QC Canada H3C 3A7

Abstract

The quality of process data in continuous-process industries significantly affects the performance and the profit to be gained from optimization, control, or cost modelling. Unfortunately, data inherently contain inaccurate information because measurements are obtained with imperfect instruments. In addition, the data may reflect the influence of ambient factors and a lack of sufficient instrumentation. Within the last couple of decades, many data-cleaning techniques have been presented to solve this problem. The papermaking industry faces many challenges in implementing these techniques. These are mostly due to the complexity of the production processes, the unavailability of suitable sensors, and the lack of installed instrumentation. The variables monitored and stored to ensure mill safety and control are often too few for model-based data validation procedures such as data reconciliation. This paper presents a very practical methodology that overcomes the problem of low redundancy in the pulp and paper operation and helps to validate plant-wide pseudo-steady-state data sets for cost modelling. There are tremendous opportunities to use these data sets effectively in practical decision-making related to process operation or design by modelling the link between business and lower-level process data. By doing this, decision-makers on a supervisory level will gain improved knowledge of their operation from both a process and a cost perspective.

[7] Aurel systems, inc

[8] Telephone +1 514 340 4711 extension 4384; fax +1 514 340 5150 Email: paul.stuart@polymtl.ca

Overall goal of the paper

The objective of this paper is to present and discuss a practical process data rectification methodology which gives pulp and paper mills adequate plant-wide information despite a lack of redundancy. The results can offer a high degree of confidence in critical parameters such as process yield and can enable data to be interpreted to make effective and insightful process decisions for cost reduction.

1. Introduction

With today's tremendous progress in information processing and database management systems, the pulp and paper industry has become information-intensive. Vast amounts of data from multiple sources and from different layers of the business are available to mill personnel. To be beneficial, these data must be systematically exploited and combined into practical decision-making information that is useful to the right person at the right time. However, full and effective use of all these data is not a common practice today. Data are mostly used in an *ad-hoc* fashion for problem-solving. A low level of trust in the quality of the data contributes to their lack of usefulness and hence to the absence of systematic tools that could take advantage of good process data. The quality of data in continuous-process industries significantly affects the performance of optimization and control activities and the accuracy of cost accounting. Unfortunately, data inherently contain measurement inaccuracies. Imperfect instruments, the influence of ambient factors (Narasimhan, 2000), lack of instrumentation, and poor calibration all contribute to poor measurements and therefore to inaccurate data.

In this document, a practical approach is discussed that overcomes the data-validation difficulties faced by the papermaking industry. The goal is to gain useful operational knowledge for well-informed decision-making. A wavelet-based signal processing technique is used to obtain a pseudo-steady-state process representation on-line. This step is carried out simultaneously by eliminating signal noise from raw measured variables at multiple scales and by identifying a pseudo-steady-state operating condition. Moreover, the quality of these pseudo-steady-state data sets is further enhanced using simulation-driven data reconciliation techniques by imposing mass and energy balances and other constraints onto them to satisfy conservation laws. The data reconciliation problem is an old industrial application which was proposed first by Kuehn and Davidson (1961) to minimize the error between measured data and the underlying process model. Since they first published their pilot solution to the linear steady-state data reconciliation problem, further studies have led to progress in this area. Crowe (1983) proposed to solve the nonlinear data reconciliation problem by successive linearization. Liebman and Edgar (1998) demonstrated that reconciliation results can be improved by nonlinear programming instead of successive linearization when solving the nonlinear data reconciliation problem. Tjoa and Biegler (1991) showed that nonlinear programming together with a contaminated normal (Gaussian) objective function other than the least-squares objective function can improve the results further. Many other developments in data reconciliation and gross error detection have been proposed in numerous papers (Johnston & Kramer, 1995, Arora & Biegler, 2001).

The most common estimator used for data reconciliation is the weighted least-squares estimator, which is very sensitive to the potential presence of systematic errors, often referred to as gross errors. If gross errors exist in the measurement data, the weighted least-squares estimator will yield incorrect estimates which will then significantly deflect reconciliation of other measurements. The critical task of identifying the presence of gross errors and estimating their values remains a challenge in practical industrial applications. Several methods to solve this problem have been proposed, for instance, the measurement test gross-error detection method presented by Tamhane and Mah (1985) and the modified iterative measurement test gross-error detection algorithm presented by Serth and Heenan (1986). Other statistical approaches have also been used, such as the generalized likelihood ratio (Narasimhan & Mah, 1987), the maximum power test (Crowe, 1992), and the principal component test (Tong & Crowe, 1995). The method proposed here exploits the statistics of the historical measured process data. The analysis of each measured variable is compared to its historical values. If a change is detected, then the current systematic error is estimated, and the biased measurement value is corrected. Data reconciliation is then repeated with the new corrected value of the measurement.

The second part of the paper summarizes the basic formulation of the data reconciliation problem to provide a mathematical basis for introducing the simulation-driven data reconciliation procedure which is covered in the third part of the paper. The fourth section is dedicated to a case study that uses pseudo-steady-state data sets extracted from a real thermo-mechanical pulping process to demonstrate a practical way to bypass the lack of redundancy in measured variables. The insufficient number of redundant variables and the constant presence of systematic errors are everyday challenges for process engineers and cost accountants in the papermaking industry. The paper concludes with a discussion and final thoughts in the fifth and final section.

2. Background: General Problem Formulation

Measured data quality affects not only the quality of high-level tasks such as optimization and cost accounting, but also the quality of any estimated process model. Therefore, reliable and consistent measurement data play an important role in process plants. Random and gross errors can result in poor-quality measured data. Data processing and data reconciliation can be beneficial in minimizing measurement errors. The general form of data reconciliation is the minimization of measurement errors subject to the constraints of the physical process. Random errors are minimized by the use of data processing techniques and refined further in a reconciliation step. However, systematic errors must be identified and estimated before reconciliation. For a steady-state application, inconsistencies between instrument values and a steady-state process model are also caused by process dynamics. In this case, data reconciliation helps to distribute the errors caused by the steady-state assumption systematically onto the whole set of variables while still satisfying the underlying process model. For purposes of illustration, Figure 1 presents a simple data reconciliation problem around a single unit, a splitter.

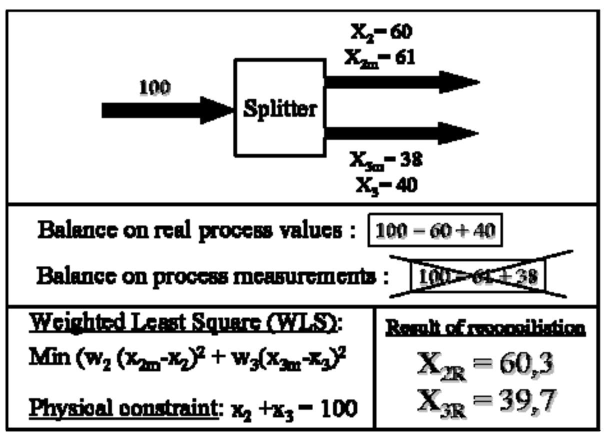


Figure 1: Example of the data reconciliation concept around a splitter.

A simplified mathematical formulation of the data reconciliation approach may be written as the weighted least-squares minimization problem of the difference between measured/unmeasured and reconciled values of variables with regard to instrument and process characteristics:

$$Min\sum_{i=1}^{K}\left(\frac{(measured/unmeasured)_i - estimated_i}{weight_i}\right)^2 \qquad (1)$$

Subject to: mass, energy, component balances

With an assumption of normally distributed random errors with no systematic errors present, this constrained minimization procedure was first introduced by Kuehn and Davidson (1961). It is important that analytical or hardware redundancy[9] of measured variables be present if data reconciliation is to be performed. The character of the problem represented by Equation 1 depends on the formulation of the constraints, e.g., linear, nonlinear, or dynamic. Furthermore, data reconciliation not only validates and estimates measured values, but also provides estimates for variables that are not being measured (often referred to as a coaptation process).

Many methods in the literature provide simplifications and solutions of the problem stated in Equation 1 by eliminating the unmeasured process variables from the problem statement. In linear data reconciliation, Crowe (1983) used a projection matrix method to decouple the unmeasured variables from the constraints. Other methods have also been used, such as a Gauss-Jordan elimination procedure (Madron, 1992) and QR decomposition (Sanchez and Ramagnoli, 1996). For nonlinear data reconciliation, the procedure is based on successive linearization of the constraints, and the resulting simplified problem is then solved using Crowe's (1983) projection matrix (Liebman 1988, Veverka, 1997). Crowe (1986) extended his previous technique to nonlinear (bilinear) processes by a two-step projection matrix technique which significantly reduced the computational effort for bilinear systems. Many other authors have addressed the computational challenges of data reconciliation for particular cases. However, there is a lack of practical applications to the pulp and paper industry, which is to some extent due to the low system redundancy of the papermaking operation. Monitoring of a sufficient number of variables to ensure redundancy is limited by installation and maintenance costs (Jacob, 2003), inaccurate measurement techniques, and the current unavailability of instrumentation. Hence, reconciling process data in the pulp and paper mills becomes a challenging and often impossible task.

In large-scale applications such as plant-wide optimization or cost modelling, the need for reliable[10] and consistent[11] plant-wide data sets is critical. To reconcile process data plant-wide, it is common practice to use a data set that consists of averaged variables over a specific, fairly

---

[9] It is said that a measured variable has hardware redundancy if two instruments are used to measure its value. On the other hand, analytical (software or spatial) redundancy of a variable is ensured when its value can be estimated in two independent ways, e.g., by a measurement and by a value from a process model.

[10] "Reliable process data" refers here to estimated variable values that are close to the true values of the process variables.

[11] "Consistent process data" refers here to data sets that are consistent with the underlying fundamental process model.

stationary time period. However, the discrepancies in the measurements are due not only to random errors (assumed to be normally distributed), but to process dynamics as well. Clearly, this type of error does not follow the assumption of a normal distribution. Bagajewicz (2000) has shown that in systems with no significant hold-ups, this error can be neglected. However, a papermaking operation is a collection of unsteady-state manufacturing processes with many hold-ups and recirculation loops; hence, significant error may be created if averaged process data are used with no systematic approach.

A different approach to plant-wide data validation is dynamic data reconciliation, which has received significant attention within the last decade, although large-scale industrial applications still remain to be developed. Clearly, the inhibiting factor is the high computational requirements of these procedures. However, with today's advances in information technology, and considering that processes are actually never at steady state, it may be better to consider using dynamic data reconciliation techniques even for near-steady-state processes (Narasimhan, 2000). On the other hand, from a practical perspective, for on-line applications, it would be wise to extend steady-state reconciliation to deal with dynamic situations (Benqlilou, 2000). Process optimization and process control would undoubtedly benefit from dynamic data reconciliation; however, for cost modelling, steady-state, not dynamic, data sets are required. In fact, as mentioned by Bagajewicz (2000), for the time being, there are more pressing problems to resolve, for example gross error detection, which is closely related with the problem of data reconciliation.

Hundreds of publications exist devoted to gross error handling, and many methodologies and techniques have been proposed. However, our capabilities to detect and correct gross errors are still limited. As for current commercial software available for gross error handling, the main technique used today is the serial elimination strategy. As mentioned by Bagajewicz (2000), vendors need to improve their strategies, for example by implementing methodologies to handle uncertainty and to enhance processing of gross errors in an industrial context.

Weighting matrix estimation

The success of data reconciliation methods generally depends on the hypothesis that the errors are normally distributed, and hence on the quality of estimation of the variance-covariance matrix. This symmetric and positive-definite matrix quantifies the uncertainties in each instrument value (Benqlilou, 2004). If the process is truly at steady state, the covariance matrix can be estimated by the direct method (Bagajewicz, 2000), which is simply a sum of standard deviations within the time of the true SS. The mean value can be calculated as:

$$\bar{y}_i = \frac{1}{n}\sum_{k=1}^{n} y_{i,k} \qquad (2)$$

and the covariance matrix can be estimated as:

$$cov(y_i, y_j) = \frac{1}{n-1}\sum_{k=1}^{n}(y_{i,k} - \bar{y}_i)(y_{j,k}\ \ \bar{y}_j) \qquad (3)$$

where *n* is the duration of true steady state and *y* is the measurement set. This direct estimation of the covariance matrix helps to correct instrument values on an optimal statistical basis. Because it is known that the process is never at true steady state, the process of estimating the variance-covariance matrix becomes more complex (Gedeon, 1984; Crowe, 1996; Chen, 1997).

A highly simplified and common practice in industrial applications is to use engineering judgment for matrix estimation by allocating uncertainty weights to each instrument. This pragmatic approach, which is also used in the current study, takes into account knowledge of the process dynamics around each particular instrumentation sub-network as well as information about each instrument's accuracy, precision, and reliability.

Gross error handling

Because data reconciliation is limited to the elimination of random errors, systematic errors must be eliminated *a priori*. Several methods are available to do this, ranging from pure statistics through neural networks to time-series screening. The efficiency and practical usefulness of the statistical approaches seem to be superior to the others. In the present work, historical knowledge about the potential locations and relative sizes of biases is used. This approach can be situated in the framework of measurement adjustment using statistical methods such as the measurement test (Mah, 1982 or Crowe, 1983). In this type of method, the data are first reconciled, and then each measurement point is tested for possible bias. If gross errors are present, then their values are estimated by solving a simple nonlinear problem (McBrayer, 1995):

$$\begin{aligned}
&\text{Min } B(y,b) \\
&\quad \text{s.t.} \quad f(y) = 0 \qquad (4) \\
&\qquad\qquad y_{min} < y_i < y_{max} \\
&\qquad\qquad b_{min} < b_i < b_{max} \\
&\qquad\qquad B(y,b) = (-)_{i-1} \ldots + [(y_i - (y_{mi} - b_i))/s_i\,]^2 + \ldots (-)_{i+1}
\end{aligned}$$

Once the values of the gross errors are known, the biased measurement value is corrected, and the data reconciliation procedure is run again. The process is repeated until convergence to a minimum value of squared error is achieved.

Industrial applications of data reconciliation

Many commercial software packages for process analysis and simulation today provide integrated functionality for data reconciliation. All these applications are for linear steady-state data reconciliation. Bagajewicz and Rollins (2002) discussed the functionality of one academic and eight commercial packages and concluded that most of them deal with material and component balances. Only a few provide the advanced possibility to connect directly to DCS systems for on-line applications. Generally, all the packages were developed for the petrochemical industry, with embedded features such as phase equilibrium constraints or model libraries for some petrochemical process units. The most popular packages in the industry are Sigmafine (OsiSoft), Datacon (Invensys), and Vali (Belsim).

Applications of any commercial software in the papermaking industry are scarce because of the dynamic nature of the process and its lack of measurement redundancy. The Sigmafine package has been used for off-line data conditioning (Jiang 2003a) and has been assessed for possible on-line application in a recausticizing plant at a kraft paper mill. The application was limited to material balances because the package cannot accommodate nonlinear constraints such as energy balances. Another package used in papermaking is the CADSIM® Plus simulation software from Aurel Systems. It has been applied to on-line energy accounting for the steam utility system in a kraft paper mill (Wasik 2007). This practical approach uses a parallel process simulator module to perform a process data validation procedure, which is referred to as *simulation-driven data rectification*. A further commercial implementation of this software involves a dynamic application to track pulp stock from batch digesters (Rankin 2009).

3. Simulation-driven data rectification

The heart of the model-based process data rectification approach proposed here is an iterative algorithm which alternates between process simulation and optimization modules. The discussion begins by presenting the overall methodology, followed by more in-depth discussions for each methodological step.

Overall methodology for plant-wide process data rectification

A sensitivity analysis of the process dynamics is performed to identify the key state variables that represent the process state identification. The whole set of measured raw process data is then extracted from the distributed control system before analog filtering. Then a data processing technique based on a wavelet transform and analysis is used to clean random noise and abnormalities from the data. Simultaneously, the processing technique analyzes the time-frequency domain for a potential steady-state occurrence of each selected variable. When a steady state has been identified, an automatic check for a potential multivariable pseudo-steady state is performed. In this way, a plant-wide process steady state can be systematically detected and used as an input to the simulation-driven data rectification technique (Korbel et al., 2011).

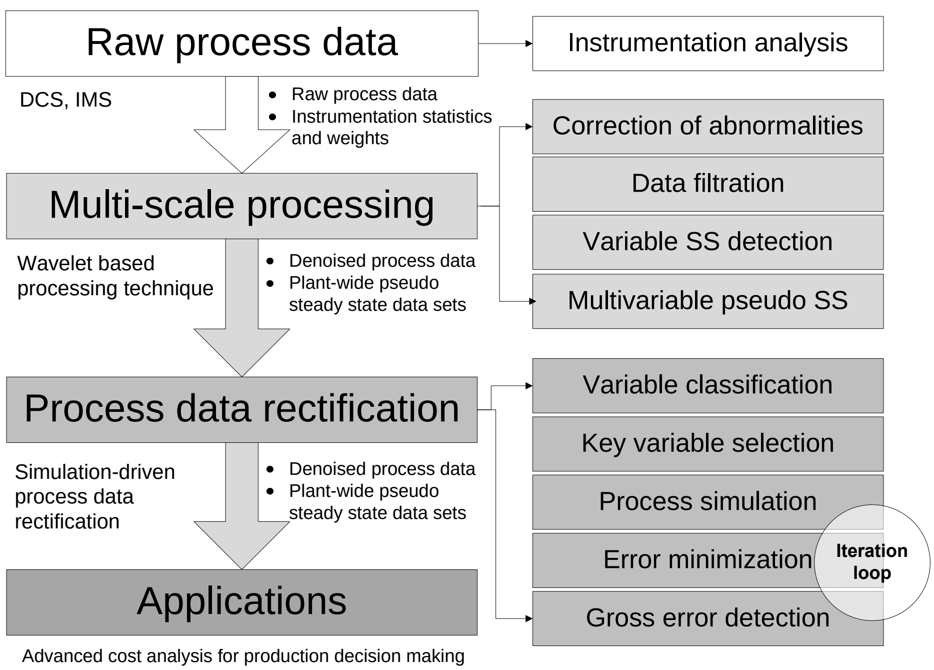


Figure 2: Overall methodology for plant-wide simulation-driven data rectification.

The data rectification procedure starts with the classification of process variables into redundant, observable, and undeterminable. The key variables are selected from the signal processing step based on the process state identification. A sufficient number of key variables are required to make a zero-degree-of-freedom process simulation feasible. The outcomes from the simulator are then compared to the inputs of the measurement matrix, and a least-squares error is calculated.

Iterative use of the optimization module leads to a solution with a minimal least-squares error value. In the iterative loop, a gross error check is performed on each instrument. If an instrument value is identified as biased, then the simple nonlinear problem (Equation 4) discussed in the second section is solved. The new estimated variable value is selected as a new input for data rectification, and the whole iterative process is repeated until a sufficiently small least-squares error is achieved.

Instrumentation performance analysis

This is the first step, to be performed before data processing and rectification techniques can be applied. Instrument documentation covering calibration, accuracy, and precision is first reviewed. Then a variance-covariance matrix of the measurement errors is estimated according to the discussion in the second section of this paper. A well-informed allocation of weights to each instrument is of critical importance to performing data rectification correctly.

Multi-scale data processing

Various parameters must be defined to use a wavelet multi-scale data processing technique correctly. Based on historical data and the dynamics of the operation, a processing scale length is defined for each variable. Four critical steps must be carefully addressed:

- Select independent and comprehensive variables for determining near-steady states;
- Establish near-steady-state criteria for each variable and for the system as a whole;
- Determine the minimum length of steady-state periods according to the process system delays;
- Determine criteria for threshold values for steady-state periods (for each variable and for the system as a whole).

This step requires profound engineering analysis of the underlying process dynamics to represent the true process trend correctly by the wavelet transformation (for more details, see Korbel et al. (a))

Process simulation

A process model is constructed by a classical flow-sheet definition using standard building blocks which describe fundamental operations such as mixing or separating, but also papermaking-specific units such as chip refiners and paper machines. The measured and simulated variables are reconciled by minimizing a weighted least-squares error while satisfying the model equations and other user-defined modelling constraints. The minimum is obtained by a simplex search algorithm. The solver is sequential for steady-state or dynamic and for linear or nonlinear systems, calculating each module in term with its output/input streams for the next module.

Data rectification with the help of the parallel optimization module

The simulation solver is sequential, but the data rectification solver is simultaneous. First, the plant operators must identify (by sensitivity analysis) the number of independent measured variables, referred to as free variables that represent the state of the operation. The free variables are input to the process simulation model (Figure 3) to estimate the whole set of process variables. The pillar of the data rectification module is the modified version of the simplex optimization technique. In the first iteration, the algorithm compares the changes in the free simulated variables to their measured values. The simulation and iteration process repeats until the minimum least-squares error between the simulated variables and the measured values is obtained. The output of the rectification process is the set of simulated variables, including rectified measured values and other calculated variables not available from measurements.

The mathematical description of the minimization problem is identical to that of a classical data reconciliation procedure, with the difference that the constraint of the minimization problem is not the underlying process model, but rather is user-defined:

$$min_{x_i} \sum_{i=1}^{n} w_i \left(\frac{y_i - x_i}{span_{x_i}} \right)^2$$

$$\text{s.t.} \quad \mathbf{f}(\mathbf{x}, \mathbf{z}) = 0$$

$$\mathbf{g}(\mathbf{x}, \mathbf{z}) \geq 0$$

where:

x_i = reconciled value;

y_i = measurement – free variables (FV);

w_i = weight;

z_i = non-measurement variables - computed variables (CV)

$span_{xi}$ = normal operating span for variable x_i ;

The vector x is subject to constraint equations, i.e. mass and energy conservation laws and specified inequality constraints. The objective function is added to the simulation to be reconciled using mathematical functions native to the simulator used. The iterative search for values of x is performed using optimization module based on a simulated annealing version of the well known simplex optimization algorithm. Furthermore, normalized values are used to calculate the objective function due to the variety of physical units met in pulp and paper operation, for instance volumetric flows which can be in the thousands and mass fractions which are between 0 and 1.

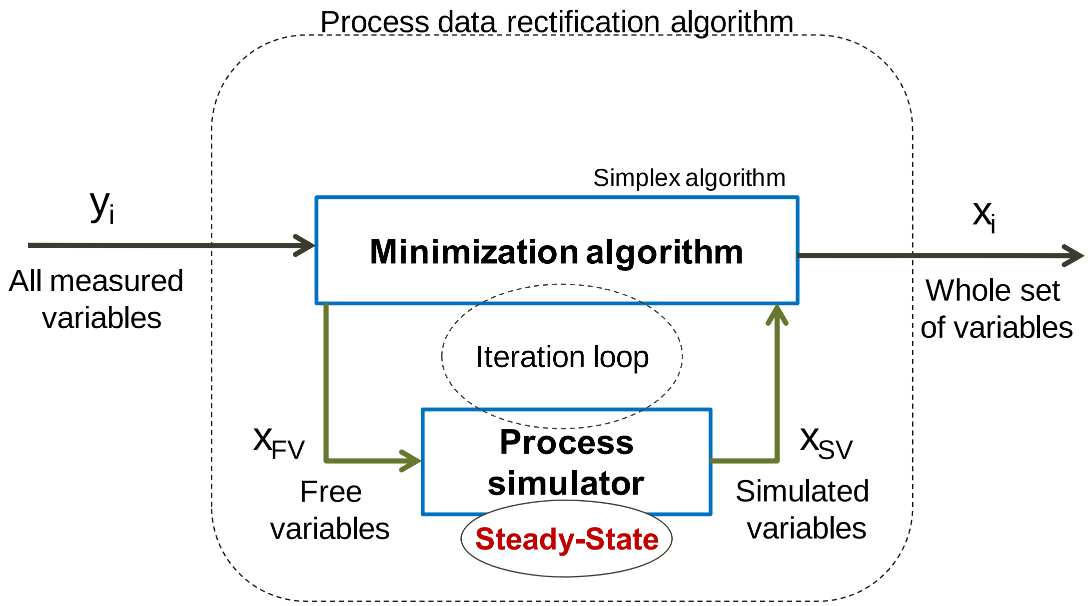


Figure 3: Iterative algorithm in simulation-driven process data rectification exploits the downhill simplex method to search for minim least square error.

4. Case study: Thermo-mechanical pulping process

Base case definition

The base case is an existing large-scale problem: an integrated thermo-mechanical newsprint mill situated in eastern Canada. Process engineers have asked the authors to provide plant-wide consistent data sets, to identify miscalibrated instruments, and if possible to estimate the values of their biases. In most newsprint mills, the thermo-mechanical pulping process (Figure 5) is a low-redundancy system, and therefore model-driven data validation using reconciliation techniques is a challenging task. The use of process simulation with an optimization module helps to overcome this lack of redundancy, but creates a high sensitivity to the quality of the process model.

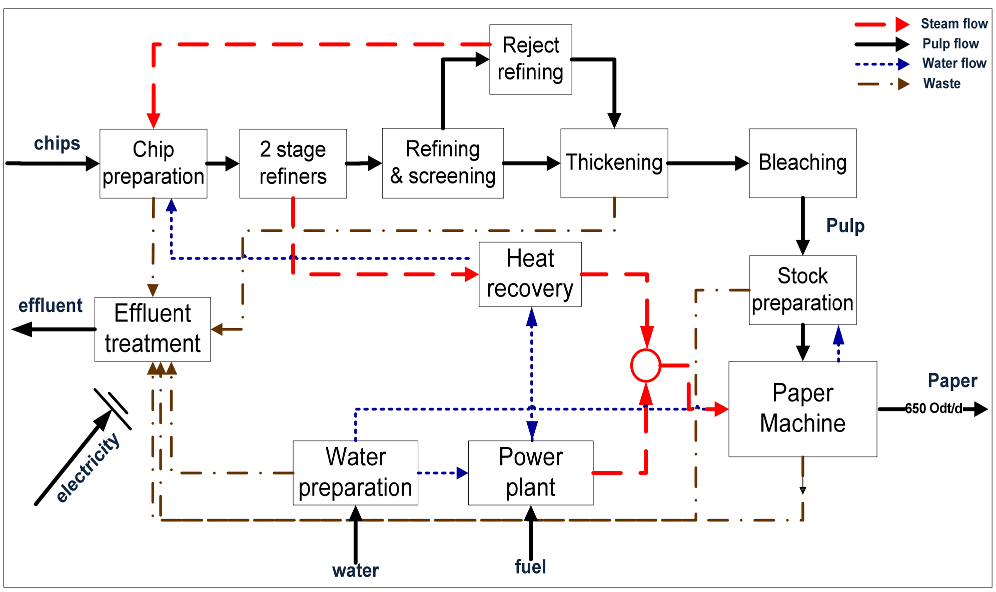


Figure 4: Block diagram of a complex thermo-mechanical pulping process with paper mill.

Case study objectives

The purpose of the case study was to demonstrate that the proposed practical methodology is capable of producing reliable and consistent plant-wide process data sets for higher-level applications such as cost modelling. Figure 5 presents the methodological steps that were followed to address the needs of mill personnel and simultaneously to answer the following research questions:

- Is it possible to apply simulation-driven data rectification to low-redundancy systems?
    - How does improved system redundancy influence the error?
- How much does this method improve data quality, with and without data pre-processing?

- Is the process data quality obtained from simulation-driven data rectification equivalent to that obtained from classical data reconciliation in the high-redundancy sections of the mill?

Selected measures

To assess the problem systematically, three measures were used to compare the quality outcomes of both methods (Benqlilou, 2006):

$$\text{Relative Error Reduction}: \; RER = \frac{\sum_i (MRE_i - RRE_i)}{\sum_i MRE_i} \times 100\%$$

$$\text{Measurement Relative Error}: \; MRE_i = \frac{|x_i - x_i^m|}{x_i}$$

$$\text{Reconciled Relative Error}: \; RRE_i = \frac{|x_i - x_i^r|}{x_i}$$

Classification of process variables

The classification of process variables in the current instrumentation network is presented in Table 2 below, followed by a description of a potential retrofit of the network to achieve, first, full system observability and second, full system redundancy.

296 variables – input/output flowrates and state variables	
<u>Existing sensor network:</u>	
	59 redundant variables (29 sensors)
	80 estimable variables (22 sensors)
	157 not estimable variables
<u>Retrofit sensor network:</u>	
A: Complete system observability	
	52 new sensors added
	92 redundant (74 sensors)
	122 estimable (30 sensors)
B: Complete system redundancy	
	98 new sensors added
	296 redundant (149 sensors)

Table 2: Variable classification and instrumentation network description.

Case study: methodological steps

The extracted raw process data were de-noised using a wavelet-based technique and simultaneously analyzed for potential pseudo-steady state. When the assumption of pseudo-steady state is met, the very same data set is then processed further using simulation-driven data rectification and classical data reconciliation techniques separately. A couple of operating divisions have a superior level of system redundancy, and both validation techniques were applied to them to assess the third research question. The system to solve is a linear problem to which both methods are applied. As mentioned, the raw process data were extracted from the information management system using an on-line steady-state detection technique based on wavelet transformation (Korbel et al., (a)) The quality of plant-wide multivariable and pseudo-steady-state data sets was discussed in this work. In the current case study, process data sets with the lowest measurement trend error that permits pseudo-steady-state detection are used.

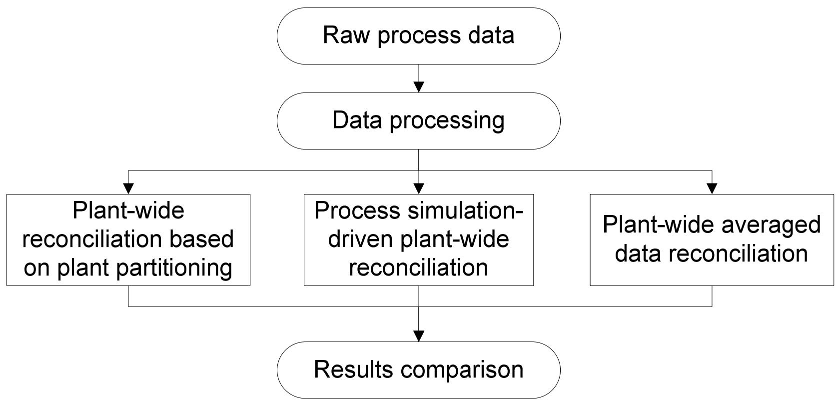


Figure 5: Three different pathways analyzed in the data validation problem.

To address the first question, a hypothetical steady state of the process operation was created. Historical information on the instrumentation network (accuracy, precision), engineering knowledge of the manufacturing operation, and an actual process simulation were used to create a hypothetical data set. After defining the problem description and measures, the methodology can be summed up in the four following steps:

- Apply classical data reconciliation (DR) and simulation-driven data rectification (SDDR) techniques to the hypothetical steady-state data set;
- Analyze the outcomes from both methods using the relative error reduction (RER) measure;
- Gradually increase the system redundancy[12], repeating steps 1 and 2;
- Analyze and address the redundancy and the error function.

[12] A systematic way to increase redundancy was selected: the instrumentation upgrade algorithm described in (Bagajewicz, 2000), with the objective function being the number of instruments.

Results and discussion

The information management system integrated into the thermo-mechanical pulping operation was used to identify a steady-state operating condition. Several candidates for steady state were extracted. The quality of each steady state and its distance from an idealized stationary trend were addressed using a multivariable measure (Korbel et al., 2011): the MTE (measurement trend error). Figure 6 represents one of the key variables (production rate) that were selected from a sensitivity analysis to identify the pseudo-steady state of the process. The mean squared error (MSE) of this key measured variable relative to the true process trend was calculated as 43. The highlighted region corresponds to the multivariable pseudo-steady-state operation that was selected for analysis. With this knowledge, a hypothetical steady-state process operation data set of all variables was created to identify the true process trend as a reference value for comparing methods.

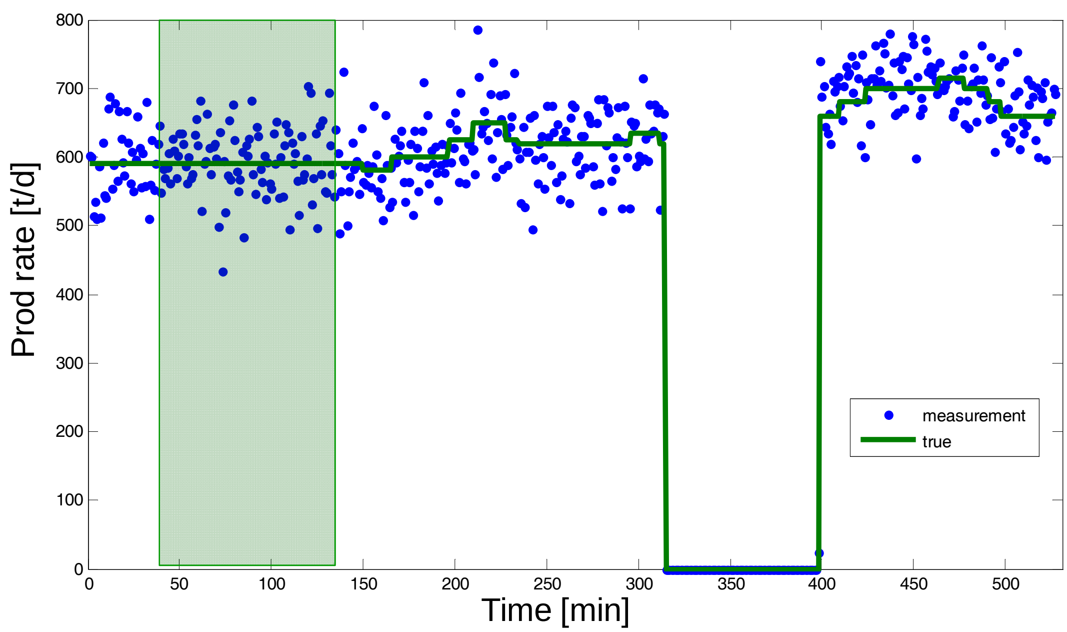


Figure 6: Example of raw measurements for production rate with a mean squared error of 43 and the hypothetical true trend.

First, the current process operations system with the actual instrumentation layout was subjected to data validation using both the classical and the simulation-driven rectification techniques to validate the possibility of using the proposed methodology. Figure 7 shows that the relative error reduction (RER of pulp volumetric flow) was very similar when each of the methods was applied to several process subsystems with some level of redundancy (e.g., Main Refining and Main

Screening, Rejects Screening). In those sections of the mill where not enough instrumentation is available, classical data reconciliation will at best result in a data coaptation process (input = output model). However, because to perform the simulation-driven approach, it is sufficient to maintain zero degrees of freedom, the relative error reduction is significantly better (close to 62% error reduction in the chip pre-treatment section of the operation) than classical reconciliation could provide (no error reduction). It is important to note that because the iterative process of error minimization is taking place between simulated and measured variables, process model quality is of critical importance. The current case study does not analyze the impact of model inaccuracy, and the model is assumed to correspond to the underlying process.

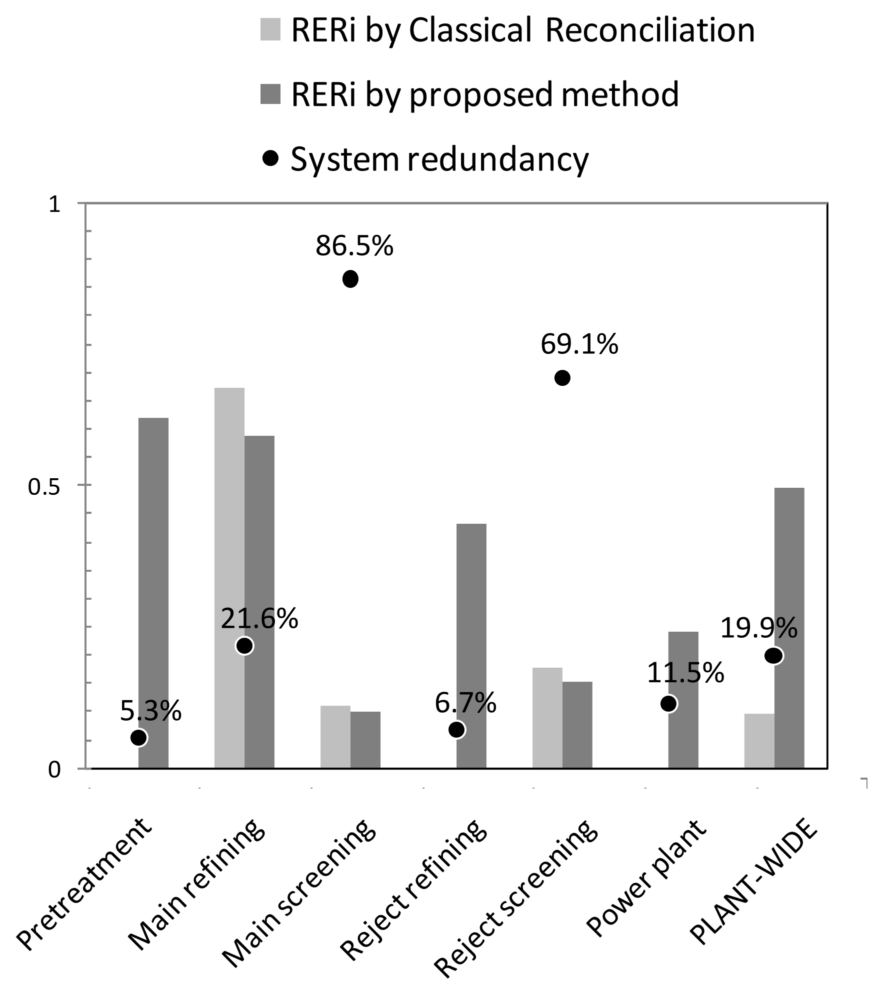


Figure 7: Relative error reduction (RER) of production rate in different sections of the process operation.

In plant-wide data validation, classical data reconciliation had reduced the error in production rate exiting the thermo-mechanical pulping process by approximately 10%. Comparing this to the simulation-driven approach that reduces error by approximately 50%, it is apparent that the estimators produced by the proposed methodology are superior if the process model is correct. Clearly, the difference in the outcomes of the two methods is due to the fact that the simulation module includes more degrees of freedom because optimization is taking place outside the module. Furthermore, the optimization module enables more practical equality and inequality constraints to be implemented. These constraints are drawn from engineering judgment related to a particular processing unit and are in many cases impossible to express, implement, or solve using the optimization formulation of classical data reconciliation.

The second part of the case study seeks to identify the relation between level of process redundancy and relative error reduction using either classical data reconciliation or the proposed method. The reference system redundancy value is that of the current operating case (19.9%). From the actual process data, a hypothetical steady-state data set representing the thermo-mechanical pulping process was created for each level of increase in system redundancy. Because new instruments were added at each step as the redundancy was increased, the values of their accuracy and precision were set to vendor-supplied values so that a new hypothetical data set could be created. This process was continued until full observability and then full redundancy of the system was achieved. At every step, the relative error reduction of each method was calculated to analyze the impact of system redundancy on the performance of each method used.

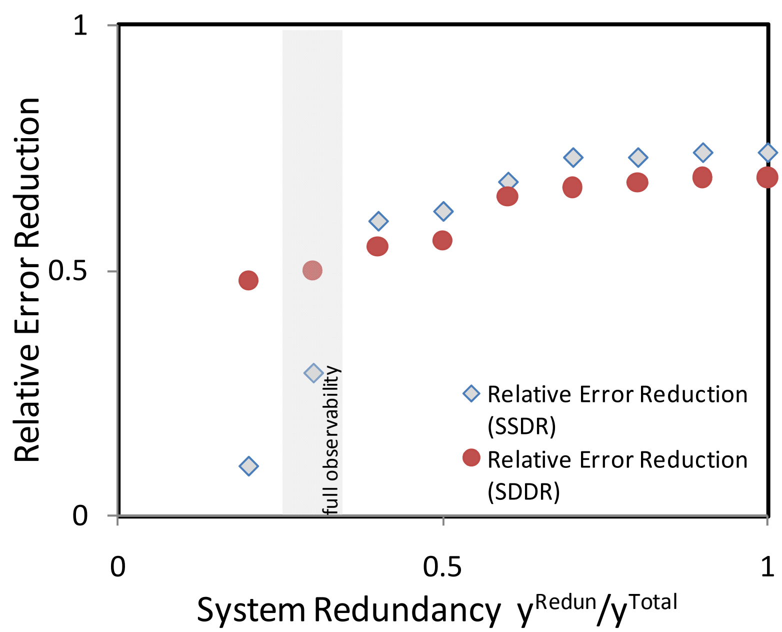


Figure 8: Relative error dependence on overall system redundancy.

Figure 8 presents an assessment of redundancy impact on relative error reduction for a single variable, pulp volumetric flow rate (production rate). A similar function can be plotted for any process variable of interest, thus creating a better understanding of the relationship between the increase of process data set reliability and instrumentation upgrades. In the case of production volume, it is shown that the low-redundancy cases are favourable to the proposed methodology. With the implementation of new instruments, classical data reconciliation achieves an approximately 70% improvement in error reduction (from 10% to close to 80%). Looking at the application of the proposed methodology, the maximum possible improvement is 20% (from approximately 50% to approximately 70%) with increasing redundancy values. Clearly, this conclusion illustrates that model quality is of critical importance. However, in reality the assumption of a perfect model never holds, and hence some errors will be introduced because of model inaccuracy. The relative error reduction converges to approximately the same values at higher redundancy. The grey stripe in Figure 8 represents the region where full observability of the system has been achieved (SR=0.286). Even though the system is fully observable and approximately 30% of variables are redundant, the classical SSDR reduces relative error by only RER = 0.292, compared to SDDR that reduces it by RER=0.501.

5. Conclusions and final thoughts

The proposed data rectification methodology uses a multi-scaled pseudo-steady-state process data set to perform an iterative search for the minimal least-squares error between simulated and measured variables. A detailed description has been provided, and the differences between this approach and the classical notion of data reconciliation have been discussed. The main differences were identified to be in the constraints of the optimization problem. Classical data reconciliation is subject to the underlying process model, whereas simulation-driven data rectification is subject to user-defined practical constraints. Data consistency is ensured by the use of a process model in the simulation module, where model quality becomes even more critical.

The case study has shown that the method is very practical and that its robustness coupled with wavelet data pre-processing is comparable to that of classical data reconciliation when high system redundancy is present. Only the proposed methodology can validate process data for low-redundancy systems where at least zero degrees of freedom are present. This is particularly useful in the papermaking industry where low-redundancy systems are common because of poor instrumentation accuracy and functionality. In these low-redundancy regions, classical SSDR cannot be fully used because redundancy is a requirement for its application.

The proposed technique may be computationally very expensive compared to classical methods. This is due to the optimization technique that iteratively searches for improved variable profiles among the simulation outcomes and compares them to the profile of the measured data set.

Finally, gross error handling is a very practical approach that uses information about each instrument from historical data sets. After reconciliation has been performed, if the squared error is very high, the most commonly occurring biased measures are checked for potential systematic error. The bias is estimated and corrected using the algorithm described here, and data reconciliation is repeated.

It is recognized here that if the simulation model is inaccurate, the outcomes from data rectification may be misleading. If the reconciled data are used in this form for higher-level analysis such as optimization or cost analysis, error propagation will be significant. Hence, the analyst using the technique must recognize this drawback and systematically analyze the results and the validity of the model.

References

Arora, N. and L. T. Biegler, “Redescending estimators for data reconciliation and parameter estimation,” Computers and Chemical Engineering, 25, p. 1585 (2001)

Crowe, C. M., “The Maximum-Power Test for Gross Errors in the Original Constraints in Data Reconciliation,” Can. J. Chem. Eng., 70, 1030 (1992).

Bagajewicz, M.J., Sanchez, M.C., Design and upgrade of nonredundant and redundant linear sensor networks, AIChE J. 45(9), 1999, 1927–1938.

Bagajewicz, M.J., Jiang, Q., Comparison of steady state and integral dynamic data reconciliation, Comput. Chem. Eng. 24(11), 2000, 2367–2383.

Bagajewicz, M.J. Process plant instrumentation: design and upgrade. Lancaster, PA: Technomic Publishing Company, 2000.

Bagajewicz M.J., D.K. Rollins, Data Reconciliation, in Instrument Engineers' Handbook (3rd Edition), Volume 3 : Process Software and Digital Networks, B.G. Liptak editor, CRC Press (2002)

Bellec, S., Jiang, T.W., Kerr, B., Diamond, M., Stuart P.R., 2007, " On-line Processing and Steady-State Reconciliation of Pulp and Paper Mill Process Data", Pulp & Paper Canada, 108:6, p. 36-40

Benqlilou, C., Data Reconciliation as a Framework for Chemical Process Optimization and Control, PhD thesis, Universitat Politecnica de Catalunya, 2004.

Cao, S., Rhinehart, R.R., An efficient method for on-line identification of steady state, Journal of Process Control 5(6), 1995, 363–374.

Chen, J., Bandoni, A. and Romagnoli, J.A. (1997). Robust estimation of measurement error variance/covariance from process sampling data. Computers and Chemical Engineering, 21, 593-600.

Crowe, C.M., Campos, Y.A.G., Hrymak, A., Reconciliation of process flow rates by matrix projection. Part I: linear case, AIChE J. 29(6), 1983, 881–888.

Crowe, C.M. (1996). Data reconciliation - progress and challenges. Journal of Process Control, 6 (2/3), 89-98.

Jacob, J., Paris, J., Data sampling and reconciliation—application to pulp and paper mills. Part II: Case studies, Appita J. 56(2), 2003b, 116–120.

Janssen, M., Laflamme-Mayer, M., Stuart, P., Survey of data management systems used in the pulp and paper industry, Proceedings, Foundations of Computer-Aided Process Operations (FOCAPO 2003), Coral Springs FL, 2003, 551–554.

Jiang, T., Chen, B., He, X., Industrial application of wavelet transform to the on-line prediction of side draw qualities of crude unit. Comput. Chem. Eng. 24, 2000, 507–512.

Jiang, T., Chen, B., He, X., Stuart, P.R., Application of a steady-state detection method based on wavelet transform, Comput. Chem. Eng. 27, 2003a, 569–578.

Jiang, T., Chen, B., Jasmin, K., Stuart, P.R. (2003b), Strategy for improving data quality for a kraft pulp mill recausticizing plant, Proceedings, Foundations of Computer-Aided Process Operations (FOCAPO 2003), Coral Springs FL , 2003b, 467–470.

Johnson, L. P. M., Kramer, M. A., 1995. Maximum Likelihood Data Rectification: Steady-State Systems. AIChE J. 41, 2415.

Korbel M., Bellec S., Jiang T., Suart R P., (a), Steady-state identification for on-line data reconciliation based on wavelet transform and filtering, submitted to C&CE 2011

Kuehn D.R., Davidson H., Computer Control II. Mathematics of Control, Chem. Eng. Process 57: 44-47, 1961.

Liebman M.J., Edgar T.F., Lasdon L.S., Efficient Data Reconciliation and Estimation for Dynamic Processes Using Nonlinear Programming Techniques, Computers Chem. Eng. 16: 963-986, 1992.

Liebman M. J., and Edgar T. F., Data reconciliation for nonlinear process, Proceedings of the AIChE Annual Meeting, Washington, DC, 1998.

Liebman and Edgar, 1988 M. Liebman and T. Edgar, Data Reconciliation for Nonlinear Processes, 1988 AIChE Annual MeetingWashington DC (1988)

Madron, F., and V. Veverka, "Optimal Selection of Measuring Pointsin Complex Plants by Linear Models," AIChE J., 38(2), 227 (1992).

Mah, R.H.S., Tamhane, A.C., Detection of gross errors in process data, AIChE J. 28(5), 1982, 828–830.

McBrayer, K.F., Edgar, T.F., Bias detection and estimation in dynamic data reconciliation, J. Process Control 5(4), 1995, 285–289.

Narasimhan and Mah, S. Narasimhan and R.S.H. Mah, Generalized likelihood ratio method for gross error identification. AIChE Jl, 33, 1987, p. 1514.

Narasimhan, S.; Jordache, C., Data Reconciliation and Gross Error Detection, an intelligent use of process data; Gulf Publishing Company: Houston, TX, 2000.

Tamhane and Mah, 1985 A.C. Tamhane and R.S.H. Mah, Data reconciliation and gross error detection in chemical process networks. Computers chem. Engng, 17 (1985), p. 409

Tjoa, I.B., Biegler, L.T., Simultaneous strategies for data reconciliation and gross error detection of nonlinear systems, Comput. Chem. Eng. 15(10), 1991, 679–690.

Tong, H., “Studies in Data Reconciliation and the Detection of Gross Errors,” PhD Thesis McMaster University, Hamilton, Ont., Canada (1995).

Sánchez M. and J. Romagnoli. Use of Orthogonal Transformations n Classification/Data Reconciliation. Comp. Chem. Engng, 20, 483- 493, (1996)

Serth and Heenan, 1986 R.W. Serth and W.A. Heenan, Gross error detection and data reconciliation in steam metering systems. AIChE Jl, 32 (1986), p. 733.

Veverka, V.V. and Madron F. (1997). Material and energy balancing in the process industries: from microscopic balances to large plants. Computer-aided chemical engineering series, vol. 7, Amsterdam, The Netherlands, Elsevier.

APPENDIX C –

ASSESSMENT AND INTERPRETATION OF ADVANCED COST DATA FOR PROCESS IMPROVEMENT AT AN INTEGRATED NEWSPRINT MILL

Submitted to International Journal of Production Economics

Assessment and interpretation of advanced cost data for process improvement at an integrated newsprint mill

Milan Korbel and Paul Stuart[13]

NSERC Design Engineering Chair in Process Integration, Department of Chemical Engineering, École Polytechnique de Montréal

P.O. Box 6079, Succ. Centre-Ville, Montreal QC Canada H3C 3A7

Abstract

The implementation of information management systems in pulp and paper companies has enabled a better understanding of both business and production processes. Even though mill engineers and accountants have incorporated the use of real-time data into their daily practices, they are often limited to using these data only for *ad-hoc* problem solving. The critical information captured in the data has not yet been made visible to decision-makers. Data trends are studied, but information is seldom extracted from the actual measured variables. It is argued that if information management systems at the mill can be exploited to their full potential, decision-making activities will be enhanced significantly by access to new and insightful manufacturing information for operational, tactical or strategic decision-making.

This paper briefly presents the structure of an operations-driven cost modeling approach that is then applied to a case study to assess the true profit margins of different production runs and operating regimes for a company's short-term benefit. The supporting pillar of this approach is the use of cleansed and reconciled real-time measured data from the operation. The flow of information from these data sets is integrated to cost data and follows the principles of ABC-like cost accounting.

The results from a case study – existing integrated newsprint mill – shows that the presented operations-driven approach dramatically increases the granularity and transparency of

---

[13] Telephone +1 514 340 4711 extension 4384; fax +1 514 340 5150

Email address: paul.stuart@polymtl.ca

manufacturing costs compared to conventional accounting techniques, while permitting comprehensive analysis of critical production and design parameters. For instance, understanding the true contribution margins of every product in the context of different manufacturing set-ups and recipes will improve planning and scheduling, enhance continuous improvement, and open the door for future margin-centric supply-chain management.

INTRODUCTION

Without a doubt, pulp and paper products are mature and standardized goods, many of which are beyond their product lifetimes. Every day, paper manufacturers are being challenged by diminishing market conditions. Various business strategies are being investigated to stay competitive in these tough market conditions. Managing the strategic level by choosing the number of facilities, their locations, and their capacities is currently helping pulp and paper companies to raise their cost curves temporarily. Managing the tactical level at the facility itself will not only tighten control of manufacturing costs at each mill, but will also provide critical information for long-term strategic planning and decision-making. However, to sustain a successful business in North America, these commodity products must first be manufactured at the lowest possible mill-level cost. Today, tactical or operating decisions are based mainly on information derived from mill benchmarking, home-grown cost accounting systems, or both. It is rare that this information is based on actual process data from information management systems.

