ENG450 –Final Year Engineering Internship

Development and Evaluation of Control Performance Assessment Indices for Alcoa's Advanced Process Control Applications

"A report submitted to the School of Engineering and Energy, Murdoch University in partial fulfilment of the requirements for the degree of Bachelor of Engineering"

Peter Baker
Student Number: 12020064
2008

Academic Supervisor: Parisa Bahri
Industry Supervisor: Ali Nooraii

School of Engineering and Energy
Faculty of Minerals and Energy

Abstract

This report details development and evaluation of potential performance measures for Advanced Process Control (APC) applications implemented across Alcoa sites. The final measure would ideally aid in the diagnosis of poor control and enable comparison between the performances of separate controllers.

In particular, the work has focused on the development of a suitable control performance index for Honeywell's Robust Model Predictive Control Technology (RMPCT – Profit Controller) as implemented on an evaporator process located at Alcoa's Kwinana alumina refinery.

Research in the field of controller performance assessment, particularly the performance of multivariate Model-based Predictive Controllers, was investigated. Existing performance indices proposed in the literature were assessed for their suitability to Alcoa's applications. For the greater part, these methods are not suited to the specific characteristics and functionality of Honeywell RMPCT.

A CPA metric entitled Event Frequency Performance Index (EFPI) is proposed in this report. It is a composite metric comprising five component metrics each of which are designed to gauge different aspects of RMPCT performance. Its stages of development are described and it is applied to seven periods of RMPCT historical data. The metric results are analysed and compared to general expectations about controller performance for these assessment periods in order to determine the utility of the proposed approach.

A historical benchmarking method for performance assessment is also proposed. This involves the identification of a period of controller operation that is known to be good and then comparing subsequent assessment periods to this benchmark. This approach is applied to three different aspects of RMPCT performance: CV limit violation, MV movement and economic optimisation. Performance indices using this method are obtained for six periods of RMPCT historical data.

Acknowledgements

I would like to thank Professor Parisa Bahri of Murdoch University and Dr. Ali Nooraii of Alcoa for providing me with the opportunity to undertake this work and for your support and encouragement.

Thanks also to Stephen Nigli and Peter Rose of Alcoa's QUASAR development team for all your help throughout the course of this project.

Table of Contents

Table of Figures

1 Introduction

1.1 Control Performance Assessment

Controller performance assessment (CPA) aims to evaluate the performance of controllers from routine operating data. It is necessary to ensure effectiveness of process control and consequently safe and profitable plant operation.

The initial design of control systems includes many uncertainties caused by approximations in process models, estimations of disturbance dynamics and magnitudes, and assumptions about operating conditions. These uncertainties can lead to plant performance that may differ significantly from the design specifications. Even if controllers perform well initially, many factors can cause their abrupt or gradual performance deterioration over time.

It is often difficult to effectively monitor the performance and diagnose problems from raw data trends as they tend to show complicated response patterns resulting from the presence of disturbances, noise or non-linearities. CPA is therefore primarily concerned with the development of statistics that are able to measure criteria that have been identified as reflecting aspects of control performance. While the diagnosis and correction of control issues indicated by these statistics may be considered an integral part of CPA, in this report CPA refers only to the application of the indices to gauge control performance.

Effective CPA is also important with regard to appropriate allocation of resources. A plant may have a number of different control assets. Maintaining them based on their respective conditions requires an effective way to determine their performance and prioritise action. In order to enable this comparison between controllers, CPA metrics' upper and lower bounds should indicate the best and worst performance a controller is capable of.

Alcoa does not currently have any standard CPA procedures in place, other than measuring Manipulated Variable (MV) utilization, which indicates the percentage of critical Manipulated Variables (MVs) the controller is using over an assessment period. It is predicted that, without the adoption of effective CPA methods, the performance of Alcoa's Advanced Process Control (APC) applications will be significantly reduced. This is especially the case given the predicted increase in the number of their APC assets.

1.2 State of CPA Research

A number of algorithms for estimating a CPA index are proposed in the literature. The conventional method involves comparing the existing controller to a theoretical benchmark such as the Minimum Variance Controller (MVC).

Harris (1989) laid the theoretical groundwork for CPA of single loop controllers from routine operating data. He proposed a comparison of the output variance term with the minimum achievable variance. Desboroug and Harris (1993) apply this idea to assessing feedback/feedforward control schemes. Harris et al. (1996) and Huang et al. (1997b) applied the generalized the minimum variance benchmark to the multivariate case based on the multivariate interpretation of the delay term, known as the interactor matrix.

Kozub and Garcia (1993) proposed more practical user defined benchmarks based on settling times, and rise times. The settling time or rise time for a process can often be chosen based on process knowledge. A correlation analysis of the operating data is used to determine whether the desired closed loop characteristics were achieved.

Tyler and Morari (1995) proposed a CPA method based on likelihood methods and hypothesis testing. Performance assessment of non-minimum phase and open loop unstable systems was also addressed by Tyler and Morari (1995). Ko and Edgar (2000) addressed the issue of cascade control system performance assessment.

Huang and Shah (1999) proposed the Linear Quadratic Gaussian (LQG) control as the benchmark instead of MVC. This technique also takes input variance into account. The input variance is often of major concern as it is frequently a utility such as steam or power with significant cost. A model of the process and the disturbances is required to do the LQG benchmarking. Kammer et al. (1996) used non-parametric modelling in the frequency domain to ascertain the optimality of a LQG controller, based on the comparison of the optimal and the achieved cost functions.

A constrained Model Predictive Controller (MPC), such as RMPCT, is essentially a non-linear controller, especially when operating at constraints. Conventional MVC benchmarking techniques which rely on linear time-series analysis are therefore infeasible. Patwhardan et al. (1998) attempted to address this issue by using the historical (control) objective function as a practical performance benchmark. This technique has been adapted to assessment of RMPCT in this report.

Ko and Edgar (2001a) propose a benchmark based on the finite horizon MVC derived from closed loop data and knowledge of the order of the delay matrix. This was extended to the constrained MPC case in Ko and Edgar (2001b). While this idea has merit, it relies on accurate data for all of the process's disturbances in order for the benchmark to be realistic. The number of unmeasured disturbances in a typical RMPCT application prohibits this. Accurate process and disturbance models are also required. Any model uncertainty will also result in inaccurate estimation of the benchmark.

Patwhardan et al. (2002) propose a performance metric based on comparison of the designed and achieved MPC objective functions. This method takes into account the structure of the controller along with its design specifications such as the weighting factors associated with different variables. While this approach is attractive, its use for RMPCT assessment is precluded by the fact that the RMPCT control objective function is not obtainable as historized data.

A data-based covariance benchmark is proposed by Yu a and Qin (2001). The scheme uses generalized eigenvalue analysis to extract the directions with degraded or improved control performance against a benchmark period. It was found that application of this method to large multivariable controllers often results in index values so large or small (from 10^{-5} to 10^{8}) that the exact level of performance improvement of degradation is difficult to interpret.

1.3 Honeywell Robust Model Predictive Control Technology (RMPCT – Profit Controller)

Honeywell's Robust Model Predictive Control Technology (RMPCT), or Profit Controller, program controls and optimizes the operation of processes that have significant interaction between variables.

The controller employs a model of the process dynamics in order to explicitly predict future process behaviour and determines the control moves necessary to bring all process variables to setpoints or within constraints. If there are any degrees of freedom remaining to the controller it adjusts the process to optimize operations, for example by maximizing product quality.

Profit Controller, as with Multivariable Model-Predictive Controllers (MPCs) generally, considers an entire process as a single entity rather than as a collection of isolated control loops. As such, it is more appropriate to the control of highly interactive variables than many single loop controllers. Profit Controller is essentially a tool to keep the process within operational restraints while optionally optimizing some performance measure.

The following is an introduction to some of the main features of Profit Controller and those that are deemed to offer some insight into the controller's performance in terms of what to expect from the controller under varying process conditions.

1.3.1 Profit Controller Implementation

RMPCT employs three types of process variables as control input and output:

Controlled Variables (CVs) are variables the controller attempts to keep at setpoint or within an Operator specified range with prioritisation given to maintaining them within their restraints.

Manipulated Variables (MVs) are adjusted by the controller in order to keep CVs within restraints and to optimize the process while not violating restraints placed on the MVs.

Disturbance Variables (DVs) are variables which, although measured, are not under control of the controller but affect the values of CVs. The controller, on the basis of feed-forward information, may predict the future effect of DVs on process response and take action to prevent CV excursions outside constraints before they develop.

RMPCT uses a process model to predict process behaviour. The overall model comprises a matrix of dynamic sub-process models which describe the effect of the MVs and DVs on CVs. Each sub-process is of a generic form that provides a reasonably accurate description of the behaviour of the majority of processes that can be found in processing industries. They contain a number of coefficients whose values determine the dynamic response of the sub-process.

The sub-process models are specified for a given process by determining the coefficient values by model identification which involves open-loop step testing. This is typically done when the controller is first commissioned.

1.3.2 Robustness Features

Profit Controller's *robustness* refers to its ability to maintain good control of highly interactive processes even in the event of significant model error. An understanding of these robustness features impact on what can be expected from Profit Controller's performance. These features include:

> **Range Control Algorithm (RCA)** as opposed to setpoint tracking. Where range control is common, performance measures such as settling time and offset are less applicable than for conventional feedback control loops. While RMPCT allows setpoints to be implemented and changed and therefore, its servo performance assessed, this is not usual under normal operation. Statistically derived measures concerning the violation of restraints and MV movement may be more appropriate.

> **Singular Value Thresholding** (SVT) is employed to correct poor conditioning of the matrix used for control calculations. The controller effectively drops any of the matrix's singular values which are below a specified threshold. This is done in order to desensitise the controller to model error and prevent excessive MV movement. One of the implications of which is that if a controller is Singular Value Thresholding it may result in no MV action being taken despite a CV being outside the desired range. While this may appear to be poor performance, it is in fact appropriate to the controller's objective, i.e. preventing overly aggressive MV movement for little benefit in CV response

1.3.3 CV Characteristics

In a typical Profit Controller process, there is significant interaction between CVs. This means that action taken to change the value of a CV may also change the value of other CVs. The controller must therefore coordinate changes to a number of MVs in order to move a particular CV as desired without causing undesired changes in other CVs.

As in conventional MPC a CV can have a setpoint that defines the desired value for the CV. It is more common, at least in Alcoa's RMPCT applications, that the CV will have a high a high and low limit that define a range of allowable features. This is one of RMPCT's robustness features. The controller will not take corrective control action provided CVs are within their limits which minimises unnecessary MV movement and makes the controller less susceptible to plant-model mismatch.

In addition to these ‘hard’ limits, it is also possible to define soft limits for each CV. These limits, defined as an offset within the hard limits specify the allowable limits for optimisation of the process. They effectively provide a buffer which allows the controller to push the CVs close to restraints while retaining the ability to absorb disturbances without violating those restraints.

CV tracking results in the controller adjusting the external (Operator-set) limit or setpoint and the internal (controller-honoured) violated limit so that there is no CV error on initialisation. The Operator must then return the limit or setpoint to the desired value.

Limit ramping adjusts only the internal, violated limit to the current CV value. The controller then returns the internal limit gradually to the external limit or setpoint. Both CV tracking and limit ramping aim to minimize the initial jolt that can result when CVs exhibit large error when control is initiated.

Limit ramping also applies when the operator makes a large change in a limit or setpoint. It minimizes the disruption by establishing the rate at which the controller moves the old limit towards the new limit.

1.3.4 MV Characteristics

Each MV has a high and low limit which the controller will never violate of. The controller will return the MV to within limits when the controller is started with the MV outside its limits (except when tracking is on) or when the operator changes an MV limit such that the MV value is outside of it.

Rate-of-change limits may also be set in order to prevent excessive MV action when an abnormal event occurs. If these are being hit repeatedly, the limits are possibly being set too small and the controller therefore has less freedom to determine the optimum trajectory.

Limit ramping for MVs determines the minimum rate at which an MV must move towards a violated limit (in the event of initialisation or the Operator changing a limit such that it is violated).

MV weighting is analogous to CV weighting. Greater MV movement weights discourage the movement of particular MVs to resolve CV error. This results in greater movement of larger MVs. When there are more MVs than required in order to meet control objectives, the controller minimizes the sum of the squared changes of the MVs, with each change multiplied by its respective MV weight.

Movement weights do not affect the speed of response or controller stability. Movement weights are only used to set priorities with regard to which MVs it is preferable to move in the event that more than one MV will suffice.

1.3.5 Feedback Performance Ratio

The feedback performance ratio is a tuning parameter defined as the ratio of the closed-loop to open-loop settling times for a CV. The nominal open-loop settling time is the gain-weighted average of the settling times for all of the sup-process models of a given CV. The nominal dead-time is gain-weighted average of the dead times for the CV.

A performance ratio is therefore used to tune controller response. A performance ratio of 1.0 means the CV is returned to zero error within the nominal open-loop settling time, while a ratio of 0.5 means it will be returned to zero error in half that time.

The performance ratio determines the inherent tradeoffs in controller performance that are associated with speed of response, model accuracy and MV movement. That is, a smaller performance ratio results in faster setpoint tracking and disturbance rejection, larger MV movement and higher sensitivity to model error. The converse is also true.

1.3.6 Degrees of Freedom (DOF)

Profit Controller maintains all CVs at setpoint or within range provided there are sufficient DOF to do so. The number of DOF is the number of MVs not at a limit, minus the number of CVs that either have a specified setpoint or are at or outside a limit.

So long as the degrees of freedom are zero or positive CV constraints can be satisfied. If they become negative it is physically impossible to keep setpoints within range.

When there are negative degrees of freedom, Profit Controller attempts to maintain a compromise by minimizing the weighted sum of the squared CV error:

$$minimize \sum_i weight_i^2 \times error_i^2 \qquad (1)$$

where i is the CV index.

In the above formula the error is the scaled CV error. This scaling results in equal increments of different CVs having equal importance on the process. Error trade-off between CVs may be influenced by specifying engineering unit give-ups for each of the CVs. Weights are inversely related to scaling factors and EU give-ups by:

$$weight_i = \frac{1}{(CV\ scaling\ factor) \times \sqrt{(EU\ give-up)}} \qquad (2)$$

The smaller the Engineering Unit (EU) give-up the more the controller attempts to minimize the error for that CV. The EU give-ups are relative to each other. That is, if CV1 has an EU give-up of 3.0 while CV2 has a give-up of 1.0, CV1 will exhibit approximately 3 units of error to every 1 of CV2’s.

EU give-ups have no effect when there are sufficient degrees of freedom to bring CV errors to zero. Further, give-ups do not affect the speed with which the controller corrects CV errors.

1.3.7 Economic Optimization

If the controller has degrees of freedom remaining to it, it is able to optimize an objective function that represents one or several aspects of the process, for example, improvement of product throughput or lower utility costs.

The controller will minimize the objective function (or maximize its negative) subject to keeping all CVs and MVs within limits.

The general form of the objective function is

$$\text{Minimize } J = \sum_i a_i CV_i + \sum_i b_i^{\,2} (CV_i - CV_{0i})^2 + \sum_i c_i MV_i + \sum_i d_i^{\,2} (MV_i - MV_{0i})^2 \tag{3}$$

where a_i and c_i are the linear coefficients of the CVs and MVs respectively, b_i and d_i are the quadratic coefficients of the CVs and MVs and CV_{0i} and MV_{0i} are the desired steady state values of the CVs and MVs.

1.4 Case Study Controller – SLAC

The evaporation area of Alcoa’s Kwinana refinery has the process objective of concentrating the Spent Liquor (SL) from the precipitation area before returning to the Digestion Feed Tanks. This is achieved by heating the SL in shell and tube heat exchangers and then flashing off water vapour by dropping the temperature and pressure in a series of flash tanks.

Evaporation Optimisation application, also known as the Spent Liquor Advanced Controller or SLAC, aims at managing the levels of the spent liquor stock tanks that feed into the evaporation units whilst optimising the evaporation building.

SLAC is the Profit Controller application that has been selected for this study and development of possible CPA methods. The first objective of the controller is to maintain safe operating conditions in the evaporation units. Constraints have therefore been included in the controller design to ensure that the operating pressures and tank levels are within safe limits. The second control objective is to maximise the total evaporation rate of the building, thereby increasing the caustic recovery, reducing refinery costs and increasing production. The third objective is to control the stock tank levels to ensure liquor stocks are balanced to maximise liquor circuit flow.

Prioritising these control objectives ensures that safe operation of the evaporation process is not compromised by the controller. The evaporation process is thus prevented from reaching safety override trip settings that would cause undesirable flow cuts.

SLAC is a large controller, even by Advanced Process Control standards. 87 CVs, 27 MVs and 16 DVs in total are used in the application.

2 Event Frequency Performance Index (EFPI)

A comprehensive Control Performance Assessment (CPA) procedure would ideally incorporate several methods that reflect different aspects of control performance. The goal of creating the EFPI is the development of a metric which combines several component metrics, each of which measures a different aspect of RMPCT performance and therefore provides a general indication of how well, or poorly the controller is performing.

The name, Event Frequency Performance Index, comes from the fact that each of the component metrics measures the average frequency of certain events, or the time the controller spends in certain states. This approach was predicted to have several advantages, not least of which is mathematical simplicity.

