INL/CON-11-21932
PREPRINT

Concept of Operations for Data Fusion Visualization

ESREL 2011

T. R. McJunkin
R. L. Boring
M. A. McQueen
L. P. Shunn
J. L. Wright
D. I. Gertman
O. Linda
K. McCarty
M. Manic

September 2011

The INL is a
U.S. Department of Energy
National Laboratory
operated by
Battelle Energy Alliance

Concept of Operations for Data Fusion Visualization

T.R. McJunkin, R.L. Boring, M.A. McQueen, L.P. Shunn, J.L. Wright, D.I. Gertman
Idaho National Laboratory, Idaho Falls, Idaho, USA

O. Linda, K. McCarty, M. Manic
University of Idaho, Idaho Falls, Idaho, USA

ABSTRACT: Data fusion for process control involves the presentation of synthesized sensor data in a manner that highlights the most important system states to an operator. The design of a data fusion interface must strike a balance between providing a process overview to the operator while still helping the operator pinpoint anomalies as needed. With the inclusion of a predictor system in the process control interface, additional design requirements must be considered, including the need to convey uncertainty regarding the prediction and to minimize nuisance alarms. This paper reviews these issues and establishes a design process for data fusion interfaces centered on creating a concept of operations as the basis for a design style guide.

1 INTRODUCTION

An architecture for the control system of a complicated industrial system or process should foster a human-machine interface (HMI) that can be refined as new facets of the control needs are discovered. Ideally, failure points or fragilities will be identified to the extent possible during the design process. Anticipated fragilities are accounted for during the design by specifying automatic systems to identify and control the system. Diagnostics and user interfaces are prioritized within the known procedures and expectations of the plant. Resilience (Rieger et al., 2009) is built into a system by accounting for the known possibilities as thoroughly as possible but is improved by enhancing alertness of the operators to unexpected states or transients in the system. The presentation of data occurs through the HMI. The design of the HMI reflects the best known arrangement of controls and sensor displays. A perfect static display would in theory provide the operator and supervisors an ideal combination of usability and convenience of data. The critical aspects would be put front and center, while parameters for optimization would be on the periphery or omitted from the display. Such secondary information would be left for strategic queries by off-line specialists or supervisors. However, if the amount of information to be displayed requires a large surface area or frame for complete presentation, the operators will have a difficult time staying abreast of the current state at a sufficient repetition rate.

Methods of merging multiple data types into a unified data source and a design of a “dash board” have been examined (McCarty et al., 2010). The statement has been made that the human operator should not have to do mental data fusion. A dashboard concept is appropriate until the process becomes large, and the contents expand to a large screen and then to a mural size display. Then the repetition and coverage of the operator scan of the display must be considered. The possibility of coercing the operator with cues about what has not scanned recently is an option.

The system and process design must consider as many known attributes as possible. Detailed analyses can be used to prioritize and make available the information. Over time the process operation becomes mature, and risks become well qualified (if not quantified). A resilient architecture allows for refinement to take place with reasonably priced retrofits to the system, instead of living with the existing system or eventually cannibalizing it for newer system designs. An adaptive approach, as identified in this paper, is being examined as a possible step towards a resilient architecture.

2 STRUCTURE OF AN INDUSTRY SYSTEM

The conceptual HMI is designed around a plant for generating synthetic fuels. The plant is captured in a complicated model. The model framework was conceived to design and test scenarios for anomaly detection, and can handle situations ranging from malfunctions of system components to sophisticated attacks via cyber or infiltrated firmware (Chen and Abu-Nimeh, 2011). The model represents a not yet-