Several surveys have been carried out to understand the growth of information technology and management systems and their use in the pulp and paper industry. A review by Fadum (1996) points out that the use of process information systems in the industry is limited, often only for troubleshooting, and that these systems should be exploited more to enhance production profitability and product quality. Shaw's (1999) survey points out the increased availability and wider use of data acquired from the process; however, no advanced use of these data was reported by the users surveyed (engineers, IT personnel, and mill managers). Yeager (2000) has shown that today's real-time data availability promotes the development of decision-making tools which could enable mill personnel to react promptly to changes. A more recent detailed survey (Janssen et al., 2003, 2004) concluded that the interpretation of data from data management systems is increasing, but that the link between different types of data is generally missing. Clearly, gaps between the functionality of these systems and the needs of papermakers are preventing their application. These gaps are mostly due to the particular nature of pulp and paper manufacturing processes, to inaccurate and biased measurements, and to users' desire to maintain their accustomed systems.

One of the major opportunities for exploiting information extracted from the data available from information management systems (IMS) in pulp and paper companies lies in the field of manufacturing cost accounting, where operating-cost-related efficiency improvements remain to be achieved. To grasp effectively the operating knowledge that resides within the cost accounting data, the underlying process characteristics must be integrated with the cost data. Because traditional cost accounting practices use top-down cost allocation per volume of production using weekly or monthly statements, some other approach needs to be used to account for process operations. Activity-based costing (ABC) and its process-based characteristics can be used for production cost modeling in the continuous manufacturing industries (Laflamme-Mayer et al., 2011). ABC is an activity-driven costing system that was first developed to overcome indirect-cost allocation difficulties when using traditional costing (Kaplan, 1989) and has since been successfully applied in a wide range of industries. After a decade of overcoming difficulties and complications with the costing framework (Turney, 2008), the ABC philosophy has been adopted by several researchers and practitioners for particular case-driven or company-specific purposes (Steen and Steensland, 1994). These descendants of Kaplan's original work are today referred to as "Sons of ABC" or "ABC-like" cost accounting methodologies. Many companies, mostly in the discrete manufacturing industry, are saving millions of dollars due to well-informed decisions that are based on results from ABC and its granular view of resource consumption. In the continuous manufacturing industries, however, only a few implementations have been done, with the majority occurring within oil and petrochemical companies. These home-grown enterprise-specific practices are often kept confidential as company know-how and are not available to the public.

In recent years, a mix of ABC and standard cost-accounting frameworks and systems has been developed for the forestry sector (Fogelholm, 2000). In particular, the development of an operations-driven cost modeling framework (an ABC-like approach) has been shown to produce critical manufacturing cost information (Laflamme-Mayer et al., 2011). The use of lower-level process data together with financial data in a "bottom-up" cost accounting concept has yielded a better understanding of complex continuous production environments such as those found in pulp and paper mills. Operations-driven cost analysis enables efficient analysis of complex cost relationships through an understanding of the efficiency of resource usage by activities and how activities are linked to final cost objects. Applications of this operations-driven framework have been demonstrated in several case studies aimed at production improvements relative to supply-chain management and at improving the state of knowledge related to retrofit design decision-making activities (Laflamme-Mayer et al., 2012; Janssen et al., 2011). In these high-level applications, the relatively long time scale used for cost modeling (weeks to years) is adequate. Further decreasing the time scale (to hours) for production cost assessment enables the tracking of actual and true product margins (Korbel et al., 2012), which is the focus of this paper.

Literature review of cost-accounting systems

Cost accounting is the supporting pillar of the accounting framework that provides critical financial information to managers for decision-making. The cost insights provided are used only internally to enable managers to find the optimal way to maximize the company's profits. On the other hand, financial accounting is used for reporting that produces information available to the public. Many companies use several different cost-accounting systems for problem-solving, cost variance analysis, and financial reporting. Even though traditional cost-accounting practices have been developed to focus primarily on financial reporting, they are also being used for internal company cost-performance analyses (primarily in the continuous manufacturing industries). However, traditional costing systems are often unable to determine accurately the actual costs of production and of related services. This situation most likely exists because the concepts of traditional costing were developed early in the last century to satisfy past industry needs for systematic and reliable cost information (Cooper, 2000). In the labour-driven industries of the last century, standard costing methods could provide satisfactory cost-control strategies. The main focus was simply to manage and control efficiency and productivity. In this context, the standard level of resource consumption is evaluated by looking at the company's past performance. Then the standard utilization of resources is compared to the actual usage to track and evaluate cost variances. Clearly, the outcomes of such a cost-control analysis are the variance values, which are presented as gains or losses due to productivity, volume, or cost variations from the budgeted or expected target (Horngren, 2006).

Because the commodity nature of pulp and paper manufacturing and its past needs correspond to the period of development of traditional costing, standard costing was adopted by the industry for cost control and reporting. This costing strategy enabled companies to develop standard manufacturing recipes for setting up the optimal process conditions for each production line and product, which is why this approach is still used today. For target performance evaluation, the actual costs that were incurred during the time period under analysis are compared to the expected standard recourse consumption within the given product recipe. However this simplified representation of the production environment is not sufficient to provide good information on how costs are generated within the mill (Laflamme-Mayer, 2011). Clearly, the simplicity of traditional cost systems is achieved with the loss of some functionality. Furthermore, the cost distortion that occurs due to arbitrary allocation of indirect costs became significant as industries were transformed more and more into automated systems, magnifying the ratio of indirect to direct costs. At this time, ABC concepts were developed to enable better overhead allocation. However, the pulp and paper sector did not adopt these advances in cost control and did not modernize its cost management tools.

Activity-based cost accounting

ABC accounting is a relatively new approach to cost control that was developed in the 1980s in response to cost discrepancies resulting from inaccurate overhead allocation (Kaplan, 1988). By simply adding an activity as a link between resource consumption and a cost object, the knowledge of costs incurred in the organization is improved significantly. The activity becomes a fundamental cost object whose value can be directly traced to cost objects such as services, products, or customers. Different opinions about the accuracy of these methods exist and no absolute measure of the validity of product costs based on any costing approach is available. Cooper and Kaplan (1988) describe ABC as a more correct means for product costing in today's industry setting (large companies) where expenses covering marketing, distribution and support are a significantly increased proportion of the total costs compared to traditional direct labor and material costs.

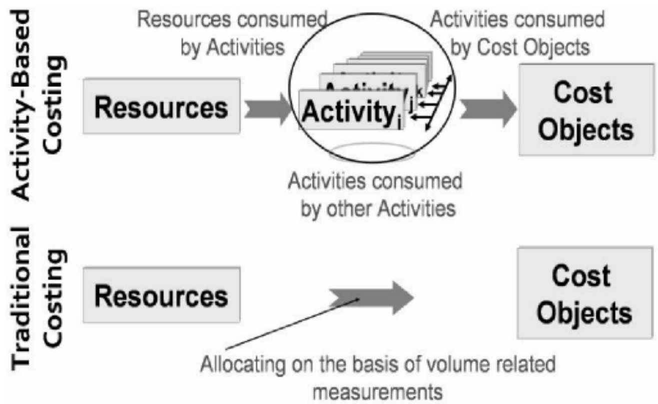


Figure 1: Activity-based costing and traditional costing.

The process-oriented character of ABC (Figure 1) shows cost generation in two logical stages, while volume- or structure-oriented traditional costs are generated in one. This fundamental principle accounts for the difference in visibility of cost data when ABC principles are used (Drucker, 2008); the traditional approach cannot encompass the critical link between actual causes and associated costs. Furthermore, advanced ABC has recently evolved into multistage systems in which individual activities can be consumed by other activities before being consumed by final cost objects, thus improving the accuracy of cost modeling even more (Emblemsvåg and Bras, 2001).

The use of the original ABC approach in the discrete-parts industry is advanced compared to the situation in continuous process environments (Horngren, 2006; Kaplan, 2004; Steen and Steensland, 1994). There is little discussion in the literature of applications of advanced costing strategies for continuous process operations (Steen and Steensland, 1994; Yeager, 1999; Fogelholm, 2000; Janssen, 2011; Laflamme-Mayer, 2011). Fogelholm (2000) discussed the difficulties of assessing product costs in the forest industry. However, standard costing principles are still being used there. The application described by Fogelholm seeks to anticipate the resource requirements for upcoming customer orders based on the dimensions, raw material content, and quantities of the product (Fogelholm, 2004). Steen and Steensland (1994) present the potential benefits from implementing advanced cost-accounting systems based on ABC principles for real-time product cost calculation. They argue that the necessary requirements for such successful costing systems must incorporate bottom-up cost and process data integration, measured process data reconciliation, system flexibility, and common data-base structures. However, the details of the proposed framework were not presented, which would clearly be necessary if the proposed costing system were to be used. Nevertheless, discussion of the needs of the continuous manufacturing industry for advanced costing systems is of critical importance. For example, an explicit relationship between operating conditions and resource consumption, which is missing in current systems, is essential for analyzing direct manufacturing costs in continuous industries. This situation arose mostly because the main focus in the late 1990s was on cost applications to address the indirect-cost allocation dilemma, and no attention was given to tracking direct manufacturing costs in these industries. Especially in the P&P industry, resolving this issue is essential because all direct costs are arbitrarily traced to the product based on standard measures, which provide no explanation of how and why these costs have actually been incurred. Therefore, a holistic costing approach is needed that would link financial information with process operating data to provide an in-depth sophisticated view of resource consumption at the process and unit levels.

Operations-driven cost modeling

Laflamme-Mayer et al. (2011) developed an operations-driven cost-modeling framework that exploits the link between financial data and the knowledge that can be captured from the vast amount of data collected by an IMS. They argue that the particular production needs of a given operation must be understood to develop cost-accounting tools adapted to a particular production environment. The effective use of cost-management systems must involve both the process operations and the product perspectives, which is the main difference from discrete-parts manufacturing. Two distinct perspectives must be assumed in a continuous manufacturing environment:

- The process perspective on costs incurred is necessary for the mainline manufacturing operations to explain actual cost generation correctly.
- The product perspective is necessary for the finishing operations (e.g., packaging) where a large amount of product diversity is created.

The combination of both process and product perspectives is required for many cost explanations, for instance, an in-depth process perspective for transition-cost calculation for a batch-level production facility. In this framework, the process perspective is described by two distinct factors: the given *production process* (design) and the *operating conditions*.

- The production process corresponds to the activity within a part of a manufacturing process. It is described by the configuration of the process, the type of processes involved (reaction, heating), and other relevant information.
- The second factor is of critical importance because the rate of resource consumption alters with changes in operating conditions. Traditional costing does not include tracing of these changes and assumes product homogeneity (Kaplan, 1988), and the corresponding costs are calculated based on recipes (Laflamme-Mayer, 2011).

The strength of this costing framework has been demonstrated on case studies which have provided information on retrofit design decision-making (Janssen et al., 2011). Another case study involving supply-chain management shows that the information provided by operations-driven grade-cost assessment breaks up the time period under analysis into segments corresponding to campaign runs. From the results, it is clear that the cost of manufacturing the same product varies significantly from one campaign to another, a fact which can be exploited in planning and scheduling for margin-centric supply-chain management (Laflamme-Mayer et al., 2012).

Operations-driven cost modeling for true product margins assessment

Campaign costs can be broken down to assess the running and transient production cost of different operating regimes and to capture the true product margins of a grade. As mentioned earlier, current mill IMS can accommodate such analysis if the measured process data are accurate. However, this is often not the case, which makes direct use of the data inappropriate or in some cases impossible. Many tools have been developed to help address this task in the processing industries (Bagajewicz, 2000); however, pulp and paper mills are a special case because of their manufacturing nature and instrumentation availability (Korbel et al. (b)). For these purposes, only plant-wide process data can be used under pseudo-steady-state conditions. To achieve a certain level of pseudo-steady-state data quality, advanced processing techniques must be used, which is the supporting pillar of this costing approach. In a nutshell, a wavelet-based multiscale method is used to identify the pseudo-steady-state of the operation. The elimination of random noise and abnormalities is carried out simultaneously. The second step is data reconciliation, which improves further the accuracy and plant-wide completeness of process data. For further information, refer to Jiang (2003) and Korbel et al. (b).

The data dissection and cost-modelling vision used in this approach can be understood from Figure 2. Traditional cost-accounting systems, as discussed earlier, permit some *ad-hoc* profitability analysis of different products. To go further and to comprehend the true manufacturing costs from an engineering perspective, process knowledge must be integrated to assess the operating costs of each manufacturing regime. At this level, the information can be used by decision-makers to avoid costly operating regimes and to try to keep the process inside the most profitable operating region.

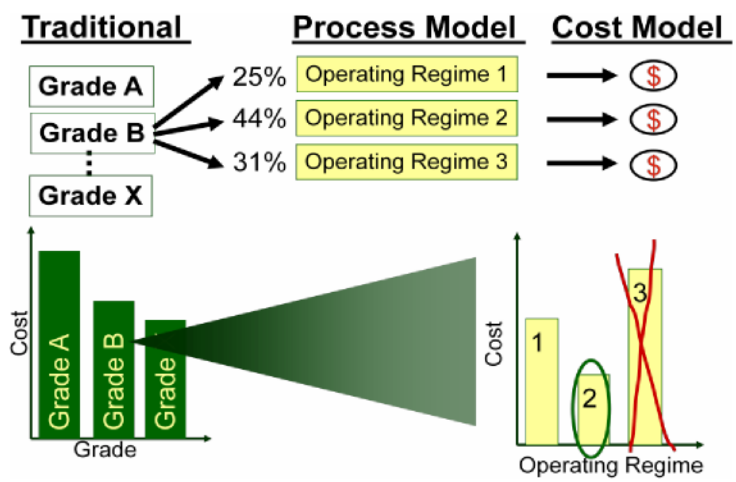


Figure 2: Smart data dissection for operations-driven cost modeling approach (Korbel and Stuart, 2012).

OBJECTIVE

The overall objective of this study was to develop a methodology which would enable the calculation of true product margins from a process perspective with process interpretation capabilities. The necessary data processing tools that are a prerequisite for cost analysis have been presented and discussed in the literature (Korbel et al. (a) and (b); Korbel and Stuart, 2012). The main goal of this paper is to demonstrate the value of this approach through an application to a case study at an integrated newsprint mill. The results are presented in comparison to outcomes from traditional standard costing to evaluate the methodology. The specific objectives addressed in this paper are:

1. to assess and characterize the manufacturing costs of a newsprint mill using this approach,
2. to demonstrate the inclusion of reconciled measured process data in smart data dissection for operations-driven cost analysis of manufacturing regimes to identify and interpret cost variances based on operational knowledge, and
3. to discuss the inclusion of the proposed method in supervisory decision-making activities for short-term benefits to the company.

CASE STUDY DESCRIPTION

The mill under analysis is an integrated thermo-mechanical newsprint mill. The manufacturing process incorporates the classical thermo-mechanical process steps, including chip pre-treatment in two atmospheric vessels followed by impregnation units, three-stage chip refining, and rejects refining. The "dirty steam" from the refiners is sent to the recovery unit for recovery, and additional steam if needed (in winter periods) is produced in three available boilers (fueled by natural gas, oil, or electricity). The thermo-mechanical facility produces a pulp quality based on paper mill demand and specifications, with its throughput being matched to that of the paper mill. Two basis weights of newsprint products are produced: 48 $g.m^{-2}$ and 45 $g.m^{-2}$.

The company is highly competitive in newsprint (in the first quartile of manufacturers), and hence its manufacturing cost efficiency is a critical issue to stay competitive under North American market conditions.

OPERATIONS-DRIVEN COST ANALYSIS

The focus of cost-model development was first to characterize the direct and indirect manufacturing costs of a newsprint mill to identify the most profitable and most costly operating regimes. This information was then interpreted to define cost variances and was then used in support of decision-making activities related to process operating improvements for the short-term benefit of the company.

The five methodological steps for cost model development and analysis are (Figure 3):

1. *Cost objectives definition*
   The first important step is to clearly define the scope and the objectives of the cost analysis procedure. It is necessary to develop guidelines for identifying relevant cost items and for characterizing the desired cost behaviour for use in decision-making.

2. *System and data dissection*

   The production system is divided into smaller subsystems to enable the qualification and quantification of different types of cost drivers. This step increases the in-depth cost analysis capabilities of the system. The separate subsystems are called *Process Work Centers* (PWCs) and are further divided into individual *Processing Units* for increased cost-tracking transparency. The system division step is usually based on rules derived from process data redundancy analysis (Korbel et al. (b)). Then, individual production runs of a given grade can be dissected into operating regimes at this stage (if this was not done in the data processing phase before cost analysis), or additional production regimes can be identified and added to the cost analysis.
3. *Driver description phase*

   This step involves intensive discussion with mill personnel to identify the cost drivers. This phase is of critical importance because it structures the shape of the cost model and the characterization and interpretation of the results.
   - Resource drivers: The characterization and measurement of the resource consumption rates of processing units and process activities are based on process data. For instance, flow measurement is a resource driver for a given flow medium.
   - Process activity drivers: These drivers characterize the linkage between operating conditions and the consumption of a resource driver. This phase identifies what information is necessary to characterize the intensity of a process activity within a process work center. For instance, the pressure in a vessel will characterize the required rate of steam flow to be input.
   - Process work center drivers: The boundaries of each PWC are defined in the second phase of the methodology. The interpretation must be intuitive to capture the cost-insight capabilities of the method clearly in a graphic user interface. The aim is to explain cost generation better at a mill-wide level. One of the important cost centers is the overhead work center, where the drivers must be clearly defined to achieve indirect-cost transparency throughout all mill departments. For instance, the work center driver for maintenance is the head count for a given subsystem of the operation.
4. *Cost model development*

   The model development follows the operations-driven cost modeling framework presented in Laflamme-Mayer et al. (2011). The supporting pillar is the integration of process and financial information based on ABC-like principles. The systematic consideration and cost aggregation of individual production processes and their operating condition into the plant-wide manufacturing operation are essential principles of this stage.
5. *Characterization phase*

   The last phase involves the characterization of the costs incurred and the interpretation of the results based on the objectives defined in phase one. Process understanding is the key element at this stage. Therefore, interaction with mill personnel is necessary to interpret the results. Sometimes, steps 2-4 will need to be repeated to arrive at a satisfactory level of in-depth cost understanding. The results of this stage, when validated, can be clearly visualized and used for decision-making support.

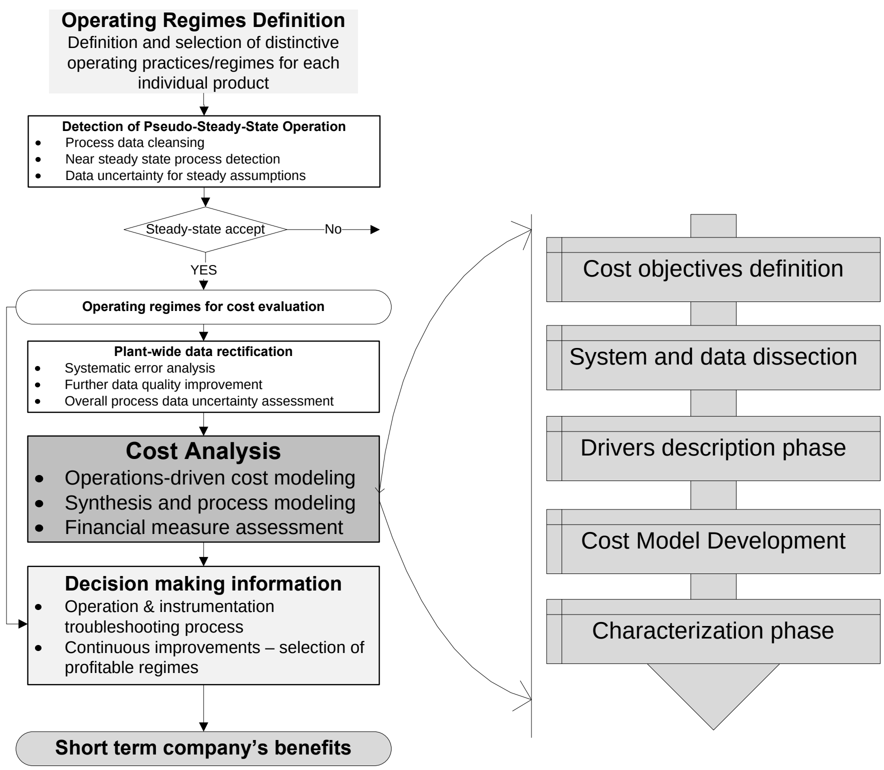


Figure 3: Overall methodological framework for true product margin assessment.

CASE STUDY METHODOLOGY

After the operations-driven cost model has been developed (including grade run campaigns and operating regime selection), the methodology used for a single manufacturing time period and a single newsprint product consists of the following four steps:

1. Production cost assessment using traditional standard costing rules
2. Production cost assessment of grade campaign runs using operations-driven cost analysis (Laflamme-Mayer, 2011)
3. Production cost assessment of operating regimes (Korbel, 2012) within individual campaign runs
4. Results interpretation from a process perspective.

RESULTS AND DISCUSSION

The advantages of operations-driven cost analysis can be presented by characterizing the mill cost of manufacturing and showing that the interpretation capabilities of the proposed method are superior to those of traditional cost-volume variance analysis. Significant inferences can be made from the results of the case-study application to achieve potential cost savings. Because the case study is a real thermo-mechanical pulping operation, the data are presented as values normalized to the standard cost estimates.

Campaign runs and operating regimes

When production campaign under analysis is run, its corresponding manufacturing costs are determined using the weighted average (based on probability of occurrence) of operating regimes and the accumulated costs during non-steady-state conditions that occurred during the run. The cost of individual operating regimes is determined based on operations-driven cost analysis. In the case of a newsprint mill, the operating regimes are characterized first by a group of seven essential process control setpoints throughout the TMP process that correspond to a given pulp freeness and pulp quality requested by the paper mill. The other parameters used to characterize regimes are the type and age of the refiner plates used in the primary, secondary, and rejects refiners. The regime definition is in the hands of the analyst and is specified at the start of the project. Because the test case mill is a simple integrated newsprint mill, the definition of regimes is limited to different operating settings.

Mill production cost characterization

The direct and indirect manufacturing costs for one of the newsprint products (48.8 $g.m^2$) are presented in Figure 4. The production costs of the grade runs and grade operating regimes are assessed by the operations-driven costing method. The results are compared to the standard costs provided by the mill accounting systems, which are based on the monthly statements within the time period analyzed. Each bar of the graph is divided by resource costs. The diamond sign within the bar corresponds to the length (in hours) of the campaign run. Clearly, the longer the campaign run, the more near-steady-states corresponding to operating regimes can be detected and analyzed, which provides more certainty about the cost information for a given production run. It is clear from Figure 4 that this approach to cost analysis provides a novel perspective on manufacturing costs compared to traditional accounting practices. These results provide further analytical granularity to the original work of Laflamme-Mayer et al. (2011) because of their extended in-depth view on resource consumptions for each different operating regime, as is clear from Figure 5.

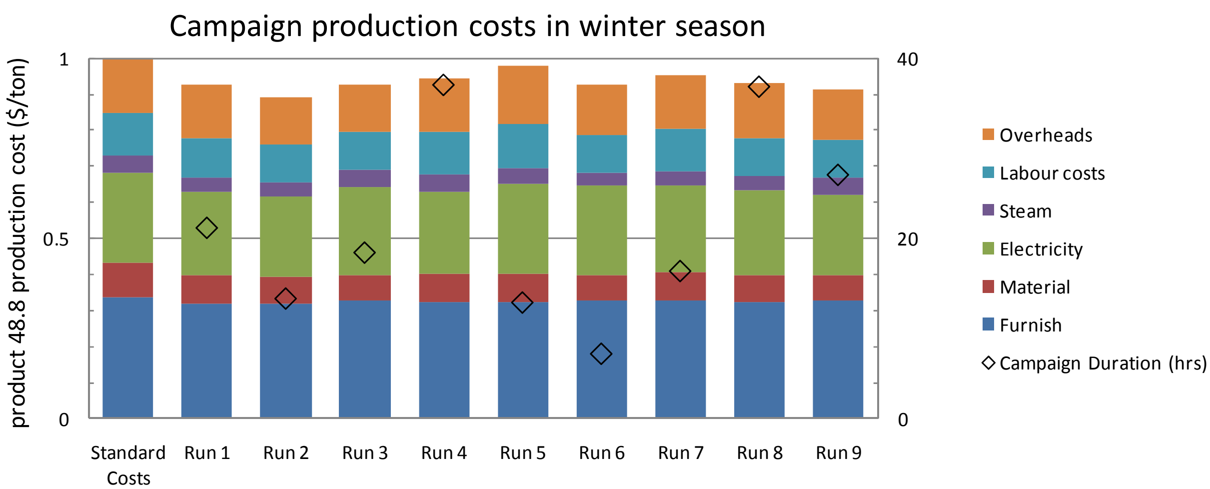


Figure 4: Manufacturing costs of a 48.8 g.m^{-2} newsprint product during the winter season time frame analyzed (production campaign runs)

The bar corresponding to standard costs (Figure 4) indicates how resources should have been consumed for the given production recipe of the given product grade. When drilling down to different operating regimes (Figure 5), the cost variance within individual campaign runs became visible and confirming that standard costing is an *ad-hoc* measure, even though for this particular time period, the variability in process running conditions are relatively small. The visibility of cost variances between different runs of manufacturing the same product, and then broken down into variances between operating regimes enhances cost characterisation of mill's production processes. As can be deduced from these two figures (4 and 5), cost analysis using operations-driven costing is more consistent with the process operation than traditional approaches. This kind of analysis can be performed in order to compare the costs from different operating runs, and analyze the impact of avoiding costly regimes within individual runs, thus providing guidance to continuous mill improvements. The difference in production costs for different operating regimes were mostly due to several factors:

- variances in specific electricity on primary, secondary and reject refiners
- the change in unit price of steam due to different factors (recovery unit efficiency, the amount of low-steam produced, the overall steam demand)

- change in chips quality (or chips ratio when new recipes for the same operating regimes were under trial periods)
- the use of chemicals (bleaching)
- overhead costs due to the change in allocation base
- maintenance costs (change in indirect cost driver due to increased/decreased maintenance hours within various campaign runs)

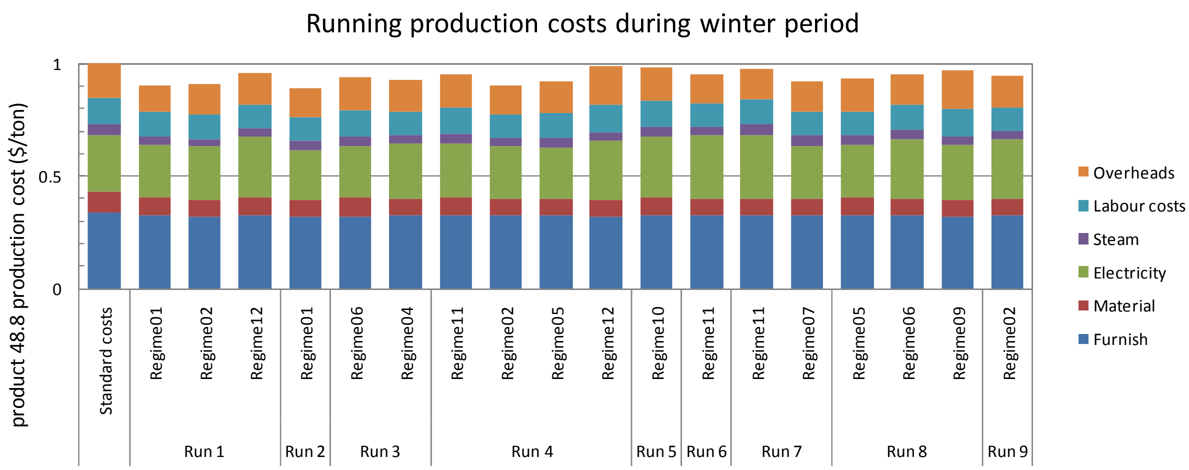


Figure 5: Manufacturing costs of a 48.8 g.m^{-2} newsprint product during the campaign-cost time frame analyzed and its corresponding operating regimes.

Figure 6 shows the results of the same production cost analysis for the summer period. Only six campaigns occurred within the time period under analysis. It is clear that the difference between standard costs and operations-driven cost values becomes even more apparent in the summer months. After closer analysis it was found that the increase in manufacturing costs for campaign run 2 and 5 was due to the increase demand of steam. In a common summer manufacturing period, recovered steam from the process is sufficient for the need of the paper mill (and the rest of the facility). However, during the campaign run 2 (and 5), additional high-pressure steam had to be produced, thus increasing unit price of steam significantly. The cause of this difference was found to be due to the unplanned maintenance of recovery unit at the energy island of the facility.

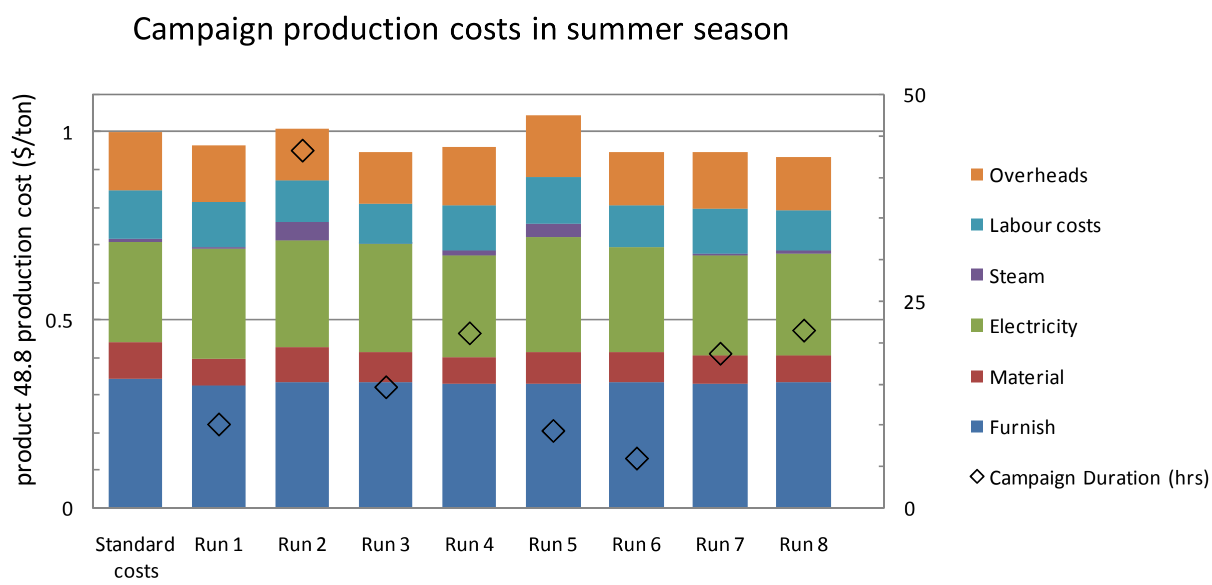


Figure 6: Manufacturing costs of a 48.8g.m^{-2} newsprint product during the summer season time frame analyzed.

The information derived from this cost analysis will have an essential impact on mill cost improvement strategies. Traditional costing does not look at these operations-driven cost drivers, which does not enable well-informed operating decisions. On the other hand, systematically looking at each individual product's true margin within each individual operating regime as well as at the complete set of operating regimes will guarantee that managers can make well-informed decisions.

<u>Mill production cost interpretation</u>

Operations-driven cost modeling not only provides a sophisticated view of manufacturing costs incurred, but also enables direct engineering interpretation of the acquired cost of manufacturing. These interpretations are based on the ability of the costing approach to drill down into a measured-data-driven characterization of the operation. As shown in Figure 7, process engineers and accountants can drill down from product-level costs through campaign and regime costs to the actual true costs of the production work centers and the corresponding rate of resource consumption by the process activities and processing units. This versatility helps to understand and interpret the cost variances that arise from changes in the rates of resource consumption due to process operations.

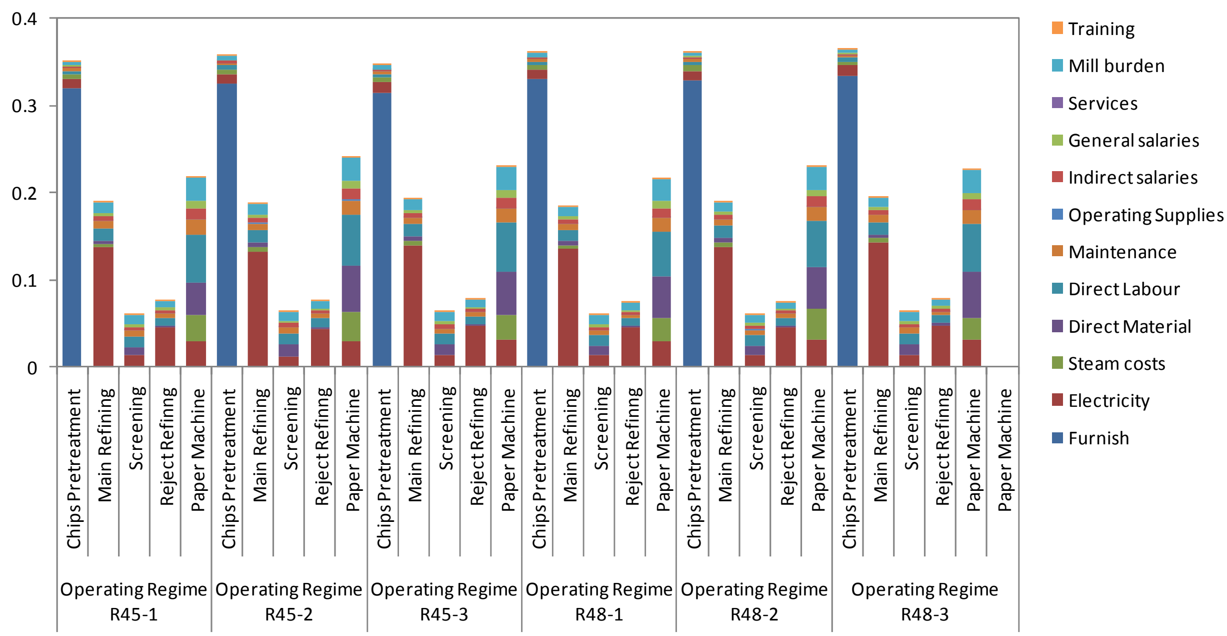


Figure 7: Production costs of different operating regimes divided into Process Work Centres and both types of costs – direct and overhead costs

Figure 7 presents the cost for each process work center during run 3 and operating regime 03 while producing the 48.8 g.m^{-2} newsprint grade. This cost information can be used and compared to the other cost results corresponding to the same operating regime. The interpretations of the cost variances can then be addressed by the process engineers and accountants at the mill. As also shown in Figure 5, the difference in manufacturing costs for the same grade is different based on operating regime chosen. Another level of analysis is an interpretation of essential causes of cost variances within the same operating regime. This information can provide well informed, process troubleshooting analysis by benchmarking the cost of individual regime:

- Manufacturing costs of the same operating regime, but within different campaign runs, can vary significantly (e.g. operating regime09 in campaign runs 1, 2 and 9)
    - This variance was identified to be caused by the increase in the use of bleaching chemicals as well as the increase in electricity consumption on primary refinery. The disturbance was caused by the change in the raw material characteristics. After further analysis it was identified that the reason was the instrumentation miss-calibration that measured slightly higher production throughput than was actually being produced (the difference in chips humidity parameter). This biased measurement caused the specific energy to be not in the optimal level, creating higher level energy consumption

  - The drop in steam unit price was caused by the increased rate of steam recovered, thus minimizing the need to produce high-pressure steam.
- Operating Regime11 cost variance between runs 6 and 7 is mainly due to unit steam price. This variance was identified to be due to excess production of high-pressure steam due to low efficiency of steam recovery unit due to mechanical cause: the microfibers present in the dirty steam coming from high consistency refiners, caused clogging of the recovery unit. This problem has not been identified for the next couple of days causing significant profit loses. One can noticed that the unit steam price fell back to its optimal values (regime09 and regime02 of runs 8 and 9).

When looking at the higher-level analysis of campaign run costs, the variance in the costs due to raw material change is not visible, making the costs of runs 1, 3, 4, and 8 look identical. When drilling down to operating regimes, clearly it is not the case that the individual process activities correspond to the trend given by the aggregated campaign level. Clearly, a distortion due to process dynamics is introduced at the level of campaign costs, making the differences in resource consumption for particular running conditions invisible.

This clear process-operation visibility provides a better understanding of the mill's cost structure. Furthermore, because process operating regimes are defined based on process conditions, the manufacturing costs of the same product can be addressed in multiple ways, making the cost distribution of a given product grade available for the first time in pulp and paper mills. Figure 8 shows the manufacturing information covering the whole set of operating regimes for the 48.8 $g.m^{-2}$ product during the winter period analyzed. Each regime is labelled by its corresponding total production cost and its probability of occurrence. The width of the bar corresponds to the cost range of the regime due to the use of multiple steady-state data sets for regime costing. The thick line inside each of the bars represents the weighted average of near-steady-state cost regimes. The colors in the red spectrum indicate costly operating regimes, whereas the bars with green colors represent more profitable regimes. The grey ones represent approximately average values. It is clear from this figure that producing the same grade under different operating regimes does create a significant variance in product margin (~$22 per ton of paper).

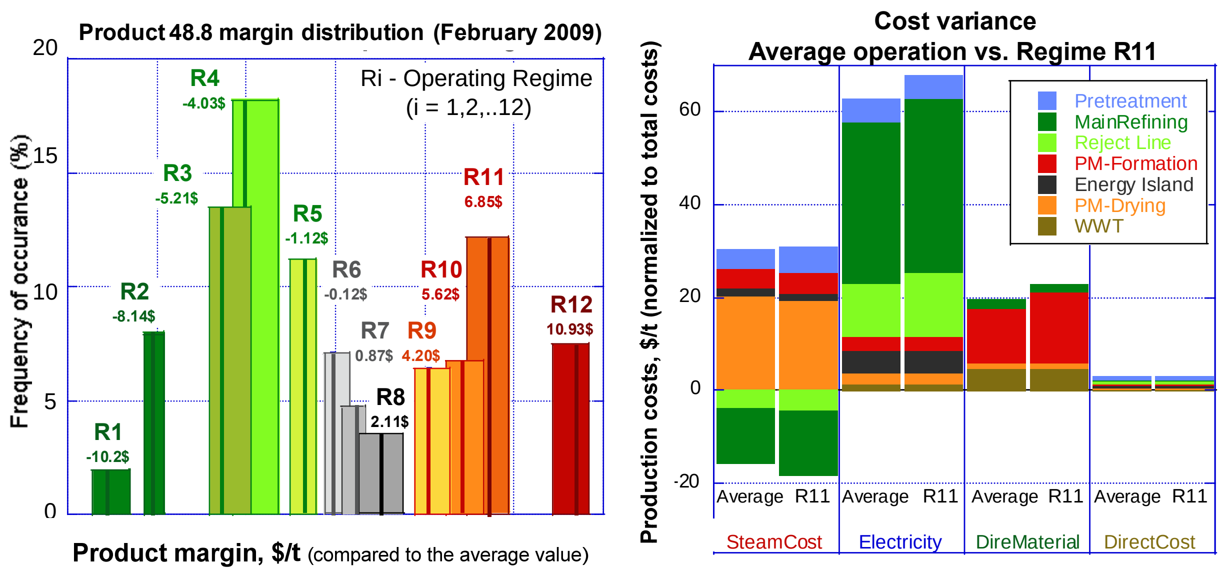


Figure 8: Wide range of product profit margins for a 48.8 g.m^{-2} product for one month of operation.

The R11 operating regime was in effect for nearly 13% of the near-steady-state operating periods (which represent only a small fraction of the overall dynamic operation). An analysis of variance of regime R11 with respect to the average value is presented in Figure 8b and indicates that the electricity costs within the primary and reject refining areas are higher than average and that the steam cost credits are also higher due to production of dirty steam from both refiners. Furthermore, a variance associated with material usage in sheet formation was also detected due to the higher use of bleaching material. After closer analysis, it was deduced that operating regime R11 corresponds to the operating situation in which the TMP mill is sequentially moving its control setpoints towards lower values of pulp freeness. This period is followed by campaigns for another grade (45 g.cm^{-2}). Higher values of pulp freeness correspond to higher electricity demand while producing more steam at each refiner.

Some costly operating regimes simply cannot be avoided due to process or raw material constraints. In particular, the costly regime R12, which is due to different raw material characteristics (chip ratio or chip size differences), could potentially be discarded if operators had this particular information. Operating conditions in regime R11, which is due to a scheduled process shift towards different pulp properties, cannot be avoided completely, but the regime duration can probably be minimized. Some other costly regimes are a reaction to a sudden need for process reconfiguration due to unknown causes or an unexpected and sudden change in raw material properties. These however, are unavoidable because process safety takes priority over profitability. Other options could possibly be explored to cope with similar situations.

In general, if the company could avoid operating regimes R9–R12, the potential cost savings could be business savers. The cost analysis based on the process regimes identified using data-processing tools provides the possibility of analyzing the cost efficiency of mill operating subsystems. These cost results, analyses, and interpretations are not possible using traditional costing methods.

CONCLUSIONS AND IMPLICATIONS

Data management systems at pulp and paper mills today are not fully exploited. Companies now have the opportunity to use these data to develop tools and methodologies to understand their production processes better. The methodology proposed here uses real-time process data that are gathered and stored by these systems. Data cleansing must take place before these raw data can be used for advanced cost analysis. After the pseudo-steady-state conditions of plant-wide operations have been identified, the data sets are ready to be used as an input into the operations-driven cost analysis framework presented here. This framework is based on activity-based costing principles which, if implemented, will help forestry managers to improve company profits. The use of lower-level measured process data from information management systems helps to improve the understanding of manufacturing cost variability due to its cost-process nature. The advantages of this approach have been demonstrated in a case study at an integrated thermo-mechanical newsprint mill. Improved understanding and visibility of the mill cost structure permitted process-based interpretation of cost variances between the summer and winter periods and helping to identify costs associated to operating changes. For instance, it was found that, the shutdown of recovery unit have a large cost impact on production costs in summer periods. The impact of the change in feedstock properties has manifested as the variance in steam, electricity and chemicals consumptions. Furthermore, it was found that, significant cost variance occurred, when operating in regime that corresponds to a different product grade. This manufacturing scenario is occurring during scheduled grade production change.

It can be concluded that pulp and paper mills should implement such systems today to improve their current cost-saving strategies. Such an approach could be used as a knowledge-based operational decision-making support tool for reducing the cost of production and helping with process troubleshooting. Furthermore, the knowledge provided will certainly help the company in its strategic decision-making activities, which will be investigated in a follow-up study.

ACKNOWLEDGEMENTS

The authors wish to thank mill personnel for their continuous support and many discussions during the project. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Design Engineering Chain at École Polytechnique de Montréal.

REFERENCES

3C Software, 2010. www.3Csoftware.com

Bagajewicz, M.J., 2001. *Process Plant Instrumentation and Upgrade*. Technomic Publishing, Lancaster, Pennsylvania.

Cooper, R., 2000. Cost management: from Frederick Taylor to the present. *Journal of Cost Management* (September/October) : 4–9.

Crama, Y., Pochet, Y., and Wera, Y., 2001. "A discussion of production planning approaches in the process industry". *Core Discussion Paper*. No 42, University of Liege.

Drucker, P. F., Collins, J., Kotler, P., Kouzes, J., Rodin, J., Rangan, V. K., et al., The Five Most Important Questions You Will Ever Ask About your Organization, p. xix (2008)

Emblemsvåg, J. and B. Bras (2001). Activity based cost and environmental management - A different approach to the ISO 14000 compliance. Norwell, MA, USA, Kluwer Academic Publishers.Fadum (1996)

Fogelholm, J., 2000. *Cost Function Modelling in the Paper Industries*. Ph.D. thesis, Helsinki University of Technology, Department of Industrial Engineering and Management.

Horngren, C.T., Foster, G., Datar, S.M., Teall, H.D., and Gowing, M. (2006). *Cost Accounting: A Managerial Emphasis*, 4th Canadian edition.

Janssen, M., Laflamme-Mayer, M., and Stuart, P.R., in FOCAPO 2003: Proceedings Foundations Of Computer Aided Process Operations. 2003. Coral Springs, FL, USA.

Janssen, M., Laflamme-Mayer, M., Zeinou, M.-H., Stuart, P., 2004. Survey indicates mills' need to exploit it systems with new business model. Pulp & Paper 78 (6), 46-51.

Janssen, M. and P. Stuart (2011). "Multi-criteria decision making for retrofit process design using process- and supply chain-level criteria." Submitted J-4 journal

Jiang, T., Chen, B., He, X., Stuart, P.R. (2003). Application of a Steady State Detection Method Based on Wavelet Transform, Computers and Chemical Engineering, Vol. 27, pp 569-578.

Kaplan, R.S. and Cooper, R., 1998. *Cost and Effect: Using Integrated Cost Systems to Drive Profitability and Performance*. Boston: Harvard Business School Press, xi, 357 pp.

Kaplan, R.S.R. and Anderson, S.R., 2004. Time-driven activity-based costing. *Harvard Business Review* 82(11): 131.

Kaplan, R.S.R., 1988. One cost system isn't enough. *Harvard Business Review* 66(1): 61.

Korbel, M., Bellec, S., Jiang, T., and Stuart R P., (a). Steady-state identification for on-line data reconciliation based on wavelet transform and filtering. Submitted to Computers and Chemical Engineering, 2011

Korbel, M., Wasik, L., and Stuart, P.R., (b). Practical methodology for plant-wide process data rectification in the pulp and paper industry. Submitted to International Journal of Production Economics, 2011

Korbel, M. and Stuart, P.R., (2012). *Cost Integration Methodology and the Forest Biorefinery: Integrated Biorefineries: Design, Analysis, and Optimization*. Taylor and Francis.

Martin Laflamme-Mayer, Matty Janssen, Paul Stuart, 2011, "Development of an operations-driven cost model for continuous processes", Part I: Framework for design and operations decision making. Journal of Science and Technology for Forest Products and Processes (J-FOR), Vol. 1, No. 1, 2011

Steen, M. and Steensland, T., 1994. Real-time profit monitoring and activity-based cost management. *TAPPI Journal* 77(2): 105–111.

Turney, P., 2008. "Activity-Based Costing: An Emerging Foundation for Performance Management", http://www.sas.com/whitepapers/indexAZ.html.

Yeager, R., 1999. Costing software helps Sappi mills make the grade more efficiently. *Pulp and Paper* 73(12): 63–66.

APPENDIX D –

PRODUCT MARGIN ASSESSMENT FOR PROCESS COST-IMPACT ANALYSIS

Submitted to Journal of Industrial and Engineering Chemistry Research

Product margin assessment for process cost-impact analysis

Milan Korbel, Jawâd Jeaidi, Paul Stuart[14]

NSERC Design Engineering Chair in Process Integration, Department of Chemical Engineering, École Polytechnique de Montréal

P.O. Box 6079, Succ. Centre-Ville, Montreal QC Canada H3C 3A7

Abstract

Information management systems in pulp and paper mills enable the development of advanced methodologies that assist the forestry sector in better controlling costs and improving profits. These systems are of especially critical importance for commodity producers such as newsprint mills, where low production costs are essential for business survivability. This paper presents the application of operations-driven cost analysis to characterize manufacturing costs in an existing newsprint mill and to use this information to analyze the cost impact of different retrofit design scenarios. This work does not aim to evaluate capital spending opportunities, but rather seeks to examine critical operating variants after integration of a new production line. The pillar of the methodology is the ABC-like operations-driven cost accounting, for which the use of reconciled real-time process data to characterise manufacturing operation is essential. The methodology is demonstrated using a case study considering three retrofit biorefinery implementations into integrated newsprint mill. The results show that the operational profitability of new integrated production lines strongly depends on the operational differences in current manufacturing regimes of core business products. These differences in manufacturing costs can be visible from a process perspective and enables assessment of individual product margins. This information is essential for margin centric supply chain of the enterprise and for exploring process flexibility to achieve an optimal product profile according to market conditions.

Keywords

COST ACCOUNTING, OPERATIONS-DRIVEN COST MODELING, PROCESS DESIGN ANALYSIS, DECISION SUPPORT, INTEGRATED NEWSPRINT MILL

---

[14] Telephone +1 514 340 4711 extension 4384; fax +1 514 340 5150

Email address: paul.stuart@polymtl.ca

INTRODUCTION

Actual product margin assessment

Pulp and paper companies produce multiple commodity products with individual specifications to satisfy multiple customers. Understanding individual product margins becomes essential to determine the optimal unit prices, thus revealing the true profitability of production. Current cost accounting systems and practices provide only *ad-hoc* assessment of these values. The common assumption of product homogeneity creates distortion to the real costs that are incurred in the time frame under analysis (Laflamme-Mayer, 2011). Standard costing methods based on standard recipes can serve only as a mill benchmark for performance evaluation. On the other hand, actual cost calculations using traditional methods provide only aggregated costs that are assessed in a top-down manner. The division of such aggregated costs into cost pools corresponding to individual products is usually volume-based and therefore incorporates various changes in process operation due to process dynamics, raw material disturbances, or both. For this reason, this assumption is often far from reality, making the estimated costs unreliable for determining the true product profitability that is critically important for decision-making. Mill personnel (accountants and engineers) recognize that the rate at which each mill generates costs may vary significantly, even when the mills are manufacturing the very same product. Determining the true contribution margins of products is clearly a challenging task for accountants in the processing industry, because both process and cost data are biased. A survey by Ernst & Young (2003) shows that 98% of respondents claim that financial reporting is distorted, with indirect costs and overhead allocation being the main areas of biased information. Almost 40% believe that the production cost data they receive are significantly inaccurate and unreliable.

With recent advances in information management systems, it is possible today to replace the *ad-hoc* cost information currently used with the true individual product margin values (Korbel and Stuart (c)). To gain these knowledgeable insights, process and financial information must be integrated into a single system. The source of process knowledge lies within the measured data, which are inherently corrupted by different types of errors. Therefore data-cleansing techniques must be used before these data are entered into higher-level analyses (Steen, 1994; Korbel et al. (b)). The modeling of this link between cost and process information will provide granular knowledge of production processes and therefore of contribution margins by assessing the costs of various operating regimes (Laflamme-Mayer et al., 2011; Korbel and Stuart (c)).

The knowledge gained from operations-driven costing systems is essential in a complex multiproduct manufacturing environment. Because of current market conditions, the forestry sector is exploring different strategies for potential business transformation, which will shift traditional commodity thinking into a multiproduct environment. Hence, production cost visibility and understanding will become even more crucial.

Retrofit cost-impact analysis

The assessment of manufacturing cost for retrofit operations is usually driven by traditional costing practices. This information is sufficient for many evaluation objectives; however, the operational details that are often needed to evaluate carefully the impact of a retrofit on the core business are lacking.

In recent years, several studies that address the involvement of process characteristics in the cost of production have been published (Greenwood and Reeve, 1992; Steen, 1994; Janssen, 2011; Laflamme-Mayer 2011; Korbel and Stuart (c)). Greenwood and Reeve have enhanced their earlier work on ABC-like costing to take better account of the appropriate cost-activity structure, product attributes, and process-based drivers. Steen discussed the needs of an efficient process-based model. The case-study-driven applications presented by Laflamme-Mayer, Janssen, Korbel, and Hytonen, which describe various operations-driven cost analysis frameworks for improved granularity and new insights into operations with different implications for operation, design, and supply chain management, provide more discussions of ABC and ABC-like analyses.

Clearly, the understanding of individual product margins in a multiproduct environment will improve the performance of planning and scheduling tasks. Simultaneously, it will provide process flexibility which can be exploited in selecting the optimal product mix based on market conditions. This new information would create a knowledge-based manufacturing plant, opening up the possibility of margin-centric supply-chain implementation for a company's long-term benefit. If this system were implemented today in a mill operation, mill personnel would benefit from the most well-informed decision-making information available for their future planning efforts.

