Improving expressional power and validation for multilevel flow models

Jan Eric Larsson[a, b, *], Jonas Ahnlund[a], Tord Bergquist[a], Fredrik Dahlstrand[b], Bengt Öhman[b], and Lambert Spaanenburg[a]

a) Department of Information Technology, Lund Institute of Technology, P.O. Box 118, 221 00 Lund, Sweden
b) GoalArt, Tunavägen 39C, 223 63 Lund, Sweden
*) Corresponding author. Tel: +46 46 192630, Fax: +46 46 192641, E-mail: janeric@it.lth.se.

Abstract. Multilevel flow modeling (MFM) is a modeling method for complex technical systems in which the goals and functions of the system are explicitly described. MFM can be used as a basis for root cause analysis, where primary root causes are separated from consequential faults, in complex fault situations.

Model representations for use in diagnostic reasoning usually describe causality, between parameters, faults, or process states. However, the causality of a system may vary depending on details in the construction, as well as over time with the process state. One contribution of this paper is a general method of describing varying causality in a simple and efficient way. The method has been tested using multilevel flow models.

Causality is visible in measurements and can be used to increase process understanding. The standard cross-correlation technique is insufficient for causality detection in industrial processes. Another contribution of this paper is a new method that can detect causality in industrial signals, and thus be used to validate the design of multilevel flow models.

Keywords: Alarm analysis, causality, complex technical systems, correlation, multilevel flow models, fault detection, root cause analysis.

1 Introduction

Large industrial plants are difficult to fully understand and operate. Modern control systems provide large quantities of process measurements, but are less well equipped to interpret the data in a way useful for the operator. The existence of unknown or hidden cross-dependencies between the process signals is one of the major causes of upsets and accidents in the process industry today. The operators cannot always understand the consequences of their actions and the alarm systems are inadequate for handling intelligent reasoning and analysis. Therefore additional techniques are required.

Fault detection, fault diagnosis, root cause analysis, and several other diagnostic tasks can be efficiently implemented using models. One such modeling concept is Multilevel Flow Modeling, (MFM). Here, the objectives, the functions, and the causality between different functions of a system are explicitly described.

1.1 Variable Causality

Causality is a central concept in diagnostic model representations. It is used in deduction on whether one fault depends on others or is a primary (or root) cause of a fault situation. For example, consider the simple lab process shown in Figure 1.

A simple example of a fault situation would be if the pump stopped. This would cause a low volume in the upper tank, which in turn would cause a low volume in the lower tank. There would be three faults, whereof one primary (the pump low flow) and two consequential (the low volumes in the tanks). By modeling the causality of the process, a diagnostic algorithm could reduce the three-alarm situation into one primary alarm.

Now consider a similar process, shown in Figure 2. The difference here is that the two smaller tanks are closed and connected to each other and the surrounding equipment with closed pipes. The same analysis as above is valid for this process too. However, the two processes differ in several other aspects.

For example, assume that the right tank is completely full. In the closed tanks case, this high volume will block up the system and cause a high volume in the left tank and a low flow through the pump. In the open tanks case, there is no such “backward” causation. Instead the water will run over the edge of the lower tank, while the upper tank and the pump are unaffected.

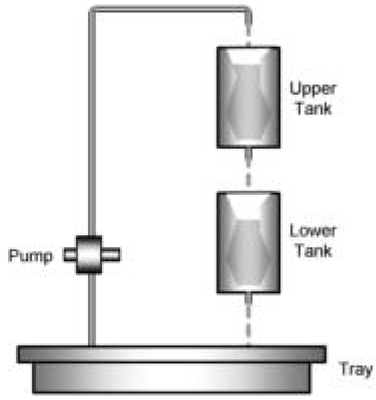

Figure 1. The open tanks process. Water is pumped from a tray tank, up into a small, upper tank with a hole in the bottom. From the upper tank, the water flows down into a lower tank, and then back into the tray tank. The purpose of this process is to train students in basic control theory, by allowing them to try different methods of controlling the level of either the upper of lower tank.