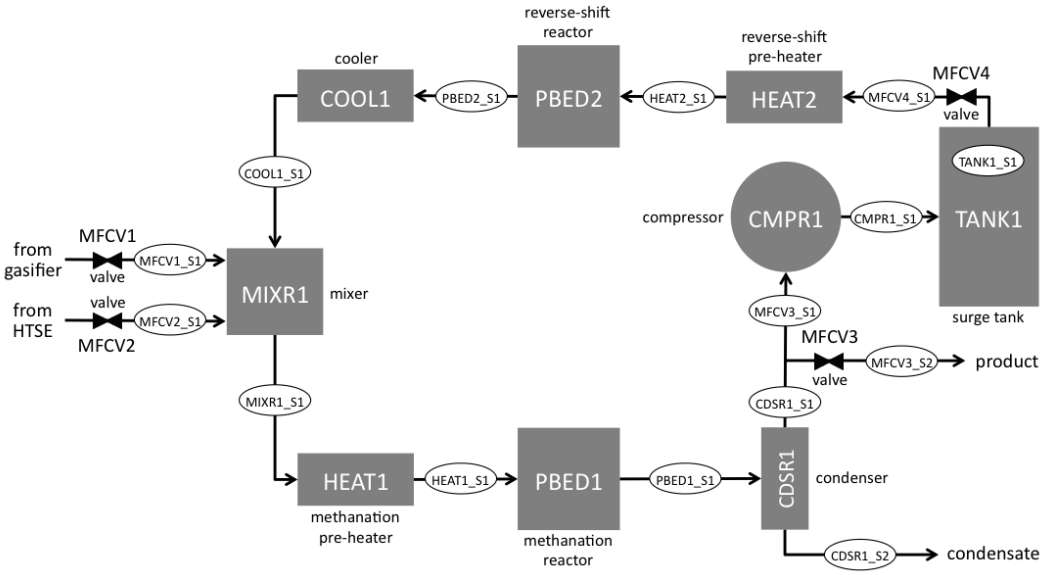


Figure 1. Process diagram of a substitute natural gas (SNG) production plant.

built part of a larger hybrid energy test facility, and thus currently has no explicit real-world plant to use for validation. The model, therefore, is the plant for the purpose of this paper, and has evolved according to the needs of the current project. Since the model has no real-world counterpart, there are also no hardened operating procedures or control and HMI designs. This presented the team with unique opportunities (good and bad) to consider a ground-up design platform, where the initial synthesis of the needed systems could be considered without an established plant interface. Mechanisms for producing scripted failures or simulating attacks were built into the model as a surrogate to real failures.

2.1 *Synthetic Fuel MatLab Model*

The target process for the current study is the synthetic fuels (synfuels) production plant shown schematically in Figure 1. In this process, syngas from gasified coal or biomass is combined with hydrogen from high-temperature steam electrolysis (HTSE) to chemically produce methane or substitute natural gas (SNG). SNG can be used as a natural gas equivalent to improve greenhouse gas emissions and improve energy security by utilizing domestic (coal) or renewable (biomass) carbon resources. SNG is a practical alternative to the "hydrogen economy," as it leverages the infrastructure, supply network, and applications that currently exist for natural gas. This and similar hybrid approaches to energy are being studied at Idaho National Laboratory's HYTEST facility (Boardman and Aumeier, 2009).

The SNG plant shown in Figure 1 involves a variety of process units and control systems. A listing of the major equipment items is provided in Table 1, with [illegible] data are collected from sensors that are distributed throughout the system. An array of measurements is taken at each sensor location, including: temperature, pressure, chemical composition, and flow rate. The sensor locations are identified in Figure 1 by oval bubbles. This system involves a large range of relevant time scales—from chemical reactions that occur on the order of milliseconds to the filling/draining of the storage tank that may take hours. This presents a variety of challenges to effectively monitor, control, and optimize the system.

Table 1. Equipment summary.

Unit ID	Unit description	Purpose
CDSR1	Condenser	Remove water formed in PBED1
CMPR1	Compressor	Maintain pressure in TANK1
COOL1	Cooler	Cool output from PBED2
HEAT1	Heater	Pre-heat inlet flow to PBED1
HEAT2	Heater	Pre-heat inlet flow to PBED2
MFCV1	Mass flow control valve	Regulate flow from gasifier system
MFCV2	Mass flow control valve	Regulate flow from electrolysis system
MFCV3	Mass flow control valve	Regulate flow of product stream
MFCV4	Mass flow control valve	Regulate flow of recycle stream
MIXR1	Stream mixer	Combine feed and recycle streams
PBED1	Catalytic packed bed reactor	Produce synthetic methane (SNG)
PBED2	Catalytic packed bed reactor	Shift carbon dioxide to carbon monoxide
TANK1	Storage tank	Provide buffer to recycle system

A computational model of the SNG plant has been developed for the MATLAB Simulink environment. This model features a fully transient description of each process unit, and simulates the proportional-integral-derivitive (PID) actions of each controller. Chemical production rates in the packed bed reactors are described by the simplified mechanism of Xu and Froment (1989). These rate expressions are valid for methanation, methane steam reforming, and water-gas shift processes over a nickel catalyst ($Ni/MgA_{l2}O_4$). Six chemical species are explicitly represented in the model: H_2, CO, CO_2, H_2O, CH_4, and N_2. The system of differential equations from the model is integrated in time using MATLAB's built-in stiff solver `ode15s` (Shampine and Reichelt, 1997).