Also, each individual metric is normalized based not only on time, but also the controller parameters, such as number of variables and limit values. It is therefore hoped that the metric can be applied consistently to different controllers without the need for scaling, as the metrics are already scaled using the intrinsic characteristics of the controller.

Six aspects of control performance are measured by the EFPI. The individual, component metrics were initially defined as follows:

1. **Constraint Ratio (CR) –** This measures how the controller uses its capacity to add value. At each interval, all MVs are checked to see whether they are at a constraint and given a value of ‘1’ if they are and ‘0’ if not. The scores are averaged over time and the resulting values for each MV are then summed.

   Similarly each CV is checked to see whether it is at a soft constraint. In the absence of historized data for the CV soft limits, a value of 5% of the CVs operating range from the hard limits was used. If the CV read value for an interval is within this range without violating the hard restraint a value of ‘1’ is assigned for that CV at that interval. Otherwise a value of ‘0’ is assigned. The result for each CV is averaged over the sampling interval and the CV results are summed.

   A normalized result for the CR is then obtained by the following calculation:

$$CR = \frac{CV_C}{CV_{C,Max}.MV_{C,Max}}(MV_{C,Max} - MV_C) \qquad (4)$$

Where CV_C and MV_C are the average number of CVs and MVs respectively hitting a constraint per interval over the assessment period and $CV_{C,Max}$ and $MV_{C,Max}$ are the total number of CVs and MVs that could be hitting a constraint at a given interval. The metric therefore penalizes for MVs hitting constraints and rewards CVs at constraints.

The metric ostensibly penalizes those controllers that are not optimizing to constraints or who are not optimizing at all. The measure is based on the assumption that a controller is at its most useful when only CV constraints are being hit. It may be useful in the diagnosis of problems arising from operators setting MV constraints too narrow and thereby limiting a controller's capacity to push CVs to optimal operating points.

2. **Economic Movement Index (EMI) –** This metric aims to measure how necessary the controller is to economic unit operation. It is defined by mapping the economic objective function to the controller MVs. The relevant MVs are identified by whether they possess a linear/quadratic economic coefficient or a non-null sub-process relationship to CVs with a linear/quadratic economic coefficient.

   These MVs are checked for a non-zero gradient at each time interval. EMI is then defined as the time-averaged ratio of those MVs that have a non-zero gradient to the total number of MVs.
   This component is based on the assumption that a controller is more economic if all MVs are pushing in an economic direction and will ideally penalize those controllers that are not used to optimize operation or that only partially use MVs.

   It may be that this component also enables inference about the degrees of freedom (DOF) available to the controller. A controller may generally be considered to be performing well in this regard if it has DOF > 0 as it has the capability to correct for disturbances. If the controller is optimizing it indicates that this is the case.

3. **Objective Function Attainment (*OFA*) –** This metric aims to measure how much value the controller is generating. It is defined as the percentage of time the current objective function value is within a certain range of the steady-state objective function. This condition is checked at every interval and if the current objective function is within the desired range of the steady-state value a score of '1' is assigned. A '0' is assigned if it is not. These scores then are averaged for the assessment period

The metric is based on the assumption that a controller is generating more profit if it spends a lot of time at its steady-state objective function value. Initially the range within which the current objective function has to fall, or the *OFA Threshold*, was 5% of the stead-state.

4. **Movement Index (*MI*) –** This attempts to measure how smoothly the controller is operating and therefore decreases with increasing MV movement. It is calculated by measuring the movement of each MV as a percentage of the maximum allowable move at each sampling instant. The maximum allowable move value will depend on the MV direction, so this is ascertained for each sampling instant. A score between '0' to '1' is assigned for each interval and each MV and the result is averaged over the assessment period and all MVs then subtracted from one.

   Gating was implemented such that if a MV is not on for a given interval , that is, it is not being used by the controller at that point, then it is not included in the metric for that interval. This prevents the metric from rewarding the controller for not moving a MV that is not being used for control.

   The metric penalises those controllers that are moving the process around significantly. For the initial EFPI implementation this is the only component metric that does not rely entirely on the frequency of certain events, as it incorporates the magnitude of MV movement as a percentage of the maximum move limits.

   It should be noted that because MI penalises MV movement, while EMI rewards movement of certain MVs, a perfect EFPI score is not possible, even in theory. However, it was believed that those controllers that push towards optimization with a minimum of MV movement may still score highly.

5. **Constraint Adherence Index (*CAI*) –** This measures how well the CV constraints are honoured. It is calculated by taking the average number of constraint violations per CV, per interval occurring over the assessment period, resulting in a value between '0' and '1'. The result is then subtracted from one. The assumption is that a controller that is not keeping the process within defined limits is neither reliable nor safe.

A final, overriding performance factor is controller **Time in Normal (TIN).** The controller parameter, *ControllerMode* is used to determine whether the controller is ON over the assessment period. A value of'1' is assigned if the controller is ON and '0' if it is not. The results are then averaged for the period.

The composite EFPI metric is defined as

$$\mathbf{EFPI} = TIN \times (0.2\ CR + 0.2\ EMI + 0.2\ MI + 0.2\ CAI + 0.2 OFA) \quad (5)$$

While each EFPI factor is given an equal weighting, it may be necessary to individually weight the variables used in the calculations to better reflect the design objectives of a controller. For example, the restraints on a given CV may have been deliberately set such that they are violated frequently. This may have been done intentionally so as to elicit a specific desired behaviour from the controller and process. In order to reflect this design objective, the Reliability of this individual CV could be given a lower weighting than others. This customization will enable better comparability between controllers.

2.1 Initial EFPI Implementation: Results and Discussion

Table 1 First run EFPI results

SLAC EFPI -All CVs/MVs								
Period	CAI	EMI	CR	OFA	MI	TIN	EFPI	EFPI w/o TIN
1/06/2007 - 30/06/2007 (G)	0.908	0.739	0.122	0.984	0.908	0.980	0.718	0.732
1/10/2007-30/10/2007 (R)	0.861	0.664	0.103	0.898	0.935	0.998	0.691	0.692
1/05/2008-30/05/2008 (P)	0.710	0.669	0.138	0.998	0.948	0.957	0.663	0.693

The EFPI metrics as defined in Table 1 were implemented on three periods of historical data for SLAC. These were initially classified as 'Good', 'Reasonable' and 'Poor' periods of controller operation, based on the amount of attention the controller was receiving during these periods, length of time since the controller was commissioned and the 'gut feel' of engineers familiar with the controller.

The results for the first run application of the EFPI, displayed in Table 1, suggest a definite overall degradation in performance between the first period (period G) and the second (R) and between period G and the third period (P). Whether performance has improved or worsened between periods R and P however, depends on whether the controller's Time in Normal (TIN) statistic is included.

The overall EFPI is calculated both with and without TIN as it is debatable whether or not it is really a measure of control performance. Despite the fact that the controller is on for a greater percentage of R than for P, R has a lower average index for the other components. This suggests that the controller has maintained other aspects of control more effectively over period P despite being active less of the time.

Further to consideration of the TIN statistic, it can be observed from the daily component averages for period P shown in Figure 1 that a trough in TIN corresponds to decreases in all other components. This level of interdependence in CPA metrics is undesirable, particularly if the end goal is a composite, 'rolled up' metric, as it results in the repetitive inclusion of certain aspects of performance.

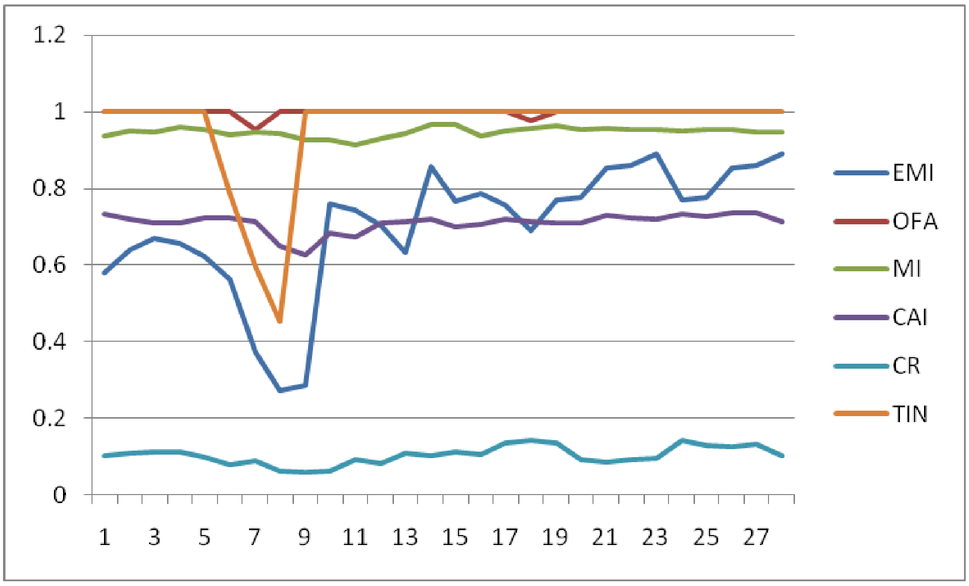


Figure 1 First run EFPI results for 1/05/2008-30/05/2008

For these reasons the TIN component should not be considered as a CPA metric but rather as a potential diagnostic. For example, if several other indicators drop below a specified level and the TIN for that period is also low, it is likely to be a root cause.

This evidences one of the problems with implementing a composite index of this type, that is outlying components can skew the overall metric such that it does not present an accurate picture of control performance. The same may be said of other components: the OFA factor for P is considerably better than for any of its other components, or for those of the other assessment periods.

The above results suggest an apparently inverse relationship between CAI and MI. That is, where the controller is reducing CV constraint violation, MV movement increases. However, this relationship is not supported by inspection of the daily averages obtained for these two components, or calculation of their correlation coefficients shown in Table 2.

Table 2 Correlation coefficients between EFPIs for three assessment periods

Period	Correlation Coefficient for CAI and MI daily figures
1/06/2007 - 30/06/2007 (G)	-0.2672
1/10/2007-30/10/2007 (R)	0.0476
1/05/2008-30/05/2008 (P)	0.4933

While a daily relationship is not supported by these figures, it does not disprove the notion that if the controller is averaging high scores for constraint adherence it is likely to be moving MVs more. In fact, the overall CAI-MI correlation coefficient for daily values for all three periods combined approaches -0.6, suggesting a reasonably strong inverse relationship between the two indices.

This relationship further highlights a key problem with a “rolled-up” metric, that the controller performance can exhibit very different characteristics which are hidden by combining the scores of different indices.

Of further note is the very low scores attained for the CR all three periods. This is less likely to indicate poor control performance than it does the inherent nature of the system being controlled. The CR is defined as

$$CR = \frac{CV_C}{CV_{C,Max}}(1 - \frac{MV_C}{MV_{C,Max}}) \tag{6}$$

Where CV_C and MV_C are the average number of CVs and MVs respectively hitting a constraint per interval over the assessment period and $CV_{C,Max}$ and $MV_{C,Max}$ are the total number of CVs and MVs that could be hitting a constraint at a given interval, in this case 87, the total number of CVs or 27 the total number MVs. The metric therefore penalizes for MVs hitting constraints and rewards CVs at constraints.

It was apparent during the course of calculating this metric that the term concerning MV constraints would have very little impact on the overall score as the average number of MVs at constraints, at each sampling interval, over each assessment period was of the order of 10^{-2}.

An additional problem was in defining what constituted a restrained CV. It was not desirable to use the CV hard constraints as a SLAC CV seldom pushes against a hard constraint without violating it. To reward those CVs that were violating limits would create a number that was the inverse of the CAI. The optimisation limits or delta-soft limits (defining an offset from the hard limits for optimisation) were the preferred values to use. However, it was determined later that many of the values for these from the process data historian were not correct.

It was decided to choose some arbitrary off-set from the hard limits, defined as a percentage of the CV operating range, and if a CV was between this point and the nearest hard limit it was assumed to be restrained.

This solution was not ideal, as those CVs with very large operating ranges would have inordinately large regions where they were assumed to be restrained. Further it was later discovered that not all SLAC CVs have hard limits; the values being used for hard limits were extrapolations by the data historian based on limits that may have once existed. This fact also significantly affected the CAI and was the first point to be addressed when revising the metrics.

Ultimately the low CR scores attained for each assessment period were a result of the fact that very few of the SLAC CVs typically operated close to the limits defined for the metric. The controller was performing very well with respect to MVs not becoming restrained but this is not evident from the scores. This is another example of combining two or more factors into a metric obscuring the true picture of controller performance.

Perhaps one of the most interesting observations for this first application the EFPI was the fact that period P, expected to exhibit the worst control performance, had an OFA score close to perfect. That is, the Current Objective Function value was within 5% of the Steady State Objective Function value for almost the entire one month period.

2.2 EFPI Revision A

A number of initial revisions to the EFPI components were performed. These revisions were primarily concerned with incorporating design knowledge into the CAI. The revised metrics were applied to the original three assessment periods. They were also applied to data obtained for a further three periods which were similarly classified as 'Good', 'Reasonable' or 'Poor'.

The revisions to the CAI were as follows:

Correct Hard Limits – Some of the hard limits initially obtained from historized data did not actually exist, or were different from the correct limits. These were corrected.

Activated Limits – Some limits are activated by other variables. For example certain flowrate limits are activated in the event of valve saturation. This was handled by gating all values when the limits weren't activated in the controller

Deliberate Violation of Limits – Several of the SLAC CVs violate one or both of their limits by design. These CVs have been excluded from the metric or had the deliberately violated CV removed

CV Weighting – The ability to weight CVs has been incorporated into the EFPI program. At this point, all CVs' CAI have a weighting of 1, excepting those that are indicator CVs only or others that are not representative of APC performance in some way. These are given a weight of zero. These zero-weighted CVs have not been excluded all-together as their individual CAI may provide useful information at the CV level, as opposed to controller level.

Spare CVs – These have been removed from the metric altogether.

All the above revisions were also applied to the CV component of CR.

2.2.1 EFPI Revision A: Results

Table 3 EFPI results for Revision A

SLAC EFPI -Revision A								
Period	CAI	EMI	CR	OFA	MI	TIN	EFPI	EFPI w/o TIN
15/04/2007 - 14/05/2007 (G1)	0.973	0.829	0.173	0.964	0.953	0.96	0.747	0.778
1/06/2007 - 30/06/2007 (G)	0.949	0.739	0.13	0.984	0.908	0.981	0.728	0.742
1/08/2007 - 30/08/2007 (R1)	0.94	0.668	0.144	0.897	0.946	0.959	0.689	0.719
1/10/2007-30/10/2008 (R)	0.948	0.665	0.145	0.896	0.935	0.997	0.716	0.718
1/04/2008 - 30/04/2008 (P1)	0.955	0.567	0.159	0.946	0.926	0.79	0.545	0.711
1/05/2008-30/05/2009 (P)	0.958	0.664	0.18	0.998	0.948	0.957	0.718	0.750

The revisions detailed in 2.2 yielded the EFPI results in Table 3. The only metrics affected are the CAI, EMI and the composite metrics. Each of these was improved significantly for the three original assessment periods. Period G remained the best overall performer.

Period P's overall EFPI is now considerably better than period R, which at the time of data collection was expected to be of reasonable performance. The original 'Good', 'Reasonable' and 'Poor' classifications for assessment periods were based on length of time since controller rebuild, the utilization figures, engineer's intuition and the attention the controller was receiving at that point. The classifications were revisited subsequent to obtaining these latest results and it was determined that during period P, controller attention and maintenance had increased significantly and the period should be reclassified as reasonable to good.

The reclassification of assessment periods is much more congruent with the performance indices obtained, the general trend of which is a gradual decrease throughout the 2007 and early 2008 before a significant improvement in May of 2008. The EFPI trend for the assessment periods is shown in Figure 2.

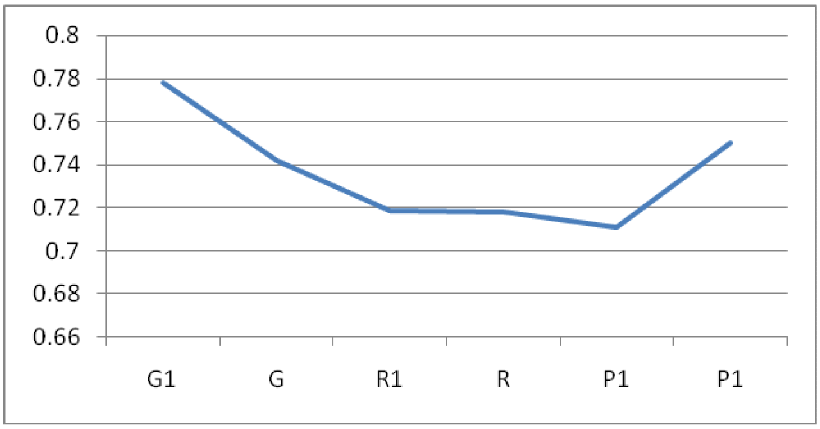


Figure 2 EFPI for six periods of SLAC operation

A correlation analysis was performed on the 5 key EFPI components and the coefficients displayed in the matrix of Table 4. The first point noted was that the inverse relationship between CAI and MI was no longer present. The relationship identified in the first EFPI implementation did not likely exist, as the CAI figures were calculated using limits that did not exist, CVs not used in the controller calculations and bad data.