OBJECTIVES

This paper is the second in a series describing an operations-driven cost modeling approach to assessing product margins for short-term (Korbel and Stuart (c)) and long-term benefits. The overall objective of this work was to develop a methodology that would enable the assessment of true product margins based on real-time process data and would also provide interpretation and tracking capabilities.

This work does not aim to evaluate capital spending opportunities, but rather seeks to examine critical operating-cost variants after integration of a new production line into a core business. The main goal of this second paper is to demonstrate the application of the operations-driven cost modeling approach and to evaluate the cost impacts of a retrofit design using the structure of an ABC-like cost accounting system. The results are presented in comparison to outcomes from a traditional standard costing system to validate and visualize the values of the proposed methodology. The specific objectives addressed in this paper are:

4. to characterize the current core-business production costs of a newsprint mill using operations-driven cost analysis of various operating regimes.
5. to characterize and interpret the cost impact of integrating new production lines into the current core business.
6. to discuss the implications of this method for both operational and strategic decision-making activities for a company's long-term benefit.

EXISTING MILL AND RETROFIT DESIGN ALTERNATIVES

The base-case mill is an existing integrated newsprint mill. The thermo-mechanical facility produces different pulp qualities based on paper mill demand and specifications, with the throughput matched to that of the paper mill. Two newsprint products with different basis weights, 48 $g.m^{-2}$ and 45$g.m^{-2}$, are produced.

The case mill under analysis is a highly competitive newsprint mill (in the first quartile of manufacturers) with limited access to biomass. Hence, they have chosen to investigate a biorefinery strategy that could be integrated into their existing operations.

Existing mill configuration

The following manufacturing steps are involved in the base-case mill:

- 1 newsprint machine with a total average production of 680 tonnes/day of newsprint,
- 1 TMP line with a total average production of 680 tonnes/day of pulp.

The following supporting processes are also part of the base-case mill configuration:

- A wastewater treatment plant processing 30,000 m^3/day
- A boiler plant producing 2500 GJ/day of steam
- A steam recovery unit in the TMP line, producing 3000 GJ/day of steam.

Forest biorefinery retrofit alternatives

Three major forest biorefinery retrofit options at an integrated newsprint mill were selected for production cost analysis:

- Cellulosic ethanol production: ~3000 empirical gallons per day ethanol production from hemicelluloses extracted before pulping
- PLA production: 11.5 tons per day of polylactic acid (PLA) production from lactic acid extracted from hemicelluloses before pulping
- Biocomposite production: 80 tons per day of biocomposite pellets produced from the blending of TMP fibres and polypropylene.

The first two retrofit options are based on the sugar platform, i.e., sugars are the feedstock for production of these biochemicals. Ethanol and PLA products share the same process design up to the fermentation unit (Figures 1 and 2).

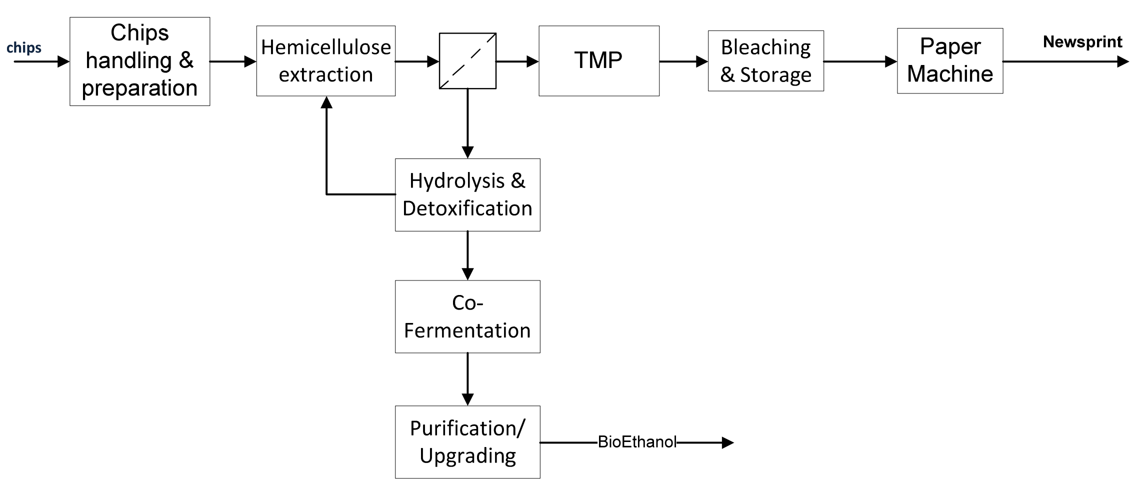


Figure 1: Simplified flowsheet of ethanol production.

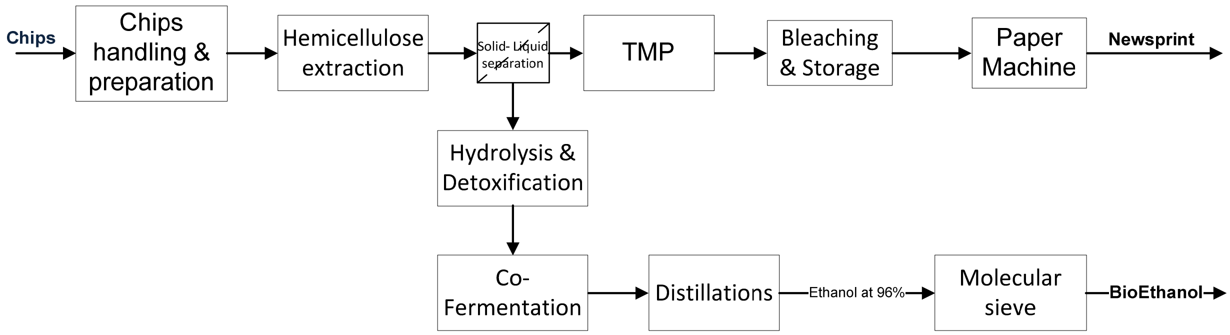


Figure 2: Simplified flowsheet of PLA production.

Fermentable sugars are extracted from the hemicelluloses which were extracted before the pulping process. This step is performed at the chip impregnation stage using the VPP (Value Prior-to-Pulping) concept (Van Heiningen et al., 2010; Huang et al., 2008). At this phase, the washed and steamed wood chips are treated with an oxalic acid solution (Houtman et al., 2011; Meyer-Pinson et al., 2004). As a result, this process is able to extract an amount of hemicellulose sugars corresponding to 3% of the incoming wood mass. The core production line must stay intact, and therefore the flow of chips supplied to the TMP line has been increased to account for the fibre losses in the new process line. The benefits of hemicellulose extraction are a decrease of approximately 25% in the specific energy in the refining (Houtman et al., 2011; Kenealy et al., 2007), because the mechanical properties of the treated chips have been altered. However, by decreasing TMP power consumption, the production of steam from the TMP recovery unit is also reduced. More clean steam must be produced in the boiler to compensate for these losses.

A hemicellulose stream is recovered after chip-liquid separation and will be prepared for the best-suited fermentation process. The sugar content of this stream should be increased through removal of water by evaporation. Polysaccharide sugars should be hydrolyzed into C_5 or C_6 monosaccharides, for example by addition of a strong acid. If fermentation inhibitors such as phenolic compounds are present, they should be removed, for example by lime addition (if lime is added, gypsum is made as a by-product and needs to be sold or disposed of).

The fermentable sugar stream is sent to the fermentation tanks, where a cocktail of the appropriate enzymes or microbes is used to convert both the C_5 and C_6 sugars into ethanol (or lactic acid depending on the end product required).

At this stage, ethanol (or lactic acid) will be concentrated and purified to meet customer requirements. For ethanol, distillation columns are used to reach the azeotropic point, and then a molecular sieve is used to purify the ethanol product to 99% purity.

For PLA production, lactic acid acts as a building block. At the exit of the fermentation process, a low-consistency lactate solution (the alkaline form of lactic acid) is produced. A poly-acid-lactic polymer is made by condensation of lactide. Several steps before polymerization are needed, including lactic-acid recovery, lactide formation, and separation of lactide isomers. This process has already been illustrated at commercial scale using Purac's technology, which uses lactic acid produced from starch. This process has a high capital cost. Isomer management before polymerization is critical for PLA quality.

The last forest biorefinery alternative is based on mechanical blending of TMP fibres with a plastic matrix to create a biocomposite material.

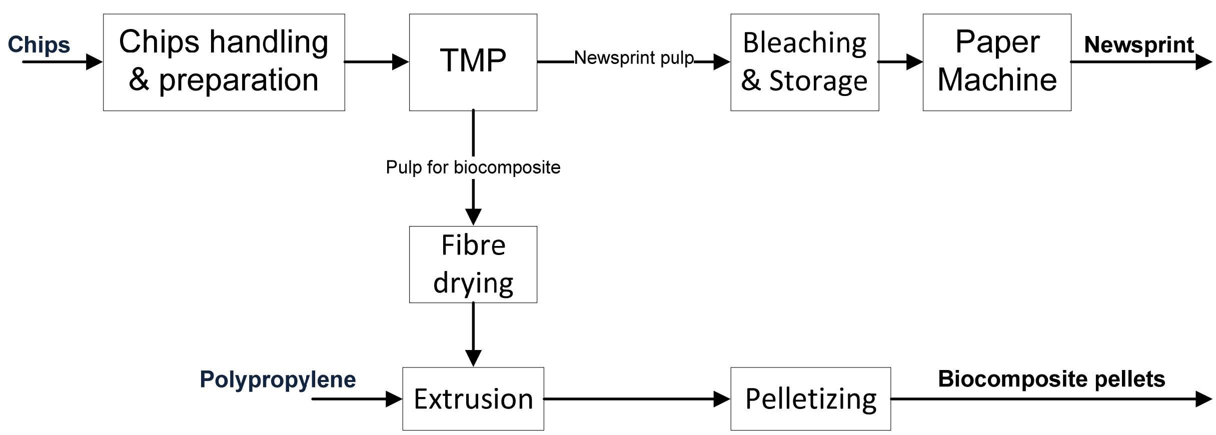


Figure 3: Simplified flow-sheet of biocomposite production.

Biocomposites are made from the blending of bio-fillers with a plastic (bio-based or otherwise). The most commonly used bio-fillers are hemp and flax fibres, wood flour, and medium-density wood fibres. The main advantage of this alternative is the sharing of the TMP asset to provide both newsprint pulp and fibres for biocomposite production. Because the base-case throughput must be maintained, increased chip input is required in the TMP refiners. Because of the increase in TMP production, steam production from the TMP plant is also increased. Pulp dedicated to biocomposite production requires drying before it can be properly blended with a polymeric matrix by extrusion. This stage of the process will consume steam. A pelletizing unit is then added before packaging.

METHODOLOGY

Using operations-driven cost modeling and plant-wide real-time data sets corresponding to near-steady-state operation, costs for each operating regime were calculated. For simplicity of illustration, only the three most common operating regimes were selected for further cost analysis. The probability of occurrence for each of these three regimes within the same set of product specifications was between 20% and 30%. The six operating regimes analyzed, which describe six manufacturing states for producing two distinct paper grades, are labelled as 48-1(2,3) and 45-1(2,3). The methodology described below was then used to perform a cost estimation of the retrofit operation (Figure 4):

5. Calculation of current core-business production costs for the three selected operating regimes was performedfor each product within the time period analyzed (March 2009).
6. A process simulation (material and energy balances) of a retrofit biorefinery option was integrated into the actual newsprint process using CADSim software. The simulation was performed for each operating regime for individual paper grades.
7. Operations-driven cost modeling and analysis using real-time and simulated process data were carried out for each retrofit scenario.
8. The outcomes and results were analyzed and presented with a clear and logical interpretation from an engineering perspective.

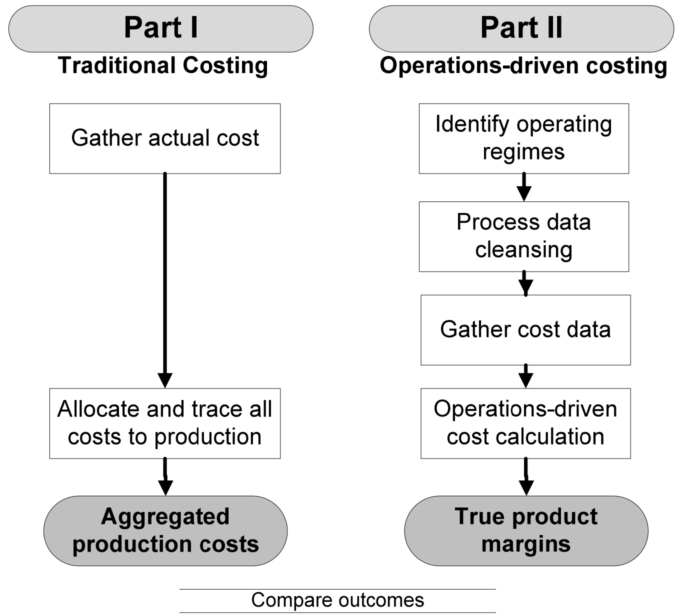


Figure 4: Overall methodology comparing traditional and operations-driven costing information (from Korbel and Stuart, 2012).

OPERATIONS-DRIVEN COST ANALYSIS

This part of the paper presents in a nutshell the necessary steps for operations-driven cost model development. A detailed in-depth discussion on model development can be found in a previous paper on operations-driven costing for short-term company benefit (Korbel and Stuart (c)).

The following five steps are required for cost model development and analysis:

6. *Cost objectives definition* - Clear objectives of the cost analysis are defined.
7. *System and data dissection* - The production system is divided into smaller subsystems called *Process Work Centers* (PWCs) in the context of an ABC-like framework.
8. *Driver description phase* – Various levels of drivers are defined:
    - Resource drivers
    - Process activity drivers
    - Process work center drivers
9. *Cost model development*
   The model development follows the operations-driven cost modeling framework presented in Korbel and Stuart (c). The basis of the method is the integration of process and cost information based on ABC-like principles
10. *Characterization phase* - The characterization, interpretation, and validation of results for decision-making are performed.

The cleansed real-time process data that represent near-steady-state operation were used as inputs for the operations-driven cost model described in (Korbel et al. (b)). Step two of cost modeling divides the mill into several activity centers, or PWCs. The mill burden was calculated in an overhead cost-center pool called the Overheads Work Centre (OWC). This division ensures the availability of detailed information about the indirect production costs for each PWC. The activity driver for the overheads center varies from one indirect cost pool to another. For instance, maintenance costs were allocated based on headcount to the PWC where the actual maintenance was carried out. Figure 4 presents the mill division into PWCs and support activity centers (OWC).

The manufacturing costs evaluated for each alternative do not account for depreciation of the new process line. This element was omitted for the sake of simplicity and because of the need to focus on understanding the impact of the new process on the core business processes using the same level of overhead cost pools.

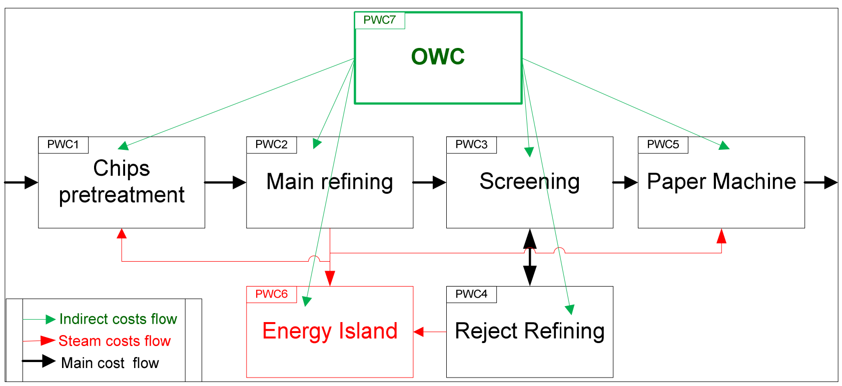


Figure 5: Mill division into PWCs and OWC.

RESULTS AND DISCUSSION

Current core-business cost analysis

This section of the study seeks to characterize and interpret the direct and indirect manufacturing costs of a base-case newsprint mill to identify the most profitable and cost-efficient operating regimes. Significant implications can be drawn from the results of the case-study application for potential cost savings for short-term benefits, which were presented in the first paper of this series (Korbel and Stuart c).

The period for analysis was chosen during the winter season (March). The process operation was analyzed and operating regimes selected. For simplicity in presenting the results of the cost analysis, three operating regimes were selected (Figure 6). Normalized results from production analysis using an operations-driven cost model were generated for the selected regimes. Comparisons were made with actual costs and with standard costs generated according to traditional standard costing practices. Figure 6 shows that the transparency and granularity of the cost results from the proposed method are superior (in terms of actual operating-cost visibility) because of the ABC-like principles used. The rates of resource consumption within each process work center are the core of the mill cost-generation procedure. When this method looks at standard costs or even calculated actual costs, it can deduce that the lack of process involvement is due to the top-down approach used, where resource cost generation is performed for the whole mill instead of for individual process units. The standard costs, which are the manufacturing costs adjusted to reflect grade recipes, do not, however, reflect reality. This information is commonly used for production cost calculation in investment design evaluation. It was found that the cost variance between operating regime is significant. The ability of the method to drill down to the process level, allows for interpretation of these differences (Korbel and Stuart (c)). Operating profit margin 45-2 (EBITDA) of a given grade (45.2 $g.cm^{-2}$) is twice as large as the one predicted by standard recipe. Whereas operating regime R45-1 is close to 3. This difference is visible also from the second part of the Figure (Figure 6) representing production costs (dissected to different cost items). The relatively low production rate of regime R45-2 is responsible for higher overhead contribution, higher electricity consumption (slightly increase in specific electricity of refiners as well as increase in indirect part of electricity costs (demand peak)). Furthermore, it was found that during the manufacturing period corresponding to operating regime R45-1, the bleaching material cost was significantly lower, when compared to R45-2, and less significant when R45-3 is considered. After further analysis, this cost variance was explained to be due to different feedstock usage (different inventory piles of the same type of chip feedstock) throughout the month.

On the other hand, the cost variance between regimes R48-1 and R48-3 (when manufacturing product grade of 48.8 $g.cm^{-2}$), is mainly due to the change in feedstock ratio. The portion of hard wood (cheaper price of feedstock fibre) was increased marginally (from usual recipe of 3% to 4%). The impact to the cost of furnish however was only marginal. On the other hand, the cost impact due to electricity consumption has caused a significant increase in the overall product cost. The electricity variance due to the increase reject throughput ratio (as a difference between operating regime R48-1 and R48-3) is due to the difference in specific electricity of reject pulp refining. The electricity impact was slightly compensated by the decrease in steam unit price due to the increase in recovery steam production; however, the net cost impact was negative to the product grade contribution margin (when operating in regime R48-1, the product margin is 4 times higher than the estimate by standard costing, whereas when operating in regime R48-3 the margin is decreased by near 1.2).

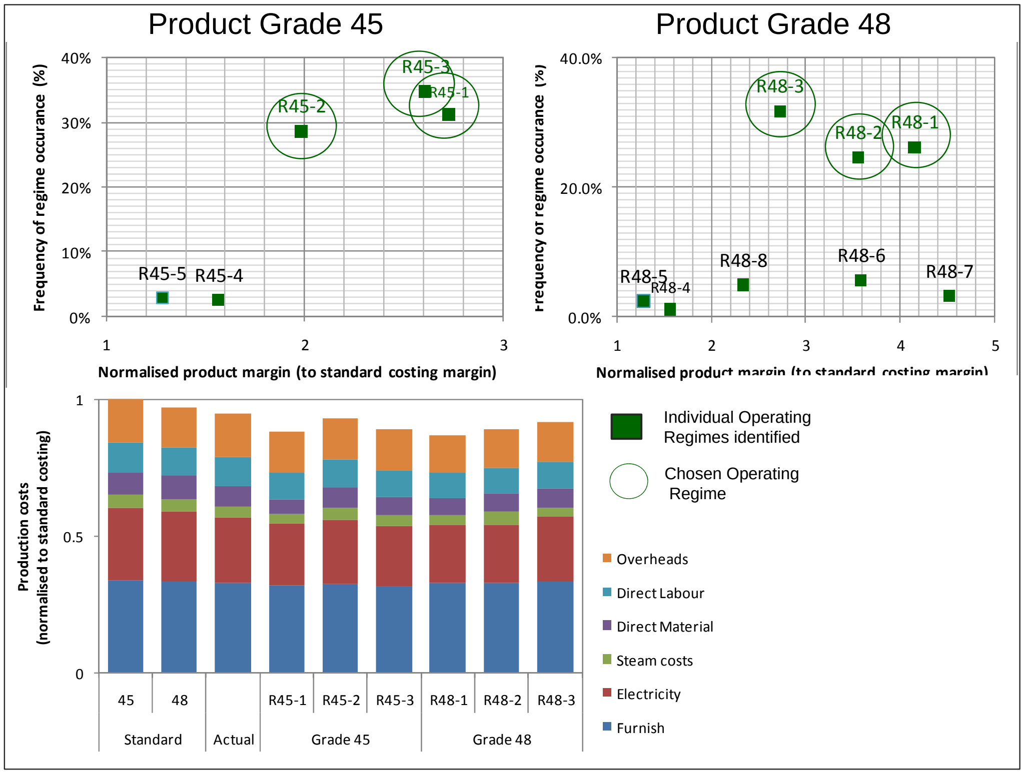


Figure 6: Operating regime selection based on frequency of occurrence (three regimes for each product were selected) for production cost analysis. The costs are compared to standard and actual costs.

<u>Production costs and profitability assessment for biorefinery scenarios</u>

The advantages of using operations-driven cost analysis over classical costing techniques for retrofit production cost analysis are presented by modelling the costs of manufacturing for each biorefinery scenario. The interpretation capabilities of the proposed operations-driven costing method are shown to be superior to traditional cost-volume variance analysis. Because the base case is an actual newsprint operation, the results are presented again as values normalized to standard costing outcomes.

The manufacturing costs of each design alternative have been divided into direct and indirect costs. These cost pools are specified based on resource drivers in PWC and OWC. Figure 7 shows the outcome from operations-driven cost analysis of a PLA strategy. The bars represent the manufacturing costs of core products in the base-case scenario and the manufacturing cost of operating regimes during and after the retrofit, including PLA. The costs are segmented into different resources. Because the production volume of PLA is very low compared to that of the paper grades (only ~1.6 % of chip throughput), the distribution of overhead from core products has a significant impact on PLA costs. The indirect costs are allocated based on the maintenance costs and the headcount for each department. Clearly, the cost impacts of the new process integration are due mainly to electricity savings (the chip pre-treatment strategy), steam price, and overhead allocation:

- The steam unit price is increased because of the lower production of recovery steam from primary high consistency pulp refiner (from 28.3 t/h to 22.1 t/h).
- The higher production costs of PLA within the second operating regime R45-2 are due to simultaneous production of grade 45.2, which lowers the production rate of PLA, increasing the ratio of indirect costs to allocation base and other operational causes discussed during the current facility characterisation.
- The highest operating margin of PLA occurs for simultaneous production of PLA and grade 45 in operating regime R45-3. The favourable production costs in operating regime R45-3, is mainly due higher production rate of PLA when compared to the other regime options. On the other hand, when looking at the current core business, this operating regime is not a favourable option of 45 grade production (Figure 9). This contra intuitive information about PLA scenario impact onto core business can be explained based on the different usage of raw material, the variance in the main pulp line production throughput and the difference in overhead allocation.

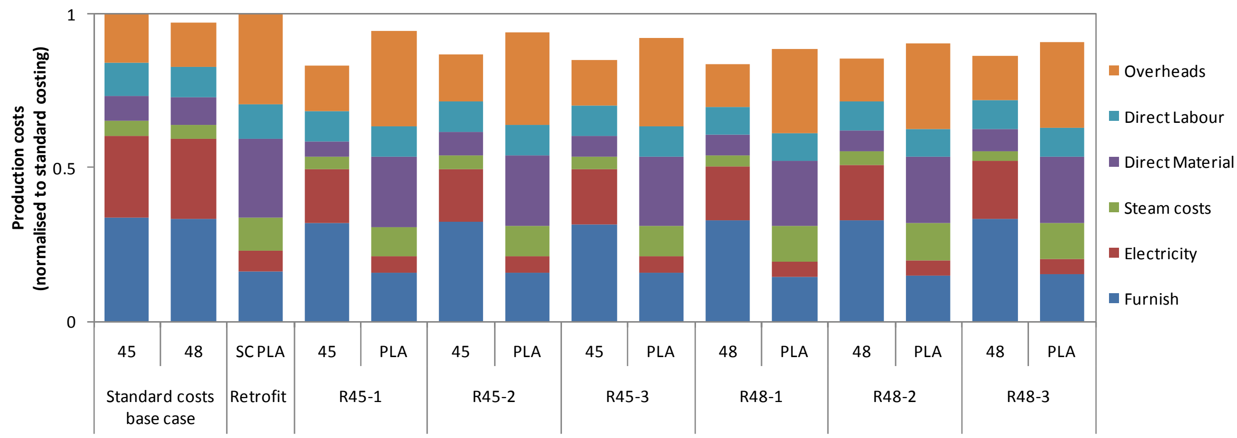


Figure 7: Production costs of the PLA option within the biorefinery scenario (paper grades and PLA) for each operating regime.

An overall summary of all three projects is presented in Figure 8. It was further found; that for all three biorefinery retrofit scenarios under consideration the impact was different in some cases. The impact of current product-cost differences (expressed within the operating regimes) on the new product line in each scenario is due mainly to steam price, electricity and overhead allocation:

- The integrated biocomposite scenario marginally influences the cost of the core business products, when compared to the base case. Capturing the cost impact is done by the use of ABC-like costing. Activity sharing, such as chips pre-treatment and primary refining is the main cause of this variance, when compared with the two VPP options (these scenario share only PWC1 activity – chips pre-treatment, whereas biocomposite option shares PWC1 and PWC2 – main pulp refining cost activity). The small change that is visible represents the increase in the price of steam per ton due to the increase in process demand. The increase in pulp throughput, from the pre-treatment step to the secondary high-consistency refiners, to enable production of 80 tons per day of biocomposite pellets increases the production of low-pressure steam recovered from the primary refiner. This production increase reduces the specific energy in the primary refiner (from 998 kWh/ODMT to 940 kWh/ODMT), which has but small impact on the overall electricity consumption. These process changes and cost impacts on core business products are minimal in the case of the biocomposites scenario.
- On the other hand, the pre-treatment of chips using acetic acid, which is common to the ethanol and PLA production biorefinery scenarios, provides significant cost savings for the core products (Figure 7).

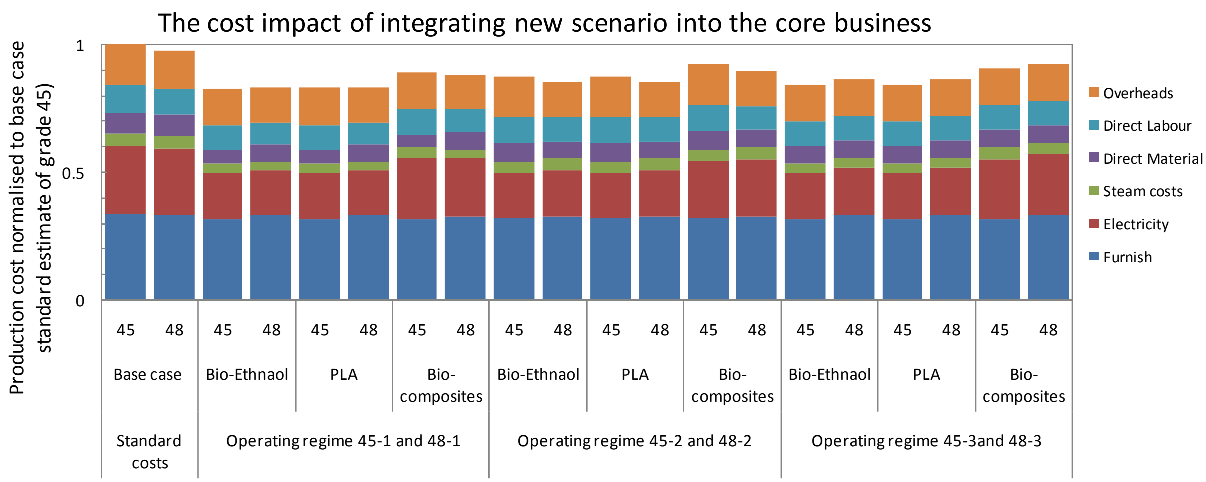


Figure 8: Comparison of the cost impact on core products for each scenario (grades 45 and 48).

Figure 7 presents the direct and indirect manufacturing costs obtained for the existing mill and for each alternative considered. The level of impact of different scenarios becomes visible when the production cost of each core product is compared for all three scenarios. From the paper products cost savings, it can be deduced that the ethanol and PLA production biorefinery options become more attractive or favourable. The increases in product margin due to the individual retrofit options are presented in Figure 9. However, the order of overall project profitability is reversed, making ethanol and PLA production uneconomical for the facility (Figure 10). The overall company's gross margin can be calculated over monthly or yearly time periods.

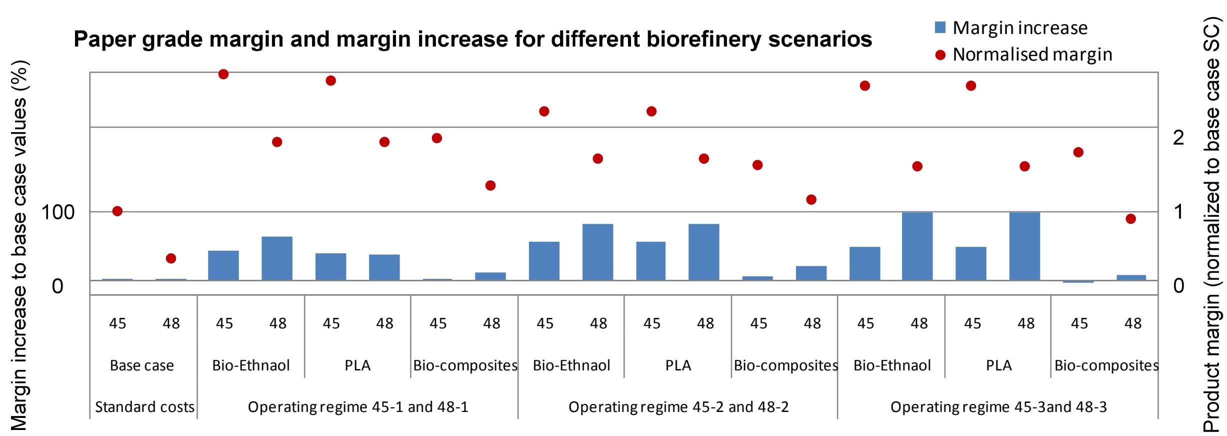


Figure 9: Contribution margins of each core product (paper grade) in different biorefinery scenarios.

From Figure 9 we can understand that the unchanged and in one case decreased margin of core product grade 45, the biocomposite option may not look attractive to decision-makers when compared to the 40% to 100% increase seen under the other scenarios. However, the overall operating margin (expressed as EBIDTA) of the whole company is the main priority of the business, and therefore the biocomposite option is the one recommended. Clearly, the simple production line of the biocomposite scenario creates significant financial value to enhance the company's overall profitability. This is due to the favourable market conditions for biocomposite only. However, when looking from core business perspective on the manufacturing cost-impact, this option has neutral or in some cases negative impact on the operating margins of individual grades. The commodity character of bio-ethanol and PLA products, and very low production capacities assumed, do not provide the facility with significant increase in monthly operating profits.

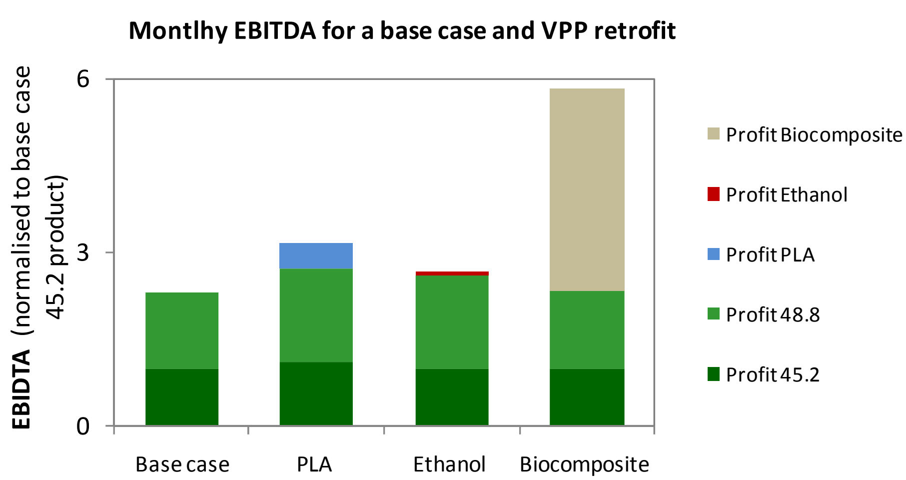


Figure 10: Overall manufacturing profitability expressed as EBIDTA (broken down by product) of each scenario under consideration.

The proposed methodology for obtaining new insights into and understanding of operating costs can be used for investment decision-making. Other financial measures, such as ROCE or IRR, must be taken into account for similar strategic company decisions. These measures can be evaluated with the knowledge derived from the ABC-like methodology presented here, and the favoured profitable scenario option can be identified for further detailed analysis. Because the chosen scenario has been analyzed using operations-driven cost analysis of various operating regimes, process flexibility considerations can be explored. For instance, information based on process flexibility and market conditions can be used to achieve further increases in company profitability. Moreover, knowledge of the true product margins can be exploited to improve mill planning and scheduling activities and to provide critical information to enable margin-centric supply-chain management.

CONCLUSIONS AND IMPLICATIONS

Today, the pulp and paper facilities are facing difficult times and require a systematic approach to finding an optimal path towards a more sustainable future through potential business transformation. To manage this transformation optimally, managers and decision-makers need to explore the powerful and robust cost-accounting systems that are today waiting to be implemented in practice. Use of lower-level process data improves the understanding of manufacturing-cost variability under different operating regimes for current and future products. The aim of this paper has been to better quantify the process-cost impacts and implications of several retrofit design alternatives for increased mill profitability in the future. For each design alternative, mass and energy balances were calculated using a simulation model linked with the core business manufacturing knowledge that was based on real-time process data. It was found that the proposed methodology provides transparent characterisation and interpretation of process-cost impacts due to retrofit projects. The unfavourable operating regimes of current core business products have shown to be potentially the most favourable options, when manufacturing two simultaneous products in the futuristic forest biorefinery scenario. These findings however, are not applicable for all the retrofit projects that were studied. Interpretation of different core operating strategies and their impact on producing parallel products in the future were presented. The essential knowledge gained from these granular cost results can be exploited in company’s strategic planning activities to enhance decision making information and provide optimal option for complex multi-product manufacturing environment. Furthermore, the actual cost knowledge of individual operating regime can be used, in combination with the biorefinery process flexibility knowledge, and market conditions, to enhance mills profitability in the future.

ACKNOWLEDGEMENTS

The authors wish to thank mill personnel for continuous support and many discussions during the project. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Design Engineering Chain at École Polytechnique de Montréal.

REFERENCES

3C Software, 2010. www.3Csoftware.com

Bagajewicz, M.J., 2001. *Process Plant Instrumentation and Upgrade*. Technomic Publishing, Lancaster, Pennsylvania.

Cooper, R., 2000. Cost management: from Frederick Taylor to the present. *Journal of Cost Management* (September/October) : 4–9.

Crama, Y., Pochet, Y., and Wera, Y., 2001. “A discussion of production planning approaches in the process industry”. *Core Discussion Paper*. No 42, University of Liege.

Ernst & Young (2003), “2003 survey of Cost Management Accounting”, IMA’s Cost management updates, issues 28, 62

Fogelholm, J., 2000. Cost function modelling in the paper industries. Ph.D. thesis, Helsinki University of Technology, Department of Industrial Engineering and Management.

Grossmann, I. 2005. Enterprise-wide optimization: A new frontier in process systems engineering. AIChE Journal 51(7):1846-1857

Greenwood, T. G. and J. M. Reeve (1992). "Activity based cost management for continuous improvement: a process design framework." Journal of Cost Management for the Manufacturing Industry 5(4): 22-40.

Horngren, C.T., Foster, G., Datar, S.M., Teall, H.D., and Gowing, M., 2006. *Cost Accounting: A Managerial Emphasis*, 4th Canadian edition.

Houtman, C. and Horn, E., 2011. Pilot trials of hemicelluloses extraction prior to thermomechanical pulp production: Part I. Submitted to *TAPPI JOURNA.*

Huang, H.J., Ramaswamy, S., Tschirner, U.W., and Ramarao, B.V., 2008. A review of separation technologies in current and future biorefineries. *Separation and Purification Technology* 62(1): 1–21.

Kaplan, R.S. and Cooper, R., 1998. *Cost & Effect : Using Integrated Cost Systems to Drive Profitability and Performance*. Boston: Harvard Business School Press. xi, 357 pp.

Kaplan, R.S.R., 1988. One cost system isn't enough. *Harvard Business Review* 66(1): 61.

Kaplan, R.S.R. and Anderson, S.R., 2004. Time-driven activity-based costing. *Harvard Business Review* 82(11):131.

Kenealy, W., Horn, E., and Houtman, C., 2007. Vapor-phase diethyl oxalate pretreatment of wood chips: Part 1. Energy savings and improved pulps. *Holzforschung* 61(3): 223–229.

Korbel, M., Bellec, S., Jiang, T., and Stuart, R P. (a). Steady-state identification for on-line data reconciliation based on wavelet transform and filtering. Submitted to *Computers and Chemical Engineering*, 2011.

Korbel, M., Wasik, L., and Stuart, P.R. (b). Practical methodology for plant-wide process data rectification in the pulp and paper industry. Submitted to *International Journal of Production Economics*, 2011.

Korbel, M. and Stuart, P.R., (c). *Cost Integration Methodology and the Forest Biorefinery. Integrated Biorefineries: Design, Analysis, and Optimization*. Taylor and Francis, 2012.

Martin Laflamme-Mayer, Matty Janssen, Paul Stuart, 2011, "Development of an operations-driven cost model for continuous processes", Part I: Framework for design and operations decision making. Journal of Science and Technology for Forest Products and Processes (J-FOR), Vol. 1, No. 1, 2011

Meyer-Pinson, V., Ruel, K., Gaudard, F., Valtat, G., Petit-Conil, M., and Kurek, B., 2004. Oxalic acid: a microbial metabolite of interest for the pulping industry. *Comptes Rendus Biologiques* 327 (9–10): 917–925.

Steen, M. and Steensland, T., 1994. Real-time profit monitoring and activity-based cost management. *TAPPI Journal* 77(2): 105–111.

Turney, P., 2008. "Activity-Based Costing: An Emerging Foundation for Performance Management", http://www.sas.com/whitepapers/indexAZ.html.

Van Heiningen, A.R.P., 2010. *Integrated Forest Products Refinery (IFPR).* http://www.osti.gov/bridge/servlets/purl/979929-FuFfqu/, accessed May 29, 2011.

Yeager, R., 1999. Costing software helps Sappi mills make the grade more efficiently. *Pulp and Paper* 73(12): 63–66.

APPENDIX E –

ON-LINE METHODOLOGY FOR OPERATIONS-DRIVEN COST ASSESSMENT OF OPERATING REGIMES USING REAL-TIME PROCESS DATA

Submitted to TAPPI Journal, 2011

On-line methodology for operations-driven cost assessment of operating regimes using real-time process data

Milan Korbel & Paul Stuart

NSERC Design Engineering Chair in Process Integration, Department of Chemical Engineering, École Polytechnique de Montréal

P.O. Box 6079, Succ. Centre-Ville, Montreal QC Canada H3C 3A7

Abstract

It is of critical importance for commodity producers to manufacture goods with the lowest cost possible in order to stay competitive in the current global market conditions. Knowing individual product margins becomes essential to determine the optimal unit prices, thus uncovering the actual operating profitability of manufacturing. This paper presents a new methodology for on-line manufacturing cost analysis using real-time process and cost data available from information management systems and process systems engineering tools, that is capable to provide this information. This methodology consists of three main steps. First, a signal processing technique based on multiscale wavelet transformation and filtering is used to analyze every segment of the plant-wide instrumentation network simultaneously. This step also cleans high-frequency noise and abnormalities from measured data and seeks to identify when manufacturing processes are operating near steady-state conditions. The second step further improves process data quality by reconciling the set of variables to the underlying fundamental process model. The plant-wide manufacturing information is updated by coaptation and correction of biased measurements. Third, this operational knowledge is integrated with financial data in an operations-driven cost model to calculate and analyze the production costs of operating regimes for the short- and long-term benefit of the company. A case study demonstrates that this methodology provides visible and transparent manufacturing-cost knowledge of a current business environment, enhancing multiple opportunities for short and long term company's benefits. This methodology can be used as a tool in day-to-day operations that would assist mill personal on multiple aspects of the organisation, such as practical process and instrumentation troubleshooting, cost-control, continuous improvements, planning and scheduling and enhance knowledge for strategic decision making activities.

Keywords

WAVELET PROCESSING, DATA RECONCILIATION, COST ACCOUNTING, OPERATIONS-DRIVEN COST MODELING, PROCESS DESIGN ANALYSIS, DECISION SUPPORT, INTEGRATED NEWSPRINT MILL

Introduction

The competitive advantage of commodity producers in the market is their ability to produce goods at the lowest possible manufacturing cost. The North American newsprint market is a very competitive environment because of many factors such as globalization, increases in energy and raw material prices, and fluctuations in paper prices. Therefore, papermaking companies are being challenged every day to stay in business. Information management systems (IMS), which have long been integrated with production processes, could help in the fight for business survivability. It has been shown that the use of IMS has increased insight into the business and production processes at many pulp and paper mills (Janssen, 2004). This survey concluded, however, that the current use of these systems is mostly for *ad-hoc* process and cost analysis. This can provide benefits only by benchmarking process trends, rather than exploiting the real and absolute values of process measurements. To derive higher-value manufacturing knowledge from these real-time measurements, advanced techniques and methodologies must be used.

The benefits of using ABC-like cost accounting for characterizing and interpreting the manufacturing costs of an operating business have been presented by Korbel and Stuart (c). A later work by the same authors explored the benefits of using the process dimension of cost information to address the future process cost impact of various integrated forest biorefinery retrofit options. This paper presents the overall framework of the multidisciplinary methodology and further details its implications for process and instrumentation troubleshooting, as well as for cost analysis and modelling of current and future production processes.

First, a brief literature review of previous work in ABC-like costing for continuous industries and in real-time data cleansing is provided. Next, current mill manufacturing practices are described, along with the forest biorefinery retrofit scenarios that were considered for future implementation. A detailed description of the multidisciplinary methodology for on-line operations-driven cost assessment of various operating regimes is presented using real-time process data. Finally, the results of a case study application of the methodology are discussed, followed by conclusions.

Manufacturing cost analysis

Current cost-control strategies using cost-variance analysis are generally based on traditional cost accounting practices. The actual causes of this variance are very challenging to determine using only top-down cost information (Steen, 1994; Janssen, 2011). Nevertheless, papermaking companies are still using traditional thinking and practices to evaluate and control their monthly costs, even though these practices were developed primarily for financial reporting (Laflamme-Mayer, 2011). Each company is using home-grown, mill-specific cost accounting methods to try to perform this task as optimally as possible. The process-based nature of activity-based costing (ABC) is being successfully exploited in other continuous manufacturing industries for this purpose, thus improving product cost visibility and hence cost-control strategies (Kaplan, 2004). ABC is an activity-driven cost accounting approach which was first developed in the 1990s in response to the increasing level of manufacturing automation to improve tracking of indirect costs. (Kaplan, 1989; Turney, 2008; Steen, 1994). In recent years, a combination of traditional and ABC accounting principles has been developed for the forestry sector (Fogelholm, 2000). In particular, the use of process knowledge in an operations-driven cost modeling framework has been shown to capture essential manufacturing information (Janssen, 2011; Laflamme-Mayer et al., 2011). The use of measured process data with financial data in a “bottom-up” cost accounting concept has yielded an improved understanding of complex pulp and paper manufacturing operations. Several applications to case studies have been presented to provide granular production cost information for supply-chain management optimization (Laflamme-Mayer et al., 2012) and retrofit design decision-making activities (Janssen et al., 2011). In these mill tactical and strategic applications, only long time-scale analyses (months to years) were performed. By reducing the cost-analysis time scale to hours, challenges related to process data quality will emerge. However, using cost assessment at these small scales will enable access to currently invisible actual product margins and their variances resulting from changes in operating strategies (Korbel and Stuart (b) and (c)). The ABC-like cost assessment of operating regimes provides transparent and granular insights into complex cost relationships by creating an understanding of the efficiency of resource usage by process cost activity and final cost object. Ultimately, the rate of resource consumption is defined by the measured process data stored in the IMS. However, these data are biased and corrupted with different types of errors. Furthermore, in the papermaking industry, the sensor network is insufficient to provide accurate information about plant-wide operation because of the lack of instrumentation and the high prevalence of sensor inaccuracy (Jacob, 2003). Therefore methodology that would make these data sets available must be used (Narasimhan, 2000; Korbel, (b)).

Process data reconciliation

The problem of reconciling data in the industrial applications was first presented by Kuehn and Davidson (1961) to minimize the error between measurements and outputs from first-principle process models. Many researchers built on this pilot data-reconciliation work to improve the robustness of the optimization algorithm (Crowe, 1983; Liebman and Edgar, 1988; Tjoa and Biegler, 1991; Arora and Biegler, 2001). To apply these classical reconciliation methods, the system under consideration must be overestimated. This necessary condition for the data-reconciliation optimization formulation ultimately enables measurements to be crosschecked and adjusted to the underlying process-model values. Weighted least-squares minimization is the most common estimator in the optimization algorithms presented in the literature. If gross errors exist in the measurements, the estimator will provide incorrect outcomes and thereby propagate errors to higher-level tasks such as cost analysis. The on-line detection of biased measurements is a challenge in practical industrial applications. Several approaches have been presented in the literature, for instance, the measurement test gross-error detection method (Tamhane and Mah, 1985) and the modified iterative measurement test gross-error detection algorithm (Serth and Heenan, 1986). Other statistically based techniques, such as the generalized likelihood ratio (Narasimhan and Mah, 1987), the maximum power test (Crowe, 1992), and the principal component test (Tong and Crowe, 1995), have shown better results.

In summary, the extensive work reported in the literature has focussed on improvements in algorithm efficiencies over the whole optimization spectrum: linear, bilinear, non-linear, and dynamic data reconciliation problems. These applications are well suited to the small, highly redundant industrial sub-systems commonly found in the chemical and petrochemical, pharmaceutical, and in some cases mining industries. However, plant-wide applications are challenging as an industrial concept because of the process dynamics involved (Bagajewicz, 2001). Dynamic data reconciliation is very computationally expensive for on-line industrial applications (Benqlilou, 2000), and plant-wide steady-state data reconciliation creates errors due to process dynamics. The papermaking industry is a special manufacturing environment which is characterized by both these types of challenges: low redundancy in measurements and highly dynamic processes. To the authors' knowledge, a methodology that would provide plant-wide steady-state data sets in low-redundancy systems has not yet been developed (Korbel, (a)).

Signal processing

To generate near-steady-state data set candidates from real-time measurements, signal processing techniques must be used. Several methods for on-line process status identification based on statistics or filtering have been presented in the literature (Cao and Rhinehart, 1995; Bakshi and Stephanopoulos, 1993), but these create data distortions when abnormalities are present. With the advent of wavelet transform theory, the signal processing field has evolved towards multidimensional analysis of trends, which enables accurate multiscale representation of functions. Flehmig et al. (1998) explored the features of wavelet transforms to approximate process measurements. Nounou and Bakshi (1999) used wavelet features to identify and remove random and gross errors. Recently, Jiang et al. (2003a) proposed a wavelet method for detection of near-steady-state periods. These methods are used for off-line signal representation. Wavelet data processing can be used to eliminate random noise and abnormalities efficiently and simultaneously to analyze a trend for steady-state occurrences. If implemented carefully and systematically, the multiscale features of wavelet transform analysis can be exploited on-line for plant-wide applications (Korbel et al. (b)).

Scope of the work

The scope of the work described in this paper includes:

- The development of a novel methodology for production-regime cost analyses based on ABC-like principles (the use of real-time process data from information management systems coupled with company financial information)
- A demonstration of the application of this methodology to improve process and instrumentation troubleshooting, as well as to improve manufacturing cost visibility and to quantify the cost variability of different operating regimes

- The application of this methodology to identify the most profitable operating regions in the current business environment and to evaluate possible future retrofit scenarios with individual product portfolio considerations

Overall Methodology

The proposed novel methodology can be divided into several main steps (Figure 1). Each step consists of several tasks as described below:

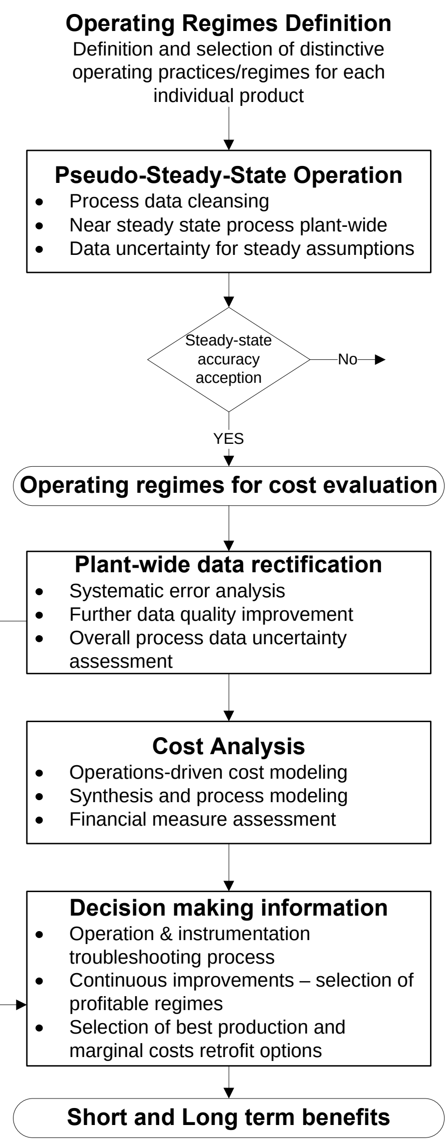


Figure 1: Overall methodology of the new production cost approach.

Operating regime definition

In this initial step, the aim is to characterize the operating differences that occur while producing the same product, thus defining the notion of an operating regime. These differences may be due to process design changes, use of different equipment, changes in production volume, or changes in control setpoint strategies. An illustrative example of a change in process regimes is presented in Figure 2. A thermo-mechanical process (operating in regime 48-01) producing a given pulp quality needs to respond to the requirement of the paper machine for a change in pulp specifications (for instance, due to frequent paper breaks). This is achieved by lowering the freeness value of the pulp using manipulated control variables STP1–STP6. These adjustments cause transient periods before the process reaches the final required freeness value (regime 48-02).