2.2 *Mechanisms for Scripting Failures*

The model has been structured such that the inputs and outputs from the system can be affected in various ways. At each control or sensor point, the user has the flexibility to interfere with the information that is transmitted to or from the controllers in four ways:

1. Report values to the HMI that differ from actual sensor readings.
2. Relay values that are different than the intended set-points from the HMI to the controller.
3. Manipulate the sensor value that feeds back from the process to the controller.
4. Change the controller command that is sent as input to the actuator.

These mechanisms provide a myriad of possibilities for representing failures and disturbances, ranging from physical failures in sensors and controllers, to elaborate sabotage scenarios where intelligent adversaries can compromise the system to varying degrees. Simulated cyber threats can be implemented in ways to delay detection until operators are unable to intervene to prevent damage to the system.

The initial scenarios examined involve controller failures that can produce results similar to those of a mechanical failure such as a stuck valve or failed heater. More sophisticated failures that represent malicious intent are under development at the time of the writing of this paper.

2.3 *Computational Intelligence Application*

Many approaches are viable for analysis of historical data. The standardized data warehouse approach for efficient analysis of existing data (McCarty et al., 2010) continues to be applied for strategic analysis, which can be incorporated into rules to provide a tactical or control advantage. New to this work are anomaly detection algorithms, which abstract a range of transients from historical data to predict future trends. This technique uses a combination of artificial neural networks and self-organizing maps to compare predicted trends to observed data. False alarms are minimized by relaxing alarm levels when observed transients are consistent with transients in the historical data. The structure and results of this method are detailed in Linda and Manic (2011).

Any variations from the predicted values are used to highlight the most interesting information in the sea of data. When a module behaves differently than the predictions from the computational intelligence engine, one of two possibilities exist. First, an anomaly may be occurring. Or secondly, the transient may simply not be contained in the training history. Either way, the operator should be alerted in a manner to increase his or her sense of urgency regarding the situation. In the case of a normal transient that has merely not been "seen" in the training data, the system should heighten the attention of the operator until a safe steady state has been achieved. If similar transients are frequently encountered and determined safe and appropriate for operations, then these data should be included in a re-training of the computational intelligence system. The outputs with the largest errors should be highlighted in the the HMI.

3 THE DATA FUSION INTERFACE

According to El Faouzi et al. (2011), "Data fusion is a collection of techniques by which information from multiple sources is combined in order to reach a better inference." In considering the design of the HMI, the presentation of the fused data is optimized for end use. Such a design process ideally makes use of first principles and practical experience from human-computer interaction and user-centered design. Yet, extensive insights and experience with such systems remains elusive, and there is currently no specific guidance to help the designer of a data fusion system to present information in an optimized or usable manner. The remainder of this paper will outline current efforts to create a style guide of design principles for the presentation of data fusion information specifically for the HYTEST system and generally for a process control context.

Process control involves an operator interacting with a control system to ensure the effective and safe startup, operation, and shutdown of a production process. Process control can take the form of manufacturing and fabrication—including especially chemical processing—to energy production and distribution. The degree of operator interaction with the control room interface varies considerably. A modern, highly automated petrochemical production system may feature an operator in a primarily monitoring role. In contrast, an all-analog power plant

control room may feature multiple operators to monitor and actively control energy production.

Current process control interfaces provide key indications on process and plant states such as temperature, flow, pressure, etc. as discussed in Section 2.1. These indications are typically provided for every available component sensor in the system. In analog process control interfaces, these sensor indicators comprise multiple panels across a control room, resulting in hundreds and sometimes thousands of indicators for the operator(s) to monitor. Digital control rooms typically employ the advantages of software windowing technology, allowing sensor readings to be displayed for only the system or components that are of interest, often coupled with an overview piping and instrumentation diagram (P&ID).