The most significant relationship suggested by the correlation matrix is between the EMI and CAI. A possible explanation for this may be found in the derivation of the Economic Movement Index. At this stage in the EFPI development it is calculated by checking the MVs that have been mapped to the objective function for a non-zero gradient at each interval. A score of ‘1’ is assigned if the MV is moving and ‘0’ if it is not.

This method is flawed in that an MV will exhibit zero movement only if it has been dropped and is not being used to control the process, otherwise it will exhibit at least some movement, however small. This implies that the EMI as it stands does not measure the economic movement of MVs, but rather the average number of MVs available to control the process. A decrease in this index then, reflects fewer degrees of freedom with which to handle disturbances which may lead to an increased frequency of constraint violation and a poorer CAI.

Although not as strong, EMI also correlates to CR, a component of which measures the average number of CVs at constraints. The degrees of freedom available to the controller also affect its ability to optimise the process, which for the greater number of CVs involves pushing them to a constraint. Thus a lower EMI suggests that the controller may not be able to do this (without violating a hard limit) and therefore incurs a lower CR index.

This weakness in the EMI derivation, along with the fact that the MVs' movement direction and relative impact on optimisation of the objective function are not incorporated, are addressed in EFPI Revision C.

Table 4 EFPI results for Revision A

	CAI	MI	CR	EMI	OFA
CAI	1				
MI	0.193069	1			
CR	0.347267	0.480067	1		
EMI	0.522823	0.306454	0.472541	1	
OFA	0.166984	0.104438	0.14414	0.15399	1

Table 5 EFPI Correlation coefficients for Revision A

The second highest correlation, between component metrics MI and CR, was also not strong and has been treated as coincidental.

2.3 EFPI Revision B

2.3.1 Potential Methods for Incorporation of Constraint Violation Magnitude into CAI

Up until this point, all the component metrics comprising the EFPI, with the exception of the MI, were based entirely on the frequency of certain defined events occurring over the assessment periods, for example, constraint violation or the Objective Function being within a certain threshold of its steady-state value. It was desired to revise the CAI such that it not only measured the frequency of constraint violation, but also incorporated the magnitude of each violation.

This significantly increased the complexity of the problem. It was desirable to maintain the CAI as a normalized index in order for it to be easily interpreted and to enable better comparability with other controllers. This is not the case with traditional measures of error such as Integral Absolute Error (IAE). For example, an IAE score for SLAC would convey very little information to someone without extensive experience and knowledge of the system. Similarly a certain IAE may be high for SLAC but low for another controller. It was therefore necessary to normalize, or at least scale the violation magnitudes on some basis that could be applied universally to other controllers. Scaling or weighting of the violation magnitudes for the individual CVs was also necessary due to the fact that some CV constraint violations are considerably more important than others.

Normalization Based on Range of Violation Magnitude

The first approach considered was similar to that taken for calculation of the MI metric which normalized each individual MVs movement at every interval based on its maximum possible movement. This approach can be expressed as

$$\delta_{i,j} = \frac{d_{i,j} - d_{\min,j}}{d_{\max,j} - d_{\min,j}} \qquad \begin{matrix} i = 1,2,\ldots,N \\ j = 1,2,\ldots,M \end{matrix} \tag{7}$$

where $d_{i,j}$ = magnitude of violation at interval i,

$d_{\max,j}$ = the maximum possible magnitude of violation for CV_j,

$d_{\min,j}$ = minimum possible magnitude of violation for CV_j, presumably zero,

$\delta_{i,j}$ = magnitude of CV_j's violation at interval i normalized between zero and one,

N = number of intervals in an assessment period and

M = number of CVs

The question is how to define $d_{\max,j}$. The theoretical maximum violation magnitude is the difference between the violated hard limit and the closest CV engineering limit, which is the absolute outer bound for that CV's region of operation. Using this value as the basis for normalization was not done for two main reasons: firstly the engineering limits are generally well outside the typical regions of operation. Using this value would scale the violation to a number so small as to be virtually meaningless, or at the least very hard to interpret. Secondly the distance of the engineering limit from the hard limit is generally unrelated to the importance of a unit violation for a given CV. Further, the SLAC engineering limits were not commonly used, accurate or available for the necessary calculations.

Defining $d_{\max,j}$ as the maximum violation incurred by CV_j for the assessment period was also considered. However the resultant metric only indicates how much time the CV spends close to its maximum violation magnitude for the period.

Scaling Based on CV Allowable Operating Range

Alternatively the violation magnitude could be scaled on the basis of the CV's allowable operating range as defined by the CV hard limits. This method is based on the assumption that if a CV has a larger allowable operating range the significance of a unit violation is less than that for one with narrower limits. The obvious drawback in this case is that not all CVs have both an upper and lower hard limit. Those that do not could be treated differently in some way but this would potentially compromise comparability between controllers as some will have more or less of these bounded CVs than others.

Scaling Based on CV Standard Deviation

The notion of scaling the constraint violations by dividing by the permissible operating range suggested a further option: that of scaling by 2 standard deviations of the CV read value. This approach assumes that greater CV variance will correspond to a wider allowable operating range and therefore less importance would be associated with a unit constraint violation.

Scaling by the standard deviation was deemed to be unacceptable for several reasons. The first being that the notion the approach is predicated upon is not correct; a wider acceptable operating range will often have no bearing on whether a violation is more or less acceptable than that for a CV with a narrower limits. This is exemplified by those CVs with no upper limits. Their allowable range may be very large resulting in a high standard deviation, but it may be considered relatively crucial that their lower constraints are not violated.

A further drawback is if the CV is exhibiting increased variance due to degradation in model quality or increased disturbances, in which case the disturbance magnitude will be scaled down as a result. In the case where the increased variance is due to model quality, the metric effectively allows greater violations for a controller which is actually performing worse, which may have been the cause of the violations.

Scaling Based on CV Average Read Value

The final method investigated for incorporating the violation magnitude into the CAI involved dividing each CV violation at each interval by a percentage of the average read value for the CV, such that

$$\delta_{i,j} = \frac{d_{i,j}}{\alpha \times N^{-1} \sum_{i=1}^{N} y_i} \qquad (8)$$

Where α = scaling percentage, initially set to 5%, and

$$y_i = CV_i \text{ measured value} \qquad (9)$$

The highest value of $\delta_{i,j}$ was capped at 1, thus a score of 1 for an interval would indicate that CV_i was violating at or greater than the maximum acceptable level. Calculation of the overall CAI for the assessment period was as per the original method: finding the average δ for each CV and then for the entire system.

The method assumes that if a CVs average value is higher, then a unit constraint violation is less important than for CVs with lower averages. So by scaling the violations by a percentage of the mean, they will be expressed as values more commensurate with their relative importance. Despite several obvious exceptions to this assumption, this method was implemented, mainly as a starting point for developing individual scaling factors for each CV.

Scaling factors were calculated with the above method then the resultant value was checked by a control engineer familiar with SLAC to ensure that the values were appropriate for both scaling and defining the maximum acceptable violation. A large number of the values did not need to be changed, but the fact that several did and that they all required verification indicated that it would have been just as, or more convenient for someone with knowledge of the process and control system to simply assign the scaling factors in the first place.

Prior to acquiring results, the procedure was updated such that the average value used was taken from all assessment periods. This was done in order to ensure that scaling was consistent for each assessment period. Also, SLAC comprises five basically identical processing units whose CVs are essentially the same, so it was desired to scale them all by the same value. Therefore, the individual average values for the corresponding CVs of different units were not applied to each respectively, but rather the median average was determined and applied to all.

2.3.2 Experimentation with Different Threshold Percentages for Calculation of OFA

The 5% range of the steady-state objective function that the current objective function must fall within for a given interval to be assigned a '1' value, was chosen arbitrarily. It was desired to apply different threshold percentages to determine whether and how significantly the selected threshold percentage affects the metric. Three thresholds were implemented, 5%, 3% and 1%.

2.3.3 EFPI Revision B: Results

Table 6 Monthly EFPI results subsequent to Revision B

Table 5 displays the monthly EFPI results for the six assessment periods having applied the revision to calculation of the CAI outlined in 2.3. Figures 3 and 4 compare the results prior and subsequent to these revisions for the CAI and overall EFPI (w/o TIN) respectively.

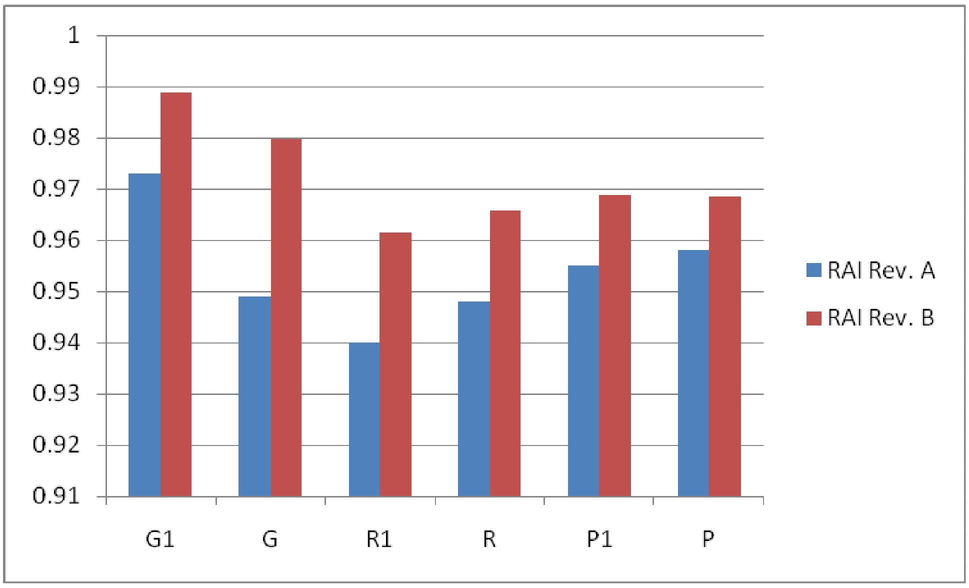


Figure 3 CAI before and after inclusion of violation magnitude

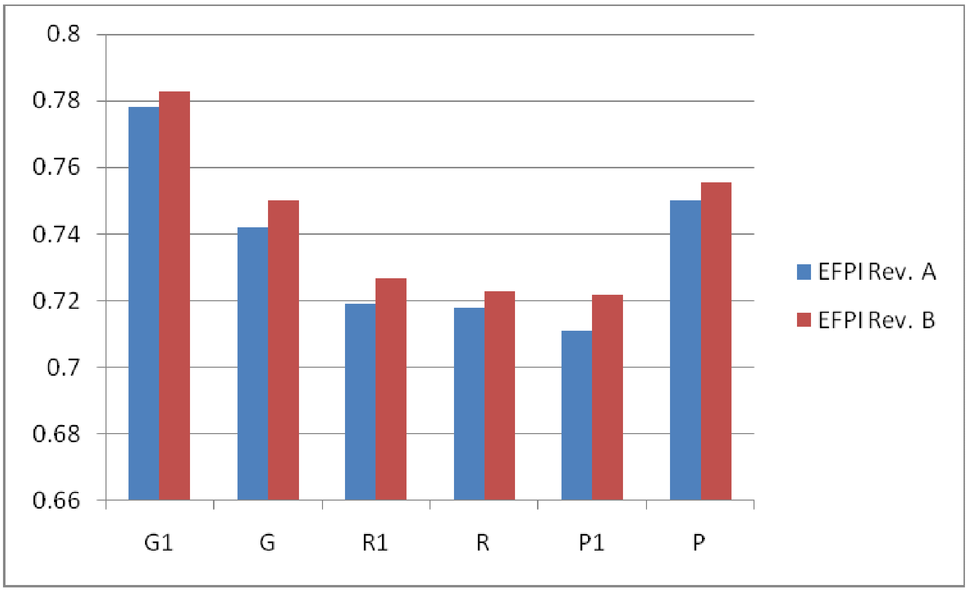


Figure 4 EFPI before and after inclusion of violation magnitude (OFA threshold 5%)

For the CAI, the overall performance trend is almost the same, however for Revision A, period P scored better than P1. After incorporating violation magnitude into the metric, however, period P1 and P CAI scores are almost identical. So while period P has a greater frequency of limit violation, the average violation magnitude during the two periods is very similar. This is potentially useful information, but would not be available without applying both versions of the metric. CAI Rev B alone will not discern whether CV limit violations are infrequent and large or frequent and small.

The CAI scores for all periods increased considerably for Revision B. This is to be expected, as when the CAI indicated the frequency of violation only, every interval where a violation occurred was assigned a ‘1’. The revised CAI assigns a ‘1’ only for those intervals where the violation is considered equal to, or greater than the maximum acceptable magnitude.

The overall RFPI has also improved for all periods, the best improvement being for P1 which now has greater parity with R and R1. For these three ‘worst performing’ periods, the effect of changing the CAI such that it effectively measures two aspects of performance, magnitude and frequency of violation, has been to reduce the discernible difference in the overall EFPI.

Changing the OFA threshold yielded the results plotted in figure 5. In addition to reducing the OFA scores for all six periods, narrowing the range generally had the effect of increasing the difference between the scores for each period, although in some cases the difference had increased. Further, the metric indicated better performance for some periods when the range was greater and the opposite when it was tightened.

While the objective function for some periods may be spending more time within a certain range of its steady state value than for other periods, tightening the range reveals that it may be spending more time at the outside limit of that allowable range. Other periods’ objective functions may spend more time outside the larger threshold, but exhibit more frequent excursions into a tighter range.

SLAC EFPI -Revision B										
Period	**CAI**	**EMI**	**CR**	**OFA (5%)**	**OFA (3%)**	**OFA (1%)**	**MI**	**TIN**	**EFPI**	**EFPI w/o TIN**
15/04/2007 - 14/05/2007 (G1)	0.989	0.829	0.173	0.964	0.945	0.765	0.953	0.960	0.746	0.782
1/06/2007 - 30/06/2007 (G)	0.980	0.739	0.130	0.984	0.937	0.747	0.908	0.981	0.725	0.748
1/08/2007 - 30/08/2007 (R1)	0.962	0.668	0.144	0.898	0.815	0.428	0.946	0.959	0.678	0.723
1/10/2007- 30/10/2008 (R)	0.966	0.665	0.145	0.891	0.754	0.296	0.935	0.997	0.691	0.720
1/04/2008 - 30/04/2008 (P1)	0.970	0.567	0.159	0.973	0.946	0.783	0.926	0.790	0.564	0.719
1/05/2008- 30/05/2008 (P)	0.970	0.664	0.180	0.997	0.983	0.763	0.948	0.957	0.717	0.752

Figure 5 OFA with thresholds of 5%, 3% and 1%

For subsequent revisions of the EFPI the OFA threshold has been selected as 3% as this appears to offer greater differentiation between good and poor performing assessment periods, without narrowing the range so much as to overly penalise even the periods where performance may be deemed good.

2.4 EFPI Revision C

2.4.1 Economic Movement Index – Revised Mapping of Objective Function to MVs

The current method of calculation for the Economic Movement Index (EMI) is unsatisfactory. This method maps MVs to the objective function via the gain array and the MV and CV linear and quadratic objective function coefficients. Sign of gain or magnitude of the coefficients is not taken into account, so although a MV that impacts the objective function may be moving, there is no way of knowing whether that movement is in an economically favourable direction. At present the index is more an indication of MV utilization.

It was desired to discern appropriate move direction based on the objective function coefficients and sub-process steady state gains and to weight each MVs individual EMI on the basis of its respective impact on the objective function.

Optimum MV Movement Direction: Linear Component of Objective Function – The linear component of the objective function was mapped to each MV in order to determine a net linear coefficient, *LC net* for each MV. The revised method used the steady-state sub-process gains and the linear coefficients for CVs and MVs to determine if the move direction of each MV at each interval was appropriate to the economic objective and to provide a weighting factor for each MV based on its respective impact on the total objective function.

For example, if MV_j has a non-null sub-process relationship with CV_i only, with a steady state gain $K_{i,j}$ and the MV and CV's linear economic coefficients are denoted a_j and b_i respectively, then the linear component of the of the objective function corresponding to MV1 can be expressed as

$$
\begin{aligned}
J_{L,MV1} &= a_j MV_j + b_i CV_i \\
J_{L,MV1} &= a_j MV_j + b_i K_{i,j} MV_j \\
J_{L,MV1} &= (a_j + b_i K_{i,j}) MV_j \qquad (10)
\end{aligned}
$$

where

$$(a_j + b_i K_{i,j}) = LCnet_1 \quad (11)$$

For an MV with non-null sub-processes with more than one CV, *LCnet* can be expressed as

$$LCnet_i = a_i + \mathbf{B} \cdot \mathbf{K} \quad (12)$$

where $\mathbf{B}$ is a column vector containing the linear coefficients of each CV and $\mathbf{K}$ is a column vector containing the steady-state gains between the MV and each CV.