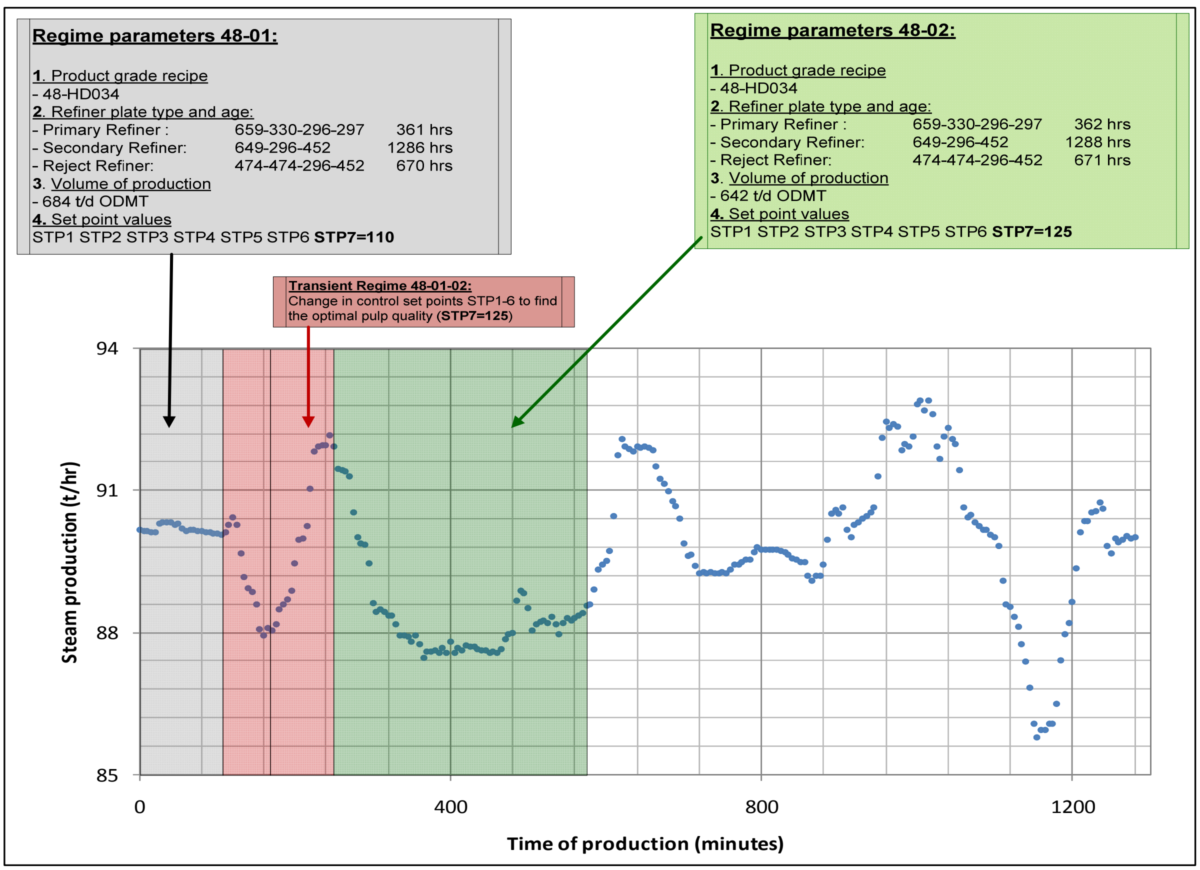


Figure 2: Change in operating regime due to required change in pulp quality.

Pseudo-steady-state operation

First, the process measurements within the time frame under analysis (corresponding to the length of the operating regime) are extracted from the IMS as a noisy signal (Figure 3). The optimal wavelet-transform parameters are adjusted iteratively to achieve optimal performance of the method (optimal data pre-processing and accurate steady-state detection). After applying a wavelet transform at the chosen scale, Gaussian noise along with abnormalities are extracted from the process trend. The de-noised signal is then analyzed for potential steady-state occurrences using a three-step simultaneous metrology (Korbel et al. (a)):

- The starting point of the steady-state period is detected using WT characteristics and its first derivative (values of the predetermined alpha parameter).
- High-frequency features of the measured signal which were not eliminated in the first step are removed by filtering, and steady-state duration is approximated.
- Finally, the steady-state end point is detected through WT feature analysis.

The multivariable near-steady state is then identified by comparing the whole set of variable states over time.

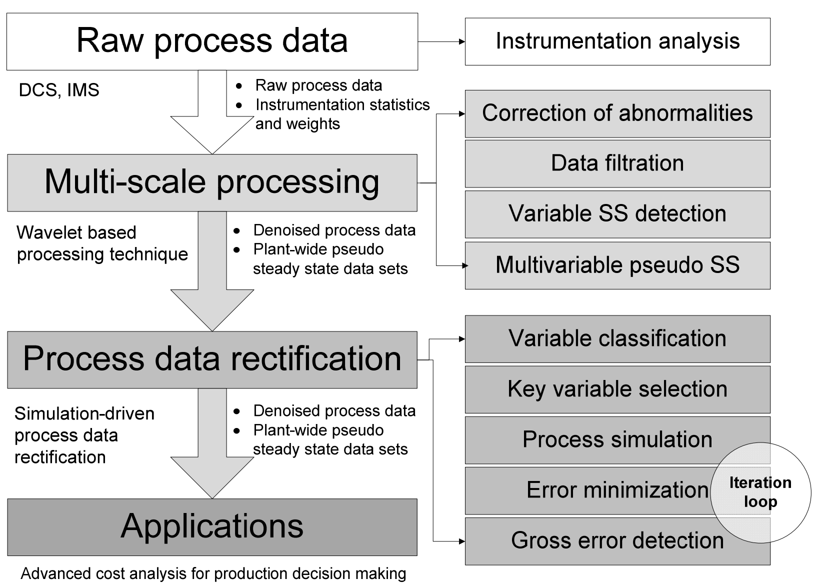


Figure 3: Methodology for steady-state detection and plant-wide simulation-driven data rectification (Korbel et al. (b)).

Plant-wide data rectification

The aim of the third step of the overall methodology is to guarantee process data validity compared to the underlying process model. During this procedure, biased measurements can also be identified. Figure 3 presents the series of steps used to apply raw data rectification to plant-wide data sets representing near-steady-state operation (Korbel et al. (b)). The principle of the method is to exploit the adaptive features of the CADSim software in the context of the papermaking industry. The additional mathematical characterization of the various processing units that are unique to pulp and paper mills creates more degrees of freedom for the data reconciliation problem. Furthermore, the real-time nature of this software makes it possible to unify wavelet signal processing and data reconciliation into one on-line methodology providing plant-wide data sets representing near-steady-state operation.

Gross error handling

The method presented here exploits the statistics of historical process data. The on-line analysis of each individual measurement is then compared to its historical values. When a change in measurement values is detected, the biased value is estimated, and the measurement is corrected. The data-reconciliation procedure is repeated by updating this new corrected measurement value.

Cost analysis: operations-driven costing

The fourth step of the methodology is to perform an ABC-like cost analysis, using the manufacturing information which was acquired using signal processing and data rectification techniques, to the real-time data for a given operating regime. The pillar of this methodology—the operations-driven cost model—was then developed to explore this novel and unique insight into papermaking production for the short- and long-term benefit of the company.

The operations-driven cost modelling approach consists of these four steps (Korbel and Stuart (c)) and was implemented using the Impact: EDCTM software:

- Characterization of the process operation based on real-time process data. The data are dissected to describe multiple operating regimes for manufacturing products in the core business and in various biorefinery retrofit scenarios

- Defining and organizing cost data and cost drivers into matrices that correspond to underlying fundamental (mass and energy) equations
- Modeling and calculation of manufacturing costs for operating regimes and biorefinery design alternatives
- Analysis, interpretation, and evaluation of the cost-model outcomes.

The core of the methodology for producing a manufacturing-cost assessment of operating regimes and the associated product-cost distribution is the ABC-like philosophy. The model was developed in the necessary detail to extract complex cost information on operating and design changes and was used to assess the production costs per tonne of a newsprint grade (as a cost object). The individual cost activities, referred to as process or overhead work centres (PWC and OWC respectively), are defined to capture and represent the production chain in the form of manufacturing sub-systems (in some cases mill departments). The direct costs are linked to these activities based on the process model and the real-time data, whereas the indirect or overhead costs are linked based on predetermined allocation rules and drivers. The PWCs consist of the following essential elements:

- The process operating criteria and characteristics describing the process regime or the retrofit design alternative
- The integration of cost with mass and energy flow along the manufacturing operation
- Specific calculations (cost-related or simply unit conversion-related) for individual PWCs
- The allocation (or assignment) of indirect and overhead costs, and
- The core ABC-like engine, with its operations-driven cost calculation.

With the use of operations-driven cost modeling and plant-wide real-time data sets corresponding to near-steady-state operation, costs can be calculated for each operating regime. The tracing and allocation of different types of costs within the costing framework is presented in Figure 4.

.

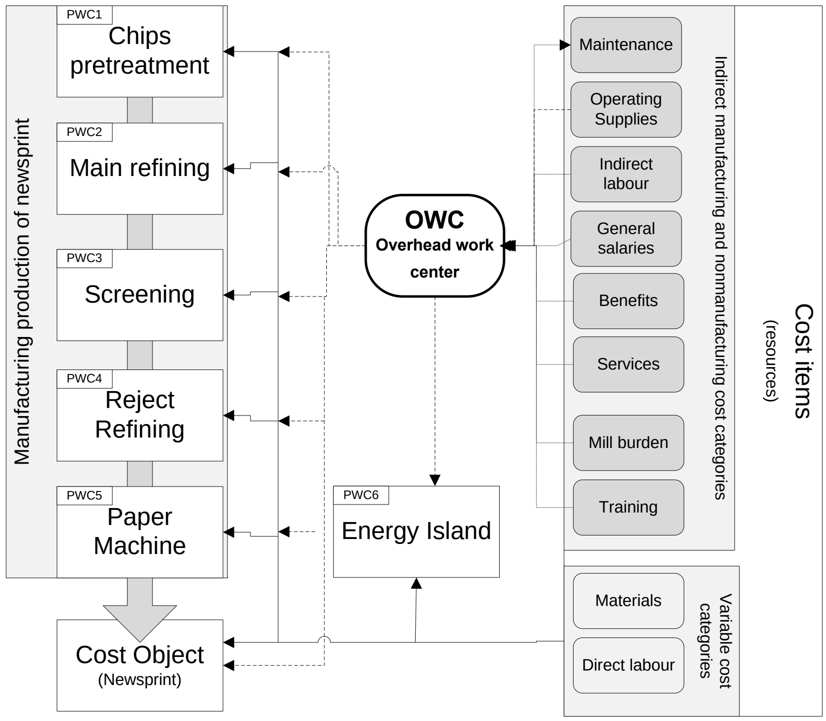


Figure 4: Definition of the process and overhead work centres which capture various cost categories within the current business base-case mill.

Operational and design problem statement

<u>Current mill operation: definition of production regimes</u>

The base-case mill is an existing integrated newsprint mill. The thermo-mechanical facility produces different pulp qualities based on paper mill demand and specifications, with the throughput matched to that of the paper mill. Two basis weights of newsprint products are being produced: 48 $g.m^{-2}$ and 45 $g.m^{-2}$.

The following manufacturing steps are involved in the base-case mill:

- 1 newsprint machine with a total average production of 680 tonnes/day of newsprint,
- 1 TMP line with a total average production of 680 tonnes/day of pulp.

Furthermore, the following supporting processes are part of the base-case mill configuration:

- A wastewater treatment plant processing 30,000 m^3/day
- A boiler plant producing 2500 GJ/day of steam

- A steam recovery unit in the TMP line, producing 3000 GJ/day of steam.

For simplicity of illustration, only the three most common operating regimes were selected to characterize the operating problem. This is justified because the probability of occurrence for each of the three regimes within the same product grade is between 20% and 30%. The six operating regimes describe six manufacturing states producing two distinct paper grades and are labelled as 48-1(2,3) and 45-1(2,3). The parameters used to describe the operating regime (Figure 2) are production volume, the type and age of the refiner plates, and the control setpoint strategy used for a given pulp freeness.

Forest biorefinery retrofit scenarios

The base-case is a competitive newsprint mill with generally limited access to chips. For this reason, the company was interested to see the impact of a forest biorefinery strategy that could be integrated into their current operations. Three distinct forest biorefinery scenarios for an integrated newsprint mill were chosen for production-cost analysis:

- Cellulosic ethanol production: 120 gallons per day ethanol production from hemicelluloses extracted before pulping
- PLA production: 11.5 tons per day of polylactic acid (PLA) production from lactic acid from hemicelluloses extracted before pulping
- Biocomposite production: 80 tons per day of biocomposite pellets produced from a blend of TMP fibres and polypropylene.

.

Results and discussion

<u>Signal processing and data reconciliation</u>

The outcomes from applying the first two steps of the methodology to the signals extracted from the IMS indicated that the proposed wavelet method is robust and provides significant improvements to the accuracy of individual measured variables. The overall accuracy of near-steady-state detection for both small and large industrial systems is improved compared to other techniques. Several measurements were identified as having biased values. This information was provided to the mill personnel, and calibration of each malfunctioning instrument was performed. Figure 4 shows one of the corrected measurements. The method was able to identify when the plant-wide operation reached near-steady-state conditions. It was concluded that the base case under study is a relatively stable production system, and multiple near-steady-state data sets were extracted for each operating regime. For instance, operating regime R45-2 was characterized by 14 identified steady-state snapshots of the plant-wide operation. The error, expressed as the standard deviation of key variables, was relatively low, as shown in Figure 5. This was the case because the manufacturing process was exceptionally stable during the month analyzed.

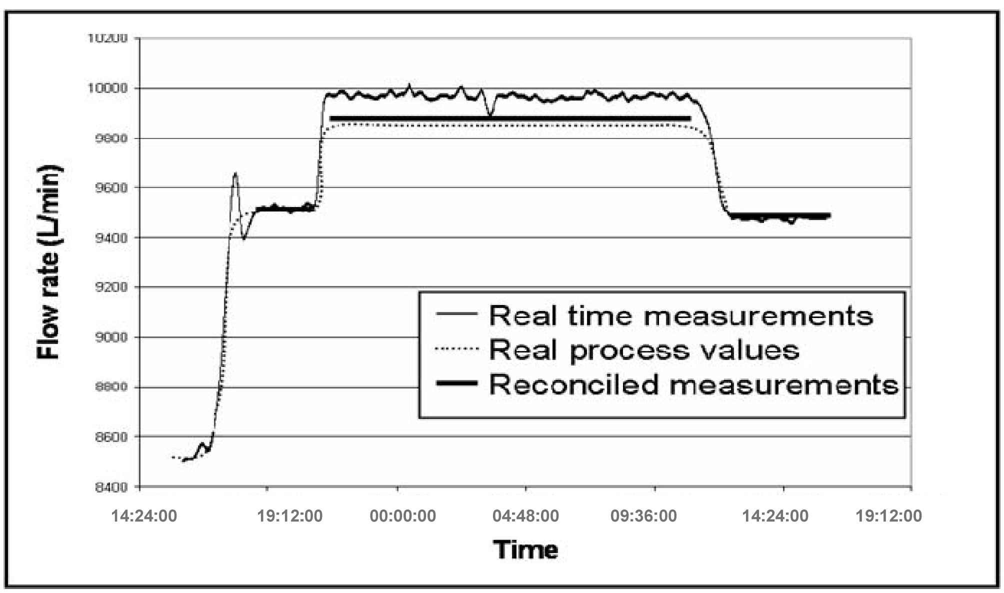


Figure 4: Example of a corrected biased measurement.

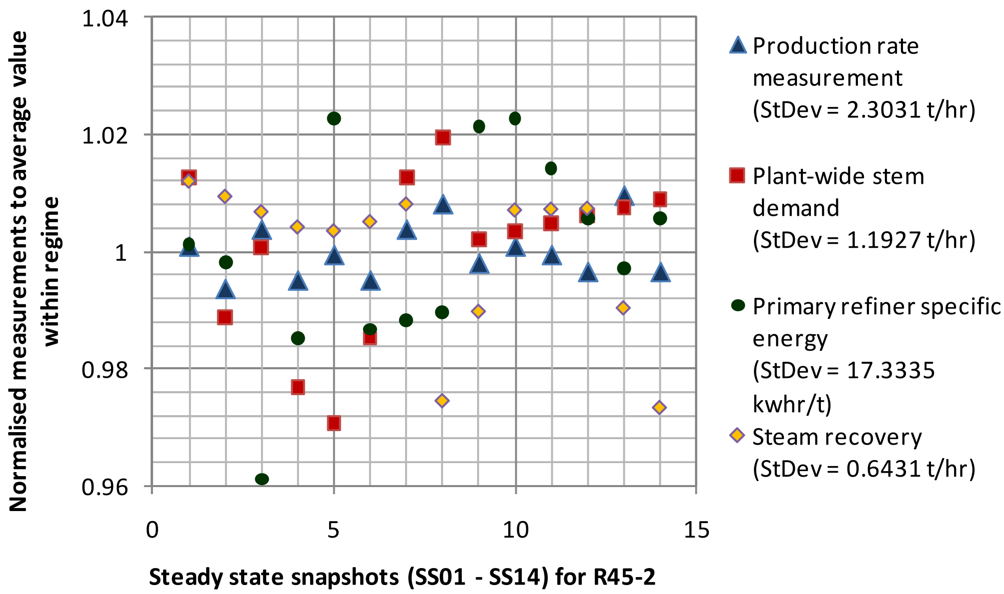


Figure 5: Variability in one of the operating regimes characterized (R45-2).

The variance in production volume has the highest impact on manufacturing costs, even though the net impact is very marginal on the variable costs. However, the impact of fixed and overhead costs can be significant because of the change in allocation base. When looking separately at each near-steady-state snapshot, it can be noted that in some data sets, the individual standard deviation for each variable is smaller, e.g., by comparing SS01 to SS05. Therefore careful analysis of each near-steady-state operation was carried out using least-squares error values (the results from the data reconciliation step), and the data sets that showed high variance were omitted from the cost calculations. This variance is mainly due to changes in raw material quality and the resulting changes in measured values. An analysis using several near-steady-state operating conditions provides a statistical representation of the costs, thus allows for validating the manufacturing costs.

Current core-business cost analysis

Using the operations-cost model, direct and overhead production costs were calculated for each PWC, for each process regime (Figure 5), and for various retrofit alternatives (Figure 6). The second type of costs that cannot be traced (but allocated) by traditional accounting was directly associated with each individual PWC based on allocation information. This association makes indirect costs behave similarly to direct costs by introducing the link between overhead cost pools and cost objects (newsprint or future FBR products) using PWCs. The costs presented were obtained by averaging the cost output of several steady-state data sets representing individual operating regimes. The error associated with each operating regime was relatively small when compared to the cost variations between individual operating regimes.

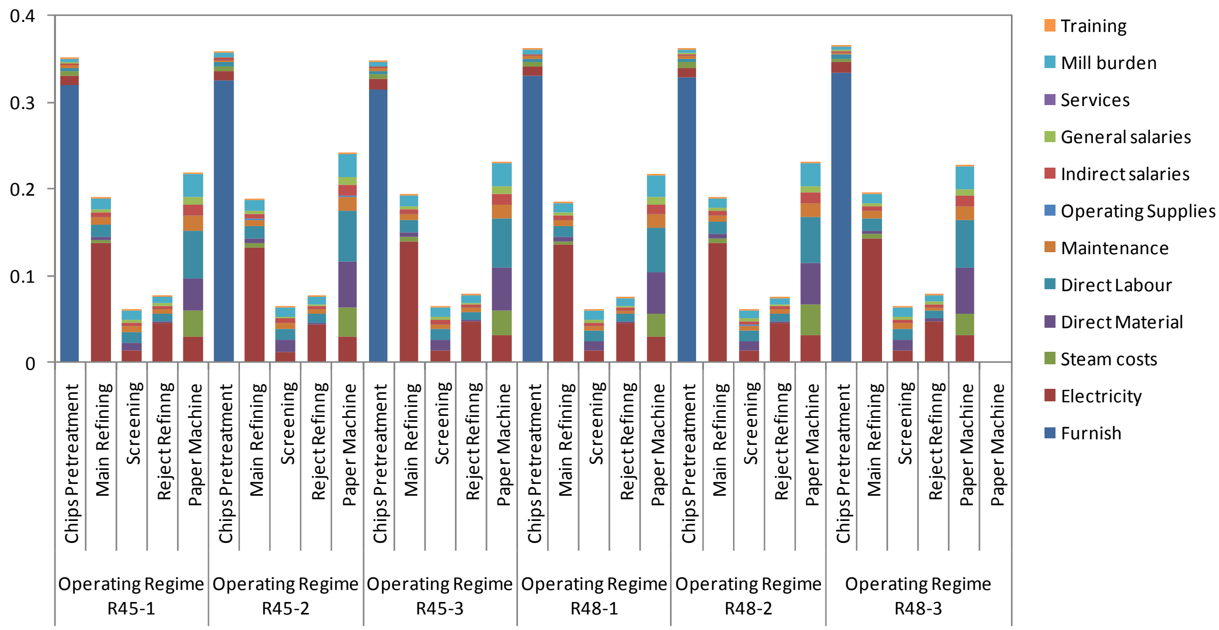


Figure 5: Production costs of different operating regimes divided into process work centres and for both types of costs (direct and overhead costs).

The three operating regimes analyzed are characterized by pulp quality and production volume, and hence the “free” steam production from refining high-consistency pulp varies among them. The main variance in costs between different products or operating regimes was interpreted as follows:

- The cost difference between two grades is mostly due to changes in pulp quality and production throughput in the paper mill. This difference causes an increase in steam demand, and therefore more high-pressure steam must be produced from the boilers (using natural gas, electricity, or oil fuels), which changes the internal unit price of steam. The unit price of steam is calculated iteratively as a ratio of high-pressure steam production price and recovered-steam production price. This iteration step is necessary because the recovered-steam price depends on interactions among multiple PWCs.

- The electricity cost variation is mainly due to the specific energy difference between operating regimes, the production rate, the ratio of reject volumetric rate to mainline production rate, and potentially changes in refiner plate characteristics. However, generally the mill uses the same types of plates for long time periods, and therefore the impact of this factor was not addressed.
- The electricity variance associated with the change in reject-line ratio (the difference between operating regimes R48-1 and R48-3) is due to the difference in specific electricity of reject pulp refining.
- The cost of fibre remains fairly constant, with slight variations due to small yield fluctuations, which becomes essential when comparing different grade recipes (e.g., when comparing the productions of 48.8 $g.m^{-2}$ and 45.1 $g.m^{-2}$ grades).

The cost increases due to steam price and its dependence on the interrelation of process activities, as well as the electricity cost increases, can be captured and interpreted only because of the operations-driven nature of the cost model. The cost model has therefore integrated the resources consumed and their related costs into the process activities in each PWC and has brought process and financial knowledge closer together. This type of analysis demonstrates that the operations-driven cost model presented here is able to unify the flows of cost and process information to increase the transparency of production costs. This kind of characterization of production costs and interpretation of variances using lower-level real-time process data has never been done before in the pulp and paper industry. Hence, the opportunity to use this approach for continuous mill improvement can be expected to minimize manufacturing costs and to increase the company's cash flow, thus providing essential competitive advantages

<u>FBR retrofit design scenario cost analysis</u>

For each design alternative, mass and energy balances were calculated using a simulation model and real-time process data for each scenario. The impact of process costs on the core paper products was analyzed using the proposed methodology. The results provide granular cost information for manufacturing, as represented by the different regimes (Figure 6). The various findings from using the proposed methodology to analyze the cost impacts can be summarized as follows:

- PLA production costs are increased when producing newsprint grades of 45 basis weight. These cost differences were identified as being due mainly to steam and overhead costs.
- The steam unit price was found to be lower because of the increased demand for high-pressure steam when producing grades of 48 basis weight. However, the specific steam usage per ton of PLA changed only marginally because the increase in the internal steam price had caused overall steam costs to increase.
- The R45-3 operating regime is the most profitable operating scenario for simultaneous PLA and 45-basis-weight grade production (the absolute cost difference is more visible from Figure 7). The advantage of the R45-3 operating regime is due mainly to the higher rate of PLA production, which can be attributed to the higher rate of parallel production of paper grade 45.

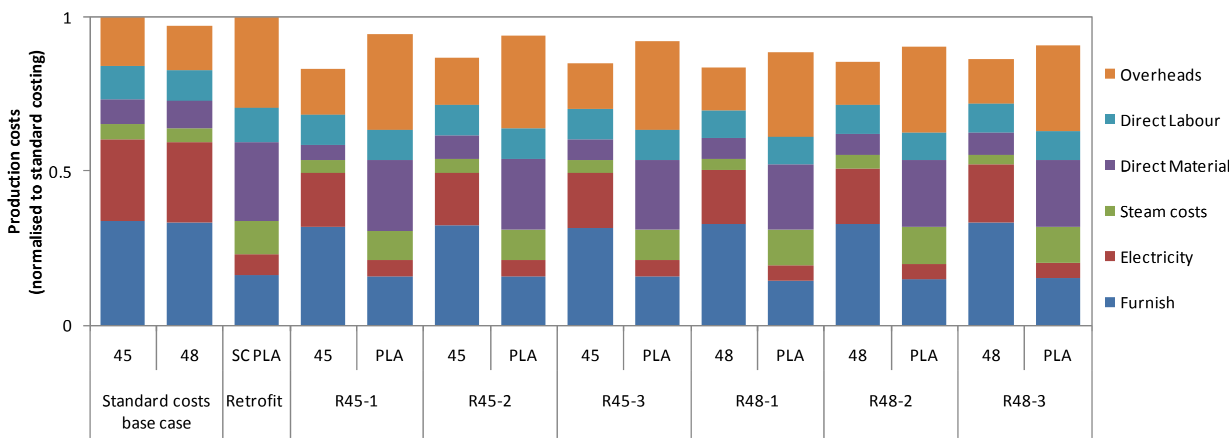


Figure 6: Production costs of the PLA option within the FBR scenario for each operating regime (paper grade costs are normalized to standard costing for grade 45, while PLA costs are normalized to PLA standard costs to make the comparison of results more visible) (Korbel and Stuart (d)).

- The majority of the variance can be observed to be due to overhead costs because the allocation base (tons of PLA produced) had increased. The specific use of steam and electricity changed only marginally due to the corresponding increase in steam demand and hence in unit steam price.
- Similarly, the R48-1 operating regime was identified to be the most profitable for parallel PLA and grade-48 production. The manufacturing costs of this scenario were significantly

lower than for any option with parallel 45-grade production. This difference is due to the increase in PLA production, thus increasing the allocation base (tons of PLA produced).

The variations in the different scenarios become even more apparent when looking at the contribution (operational) margin of each product for different scenarios characterized by operating regime (Figure 7). The values of the changes in actual or true product margins (normalized to the base case calculated by standard costing) provide information which can be explored in strategic decision-making. The near-zero or negative margin increase compared to the base case for simultaneous biocomposite production in regime R45-3 indicates that this biocomposite option may not be very attractive. PLA production appears very attractive when looking at only this potential decision-making parameter; in some cases, the margin increases by over 200% (PLA and grade-45 production under R45-1 and R45-3 regimes). However, analysis of the overall EBIDTA per month of each project indicated that the biocomposite option is far more attractive (twice as much cash flow per month as any of the VPP options).

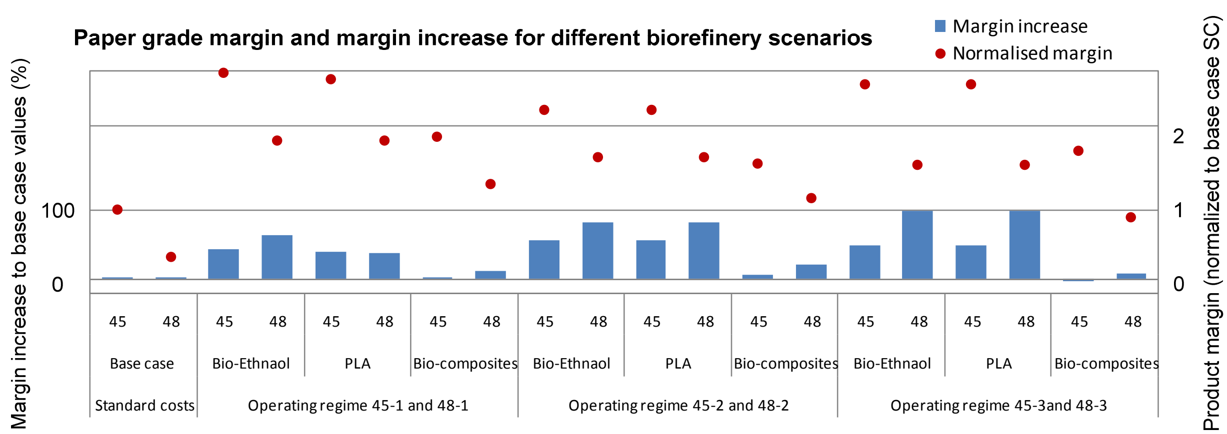


Figure 7: Contribution margins of each core product (paper grade) under different biorefinery scenarios (Korbel and Stuart (d)).

CONCLUSIONS AND IMPLICATIONS

Individual pulp and paper mills have tremendous opportunities to increase their level of competitive advantage by developing tools and methodologies to exploit their information management systems. These systems have gathered vast amounts of process and cost data which are not now fully exploited. The use of the activity-based costing philosophy and its variations could help managers improve forest-company profits based on these valuable data. This paper presented ABC-like methodology for on-line manufacturing cost assessment that is based on using this real-time process and cost data from information management systems. The supporting pillar of the method is an on-line technique that is able to detect near-steady-state operation and to establish that the steady-state data sets are relatively accurate. The accuracy and validity of the operating-regime representation with the use of near-steady-state data increases with the number of near-steady states identified. When wavelet signal processing is combined with a data reconciliation method, the analysis can provide a complete set of plant-wide reconciled data representing operating regimes.

It was found that this methodology provides visible and transparent manufacturing-cost knowledge of a current business environment, enhancing multiple opportunities for short and long term company's benefits. This methodology can be used as a tool in day-to-day operations that would assist mill personal on multiple aspects of the organisation. The strategy of continuous mill improvement can benefit from this information, process flexibility can be explored when tracking paper prices, a margin-centric supply chain will benefit from providing actual product margins for each operating regime, process-driven explanation of cost variances on a daily basis will enhance cost-control practices, and outcomes from signal and data processing will provide enhanced instrumentation and process troubleshooting. These are but a few essential examples of the potential implications of the proposed method for short-term facility's benefits. Furthermore, the analysis of potential retrofit design scenarios would benefit from having an operations-driven cost model based on ABC-like principles in place, providing systematic integration of process and cost data. The model focuses on process costs instead of only on product cost assessment and is able to analyze the implications of operating regimes for future retrofit design options. This leads to the identification of profitable operating conditions in both current and future manufacturing environments and makes possible increased profitability of retrofit design alternatives.

REFERENCES

3C Software, 2010. www.3Csoftware.com

Abramovich, F., B.W. Silverman, Wavelet decomposition approaches to statistical inverse problems, Biometrika 85 (1) (1998) 115–129.

Arora, N. and L. T. Biegler, "Redescending estimators for data reconciliation and parameter estimation," Computers and Chemical Engineering, 25, p. 1585 (2001)

Bagajewicz, M.J. Process plant instrumentation: design and upgrade. Lancaster, PA: Technomic Publishing Company, 2000.

Bakshi, B.R., STEPHANOPOULOS, G. (1993). Representation of Process Trends III. Multiscal Exraction of Trends from Process Data. Computers and Chemicals Engineering, Vol.18, pp 267-302.

Bellec, S., JIANG, T., KERR, B., DIAMOND, M., STUART, P.R. (Corrected version submitted in 2006). On-line Processing and Reconciliation of Pulp and Paper Mill Data, Pulp & Paper Canada.

Benqlilou, C.; Tona, R. V., Espuña, A., Puigjaner, L. (2001), On-line application of dynamic data reconciliation. Proceedings of the Conference on Process Integration, Modeling, and Optimization for Energy Saving and Pollution Reduction PRESS 2001, AIDIC:□ Milano, Italy, May 2001; p 403.

Bhat A., SARAF N.,(2004). Steady-State Identification, Gross Error Detection and Data Reconciliation for Industrial Process Units. Industrial and Engineering Chemistry, Vol 43, pp 4323-4336.

Cao, S., RHINEHART, R.R.(1995). An Efficient Method for On-line Identification of Steady State. Journal of Process Control, Vol.5, No.6, pp 363-374.

Crowe, C.M., Campos, Y.A.G., Hrymak, A., Reconciliation of process flow rates by matrix projection. Part I: linear case, AIChE J. 29(6), 1983, 881–888.

Dabros, , M., PERRIER, M., FORBES, F., FAIRBANK, M., STUART, P.R. (2005). Model-Based Direct Search Optimization of the Broke Recirculation System in a Newsprint Mill, accepted to Journal of Cleaner Production.

Flehmig F., WATZDORF, R.V., MARQUARDT, W. (1998). Identification of Trends in Process Measurements using the Wavelet Transform, Computers and Chemicals Engineering, Vol.22, pp 491-496.

Bagajewicz, M.J., 2001. Process Plant Instrumentation and Upgrade. Technomic Publishing, Lancaster, Pennsylvania.

Cooper, R., 2000. Cost management: from Frederick Taylor to the present. Journal of Cost Management (September/October) : 4–9.

Crama, Y., Pochet, Y., and Wera, Y., 2001. "A discussion of production planning approaches in the process industry". Core Discussion Paper. No 42, University of Liege.

Ernst & Young (2003), "2003 survey of Cost Management Accounting", IMA's Cost management updates, issues 28, 62

Fogelholm, J., 2000. Cost function modelling in the paper industries. Ph.D. thesis, Helsinki University of Technology, Department of Industrial Engineering and Management.

Grossmann, I. 2005. Enterprise-wide optimization: A new frontier in process systems engineering. AIChE Journal 51(7):1846-1857

Greenwood, T. G. and J. M. Reeve (1992). "Activity based cost management for continuous improvement: a process design framework." Journal of Cost Management for the Manufacturing Industry 5(4): 22-40.

Horngren, C.T., Foster, G., Datar, S.M., Teall, H.D., and Gowing, M., 2006. Cost Accounting: A Managerial Emphasis, 4th Canadian edition.

Houtman, C. and Horn, E., 2011. Pilot trials of hemicelluloses extraction prior to thermomechanical pulp production: Part I. Submitted to TAPPI JOURNA.

Huang, H.J., Ramaswamy, S., Tschirner, U.W., and Ramarao, B.V., 2008. A review of separation technologies in current and future biorefineries. Separation and Purification Technology 62(1): 1–21.

Janssen, M., Laflamme-Mayer, M., and Stuart, P.R., in FOCAPO 2003: Proceedings Foundations Of Computer Aided Process Operations. 2003. Coral Springs, FL, USA.

Janssen, M., Laflamme-Mayer, M., Zeinou, M.-H., Stuart, P., 2004. Survey indicates mills' need to exploit it systems with new business model. Pulp & Paper 78 (6), 46-51.

Janssen, M. and P. Stuart (2011). "Multi-criteria decision making for retrofit process design using process- and supply chain-level criteria." Submitted J-4 journal

Jacob, J., Paris, J., Data sampling and reconciliation—application to pulp and paper mills. Part II: Case studies, Appita J. 56(2), 2003b, 116–120.

Jiang, T., Chen, B., He, X., Industrial application of wavelet transform to the on-line prediction of side draw qualities of crude unit. Comput. Chem. Eng. 24, 2000, 507–512.

Jiang, T., Chen, B., He, X., Stuart, P.R., Application of a steady-state detection method based on wavelet transform, Comput. Chem. Eng. 27, 2003a, 569–578.

Jiang, T., Chen, B., Jasmin, K., Stuart, P.R. (2003b), Strategy for improving data quality for a kraft pulp mill recausticizing plant, Proceedings, Foundations of Computer-Aided Process Operations (FOCAPO 2003), Coral Springs FL , 2003b, 467–470.

Kaplan, R.S. and Cooper, R., 1998. Cost & Effect : Using Integrated Cost Systems to Drive Profitability and Performance. Boston: Harvard Business School Press. xi, 357 pp.

Kaplan, R.S.R., 1988. One cost system isn’t enough. Harvard Business Review 66(1): 61.

Kaplan, R.S.R. and Anderson, S.R., 2004. Time-driven activity-based costing. Harvard Business Review 82(11):131.

Kenealy, W., Horn, E., and Houtman, C., 2007. Vapor-phase diethyl oxalate pretreatment of wood chips: Part 1. Energy savings and improved pulps. Holzforschung 61(3): 223–229.

Korbel, M., Bellec, S., Jiang, T., and Stuart, R P. (a). Steady-state identification for on-line data reconciliation based on wavelet transform and filtering. Submitted to Computers and Chemical Engineering, 2011.

Korbel, M., Wasik, L., and Stuart, P.R. (b). Practical methodology for plant-wide process data rectification in the pulp and paper industry. Submitted to International Journal of Production Economics, 2011.

Korbel, M. and Stuart, P.R., (c). Cost Integration Methodology and the Forest Biorefinery. Integrated Biorefineries: Design, Analysis, and Optimization. Taylor and Francis, 2012.

Laflamme-Mayer, M., Matty Janssen, Paul Stuart, 2011, "Development of an operations-driven cost model for continuous processes", Part I: Framework for design and operations decision making. Journal of Science and Technology for Forest Products and Processes (J-FOR), Vol. 1, No. 1, 2011

Liebman and Edgar, 1988 M. Liebman and T. Edgar, Data Reconciliation for Nonlinear Processes, 1988 AIChE Annual MeetingWashington DC (1988)

Meyer-Pinson, V., Ruel, K., Gaudard, F., Valtat, G., Petit-Conil, M., and Kurek, B., 2004. Oxalic acid: a microbial metabolite of interest for the pulping industry. Comptes Rendus Biologiques 327 (9–10): 917–925.

Nounou, M., Bakshi, B.R. (1999). On-line Multiscale Filtering of Random and Gross Errors without Process Models. AIChE Journal, 45:5, pp.1041-1058.

Serth and Heenan, 1986 R.W. Serth and W.A. Heenan, Gross error detection and data reconciliation in steam metering systems. AIChE Jl, 32 (1986), p. 733.

Steen, M. and Steensland, T., 1994. Real-time profit monitoring and activity-based cost management. TAPPI Journal 77(2): 105–111.

Narasimhan and Mah, S. Narasimhan and R.S.H. Mah, Generalized likelihood ratio method for gross error identification. AIChE Jl, 33, 1987, p. 1514.

Tamhane and Mah, 1985 A.C. Tamhane and R.S.H. Mah, Data reconciliation and gross error detection in chemical process networks. Computers chem. Engng, 17 (1985), p. 409

Tjoa, I.B., Biegler, L.T., Simultaneous strategies for data reconciliation and gross error detection of nonlinear systems, Comput. Chem. Eng. 15(10), 1991, 679–690.

Tong, H., Crowe, C.M, "Studies in Data Reconciliation and the Detection of Gross Errors," PhD Thesis McMaster University, Hamilton, Ont., Canada (1995).

Turney, P., 2008. “Activity-Based Costing: An Emerging Foundation for Performance Management”, http://www.sas.com/whitepapers/indexAZ.html.

Van Heiningen, A.R.P., 2010. Integrated Forest Products Refinery (IFPR). http://www.osti.gov/bridge/servlets/purl/979929-FuFfqu/, accessed May 29, 2011.

Yeager, R., 1999. Costing software helps Sappi mills make the grade more efficiently. Pulp and Paper 73(12): 63–66.

APPENDIX F –

Cost integration methodology and the forest biorefinery

Chapter of a book: El-Halwagi, Mahmoud and Paul Stuart, Integrated Biorefineries: Design, Analysis, and Optimization, CRC Press/Taylor & Francis (2012)

Cost integration methodology and the forest biorefinery

Milan Korbel and Paul Stuart[15]

NSERC Design Engineering Chair in Process Integration, Department of Chemical Engineering, École Polytechnique de Montréal PO Box 6079, Succ. Centre-Ville, Montreal Que., Canada H3C 3A7

Abstract

In today‘s competitive global market, the corporate strategy is to improve the cost curve by shutting down unprofitable facilities and constantly cutting manufacturing costs. However, to survive and stay competitive, it is of crucial importance that the papermaking and forest industries exploit the other side—the revenue side—of the profit picture. Improving profit margins through cost reduction has its limits, whereas revenue growth tends to compound over time. Today the forest industry has a tremendous opportunity to enhance mill revenues by introducing sustainable alternatives for manufacturing bulk and fine chemicals to their core business.

The cost-reduction strategy certainly helps short-term business survival and should be exploited to its limits. By knowing and understanding the true profit margins, both cost-cutting strategies and critical decision-making will be enhanced. To do this, the current cost structure and cost accounting practices must be refined to account for operational knowledge and resource consumptions. Activity-based costing (ABC) has become widely accepted as the standard approach to providing a complete picture of how the organization‘s resources are consumed. Even though the implementation of an ABC strategy in the processing industries is a laborious and challenging task, the pulp and paper sector would benefit significantly. This is especially true with an operations-driven costing approach—the "son of ABC" that exploits lower-level

---

[15] Telephone +1 514 340 4711 extension 4384; fax +1 514 340 5150

Email address: paul.stuart@polymtl.ca

process data to assess resource consumption by different manufacturing activities. By understanding costs from a process perspective, the everyday knowledge of product margins would be visible to decision makers for critical decisions. This manufacturing information could be then exploited to analyze the optimal pathways towards enhancing the revenue side by looking at retrofit design options.

Chapter Objective: The overall goal is to discuss the state of the art in cost methodologies for discrete and continuous industries, focusing on forestry: how can we exploit cost information to address a biorefinery implementation more effectively. A case study example demonstrates the importance of these advanced cost integration methodologies.

Contents

1. Introduction

In today's global environment, manufacturing companies, especially commodity manufacturers, are being challenged to define new business strategies to sustain their competitive position in the market. In recent years, the most common strategy for the North American forest industry has been the unpopular closure of unprofitable operations. This arguably straightens up enterprise-level cost curves to some extent, but will probably not sustain the company's future. Another common and yet critical strategy over the short term is to minimize production costs, for instance through continuous improvements. This exercise is typically carried out by process engineers who set up benchmarking performance parameters for individual processes or for the whole mill. Undoubtedly, well-defined process-based benchmarking is helping many mills to improve process efficiency and hence to tighten up their monthly spending. On the other hand, cost-based performance measures, generally established by mill cost accountants, assist in evaluating the difference between current production costs and their expected values, which are the budget values. The cause of this variance is difficult if not impossible to determine from only cost information. Obviously, each mill has developed its own particular cost accounting practices, and some may even claim that a certain level of process information is involved in their cost analyses. From practical experience, it is evident that the communication link between cost accountants and process engineers is generally missing.

The ultimate goal of both cost variance analyses and process benchmarking is to achieve better control of a company's costs and hence to maximize its product margins. The task of determining true or actual profit-margin values for all products is rather difficult, especially in a multi-product environment where this information becomes crucial. In the forest industry, most accountants are using the traditional way of determining these values, which typically provides monthly or weekly volume-based valuations. This result, however, incorporates a wide range of operating practices and production recipes, mill interruptions, and paper machine breaks. Therefore, traditional costing is a somewhat *ad hoc* procedure and should be used with caution in decision-making activities. Ideally, values corresponding to each operating situation in the mill would be accessible. To capture such information, mill personnel must integrate financial and process knowledge into one costing system, for instance using an operations-driven costing approach (Janssen and Laflamme-Mayer, 2006). In practice, this would enable operators to avoid expensive operating regimes or accountants to understand their product margins and cost variances from a process perspective. The pillar of this approach is the principles of activity-based costing (ABC), which are generally accompanied by a certain level of complexity when setting up the system. This is one of the reasons why forest products companies do not yet recognize similar costing systems for problem-solving as an addition to their current traditional system used for financial reporting.

"One cost system isn't enough." Robert S. Kaplan, 1988

The short-term values of operations-driven (ABC-like) cost accounting practice would help companies to understand and interpret their resource consumption, hence to minimize their current core business production costs. Moreover, the understanding of individual product margins in multi-product environment will improve planning/scheduling tasks and simultaneously will provide an input to identify the optimal product mix ratio. This new information would create a knowledge-based manufacturing opening the possibility towards margin-centric supply chain implementation for long-term company benefits. If potential business transformation is being considered, the mill personal would benefit from the most well informed decision-making information at hand.

This chapter deals briefly with the state of the art in today's cost accounting practice, focusing on the continuous manufacturing environment. The first part of the chapter seeks to introduce the reader briefly to cost accounting principles. With this basic knowledge, the current state of the art in cost accounting practices in processing industries can then be presented. Next, the integration of academic knowledge about costing with process engineering knowledge will be discussed, and the first part of the chapter finishes with a discussion of how the current forest products industry would benefit from such integration. To provide the reader with a full understanding of the problem, the last part of the chapter provides a concrete case study example comparing current and leading-edge costing strategies.

2. Understanding Production Costs

2.1. Current cost accounting practices and what we are missing

Cost accounting is the heart of the accounting framework that provides valuable financial insight to decision makers. The information provided is confidential and is used only internally to help managers find the optimal way to maximize the company's profits. The environment and the outcome of decision-making activities is the cost accounting system. Various companies may use several different cost accounting systems for problem-solving. Considering that the limits of practice are entirely within the company's control, the prepared cost reports can be based on whatever rules, standards, or rational bases are chosen. Cost accounting information is commonly used in the second pillar of an accounting framework, financial accounting. This branch of accounting deals with public corporate information used solely for a company's financial statements, and its preparation must follow generally accepted accounting principles (GAAP).

The general elements of cost accounting can be divided into three pools: material, labor, and overhead costs. Direct material and labor costs are generally variable costs and are a function of the number of units manufactured or sold. Overhead costs, on the other hand, are fixed costs that do not change with the level of production. For instance, management salaries, rent, or depreciation expenses do not vary from month to month, even though the rate of production is never the same. The ability to track these various cost elements accurately determines the value of the accounting system to final decision-making activities. In the early 20th century, this task was not difficult because overhead costs were negligible compared to material and labor. However, it became more complex to account correctly for indirect and overhead costs once the face of manufacturing had shifted from a labor-intensive to a machine-intensive environment.

The ultimate focus in every organization is to control costs. Often a company chooses to use only one costing system, even though there are several approaches available. The most commonly used in today‘s industrial practice are:

- Cost-volume-profit (CVP) analysis,
- Standard cost accounting,
- Throughput accounting,
- Lean accounting,
- Resource consumption accounting (RCA), and
- Activity-based costing.

The first two systems, cost-volume-profit and standard costing, are often referred to as *traditional* or *normal* costing and are used extensively in the pulp and paper sector. This traditional approach was created for the needs of the early industrial era when the total costs were dominated by variable elements. The overhead and other indirect costs are accounted for based on simple volume-based measures such as labor or machine hours. Therefore, a product with a low level of labor hours is allocated less overhead cost. However, the actual costs may be very different if this product requires special attention or testing. The resulting unit production cost becomes even more distorted when overhead and other indirect costs begin to dominate overall manufacturing costs. Then it is strongly recommended that other supplementary costing systems be used.

The four principles stated below are relatively new in management accounting. Throughput accounting was developed for the enterprise-wide level, to help identify factors that limit the enterprise in achieving its established goals (Eliyahu 1992). In lean accounting, the essential philosophy is to preserve value with less work. This approach was developed for the car industry which was aiming to eliminate waste while simultaneously minimizing production costs and time, using techniques such as poka-yoke (Robinson 1997) or value-stream mapping (Rother 1999). Resource consumption accounting (RCA) is a fully integrated and complex managerial approach that uses available state-of-the-art methods. The combination of the German *Grenzplankostenrechnung* (GPK) cost management system and activity-based costing principles create a system that can be used and interpreted by non-accountants. An extensive discussion of each approach is beyond the scope of this chapter; the interested reader is referred to, for example, Horngren (2006). The pertinent activity-based costing principles are discussed in more detail later in this chapter.

As mentioned earlier, forest industry accounting practices are dominated by traditional costing because of its simplicity and the wide understanding of this approach among accountants. An important part of standard costing is a variance analysis. By breaking down the overall variance into the three pools listed below, this analysis helps managers identify where the difference between actual and budgeted costs has occurred:

- Labor-cost variation
- Material-cost variation
- Volume variation.

This information helps managers to identify the source of the overall cost variance, but not the cause of it. For instance, if the variance is largely due to material-cost variation, accountants with the help of process engineers need to drill down into historical process data to interpret the variance and take appropriate action.

The problem is that traditional costing considers all costs as variable with regard to production volume. This often creates inaccuracy in fixed costs whenever the volume of production changes.

Furthermore, arbitrary rather than cause-and-effect overhead allocation makes the traditional approach highly inappropriate in a multiproduct environment. Another problem in the current general accounting profession, not only in the forestry sector, is the emphasis on financial accounting. Most of the time, decision-makers must create their own cost analysis based on financial accounting reports. However, these statements contain aggregated and distorted costs with no activity data incorporated, leading to poorly informed decisions. There are a few existing advanced systems at the academic level or already being used by advanced processing industries such as the petrochemical sector. The pillar of these approaches is the principles of activity-based costing, which is briefly discussed in the following section.

2.2. Cost allocation and activity based costing

Activity-based costing (ABC) is a relatively new philosophy that emerged in the 1980s in response to overhead allocation discrepancies (Kaplan 1989). By simply adding an activity as a link between resource consumption and a cost object, the knowledge of costs incurred in the organization is improved significantly. The activity becomes a fundamental cost item whose value is directly assigned to the final cost objects such as products and customers. In other words, the rate of resource spending is traced to an activity, and the activity is then traced to the product, as shown in Figure 2.2.1.

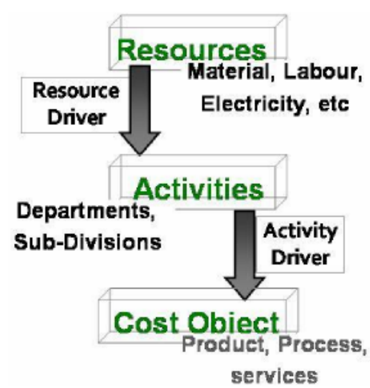


Figure 2.2.1: Resources Consumed by Activities and then by Cost Objects.

The ultimate advantage of using ABC is that it attempts to assign all costs to final cost objects, including marketing, engineering, and administrative costs. This added ability to trace indirect costs directly enables accountants to track overheads rationally and as closely as they track direct costs. This is done by making use of so-called *drivers*. As shown in Figure 2.2.1, resources are linked to activities by resource drivers, and similarly activities are linked to cost objects by cost drivers. According to this definition, resource drivers determine the amount of a resource consumed by each activity, while activity drivers specify how different cost objects (products, customers) consume these activity costs. Labor hours, kWh, and number of shipments are examples of resource drivers, whereas number of customers and number of products are examples of the second stage, the activity driver. The difference between these drivers is that the former focuses on why things happen and the latter on what happens (Emblemsvåg and Bras 2001). The implementation of an ABC system may be a complex and expensive task, and therefore it is important to determine the minimum number of appropriate drivers that will meet accounting objectives.

As shown in Figure 2.2.2, the process-oriented character of ABC means that it is implemented in two simple and logical stages, while structure-oriented traditional costing is implemented in one. This fundamental principle is the basis for increasing the accuracy of the cost data (Drucker, 1996). Traditional costing cannot encompass this critical linkage between actual causes and associated costs. Furthermore, advanced ABC has recently evolved into multistage systems where individual activities can be used by other activities before being used by final cost objects, thus enhancing even more the accuracy of cost modeling (Emblemsvåg and Bras, 2001).

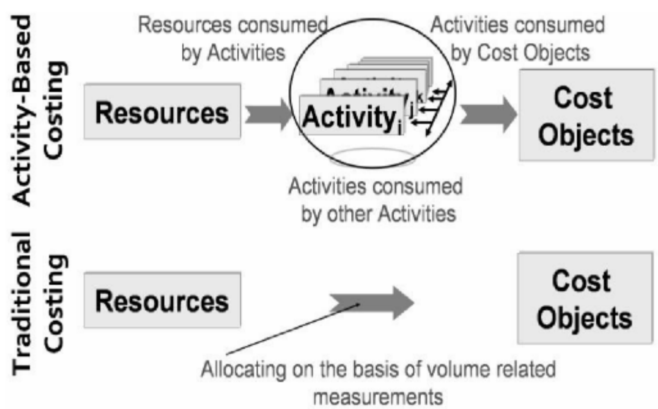


Figure 2.2.2: Activity-based costing and traditional costing (Korbel and Stuart (c)).