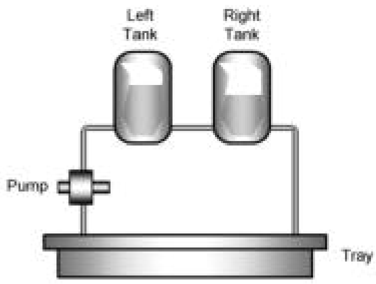

Figure 2. The closed tanks process. The two smaller tanks are *closed* and connected to each other and the surrounding equipment with *closed* pipes.

A way of handling this problem is to describe the two systems using different kinds of modeling entities, with different causal properties. For example, it would be possible to envision a modeling language with concepts such as open and closed tanks, open and closed connections, pumps that are affected by blocks, pumps that are strong enough to force fluid through blocks, etc. The drawback of such an approach is obvious. There will be a need for a very large number of objects, often similar, but with slight differences in causal properties.

Assume further that we have a system of closed tanks with lids, which can be open or closed. Then different causality is present depending on the state of the lids. In another case, the tanks may have an overflow protection, which is either working or broken. Tanks may have underflow protection or not, batteries may have automatic recharging or not, etc., and these may depend on time, operating state, fault status, and so on.

This paper presents another solution to this kind of problem, namely letting the connection between the objects contain information on which causal connections are active at a certain moment. In this way, a small number of modeling objects are sufficient, and the model can easily be given time-dependent behavior.

1.2 Model Validation Using Measured Process Signals

The main difficulty with model-based systems is that the models themselves require an effort to design and validate. If the model does not agree with the real system, the diagnostic results may be wrong, and indeed, it may be better to avoid using the diagnostic system. Thus, models *must* be correct and validated at the design stage, and during the whole lifetime of the system.

Here, we present a correlation algorithm, which can, in a general sense, detect whether there is a dependency or causality between two measured signals in, for example, an industrial plant. By using this algorithm, it is possible to validate a causal model, such as an MFM model, both during design and operation. By comparing the MFM model with the result of the correlation algorithm, it is possible to detect inconsistencies, meaning that either the model is incorrect, or something is wrong in the real system.

There are, of course, well-known correlation algorithms in literature, but simply taking industrial plant signals and feeding them into a standard correlation algorithm does not, as we will see, provide useful results.

2 Multilevel Flow Models

The methods presented have been implemented and tested using one specific modeling representation, multilevel flow models (MFM). MFM are graphical models of goals and functions of technical systems. The goals describe the purposes of a system or subsystem, and the functions describe the capabilities of the system in terms of flows of mass, energy, and information. MFM also describes the relations between the goals and the functions that achieve those goals, and between functions and the sub-goals, which provide conditions for these functions. MFM was invented by Morten Lind at the Technical University of Denmark, see [25]. Several new algorithms and implementations were contributed by Jan Eric Larsson at Lund Institute of Technology, see Larsson [15-18].

MFM provides a good basis for diagnostic algorithms. The work of Larsson [15] presents several algorithms based on MFM. Measurement validation checks consistency between redundant sensor values, and can discover flow leaks, sensor failures, and other measurement errors. The alarm analysis algorithm analyses any (multiple) fault situation and can tell which faults are primary and which faults that may be consequences of the primary ones. Fault diagnosis uses sensor values and queries to the operator to discover the faults of the target system. Other algorithms have been developed later. The failure mode analysis uses MFM with added timing information to predict the consequences of failures. It can be used both during the design phase of a plant and in real-time during actual operation [31, 34]. The fuzzy alarm analysis works in a way similar to the discrete alarm analysis, but is based on fuzzy logic, which makes it more robust when faced with noisy signals close to decision boundaries, see [4-6, 19]. Sensor fault detection spots inconsistencies between signals and can find faulty sensor values [35].

3 An Example of an MFM Model

MFM has been thoroughly explained in [25] and [15, 17]. Here a small example will be given, to show the basics of MFM modeling. We will first use the open tanks process of Figure 1. It should be noted that the open tanks process is just a toy process; MFM is usually used to model large complex systems, for example, a nuclear power plant or the human body.

3.1 The Lab Tanks Model

Each flow in an MFM model is described by flow functions. There are nine different flow functions defined in MFM. Of these nine functions only four are needed to describe the open tanks process.