The values calculated for $LCnet$ for each MV were used to determine whether they had moved in an optimal economic direction for each 35m interval for the seven assessment periods. A score was assigned for each MV at every interval based on $LCnet$ as follows

If $LCnet_i < 0$

$$\begin{aligned} I &= 1 && \text{for } \Delta MV_i > 0 \\ I &= 0.5 && \text{for } \Delta MV_i = 0 \\ I &= 0 && \text{for } \Delta MV_i < 0 \end{aligned}$$

If $LCnet_i > 0$

$$\begin{aligned} I &= 1 && \text{for } \Delta MV_i < 0 \\ I &= 0.5 && \text{for } \Delta MV_i = 0 \\ I &= 0 && \text{for } \Delta MV_i > 0 \end{aligned}$$

and if $LCnet_i = 0$, MV_i is not included in this metric. (13)

Because the controller optimizer attempts to minimize the objective function, a negative net linear coefficient for a given MV means that maximising MV is desirable. Conversely a negative coefficient implies that there are economic benefits to be gained by decreasing the MV

An average overall score of 1 for an MV indicates that the MV been moving in an economically favourable direction for the entire assessment period, while a score of zero indicates that it has been moving in a direction that minimizes economic benefits. A score of 0.5 indicates that there has been no net movement in either direction.

The assignment of a neutral score of 0.5 to an MV for an interval in which no net movement was exhibited is somewhat problematic in that zero movement can mean one of several things about the controller's operation. For example, for the case where the MV is not being used to optimize but is nevertheless not moving away from an economic optimum a neutral score is appropriate. However, if the MV has optimized to a constraint (a soft limit) and is incapable of moving further it will receive a neutral score despite the fact that the controller is performing to the best of its capabilities under the given circumstances. The latter scenario may or may not be the result of operator-set MV limits being set too tightly which is something that should be investigated in the course of diagnosis in the event of a poor criticality score.

Further to the issue of assigning a neutral score in the event of zero movement, it was questioned whether it was appropriate for the case where the MV was not being used for control. It was decided to still assign 0.5 as the economic benefits being accrued as a result were the same as if the controller was not optimizing. However, this is another factor that would need to be investigated in a diagnostic phase.

The overall Economic Movement Index (EMI) is calculated by taking the weighted average of the individual scores for each MV over the assessment period, where each weighting factor is defined as the net linear objective function coefficients for the individual CVs. The EMI is therefore more sensitive to those MVs whose values have a greater impact on the economic objective function and will not include those MVs that do not have individual economic coefficients or are not mapped to CVs that do.

The above weighting approach was considered valid with regard to the linear component of the economic objective function because at any point in an MVs operating range an incremental increase/decrease will increase or reduce the overall objective function value as if the MV had started form any other point. That is, the partial derivative of the linear component of the objective function with respect to an MV is a constant, *LCnet*. This is not true for the quadratic components of the objective function as discussed below.

Optimum MV Movement Direction: Quadratic Component of Objective Function

The EMI was extended to include mapping of MVs to the quadratic component of the objective function in order to determine whether their movement is in the most economically favourable direction. This task presented an increased level of complexity as the optimal MV movement direction as defined by the quadratic coefficients is dependent on the MVs current position.

While the linear objective function is typically used for product value optimization, the quadratic objective is used to push the process to a defined 'ideal operating point', defined for each CV and MV as desired resting values, CV_0 and MV_0.

If CV_i has a quadratic coefficient c_i, then the quadratic component of the objective function associated with this CV is

$$J_{Q,CVi} = c_i (CV_1 - CV_{0,i})^2 \tag{14}$$

If CV1 has a non-null sub-process relationship with MV1 only, with steady-state gain $K_{i,j}$, then the above term can be expressed as a function of MV1 such that

$$J_{Q,MVj} = c_i^{\,2} (K_{i,j} MV_j - CV_{0,i})^2 \tag{15}$$

Including the linear component provides the complete term for the objective function associated with $MV_{1.}$ The objective coefficients for MVs have not been included in this derivation as the SLAC MVs have none.

$$J_{MVj} = (a_j + b_i K_{i,j}) MV_j + c_i^{\,2} (kMV_j - CV_{0,i})^2 \tag{16}$$

Expanding the quadratic and combining terms yields

$$J_{MVj} = c_i^{\,2} K_{i,j}^{\,2} MV_j^{\,2} + (a_j + b_i K_{i,j} - 2c_i^{\,2} CV_{0,i}) MV_j + c_i^{\,2} CV_{0,i}^{\,2} \tag{17}$$

Taking the derivative with respect to MV_i gives

$$\frac{dJ_{MVj}}{dMV_j} = 2c_i^{\,2} K_{i,j}^{\,2} MV_j + (a_j + b_i K_{i,j} - 2c_i^{\,2} CV_{0,i}) \tag{18}$$

Setting the above term equal to zero and solving for MV_i yields the extremum for the objective function to MV_i curve, which further inspection shows is a minimum for all SLAC MVs This point is the value for MV_i that minimises the part of the objective function mapped to $MV_{i.}$

Therefore

$$MV_{i,optimum} = \frac{(a_j + b_i K_{i,j} - 2c_i^{\,2} k_{i,j} CV_{0,i})}{2c_i^{\,2} K_{i,j}^{\,2}} \tag{19}$$

Having obtained $MV_{i,optimum}$ for each MV which is mapped to a quadratic term in the objective function, scores, I for MV movement can be determined for each MV at each interval

If $MV_i < MV_{i,optimum}$

$$I_i = 1 \quad for\ \Delta MV_i > 0$$
$$I_i = 0.5 \quad for\ \Delta MV_i = 0$$
$$I_i = 0 \quad for\ \Delta MV_i < 0$$

If $MV_i >$ $MV_{i,optimum}$

$$I_i = 1 \quad for\ \Delta MV_i < 0$$
$$I_i = 0.5 \quad for\ \Delta MV_i = 0$$
$$I_i = 0 \quad for\ \Delta MV_i > 0$$

or if $MV_i =$ $MV_{i,optimum}$

$$I_i = 1 \quad for\ \Delta MV_i = 0$$
$$I_i = 0 \quad for\ \Delta MV_i \neq 0$$

(20)

The same issues associated with assigning a neutral index arise as in the purely linear case. A more difficult problem however is the question of how to weight those MVs with quadratic components when rolling them into the overall EMI.

Figure 6 shows the objective function mapped to SLAC MV7 plotted against MV7 values.

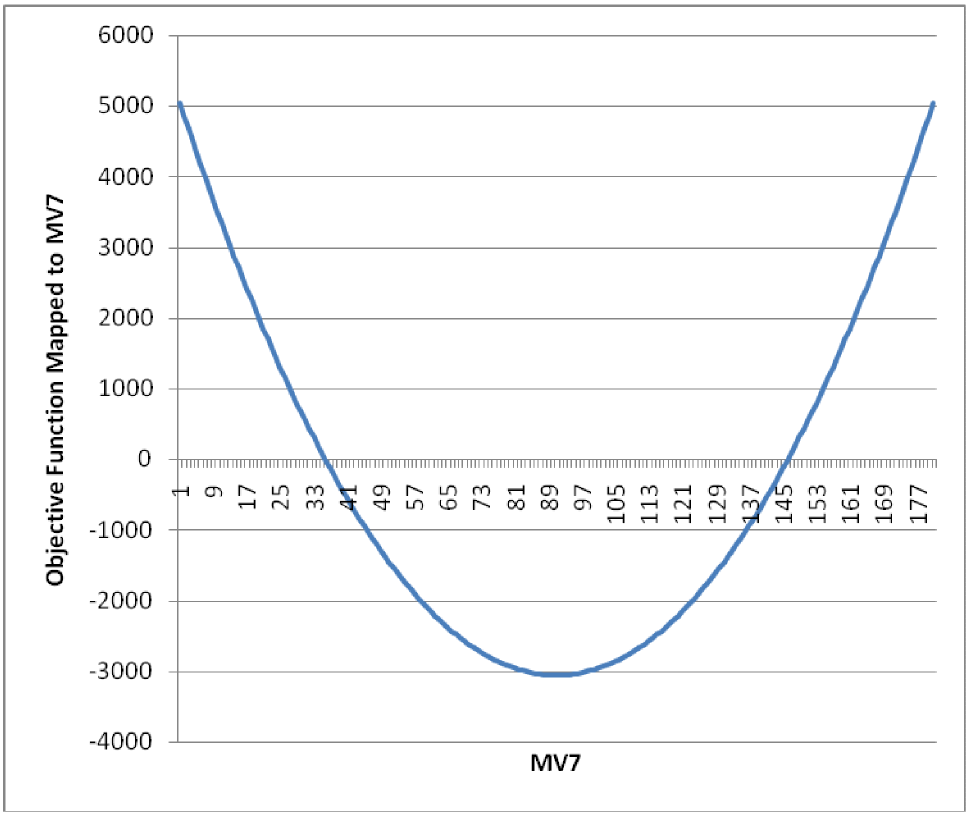


Figure 6 Objective function mapped to MV7

The impact a change the MV will have on the overall objective function value will vary depend upon the MVs current proximity to the minimum, with a change further away having considerably more bearing than one close to it. The method for weighting individual MV EMIs for an assessment period has therefore been to take the value of the partial derivative of the objective function with respect to each MV (as derived in equation 9) at each interval, and averaging them over the period.

Thus

$$Weight_i = \frac{\sum_{j=1}^{N} 2c_i{}^2 K_{i,j}{}^2 MV_j, k + (a_j + b_i K_{i,j} - 2c_i{}^2 CV_{0,i})}{N} \qquad (21)$$

where N is the number of intervals in the assessment period.

This is essentially the same as using the *LCnet* to weight the MVs with only linear objective terms; however the changing slope of the quadratic term requires averaging of the derivative values for each interval. While weighting the index assigned to each individual move on the basis of the corresponding read value would provide even more accuracy, it increases the computational burden significantly and the average has been considered adequate at this stage.

2.4.2 Additional Assessment Period

In light of the 6 month gap between periods R and P2, it was desirable to apply the EFPI metric to a 7th period between these two. It was believed that this would help validate previous results and the general performance trend for all periods.

2.4.3 EFPI Revision C: Results

Table 7 EFPI results subsequent to Revision C

SLAC EFPI -Revision C										
Period	CAI	EMI	CR	OFA (5%)	OFA (3%)	OFA (1%)	MI	TIN	EFPI	EFPI w/o TIN
15/04/2007 - 14/05/2007 (G1)	0.989	0.510	0.173	0.964	0.945	0.765	0.953	0.960	0.685	0.714
1/06/2007 - 30/06/2007 (G)	0.980	0.504	0.130	0.984	0.937	0.747	0.908	0.981	0.678	0.692
1/08/2007 - 30/08/2007 (R1)	0.961	0.520	0.144	0.898	0.815	0.428	0.946	0.959	0.649	0.677
1/10/2007-30/10/2008 (R)	0.966	0.515	0.145	0.891	0.754	0.296	0.935	0.997	0.661	0.663
1/02/2008 - 28/02/2008 (P2)	0.960	0.512	0.135	0.885	0.777	0.384	0.924	0.991	0.655	0.661
1/04/2008 - 30/04/2008 (P1)	0.969	0.509	0.159	0.973	0.946	0.783	0.926	0.790	0.554	0.702
1/05/2008-30/05/2008 (P)	0.968	0.504	0.180	0.997	0.983	0.763	0.948	0.957	0.686	0.717

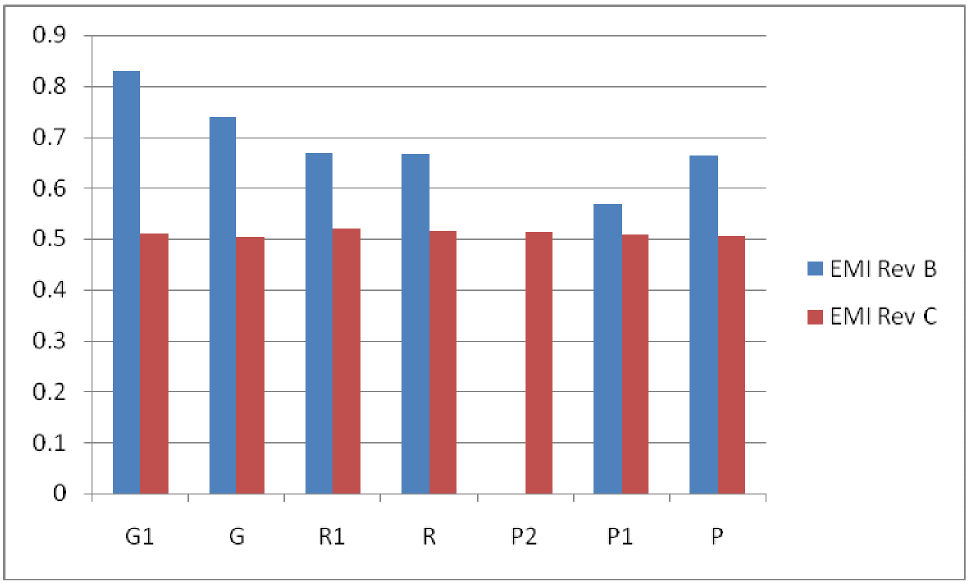


Figure 7 EMI results before and after Revision C

Figure 7 shows the economic movement index before and after inclusion of movement direction. The EMI values for all periods are all slightly larger than 0.5 indicating that overall the controller has spent more time pushing MVs in an economic direction. While the values are all very close to 0.5 and there is very little differentiation between different scores, the value of the economic objective function has changed considerably between periods, as shown in Table 6. This appears to indicate that the revised EMI does not accurately reflect how effectively the controller is using MVs to optimise the process.

The suspected reason for this is the small, equally bi-directional MV movement which occurs almost continuously. While this movement may not affect the overall optimisation as it averages to zero, it occurs so frequently that it will dominate the metric result, bringing it close to 0.5 and obscuring MV move values which more truly reflect whether the controller is using MVs to optimize.

Incorporating the magnitude of the move values into the metric was originally thought to be a possible solution to this problem. This however poses another problem: the move magnitudes associated with optimization are generally smaller than those calculated for regulatory control. Thus, if the process is experiencing higher levels of upstream disturbances and the controller is forced take regulatory action which results in MV movement away from an optimum, a metric incorporating move magnitude will score less despite the controller performing as designed and to the best of its capabilities. It results in over-penalising the controller on the basis of process performance, as opposed to control performance.

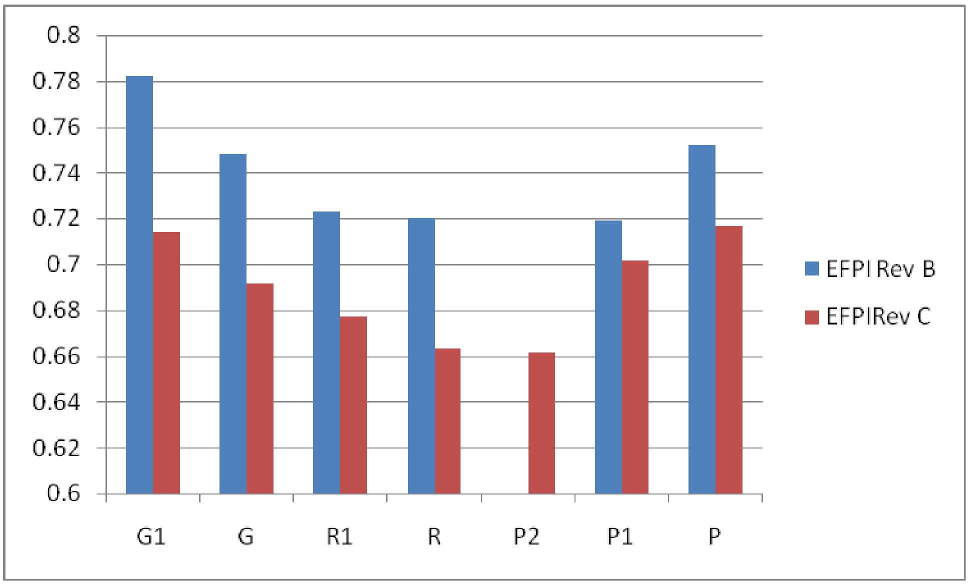


Figure 8 EFPI before and after Revision C

Figure 8 shows the overall EFPI for the seven assessment periods before and after inclusion of the revised EMI. The lower EMI scores have significantly decreased the composite metric for each period. Further, the difference between the scores for each period has been reduced, as a score of 0.5 for the EMI for each period effectively removes its effect from the overall score.

2.5 EFPI Final Results and Analysis

No further revisions were made to the EFPI subsequent to Revision C. The final overall results for each assessment period are displayed again in Table 7.

Table 8 EFPI final results

SLAC EFPI -Revision C										
Period	**CAI**	**EMI**	**CR**	**OFA (5%)**	**OFA (3%)**	**OFA (1%)**	**MI**	**TIN**	**EFPI**	**EFPI w/o TIN**
15/04/2007 - 14/05/2007 (G1)	0.989	0.510	0.173	0.964	0.945	0.765	0.953	0.960	0.685	0.714
1/06/2007 - 30/06/2007 (G)	0.980	0.504	0.130	0.984	0.937	0.747	0.908	0.981	0.678	0.692
1/08/2007 - 30/08/2007 (R1)	0.961	0.520	0.144	0.898	0.815	0.428	0.946	0.959	0.649	0.677
1/10/2007-30/10/2008 (R)	0.966	0.515	0.145	0.891	0.754	0.296	0.935	0.997	0.661	0.663
1/02/2008 - 28/02/2008 (P3)	0.960	0.512	0.135	0.885	0.777	0.384	0.924	0.991	0.655	0.661
1/04/2008 - 30/04/2008 (P1)	0.969	0.509	0.159	0.973	0.946	0.783	0.926	0.790	0.554	0.702
1/05/2008-30/05/2008 (P)	0.968	0.504	0.180	0.997	0.983	0.763	0.948	0.957	0.686	0.717

The total EFPI and its component metrics are plotted in Figure 9. It should be noted that for those months for which data was not obtained the results have been interpolated.