In a continuous manufacturing context, the process-oriented character of an ABC system and the causal relationships between cost drivers and activities make the method highly suitable for modeling and analyzing costs. The availability of real-time cost and process data from information management systems (IMS) makes ABC easier to implement. It must be made clear that ABC is a cost accounting system that can help managers understand their actual costs and improve their profits efficiently. Traditional methods are complementary to the financial reporting prepared according to GAAP.

ABC-like cost accounting

A cost accounting system that is used by a wide spectrum of industries is the *resource consumption accounting* (RCA), whose development has been strongly influenced by German cost accounting and ABC principles. The structure is very close to variable costing, a well-documented method discussed in cost accounting textbooks, but rarely used by industry. RCA and its variations are extensively used by advanced processing industries such as mining, petrochemicals, and chemicals. Often their costing methods are confidential and inaccessible to the public or to researchers. In general, RCA is based on three fundamental pillars (for further details, the reader should refer to Friedl (2005) or David (1999)):

- *View of resources*: The use of a high volume of cost pools establishes a clear linkage between resource spending and a company's costs and revenues;
- *Quantity-based model*: The value of the costing system is created in this pillar by the use of operations data and models. Traditional costing uses the output of variance analysis with dollar values, creating a fixed-costs bias. By contrast, RCA exploits causal operational relationships;
- *Cost behavior*: Understanding the nature of costs is a very important aspect of the third pillar of RCA. The clear distinction between direct, indirect, variable, and fixed costs is based on aggregating pools.

There have been significant changes in recent years, although not well documented; some forestry companies are approaching now ABC-like costing for improved decision-making activities. For example, Fogelholm (2000) has discussed the difficulties of product costing in the paper making industry and its potential industrial application. This approach is now a pillar of Metso Automation's MetsoDNA (Dynamic Network of Applications) that some companies are presently using for product-customer decision making as well as it helps their budgeting activities. The application seek to anticipate and determine the resource requirements for the next individual customer orders based on current raw material content, dimensions and quantities of the paper product (Fogelholm, 2004).

Some academic cost accounting frameworks have been developed based on ABC philosophy with potential industrial applications. For example, an approach that integrates ABC principles with environmental metrics to perform analytical economic and environmental assessment for decision-making activities was developed by Emblemsvåg and Bras (2001). Their activity-based cost and environmental management (ABCEM) system is extensively discussed in their 2001 paper. The use of an uncertainty variable introduces extra complexity and versatility into the system. The ABCEM has been applied to a wide range of industries including furniture, carpets, and supply vessels, where it has provided insights and highlighted potential areas for improvement.

Lastly, a sophisticated ABC-like approach that integrates process and cost information into one system, *operations-driven costing* (Janssen and Laflamme-Mayer 2006), is the core of this chapter. This method will be used for cost assessment in the case study part, and its results will be compared to the outcomes from traditional costing. The basis of this approach is in making a link between costs and process operations data using principles similar to those of activity-based costing. This approach is similar to RCA in some aspects, but is more versatile because it includes an in-depth engineering understanding of the process operation. The following section discusses this approach in more detail.

2.3. Operations-driven costing approach

The *operations-driven costing approach* (ODCA) is an interdisciplinary approach developed by accountants and process engineers in the pulp and paper industry. Understanding the cost of process performance is a critical success factor for paper mills. Chemical and process engineers are concerned with developing systematic tools and methodologies for both optimal design and optimal process operation (Figure 2.3.1). These procedures range from nano to industrial scales (Puigjaner, 2006). The concept can be understood from the supply chain point of view, where on the one hand, product quality is determined on the nano or micro scales, and on the other hand, the desired product properties are determined by its functionality and structure.

In the pulp and paper industry, fiber micro properties influence the quality of pulp and paper products. On the macro industrial scale, a reflection of the micro complexity of the fiber structure can be brought to light using information extracted from real-time data through IMS. The most practical way of doing this in the pulp and paper industry is to develop tools and methodologies for macro or mega scale applications that are based on real-time data and that reflect the meso and micro scales according to the general chemical-engineering definition of complexity levels.

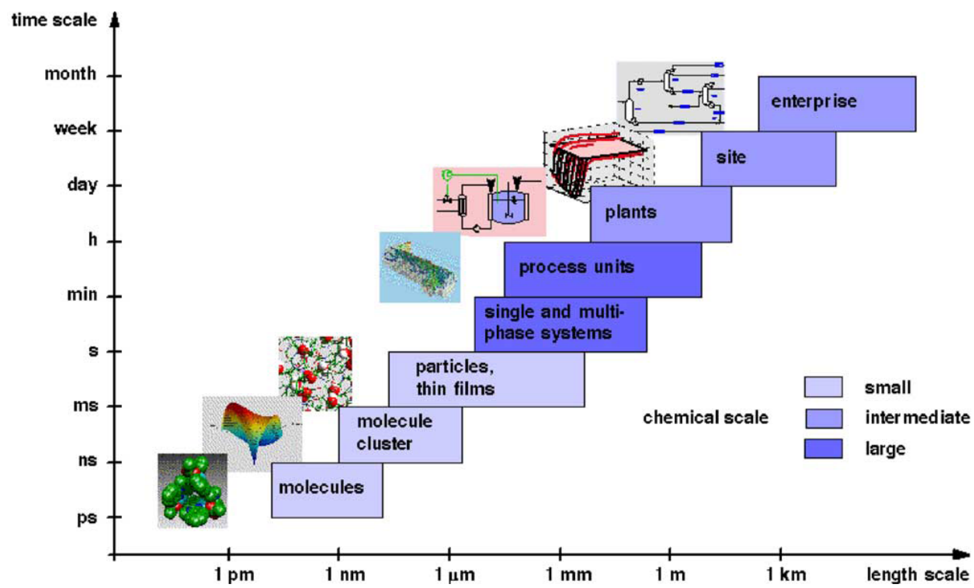


Figure 2.3.1: From micro scale to macro scale complexity or the “chemical supply” chain. (Grossmann, 2005)

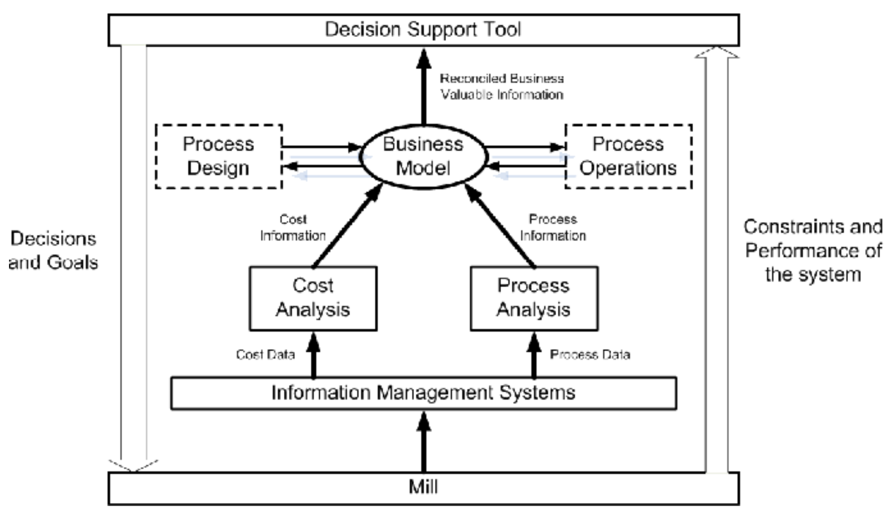
Actual product margins

Paper producers have multiple products satisfying numerous customers with different needs and corresponding product specifications and prices. It is therefore crucial to understand individual product margins; however, current cost practices and systems provide only approximate values that are based on time-framed (usually monthly) spending information. As discussed earlier, current practice involves the use of conventional accounting systems that aggregate costs over the manufacturing period and uses a standard recipe, e.g. it is based on experience from the process operation. This overall cost information incorporates various changes in process operation due to mechanical (process) or raw material disturbances. Engineers and accountants recognize that within the manufacturing period, the generation of cost differs from one product to another as well as within the production process for the same product. However, it is not a simple task to determine these cost variances. First, current accounting practices cannot accommodate cost information from a process perspective, and second, the discrepancies in the current cost data are significant. In 2003, a survey by Ernst & Young and the Institute of Management Accountants indicate that 98% of respondents claim that cost reporting is distorted, with indirect costs and overhead allocation being the main biases reported, and almost 40% believe that the cost data they receive are significantly inaccurate.

With the current state of academic knowledge, it is possible to replace the current inaccurate margin information by the true values of product margins. However, to implement this change, actual operating knowledge must be involved. The use of IMS becomes pertinent to extracting process measurement data that provide knowledge about the underlying process. However, the lack of reliability in certain measurements as well as the lack of instrumentation on site makes this task very challenging. Many older mills have these difficulties plus a lack of process data redundancy[16]; however, in these cases, there are certain ways to create redundancy and proceed with ODCA (Korbel et al. 2011).

Operations-driven cost modeling framework

Janssen and Laflamme-Mayer (2006) developed an operations-driven cost modeling framework to provide in-depth understanding of resource consumption by integrating process and cost data. The bottom-up structure (Figure 2.3.2) provides mutual communication between different business levels. The resulting generic framework can be used to enhance the understanding of manufacturing processes both for design and for operational decision support.



[16] Redundant measurement data points offer some level of possibility of crosschecking for potential measurement biases.

Figure 2.3.2: Overview of the bottom-up process-based approach (from Laflamme-Mayer, 2011).

Later, Laflamme-Mayer (2008) presented in his thesis an application of operations-driven cost modeling to assess the production costs for different product campaigns. This information was then used for planning and scheduling and optimization of high-level supply-chain analysis. The understanding and differentiation of product margins for each campaign can be used to enhance the current *ad hoc* representations of product margins. This versatile view of manufacturing costs in the paper industry is revolutionary and has tremendous value for reducing production costs.

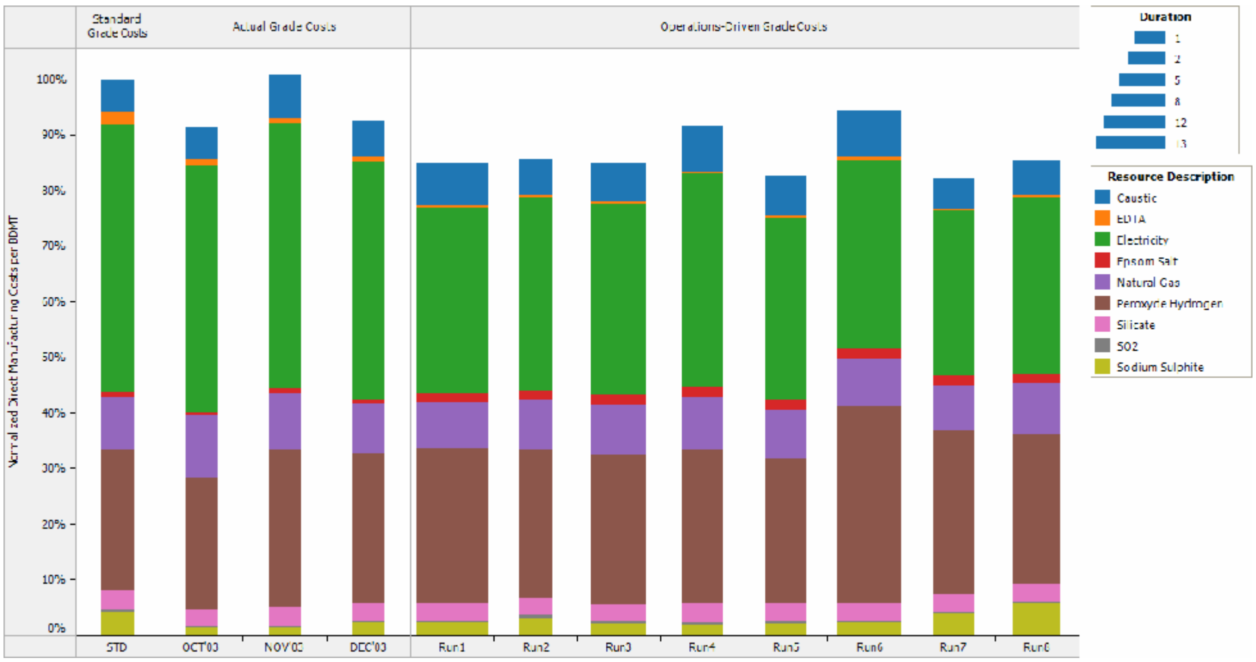


Figure 2.3.3: Comparison of standard cost, actual grade cost, and operations-driven grade costing (from Laflamme-Mayer, 2011)

Figure 2.3.3 compares standard costing with actual and operations-driven costing information. Standard costs represent how the resources should have been used to manufacture a particular grade; actual grade costs are the true resource consumptions at the end of the three-month period. The operations-driven grade-cost assessment breaks up the three-month period into segments corresponding to campaign runs. From these results, it is clear that manufacturing the same product varies significantly from one campaign to another.

Cost assessment of operating regimes

With the use of advanced data processing tools and methodologies, campaign costs can be broken down further to assess the cost of different operating strategies. Every paper product is manufactured according to its production recipe. However, within this recipe, different operating strategies can be followed by operator choice or as a result of natural process-material interactions. These strategies, referred here as operating regimes, are driven by the process design characteristics and operating practices. For instance, the use of different chip-refining plates, the control setpoint strategy for freeness control, and the open or closed nature of process loops and units are examples of operating regimes. Recent advances in plantwide acquisition systems, which capture real-time data from the pulp and paper operation, provide an opportunity for creative ideas to improve knowledge-based manufacturing and to support decision-making activities. In this context, it is proposed to exploit the operations-driven cost modeling framework to assess the costs of different operating regimes (Korbel, 2007).

The overall structure of this cost-modeling vision can be understood from Figure 2.3.5. Traditional cost accounting procedures permit *ad hoc* profitability analysis of different products (grades). To go further and to understand the actual costs incurred from a chemical engineering perspective, a process model should be used to assess the profitability of individual operating regimes with the probability of occurrence of each regime. At this level, the information can already be used by decision-makers to choose the most profitable operating regimes and to eliminate costly ones.

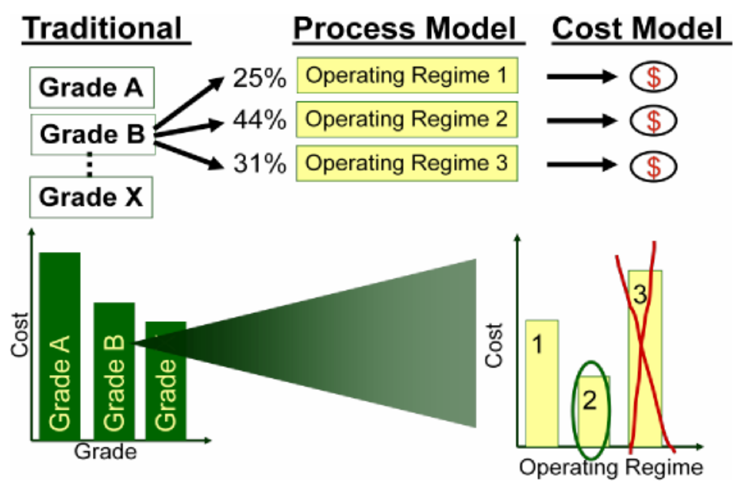


Figure 3.3.4: Smart data dissection for operations-driven cost modeling approach.

For these purposes, it is important to use a plant-wide process data set that is extracted from the operation under near-steady-state conditions. Furthermore, the process operation must stay close to steady state for a satisfactory time period to guarantee reliable and accurate information. To achieve a satisfactory level of data quality, advanced cleansing techniques must be used. The mathematics of these methods is beyond the scope of this chapter; interested readers should refer to Jiang (2003) and Bagajewicz (2001). Briefly, these data cleansing techniques improve the accuracy, reliability, and completeness of a given measurement data set by correcting different types of errors. First, a wavelet-based processing technique is used to identify the state of the process. When the system is at pseudo-steady-state, elimination of random white noise and abnormalities is carried out. The second step is data reconciliation, which helps to improve further the accuracy and completeness of the data and the level of compatibility of the data set with the process operation. There are many challenges in carrying out such an analysis in the pulp and paper industry, but they can be overcome using a smart and highly practical approach (Korbel, 2007).

Because pulp and paper facilities operate in an item-based or order-driven manufacturing environment, the use of a regime costing system would create sustainable value for the corporation. Not only would savings in manufacturing costs be achieved, but also high-value supply-chain modeling and potential transformation of the business to a biorefinery would benefit from these valuable insights into production knowledge. What kind of information a company can access through these methods is described in a real case-study example in the next section of this chapter.

3. A case study: Integrated newsprint mill and its business transformation

3.1. Problem statement

The objective of this case study is to compare the accounting practices currently in use with others awaiting industry implementation. This will make it possible to understand the value in using operations-driven cost modeling on an everyday basis to improve the cost knowledge of today's core business (short-term values) as well as to help assess the potential business transformation of the forestry sector and its margin-centric supply chain (long-term values). To clarify the differences in the two costing approaches and their consequences, a concrete example is needed. A cost analysis by conventional and operations-driven cost accounting was performed on a real single-line newsprint mill situated in Canada.

3.1.1 Case study description

Current core business

The current business is a simple single-line integrated thermomechanical newsprint production facility located in Canada. Figure 3.1.1 shows a simplified block diagram of the base-case process operation, including various manufacturing steps such as chip refining, pulp screening, and bleaching before papermaking.

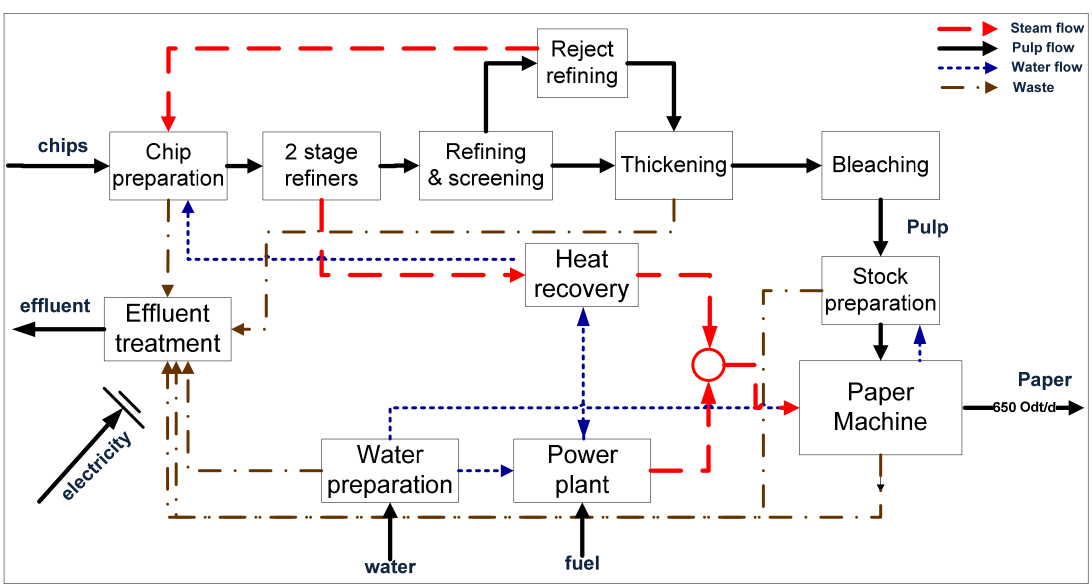


Figure 3.1.1: Block diagram of the base-case TMP pulp and paper process.

The process first heats chips by process steam in an atmospheric vessel at approximately 110°C, after which the pretreated chips are washed in hot circulating water from elsewhere in the plant. The softened warm chips are fed by a plug screw feeder into a pressurized preheater with relatively low pressure and temperature. After the preheater, the chips pass through a second plug screw feeder on their way into the first refiner, which operates at relatively high pressure and temperature. After first-stage refining, the pulp is driven to a steam separator, from which it is fed to the secondary refiner, which operates under approximately the same conditions as the first stage refiner. Then the pulp is again being separated from the steam in the second separator. After the pulp passes through the plug screw feeder at the bottom of the second separator, it falls down into a pulper for removal of latency. Screening and reject refining follow, then bleaching before storage and transport to the paper mill (Sundbolm, 1999). The paper mill produces two types of products: 48 $g.cm^2$ and 45$g.cm^2$ newsprint.

Business transformation: PLA production

The studied company is highly competitive newsprint (in the first quartile manufacturer) and has a limited access to the biomass. Hence, they have elected to go for a biorefinery strategy that integrates into their existing processes. Thus with the relatively low amounts of hemicellulose at hand and using VPP (value prior to pulping) process, the mill could develop a robust business model by manufacturing a higher added value product, such as PLA (polylactic acid). Therefore the potential business transformation was inspired by the combination of *VPP* and *Purac* technologies, e.g. the well-established cellulosic technology from *Biopulping International* (BPI) to extract organic chains from pulp is assumed to be integrated with Purac technology to produce a specialty product, polylactic acid (PLA).

It is assumed that a mass corresponding to 3% of the incoming dry chips by weight is extracted before pulping, while paper production is kept unchanged with respect to the current core business. This means that the yield of paper production drops to 95% from its current 98% value. The new process design is depicted by the simplified process flow diagram in Figure 3.1.2. The new process flow design can be summed up[17] in point form as follows:

- The thermomechanical pulping and newsprint production lines are unchanged;
- Before entering the TMP process, wood chips are subjected to an oxalic acid pretreatment in an impregnator, producing a solution of oligomers by the preferential extraction of hemicellulose and other wood constituents from chips;
- The extracted and separated sugars are then fed into a simultaneous saccharification and cofermentation (SSCF) unit. In this processing step, enzymatic hydrolysis of the extractives takes place together with fermentation. Acetic acid is also produced as a byproduct of the extraction process (The efficiency of conversion is greater than 95% on carbohydrate substrate, Datta et al., 1995)..
- The flow of lactic acid is then brought to the synthesis unit where etherification and catalytic synthesis of lactic acid into cyclic ester (*Lactide*) is carried out

- The granules of *Lactide* are then brought to the synthesis unit, where the process of catalytic and thermolytic ring-opening polymerization of lactide to polylactide takes

[17] For more details on the BPI and PURAC technologies, the reader should refer to Hunt, 2004 and Gruber & O'Brien, 2002 respectively

place.

- After the finishing and granulation steps, the final product, polylactic acid, is extracted and stored, ready for shipping to customers.
- As shown in Figure 3.1.2, the BPI-PURAG-TMP process model also incorporates the TMP process shown in Figure 3.1.1. However, with the introduction of VPP, some TMP operating parameters may require adjustment. For example, VPP chip pretreatment may result in reduced energy requirements for TMP refining.

Table 3.1.1.— List of assumptions

Unbleached [illegible] production	650 t/d
VPP	
Extraction (C5, C6, CX [illegible])	865 kg/h
Energy [illegible] consumption [illegible]	25%
PURAG	
PLA production volume	700.4 kg/hr
Fermentation efficiency	95%
Purification efficiency	98%
Assumed PLA market price	2.0 $/kg

It is also recognized that different wood species yield different amount of extractives and consequently different quantities of carbon chains from the impregnating unit. Current mill furnish consists of high- and low-density chips from black or white spruce respectively. Furthermore, a small amount of hardwood is added to the final furnish. These information must be carefully monitored since the chips ratio used can influence manufacturing costs significantly.

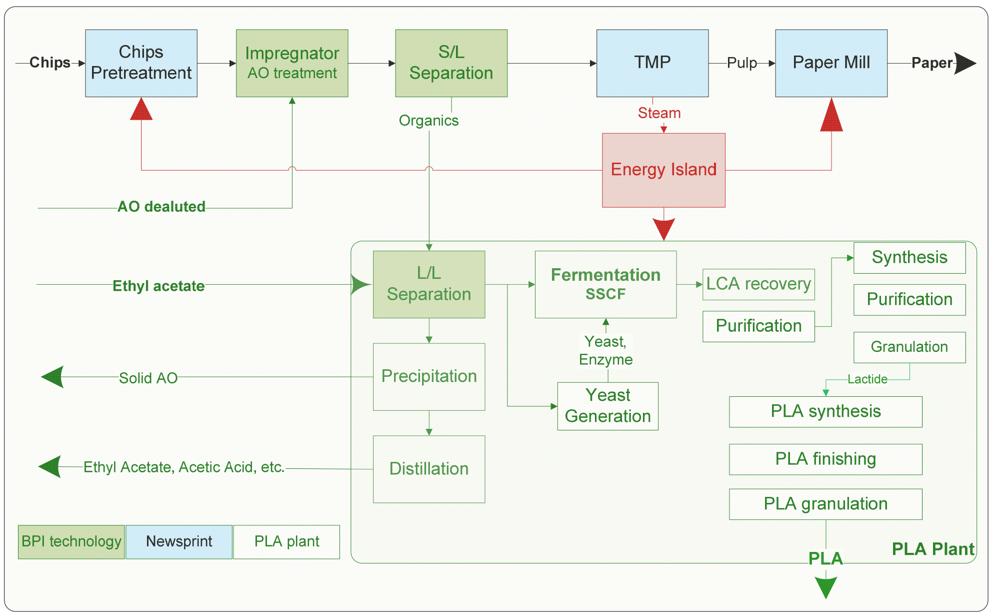


Figure 3.1.2: Block diagram for the potential business transformation of a current newsprints

3.2. Overall Methodology: from low-level data to smart decisions

The overall methodology consists of two parallel blocks of activities aiming to assess the production cost and profit margin of the same products and to compare the outcomes. The first block represents the use of traditional accounting practices, whereas the second block represents the multidisciplinary union of process and accounting tools which make up the ODCA method. Figure 3.2.1 presents the methodological steps that make up each block.

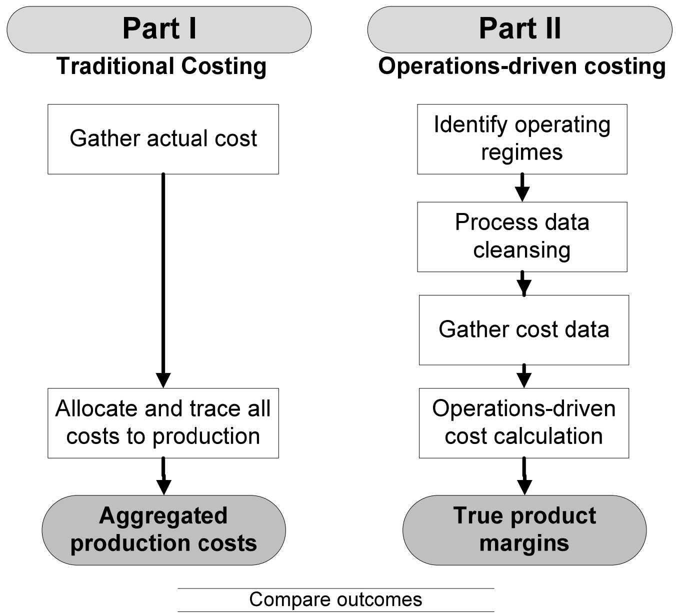


Figure 3.2.1: Overall methodology comparing traditional and operations-driven costing information

The first block: Traditional current practice

Traditional practices, as discussed to some extent throughout the first part of this chapter, follow a so-called top-down approach. One can simply state that the traditional cost of a product is an aggregation of liabilities and bills that the company receives with regard to the inventory state, divided by the total production within the period analyzed.

In this phase, actual costing will be used to represent the traditional way of assessing individual product margins. The following methodological steps were followed:

1. Aggregate all costs incurred during the manufacturing period to be analyzed (one month)
2. Gather information about the tonnage of production

3. Allocate all costs to the level of production.

The second block: Operations-driven cost modeling

Operations-driven cost modeling fully exploits low-level process data in a bottom-up fashion. Consistent plant-wide process data sets are integrated with business data in a model whose structure was inspired by an activity-based costing philosophy. The construction procedure for this model involves the identification of multiple parameters in the steps shown in Figure 3.2.2.

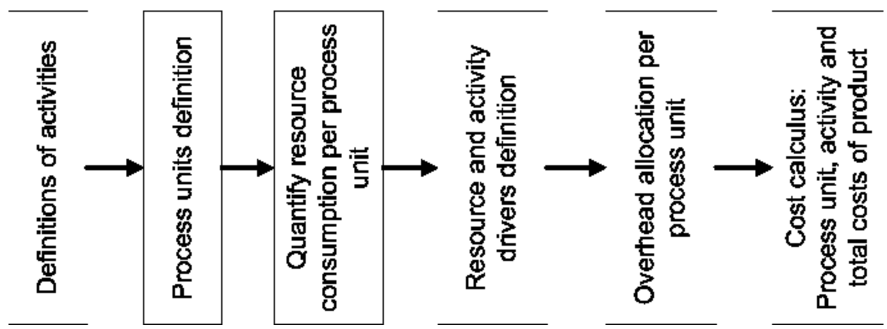


Figure 3.2.2: The ODCA modeling strategy.

The definition of activities involves the division of an operation into a final number of operating activities (cost centers), each of which is an aggregation of various processing units. Each of the cost centers is characterized by two key parameters: the design layout characteristics and the operating-state knowledge. The flow and combination of information is presented in Figure 3.2.3. The resource, activity, and cost drivers are defined by operating knowledge corresponding to resource consumption rate, activity performance, and individual cost-center contributions respectively. The integration of cost and process data is done within the cost center and is simply defined as a step-by-step cost-calculation recipe. Similarly, overhead costs are treated in a separate cost center which accurately traces them to different operating activities and to the final cost object. Finally, the sum of the cost activities represents total millwide operating costs.

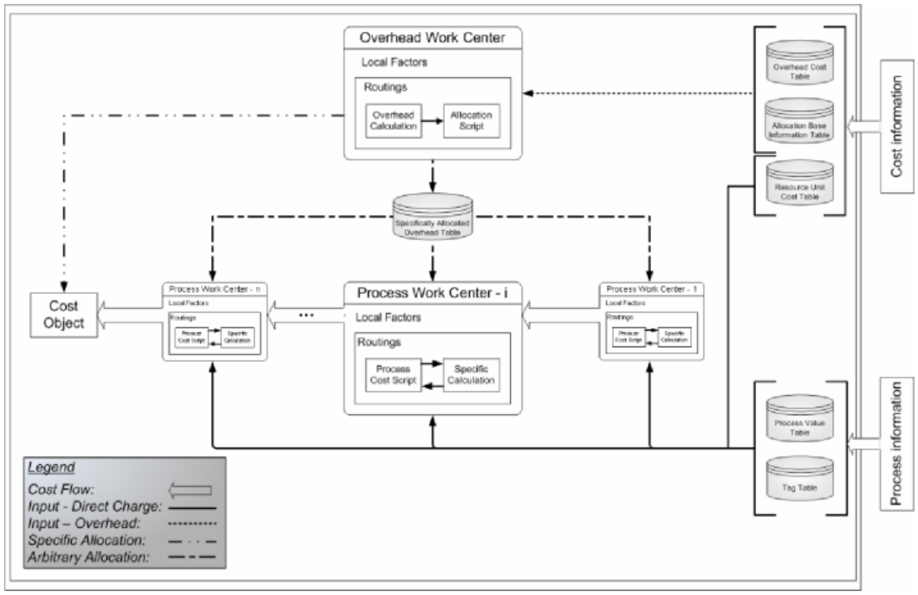


Figure 3.2.3: Work center structure within the operations-based cost model (from Laflamme-Mayer, 2011)

The following section describes the two main phases and their corresponding steps in the operations-driven activity costing block (Figure 3.2.4):

Plant-wide consistent manufacturing information

The second-block analysis starts by defining the set of operating regimes that are used to manufacture a given product. The criteria for distinguishing a regime are given by changes in process design characteristics (type/age of refining disk, opening/closing valves) as well as by variations in operating characteristics (different control setpoints or strategy, production rate). Within each defined operating regime, pseudo-steady-state data sets are identified for subsequent plantwide data reconciliation. This phase of the methodology provides profound understanding of the underlying manufacturing processes in the form of reconciled near-steady-state data sets (Korbel et al(a)).

Operations-driven production costs

The definition of relevant cost information for a given analysis is an important step to facilitate cost calculation and improve cost transparency[18]. The process data sets from previous phase are

[18] All cost items identified as irrelevant at this step were excluded from the first-block activities as well.

used to calculate the product cost simply as the sum of all production activities (cost centers) across the plant. Precision and validity of the production costs are ensured by running the same cost model with different process data sets[19]. The output of the analysis is a product cost distribution for a given operating regime.

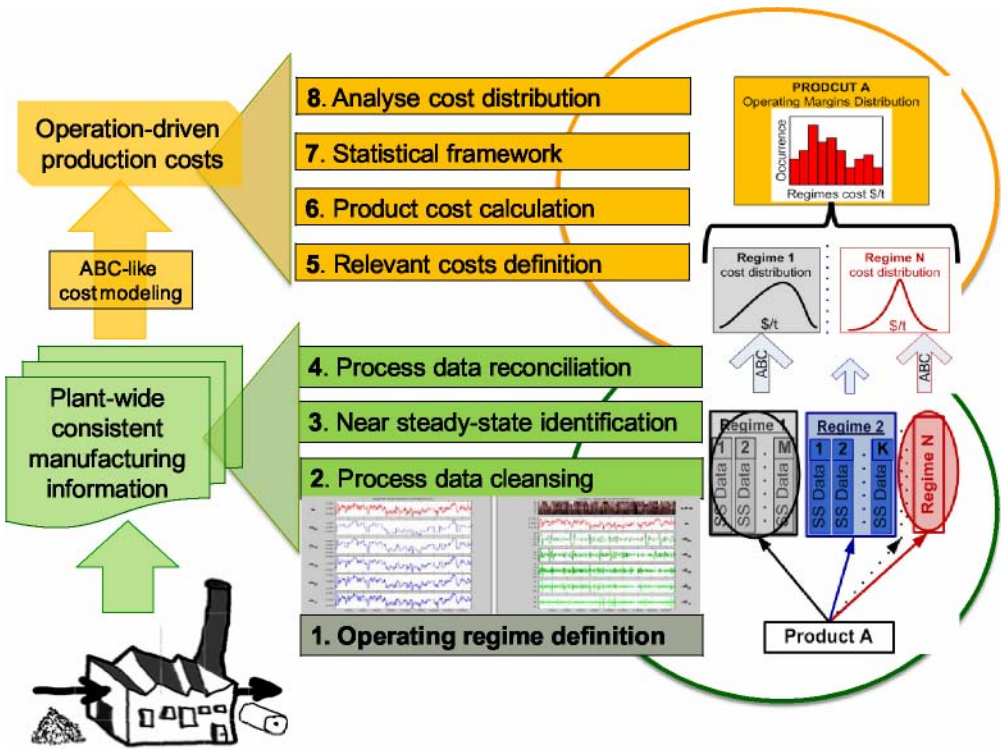


Figure 3.2.4: Step-by-step activities within the methodology to address individual product margins (Korbel et al (c)).

[19] These data sets were extracted during production of the same product within one operating regime.

3.3. Traditional costing and ODCM: economic results and discussion

The economic results of the case study are presented and discussed in two phases. In the first phase, the familiar cost picture based on traditional practices is compared to the operations-driven results for a current papermaking process. The second phase goes through a similar type of analysis, but with respect to a more complex production environment—the business transformation. The comparison of results, interpretation of the data, and potential improvement strategies are discussed at the end of each phase. With this in mind, let us proceed to the first phase.

3.3.1 Analysis of the current core business

The first bar in Figure 3.3.1 presents the monthly aggregated resource spending as calculated using traditional costing. This enables cost accountants to understand and differentiate the resource contributors to the total product cost. A predetermined (based on standard costing) volume-based parameter is often used to multiply this overall production costs to differentiate individual product profit margins (second and third bars). This operation ultimately distorts the results. Another approach may be to use standard costing, although the cost variance must be properly prorated to inventories or cost of goods sold to meet GAAP requirements. This *ad-hoc* approach to distinguishing actual costs for each product in a continuous-process industry is inaccurate and can lead to wrong decisions.

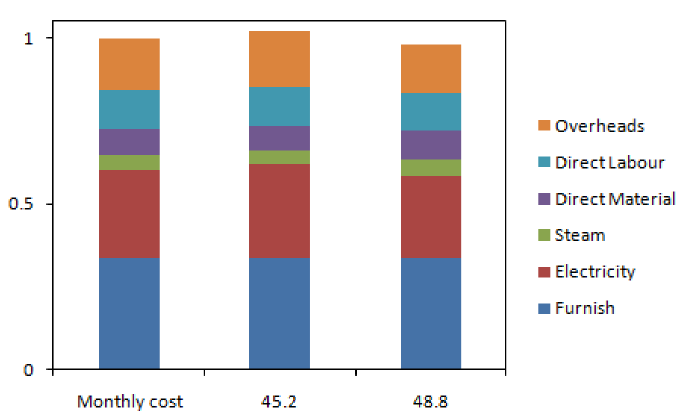


Figure 3.3.1: Traditional cost analysis results for a current core business.

The same accounting period was analyzed using operations-driven cost modeling, as shown in Figure 3.3.2. For the sake of simplicity, only three (the most common) operating regimes are presented per product. However, the number of different regimes that the mill operates in within a single month is significantly larger[20]. The outcomes of the cost model are normalized to the values acquired by traditional costing (the first bar in Figure 3.3.2). The cost of each product (second and third bars) is calculated as the weighted average[21] of a corresponding set of operating regimes (bars 4–9). The traditional costs represent a monthly aggregation of costs, while the operating-regime costs represent the actual running costs.

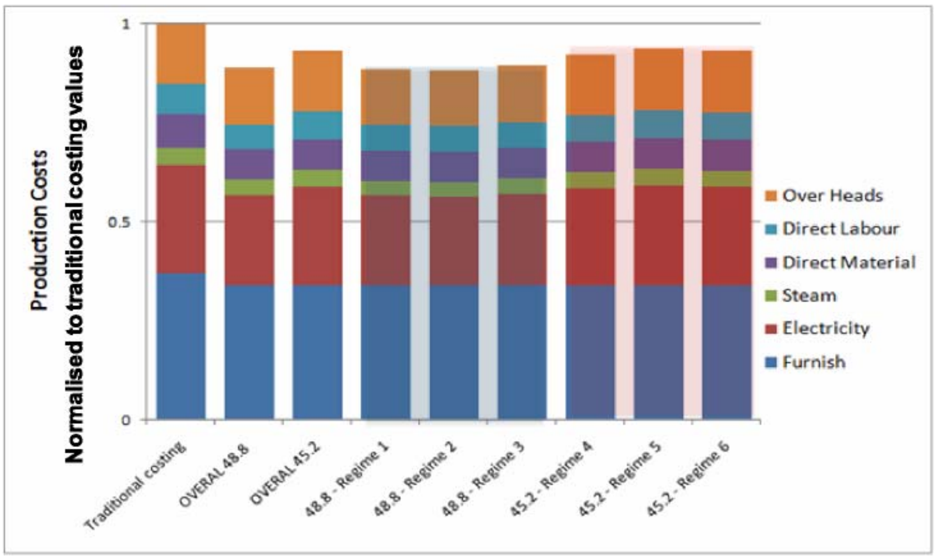


Figure 3.3.2: Operations-driven cost analysis of production activities (base-case scenario).

It is clear from Figures 3.3.2 and 3.3.3 that the knowledge acquired from operations-driven cost analysis can have a tremendous impact on mill improvement strategies. The information about each product's actual margin within each operating regime as well as over the complete set of operating regimes will guarantee that managers can make well-informed decisions. For interpretation purposes, one can simply drill down from product costs through regime costs to

---

[20] For this case study, twelve and seventeen operating regimes were identified for products of 48.8 and 42.5 $g.cm^{-2}$ respectively. However, the production time in most of these regimes was less than 1% of the total production time

[21] The weighting factor used is the ratio between the time that the mill operates in a given regime and the total operating hours.

the actual cost of manufacturing activities and their corresponding resource consumptions by individual processing units (Figure 3.3.3). This functionality helps understand and interpret the cost variances that arise from changes in resources consumption rates. The variance analysis that a cost accountant usually performs can now be well interpreted without the need to dig into data storage systems and search for answers in piles of data.

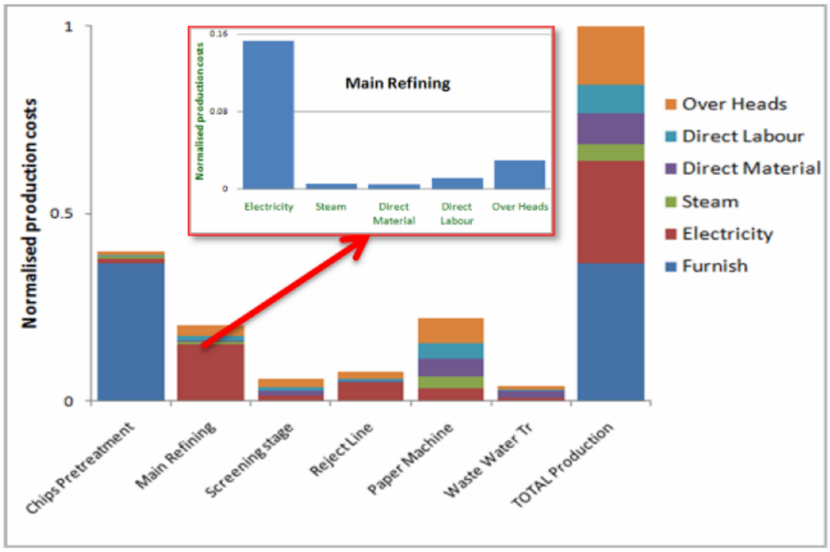
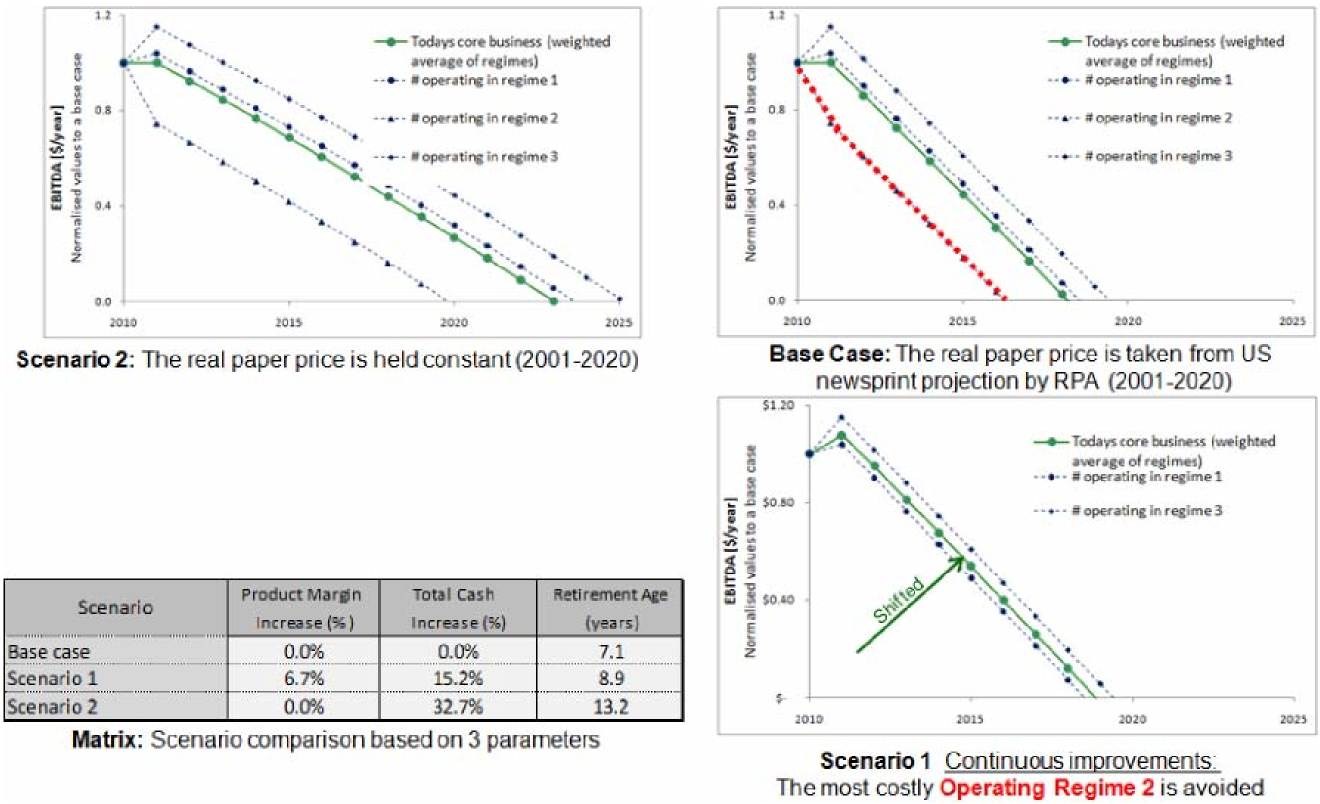


Figure 3.3.3: Operations-driven cost analysis of a current core business producing a 48.8 g.cm^{-2} grade in operating regime R3.

As has been briefly outlined in the context of the chapter, the apparent process improvement which can be gained from these results is to avoid operating in regimes R3 and R5 when producing grades 45.2 g.cm^{-2} and 48.8 g.cm^{-2} respectively. This possibility could be explored further by analyzing what would be the potential impact of this improvement on the whole business cash flow (Figure 3.3.4). However, it must be recognized that some operating regimes possibly cannot be bypassed due to inherent process-material characteristic interactions. Figure 3.3.4 presents a tree of possible future scenarios and their interpretations in the matrix. Each of these scenarios has been analyzed under the following assumptions:

- Constant amount of paper products sold to customers,
- Yearly increase in production efficiency due to operational improvements,
- Yearly labor and raw material cost increases,
- Selling price for scenario 1 and the base case is taken from predictions by RPA (2001–2020); the selling price is held constant for scenario 2.

The results show that, solely by avoiding mill operation in regime R3 (the process from base case to scenario 1), an increase of 15.2% in total cash within the analyzed business period may be achieved. Furthermore, the product margin of the 48.8 g.cm^{-2} grade is increased by nearly 7%, enabling newsprint production to continue more than 22 months longer than in the base case (under the assumptions listed above).



Scenario	Product Margin Increase (%)	Total Cash Increase (%)	Retirement Age (years)
Base case	0.0%	0.0%	7.1
Scenario 1	6.7%	15.2%	8.9
Scenario 2	0.0%	32.7%	13.2

Figure 3.3.4: EBID TA forecast and the potential impact of regime costing on product margin and total cash (Korbel and Stuart (c))

To provide a complete understanding of the complexity of cost modeling, Figure 3.3.5 shows the manufacturing information covering the whole set of operating regimes for the 48.8 $g.cm^{-2}$ product during the period analyzed. Each regime is labeled by its corresponding total production cost and its probability of occurrence. The width of the bar represents the cost range of the regime due to the use of multiple steady-state data sets for regime costing. The thick line inside each bar represents the weighted average of plantwide steady-state cost snapshots.

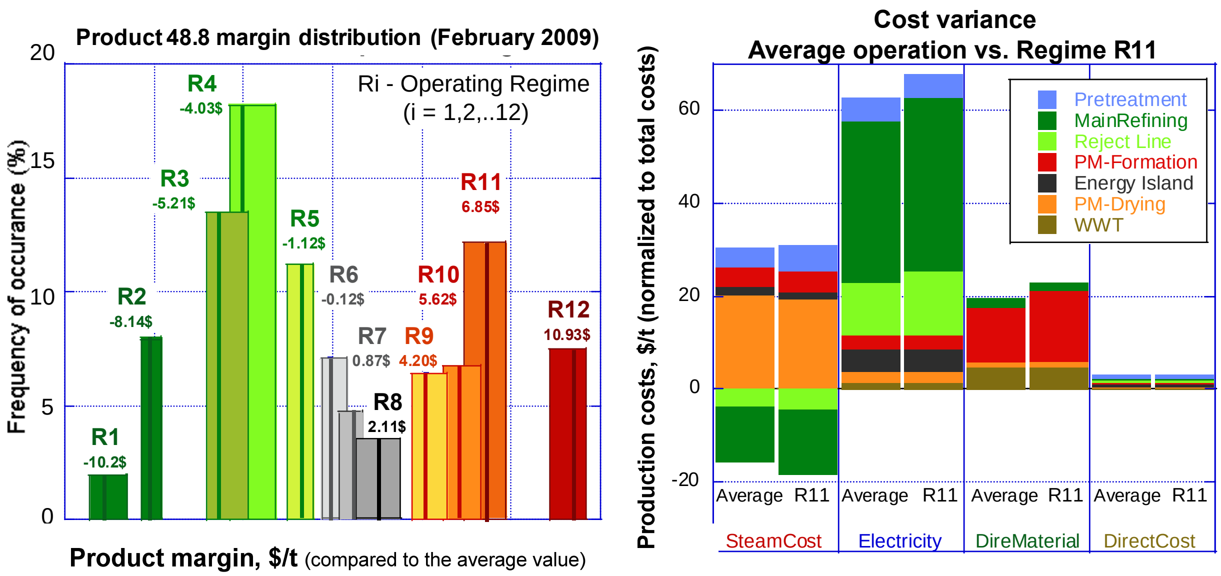


Figure 4.3.5: Wide range of product profit margins for a product 48.8 g.cm^{-2} within a month of operation (Korbel and Stuart (b))

4.3.2. Analysis of a potential business transformation: a multi product environment

The second phase of the case study involves assessing the individual profit margins and the potential changes in their values after retrofit design. The level of newsprint sales is assumed to stay constant. However, after the new business integration, the profit margin is significantly modified for each grade, as will be seen from the results. It is worth reminding the reader that by introducing the polylactic acid product into the business model, the simple production environment has shifted towards a more complex, simultaneous multiproduct environment. Traditional costing fails to assess actual product margins in a multiproduct industry.

An accountant using traditional costing tools would predict the cost of each product by using standard costing similar to the analysis of a current core business (Figure 3.3.1). This information has always been sufficient for providing investors with various financial criteria, such as ROCE or IRR. With these values, more complex economic analyses can be performed to predict the potential future business cash flow. However, the internal cause-and-effect relationships due to technology integration cannot be understood with traditional thinking about aggregated production costs.

An accountant using the ODCM toolbox would try to understand the cost and process impacts of the new business integration on existing core business products. With the use of current manufacturing knowledge from a process-cost perspective and a simulation of the new process operation, this impact can be analyzed efficiently. The results shown in Figure 3.3.7 are normalized to the traditional cost values for easy comparison. The first two bars represent the impact of business transformation by means of a traditional cost comparison of core newsprint products. It is clear that the main impact on production arises from electricity savings and overhead sharing with the new facility. The ODCA results offer a robust and complex cost analysis of future production with an understanding of cost-process impacts on core production during different operating regimes. The interpretation of the base-case retrofit variance can be understood by drilling down into the actual cost items and resource consumptions within operating activities. For instance, the production costs in operating regime R2 have been reduced due to significant electricity savings; however, steam costs were increased due to the reduced steam recuperation rate.