- The source function represents a component's capability of providing mass, energy, or information. The graphical representation of the source is the leftmost symbol labeled "Tray" in Figure 3.
- The transport function represents a component's capability of moving (transporting) mass, energy or information. The graphical representation of the transport is the symbol labeled "Pump" in Figure 3.
- The storage function represents a component's capability of storing mass, energy, or information. The graphical representation of the storage is the symbol labeled "Upper" in Figure 3.
- The sink function represents a component's capability of draining (consuming) mass, energy, or information. The graphical representation of the source is the rightmost symbol labeled "Tray" in Figure 3.

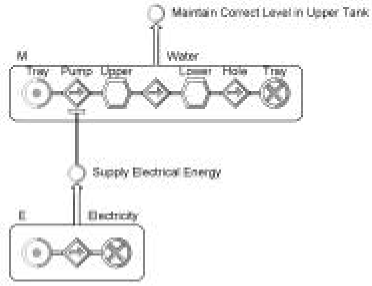

Figure 3. An MFM model of the tanks system. The tray tank is described by a source function, the pump by a transport, the upper and lower tanks by storages, the gravitational outflows by transports, and the tray tank also by a sink function. In addition, the electrical power supply is shown in a sub-flow, which is a condition for the pump to move water. The goal of the water flow is to "Keep correct level in upper tank" and the goal of the lower network is to "Supply electrical energy."

An MFM model of the tanks process is shown in Figure 3. MFM functions in the same flow are grouped together in networks, the networks are shown as rectangles surrounding the functions in Figure 3. The upper network represents the water flow, and the lower network represents the flow of electricity. It should be noted that MFM describes how different flows enable each other. In the simple example in Figure 3, it can be seen that the electrical energy flow is a necessary condition for the pump and thereby for the water flow through the process.

Exactly the same MFM model, as the one shown in Figure 3, is used to describe the closed tank process.

3.2 A Simplified Heating Plant

Here, we will give a slightly larger example, a part of the main systems of a bio-fueled district heating plant. A simplified process graph is shown in Figure 4 and the corresponding MFM model is given in Figure 5.

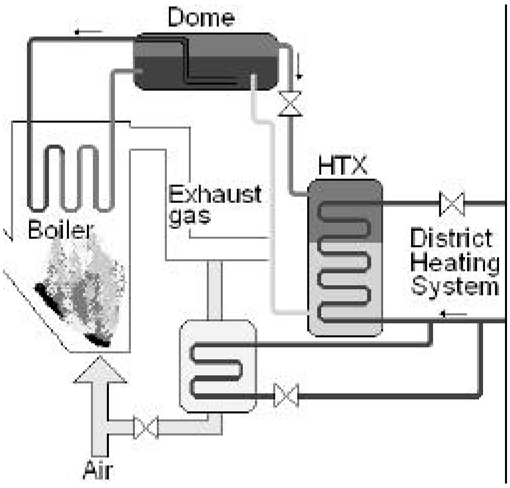

Figure 4. A process graph of the main systems of the test plant. There is a boiler where water is heated by burning bio-mass, a dome where hot water is stored, a heat exchanger, where boiler water is used to heat district heating water, and another heat exchanger (the economizer), where thermal heat in the exhaust is used to heat district heating water.

Again, it should be noted that MFM describes how different flows enable each other, and which flows, volumes, etc., depend on each other. For example, the MFM model in Figure 5 shows the following:

- The temperatures of the dome and the heat exchanger are closely dependent.
- The temperature of the heat exchanger and the incoming district heating water are not dependent.
- The boiler temperature depends on the oxygen concentration in the boiler.

It is also important to observe that the given example is still a very small illustration. MFM is designed specifically to handle large models with thousands of objects. Hopefully, the two examples give a flair for what a larger MFM model looks like. There are more networks and more flow functions, but no qualitative differences. The largest MFM model in existence describes a part of a Swedish nuclear power plant and comprises some 5000 components and 1000 sensor points.