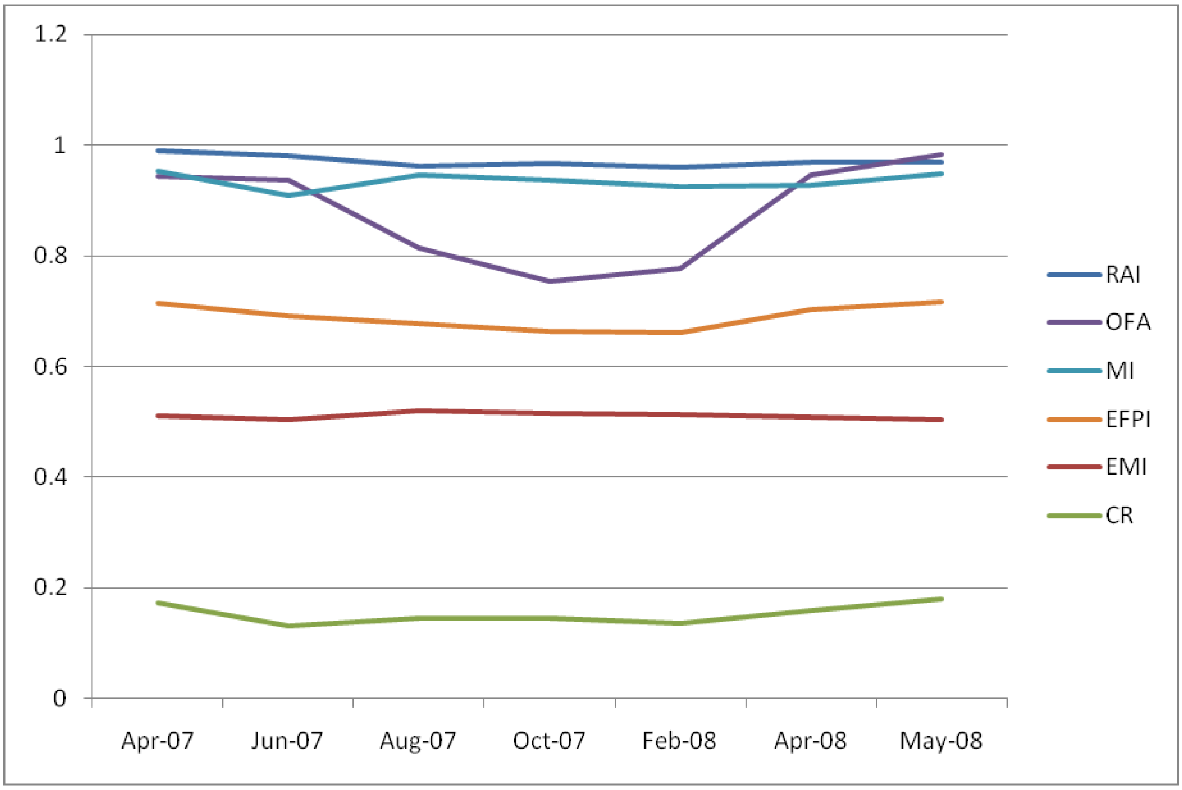


Figure 9 EFPI monthly results

The general trend for each metric except for the EMI and MI was congruent with what was expected for the assessment periods. That is, control performance for the post-commissioning period (Periods G1 and G) was expected to be very good, followed by a gradual decline in performance due to factors such as the degradation of model quality (Period R1). From October 2007 to February 2008 (Periods R and P3) there were no control engineers permanently on site and SLAC performance was expected to be at its worst due to a lack of general maintenance and attention. By April of 2008 SLAC was being maintained by a control engineer permanently on site and control was expected to have been improving. These expectations were generally reflected by the results obtained.

The exceptions to this were the EMI and MI components. The EMI actually suggests a curve that moves in the opposite direction to that expected, while the MI after a short initial decline for the second period, improves, declines marginally over the subsequent months and improves in the final period.

2. 5.1 Reporting Frequency

While the monthly figures shown in Figure 9 may reasonably reflect control improvement or degradation after the fact, monthly information on control performance is of little use in identifying problems and taking appropriate steps to deal with them before process productivity is severely impacted. An important aspect of control performance assessment is determining the optimum reporting frequency which enables diagnosis and correction of control problems within an acceptable time-frame.

Figures 10, 11, 12, 13, 14 and 15 show the daily averages for the CAI, CR, EMI, MI and EFPI respectively, for all seven assessment periods. Again, despite the fact that the data is not absolutely continuous because it was not obtained for several months, the daily scores for each assessment period are shown as a contiguous plot in order to better identify general trends.

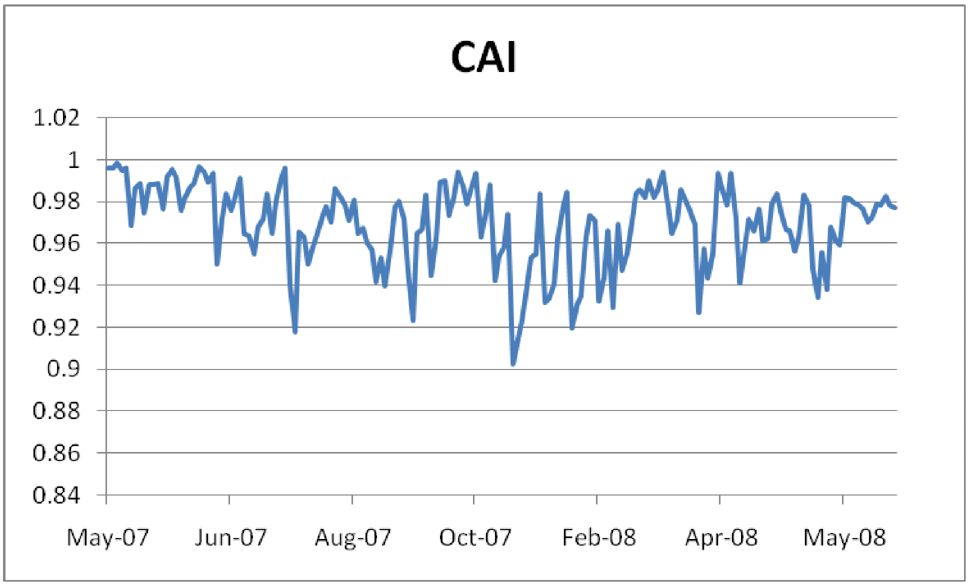


Figure 10 Daily CAI Values

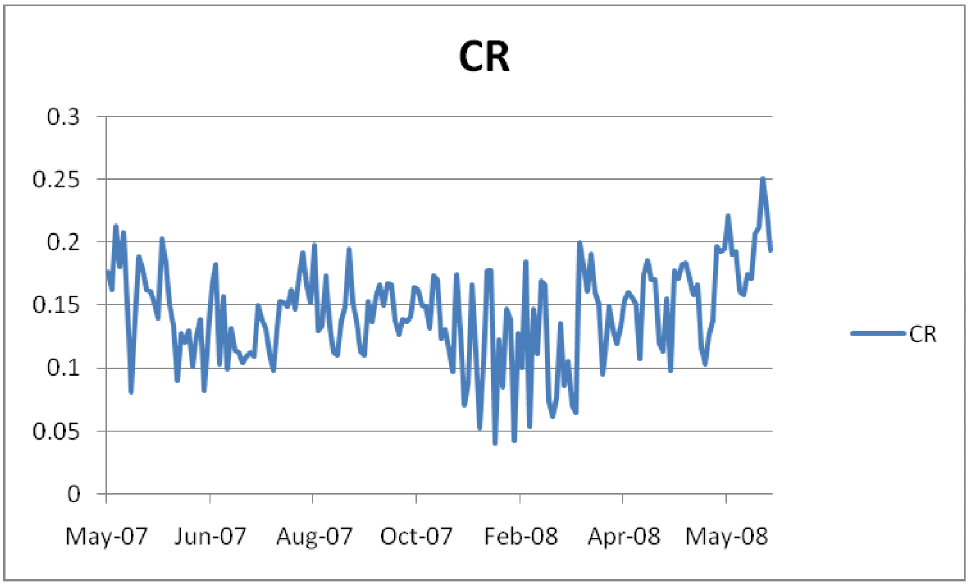


Figure 11 Daily CR values

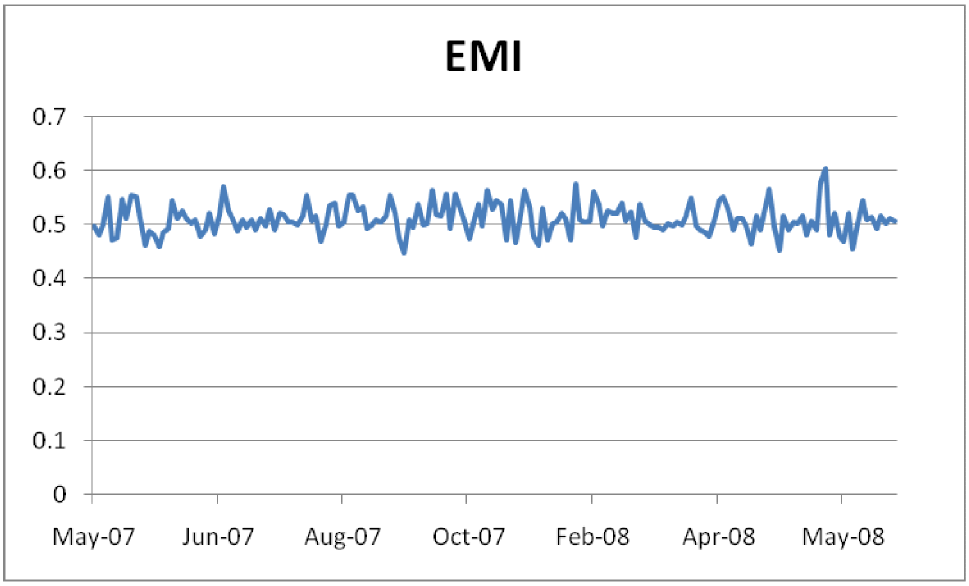


Figure 12 Daily EMI values

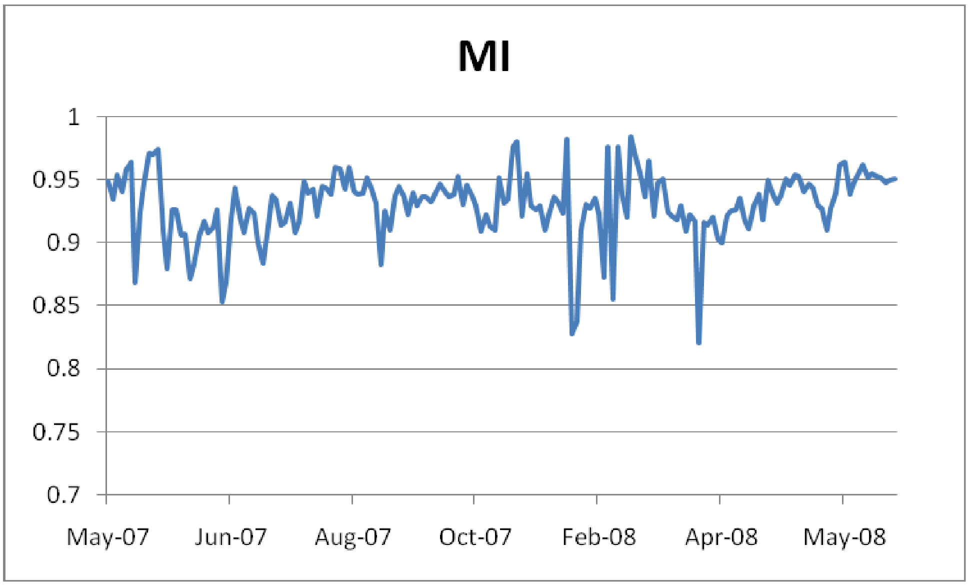


Figure 13 Daily MI values

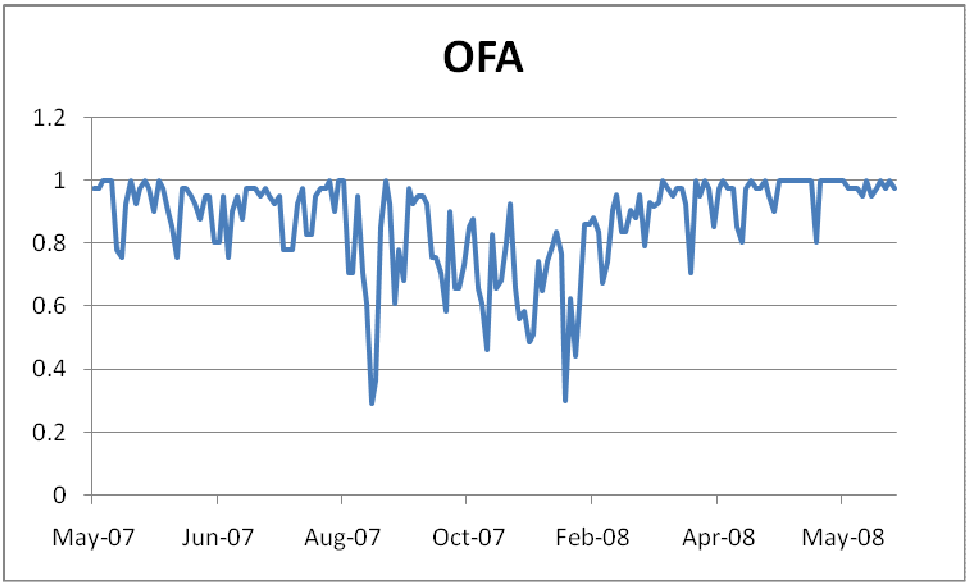


Figure 14 Daily OFA values

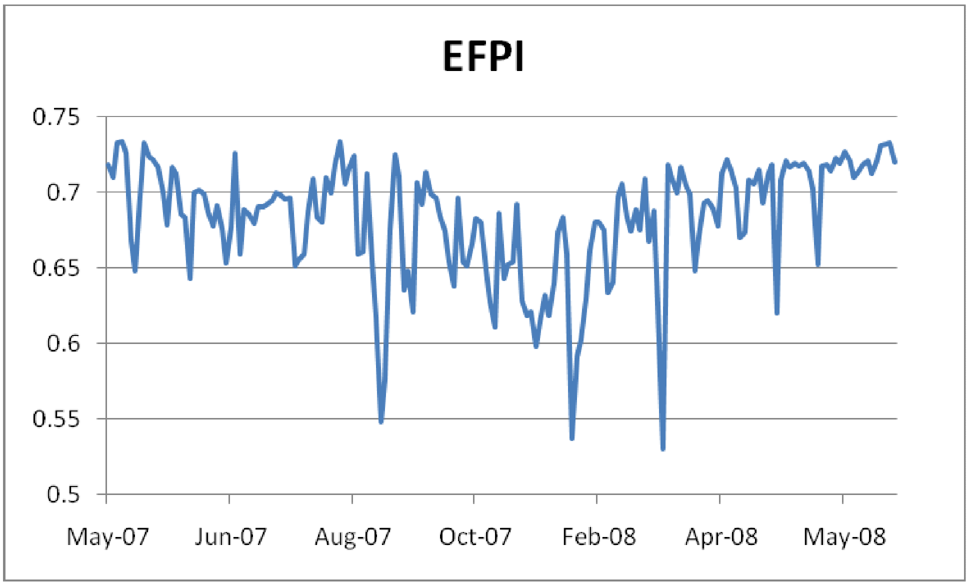


Figure 15 Daily EFPI values

While the general, overall trends for each of these components can still be discerned from these daily plots when viewed for all the assessment periods together, they offer little value with regard to identifying trends in performance improvement or degradation on a short-term, actionable level. Due to the 'noise' associated with each component, a significant decrease in an EFPI component on one day does not indicate a negative trend in controller performance.

Daily reporting and interpretation of these indices may be of value with regard to identifying short term, temporary but frequently occurring control problems, such as operators setting MV limits too tightly, resulting in downward spikes in the CAI on certain days. This could be an occurrence associated with one operator in particular and by determining the days on which this occurs, the operator may be identified and advised of the problem.