With the complexity of multiple indications for the operator to monitor, it is crucial that the operator is aware of malfunctions or undesirable transients in the system. As such, alarms are provided to alert the operators of anomalous conditions to direct the operator to resolve the problem in a timely manner. In many cases, however, the alarms feature rigid setpoints, resulting in a variety of false or nuisance alarms. Moreover, alarms are often coupled to individual sensors. Any disruption in the system process can result in a cascade of alarms as multiple component sensors report the disturbance. Such an alarm flood can actually undermine the purpose of alarms, namely to focus the attention of the operator on a problem. Each concurrent alarm dilutes the ability of the operator to focus his or her response on the problem at hand.

Data fusion may encompass both sensor input and alarms. In terms of sensor input in a process control interface, data fusion represents the attempt to group multiple component sensor readings into a high-level system indication. For example, separate indications for temperature, flow, and pressure might be merged into a single gauge. Alternately, readings from multiple temperature sensors might be aggregated into a single indicator of overall temperature range. Should individual sensor readings fall out of the desired range, it is important that the operator be given enough information to respond. If, for example, a single gauge enveloping temperature, flow, and pressure indicates an anomaly, the information provided must guide the operator to readily diagnose the source of the anomaly. Likewise, if an aggregate temperature reading shows a higher than expected level, it is important that the operator be able to determine which specific temperature sensor is reporting high so that appropriate response can be taken.

For alarms, data fusion takes the form of aggregating multiple alarms into a single alarm. Two current approaches accomplish such aggregation: alarm filtering and root cause alarms. Alarm filtering is typically used as a way to eliminate nuisance alarms. Fixed alarm setpoints are modified according to the context of the process, thereby eliminating many irrelevant alarms. For example, during process startup, many standard indications will trigger a low-low alarm, because they have not yet reached their normal operating range. Low-low alarms designed for normal (energized) operations are meaningless during startup mode. More importantly, alarms that are actually important in detecting anomalies during startup may be masked by nuisance alarms. Alarm filtering may also be applied for the purposes of data fusion. Rather than annunciating alarms across a range of impacted sensor readings, alarm filtering may simply provide a group alarm for a series of related sensors. The same limitation applies as with the fusion of component sensor data—by eliminating some information from the indication or alarm, it may hinder the ability of the operator to pinpoint the specific component causing the alarm. For this reason, root cause alarm systems have been developed. These alarms require a network of interdependencies in order to determine which alarms are interconnected. When a pattern network is detected, the alarm system traces through the recent alarm history to determine which alarm was the initiator for the series of related anomalies. Instead of providing a cascade of alarms, the root cause alarm highlights the alarm that triggered the malfunction. In most cases, by addressing the root problem that first initiated a series of subsequent malfunctions, it is possible to solve the problem and eliminate subsequent downstream malfunctions.

4 PREDICTOR SYSTEM AS DATA FUSION

Our discussion on data fusion has thus far focused on HMIs for existing sensor data found in process control systems. Predictor systems have recently been introduced into process control systems (Baigorria et al., 2003) and other systems (Sproles and Bavuso, 2003) as an additional source of data available to assist the operator in controlling the system. Uniquely, predictor systems anticipate future states of the system based on currently available information. This look-ahead is based on anomaly detection algorithms using a variety of neural-fuzzy architectures as discussed in Section 2.3.

A predictor system includes the challenges of data fusion interfaces for existing sensor indicators—the tradeoff between displaying parsimonious indications and providing precise diagnostic information to the operator, and the challenge of down-selecting the most appropriate or relevant alarms. In addition, a predictor system presents new interface issues for data fusion. Most noteworthy of these issues is the fact that a predictor system is an uncertain indication. While the operator may assume a high degree of system integrity and sensor reliability with con-

ventional data fusion, the operator is confronted with the new challenge that the predictor system provides a probabilistic, extrapolated outcome for the process control, but there is no guarantee that such an outcome will actually occur. Essentially, the predictor system must win and maintain operator trust (Li et al., 2004).

For the HMI designer, the challenge on the one hand is to convey the uncertainty associated with a prediction and on the other hand to prevent the introduction of yet more nuisance alarms to the operator. The challenge becomes designing a predictor system that is useful to the operator, that provides reliable indications of future plant states, that alarms only when there is a high degree of certainty regarding a pending system anomaly, and that synthesizes disparate information sources. A summary of interface design issues for data fusion from existing sensor data and predictor systems is found in Table 2.