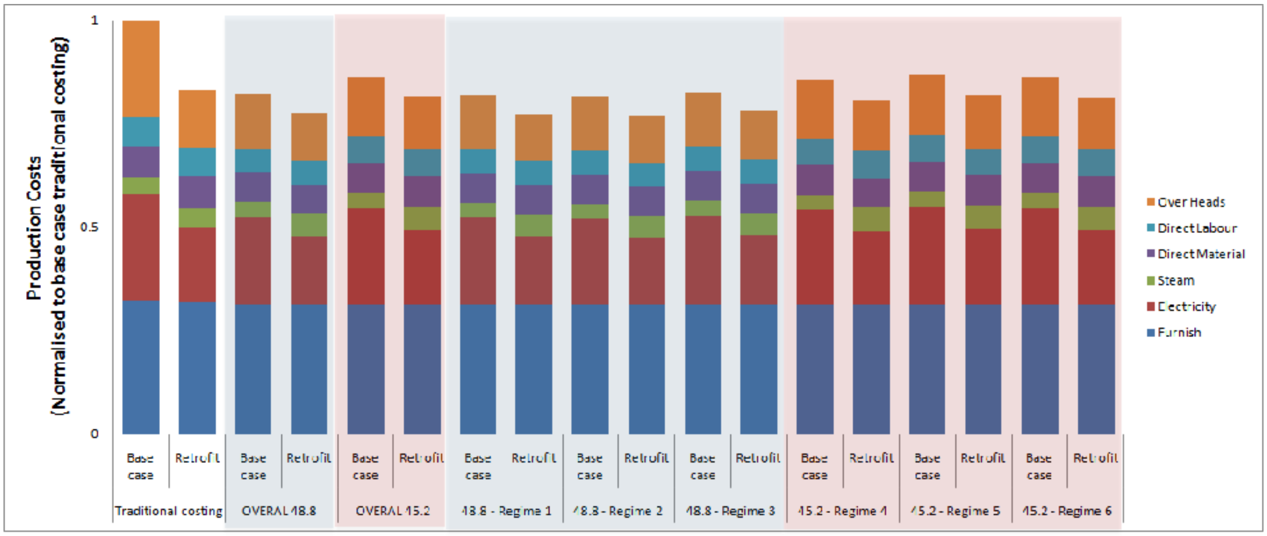


Figure 3.3.7: Operations-driven cost analysis of production activities: cost impact on a ***core business*** *after business transformation.*

Figure 3.3.8 illustrates the actual production costs of each product, including PLA, for different operating regimes. The cost of PLA can vary significantly from one operating regime to another. The energy (steam costs) and material usage are the items that to some extent cause the variance in production costs. After closer analysis, it became apparent that this variance is largely due to the more expensive unit steam price when manufacturing the 48.8 $g.cm^{-2}$ grade in operating regime R2 because of refiner plate-gap differences. Ultimately, this production state should be avoided in the future. Similarly to the first part of the case study, the EBITDA forecast can be used to assess different future scenarios and their impact on PLA production cost and total cash flow.

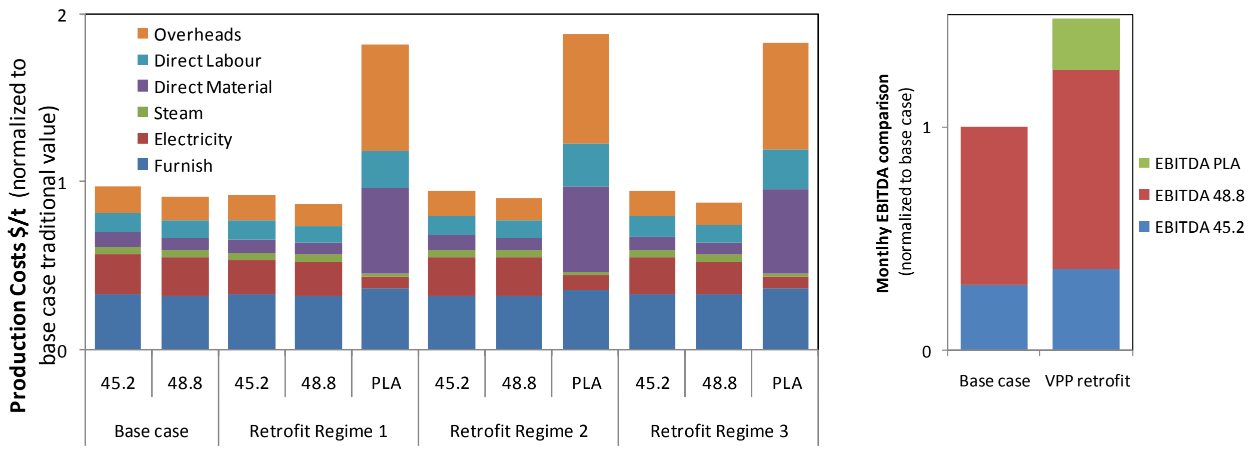
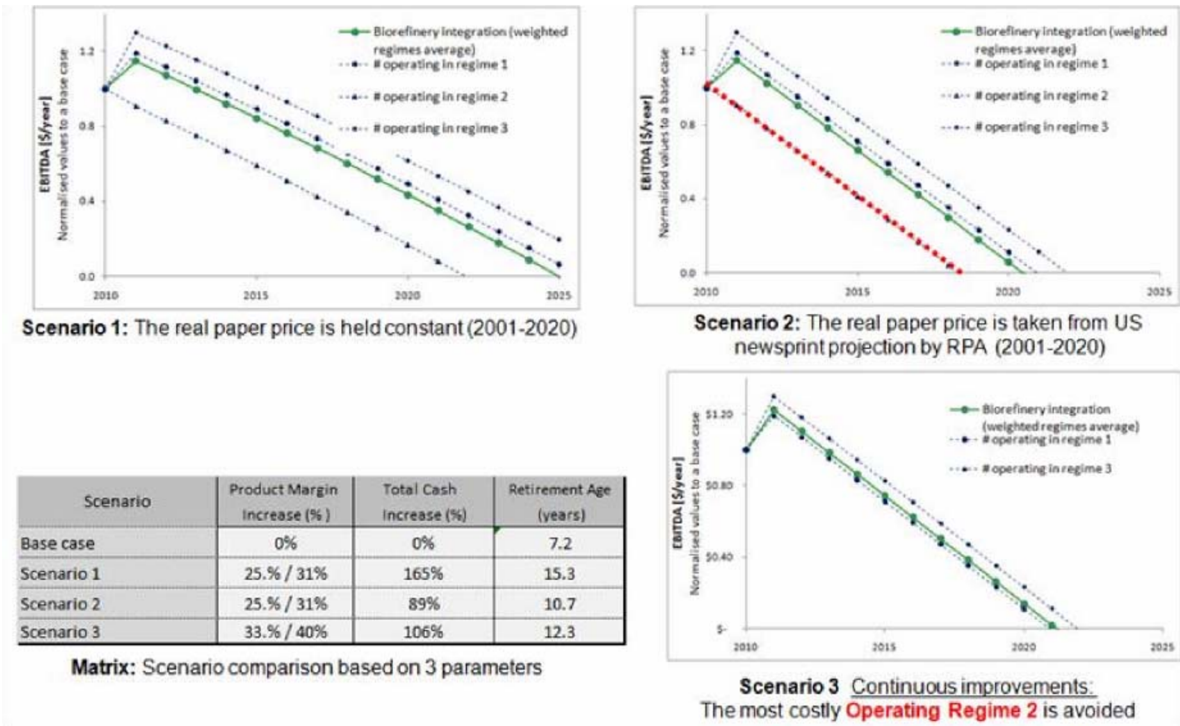


Figure 3.3.8: ODCM: The production cost of every product within a tree of operating regimes after business transformation and EBITDA comparison to a base case.

The assumptions for predicting EBITDA over time are identical to those used in Section 4.3.1, except that the price of newsprint is held constant in scenario 1 and is set equal to that predicted by the RPA (Figure 3.3.9) in scenarios 2 and 3. The assumption of constant newsprint price after biorefinery implementation shows a business cash flow increase of more than 165% compared to that of the current business. This will make the new business model break-even in more than 15.3 years from the present (8 years more than the base case).



Scenario	Product Margin Increase (%)	Total Cash Increase (%)	Retirement Age (years)
Base case	0%	0%	7.2
Scenario 1	25.% / 31%	165%	15.3
Scenario 2	25.% / 31%	89%	10.7
Scenario 3	33.% / 40%	106%	12.3

Figure 3.3.9: EBID TA forecast and the potential impact of regime costing on product margin, total cash, and retirement age of a business.

3.4. Concluding remarks

The work presented in this chapter seek to provide information to the reader about advances in cost accounting systems and the potential benefits of using new approaches instead of the current old-fashioned systems for decision making activities. Today, the forestry sector faces difficult times and requires a systematic approach to finding an optimal path towards a more sustainable future through potential business transformation. To manage this transformation optimally, managers and decision makers need to explore the powerful and robust cost accounting systems that are today waiting to be implemented in practice. Other industries, such as petrochemical companies or automobile manufacturers, are far more advanced in their costing systems. The use of the activity-based costing philosophy and its variations could help managers improve forestry company profits throughout the world. One particular approach that has been developed and designed to help the forestry sector is operations-driven cost modeling. The use of lower-level process data improves the understanding of manufacturing-cost variability due to different operating regimes for current and future products. The idea has been presented throughout the chapter that forestry companies should implement ODCA systems today to enhance their current cost-savings strategies as well as to identify and sustain the best operating scenarios in the future multiple-product environment.

References

Bagajewicz, M.J., (2001). *Process Plant Instrumentation and Upgrade*. Technomic: Lancaster, Pennsylvania.

Bahri, P.A., Bandoni, A., Romagnoli, J. (1996). Operability assessment in chemical plants. *Computers and Chemical Engineering* 20 (Suppl. B), S787.

Bartko, J.J. (1991). Measurement and reliability: statistical thinking considerations. *Schizophr Bull.* 17(3): 483–489 (1947873).

Brown, C.D. (1999). *Accounting and Reporting Practices of Nonprofit Organizations: Choices and Applications.* New York: American Institute of Certified Public Accountants.

Dansereau, L.P., El-Halwagi, M.M., Stuart, P. (2009). Sustainable supply chain planning for the forest biorefinery. In: *Design for Energy and the Environment: Seventh International Conference on the Foundation of Computer-Aided Process Design*, Breckenridge, Colorado, USA, p. 1101.

Drucker, P. (1995). Managing for business effectiveness. *Harvard Business Review* (May-June), 54–55.Emblemsvåg, J. and B. Bras (2000). Activity-Based Cost and Environmental Management: A Different Approach to the ISO 14000 Compliance. Kluwer Academic Publishers, Boston. pp. 317. ISBN: 0-7923-7247-6.

Friedl, G. and Kupper, H.-U. (2005). Relevance added: ABC with German cost accounting. *Strategic Finance,* (June), 56–61.

Fogelholm, J. (2000) 'Cost Function Modelling in the Paper Industries.' Report No.11. Helsinki University of Technology, Department of Industrial Engineering and Management. 2000. ISBN 951-22-5009-8

Fogelholm, J. (2004) 'Reporting and simulation of economic steering information', 27th Annual Congress of the European Accounting Association, Prague, Czech Republic.

Gale, R. (2006). Environmental costs at a Canadian paper mill: a case study of Environmental Management Accounting (EMA). *Journal of Cleaner Production* 14(14), 1237–1251.

Gruber, P., O'Brien, M. (2002) Polylactides "Natureworks® PLA". In: Doi, Y., Steinbüchel, A. (eds) Biopolymers- polyesters III applications and commercial products. Vol 4. Weinheim: Wiley-VCH p235-249. ISBN 3-527-30225-5

Garg, A., Ghosh, D., Hudick, J., and Nowacki, C. (2003). Roles and practices in management accounting today: results from the 2003 IMA – E&Y survey. *Strategic Finance* 1–6.

Grossmann, I.E., "Challenges in the New Millennium: Product Discovery and Design, Enterprise and Supply Chain Optimization, Global Life Cycle Assessment," Computers and Chemical Engineering, 29, 29-39 (2005).

Goldratt, E.M. and Cox, J. (1992). *The Goal: A Process of Ongoing Improvement*. North River Press; 2nd revised edition (1992).

Horngren, C.T., Foster, G., Datar, S.M., Teall, H.D., Gowing, M. (2006). *Cost Accounting: A Managerial Emphasis*, 4th Canadian edition.

Hunt, C., W. Kenealy, E. Horn, and C. Houtman. 2004. A biopulping mechanism: creation of acid groups on fiber. Holzforschung 58:434-439.

Jiang, T., Chen, B., He, X., Stuart, P.R. (2003). Application of a steady-state detection method based on wavelet transform. *Computers and Chemical Engineering* 27, 569–578.

Kaplan, R.S. and Atkinson, A.A. (1989). *Advanced Management Accounting*. Englewood Cliffs, NJ: Prentice-Hall.

Keys, D.E. and Van Der Merwe, A. (1999). Cost management: what insights does German cost management have for U.S. companies? *Management Accounting Quarterly* 1 (Fall): 1–8. ISSN 1092-8057.

Korbel, M., Wasik, L., Stuart, P.R., "Practical methodology for plant-wide process data rectification in the pulp and paper industry" Submitted to J-4 journal

Korbel, M. Stuart, P.R., (b)"Assessment and interpretation of advanced cost data for process improvement in an integrated newsprint mill", submitted to International Journal of Production Economics

Korbel, M. Stuart, P.R., (c) “Online methodology for operations-driven cost assessment of operating regimes using real-time process data”, submitted to TAPPI journal

Laflamme-Mayer, M., Janssen, M., Stuart, P.R, “Development of an operations-driven cost model for continuous processes” Part I: Framework for design and operations decision making, submitted to J-4

Levin, H.M. and McEwan, P.J. (2000). Cost-Effectiveness Analysis: Methods and Applications (2nd editon) Thousand Oaks, CA: Sage.

Mansoornejad, B., et al. (2010). Integrating product portfolio design and supply chain design for the forest biorefinery. *Computers and Chemical Engineering* doi:10.1016/j .compchemeng.201 0.02.004.

Mueller, D.C., ed. (1990). *The Dynamics of Company Profits: An International Comparison*. New York: Cambridge University Press.

National Association of Accountants (1985). *Statements on Management Accounting (4B): Allocation of Service and Administrative Costs*. Montvale, NJ: National Association of Accountants.

Robinson, H. (1997). Using poka-yoke techniques for early defect detection. http://facultyweb.berry.edu/jgrout/pokasoft.html. Retrieved May 4, 2009.

Rother, M. and Shook, J. (1999). Learning to See: Value-Stream Mapping to Create Value and Eliminate Muda. Lean Enterprise Institute, 1999. ISBN 0966784308.

Sammons Jr. N.E., Yuan W., Eden M.R., Aksoy B., Cullinan H.T. (2008): “Optimal Biorefinery Product Allocation by Combining Process and Economic Modeling”, Chemical Engineering Research and Design, 86(7), pp. 800-808.

Sundholm, Jan. (1999).“Mechanicul Pulping”, Tappi. ISBN 952-5216-05-5 (book 5) published by Fapet Oy. Printed by Gummerus Printing, Jyväskylä, Finland 1999.

U.S. Census Bureau (1998). *Quarterly Financial Report for Manufacturing, Mining, and Trade Corporations*. Series QFR-98-3. Washington: GPO. Available from www.census.gov.

Willson, J.D. and Colford, J.P. (1990). *Controllership: The Work of the Managerial Accountant*. New York: Wiley.

APPENDIX G –

PROCESS AND BUSINESS DATA RECONCILIATION IN THE PULP AND PAPER INDUSTRY

Conference paper, TAPPI Conference, Innovations in Engineering, Pulping & Environmental, Jacksonville, FL, 2007

Process and Business Data Reconciliation in the Pulp and Paper Industry

Milan Korbel[1], Laurent Garrigues[2], and Paul Stuart[1]

1-NSERC Environmental Design Engineering Chair, Department of Chemical Engineering, Ecole Polytechnique – Montreal

2-OSIsoft Canada ULC, 1155 University Street, Montreal

ABSTRACT

Knowledge-based data processing in the pulp and paper industry will become increasingly critical in the coming years for mills to remain competitive. For example, mills will make process change decisions (design and/or operations) based on plant-wide optimization, and real-time optimization will become increasingly common. In order to successfully implement such advanced knowledge-based decision making systems, it is essential to have reliable process and business data.

Pulp and paper processes have a highly dynamic character but despite this, the representation of pulp & paper processes can be based on steady-state conditions and this information used in optimization models. Obtaining a reliable steady-state process representation is not straightforward. Measured data must be processed in order to eliminate signal noise so that steady-state identification can be performed. Data reconciliation techniques can then be used to detect gross errors, eliminate measurement bias, and reduce random errors associated with raw measurements by imposing mass and energy balance constraints to satisfy conservation laws. This study describes techniques that have been reported to make plant-wide process data more reliable, and the use of this data for performing plant-wide real-time optimization through coupling with business models. Linking reconciled process data with business data as an input into a process-based business modeling framework provides mills with a valuable tool for operating decision support. The paper reviews how these tools might be applied to optimize mill operations.

INTRODUCTION

Increased global competition as well as continual increase in energy and raw material costs is currently forcing pulp and paper sector to continuously improve operation efficiencies, define new business and technology strategies, while keeping producing high quality products. One set of possible applications that deliver on these needs is the development of techniques and methodologies to obtain business information captured in the process and cost data tagged and stored by the Plant Information Management Systems (PIMS).

It has been few decades since various processing plants have integrated their plant information management systems (PIMS) into operation and they are today commonly used on a daily basis by different users in the enterprise. PIMS are both information management system and hierarchical database supporting high transaction volume and real time processing capabilities. Oil and gas refineries were among the first to adopt and take advantage of PIMS by developing various methodologies and applications to gain a profound understanding and knowledge on their daily operation. The use of PIMS extended rapidly to various manufacturing branches, e.g. chemical and food industry, metallurgical and mining facilities, pharmaceutical production and others. Pulp and paper industry followed the trend and nowadays most of the pulp & paper mills in North America leverage the capabilities of these information management systems providing mill personnel with critical information of the whole plant complex.

The use of real-time plant-wide data has found its benefit in many fields as it simplifies and improves data accessibility, therefore opening new opportunities to improve process operation, to detect process problems faster as well as to identify more profitable process conditions. It has been adopted by the pulp and paper industry for relatively long time and yet its potential has not yet been fully exploited. Good examples are the integration real time process data for business process analysis or supply chain management which are both overlooked or even unknown to pulp and paper mills managers [1] [2].

Focus should be given not only on the basic continuous improvements in production and maintenance, but as well on in-depth process optimization approaches. For example, the development of methodologies such as business process modeling based on the coupling of process and business data could provide an eligible and flexible tool to decision makers for a real time decision making. Moreover there is an opportunity to use PIMS by developing inter-organizational systems that would improve the communication in the supply chain [3].

Since efficiency of such process-driven applications strongly depends on data quality as well as data availability, there is a need to ensure a steady supply of reliable and accurate data into the business models. However raw process data often carry different errors and inconsistencies (instrument miscalibration or malfunction, power supply fluctuation, process noise, etc.), that limits usability. Numerous data cleansing techniques are available to improve data quality. Jiang [4] [5] and [6] developed a robust data processing technique consisting of a 4-steps methodology (detection of abnormalities, data filtration, steady-state detection, and data reconciliation) and adapted it to on-line steady-state identification using wavelet transforms. Bellec [6] [7] further improved the technique using statistical theories and showed promising results in increasing the data accuracy. This method is also useful for process state estimation and subsequently resolving and reconciling inconsistencies among different data sources. Since plant-wide steady-state operating condition detection is very challenging, the use of this technique is still limited to specific sections of the plant, where different steady-state operating regimes can be identified and reconciled.

Due to the highly dynamic character of the pulp & paper processes and the numerous interactions between its manufacturing processes, it is a real challenge to improve the operation performance in a real-time. Therefore most of the representation of pulp & paper processes usually assumes steady-state conditions. These conditions are used in simulations and optimization models. Today, and to comply with such an assumption, data inputs into reconciliation systems for plant-wide applications (real-time optimization), are in the form of average values over certain time period, for example 24 hours span to generate daily accounting reports, hence no real-time insight into cost spending by production facilities.

This paper aims to introduce the idea behind on-line plant-wide process and cost data reconciliation and how it might be applied to optimize in real-time mill operations.

BACKGROUND

Pre-processing and Reconciliation of Data

One of the most crucial challenges in processing plants today is to deliver meaningful plant-wide material and energy balances being in agreement with instrumentation values. This is highly due to various errors present in measurement values including random errors and inconsistencies caused by diverse irregular events. In order to obtain decent steady-state data, measured data need to be processed before steady-state identification can be performed. There are various techniques available for data processing and steady-state detection. Most of them are based on recurrent statistical and regression analysis over a predefined time window which is not always appropriate for on-line use.

A new wavelet-based process steady state detection method [4] elaborated for on-line application [5] has been used for real-time data denoising and steady-state identification in order to resolve more accurate discrepancies among data by reconciliation techniques. The core of this technique is based on wavelet multi-scale analysis and data filtering.

In this proposed methodology wavelet denoising and trend analysis is implemented as a first step before reconciling identified steady-state dataset. Furthermore this data pre-processing not only enhances data reconciliation performance, but ensures that reconciliation itself can be carried out independently of the data sampling time. A commercial product, Sigmafine from Osisoft was used for data validation and reconciliation.

Bottom-up cost accounting

Process-driven decision making can be improved by making use of process integration techniques. Process integration emphasizes the functional relation between individual process units and the enterprise as a whole. Mills adopting advanced methodologies such as bottom-up operations-driven business models based on integrating process and accounting data would gain insight into their daily cost spending. With this new view on the process, mills would be capable of identifying new opportunities to make improvements in their operations.

The process-based business framework [4] integrates process and cost information. The advantage of linking these different data sources lies in improving data consistency between plant floor operations and accounting therefore improving the costing system itself.

Usually cost management systems perform cost allocation from aggregated level to detailed level which can easily misrepresent the actual cost figures. In contrast, the scope of this methodology is based on a so-called bottom-up approach, where the relationships between detailed-level process data from the plant floor and accounting data are modeled. The information management system supplies the real time process data. Likewise, general ledger or other accounting systems are the source for cost information. This information based on in-depth resource consumption is then used for the cost calculation. Furthermore the manner in this approach for tracking resource consumption and cost allocation employs multi-level activity based costing system with the functionality that closely coupled process cost centers are connected together to form fluid business process analysis.

Industry application

Today refineries have implemented the reconciliation systems to be performed before further process data use, for example: yield and performance evaluation, planning and scheduling, process optimization and advanced process control. The execution of data reconciliation is usually tailored for daily, hourly or even minute time period with the use of updated averaged values from the previous time window. However, the use of such packages is limited to stable processing plants with high degree of redundancy.

The use of such systems in pulp and paper industry is very limited due to not enough redundant measurements present as well as its high process dynamic nature. The work of Jacob and Paris [8] [9] deals with data reconciliation and highlight data collection challenges in pulp and paper mills. Lately Stettler [10] applied the reconciliation techniques to sulfite wood pulping process for energy utilization with subsequent techno-economic evaluation. Where the use of utility system data reconciliation with the process models decreases the need of plant measurements.

Plant-wide real-time data reconciliation and cost analysis have not yet been implemented by the pulp and paper mills. If data from PIMS should be used for further applications in pulp and paper industry, there is a high demand on developing new techniques in order to ensure the quality of data.

OBJECTIVES

The principal goal of this study is to develop a practical methodology for making available on-line reconciled process and business data in a form suitable for advanced decision making. In reaching for this goal, the following challenges need to be overcome:

- *Process data analysis level*: need to define a methodology integrating on-line process trend wavelet analysis method with data reconciliation techniques for real-time plant wide process data validation / estimation,
- *Cost data quality:* Define strategy for business data reconciliation in order to ensure consistency in cost data among different cost sources.
- *Business data analysis level*: Process-based business framework is used to evaluate actual costs of reconciled process work centers (snapshots of the mill). The challenge is to know how to use the output information to real-time plant-wide process optimization,

PROCESS DATA TREATMENT

Before on-line process measurements can be used for further analysis, they have to be treated by diverse techniques in order to improve their accuracy. In this study wavelets are used in both denoising of measured process data and in extracting the refined process trend. The identified steady-state data set is then used by reconciliation procedure to make this data set corresponding to its process model.

<u>*Process trend extraction using wavelet analysis*</u>

The recent strike of continuous wavelet transform (WT) into time series analysis has revealed its benefit also for process trend estimation in industry. WT decomposes a one-dimensional time series into two-dimensional time-frequency space, in other words, the observable process signal *f*(t) (defined by measurement values) comprises two unobservable components, the desired process trend *T*(t) and process noise *N*(t) (stochastic component).

$$f(\mathrm{t}) = \mathrm{T(t)} + \mathrm{N(t)}$$

The wavelet-based multi-scale analysis method developed by Jiang [4] is very efficient for abnormality detection and denoising the real-time process data thanks to dividing the real-time process signal into diverse frequency components at different scales, e.g. it uses the fundamental idea to represent the series of measurements as limit successive approaches at different frequencies (Jiang et al 2000):

$$f(\mathrm{t}) = f(0) = \sum_{i\in I_0} c_{0,i}\varphi_{0,i} = \sum_{i\in I_1} c_{1,i}\varphi_{1,i} + \sum_{i\in k_1} c_{1,i}\psi_{1,i} = \dots = \sum_{i\in I_j} c_{j,i}\varphi_{j,i} + \sum_{j=1}^{J}\sum_{k\in k_1} d_{j,k}\psi_{j,k},$$

Where $\sum_{i\in I_j} c_{j,i}\varphi_{j,i}$ is the smoothed signal representing the low frequency part of the original signal with so-called mother coefficients $c_{j,i}$ and

$\sum_{j=1}^{J}\sum_{k\in k_1} d_{j,k}\psi_{j,k}$ is the detail signal representing the high frequency components with father coefficients $d_{j,i}$. Individual isolated components are then analyzed and modified by altering its coefficients $c_{j,i}$, $d_{j,i}$ to $c'_{j,i}$, $d'_{j,I}$. By thresholding, the coefficients below given threshold value are removed. The validated ones then serve for process trend reconstruction using the inverse wavelet transformation. The figure 1 shows an example of trend decomposition by this technique. For more information refer to Jiang [4] or Mallat and Hwang [11].

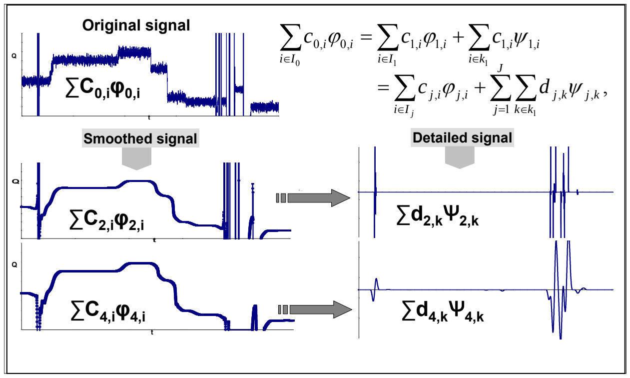


Figure 1: Multi-scale decomposition of real-time measurements

On-line application

On-line denoising and steady-state detection technique presented in the study by Bellec [6] taking the advantages of 1st and 2nd WT features, consists of 4 simultaneous segments as follows:

- Using WT features, abnormalities and high frequency noise are removed,
- The starting point of steady-state is detected using the WT properties and its first derivative,
- Steady-state duration is estimated by law pass filter based on historical data analysis for probability function,
- The end point or a drift from the steady-state period is determined by WT analysis as well.

Since all the segments of the method are performed simultaneously (the steps 1, 2 and 4 use the same features), there is no need for extensive computation which demonstrates the robustness and on-line application of proposed methodology.

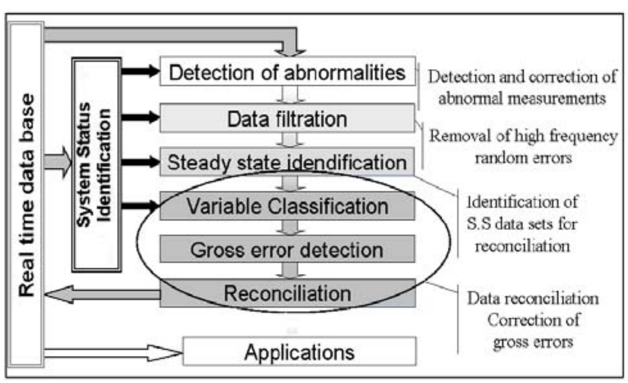


Figure 2: General methodology for on-line data correction (from Bellec [7])

De-noised extracted trends are expectedly more accurate than the measurement values. However their values might not be in agreement with process model constraints. Data reconciliation techniques are used to resolve such discrepancies.

Data Reconciliation

Data reconciliation is a well-established subject not only in the data processing field. By employing redundancy of process data, it is simply an approach to adjust values of process measurements to their process constraints. In other words, data reconciliation technique assumes that process variables are linked through a connectivity model, usually through material and energy balances. To formulate the data reconciliation approach one can write the problem as follows

$$Min\sum_{i=1}^{K}\left(\frac{measured/unmeasured_i - reconciled_i}{tolerance_i}\right)^2$$

Subject to: Mass, Heat, Component Balances

With a normal distribution of measurement errors being assumed diverse methods were published to simplify the issue above in order to eliminate unmeasured process values from the problem statement: Projection matrix method [12] Gauss-Jordan elimination procedure [13], QR decomposition [14], nonlinear DR using successive linearization [15] [16] [17], and many others. Obviously numerous studies have been done on linear or non-linear steady-state DR, but there is far less dealing with dynamic data reconciliation. This is mostly due to high computational effort which is thought to be impractical for engineering practice. However, with the advance in the information technology and given that processes are actually never at steady-state, it is better to consider applying dynamic data reconciliation even for “steady-state” processes [18]. On the other hand, from the perspective of minimizing the computational effort for on-line applications it is wise to extend steady-state DR to deal with dynamic situations [19]. In fact as mentioned by Bagajewicz [20] for the time being there are more pressing problems to resolve, for example gross error detection which is in close relation with the problem of data reconciliation.

Gross error detection

Data reconciliation techniques are based on the hypothesis that data are corrupted only with random errors excluding any potential systematic errors present. If the reconciliation is carried out on the measurement data set where gross error is present, the results will lead to inaccurate value estimations. The source of such systematic errors may be due to instrument miscalibration, drifting, biases or total failure of the instrument. Additionally, inconsistencies related to inaccurate formulation of process models can produce similar outcome as wrong instrument performance.

There are generally two principal issues linked to gross error (GE) handling, e.g. the gross error detection and its value estimation. Furthermore in any proper GE detection technique, the four following points should always be addressed:

- *GE detection* – possibility to identify the existence of one and/or multiple GEs in the measurements,
- *GE location* – possibility to locate one or multiple gross errors.
- *GE identification* – possibility to determine the GE type,
- *GE estimation* – possibility to estimate the magnitude of GE.

The gross error handling problem is well documented in the literature. Many methodologies and techniques have been proposed in order to satisfy the four above requirements. For more background on GE refer for example to [18] [20].

With regard to the current commercial software available for gross errors handling, the main technique used today is serial elimination strategy. As mentioned in [20] the vendors could improve their strategies by, for example, implementing methodologies to handle uncertainties in order to enhance industrial gross errors treatment.

On-line Plant-wide Data Reconciliation

Although there is a vast amount of publications on data reconciliation, there are just a few that focus on the large plant-wide applications and none of them, in our knowledge, for pulp and paper industry. All of the commercial softwares available today are based on steady-state DR, namely Sigmafine from OsiSoft, Adviser from Aspentech, Datacon from Simsci, and Recon from ChemPlant. The use of these software packages is therefore limited to stable plants (in plant-wide application).

In order to perform plant-wide steady-state process data reconciliation, available reconciliation systems use averaged values of process measurements over a certain time period in order to satisfy their “steady-state” assumption. Since the averaged process data carry not only random errors, but also errors from process variations, the results may lead to incorrect conclusions. However, the study by Bagajewicz [20] proved that in systems with no hold-ups this error can be negligible. Furthermore, by making use of averaged values one cannot really perform data reconciliation in real time with high precision. Hence, the applications of wavelet transform features for on-line data denoising and steady-state identification in combination with data reconciliation techniques could reform the accuracy in real-time process data.

Challenges in data reconciliation

Mostly DR studies have been focused on the specific particular unit or on a small process subsystem in a steady state or a dynamic way. Although there are several that deal with plant-wide data reconciliation and optimization problems, they are just under quasi-steady-state hypothesis.

There are three rational possible ways to put plant-wide steady-state data reconciliation into practice:

- Reconcile averaged process measurement values,

- Identify plant-wide steady-state operating regime to perform reconciliation,
- Identify steady-state regimes on a smaller segments of the processing plant,

The first option using averaged values is limited to linear systems with no hold-ups. Furthermore, the execution can be performed only in specific time periods. The second option is obviously inaccessible for pulp and paper industry due to its high dynamic nature. The last alternative is being investigated and examined for its application and suitability for pulp and paper mills. We seek to define methodology starting with TMP integrated mill.

In our knowledge, there is virtually no existing study that would address hierarchical decomposition of a process plant into different subsystems with on-line process state identification, and that would deal with the challenge of defining the potential "bridge" between separate reconciled subsystems to ensure plant-wide consistency in measured process values.

COST ACCOUNTING SYSTEMS

Business managers, to make the best possible decisions for their mill, need accurate information which is provided by management accounting systems. In present complex business environment cost accounting plays a key role in making business decisions focusing on long term profitability.

Both financial and cost accounting is part of the financial management systems in the enterprise. Whereas financial accounting reports organization's financial statements to investors, regulators, suppliers, banks and other outside parties, the cost accounting determines and reports financial and other types of information relating to the cost of resources to help managers in meeting their goals. Furthermore, financial accounting focuses only on the past-oriented performance reports, while the future-oriented character of cost accounting plays an important role in planned continuous reduction of costs.

It is quite appropriate here to define some terms that will be used:

- *Cost object* can be an activity, a customer, part of the process, a product, a service or anything for which management needs a separate measure and accumulation of costs,
- *Manufacturing costs* can be assigned to the cost object either direct or indirect way,
    - *Direct manufacturing costs* are directly linked to the cost object and are traced in a cost-effective manner, for example using mass and energy balances,
    - *Indirect manufacturing costs* are not easily traced in a cost-effective manner due to their indirect relationship to the cost object. They have to be allocated to the cost object using cost allocation method specifying some formula-relationship, also called basis of allocation (overhead costs),
- *Non-manufacturing costs* have the weakest relationship with a cost object and have to be assigned arbitrary (overhead costs),
- *Basis of allocation* methodically links non-manufacturing costs, an indirect costs or indirect cost pools to the same cost object. It can have a financial character, like the direct manufacturing labour costs or a nonfinancial one like the kilograms of chips used,
- *Cost pool* is an accumulation of costs with one or more mutual characteristics. Costs from cost pools are assigned to different cost objects or even cost pool as a whole can be assigned to the cost object.

Traditional Cost Accounting Practice

Today, despite the known disadvantages, many companies are still using traditional cost accounting systems (TCA) in their business organization. Actually, it started when cost accountants in order to allow the indirect cost to be accounted, were arbitrarily adding a rough percentage into the direct costs. However, when the proportion of overhead costs had grown (mostly due to industry automation), this method became improper. Therefore one of the most highlighted disadvantages of traditional cost accounting system is that it uses only one allocation base. In this way all overhead costs are allocated in the same way to all cost objects despite the potential variance in its resource consumption. Furthermore, administration, financing or marketing costs are summed up in cost pools and assigned to a specific time period and not allocated or assigned to any cost object.

There are several approaches of TCA. The two most used types of traditional costing systems are direct/variable and full absorption costing systems. Where indirect manufacturing costs are allocated to cost object on the basis of production volume related measurements. In other words, in TCA systems, direct material and direct labour costs are directly charged to the cost object, whereas overheads are handled as indirect costs and are allocated usually with a single, plant-wide, predetermined overhead driver (direct labour). Since the direct labour does not drive the production costs anymore, it is quite inadequate to employ TCA in modern companies with the intention to accurately trace the factual cost of the cost object. Consequently, managers are often making decisions based on inaccurate information. Despite the above facts, TCA is still used and quite appropriate to perform costing in firms, where the manual labour significantly dictates final product cost.

In order to ascertain the true value of the cost object, there was a need to define a new cost management system. Activity based costing emerged in manufacturing sector of the United States in the late 70s, in order to improve enterprise costing systems by tracing costs more effectively.

Activity-Based Costing

Activity based costing (ABC) is relatively new philosophy for cost accounting introducing activities as a link between resources and a cost objects. It clarifies more accurate the costs incurred in the organization by focusing on a single activity as the fundamental cost object. Then using this activity cost as a basis to assign costs to other cost objects, such as products and customers. In doing so, ABC systems intent to assort indirect costs and indirect pools as direct ones which gives more in-depth understanding on the costs incurred. In other words costs are traced to activities and these costs, in a second phase, are traced to the products that use these activities as it is shown in the figure 3.

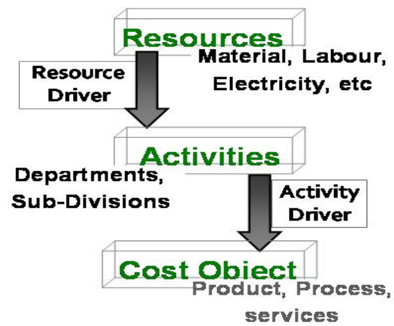


Figure 3: Resources consumed by Activities consumed by Cost Object.

Another advantage of ABC over TCA methods is that it attempts to assign all costs to cost object including marketing, engineering, administrative costs, etc. To do such cost tracking, ABC introduces assignment of costs additionally to traditional cost allocation and direct cost attribution options. This cost assignment ability allows accountants to rationally track overheads incurred (traditionally arbitrary allocated) as nearly as direct costs. This is done by making use of so called drivers. As it is shown in figure 3, resources are linked to activities by resource drivers and similarly activities are linked to cost object by use of cost drivers. Per definition, resource drivers determine the amount of a resource consumed by each activity, while activity drivers specify how different cost objects (product, customer) consume these activity costs. Labour hours, kWh and number of shipments is an example of resource driver, whereas number of customers and number of products is an example of the second stage driver or activity driver. For the sake of simplicity, it is always important to determine the right amount of appropriate drivers that would meet the accounting objective.

The key is to define appropriate cost pools and drivers for different types of cost. Before implementing ABC method in the enterprise, both drivers are often specified by a survey of managers or other mill personnel who have subject matter expertise. With a smart driver selection, accountants are now able to better track overhead costs. The result is elaborated and consistent cost allocation. Besides, ABC method gives appropriate information on resource consumption to decision makers in order to reduce more efficiently their operational cots.

As shown in Figure 4, the process oriented cost assignment using ABC is performed in two stages, while structure-oriented TCA in one. Furthermore, today advanced ABC have evolved into multi-stage systems where individual activities can be used by another activities before being used by final cost object enhancing even more the accuracy of cost modeling.

The advantages of ABC philosophy found its use also in the real time cost monitoring application proposed by Steen [21]. This cost model uses bottom-up concept for consumption of resources with the feature to track and analyze the cost variances by attributing production cost to production and quality. This method was applied to paper manufacturing example with the emphasis on data quality.

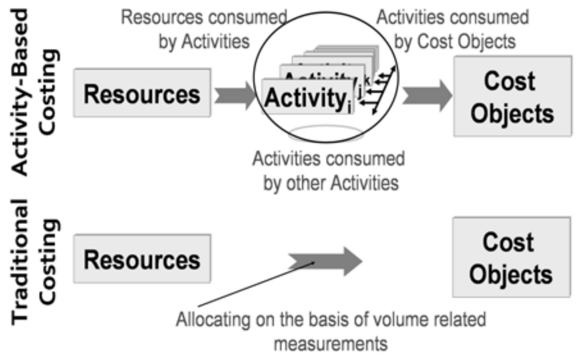


Figure 4: ABC versus TCA

Operation-driven cost accounting

Cost information in pulp and paper mills is mostly evaluated at the end of each month, what doesn't give any flexibility to managers to analyze the potential causes of cost variance. Hence operation-driven cost and process analysis tool such as process-based business modeling [3] [22] brings substantial benefits to this area. This model combines process engineering and cost accounting principles which creates new, in-depth representation of manufacturing processes.

The emphasis is on the cost modeling of continues processes, where knowledge from both process design and operating conditions are considered together in an element called *Process Work Center* (PWC). Process design defines the physical side of PWC which comprises the specification of all process units used within, whereas operating condition describes the activity performed by PWC. Operating conditions are determined by detailed-level information (process measurements such as temperature, flow rate, consistency, etc.). Subsequently, given amount of well defined interconnected PWCs can easily embody and formulate a whole processing plant.

In order to model production overhead and non-manufacturing costs more accurate, the proposed business model framework exploits activity based costing principles. It introduces *Overhead Work Center* (OWC) which has similar attributes to PWC. In a nutshell, cost calculations and analysis are accomplished in every PWC separately including potential OWC contribution. Consequently the calculated costs flow between neighbors PWCs accumulating in the *Overall Cost Object*, which determines the desired final cost object.

The term operation-driven cost accounting is due to the fact that both operating conditions and process design dictate the cost accounting procedure using ABC principles.

Cost data reconciliation

Since cost information for business model may come from different sources, the term of cost reconciliation has emerged because of inconsistencies present between them (usually between manufacturing account and general ledger account). The variance in cost data might be also due to the variation between estimated and actual costs or between “costs accounted for” and “costs to be accounted for” for the next time period in general ledger. However, these differences are generally easily corrected by using weighted-average method.

The reconciliation process is carrying out on a regular basis ending with a final reconciliation for each period in order to prepare consistent financial reports. The possibilities of identifying coherent and more frequent cost reconciliation procedures are being analyzed in order to acquire the quality in combining process and cost data.

PROPOSED OVERALL METHODOLOGY

In order to address and investigate the potential applications of proposed methodology the following challenges and overall steps are of concern:

Define approach for on-line plant-wide process data reconciliation:

- Divide the process operation into subsystems in a way suitable for steady-state identification,
- For each process subsystem perform:
    - Process data denoising and process trend extraction using wavelet-based multi-scale data processing technique,
    - By steady-state detection technique determine possible operating regimes,
    - Perform steady-state data reconciliation and gross error detection to ensure data consistency,
- Define the approach to merge reconciled process data to plant-wide level in order to eliminate discrepancies in process measurements,

Define the approach for systematic cost data reconciliation

- Recognize the different sources of cost data in order to analyze the potential discrepancies among them,

Employ the reconciled process and cost data in process-based business model:

- The well elaborated process-based business framework is used to assess and evaluate the costs of each process subsystems (PWCs) for further analysis,
- Since business model unifies the process and cost information, it forms a good basis for reconciliation of different information. Relationship between process and cost data as well as between PWCs for different plant-wide process conditions can then be modeled and analyzed.

POSSIBLE BENEFITS AND APPLICATIONS

The possible applications and further benefits associated with real-time plant-wide application of proposed methodology are here outlined:

- With better understanding of the process operation, the equipment efficiency can be estimated more precisely. Then plant benefits such as improvement in maintenance for both instruments (calibration) and equipment (cleaning) will emerge,
- Improved instant identification of the cause of the process problems, for example localizing process leaks, product loss and instrument fault detection. With the possibility to tract the origin of the problem back in time,

- Enhanced process control will be achieved when on-line reconciliation results are used to update process status and overall balances. Moreover, denoising of process data using wavelets can help process/control engineers in maintaining process closer to the optimum,
- Once steady-state regime is identified, its information such as: occurrence, duration and the transient period between them can be stored and analyzed. This valuable information can be then used for on-line applications, for example to tune up cutting scale parameters for wavelets as well as to help identify correct operating conditions,
- Furthermore, using such accurate process data in combination with business, quality and environmental data will form a knowledge base for continues improvements in production process. Process-based business model could be used to assess and evaluate different operating regimes in order to select the most profitable ones,
- This technique would provide reliable data for further process integration applications, particularly variability analysis and real-time optimization. Maximizing expected profit simultaneously with minimizing the environmental impact and still meeting the quality requirements would be improved by developing real-time optimization techniques based on business model framework. The results would lead to improved overall process efficiency, such as: improved production efficiency, costs, quality, environmental as well as safety,
- If continuously used, the knowledge from process operation history might be used to predict cost of the following PWC_{i+1} based on the former PWC_i cost analysis. This followed up understanding might be very useful in further applications,
- If such a methodology were to be developed as a potential real-time plant-wide application for pulp and paper industry, the decision making procedures and so mill profits would be enhanced significantly.

CONCLUSION

Pulp and paper mills seek to improve their operation performance in order to increase production and reduce costs while still meeting the quality requirements. Developing tools and methodologies to extract the relevant knowledge from information management system could be one way of doing it. In spite of today’s power in data acquisition, it is still quite difficult to implement functional online data validation system. This is due to various errors present in process and business data that have to be removed since the effectiveness of plant-wide applications depends distinctly on data quality. Integration of on-line multi-scale wavelet analysis for process trend extraction with on-line steady-state data reconciliation could potentially evolve to real-time plant-wide data validation tool. Which can be further applied to process-based business model to better represent a complex pulp and paper dynamic environment from both cost and process perspectives. This could bring a new resource for decision makers as well as for real-time optimization of pulp and paper mills.

REFERENCES

[1] Janssen, M., Laflamme-Mayer, M., and Stuart, P., "Survey of data management systems used in the pulp and paper industry," FOCAPO 2003: Proceedings Foundations Of Computer Aided Process Operations, Coral Springs, FL, USA, pp. 551-554, 2003.

[2] Janssen, M., Laflamme-Mayer, M., Zeinou, M.-H., and Stuart, P. R., Survey Indicates Mills' Need to Exploit IT Systems with New Business Model Pulp & Paper, vol. 78, no. 6, pp. 46-51, 2004

[3] Laflamme-Mayer, M., Linkewich, J., and Stuart, P. R., “Methodology for using realtime process data for supply chain optimization in the pulp and paper industry," 91st Annual Meeting Preprints-Book C, 10 Feb. 2005, pp. 279-286, 2005.

[4] Jiang, T., Chen, B., He, X. (2000). Industrial Application of Wavelet Transform to the On-line Prediction of Side Draw Qualities of Crude Unit. Computers and Chemical Engineering, Vol.24, pp 507-512.

[5] Jiang, T., Chen, B., Jasim, K. Stuart, P.R. (2003a), “Strategy for Improving Data Quality for a Kraft Pulp Mill Recausticizing Plant”, FOCAPO, p 467-470.

[6] Jiang, T., Chen, B., Xiaorong, H. Stuart, P.R. (2003b), “Application of Steady State Detection [1] [9] Method Based on Wavelet Transform”, Computers & Chemical Engineering, vol 27, pp 569-578.

[6] Bellec, S., Taiwen Jiang, Kerr, B., Diamond, M., and Stuart, P., "On-line processing and steady state reconciliation of pulp and paper mill process data," 91st Annual Meeting Preprints-Book B, 9 Feb. 2005, pp. 105-110, 2005.

[7] Bellec, S., Jiang, T., Kerr, B., Diamond, M., Stuart, P.R., (Corrected version submitted in 2006). On-line Processing and Reconciliation of Pulp and Paper Mill Data, Pulp & Paper Canada.

[8] Jacob, J. and Paris, J. (2003a), “Data Sampling and Reconciliation – Application to Pulp and Paper Mills. Part I: Methodology and Implementation”, Appita J., 56(1), 25-52.

[9] Jacob, J. and Paris, J. (2003b), “Data Sampling and Reconciliation – Application to Pulp and Paper Mills. Part II: Case Studies”, Appita J., 56(2), 116-120.

[10] Stettler, M., Marechal, F., Efficient energy conversion in the pulp and paper industry: application to a sulfite wood pulping process. Project Report 2007.

[11] MALLAT, S., HWANG, W.L. (1992). Singularity Detection and Processing with Wavelets. IEE Transactions on Information Theory, Vol.38, pp 617-641.

[12] Crowe, C.M., Campos, Y.A.G., and Hrymak, A. Reconciliation of process flow rates by matrix projection. Part I: linear case, AIChEJ., 29(6):881-888, 1983.

[13] Madron, F., and Veverka, V., Optimal Selection of Measuring Points in Complex Plant by Linear Models, AIChE J., 38(2):227 – 236, 1992.

[14] Sanchez, M.C. and Romagnoli, J.A., Use of orthogonal transformation in data classification-reconciliation, Comp. Chem. Engng., (20):483-493, 1996.

[15] Knepper, J.C. and Gorman, J.W., Statistical Analysis of Constrained Data Sets, AIChE J., 26(2):260-264, 1980.

[16] Liebman, M.J. and Edgar, T.F., Data reconciliation for nonlinear processes. AIChE Annual Meeting, Washington, 1988.

[17] Veverka, V. V., & Madron, F. (1997). Material and energy balancing in the process industries: From microscopic balances to large plants. In Computer-aided chemical engineering series (Vol. 7). Amsterdam, The Netherlands: Elsevier.

[18] Narasimhan, S., & Jordache, C. (2000). Data reconciliation and gross error detection: An intelligent use of process data. Houston, TX: Gulf Publishing Company.

[19] Benqlilou, Chouaib, Data Reconciliation as a Framework for Chemical Processes Optimization and Control, PhD thesis, Universitat Politecnica de Catalunya, 2004.

[20] Bagajewicz, M. (2001) "Process Plant Instrumentation and Upgrade", Technomic Publishing company, Lancaster, Pennsylvania.

[21] Steen, M., Steensland, T., Real-time profit monitoring and activity-based cost management, Tappi journal, 1994, vol 77, pp 105-111.

[22] Janssen, M., Stuart, P.R. (2004b). "Business Model Framework for Using Real-Time Process Data for Design Decision Making". Proceedings, C-DEN Design Conference, Montreal, QC. Comm. 07B5.

[23] Horngren, Charles T., Foster, G., Datar, S. M., Teall, H., and M. Gowing, Cost Accounting: A Managerial Emphasis, Fourth Canadian Edition, Prentice Hall Canada, 2006.

[24] Emblemsvåg, J. and Bras, B.A. (2001) Activity-Based Cost and Environmental Management – A Different Approach to the ISO 14000 Compliance. Boston / Dordrecht / London, Kluwer Academic Publishers

APPENDIX H – INSTRUMENTATION NETWORK IN THE NEWSPRINT MILLS

Although the design of an instrumentation network in the pulp and paper industry is constrained by the costs, and is focused mainly on monitoring the controlled variables, also some secondary variables are measured. And contrary, some necessary controlled variables are not measured directly in certain mill areas lack monitoring system completely. This is typical for the pulp and paper industry since some properties/variables are very hard or even impossible to measure.

The main goals of the instrumentation network in the newsprint processes are usually to monitor:

- the quality of the product (pulp freeness and consistency, paper characteristics)
- the environmental impacts (emissions monitoring, effluent treatment system)
- the operation safety (fault occurrence)
- the performance for the boiler or steam turbine in order to detect gradual changes
- the rate of production for production accounting

The key parameters affecting the process performance and pulp quality are the specific energy consumption and the refining intensity. On the figure 2.5 it is shown how these key parameters of the TMP refining interact. It is very challenging to estimate these parameters based on monitoring variables. The refining intensity is increased when the refiner gap clearance is reduced causing the increase of energy input of single impacts on refining, which is not difficult to control. The second parameter - the specific energy consumption is the major factor influencing process performance and pulp quality. Its value is determined by the motor load and by the production rate of fibres through refiner. The rate of production is a difficult parameter to measure and control because the plug screw feeder (the chip feeding equipment), is based on volumetric feed of material into the refining line. If an accurate measurement of the chip flow is to be determined, the combination of the given volumetric feed rate with an accurate measurement of the bulk density as well as the chip moisture is essential. However, accurate and reliable instruments for moisture and bulk density measurements of chips are not available. Furthermore, the measurement of the production rate based on a pulp consistency and the pulp mass flow after the secondary or tertiary refining is characterized by similar problem, since the consistency measurement is very sensitive to pulp quality (fibre size distribution), air content and temperature.