Alternatively weekly reporting may be preferable. Figures 16, 17, 18, 19 and 20 show the weekly averages for the CAI, CR, EMI, MI and EFPI respectively, for all seven assessment periods:

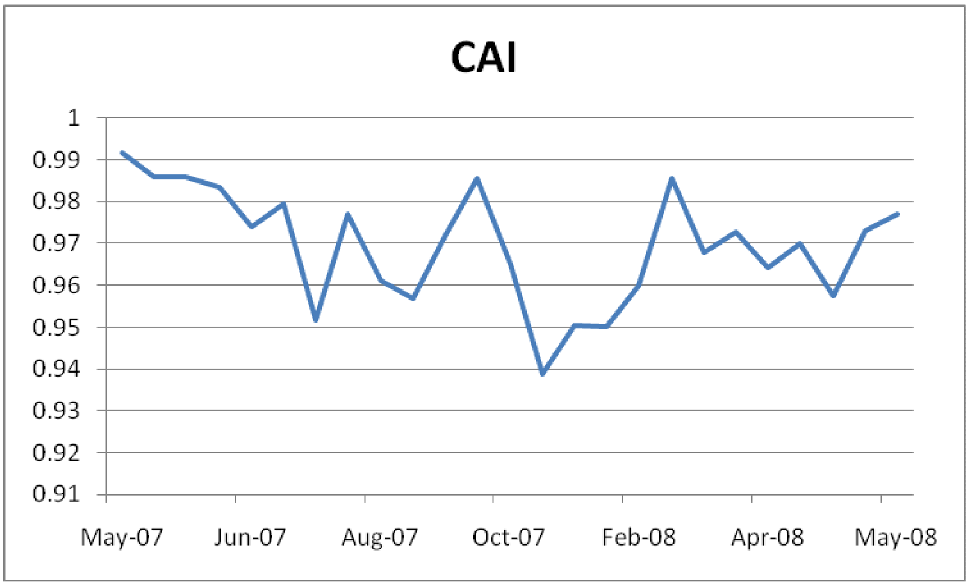


Figure 16 Weekly CAI values

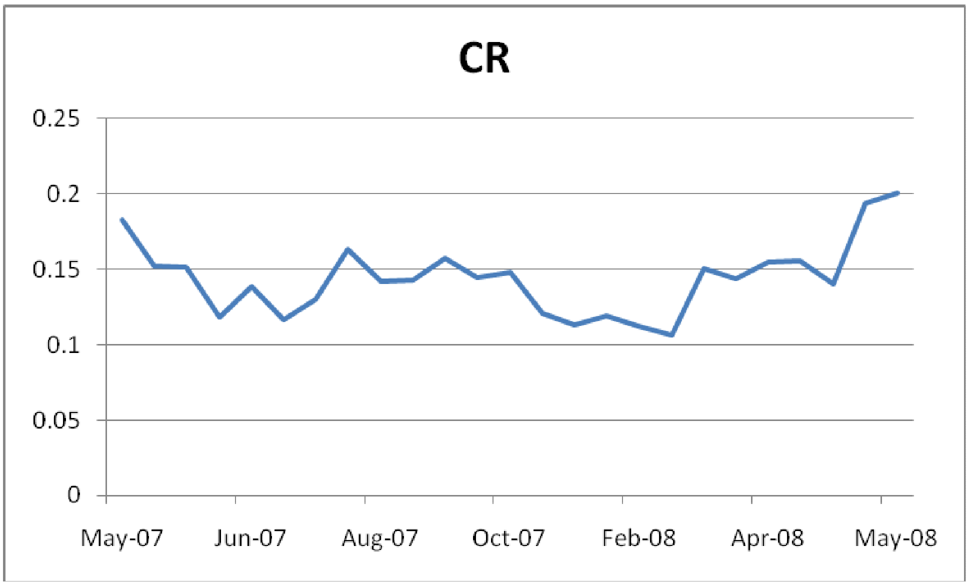


Figure 17 Weekly CR values

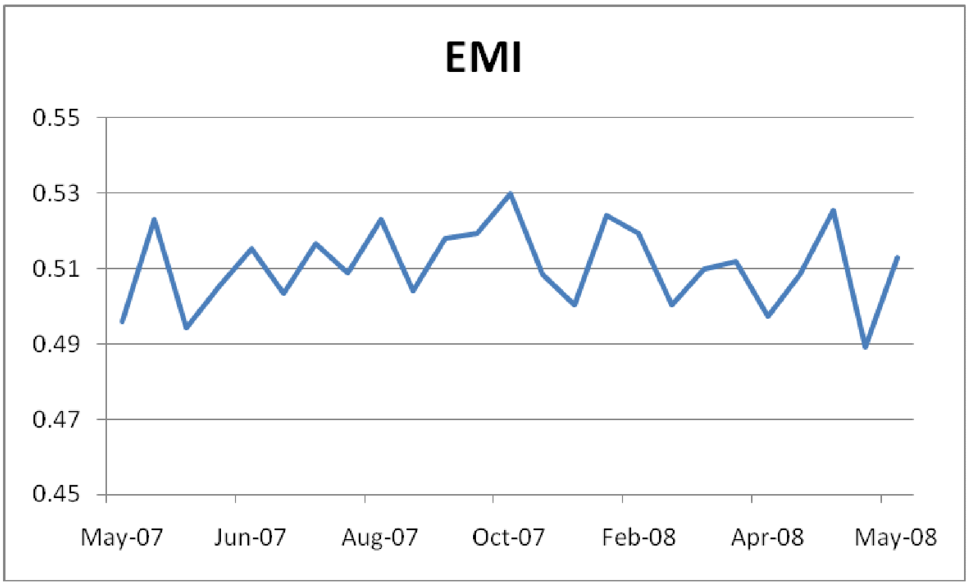


Figure 18 Weekly EMI values

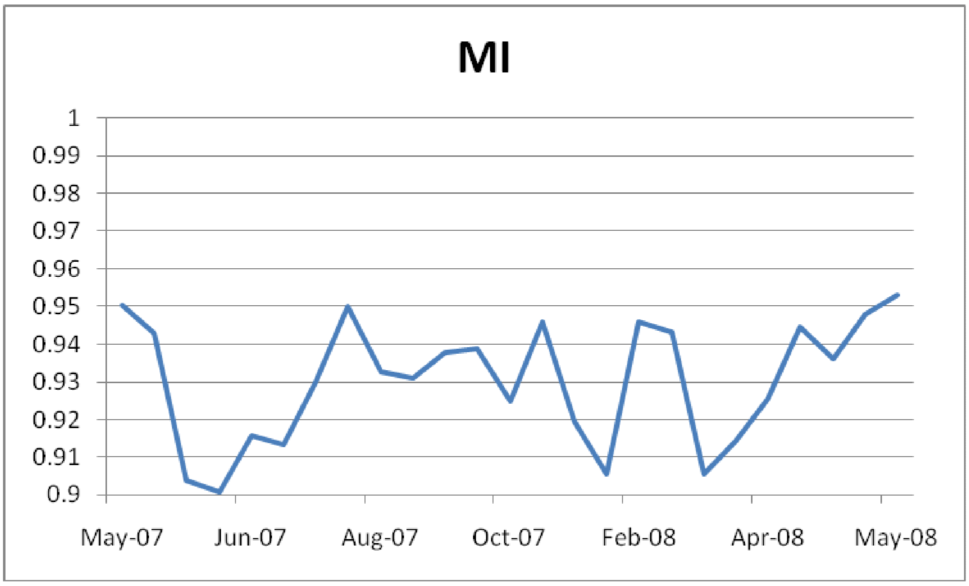


Figure 19 Weekly MI values

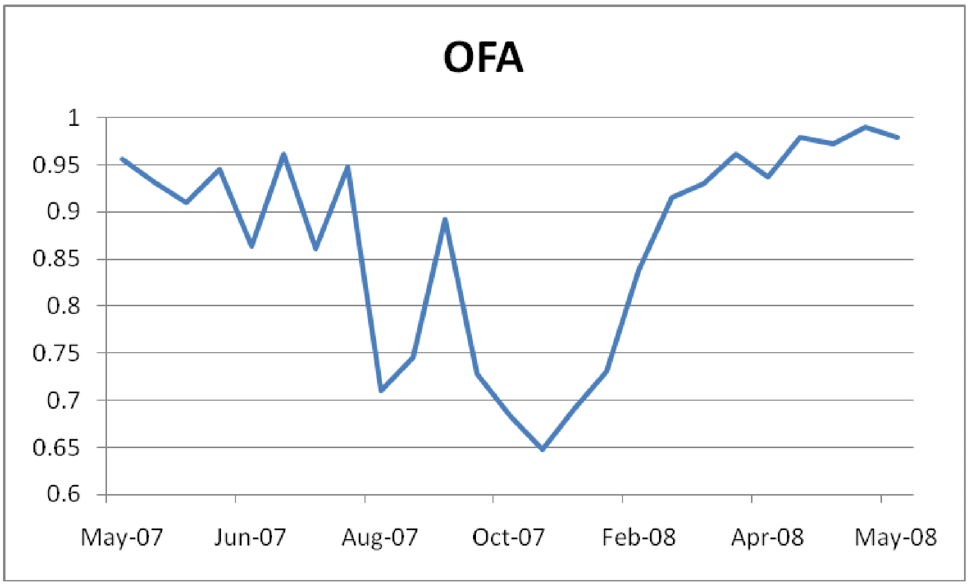


Figure 20 Weekly OFA values

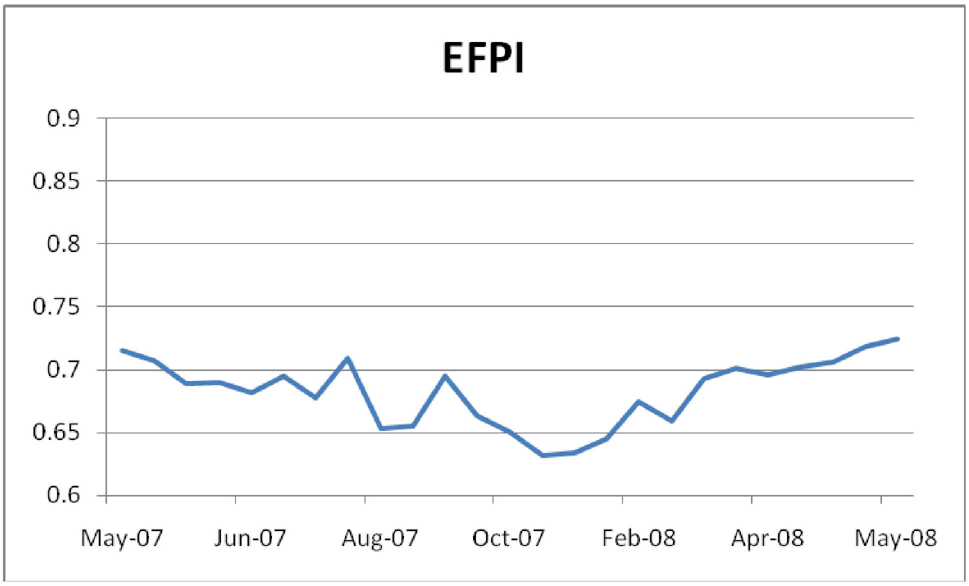


Figure 21 Weekly EFPI values

The noise associated with each metric is significantly reduced from the daily averages. The EFPI and its components each evidence spikes or troughs that do not fit the trend, but generally a few weeks worth of EFPI information could indicate the overall trajectory of each metric.

The most notable exception to this is the MI component for which it is difficult to verify any sort of trend without more than several months' data. This metric is susceptible to factors such as an increase in disturbance frequency and size and inappropriately set operator limits. This could possibly explain the high, virtually stochastic variation in the metric over the assessment periods, although the same could be said of the other components.

2.5.2 EFPI Trends and Relationships

	CAI	CR	EMI	MI	OFA	EFPI
CAI	1					
CR	0.370818	1				
EMI	-0.28147	-0.19595	1			
MI	0.062157	0.431007	0.014164	1		
OFA	0.588947	0.444703	-0.36281	0.096072	1	
EFPI	0.588976	0.696213	-0.23495	0.239923	0.923191	1

Table 9 Correlation coefficients between EFPI components

The correlation coefficients between individual EFPI components may be used, to indicate whether the trends of each metric are congruent with each other or not.

The correlation matrix for the weekly averages of the EFPI and its component metrics (Table 8) confirms the general conclusions based on a visual inspection of the results. That is, the CAI, CR, OFA and overall EFPI show the general trajectory expected for the year of assessment periods. The MI does not have a significant correlation to any of the other metric components except a relatively weak positive relationship with CR, while the EMI exhibits a trajectory which is the inverse of that expected for the overall performance during the assessment periods.

The MI's apparently positive relationship with CR is to be expected if a larger amount of MV movement, which corresponds to a lower MI, results in MVs hitting limits more frequently. If this is the case, and MVs are becoming constrained more of the time then a lower CR will result provided that this effect is not outweighed by an increase in the number of CVs at limits.

Of interest is the apparently inverse relationship between the EMI, which measures the amount of time spent pushing MVs towards an economic optimum and the OFA which measures the amount of time the objective function spends within, in this case, 3% of the steady state objective function. While at first this seems counterintuitive, it can be explained by interpreting a high OFA score as the situation where a significant number of MVs have been pushed close to their optimisation limits and therefore cannot move further in that direction which will result in a decrease in the EMI.

This suggests that it may be important to examine the results for these two metrics together. If the OFA score is low but the EMI is high, it may be that the controller is in the process of optimising and the OFA can be expected to rise. However, if both the EMI and OFA are low, the controller may not be free to optimise and other factors such as tight MV limits or process problems or disturbances need to be investigated.

The above explanation does not account for the apparently inverse relationship between the EMI and CAI. The controller will only optimise when it has non-negative degrees of freedom, defined as:

DOF = No. of MVs not at a constraint – No. CVs at a setpoint or within limits

Thus with a higher level of constraint violation, that is a lower CAI, we should expect the controller to be optimizing less of the time, resulting in a lower EMI.

However, this negative CAI-EMI relationship could possibly be interpreted as occurring as a result of degraded model quality. If the controller has DOF with which to optimize the process, but its ability to predict the resultant CV output is compromised by poor model quality, then it may push MVs to values which the model prediction indicates will not cause CV violations, but in fact will.

2.5.3 Significance and Sensitivity of EFPI and Individual Components

It is clear from plotting the EFPI components over the assessment periods and the correlation coefficients between individual components, that the component whose trend bears the strongest resemblance to the EFPI is the OFA index. This is because it exhibits the largest variations between assessment periods and therefore significantly influences the trajectory of the composite metric. This is in contrast to the EMI which exhibits the least variation and thus influences the shape of the EFPI plot very little, although it does offset it somewhat.

This evidences one of the weaknesses of the EFPI metric and also suggests problems with comparability between the individual component results and those that may be obtained for other controllers. That is, the significance of the effect of each component on the overall metric and the different levels of sensitivity each component has with regard to changes in controller performance.

Significance

The CR index, for instance, was consistently low for all seven assessment periods, ranging between around 0.05 and 0.25. This may not necessarily mean that the controller is performing consistently poorly; it may be more accurate to assume that these CR scores are a representative sample which includes the best and worst of what the controller is capable of in this regard. For example a large number of CVs means that the likelihood of a significant number of them being at a soft constraint most of the time is small. A score or 0.25 may therefore be the best that could reasonably expected for the SLAC CR index, while a score of 0.8 might be expected when applied to a different controller.

Similarly, the EMI evinces little variation around its mean as a result of the previously discussed small, bidirectional MV movement occurring frequently. This suggests that, for this controller, the most significant figure of the EMI score may be the second or third decimal place. Conversely, the OFA index (with a threshold of 3%) exhibits variations as large as 0.3.

Combining these components into a composite metric as an unweighted average can therefore result in obscuring different aspects of control performance by hiding variations in certain metrics that may be significant. Or the overall metric may appear to indicate control performance which is better or worse than it actually is due to the inclusion of a component metric that, due to the inherent characteristics of the controller and process, is consistently high or low or exhibits large or small variation.

An alternative approach to rolling each component metric into a composite index is to normalise each individual index value on the basis of its expected minimum and maximum values prior to combining them. This would scale each index on the basis of the controller's expected capabilities. Values and variations in each metric would therefore be comparable to each other and the composite measure would be more equally representative of the five aspects of control performance.

This approach however, would require *a priori* knowledge of the expected maximum and minimum values of each index which would not be available before applying the metrics to a sufficiently long period of historical data. How long exactly would need to be determined and would likely be different for separate controllers.

Sensitivity

The issue of metric variability is closely related to the sensitivity of the component metrics, that is, how well a change in the metric reflects changes in the aspect of controller performance it is designed to measure. This sensitivity varies between individual metrics and very likely between the same metrics for different controllers.