Table 2. Data fusion interface issues.

Design Issue	Domain
Selection of appropriate sensor data to provide operator with good system monitoring overview (which sensor data to fuse)	Sensor data
Anomaly presentation using fused data displays that allow operator to pinpoint problem areas	Sensor data
Alarm groupings that allow operator to trace root cause of alarm	Alarm filtering
Presentation of uncertainty information on predictive states to allow operator to judge the potential for anomaly	Predictor system
Alarm presentation for highly likely predicted anomalies	Predictor system

5 CONCEPT OF OPERATIONS FOR DATA FUSION INTERFACES

5.1 *Definition*

The Institute of Electrical and Electronics Engineers (IEEE; 2007) defines a concept of operations (ConOps) as a document that:

> … describes system characteristics of the to-be-delivered system from the user's viewpoint. The ConOps document is used to communicate overall quantitative and qualitative system characteristics to the user, buyer, developer, and other organizational elements (e.g., training, facilities, staffing, and maintenance). It describes the user organization(s), mission(s), and organizational objectives from an integrated systems point of view.

The ConOps provides a working document that defines the scope of the design of a system, fine-tunes that design into a specification, and refines the design implementation through verification and validation, as depicted in Figure 2.

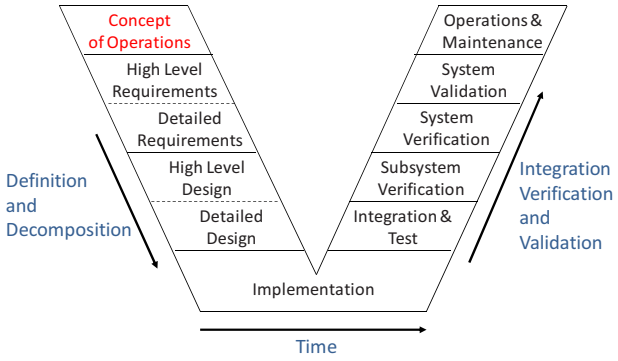


Figure 2. ConOps phases (from US Department of Transportation, 2004)

5.2 *Specific Concept of Operations*

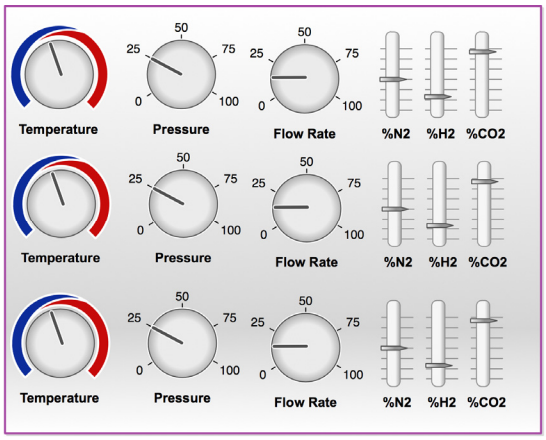
The HYTEST system includes different types of potential process control interfaces, ranging from traditional sensor-based analog-style gauges (see Figure 3a), to digital text-only indications (see Figure 3b), to data visualizations incorporating a predictor system that provides status look-ahead information (see Figure 3c).

The ConOps outlined in Table 3 has been developed to serve as guidance for design of the HMI. While the specific implementation of the HYTEST system serves the purposes of defining and refining the HMI, the purpose of the ConOps is ultimately to develop generalized principles that will serve as a formal style guide for future data fusion interface development, especially where data fusion includes predictor systems. Note that at the time of this writing, the ConOps only encompasses high-level requirements translated into design specifications. As the project continues, the high-level requirements will be refined into detailed requirements and design specifications in the next phase of ConOps development.