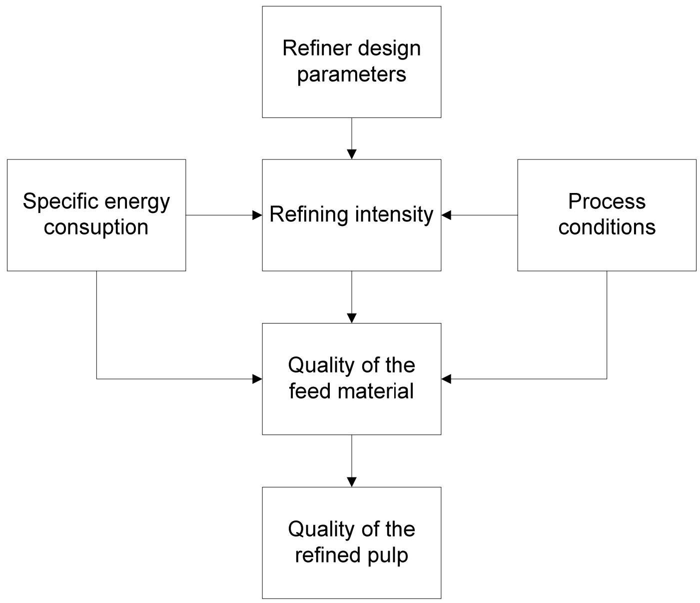


Figure 1: Interaction of the key parameters of TMP refining

Control of refining operation

The basic goal of the refiner controls is to balance the effect of change in raw material properties. This procedure is affected by the accuracy and precision of measured variables which makes the control a difficult task. Typical controlled and manipulated variables are summarised in the figure 2. Generally, there is no common control strategy for all TMP plants due to different operating strategies, process equipments and raw materials (Sundbolm, 1999).

Manipulated variables

1. Plate Gap
2. Chip feed speed (throughput)
3. Dilution rates

Controlled variables

1. Consistency
2. Specific energy consumption
3. Freeness
4. Long fiber content

Figure 2: A typical controlled and manipulated variables for TMP refining

APPENDIX I – ADDITIONAL METHOD DESCRIPTION

<u>Objectives of this section</u>:

- To describe the method with output and input variables between calculation steps.
- To present a clear description of methodology within each calculation step.

The overall method presentation is described according to the methodological steps that are shown in Figure 1. The knowledge acquired from process understanding serves to optimally select a set of key variables from the whole measurement data set that represent an operating condition of the plant-wide production. The case study process model is presented in detail as flow sheet diagrams (Figures 2 to 11 of this section), highlighting the measured variables and those that are selected as key variables. This set of key-variables is utilized in the second step of the method, to identify a near steady-state condition that represents the given production regime using the wavelet processing technique. With the identified near steady-state periods, the whole measurement data set is processed using wavelet transform, in order to eliminate random noise and abnormalities and improve the performance of process data reconciliation step. The terms and mathematical formulation of the data reconciliation (error minimization) problem are presented (section. The whole measurement data set is presented by classifying variables into redundant, estimable and non-estimable groups respectively. The final stage of the reconciliation model presentation is a discussion of the solver functionality followed by an example.

The use of relatively complex data treatment procedure is justified by acquiring novel cost information on various production regimes. Several cost items that were used to assess production knowledge are classified into two groups: direct and indirect cost items. The operations-driven cost model is discussed further to clarify the ABC-like character of the approach and to discuss the necessity to use indirect costs allocation.

The overall method presentation is finalized by discussing its valuable implications for new types of decision making, for short and long term company's benefits.

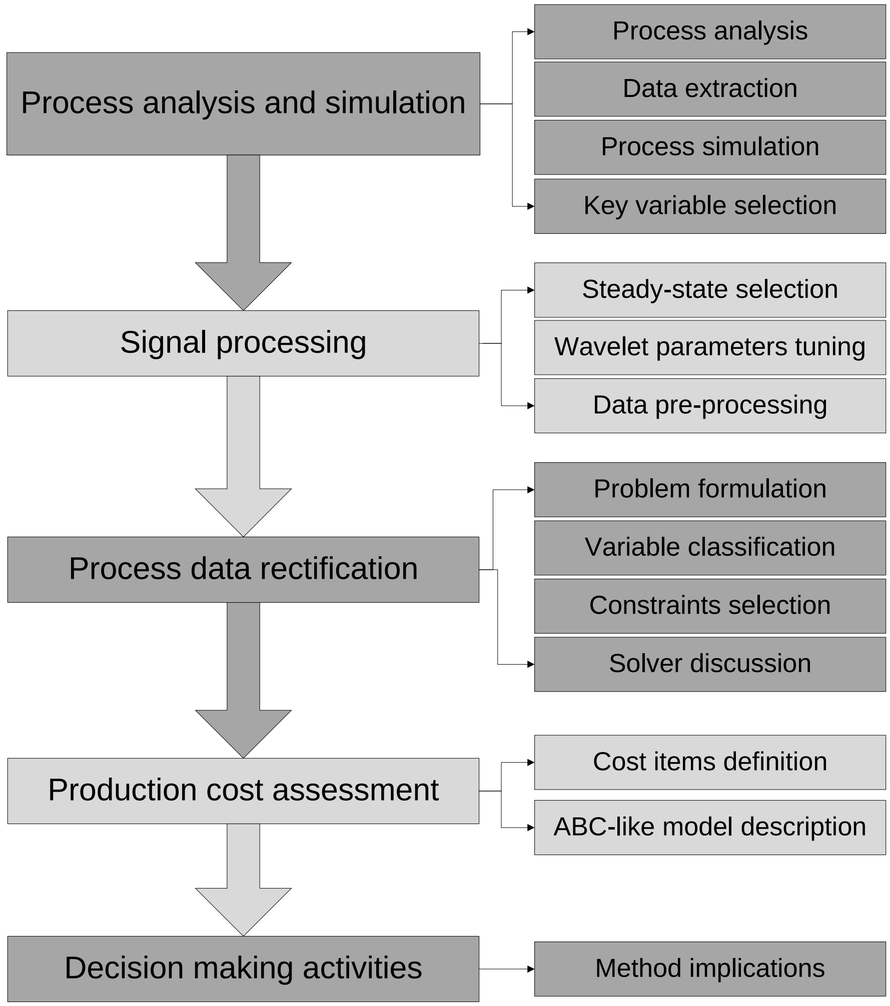


Figure 1**:** Overall method presentation

Process analysis and simulation

Process analysis

It is necessary to comprehend the production processes by analyzing process flow sheets, PID diagrams and production practices prior to developing the simulation model and applying the proposed method. According to the acquired knowledge a first version of the simulation model was developed using WinGEMS software, later, it was then transferred over to CADSIM software for further development.

Data extraction and Process Simulation

The process simulation was built using CADSIM software with the possibility to reconcile process data. For better visualization purposes, simulation charts from Metso Process Simulator (Metso WinGEMS 5.3 ©) are presented in the following section. The plant-wide operation was divided into six process work centers. This dissection of production processes provides in-depth tracking of resource consumption to different parts of mill.

The flow-sheets presented contain large amounts of information including names and types of units within the process, as well as numerical values of corresponding process streams. These actual values are irrelevant for the purpose of model presentation; however they have been included as a part of the model visualization.

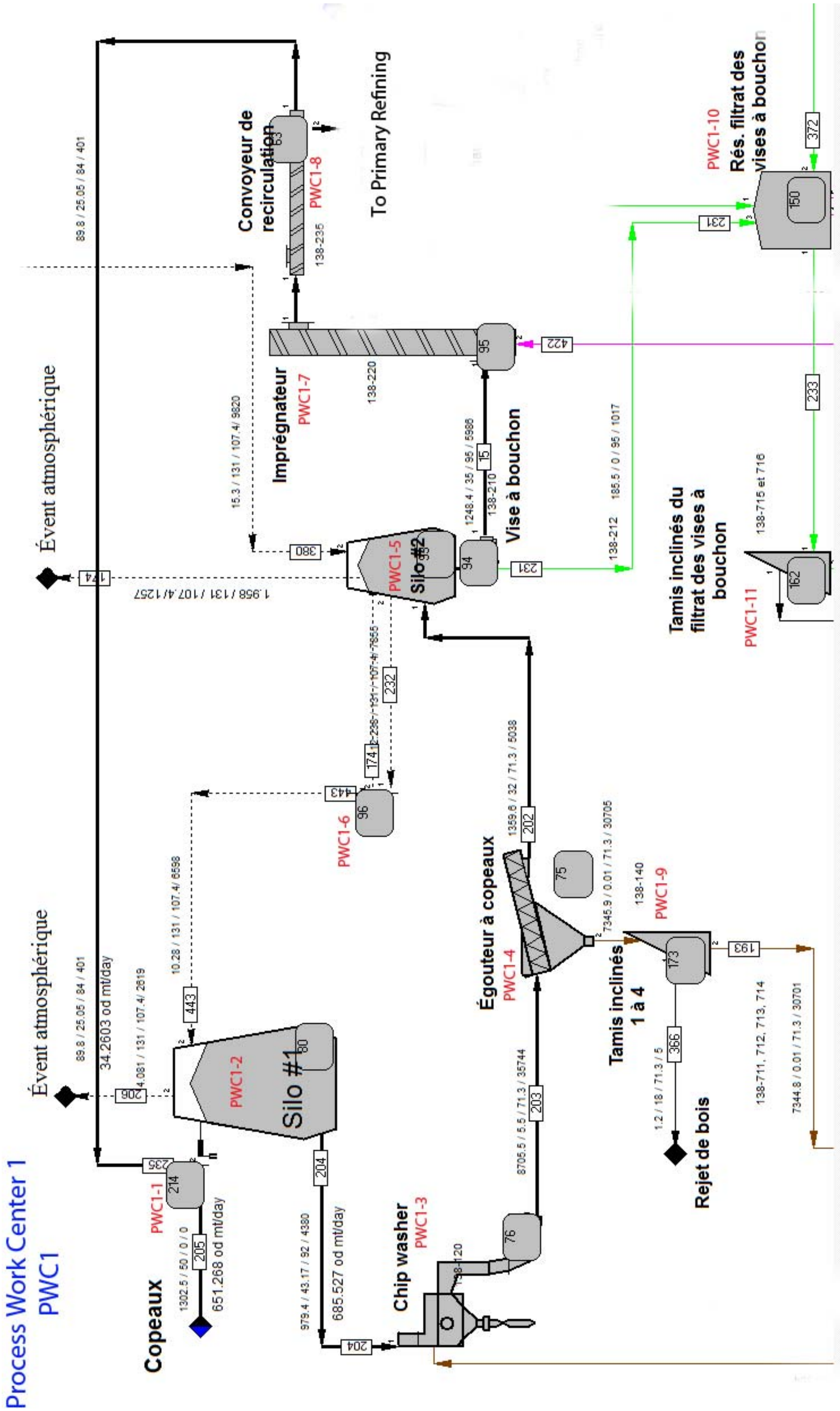


Figure 2 Process flow diagram - chips pretreatment

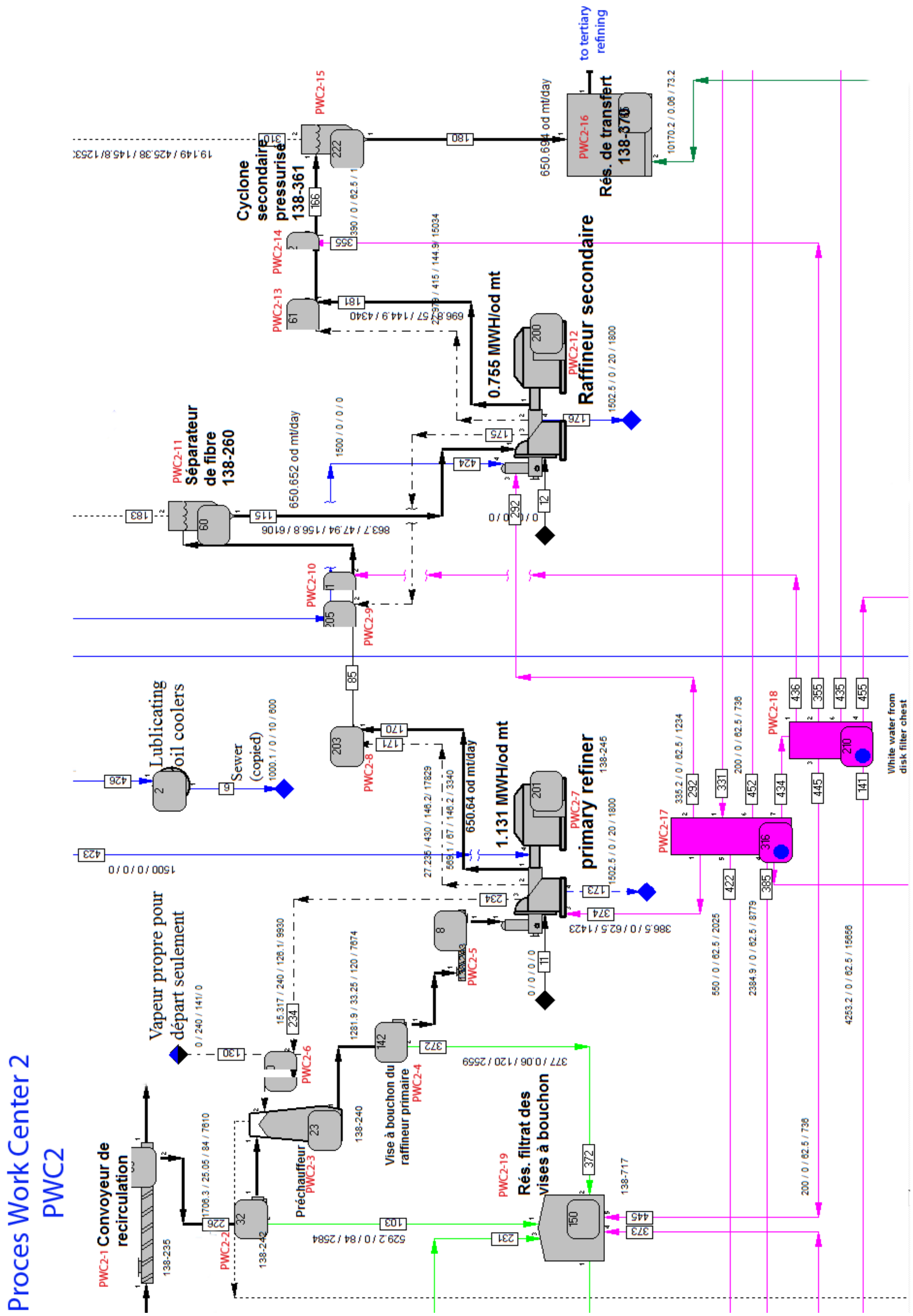


Figure 3 Process flow diagram - Main refining line

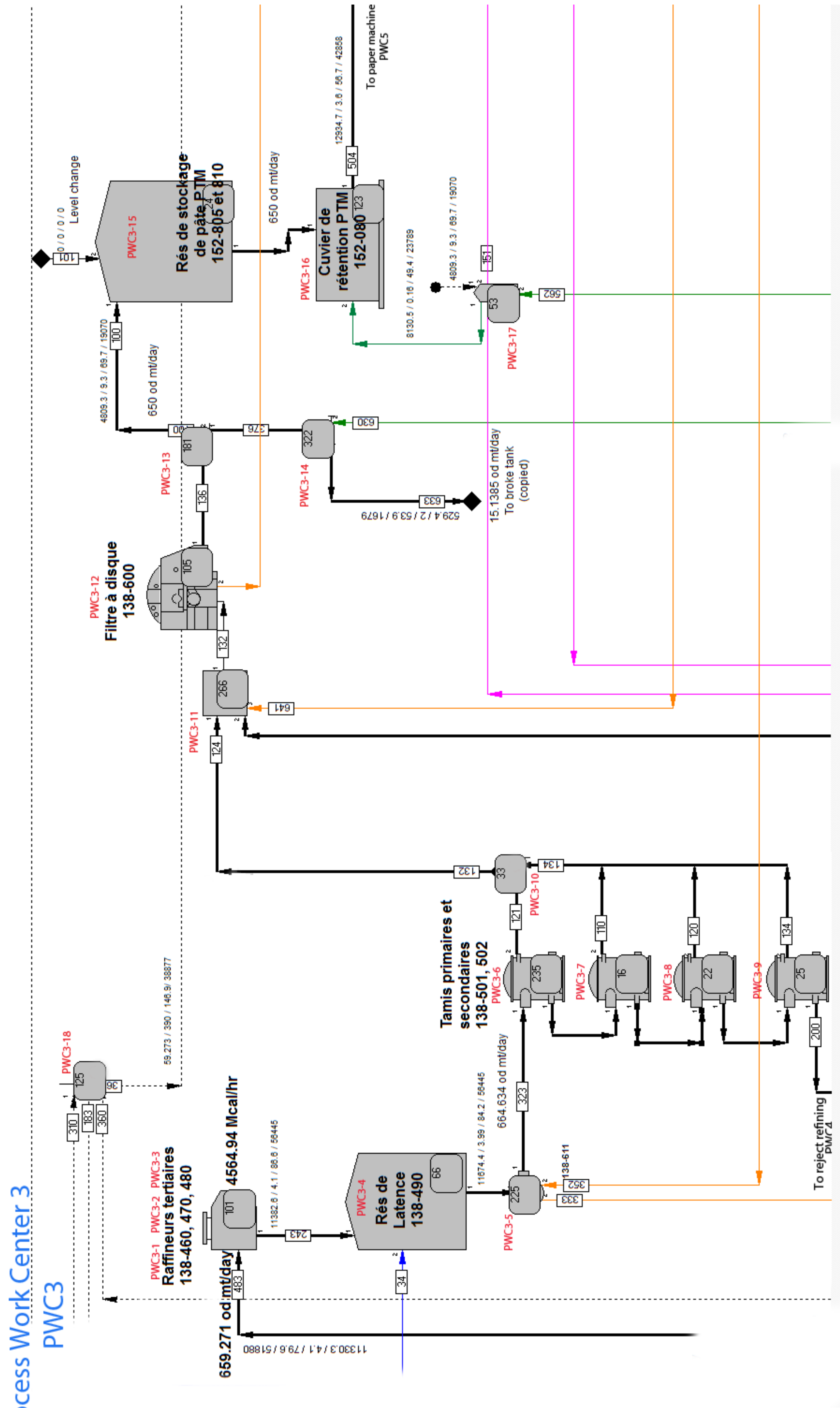


Figure 4 Process flow diagram - Screening

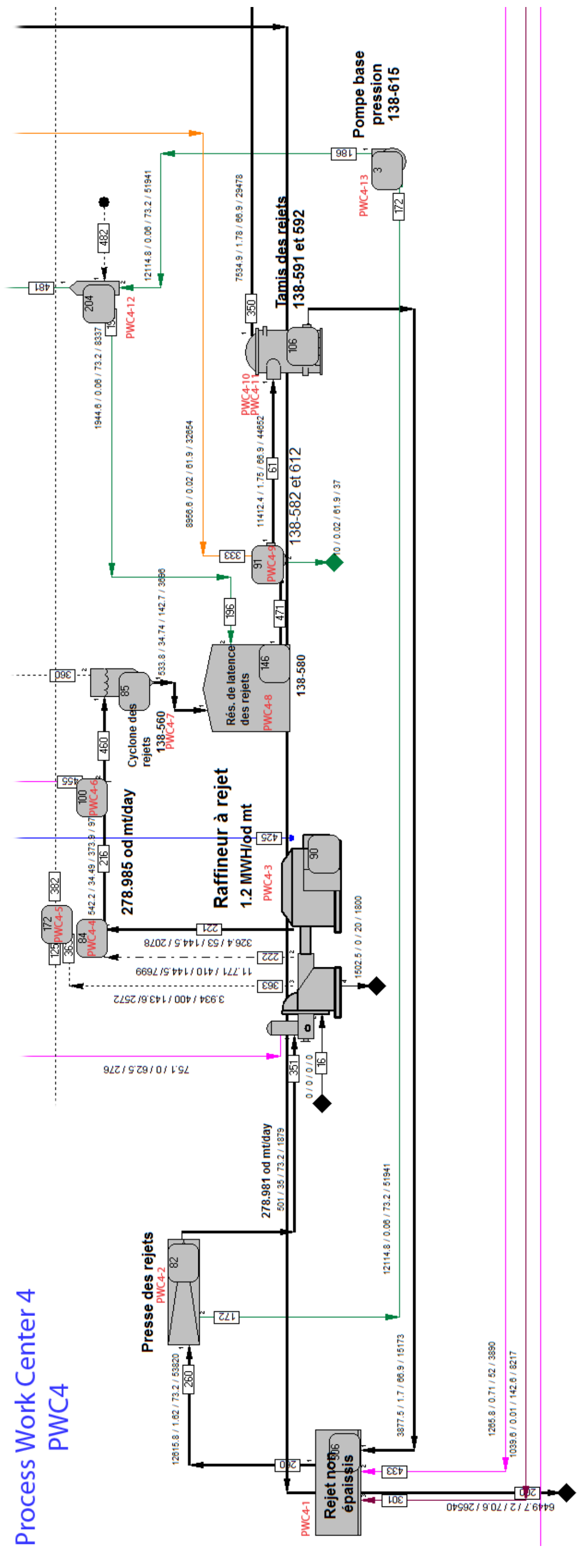


Figure 5 Process flow diagram – Reject refining line

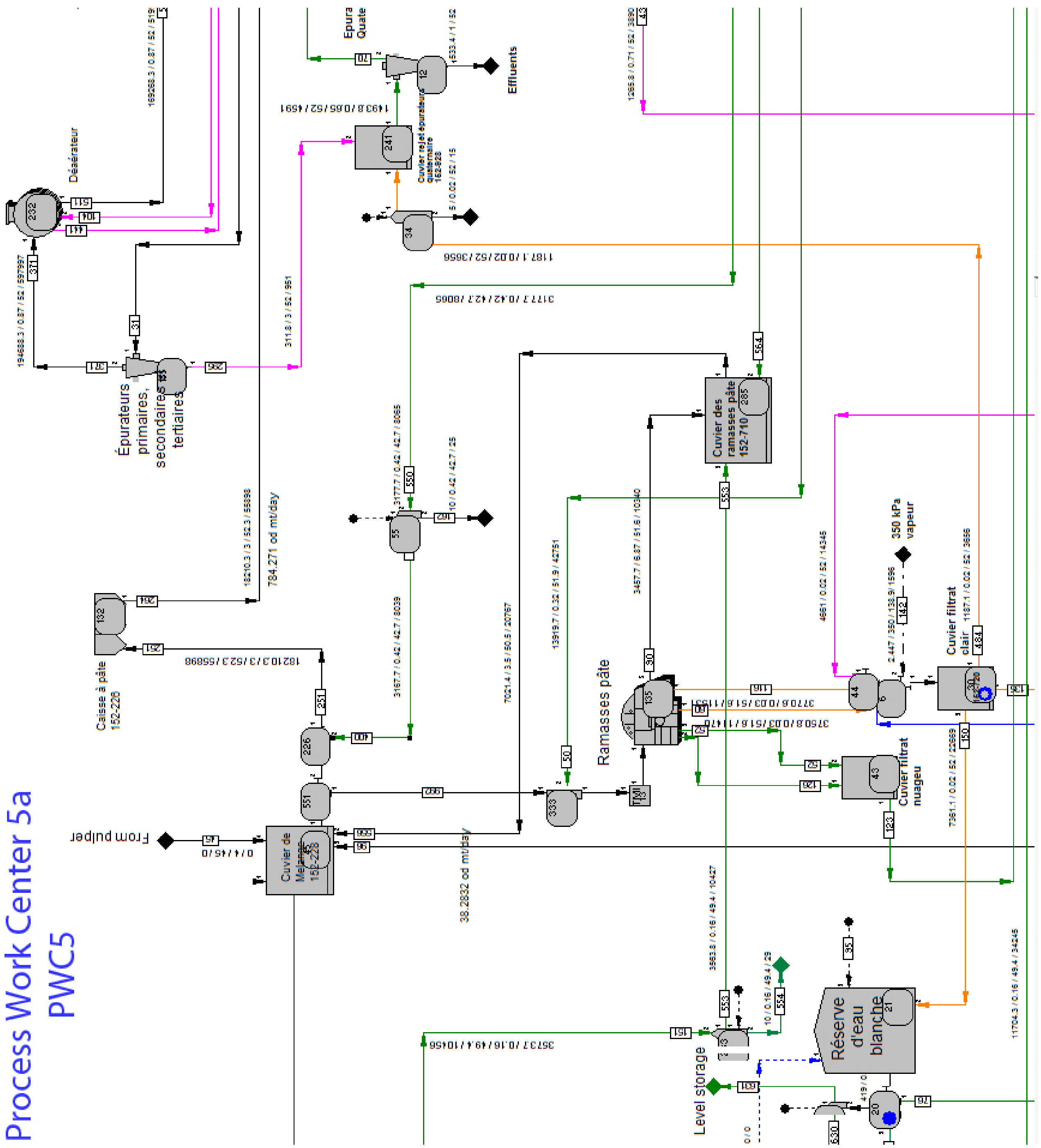


Figure 6 Process flow diagram - paper machine

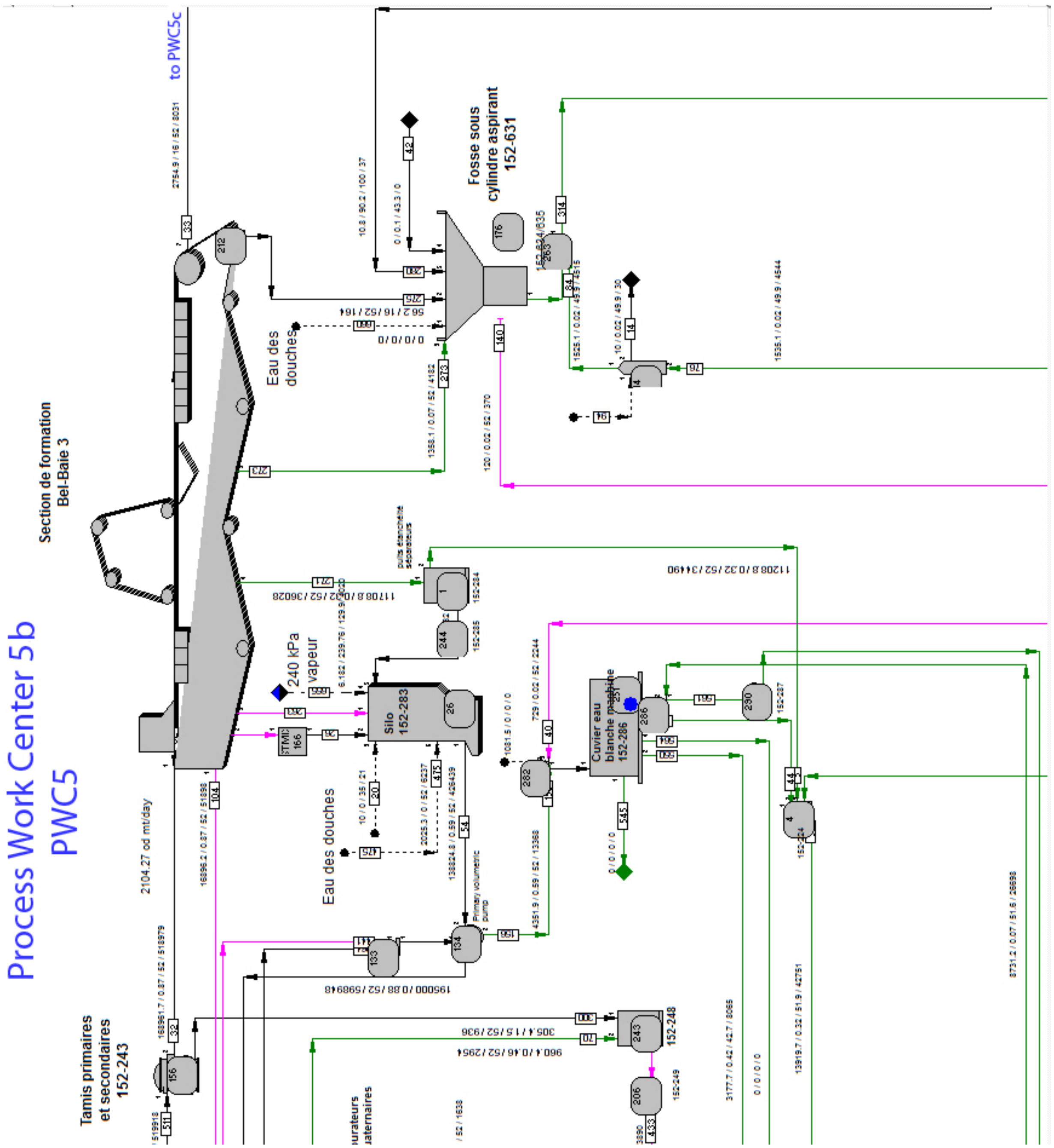


Figure7**:** Process flow diagram - chips pretreatment

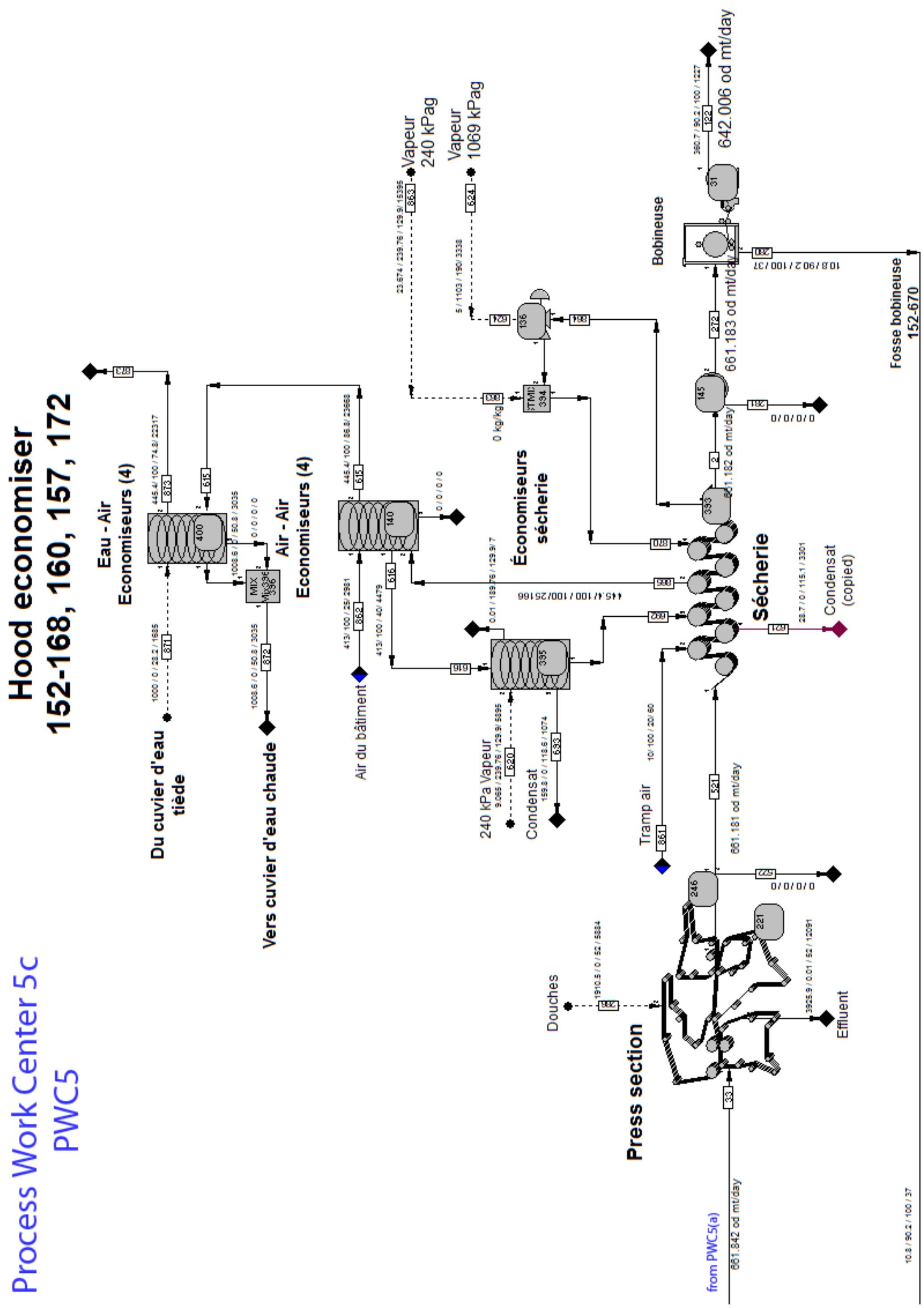


Figure 8**:** Process flow diagram - paper machine

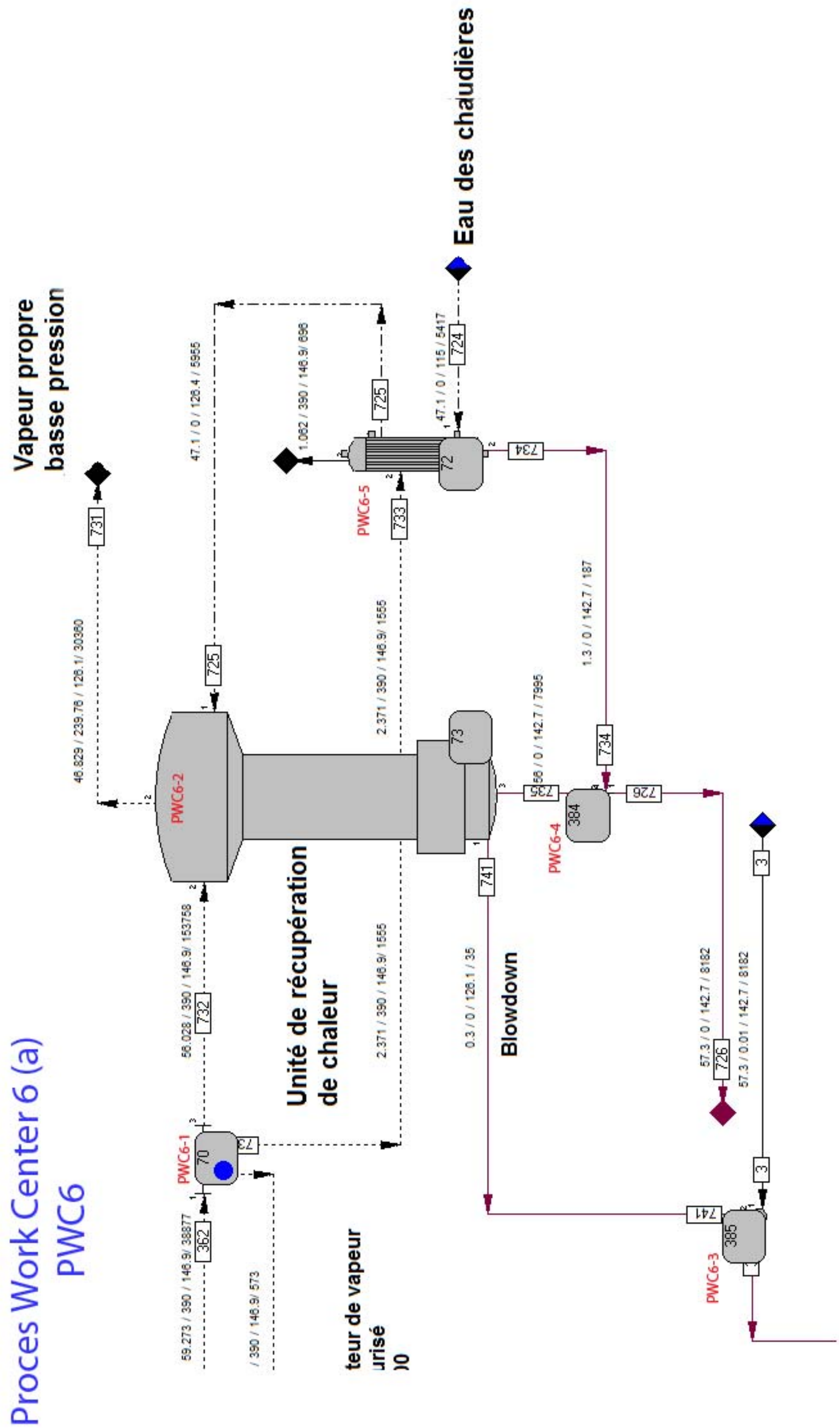


Figure 9 Process flow diagram - Energy Island

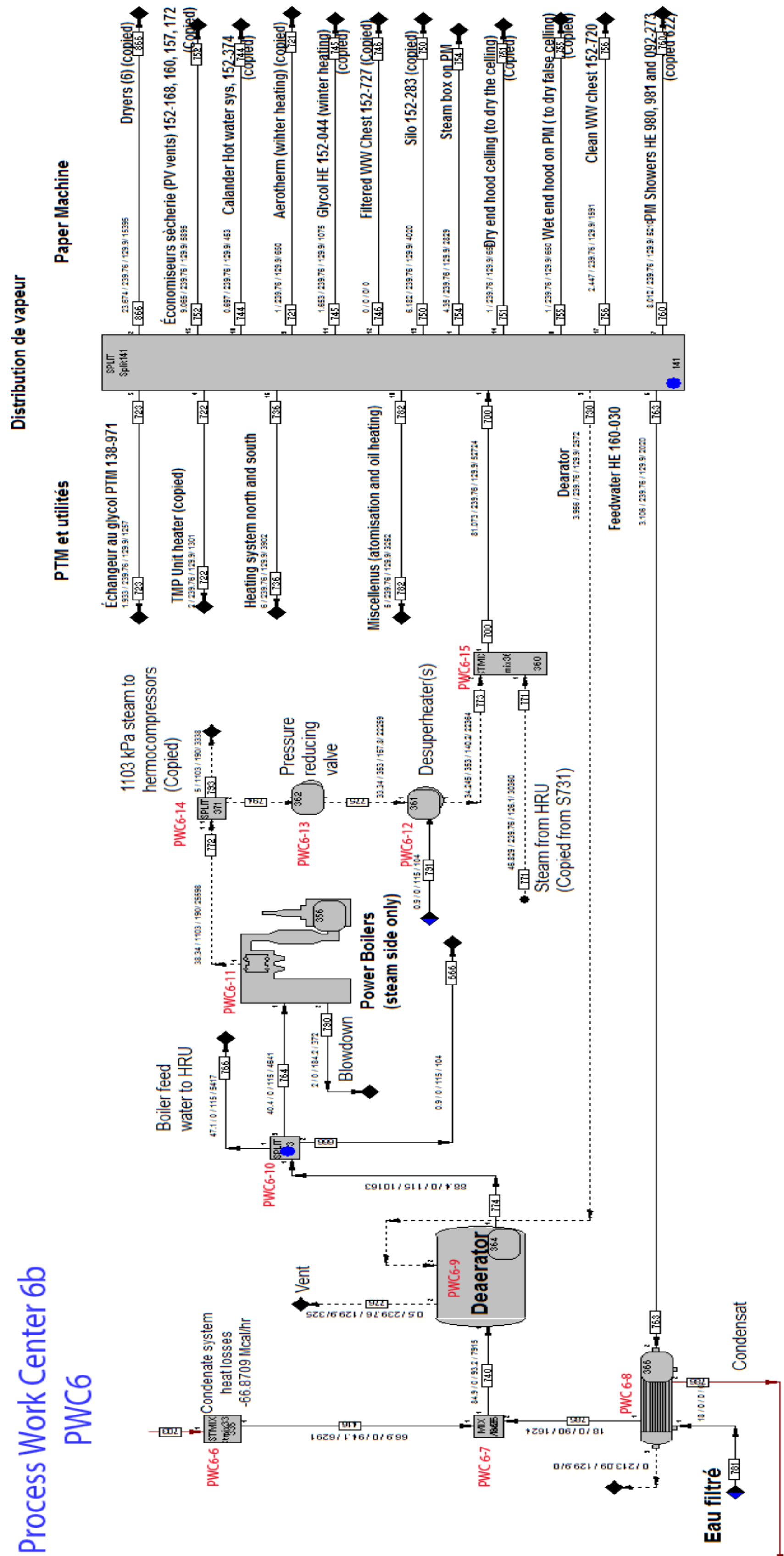


Figure 11 Process flow diagram - Energy Island

LEGENDE

White water 0.07%	Steam 240 Kpag
White water 0.04%	Steam 1070 Kpag
White water 0.01%	Air
Fresh water	Filtrat
Pulp, paper and chips	
Hot water	
Warm water	
Chip washing water	Condensat

Units for liquid and pulp

lpm/ %OD / deg C / Mcal/h

Steam unit:

mt/hr /KPa / deg.C/ Mcal/h

Figure 12 Legend

Key variable selection

The large-scale and complexity of processing plants challenges the occurrence of a steady-state condition. The process state representation is performed by using monitored variables. However, the use of the whole set of measured variables makes it challenging to actually identify the near steady-state period of a large scale process. Thus it is essential to select a subset of measurement variables to represent a state of the process. In practice, this is overcome by using knowledge of the process operation, i.e., an analyst has pre-selected a set of variables that represent the process state. This pre-selection is based on the precept that not all variables are equally important; certain variables contain more information about the process state than others and are hence more useful for a given purpose. The key variables are assumed to be non-correlated and should cover the system as completely as possible. In this work, such variables are referred to as *key variables*.

There is no unique way to select the set of key variables that characterize process states. In this work the selection was carried out with the use of engineering judgment about the underlying production processes and process variables. The judgment is drawn from understanding of the process dynamics involved with the particular process under study. These key variables provide direct information about the state of a process. Processing plants are commonly operated in multiple states and frequently switch between them. A fundamental precept in this work is that measured key -variables depend on the operating state of the process. For instance, when considering a thermomechanical pulping process, the measurement of production-rate is critical, and should be selected as a key variable since changes in the pulp flow will inherently result in process changes (specific energy of refining). The process of selection of key variables for thermomechanical newsprint process has taken into account following measurements:

- Main line pulp flow and consistency indicators
- Measurements of level in various pulp tanks
- Paper machine speed
- Steam volume rate

The summary of the selected key variables is listed in the table 5.1. The type of Process Work Center (PWC) and Input / Output processing units correspond to the simulation model presentation in flow-sheets (Figure 2-10 of this Appendix I).

Table 1: List of selected key variables for steady-state detection

No	Key variable for steady state detection	PWCi	Output unit	Input unit
1	Production volume (chips throughput)	PWC1	PWC1-1	PWC1-2
2	Level Silo 1	PWC1	PWC1-2	-
3	Level Silo 2	PWC1	PWC1-5	-
4	Production volume (chips throughput)	PWC1	PWC1-5	PWC1-7
5	Production volume (chips throughput)	PWC2	PWC2-1	PWC2-2
6	Production volume (chips throughput)	PWC2	PWC2-5	PWC2-7
7	Consistency measure	PWC2	PWC2-7	-
8	Stem flow rate	PWC2	PWC2-7	PWC2-6
9	Production volume (pulp throughput)	PWC2	PWC2-11	PWC2-12
10	Level transfer chest	PWC2	PWC2-16	-
11	Production volume (pulp throughput)	PWC3	PWC2-16	PWC3-1
12	Production volume (pulp throughput)	PWC3	PWC2-16	PWC3-2
13	Production volume (pulp throughput)	PWC3	PWC2-16	PWC3-3
14	Consistency measure	PWC3	PWC2-16	-
15	Level measure - Latency chest	PWC3	PWC3-4	-
16	Production volume (pulp throughput)	PWC3	PWC3-10	PWC3-11
17	Level measure - pulp tower	PWC3	PWC3-15	-
18	Level measure -reject storage	PWC4	PWC4-1	-
19	Production volume (pulp throughput)	PWC4	PWC4-2	PWC4-3
20	Level measure - latency chest	PWC4	PWC4-8	-
21	Production volume (pulp throughput)	PWC4	PWC4-11	PWC3-11
22	Production volume	PWC5	PWC3-16	PWC5-1
23	Consistency measure	PWC5	PWC3-16	-
24	Production volume (paper production)	PWC5	PWC5-19	PWC5-24
25	Steam flow	PWC5	PWC5-24	PWC5-26
26	Steam flow	PWC6	PWC6-1	PWC6-2
27	Steam flow	PWC6	PWC6-2	PWC6-5
28	Steam flow	PWC6	PWC6-11	PWC6-14

Signal Processing

This part of the thesis revision document discusses further the functionality of signal processing step of the method.

The first block of activity of the method is careful cleansing of real-time process data from high-frequency noise, abnormalities and identification of when the process is near steady-state operation. The technique used involves multi-scale wavelet decomposition of measurement trends. Two essential steps are required to initialize the wavelet technique:

- Gathering of information on each individual measurement point and analysis of its historical values to identify the optimum wavelet transform (WT) cutting scale for each variable, and
- Analysis of historical data to determine the optimal steady-state values of the detection parameters (alpha parameters).

According to the two steps described above, the sensor network at the case mill was analyzed, and each measurement point was characterized by its accuracy and precision values. Multiple decomposition trials and tests of each measured variable trend were carried out to identify the two essential scaling and parameter values.

Only the selected key variables that represent the process state are first processed to cleanse random and abnormal errors. Simultaneously, the processing technique analyzes the time-frequency domain for a potential steady-state occurrence of each selected variable. When a steady state has been identified, an automatic check for a potential multivariable pseudo-steady state is performed. In this way, a plant-wide process steady state can be systematically detected and used as an input to the simulation-driven data rectification technique. Within the identified steady-state period, the whole set of measured raw process data is extracted from the information management system (non-compressed data sets). Then a wavelet data based processing technique is used to clean random noise and abnormalities from the set of data.

The two simultaneous tasks of data pre-processing and pseudo-steady-state detection and tuning corresponding parameters are discussed further in the following section.

Wavelet parameter tuning for data pre-processing and steady-state detection

The measurement data is extracted from the information management system (IMS) as a noisy signal. The process of extracting measurement values must include the uncompressing of data from storage. The optimal WT scale is analyzed using an iterative procedure to find the optimal values. After applying a wavelet transform of the chosen scale (data pre-processing), Gaussian noise and other abnormalities are discarded from the process trend (Figure 5.13). The process status should be estimated at the proper scale. If the scale is too small, the WT representation of the signal is dominated by noise and temporal features and thus distinct trends in process measurements are difficult to distinguish. On the other hand when the scale is too large, the measurement trends will be smoothed too much creating distortion, and the steady-state detection may not sharply reflect the variation of the process.

For illustration purposes, a process trend wavelet representation of a steam measurement is shown on the Figure 5.12. The analyst must be careful to not distort the true signal, and at the same time, must extract as many random errors as possible. Hence the scale choice of 4 is selected because the trade-off between scarcity of the detailed signal and distortion of process trend.

An example of a time period between 300-350 hours of the presented trend is analyzed for tuning the steady-state parameters. According to the extensive definition in Appendix A of this thesis, several parameters must be tuned for steady-state detection of each measured key variable. The three critical parameters (α_1, α_2 and α_3) are determined according to the degree of fluctuation of the measurements as well as the sensitivity of key variables to process operation. For practical industrial applications, they can be selected from the historical knowledge of the measurements as follows:

- Select successive measurement under steady-state
- Perform multi-scale WT and second order WT
- Compute the standard deviation of the WT (σ_w)
- Compute the median of the second order WT (σ_{ww})

Then the threshold values for steady-state detection are calculated as follows:

$$\alpha_1 = \sigma_w$$

$$\alpha_2 = \sigma_{ww}$$

$$\alpha_3 = \lambda\sigma_w$$

where λ is an adjustable parameter whose value is around 1.0, and is determined according to the variability of the measurements and their sensitivity to the operation.

Data pre-processing

Within the identified steady-state period, each measured variable is pre-processed. The knowledge from historical data (from information management systems at the mill) is used to tune appropriate thresholding parameters. Thresholding is the process of discarding values that are below a threshold and keeping values over the threshold, which is often used in data compression or image processing. This process of tuning is done by applying the wavelet algorithm to each measurement separately. The detection of the optimal parameters is described in Appendix A. In this work the thresholding for noise removal is based on the algorithm presented by Jiang et al. (2000).

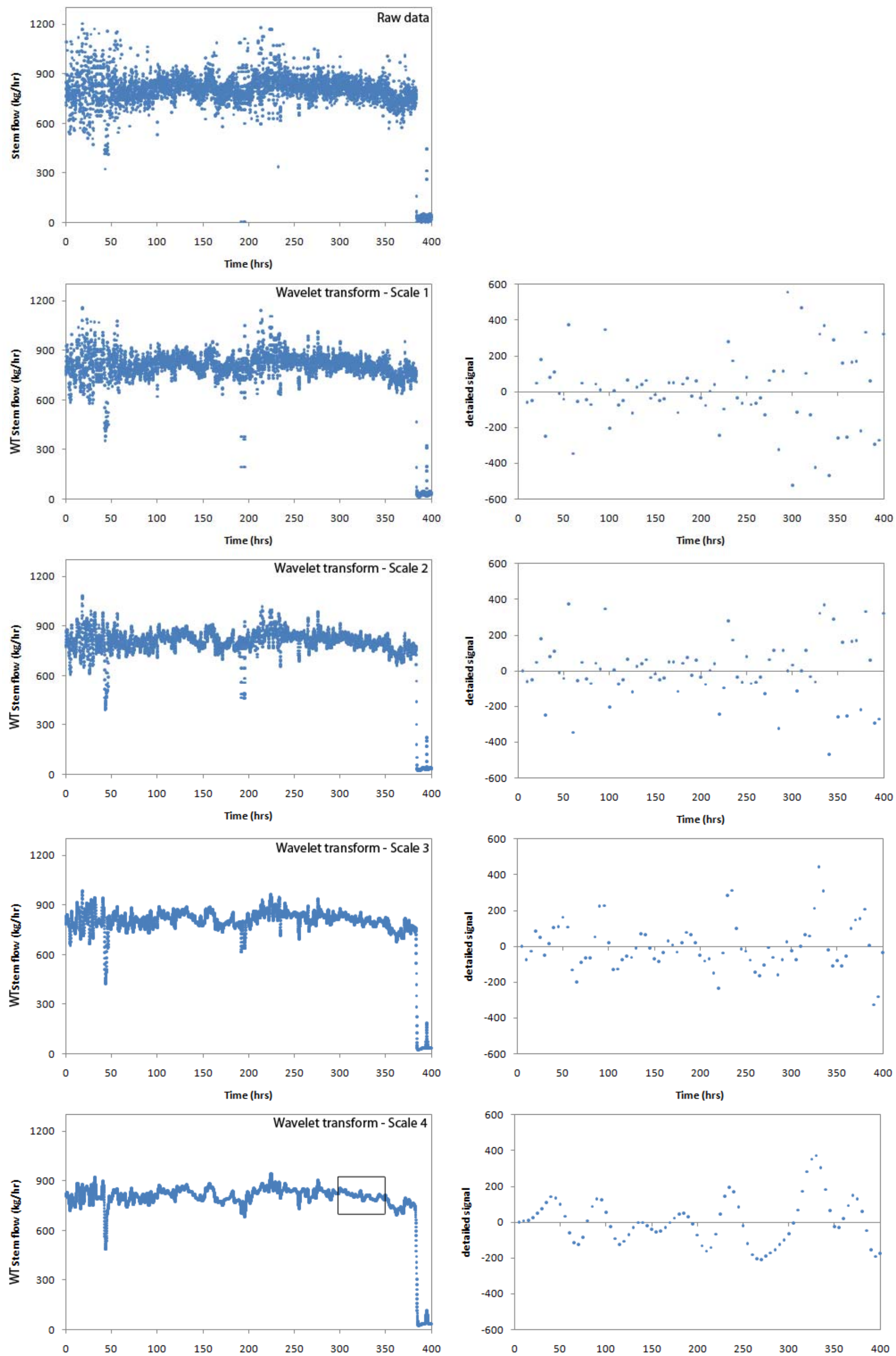


Figure 13: Wavelet decomposition of a steam measurement (scale 1-4)

Process data reconciliation

This third section of the overall method presentation discusses the second pillar of the method: steady-state data reconciliation using a simulation driven solver. The need to apply a data reconciliation step is justified by the imbalanced nature of monitored variables. It is important that a balanced set of process data is used to represent the operating condition since the state of a pulp and paper process at any given time can be, in principle, described by the values of a number of parameters such as stream properties (flow rates, concentrations, temperatures, etc) and equipment operating parameters (dilution rates, separation efficiencies, retention in the paper machine, etc). Those parameter values must satisfy the equations of mass and energy conservation. However, in practice, the true values of these variables are unknown, only measured values of some of those parameters are known. The measured values are not equal to the true process parameters values because of imprecision or inaccuracy of monitoring and recording devices, instrument disfunction or poor sampling techniques. Furthermore, since a manufacturing pulp and paper process is never at rigorous steady-state because of a variety of disturbances and since measurements are not taken instantly, they do not represent a single process state. Among frequent process disturbances are, for example, feedstock variability, equipment instability, sheet breaks in the paper machine, equipment change over, etc. As a consequence, measured process variables and parameter values do not constitute a coherent set of data, i.e. a set of data satisfying the conservation equations. This is a major obstacle when trying to represent real process operation for analysis purposes. In practice, it is assumed that the reconciled data constitutes a reliable approximation of the real process and can therefore be safely used to represent a process operation under study.