Figure 22 shows a hypothetical example where CAI indicates all CVs are initially within limits, then over time one CV progressively violates a limit for 1% of the assessment period, then 2% etc. until it is outside its constraints continuously. Similarly plotted is the EMI function with all 27 MVs initially moving economically 100% of the time, followed by one MV moving in the wrong direction 1% of the time, then 2% and so on, until it is moving in the wrong direction for an entire period while all other MVs are still moving towards their optimum

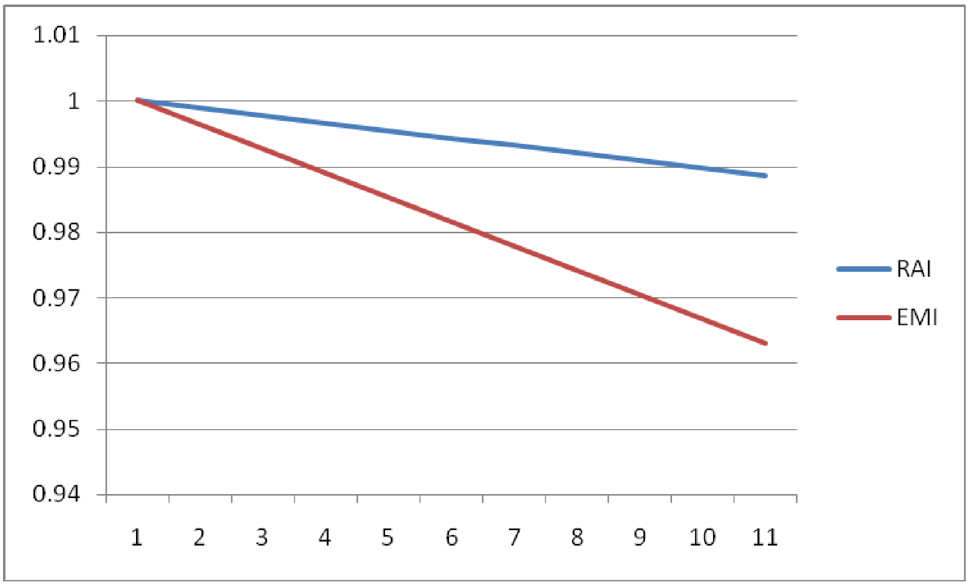


Figure 22 Sensitivity comparison between CAI and EMI

If the system under control is highly non-square in favour of the CVs, the metrics derived from CV parameters are less sensitive to changes in what they are measuring than those concerning MVs. This is fairly unavoidable without some form of scaling of the indices; however it becomes problematic when the metrics are combined, unweighted and unscaled in a composite metric.

The issue of sensitivity also raises concerns regarding the comparison of metrics between controllers. The EFPI metrics for a large controller such as SLAC will be less sensitive to changes than a smaller one. This may or may not be appropriate to the importance associated with changes in each controller's performance, but it must be taken into account when comparing their respective scores.

3 Historical Benchmarking

One of the key difficulties in developing CPA indices is ensuring that they adequately reflect a controller's true performance capabilities. For instance, an index's theoretically possible values range from zero to one, but in reality the controller's behaviour, when gauged by this metric, will only register from 0.4 to 0.6. For the EFPI indices this challenge has manifested as determining the basis on which the metrics for individual variables at each interval should be scaled.

Incorporating the magnitude of limit violation at each interval into the CAI posed difficulties because it was hard to define what the maximum 'acceptable' magnitude should be in any statistical, consistent way. The theoretical at maximum violation was often unrealistic and what may be considered a completely unacceptable violation was often well within this theoretical upper bound. Similarly, although the maximum MV move limits were used to normalize the move values each interval for the MI, these limits are often set very high and are only hit in the event of emergency.

The vast majority of the academic work in CPA has been concerned with defining benchmarks for the upper bounds of some aspect of controller performance. The current performance can then be gauged against this benchmark, indicating whether the controller is performing to the best of its capabilities.

A number of such CPA methods involve a mathematical derivation of the process output if it was under some form of 'ideal' control, thereby establishing the theoretical best control performance that could be achieved for the process. For the various reasons discussed in 1.2, these methods have mostly been deemed unsuitable for application to RMPCT applications.

Rather than establishing a theoretical upper performance bound, it is possible to benchmark some period of operation that is considered to be very good on the basis of some criteria. Subsequent results can then be compared to this benchmark period thereby gauging the control performance on the basis of the best the controller has previously been capable of.

This method is extremely attractive given its simplicity to implement and interpret. It enables much better comparability between the performances of different controllers, as they are being gauged on a scale of what they are historically capable of. It was applied to the two aspects of control performance that presented the most difficulty for the EFPI: CV limit adherence and MV move minimisation. The post commissioning period from 1/06/2007-

30/06/2007 was identified as the benchmark period for SLAC as control performance during this time was identified by engineers to be highly satisfactory.

3.1 Constraint Adherence Benchmark

Patwardhan et al. (1998) propose calculating the following quantity which is based on the least-squares control calculation employed in conventional model predictive control. This value is then compared for the benchmark and assessment periods.

$$J_{BM} = E\{(w-y)^T Q (w-y) + \Delta u^T S \Delta u\} \quad (22)$$

where $E(.)$ denotes the expectation operator, w, y and u are the measured values of the setpoints, CVs and MVs during the period of good performance. The matrices Q and S are matrices that can be used to weight the output error and MV moves respectively.

This notion of error as an offset from a setpoint does not often apply to RMPCT and certainly not to SLAC which employs range control exclusively on all variables. The analogous parameter however is the magnitude of violation of CV hard limits. Thus we can define a quantity

$$V = E(R^T Q R) \quad (23)$$

where R is the vector of CV limit violation magnitudes at a sampling instant and Q is a diagonal matrix whose non-zero entries are the engineering unit (EU) give-ups for each CV. EU give-ups are set based on the relative importance of keeping a CV within constraints. V then provides a measure of the scaled, average CV limit violation over a period.

Calculation of V for both the benchmark and assessment periods and taking the ratio of the two yields and index which indicates whether there has been any significant improvement or degradation of the controller's performance with regard to keeping CVs within constraints.

$$I_V = \frac{V_{Current}}{V_{Benchmark}} \quad (24)$$

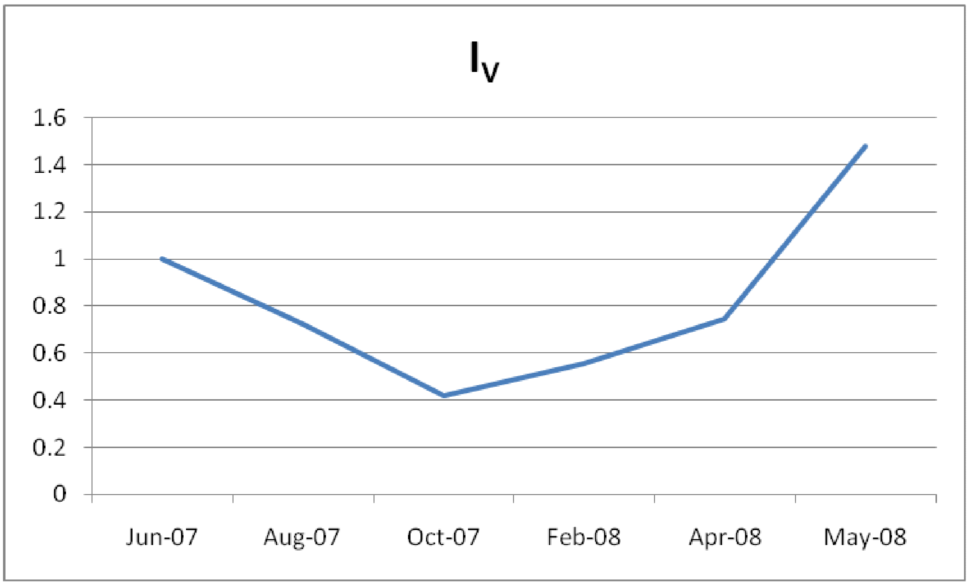
A value less than 1 indicates a higher level of constraint violation while a value greater than 1 indicates that there has in fact been an improvement in the control performance. This approach is therefore very easy to interpret provided that the benchmark period has been selected appropriately.

This index was applied to the five assessment periods following the benchmark period of June 2007. The results are shown in Table 9 and plotted in Figure 23.

Table 10 I_v values for 6 assessment periods

Period	V	I_V
Jun - 2007 (Benchmark)	136176	1.000
Aug - 2007	187755	0.725
Oct - 2007	325717	0.418
Feb - 2008	246163	0.553
Apr - 2008	183470	0.742
May - 2008	92031	1.480

Figure 23 I_v values for 6 assessment periods



The I_V trend concurs with the control performance expected for these six periods. Performance falls from the benchmark and remains low until the second last period which saw some improvement. By the final period performance has improved significantly and is in fact better than the benchmark period.

This final period may now be set as the new benchmark. By resetting the benchmark with every improved score, the metric's ability to gauge performance on the basis of what the controller is capable of is improved over time.

3.2 MV Movement Benchmark

A historical benchmarking method was also applied to SLAC's MV movement over the assessment periods from June 2007 to May 2008. The second term of equation 17 was used to calculate a quantity representing the average, scaled MV movement for a given assessment period:

$$M = E(\Delta u^T S \Delta u)$$

where Δu , as in the MPC context, is the vector of control moves for each MV at each interval. *S* is a diagonal matrix containing the MV movement weights for each MV. These weights are used to discourage the use of particular MVs in resolving CV error. There is a slight distinction between these weights and the move suppression factors employed in conventional MPC. Movement weights are only used to set priorities with regard to which MV to use when more than one can do the job. If there are redundancies in the MVs, the movement weights have no affect on movement or speed of response.

As with the MI component of EFPI, it was desired to exclude individual MVs from the metric for an interval if they were not being used to control the process. This involved removing MV move values that were not being used from vector Δu at each interval and adjusting the weighting matrix *S* accordingly.

Once again, by taking the ratio of *M* for the benchmark and subsequent assessment periods, a value is obtained which indicates whether the controller is moving the process around more or less than for the benchmark period. Results are shown in Table 10 and plotted in Figure 24.

Table 11 I_M values for 6 assessment periods

Period	M	I_M
Jun-07 (Benchmark)	228.645	1.000
Aug-07	86.669	2.638
Oct-07	148.325	1.542
Feb-08	160.199	1.427
Apr-08	66.049	3.462
May-08	74.434	3.072

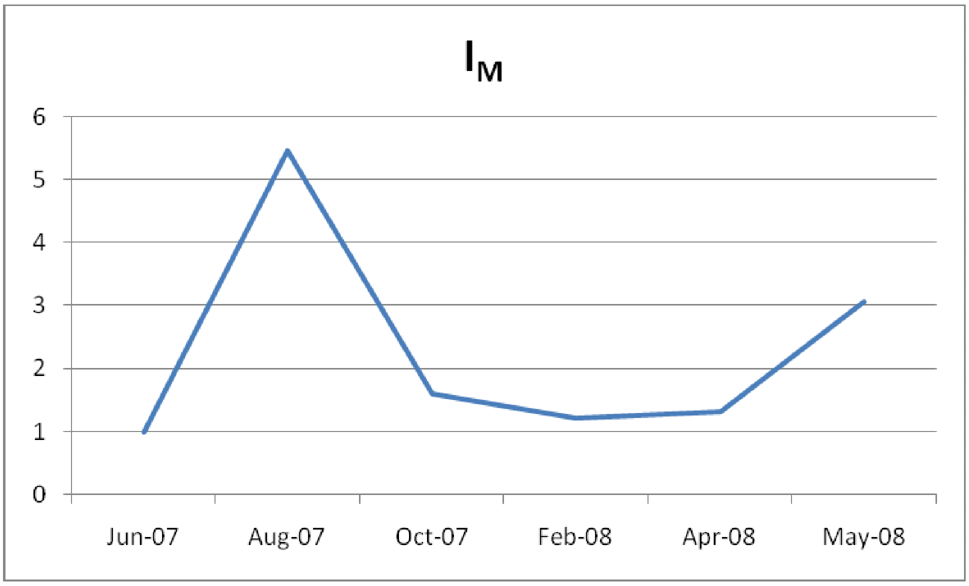


Figure 24 I_M values for 6 assessment periods

I_M does not reflect the overall control performance expected for these periods, nor do these values agree with those obtained for the MI, although the overall trend is one of improvement for the year. However, it is assumed to be a better reflection of the actual increase or decrease in MV movement over time than the MI component of EFPI.

It appears from these results that the benchmark for M should be reset to the value obtained for Aug 2007. Observing the value of I_V for this period shows that the controller is not keeping CVs within constraints well at this time. This may indicate that the controller may not be able to move this little without allowing significant limit violation, which raises the question of whether selecting different benchmark periods for different CPA components is appropriate, given that at some level they may be mutually exclusive.

3.3 Economic Objective Function Benchmark

Historical benchmarking was also applied to the objective function values obtained by the controller in order to gauge the economic benefits, as defined by minimisation of the economic objective function, the controller is generating. The quantity to be benchmarked is defined as

$$J = E(Abs(C)) \qquad (25)$$

where C is the current objective function value at each interval of the assessment or benchmark period. The controller attempts to minimise this value and for SLAC it is invariably a minimum, hence taking the absolute value of C. The economic index is then defined as

$$I_{ObjFcn} = \frac{J_{Benchmark}}{J_{Current}} \qquad (26)$$

I_{ObjFcn} for the benchmark and 5 assessment periods are shown in Table 11 and Figure 25.

Period	Abs(C)	I_{ObjFcn}
Jun-07 (Benchmark)	8060	1
Aug-07	7638	0.948
Oct-07	8081	1.003
Feb-08	6016	0.746
Apr-08	13910	1.726
May-08	7959	0.987

Table 12 I_{ObjFcn} values for 6 assessment periods

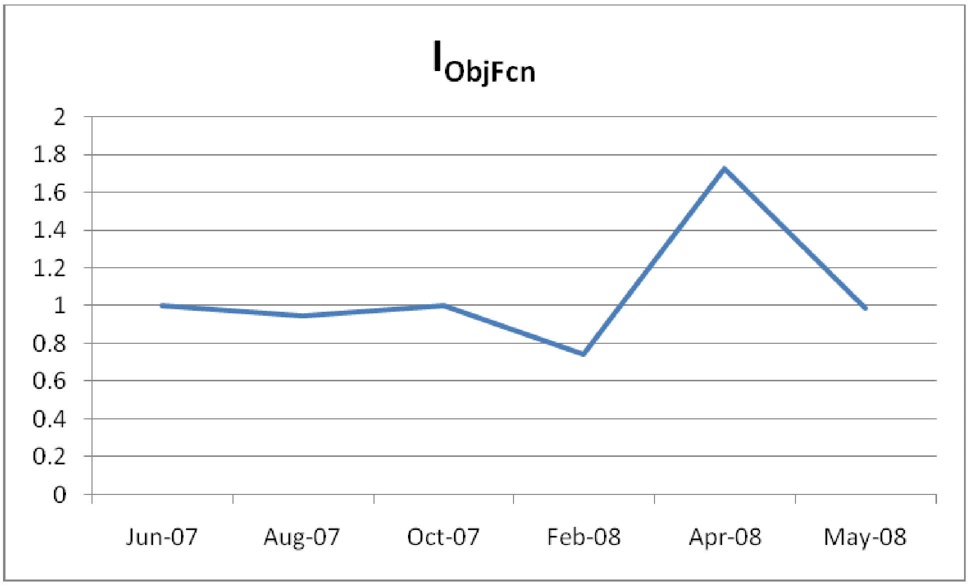


Figure 25 I_{ObjFcn} values for 6 assessment periods

As with I_M these results do not reflect what was expected for these periods. The objective function was fairly constant until early 2007 at which point a significant decrease in the controllers ability to optimise occurred, whether due to an increase in disturbances or operator set limits being placed to tightly. Between February and April 2008 the controller was able to minimise the objective function to an order of magnitude less than the other periods, before it returned to a value on par with the benchmark. Although this period was expected to be exhibiting improvement, this spike is yet to be satisfactorily explained.

3.4 Composite Metric Based on Historical Benchmarking

It is possible to combine the above three metrics into a composite measure of control performance. This presents the same fundamental problem as that identified for the EFPI rolled up index, which is that a single metric cannot adequately provide a complete picture of the various different aspects of control performance.

An additional difficulty is selecting the criteria for defining the benchmark period, given that control performance may be regarded as good in one respect and bad in another for a given period. If one benchmark period is chosen for all three aspects a component that may have been performing particularly badly for that period will exhibit inordinately large scores for periods where it was performing well thereby obscuring other aspects.

An overall metric, *I*, was defined by taking the unweighted average of all three components. The results are shown in Table 12 and Figure 26 and they show that the components have effectively balanced each other out, conveying much less information about overall performance than the three discrete metrics.

Table 13 I values for 6 assessment periods

Period	I
Jun-07	1
Aug-07	1.0246
Oct-07	1.236
Feb-08	1.085
Apr-08	1.287
May-08	1.024

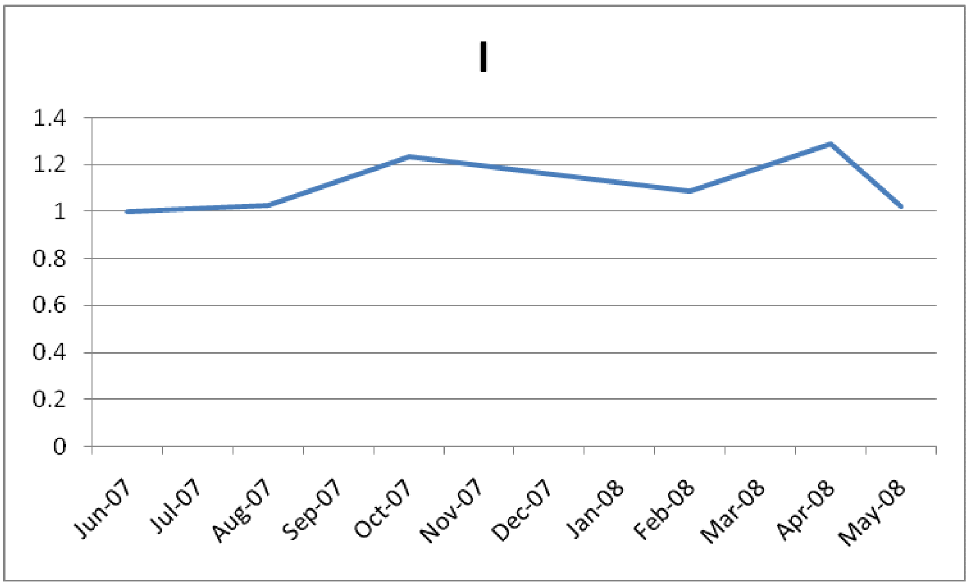


Figure 26 I values for 6 assessment periods

3.5 Prediction Error Diagnostics

While the focus of this project is on the assessment of control performance as opposed to diagnosis of performance issues, benchmarking of performance indicators suggests a possible method to diagnose degradation in control.