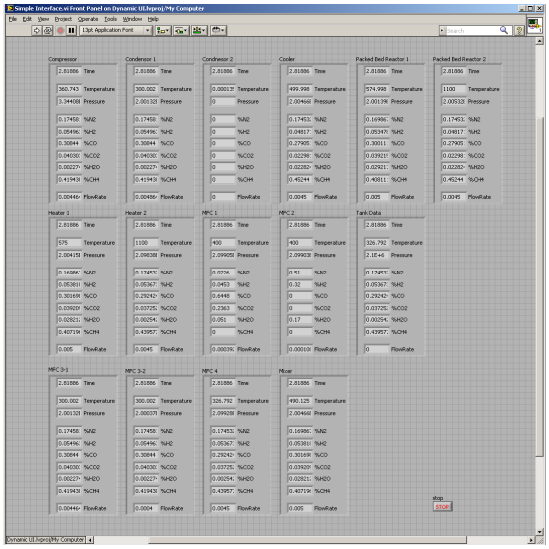
5.3 *ConOps Verification and Validations*

The initial high-level ConOps provides a suitable starting point to begin drafting a process control interface. However, the high-level description leaves many unanswered questions about the optimal manner of presenting information in the interface. Before a detailed design specification can be developed—and before the detailed design specification can be generalized into a style guide—it is important to test design assumptions. We have adopted an iterative design and test strategy to validate and refine the high-level design specification into the detailed design specification.

(a)



(b)



(c)

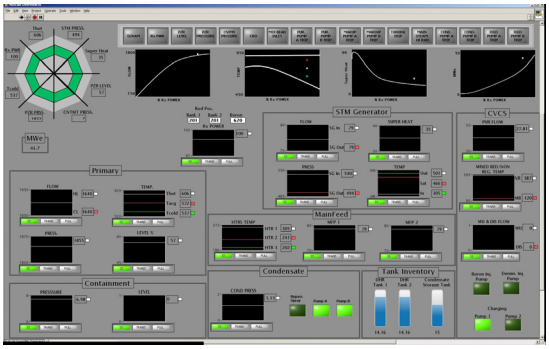


Figure 3. Process control mockup displays for traditional Analog gauges (a), digital text-only indicators (b), and a data fusion display with predictor system (c).

An operator-in-the-loop study using process experts will be performed using variations of the process control interfaces depicted in Figure 3. Sample data streams from the HYTEST system will be played back on each of the three interfaces. The experimental manipulations will include:

- *Interface Type:* The basic component-level sensor data using analog-type gauges and digital text readouts will be contrasted with a data fusion interface that aggregates component sensor data at the system level.

Table 3. High-level concept of operations for data fusion.

Number	High Level Requirements	High Level Design Specification
1	Display basic component sensor data for monitoring	Identify crucial components; Display sensor data as gauge, text, or chart
2	Integrate component data to provide system-wide overview	Identify candidate system data points that combine multiple component sensor data; Display fused system sensor data as gauge, text, or chart
3	Provide anomaly pinpointing	Provide clear indication of anomaly range on component sensor indicators; Provide component-level anomaly data in fused display
4	Provide alarms for critical anomalies	Identify alarm setpoints and operating mode dependencies; Prioritize alarms; Display alarms; Filter low priority and nuisance alarms
5	Integrate predictor system anomaly detection	Provide display that links current and predicted state of sensor data; Display anomaly range; Display uncertainty information to help operator determine if prediction is actionable; Allow operator to judge usefulness of predicted information or update predictor algorithm based on actual outcome

- *Predictor System:* Each of the interface types will be equipped as real-time readouts only or as real-time readouts coupled with a data from a predictor system.
- *Alarm Type:* The interface types with and without predictor systems will be equipped with either a simulated lightbox-style annunciator or a digital alarm list.
- *Scenario:* Several types of anomaly scenarios will be presented across each of the conditions.

The study will employ a within-subject design, during which the following performance data will be collected:

- *Response:* The interface will feature a fixed-response selection of possible operator actions to measure operator accuracy and response time.

- *Visual Gaze:* Eyetracking data will be collected to measure the areas of visual fixation and overall scanning pattern, which provide information about which visual items are particularly salient to the operator as well as the paths most employed in visual search, respectively.
- *Subjective User Impression:* While it is important to distinguish between what the user desires vs. what the user needs (Yung and Anttila, 2007), data related to user satisfaction with a particular interface and overall user preference can be helpful in shaping future design decisions.
- *Debrief:* The process experts will be interviewed subsequent to reviewing all experimental conditions to determine any additional insights that should figure into the detailed design.