The first part of this section describes the general problem formulation of data reconciliation. The discussion and presentation of the system under study is then explained in more detail, i.e. selection of values for weighting matrix assessment, variable classification, constraint selection and finally the solver presentation.

Problem formulation

In general, the optimal estimates for process variables by data reconciliation are solutions to a constrained least-squares or maximum likelihood objective function, where the measurement errors are minimized with process model constraints. With the assumption of normally distributed measurements, a least-squares objective function is conventionally formulated for the data reconciliation problem. At process steady state, the reconciled data are obtained by:

Min$_y$ (J) $\quad J(\hat{\mathbf{y}},\hat{\mathbf{z}}) = (\mathbf{y}-\hat{\mathbf{y}})^T \mathbf{V}^{-1}(\mathbf{y}-\hat{\mathbf{y}})$

Subject to $\quad \mathbf{f}\ (\hat{\mathbf{y}},\hat{\mathbf{z}}) = \mathbf{0}$

$\quad \mathbf{g}\ (\hat{\mathbf{y}},\hat{\mathbf{z}}) \geq \mathbf{0}$

where

y is an M×1 vector of raw measurements for M process variables,

$\hat{\mathbf{y}}$ is an M×1 vector of estimates (reconciled values) for the M process variables,

$\hat{\mathbf{z}}$ is an N×1 vector of estimates for unmeasured process variables, **z,**

V is an M ×M weighting matrix of the measurements,

f is a C×1 vector describing the functional form of model equality constraints,

g is a D×1 vector describing the functional form of model inequality constraints which include simple upper and lower bounds.

The models employed in data reconciliation represent variable relationships of the physical system of the process. The reconciled data takes information from both the measurements and the models. In reconciling steady-state measurements, the model constraints are algebraic equations.

Weighting matrix

In data reconciliation theory, the weighting matrix is expressed as a standard deviation of each measurement variable. This is justified when only random errors with normal distribution corrupt the measured variable. However, in reality the errors are not only random noise but also systematic errors that are caused by dynamic nature of process operation. These small biases must be reconciled in order to satisfy material and energy balances. On the other hand, large biases in measurement that are caused by incorrect instrument readings must be identified and corrected before they can be used in data reconciliation.

Another approach that is a common practice in industrial applications is to use engineering judgment for matrix estimation by allocating uncertainty weights to each instrument (Narasimhan and Jordache, 2000). This pragmatic approach, which is also used in the current study, takes into account knowledge of the process dynamics around each particular instrumentation sub-network as well as information about each instrument's accuracy, precision, and reliability. The weighting aspect enables the analyst to provide levels of confidence between various measurements, i.e. the expert can specify that some measurements are likely to be more reliable than others.

Process Variable classification

In order to better assess input and output variables going to and coming from the data reconciliation model, it is important to clarify the concept of variable classification in data reconciliation techniques. Measured variables are classified as redundant and non-redundant, whereas unmeasured variables are classified as observable and non-observable. The general classification of process variables is shown in Figure 13 followed by the classification of case study process variables.

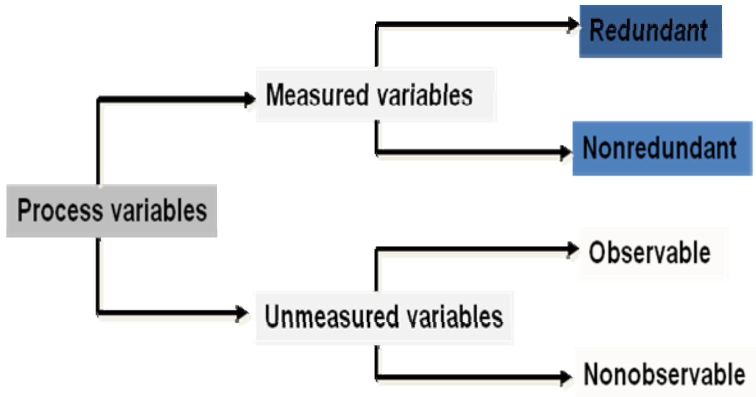


Figure 13: Classification of process variables

Definitions:

- A **redundant variable** is a measured variable that can be estimated by other measured variables via process models, in addition to its measurement.
- A **non-redundant variable** is a measured variable that cannot be estimated other than by its own measurement.
- An **observable variable** is an unmeasured variable that can be estimated from measured variables through physical models.
- A **non-observable variable** is a variable for which no information is available.

Table 2: Case study - classified variables

	Variables	Measured	Redundant	Non-redundant	Observable	Non-observable
TOTAL	1261	386	67	319	639	236
Streams	786	212	54	158	445	129
Temperatures	394	155	13	142	157	82
Consistencies	81	19	0	19	37	25

Constraint selection

Two types of constraints are specified in the data reconciliation procedure:

- Equations of process units in simulator
    - mass and energy conservation laws
    - unit specific equations (screener ratio of reject and accept flows, steam generation from high consistency refiners, etc)
- Lower and upper boundaries of each measured variable

Equations specific to processing units:

First, main equations and unit specific equations are computed by the simulator for each type of unit:

- Splitter (process unit that splits production streams)
- Silo vessels (chips pretreatment by process low pressure steam)
- Heat exchangers
- Impregnation unit (chips pretreatment)
- Cyclone (separation of fibers)
- Disk filter (separation of fines from main pulp line)
- Boiler (high pressure steam generation unit)
- Recovery unit (recovery and cleaning of process low pressure steam)
- High consistency refining
- Low consistency refining
- Screeners
- Pumps
- Paper machine – forming section
- Paper machine – press section
- Paper machine – drier section

Solver description

In this work, data reconciliation was performed using commercial software (CADSIM Plus academic version 2.4). The simulation solver is sequential, but the data rectification solver is simultaneous. First, the analyst selects a number of independent measured variables, referred to as free variables that satisfy system's zero-degrees of freedom. The free variables are inputed into the process simulation model to estimate the whole set of process variables. The pillar of the data rectification module is a modified version of the simplex optimization technique. In the first iteration, the algorithm compares the changes in the free simulated variables to their measured values. The simulation and iteration process repeats until the minimum least-squares error between the simulated variables and the measured values is obtained. The output of the rectification process is the set of reconciled measured values and other calculated variables not available from measurements.

Simulation context

Process simulation refers to the process of solving a system of nonlinear equations that represent the process operation under the fundamental laws of mass and energy conservation. One can represent a process using a set of "m" equations with "n" unknowns, with the difference "n-m=d" representing the "d" degrees of freedom of the system (DOF). The simulator can find a solution when $d = 0$. Generally, two types of solving procedures exist: simultaneous or sequential.

In the case of simultaneous simulation, the solution can only occur when "d" over "n" combinations of unknown variables are specified. The analyst must specify the "d" variables according to knowledge about the operation. In practice some combinations of specified variables will lead to a very difficult system to be solved, sometimes even impossible to resolve.

As sequential simulation is of concern, the large problem is reduced into a set of smaller sub-systems being defined by the equations describing a single operating unit of the whole process. Conceptually the global set of "m" equations is broken down to "p" smaller local sets of equations. Hence in sequential simulation the choice of which of the combinations is to be used provides better chances of finding a solution.

Since the simulator in this work is a sequential solver for data reconciliation, the notion of sequential versus simultaneous procedure becomes critical since it dictates which variables are to be specified. It is possible to draw an analogy between the notion of observability of a system as defined by the choice of its measured variables and the notion of degrees of freedom (DOF) of a simulated system defined by the choice of its specified variables. Since a simulation engine can solve only when the DOF are zero, it is possible to assume that a simulation is always "observable". The CADSIM simulator requires a set of "d" variables to be specified (free variables, FV) and computes the remaining "n-d" variables (computed variables, CV) using first principle equations. When using data reconciliation terminology and assuming the values provided for the free variables come from process measurements, the simulation can be assumed to bea non-redundant system.

For a data reconciliation procedure, an analyst must choose from the whole set of measured variables, a set of redundant measured variables. Since the simulation is "observable" the system will then have the flexibility to associate these redundant measurements on any combination of FV or CV in a simulation. If these measurements are attached to only some of the FV in the simulation, the data reconciliation problem can be classified as "redundant" for every such measurement. If the set of measurements is associated to combinations of the FV and CV in the simulation, then the data reconciliation problem is "redundant" if the CV measurements have a functional relationship with some of the FV measurements. Finally if the measurements are attached to only some of the CV in the simulation, then the data reconciliation problem is "non-redundant" for every such measurement.

Reconciliation strategy

Within the framework of the method implementation, a minimization of the weighted square error between the simulated values, which are necessarily balanced, and the corresponding measured variables, which are inherently unbalanced is carried out:

$$min_{x_i} \sum_{i=1}^{n} w_i \left(\frac{y_i - x_i}{span_{x_i}} \right)^2$$

$$\text{s.t.} \quad \mathbf{f}(\mathbf{x}, \mathbf{z}) = 0$$

$$\mathbf{g}(\mathbf{x}, \mathbf{z}) \geq 0$$

where:

x_i = reconciled value;

y_i = measurement – free variables (FV);

w_i = weight;

z_i = non-measurement variables - computed variables (CV)

$span_{xi}$ = normal operating span for variable x_i ;

The vector x is subject to constraint equations, i.e. mass and energy conservation laws and specified inequality constraints. The objective function is added to the simulation to be reconciled using mathematical functions native to the simulator used. The iterative search for values of x is performed using optimization module based on a simulated annealing version of the well known simplex optimization algorithm. Furthermore, normalized values are used to calculate the objective function due to the variety of physical units met in pulp and paper operation, for instance volumetric flows which can be in the thousands and mass fractions which are between 0 and 1.

Downhill Simplex Method

The downhill simplex method is an optimization algorithm created by Nelder and Mead (1965). This heuristic method does not make any assumption on the objective function to minimize. In particular, the objective function must not satisfy any condition of differentiability. It relies on the use of simplices, *i.e.* polytopes of dimension n+1. For instance, in two dimensions, a simplex is a polytope with 3 vertices, most commonly known as a triangle. In three dimensions, a simplex is tetrahedron.

The fundamental mechanism of the downhill simplex method is the following: The algorithm starts from an initial simplex. Each step of the method consists in an update of the current simplex. These updates are carried out using four operations: reflection, expansion, contraction, and multiple contractions. To be more precise, by successive iterations, the process involves the evaluation of the simplex point where the objective function is maximal in order to substitute this by reflection with respect of gravity centre of N other points. If the value of the objective function is lower than the value of other points, the simplex is expanded in this direction. If the value of the objective function is higher, the local minimum function is supposed to be closer , and the simplex is reduced. Figure 14 shows the four fundamental operations on a 3-simplex (two dimensions plus one dimension for objective function), where x1 is the simplex point corresponding to the maximum value of the function.

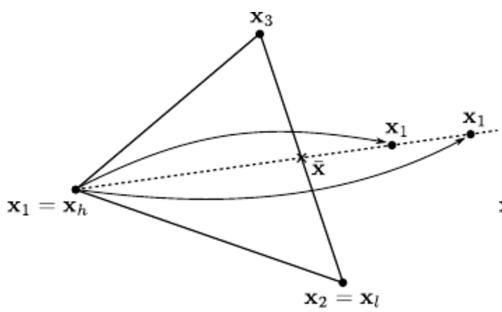


Figure 16: Illustration in 2 dimensions of the four fundamental operations applied to the current simplex by the downhill simplex method: (a) reflection and expansion, (b) contraction, and (c) multiple contractions.

Let f: R^n -> R, be the function to minimize and let $\{x_0, \ldots, x_n\}$ be the current simplex (xi ∈ R^n ∈ Rn for all i ∈ [0, n]). Let h ∈ [0, n] be the index of the `worst vertex', *i.e.* the value *h* = arg maxi f(xi) and l ∈ [0, n]) be the index of the `worst vertex', *i.e.* the value l = arg min f(xi). The downhill simplex method is detailed in the following algorithm.

Algorithm 1: Downhill simplex of [Nelder and Mead, 1965].

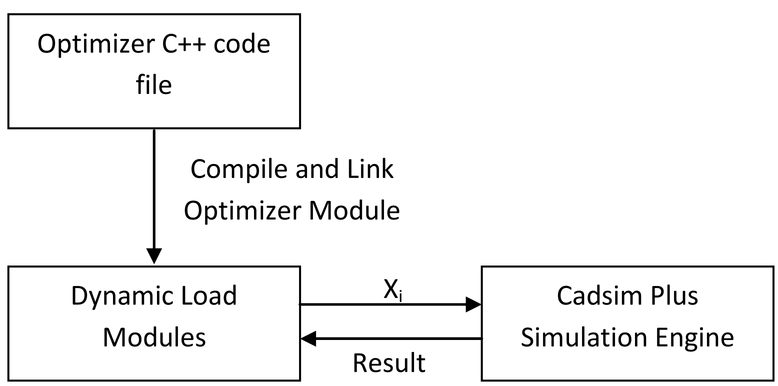
```
input  : the cost function f : ℝ^n → ℝ
         {x_i}_{i=0}^n an initial simplex
output: x*, a local minimum of the cost function f.
begin
    k ← 0 ;
    while STOP-CRIT and (k < k_max) do
        h ← arg max_i f(x_i) ;
        l ← arg min_i f(x_i) ;
        x' ← (1 + α)x̄ − αx_h ;
            where α > 0 is the reflection coefficient ;
        if f(x') < f(x_l) then
            x'' ← (1 + γ)x' − γx̄ ;
                where γ > 1 is the expansion coefficient ;
            if f(x'') < f(x_l) then
                x_h ← x'' ;                              /* expansion */
            else
                x_h ← x' ;                               /* reflection */
        else if f(x') > f(x_i), ∀i ≠ h then
            if f(x') ≤ f(x_h) then
                x_h ← x' ;                               /* reflection */
            x'' ← βx_h + (1 − β)x̄ ;
                where 0 < β < 1 is the contraction coefficient ;
            if f(x'') > f(x_h) then
                x_i ← (x_i + x_l)/2   ∀i ∈ [0, n] ;      /* multiple contraction */
            else
                x_h ← x'' ;                              /* contraction */
        else
            x_h ← x' ;                                   /* reflexion */
        k ← k + 1 ;
    return x_l
end
```

The stopping criterion is defined by:

$$\sqrt{\frac{1}{n+1}\sum_{i=0}^{n}\left(f(\mathbf{x}_i)-\overline{f(\mathbf{x}_i)}\right)^2} \leq \varepsilon,$$

With f(xi) the average of the values, and ε a predefined constant. This criterion has the advantage of linking the size of the simplex with an approximation of the local curvature of the objective function.

As can be seen in figure below, implementation of the optimizer is achieved with the use of a programmatic interface enabling the design of custom computational modules to be executed within the commercial software, CADSIM Plus (Wasik, 2002).



Integration of optimizer module into CADSIM Plus simulation software

Figure 15: Integration of optimizer module into CADSIM Plus simulator software

Advantages of the solver approach

- General use.
- Simplicity.
- Efficiency for non differentiable function.
- Geometrical interpretation.
- Certainty for a set of decreasing values.

Inconvenient of the solver approach

- Difficulty if the minimum is close to a border.
- Arbitrary choice of initial simplex.
- Decrease in performance when number of dimensions increases.

Production cost assessment

The third pillar of the overall method is operations-driven cost modeling framework for product costing. A three-step analysis sequence is applied in the development of the cost models of the current and retrofit design alternatives:

- Development of the base case model,
- Validation of the base case model using process data and financial statements and reports, and
- Development of the cost models of the retrofit cases.

Steady-state process simulation models – either existing process simulation models of the current facility, or new/modified models based on the flowsheet synthesis – provide the cost models with resource and activity drivers that are based on the mass and energy balances. The integrated designs' mass and energy balances and flowsheets provide utility demands and constraint information of the current systems needed in the retrofit scenarios.

The mass and energy flows between defined cost model activities represented by the steady-state simulation model are transferred to the cost models using driver and production rate tables. These cost models were developed in this work using Impact: EDC™ from 3C Software Inc[1].

The two inputs into the cost model are the resource consumption rates provided by reconciled process data and cost items (resource unit prices). The cost items are divided into direct and indirect costs, where for the later one a table of allocation rules is specified. The following section defines various financial inputs into cost model in more details. The cost model presentation is then finalized in the second section by presenting the functionality and calculation engine of the model.

[1] http://www.3csoftware.com

Cost items definitions

The first part of cost-model development is to characterize the direct and indirect manufacturing costs of the studied facility. The steady-state process simulation model and real-time data are used to define resource and activity drivers of the cost model for current business and for the forest biorefinery options.

The cost items are divided into two categories:

- Direct costs (rate of consumption is tracked by process data), and
- Indirect costs (rate of consumption is calculated based on allocation rules)

Most of the indirect costs allocation rules are chosen to be on the basis of head count (number of direct labor per departments). The only difference in the basis of indirect cost allocation is a maintenance cost (maintenance labor and material). This was done in order to associate maintenance cost to departments that require more maintenance hours than others. It must be mentioned here that changes to indirect cost allocation rules will not result in a change in the total indirect cost of product. These rules of allocation were chosen from discussions and practices of current accounting procedures at the case mill. This way the communication and validation process of cost results with mill accountant was improved. It is recognized that there are several other allocation options for indirect costs, for instance based on machine hours, direct labor hours, or indirect cost items could be not included in the cost model. The choice of allocation basis does not influence the goal of this work which is to analyze the performance of the production processes within different operating practices.

Table 3: List of cost items input into the cost model

Bleaching Aid	Maintenance labour
Retention Aid	Wages -Misc.
Machine fabrics	Training
Machine clothing	Maintenance materials
Wrappers and Heads	Operating Supplies
Cores	Major maintenance
Finishing/shipping supplies	Salaries
Production supplies	Benefits
Refiner plates	Company pension
Sludge disposal	Insurance
Chemicals -Effluent	Property taxes
Other chemicals	Salary contingency
Oil for Steam	Variable pay
Gas for steam	Process control
Electricity for steam	Professional Services
Electricity	General Overhead
Other energy	

The use of all mill indirect cost items were chosen in order to represent the production costs on the same basis as it is currently being practiced at the mill for validation purposes. The selection depends on the analyst's judgment and hence some of the cost items could be omitted if desired. For instance, such a situation can occur if the direct cost is analyzed only in relation to process performance. On the other hand the use of indirect cost allocation points indirectly to the capacity utilization of the process. This way the difference between running operating conditions and shut downs of the production (or part of it) can be understood, i.e. the portion of indirect costs when operating at full capacity is smaller than when during shutdown of TMP line.

In summary the use of indirect items better represent the actual production costs with the ability, for instance to track maintenance costs to different departments based on actual maintenance hours. The indirect costs are calculated in the process work center that is called Overhead Work Center. Figure 16 represents the classification of cost items used in this work and their relation to process work centers at the case mill.

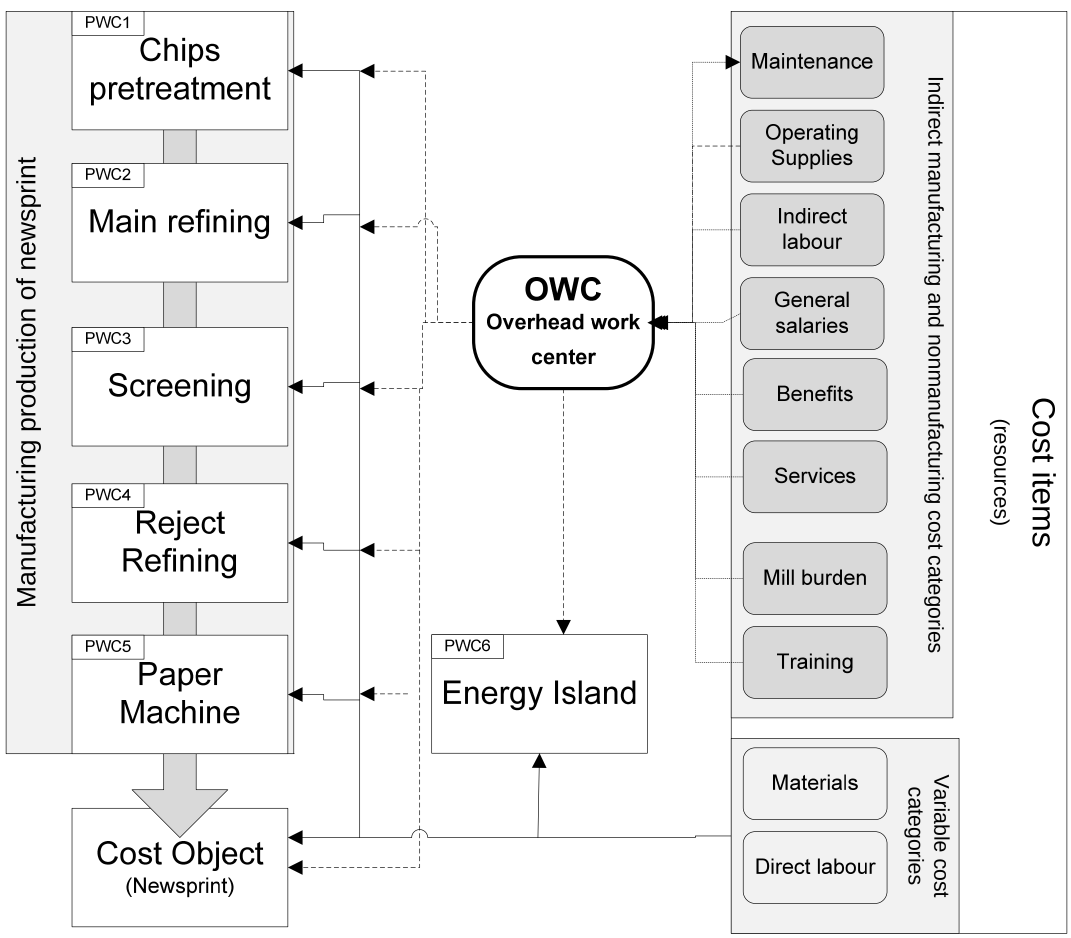


Figure 16: Definition of the process and overhead work centers which capture various cost categories within the current business base-case mill (from thesis page 74).

Operations-driven cost model

The costing framework in this work is inspired by the activity based costing principles. Many of the resource drivers and cost object activity-drivers (activity consumption by cost objects) in chemical processes are based on continuous material flows. The ABC-type definition of them can be defined similar to the classical ABC accounting:

- In classical ABC accounting the main focus is on improving indirect costs tracking in order to capture different levels of activities requirements for different types of cost objects.
- In the operations-driven costing, the ABC-like approach is an analogy to indirect cost tracking (by the classical ABC principles) with the direct costs. The activity in the processing plants is defined by two key aspects: the process design (material and energy balances) and the operating condition (actual values of resource consumption), i.e. the same processing units will result in different rate of resource consumption when operating conditions change.
- The results of an operations-driven cost model are product costs and activity costs. These can be further used in performance analysis of the production by identifying more profitable regimes of operation.

Cost assessment within process work centers:

Various drivers used in the cost calculation come from reconciled process data (resource drivers for direct costs) or from allocation table (indirect costs drivers). The general equations for calculating costs within each process work center can be summed up as follows:

$$\text{Product cost}^{(\$/t \text{ or } \$/hr)} = \sum_1^6 PWC_i$$

where PWC_i is cost of a process work center, and it is calculated

$$PWC_i = DC_{PWC_i} + OH_{PWC_i}$$

where DC and OH are direct costs and overhead costs respectively

$$DC_{PWCi} = \sum_{i=1}^{n} RD_i UC_i$$

$$OH_{PWCi} = \sum_{j=1}^{m} \frac{OH_j}{AB_j}$$

Where RD_i and UC_i are resource driver and unit price of resource for given direct cost respectively; OH_j and AB_j are an over head item (indirect cost item) and its allocation rule respectively.

An example of cost calculation of a sub activity (primary refiner) within a process work center is illustrated on Figure 17 The resource driver is the electricity consumption and electricity load with their associated unit costs. The results of this calculation is the cost of electricity consumed by primary refinery for a given production regime.

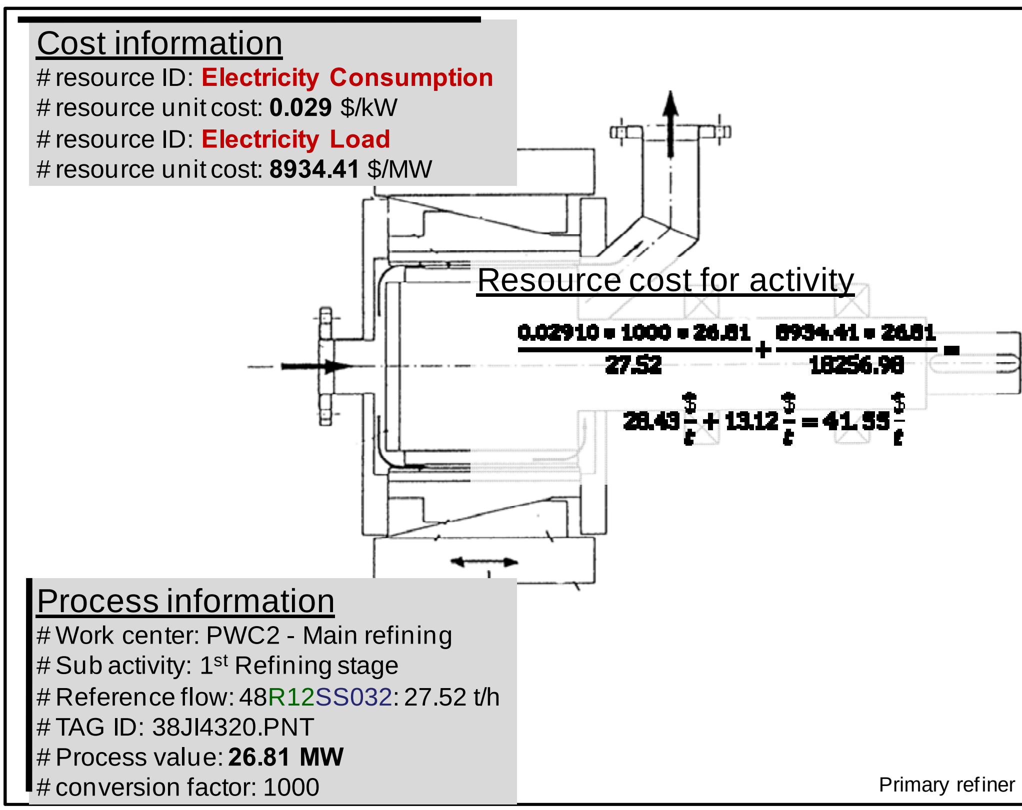


Figure 17: Illustrative example of cost calculation (electricity) for an activity - Primary refining

Decision making

The characterization of operating conditions by cleansed real-time processed data is used to enhance instrumentation and process troubleshooting:

1. Operational decision making based on near steady-state data occurrence:
    a. Piles of preprocessed data sets corresponding to a near steady-state condition can be used to characterize the duration of stable and unstable operating conditions that are caused by external perturbations (change in raw material characteristics). This characterization can enhance the understanding of chip quality impact into operating performance, i.e. how often these perturbations occur and what are the associated causes of change in quality of raw materials.
2. Operational decision making driven by reconciled steady-state data sets
    a. Reconciled near steady-state data sets are used to identify biased instruments. This process helps with the instrumentation maintenance to calibrate individual sensors to measure accurately. The use of reconciled data sets can be exploited to identify unwanted "product leaks" within process activities if they occur.
3. Operational decision-making based on cost of operating regimes:
   The assessment and interpretation of operating costs for individual operating regimes can be used to identify costly regimes. This unique knowledge provides an opportunity to avoid producing in this costly operation if possible. Thus several operational decisions can be made:
    a. Identification of profitable operating regimes: The need to change operating parameters due to shift in planned internal product quality requirements (pulp freeness).
    b. Process troubleshooting: Manufacturing cost of a given operating regime is set as a benchmark. When production cost is significantly changed for the same operating parameters, the variance can be analyzed and interpreted from a process perspective and thus indentify process problems.
4. Strategic decision making for future operation performance evaluation

a. The cost model for assessing regime operating costs can be used to compare operating costs of retrofit scenarios and thus help selecting the best option and best operating practice in the new business strategy.
b. The use of EBITDA forecast of the new business strategy provides new decision-making insights by amplifying the difference in cash flow due to variance in cost of operating regimes.

5. Operational decision making driven by cost distribution of operating regime
    a. The cost distribution for a given product within accounting period can enhance understanding of cost variances due to operating efficiency.
    b. Continuous improvements: The cost analysis of operating regimes allows identifying the best operating practices for given product. The best operating practice can be used as a method that has consistently shown results superior to those achieved with other means, and that is used as a benchmark. In addition, a "best" practice can evolve to become better as further improvements are discovered. Best practice is a feature of accredited management standards such as ISO 9000 and ISO 14001[3]

APPENDIX J –METHOD VALIDATION AND TESTING

The process of testing of the overall method consists of four parts: process simulation, signal processing, data reconciliation and cost model testing. Individual tests are essential parts of the model development process since every model only approximates reality in the face of scientific uncertainties.

Simulation model testing

The validity of the developed process simulation model has been tested during the course of this work. The simulation was built in two commercial software packages: WinGEMS and CADSIM Plus.

The state-of-the-art methodology to develop a valid simulation should include following steps[2]:

- Formulate the problem, including objectives
- Interview appropriate subject matter experts
- Interact with the decision-maker on a regular basis - to ensure that the correct problem is being solved and to promote simulation credibility
- Validate components of the simulation – using quantitative techniques

Document the conceptual model – critical for current and future applications of the simulation

- Perform a structured walk-through of the conceptual model – for a nonexistent system; this may be the single most-important validation technique
- Perform sensitivity analyses to determine important simulation factors and risks
- Validate simulation results – analyzing simulation output data using various techniques

[2] http://vva.msco.mil - Verification, validation and accreditation Recommended Practices Guide (The purpose of VV&A is to assure development of correct and valid simulations and to provide simulation users with sufficient information to determine if the simulation can meet their needs.)

The following steps for the validation of the TMP simulation model have been carried out: description of the objectives; meetings on a regular basis with the production engineer, the control engineer and a process engineer of the mill to check principles of operation, process flow diagrams, unit operations, equations and calibration of parameters, specifications and results from simulation. Sensitivity analysis has been performed with engineers even though this is not reported in a formal document. Several mill engineers reviewed simulation output for reasonableness. In the near future, the complete documentation about simulation model will be provided to the mill with the technology transfer process.

The use of statistical tests

A number of statistical tests have been suggested in the validation literature for comparing the output data from a stochastic simulation with those from the corresponding real-world system [Shannon, 1975, p. 208]. However, the comparison is not as simple as it might appear, since the output processes of almost all real-world systems and simulations are non-stationary (the distributions of the successive observations change over time) and auto-correlated (the observations in the process are correlated with each other). Thus, classical statistical tests based on independent, identically distributed observations are not directly applicable. Furthermore, it is questionable whether hypothesis tests, as compared with constructing confidence intervals for differences, are even the appropriate statistical approach. Since the simulation only approximates the actual system, a null hypothesis that the system and simulation are the "same" is clearly false. It is more useful to ask whether or not the differences between the system and the simulation are significant enough to affect any conclusions derived from the simulation.

Several statistical tests could however be performed, based on the analysis of variance and concepts of quality control.

A documented statistical test to check the validity of the TMP model and confirm the literature (Bagajewitcz, 2000) about the relevancy of identifying steady states for data reconciliation is proposed here.

This test would involve following steps:

- Register data from PI system, one set of data corresponding to one set of operation conditions for one period. Filter and average each set of data.
- Evaluate the standard deviations of relative errors between measured and simulated variables. Evaluate a function resulting from the sum of all the variances.
- Check model relations linking variables with higher standard deviations. Modify eventually these relations. Reevaluate the function after the change.
- Compare the function value with the result from reconciliation of data selected at steady state. Function value resulting from data selected at steady state should be lower than that resulting from data average for a set of operation conditions.

An informal test to check the relevancy of identifying steady states and selecting the corresponding data for reconciliation, instead of just averaging variables linked through non linear equations and tank hold-ups, has been performed during this study, although this has not been documented.

Signal processing

Several tests have been done informally during the execution of this work in order to analyze the method's ability to de-noise signal as well as to identify near steady-state conditions.

Optimal scale of wavelet decomposition

The performance of denoising by wavelet transform can be assessed by calculating the mean-square error (MSE) between the raw data and the processed signal. Since the process signal is a representation by wavelet (the inverse of detailed signal over the wavelet coefficients – discussion on the subject is in Appendix A, page 121-123), the MSE will be decreased with the increase in scale of decomposition. However, at a certain scale this process will remove not only noise but also trend features of the signal which will clearly increase the value of MSE. Hence the optimal scale can be found mathematically by selecting the first minimum of MSE reached. Mathematically, the different steps to find the optimal scale are as follow (Benqlilou, 2004):

1. The scaling and wavelet coefficients, $u_{l,v}$ and $w_{l,v}$ respectively, at various scales l and dilatation v are obtained by taking of the DWT over $y(t)$.
2. The *approximation* component, A_l, at each scale l is reconstructed through:

$$A_l = \sum_{l=1}^{L} u_{l,v}\Phi_{l,v}(t) \qquad (1)$$

3. The power P_l contained in the difference between $y(t)$ and A_l at each scale l (da_l) is calculated through, Eq. (2).

$$P_l(da) = \sum_{l=1}^{L} |y(t) - A_l|^2 = \sum_{l=1}^{L} |da_l(t)|^2 \qquad (2)$$

4. The variation of power, Eq. (2), between successive scales is computed using Eq. (3). As the dyadic scale increases the power due to the noise calculated in Eq. (2), decreases rapidly until it reaches a first minimum. The optimal scale L_m corresponds to the first minimum encountered

$$\Delta P = P_l(da) - P_{l-1}(da) \qquad (3)$$

5. At this scale L_m a first thresholding based on setting to zero all the $w_{l,v}$ greater than L_m is performed. Then, a second thresholding over the remaining coefficients is performed through Visushrink methods reported by Nounou and Bakshi (1999).
6. The de-noised signal is obtained by taking the inverse of DWT as shown in Eq. (4).

$$\hat{y}(t) = \sum_{v=-\infty}^{\infty} u_{L_m,v}\Phi_{L_m,v}(t) + \sum_{l=1}^{L_m}\sum_{v=-\infty}^{\infty} w^{*}_{l,v} \cdot \Psi_{l,v}(t) \qquad (4)$$

In the course of this work the tests of choosing different wavelet scale were not conducted. It was assumed that the best performance of data reconciliation step will be achieved if optimal scale is selected for each variable. However in future work, the following procedure can analyze the impact of scale selection on reconciliation results:

1. Create a hypothetical data set of with known random, systematic errors and known steady-state periods.

2. For given set of independent variables perform wavelet transform with an optimal scale choice (Selected according to the discussion above)
3. Analyze the set of variables for steady-state occurrence and evaluate the performance by Error I and II testing (discussion in Appendix A of thesis).
4. Evaluate the impact on data reconciliation outcomes by comparing the size of error from true values.
5. Repeat points 2-4 with different scale of wavelet transform for each independent variable.

Data reconciliation

Testing of the data reconciliation model involves sensitivity and scenario analysis of critical parameters used in the model. A change in the value of certain parameters of the reconciliation procedure affects how the variation (uncertainty) in the output of the model can be attributed to different changes in the inputs of the model. The focus of test analysis of the reconciliation model can be divided into three categories:

1. *The impact of model parameters on reconciliation output*
2. *The impact of independent variables choice*
3. *The impact of weighting choices*

More details on how this must be addressed is given in the following sections:

The impact of model parameters on reconciliation output

Various model parameters must be tested in order to assure the overall model validity when performing data reconciliation to represent an operating regime as close to reality as possible. Parameters such as ratio of steam generation from high consistency refiners to the electricity load, were analyzed for each operating regime during within this study. The design test for statistically summarizing parameter validation over each operating regime is envisioned as a five-step procedure:

1. Select a single data set representing near steady-state operation within the regime.
2. Select a value of the parameter from its range (for instance, real values of ratio of steam to electricity load are in the range of 0.8-1.2).
3. Perform a data reconciliation and plot normalized mean square error.

4. Repeat steps 1 through 3.
5. Detect which value of the selected parameter has the lowest computed NMSE value and which parameters have NMSE values that are significantly different from the optimal value.

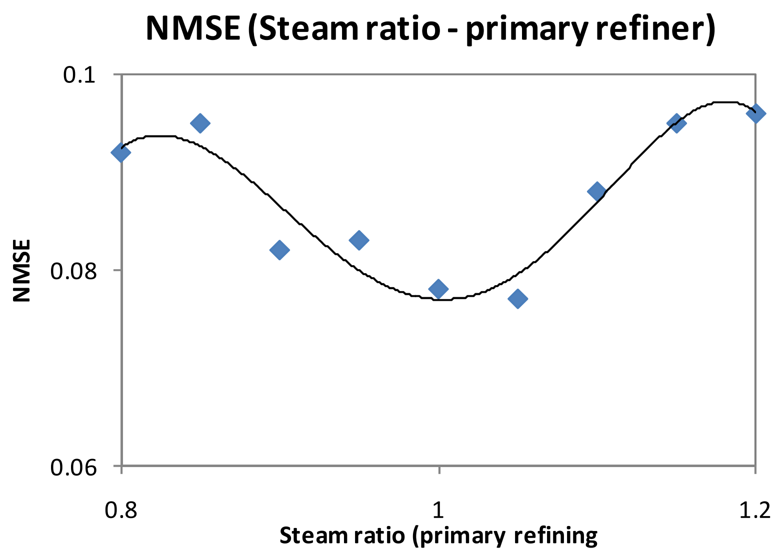


Figure 18: sensitivity analysis of model parameters - Steam ratio in primary refiner

Figure 1 shows an output of a single parameter sensitivity analysis to the reconciliation outcome. The value of the parameter that yields to the lowest NMSE is selected. In the case presented in Figure 6.1 for a given operating regime, the ratio of steam generation in primary refinery was set to 1.03. Various parameters that are specific to the processing units in the simulation model were similarly analyzed for different operating regimes. Hence it can be assumed that this process has ensured the validation of the reconciliation model. Additional tests could be performed in order to select the best reconciliation strategies.

Scenario testing of independent variable choice

As discussed in the model presentation section of this document, selection of subset of measured free variables is carried out by an analyst, and utilizing the expertise concerning the underlying process and monitored variables. In the course of this work a fixed set of measured variables were associated to the model to create a maximum redundancy of the system as possible. The process of variable selection was however limited to the lack of measurements and thus many of the sensor associations could not be altered to perform sensitivity test with experimental results. For future work, the following test of the choice of variables is presented:

1. Create a hypothetical data set that represent a real steady-state operating condition.
2. For a given set of measurement variables, identify independent variables based on the accuracy of the instruments.
3. Perform a data reconciliation step and evaluate results by the least square value and relative error measure.
    a. Since the output of reconciliation can vary, this step will should be performed several times for assuring adequate justification of results on statistical basis.
    b. The results can be presented on several critical variables (significant from a cost perspective – see Table 1 bellow), i.e. production volume, steam flow, electricity consumption.
4. Set different independent variables based on engineering judgment and repeat tasks 2-3.
5. Identify the best performance option of the reconciliation model with corresponding choice of independent variables and interpret/justify the outcomes.

Table 4: The relative error improvement by data reconciliation from the first choice of independent variables (based on instruments accuracies)

	True value	Measured value	Reconciled value	Relative error (measurement)	Relative error (reconciled)
Production (Primary refiner) t/hr	689	631	685	8.42%	0.58%
Production (Reject refiner) t/hr	376	340	375	9.57%	0.27%
Specific energy (primary refiner) kwhr/t	970	940	969	3.09%	0.10%
Stem flow (reject refiner) t/hr	9.43	11.45	9.51	21.42%	0.85%

The impact of weighting choices

The choice in measured variables comes down to the analyst's judgment of the process and instrumentation network. The expert of the system specifies the best (to his knowledge) individual values. The selection of initial weights to represent uncertainty in individual measurement was based on calculating of coefficient of variation. This coefficient is the standard deviation normalized to the mean (σ/μ) which reduces the possibility of inputs that take on large values were given undue importance (Cullen and Frey, 1999). This process has served as preliminary screening tools without additional model iterations while indicating proportionate contributions to output uncertainty. During the course of this work, these values were adjusted as a learning experience by understanding the process operation. Currently, the final matrix of weights reflects the uncertainty in each instrument reading, as well as the uncertainty caused by the natural process dynamics.

To support the choice of weights, a sensitivity analysis was carried out on every measurement. For illustration purpose the results are presented in the tables (2 to 5). The sensitivity analysis of the weighting choice has followed a four-step procedure:

1. Selection of a single data set representing near steady-state operating conditions of a production regime.
2. For given set of measurement variables, define a weighting matrix based on the accuracy of the instruments (classical selection of members of covariance matrix)
3. Performance of a data reconciliation step and evaluation of results by relative error.
4. Set different values of weights to selected key variable and repeat step 3.

Table 5: Output from data reconciliation process

Variable	Tag Name	Value	Meas.	Recon.	Weight	RE	Contr.	Information
Prod costs	Comp.	**283.5**	-	-	-			
Volume flow	FIC4318		28.37	28.99	0.90	0.022	15.4%	Steam recovery
Volume flow	FIC4114		43.12	43.33	0.95	0.005	3.5%	Boiler steam
Mass Flow	138-221		645.34	614.18	**0.05**	0.051	36.1%	Primary refiner
Mass Flow	138-243		612.98	613.30	0.8	0.001	0.4%	Sec Refiner
Mass Flow	138-261		232.98	229.60	0.65	0.015	10.5%	Reject refiner
Electricity	JI4320		26.67	26.78	0.85	0.004	2.9%	Primary refiner
Electricity	JI4720		16.45	17.09	0.8	0.038	26.8%	Sec Refiner
Electricity	XIC4967		84.70	85.23	0.9	0.006	4.4%	Tertiary refiners

Table 2 summarizes the outcomes of data reconciliation with selected weighting values for each measurement. Sensitivity analysis reveals how the accuracy of some estimates could be improved. For instance, table 2 shows the sensitivity analysis results for the production costs calculation in the pulping line (the presented value is only an illustration of an actual value). The first line in the table reports the measured value, and reconciled accuracy of this variable. The next rows in the table identify the measurements that have a significant influence on the validated value of production costs. For instance, 36.1% of the uncertainty on production costs comes from the uncertainty of variable 138-221 – mass flow of pulp prior to primary refining stage. The contribution of uncertainty from each variable is based on relative error for sensitivity analysis of weighting choice. In order to capture the actual contribution of individual variable to cost uncertainty, the measurement variance σ due to process dynamics must be included (Figure 4).

In Table 3, the sensitivity output from varying trust in the pulp mass flow measure, is summarized and presented for three variables. The sensitivity is illustrated as an impact of weight selection to the level of spread of uncertainty across other variables. For instance the change of weight from 0.05 (wrong measurement) to 0.5 would increase the uncertainty of steam measure (flow FIC4318) among others that aren't presented in the table. However, the true sensitivity of variable must include the impact of process variability as well as cost sensitivity. This is discussed and presented in a similar example in the cost model testing section (section 4).

Table 6: Weighting choice and the sensitivity to the output uncertainty

Variable	Tag Name	Value	Meas.	Recon.	Weight	RE	Contr.	Info
Prod costs	Comp.	**283.467**	-	-	-	-	-	
Volume flow	FIC4318		28.37	28.99	0.90	0.022	**15.4%**	Steam recovery
Volume flow	FIC4114		43.12	43.33	0.95	0.005	**3.4%**	Boiler steam
Mass Flow	138-221		645.34	614.18	**0.05**	**0.051**	**36.3%**	Primary refiner

Variable	Tag Name	Value	Meas.	Recon.	Weight	RE	Contr.	Info
Prod costs	Comp.	**283.467**	-	-	-	-	-	
Volume flow	FIC4318		28.37	28.99	0.90	0.022	**19.4%**	Steam recovery
Volume flow	FIC4114		43.12	43.33	0.95	0.005	**4.3%**	Boiler steam
Mass Flow	138-221		645.34	631.22	**0.50**	**0.022**	**20.0%**	Primary refiner

Variable	Tag Name	Value	Meas.	Recon.	Weight	RE	Contr.	Info
Prod costs	Comp.	**283.467**	-	-	-	-	-	
Volume flow	FIC4318		28.37	28.99	0.90	0.022	22.8%	Steam recovery
Volume flow	FIC4114		43.12	43.33	0.95	0.005	5.1%	Boiler steam
Mass Flow	138-221		645.34	642.11	**1.00**	**0.005**	**5.3%**	Primary refiner

The impact of steady-state assumption using average variables

The test of steady state assumption is presented in the thesis (Executive summary section – signal processing, page 64-66 and Appendix A). In summary the following steps were executed:

1. Identify near steady-state data set and
2. Perform data reconciliation and evaluate relative errors
3. Relaxing the steady-state assumption - Set steady-state detection parameters to include process dynamics in the stationary assumption
4. Perform data reconciliation and evaluate relative errors
5. Repeat 3-4 for different values of steady-state parameters
6. Average values of measurements within the analyzed time frame
7. Perform data reconciliation and evaluate relative errors
8. Evaluate the outcomes from different procedures and discuss the impact on cost analysis

However extended analysis could be performed to capture the plant wide operation. Furthermore, the use of different data sets with more dynamic trends than it was used would potentially highlight further the necessity to select a data set that correspond to near steady state condition as close as possible.

Operations driven model

According to the suggested guidance for sensitivity testing of complex models, the analyst should first apply non-intensive analysis methods to understand and select only inputs that create the most sensitivity. The application of more intensive techniques should then be applied to this smaller subset of inputs.

Variable sensitivity analysis

Simple sensitivity analysis of each of the variables used for calculating production costs was carried out by the following procedure:

- Step change in input of analyzed variable
- Analyze the output of the model
- Create sensitivity outcome – ratio of change in input variable to change in model output

The outcome from the analysis is the sensitivity relationship of each measurement variable to the outcome of the model. These functional relationships are linear (Figure 2). For instance a small change in electricity resource (main refining line) will create larger change in the outcome of the model than process or boiler steam resources. On the other hand, a change in production rate (near 20%) will result in the change of 10% in the output (cost of production).

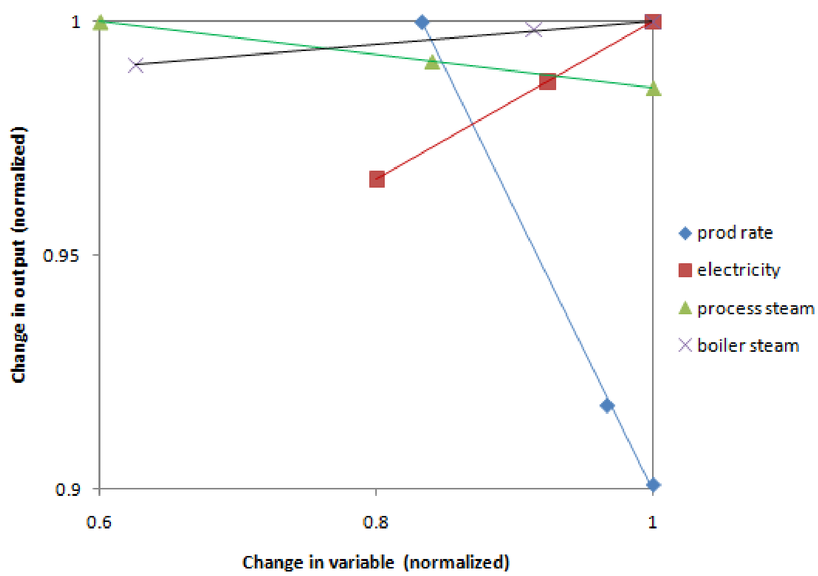


Figure 2: Simple sensitivity analysis of four variables, production rate, electricity (main refining), process steam (outlet of recovery unit) and boiler steam (outlet of boilers)

The second step of the sensitivity analysis of the overall method involves the combination of uncertainties due to several types of errors (due to measurement procedures, process dynamics, external causes, etc):

- σ_{WT} of each measurement represents variation from steady state due to process dynamic. It is expressed as the variance of preprocessed signal by wavelet transform.
- RE – relative error of raw measurement to its reconciled value.
- The total error within each variable can be expressed as a combination of both type of errors and it is calculated as the square root of the sum of squares of both errors:

$$\sigma_{tot} = \sqrt{(\sigma_{WT}^2 + RE^2)}$$

- The final sensitivity value is expressed as a Variance = S.σ_{tot}, which is a multiplication of variable sensitivity with the total error of variable.
- Contribution of variable error to the uncertainty of the outcome can be then calculated for each variable as Variance (i) / Sum (variances).

Table 6.4 summarizes the outcomes of the sensitivity analysis. When compared to the sensitivity analysis of the reconciliation model using only relative error, it can be noticed that electricity resource becomes the most significant. For instance, 46% of the uncertainty on production costs comes from the uncertainty of variable JI4320 (electricity consumption at primary refining).

Table 7: Sensitivity analysis of production costs

Variable	Tag Name	Value	Meas.	Recon.	σ_{WT}	RE	Sen.	σ_{Tot}	Var.	Contr.	Info
Prod. costs	**Comp.**	**283.5**	-	-	-		-	-	-	-	
Volume flow	FIC4318		28.365	28.992	0.67	0.022	0.32	1.09	0.35	6.79%	Steam recovery
Volume flow	FIC4114		43.120	43.330	1.10	0.005	0.22	0.91	0.20	3.90%	Boiler steam
Mass Flow	138-221		645.341	614.178	18.61	0.051	0.41	0.05	0.02	0.43%	Primary refiner
Mass Flow	138-243		612.982	613.298	7.98	0.001	0.40	0.13	0.05	0.98%	Sec Refiner
Mass Flow	138-261		232.980	229.601	3.38	0.015	0.14	0.30	0.04	0.81%	Reject refiner
Electricity	JI4320		26.671	26.781	0.33	0.004	0.79	3.02	2.39	46.56%	Primary refiner
Electricity	JI4720		16.445	17.088	0.24	0.038	0.42	4.15	1.74	34.01%	Sec Refiner
Electricity	XIC4967		84.701	85.228	0.33	0.006	0.11	3.04	0.33	6.52%	Tertiary refiners

Resource price testing

Another test that could be performed is is the sensitivity of resource prices. This may or may not be relevant according to the decisions based on the proposed method. Two levels of decision making are of focus:

1. Operational decision making to evaluate cost-process performance analysis

In this case the selection of price was taken from historical knowledge. Only comparison of process efficiency is made and hence the difference choice would not influence the comparison process.

2. Strategic decision making based on futuristic operational performance

In the case of strategic decision making, the prices of the future resources may play an important role when comparing various processes design alternatives. In the presented results the prices were taken based on RPA (2001-2020)[3]

[3] Renewable Resources Planning (RPA) which include long-term projections for forest product consumption

Additional References

J. A. Nelder and R. Mead, A simplex method for function minimization, *The Computer Journal*, 7: 308-313, 1965

Florent Brunet, Contributions to Parametric Image Registration and 3D Surface Reconstruction (Ph.D. dissertation, November 2010)

Cullen, A.C., and H.C. Frey. 1999. Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs. ed. 326. New York: Plenum Press.