4 Advantages of MFM

All algorithms described in [15] are based on discrete logic where the "sensor" values are low, normal, or high, and the resulting values are consistent or inconsistent, working or failed, primary or consequential, etc. In other

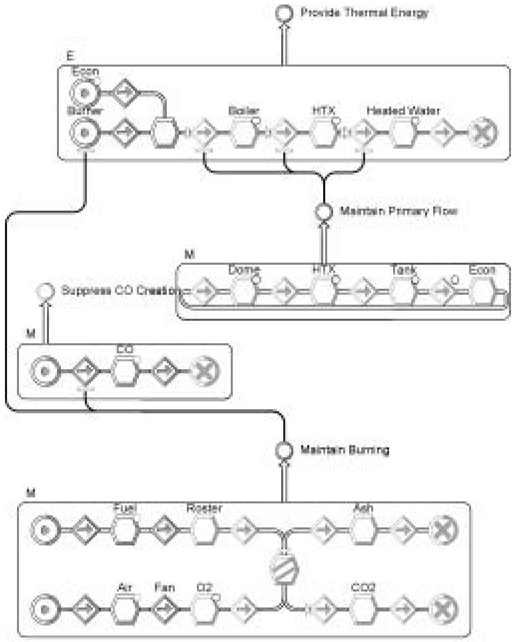

Figure 5. A MFM model of the power plant. The topmost flow describes the flow of thermal energy from the boiler via the dome and the heat exchanger to the district heating system. Below it is a description of the mass flow of water through the plant's circulation system, the so-called primary flow. There is also a small flow describing the generation of carbon monoxide under less than optimal burning conditions, and finally, a description of the mass flows involved in the burning itself, an example of an exothermal chemical reaction.

words, MFM uses a linguistic interpretation of logic variables, just as rule-based expert systems and systems based on fuzzy logic. In addition, the MFM algorithms all operate by searching in fixed graphs and together with the discrete logic, explicit means-end concepts, and graphical nature of MFM, this gives several advantages:

- It allows a complete and efficient solution to the problem of root cause analysis [15, 18] and extends on older techniques such as fault trees [10].
- The explicit description of goals and functions gives a *small semantic gap* between the diagnostic task formulation and the knowledge representation.
- The graphical representation provides strong support for knowledge base overview and consistency.
- The high level of abstraction makes knowledge acquisition, knowledge engineering, and knowledge base validation and support orders of

magnitude easier than with standard rule-based reasoning.

- The graphical nature of the models (limited fixed graphs) allows the algorithms to have good real-time properties, such as an easily computed worst-case time, low memory demands, and high efficiency.
- The high level of abstraction allows the algorithms to be very fast. A worst-case fault diagnosis on a large nuclear power plant model takes less than 4 milliseconds, (worst case) [32].

MFM has been used for medical diagnosis, [20-22]] and nuclear power, [13, 23, 32, 33]. The research of Larsson has spawned the spin-off company GoalArt, which has performed several successful projects, in medical equipment, conventional and nuclear power, and electrical and control systems.

5 Modeling Causality in MFM

The main strength of MFM is the ability to describe very complex systems using a small number of modeling elements (functions). For each of the legal connections between the MFM functions, a set of causal rules has been defined. These causal rules describe how qualitative states of the functions affect each other. The idea of implicit causal rules was invented in [15].

For example, the connection of a transport and a storage (for example, a pump and a tank) implies the following four causations:

1. Transport low => storage low
2. Transport high => storage high
3. Storage low => transport high
4. Storage high => transport low

The causal rules are used with one or more diagnostic methods, such as, alarm analysis, discrete sensor validation, or failure mode analysis.

5.1 Inhibiting Causal Dependencies

In order to explain the inhibition of causal dependencies, consider a simple example. We will use the open tanks process again. In the closed tanks example, all possible causal dependencies are active. Thus, if the right tank overfills, it will stop the outflow from the left tank, which will become overfull too, which will stop the pump flow, and so on. The complete set of causal interactions is shown in Figure 6, where each ellipsis describes one piece of equipment (in fact, flow function). The arcs show causal influences from one state to another.

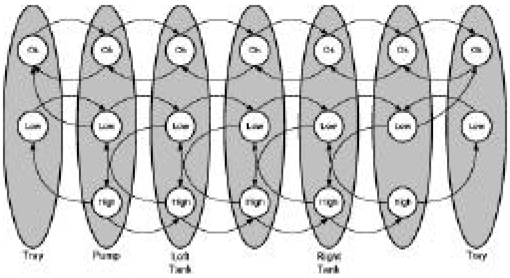

Figure 6. The complete set of causal dependencies for the closed tanks process. Each ellipsis describes one flow function, with the three possible states of OK, failed low, and failed high. The arcs show causal influences from one state to another. A low fault in the leftmost ellipsis (the source function of the tray tank) will cause a low fault in the second ellipsis (the transport function of the pump). This is described by the arc from the "low" in the leftmost ellipsis to the "low" in the second ellipsis, and so on.