Because a limited number of process experts for the HYTEST system exists, each participant in the study will view all experimental manipulations. A final validation will attempt a greater number of participants. The elements of the experiment will be randomized to prevent response bias or carryover effects. Given the low power inherent to the study design, no attempt will be made to apply inferential statistics to the data. It is understood that a certain degree of study artificiality exists due to the presentation of process control playback outside the context of controlling the process by the operator. Such experimental artifacts are understood as a necessary tradeoff between external validity and the need to collect basic performance data to improve the interface design.

As previously noted, the findings from the operator-in-the-loop study will be used to refine the HYTEST HMI. The goal is to ensure that data fusion interfaces result in at least as good of performance as basic component displays and that the predictor system enhances anomaly detection by operators. Should these findings not be borne out in the results of the verification and validation, additional refinements will be made to the design of the interface prior to determining the detailed design specification.

6 CONCLUSIONS

This paper has presented a process that is being used to arrive at a style guide for data fusion interfaces in process control as well as for the inclusion of predictor system data in data fusion interfaces. Currently, no clear guidance exists to determine the optimized presentation of fused sensor data in process control. By employing a concept of operations approach to data fusion interface design, initial design guidance has been crafted. This guidance will be validated to arrive at the detailed guidance that will serve as the basis of the design style guide.

7 ACKNOWLEDGEMENT

This work is supported by the U.S. Department of Energy under DOE Idaho Operations Office Contract DE-AC07-05ID14517, performed as part of the Instrumentation, Control, and Intelligent Systems Distinctive Signature (ICIS) of Idaho National Laboratory.

8 REFERENCES

Baigorria, L.A., Postigo, J.F., Mut, V.A., & Carelli, R.O. 2003. Telecontrol system based on the Smith predictor using the TCP/IP protocol. Robotica, 21(3): 303-312.

Boardman, R.D., & Aumeier, S.E. 2009. INL Hybrid Energy Systems Development and Testing Facilities Plan, INL/INT-09-16960. Idaho Falls: Idaho National Laboratory.

Chen, T.M., & Abu-Nimeh, S. 2011. Lessons from Stuxnet. Computer, 44(4): 91-93.

El Faouzi, N.E., Leung, H., & Kurian, A. 2011. Data fusion in intelligent transportation systems: Progress and challenges—A Survey. Information Fusion, 12(1): 4-10.

IEEE. 2007. IEEE Guide for Information Technology – System Definition – Concept of Operations (ConOps) Document, IEEE Std 1362-1998, R2007. Piscataway, NJ: IEEE.

Jung, Y., & Anttila, A. 2007. How to look beyond what users say that they want. CHI Extended Abstracts on Human Factors in Computing Systems, 1759-1764.

Li, X., Valacich, J.S., & Hess, T.J. 2004. Predicting user trust in information systems: A comparison of competing trust models. 37th Annual Hawaii International Conference on System Sciences (HICSS), Track 8, Vol. 8.

Linda, O., & Manic, M. 2011. Anomaly detection for resilient control systems using fuzzy-neural data fusion engine. Proceedings of the 4th Annual International Symposium on Resilient Control Systems, in press.

McCarty, K., Manic, M., Cherry, S., & McQueen, M. 2011. A temporal-spatial data fusion architecture for monitoring complex systems. Proceedings of the 3rd IEEE Conference on Human System Interactions, 101-106.

Rieger, C.G., Gertman, D.I., & McQueen, M.A. 2009. Resilient control systems: Next generation design research. Proceedings of the 2nd IEEE Conference on Human System Interactions, 632-636.

Shampine, L.F, & Reichelt, M.W. 1997. The MATLAB ODE Suite. Society of Industrial and Applied Mathematics Journal of Scientific Computing, 18: 1-22.

Sproles, D.W., & Bavuso, S.J. 2003. HiRel: Hybrid Automated Reliability Predictor (HARP) Integrated Reliability Tool System (Version 7.0) HARP Output (HARPO) Graphics Display User's Guide. NASA Langley Technical Report.

US Department of Transportation. 2004. Developing and Using Concept of Operation in Transporation Management Systems. Washington, DC: US Department of Transportation.

Xu, J., & Froment, G.F. 1989. Methane steam reforming, methanation and water-gas shift: I. Intrinsic kinetics. American Institute of Chemical Engineers Journal, 35: 88-96.