The improvement or degradation of an aspect of control performance can be measured, not only for the overall controller, but also for the individual variables. For instance, the difference between the average constraint violation of each CV between the benchmark and subsequent assessment periods can be measured. This was done for the historized assessment periods and the results for each period normalized, such that those CVs exhibiting the largest inflation of violation from the assessment period scored '1', while those who exhibited the least, or actually showed improvement scored a '0'. The results are shown in Table 13. The results are colour coded where red indicates the worst degradation, orange moderate and green least or improvement in staying within restraints.

It is a simple matter then to identify which of these CVs exhibits the worst degradation from the benchmark. In order to determine whether degradation of model quality has played a part in increased restraint violation, the average prediction error was calculated for badly performing CVs. This was done by taking the change in the unbiased prediction for these CVs and comparing it to the change in there measured values.

$$Average\ Prediction\ Error = E(Abs(\Delta\ Unbiased\ Prediction - \Delta CV\ Read\ Value)) \tag{27}$$

Differences in predictions are used because the unbiased prediction does not take into account unmeasured disturbances (as opposed to the biased prediction) so there is often a large offset between this and the actual value, although if model quality is good, their respective trajectories should be very similar.

The normalized increase in constraint violations from the benchmark to period P2 suggests CV77 is the worst comparative performer for the period. Comparison of the prediction errors for the CV between the benchmark and assessment period reveal a 70% increase in prediction error, suggesting a significant degradation in model quality may be a possible root cause for the increase in restraint violation. Similarly, the normalized differences suggest CV54 as the most significant worst performer for period P1, and a comparison of the average prediction errors again shows a decrease in prediction accuracy from the benchmark.

An increase in the average prediction error does not rule out other root causes of poor performance. However, a small or zero increase in prediction error can rule out model quality as a contributing factor.

Table 14 Normalized inflations/decreases in CV limit violation

	CV1	CV2	CV3	CV4	CV6	CV7	CV8	CV9	CV10	CV11	CV12	CV13	CV14
Aug-07	0.573732	0.570925	0.573798	0.574738	0.57386	0.807434	0.574033	0.618729	0.572912	0.573837	0.574033	0.574033	0.574033
Oct-07	0.262158	0.265933	0.265264	0.261709	0.263344	0.34396	0.263455	0.330794	0.262582	0.263302	0.268358	0.263455	0.263455
Feb-08	0.226295	0.228375	0.228519	0.23098	0.228025	0.218074	0.227726	0.220518	0.259369	0.227914	0.227726	0.227726	0.227726
Apr-08	0.362062	0.359967	0.363039	0.362871	0.36256	0.373566	0.362644	0.356397	0.361956	0.365373	0.368646	0.362644	0.362644
May-08	0.577391	0.579334	0.581715	0.578839	0.578559	0.567057	0.578738	0.651629	0.577615	0.578542	0.578738	0.578738	0.578738

	CV15	CV16	CV17	CV18	CV19	CV21	CV22	CV23	CV24	CV25	CV26	CV27	CV28
Aug-07	0.663056	0.574033	0.710312	0.574033	0.546472	0.573992	0.777395	0.573704	0.570935	0.574064	0.573687	0.574033	0.574033
Oct-07	0.550532	0.263455	0.551077	0.263455	0.238693	0.263419	0.382683	0.263198	1	0.264529	0.263185	0.263747	0.263455
Feb-08	0.217431	0.227726	0.234501	0.227726	0.208599	0.227693	0.224734	0.227486	0.225871	0.227726	0.227474	0.227726	0.227726
Apr-08	0.35272	0.362644	0.363246	0.362644	0.344026	0.362616	0.52252	0.362442	0.363127	0.362644	0.362431	0.364008	0.362644
May-08	0.56428	0.578738	0.574086	0.578738	0.551049	0.578693	0.575597	0.578409	0.705926	0.578738	0.578392	0.578755	0.578738

	CV29	CV30	CV31	CV32	CV33	CV34	CV36	CV37	CV38	CV39	CV40	CV41	CV42
Aug-07	0.574033	0.868201	0.574033	1	0.574033	0.544568	0.574167	0.515082	0.574033	0.473182	0.574033	0.557521	0.574033
Oct-07	0.263455	0.561546	0.263455	0.675515	0.263455	0.270522	0.263437	0.170727	0.263455	0.224574	0.263455	0.250605	0.264018
Feb-08	0.227726	0.225128	0.227726	0.235256	0.227726	0.202715	0.22779	0.136801	0.227726	0.177952	0.227726	0.215719	0.227768
Apr-08	0.362644	0.360452	0.362644	0.872146	0.362644	0.341551	0.36263	0.305039	0.362644	0.528883	0.362644	0.358466	0.366265
May-08	0.578738	0.575159	0.578738	0.577744	0.578738	0.544273	0.57874	0.477854	0.578738	0.497591	0.578738	0.574294	0.578738

	CV43	CV44	CV45	CV46	CV47	CV48	CV49	CV51	CV52	CV53	CV54	CV55	CV56
Aug-07	0.574033	0.574033	0.572173	0.574033	0.224766	0.574033	0.572772	0.574033	0.417119	0.574033	0.51935	0.570728	0.565536
Oct-07	0.263455	0.263455	0.262007	0.263455	0	0.263455	0.265694	0.263455	0.253278	0.263455	0.366457	0.260883	0.256842
Feb-08	0.227726	0.227726	0.242971	0.227726	0	0.227726	0.254656	0.227726	0.291318	0.227726	0.428819	0.274299	0.26866
Apr-08	0.362644	0.362644	0.467209	0.362644	0.240516	0.362644	0.356749	0.362784	0.264061	0.362644	1	0.360616	0.35743
May-08	0.578738	0.578738	0.877399	0.578738	0.26095	0.578738	0.570295	0.578738	0.441058	0.578738	0.553044	0.575427	0.570224

	CV57	CV58	CV59	CV60	CV61	CV62	CV63	CV64	CV66	CV67	CV68	CV69	CV70
Aug-07	0.574033	0.574033	0.574033	0.672316	0.574033	0	0.574033	0.575089	0.574231	0.668925	0.574033	0.576946	0.690974
Oct-07	0.278122	0.263455	0.263455	0.603964	0.263455	0.203621	0.263455	0.264374	0.263455	0.432069	0.263455	0.294334	0.288845
Feb-08	0.228003	0.227726	0.227726	0.784945	0.227726	0.311295	0.227726	0.227191	0.227726	0.479003	0.227726	0.226795	0.310319
Apr-08	0.36349	0.362644	0.362644	0.405763	0.362644	0	0.362644	0.363391	0.362644	0.367848	0.362644	0.377341	0.362372
May-08	0.578738	0.578738	0.578738	0.625507	0.578738	0	0.578738	0.578639	0.578738	0.667594	0.578738	0.837918	0.578294

	CV71	CV72	CV73	CV74	CV75	CV76	CV77	CV78	CV87
Aug-07	0.662987	0.576045	0.574033	0.574033	0.865045	0.574033	0.940548	0.574033	0.573589
Oct-07	0.282544	0.268995	0.263455	0.263455	0.778028	0.263455	0.871141	0.263455	0.279906
Feb-08	0.329311	0.231401	0.227726	0.227726	0.935524	0.227726	1	0.227726	0.231089
Apr-08	0.366998	0.364	0.362644	0.362644	0.366676	0.362644	0.409441	0.362644	0.403054
May-08	0.574926	0.578738	0.578738	0.578738	0.949553	0.578738	1	0.578738	0.581775

4 Conclusions and Future Needs

Several potential methods for the Control Performance Assessment of Honeywell's Profit Controller, as used in Alcoa's refinery operations, have been researched or developed and evaluated in this project.

Methods for CPA proposed in several academic studies have been researched and qualitatively evaluated for the suitability of application to RMPCT. The majority of the research in the field has focused on CPA for SISO systems or unconstrained multivariate control systems and thus the proposed benchmarking methods do not take into account the non-linearities associated with multivariable restrained systems. Further, most of these methods focus on calculating the error variance of the system under some form of ideal control which does not often apply to Profit Controller which typically uses range control as opposed to setpoints. Also, these solutions are only obtainable when the process disturbances are known and do not account for unmeasured disturbances. Model-based approaches do exist for benchmarking the performance of MPC which explicitly handle restraints. However these rely on being able to obtain the value of the control calculation objective function at every sampling interval and this was not possible with Profit Controller.

A composite CPA metric initially comprising six separate performance indicators was proposed, developed and evaluated by application to seven periods of historical data for which *a priori* knowledge of the controller's performance was available. The aspects of performance each of these were designed to measure were as follows:

- CAI – How well the controller keeps CVs within defined limits;
- MI – How much the controller moves MVs around;
- OFA – How much value the controller can generate and model quality;
- EMI – The extent to which the controller uses MVs to economically optimize the process;
- CR – How the controller adds value to the process by pushing CVs to constraints and retaining availability of MVs; and
- TIN – The amount of time the controller is on.

The last of these, TIN was later excluded from the overall metric as it was deemed more a diagnostic rather than performance assessment tool.

The major revisions made to these metrics in the course of their development are outlined. The first of these revisions highlighted that a CPA tool must be highly flexible in its configurability. The CAI for instance, could not be applied indiscriminately to every SLAC CV, as a number of them were spares, were indicative only, or deliberately violated limits. It was necessary to include this knowledge of the controller into the metric in order to obtain a result that was reflective of performance.

The majority of the EFPI component metrics were based on the frequency of certain events over the control period. This enabled the simple calculation of normalized metrics which it is fairly reasonable to assume could be applied to different controllers. The exception to this was the MI and subsequent to Revision B, the CAI. The MI normalized the magnitude of each MV's move value on the basis of its maximum allowable value. These maximum values are in practice set very large for emergency contingencies and thus do not provide a good basis for scaling the controller's MV moves. This is not a problem if the controller's performance is only being evaluated with respect to its previous performance, but it compromises the ability to compare the metric across different controllers.

The CAI presented the same problem. A satisfactory solution for incorporating the magnitude of CV limit violation into the metric and normalizing or scaling it on a statistical basis or on the basis of the controller's parameters was not found.

The frequency most appropriate for reporting of the EFPI and its components was investigated. Monthly reports can be used to identify long-term historical trends but are little use in predicting performance trends into the future. The daily averages for the metrics, because of their high variability or 'noise', offer little value with regard to identifying trends in performance improvement or degradation on a short-term, actionable level. The most appropriate reporting frequency, for the EFPI indicators is approximately weekly, as this enables trends to be identified which may be able to predict the future trajectory of control action allowing diagnosis and corrective action to be taken.

Results for the EFPI reveal the principle problem with a composite metric of this type. Combining the individual components into a single number yields an index that provides very little real information about the controller performance. Components that are consistently low or high can skew the overall metric. The low scores obtained for CR, for example, are less likely to indicate consistently poor performance but rather that these values are representative of what the controller is realistically capable of.

Alternatively the changes in the different metrics may be hidden by others. Metrics with low sensitivity, or smaller ranges of scores should be scaled such that changes in these scores are better reflected in the overall metric. For example a change of 0.01 in the EMI is likely as significant as a change of 0.1 in the OFA index. Similarly those metrics that are consistently low or high should also be scaled so as to better convey whether the controller is performing to the best of its abilities or not.

The notion of whether a metric incorporates realistic expectations of the actual capabilities of the controller led to development and implementation of a historical benchmarking approach whereby aspects of control performance were gauged relative to what the controller had previously achieved. This approach was applied to three aspects of control performance: keeping CVs within restraints, minimising MV movement, and minimising the economic objective function.

The results are very easy to interpret, as they simply indicate how well the controller has performed relative to a period of operation that was satisfactory. It is recommended that this method be used in the EFPI to replace the CAI and MI as it removes the need to find an appropriate basis on which to scale the CV limit violations and MV move magnitudes. If, however, the frequency of limit violation is also desired, the CAI in its original form should be retained.

The results obtained for the EFPI show a trajectory that generally reflects what was expected in terms of control performance for the assessment periods. However, the rolled up metric conveys very little information without also observing its component metrics.

More work is required to ensure that these individual metrics produce results that are comparable between different controllers. The measures should ideally be applied to several different controllers and scaling methods for the metrics further investigated.

The relationships between the different metrics developed need to be further studied to improve their utility as a diagnostic tool. There is not a currently a good understanding of why some scores may be low and others high in differing combinations. It is believed that a detailed study of these relationships, combined with more information about what was occurring in the process and with the controller during assessment will enable a better interpretation of the information the metrics convey. In particular, this further work should focus on investigation of the metrics in combination with process disturbances, changes in operator set limits and prediction error as an indicator of model quality.

An inordinate amount of time was spent on applying revisions in the metrics to the historized data. Microsoft Excel was used for this as it was also used to recover the historized data for assessment. If further research is to be undertaken in this area it would be advisable to write a program with Matlab or some other mathematical program which can be coded to apply changes quickly to performance metrics.

Finally, it is necessary to define further criteria for the evaluation of different metrics. At present, the main criterion is whether metric results concur with what was expected from the controller during assessment periods. It would be particularly desirable to investigate the relationship between the performance indices and other financial indicators of process performance.

Bibliography

Desborough, L. and Harris, T.J., Performance Assessment Measure for Univariate Feedforward/Feedback Control, *Can. J. Chem. Eng.* **71,** 605-616 (1993).

Harris, T.J., Assessment of Closed Loop Performance, *Can. J. Chem. Eng*. **71**, 605-616 (1989).

Harris, T.J., Bordeau,F. and MacGregor J.F., Performance Assessment of Multivariable Feedback Controllers, *Automatica* **32**, 1505-1518 (1996).

Hay, P.S.,(1994), Basic Principles of the Bayer Process: Process Flowsheet and Chemistry, *Alcoa of Australia Ltd. Pinjarra Works.*

Honeywell Industrial Automation and Control, Profit Controller (RMPCT) Concepts Reference, Phoenix AZ (2001).

Huang, B., S.L. Shah and K.Y. Kwok, Good, Bad or Optimal? Performance Assessment of Multivariable Controllers, *Automatica* **33**, 1175-1183 (1997b).

Jelali M., (2006), An overview of Control Performance Assessment Technology and Industrial Applications, *Control Engineering Practice,* **6** (2006)

Kammer, L.C., Bitmead, R.R. and Barlett, P.L., Signal-Based Testing of LQ-Optimality of Controllers, from *Proceedings of the 35th Conference on Decision and Control*, Kobe, Japan (1996)

Ko, B.S.and Edgar, T.F., Performance Assessment of Cascade Control Loops, *AlChE J*. **46**, 281-291 (2000).

Ko, B.S. and Edgar, T.F., Performance Assessment of Multivariable Feedback Control Systems, *Automatica*, **37**. 1363-1371 (2001a).

Ko, B.S.and Edgar, T.F., Performance Assessment of Constrained Model Predictive Control Systems, *AIChE Journal*, **47**, 899-905 (2001b).

Kozub, D. and Garcia, C.E., Monitoring and Diagnosis of Automated Controllers in the Chemical Process Industries, from *Proceedings of the ALChe Annual Meeting*, Chicago IL (1993)

MacArthur, J.W., “RMPCT: A New Robust Approach to Multivariable Predictive Control for the Process Industries”, Honeywell Inc Industrial Automation and Control, Phoenix, AZ

Patwhardan, R.S., Emoto, G., Fujii, H. and S.L., Shah, Performance Assessment of Model Predictive Controllers: An Industrial Case Study, *AlChE Annual Meeting*, Miami, FL (1998)

Shah, S.L., Patwardhan, R.S. and Kent, Z.Qi, Assessing the Performance of Model Predictive Controllers, *The Canadian Journal of Chemical Engineering,* **80**, 954-966 (2002).

Shah, S.L., Patwardhan, R.S. and Huang, R.S., Multivariable Controller Performance Analysis: Methods, Applications and Challenges, from *Proceedings of CPC VI*, Tucson, AZ (2001).

Tyler, M.L., and Morari, M., Performance Monitoring of Control Systems using Likelihood Methods, from *Proc. American Control Conf*., 1245-1249 (1995)

Yu, J. and Qin, S.J., Statistical MIMO controller Performance Monitoring. Part I: Data-driven Covariance Benchmark, *Journal of Process Control,* **18,** 277-295 (2008a)

Yu, J. and Qin, S.J., Statistical MIMO controller Performance Monitoring Part II: Performance Diagnosis, *Journal of Process Control,* **18**, 297-319 (2008b).