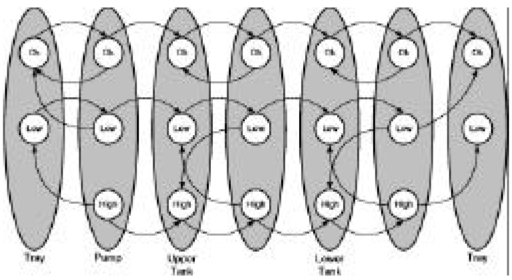

Figure 7. The set of causal dependencies for the open tanks process. Here, some of the dependencies, which occur in the closed tanks case, are missing.

However, when the tanks are not closed, they can be overfull, but the outflows of upstream equipment will not be hindered. In the open tanks case, this means that one of the causal dependencies (from an lower tank high to a upper tank outflow low) will not be valid. In the same way, several other causal dependencies are not valid for the open tanks system. In Figure 7, these casual dependencies have been removed. That is, Figure 7 shows the causal dependencies of the open tanks system.

By explicitly activating or deactivating the casual dependencies of each connection between different functions (or components), similar physical systems with slight differences in causality can be modeled efficiently.

5.2 Variable Causality in MFM

The current method is generally valid for qualitative physics representations, but it was developed with MFM in mind.

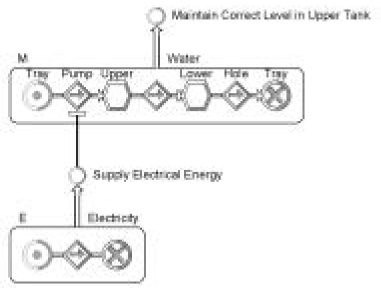

Figure 8. An MFM model of the open tanks process with some active and some inactive causal dependencies. Note the small “shark fins” between some of the transport and storage functions.

In basic MFM it is assumed that all connected functions affect each other in both directions. For example, current MFM assumes that the model in Figure 3 has the causality of the real system in Figure 2, that is, the closed tanks system. There is no way of describing the difference between the systems with open and closed tanks.

As mentioned, one solution is to have more MFM symbols to represent various types of objects such as open tanks, closed tanks, centrifugal pumps, etc. This solution quickly becomes unmanageable, since it may be difficult to find the appropriate symbol to use in a specific system, and the modeling effort becomes large.

The other solution, which is described in this paper, is to add an attribute to the causal relation indicating in which states the relation is active or not. Specifically, in MFM models, this may be indicated graphically by adding one new modeling element.

By selecting casual connections in the MFM model graph, the user can activate or deactivate casual dependencies. The activation state is shown with arrows, so called “shark fins,” see Figure 8. The implementation is simple. Each diagnostic algorithm uses the set of causal rules, and if a causal dependency is deactivated, the corresponding rule is not used in the reasoning.

5.3 The General Case

In the examples above, three values have been used to describe the working state of each object: OK, failed low, and failed high. However, it is possible to base a qualitative physics on more values, such as very low, low, slightly low, etc. It may also be that one object has a different number of states than the one it is connected to. In a general case, a causal connection implies n*m potential dependencies, each in either right or left direction, for a total of 2*n*m arcs. The input matrix shown in Figure 9 can be used to switch all 2*n*m dependencies on or off.

	Obj. 2 State 1	Obj. 2 State 2	...	Obj. 2 State m
Obj. 1 State 1	← →	← →	...	← —
Obj. 1 State 2	← —	← —	...	— →
...	...	...	...	...
Obj. 1 State n	— →	— →	...	← →

Figure 9. A general inhibition matrix for a connection consisting of 2*n*m potential causal dependencies.

5.4 Time-Varying Causality

Another complication concerning variable causality is that of dynamical behavior. The validity of causal dependencies may vary according to time and process state. For example, a closed tank may have a lid, which when opened turns the tank into an open one.

In the proposed solution, each arc (or place in the input matrix) can be connected to a logical signal, which in turn may be a function of time and/or operating state. In this way, it is possible to use the same model to capture time-varying causality, such as tanks with opening lids, batteries with over-charge protection units that break, etc.

5.5 A Comparison with Petersen

The problem of varying causality in MFM was first observed in [15]. The solution suggested there was to use different causal rule sets for different domains, which probably is an intractable solution in practice. Petersen introduces two kinds of causal MFM connections, with different causal inhibitions. Petersen’s solution contains the seed of the method proposed in this article, but goes only part of the way [30].

To illustrate the difference between Petersen and the current method, assume that an MFM transport is connected to a storage. The possible casual dependencies are:

1. Transport low => storage low
2. Transport high => storage high
3. Storage low => transport high
4. Storage high => transport low

In Petersen's first case, all four dependencies are active, while in his second case, dependency 3 is deactivated, while the others are still active. Petersen recognizes only two different cases. In this way, Petersen's solution is limited to using two cases for any 2 * n * m choice. Furthermore, Petersen only discusses static differences, described once and for all in the model, and he ignores dynamical changes in causality. The method proposed in this paper handles all cases and is a completely generalized solution to the problem.

6 A Correlation Method for Industrial Signals

We now turn to the problem of using measured signals to validate MFM models. The general idea is to find causal dependencies between measured signals and compare these "correlations" with the dependencies stated by the MFM model of the process that generated the data sets.

To examine if two signals (time series) x(t) and y(t) are correlated, one can use the mathematical cross-correlation function defined by

$$r_{xy}(k,l) = E\{x(k)y(l)\}$$

and the cross-covariance function defined by

$$c_{xy}(k,l) = E\{(x(k) - m_x(k))(y(l) - m_y(l))\}$$

where m_x and m_y are the mean values of the signals. However, there are drawbacks with these methods. The result of the cross-correlation is dependent on several variables, for example, the length of the series, the magnitude, and the zero-level. Therefore the result is sensitive to scaling, and major disturbances in the signal have a large impact. In process signals, the most interesting disturbances to detect are small variations that are precursors to larger problems.

A new "correlation" method has been developed for use on process signals off line. The method was designed with the following in mind:

- It should provide a measure of how strongly *dependent* two signals are, or in other words, how strong the causal connection between them is.
- The causal connection should provide a basis for the understanding of the process.
- If the measured causality disagrees with the model description, it can point out the need to update the model to match the reality, that is, the correlation algorithm can be used to maintain the model in a validated state.

In order to provide these properties, the correlation algorithm must be designed in a special way. We are not interested in a general signal correlation, but in the correlation between specific signal properties, such as small variations that could indicate larger problems in the future, large variations that indicate process upsets, how often the signals move outside safe levels, so called alarm limits, etc.

The proposed method is called Multiple Local Property Correlation (MLPC) and is based on the following, easily derived, general signal properties:

- Local maximum value (magnitude).
- Local minimum value.
- Local change in the variance of the signal.
- Local maximum change in the difference of two consecutive samples.

To handle negative correlation the local maximum and the local minimum values are merged into one property called maximum distance.

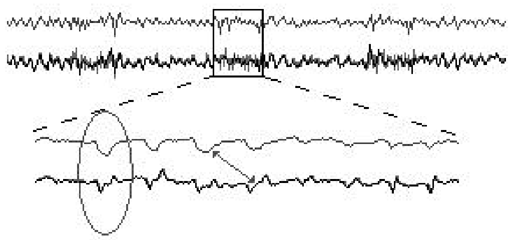

Figure 10. Small variations appear at the same time in both signals. The match inside the circle is of interest but matches that are far from each other are of less importance.

One assumption is important. A change in one of the properties above is only of interest within a short time period. The validity of this assumption depends on process characteristics but is normally true. A change in a pressure measurement should not be correlated with the opening of a valve one week before. Instead, we search for the matching properties within a small time interval (typically a few hours or less), see Figure 10.

It is important that the method is capable of finding small changes even though the signal's measurements are clearly different in mean values, amplitude of variance, and so on. To achieve this, we cut the two signals into equally sized segments. Different properties are then extracted from each segment and compared with corresponding properties in the other signal. These comparisons are then used to create one bar chart for every property. Strongly correlated signals will have almost the same distribution in all of their bar charts. More weakly correlated signals will have different distributions, but the correlated properties will still be detectable.

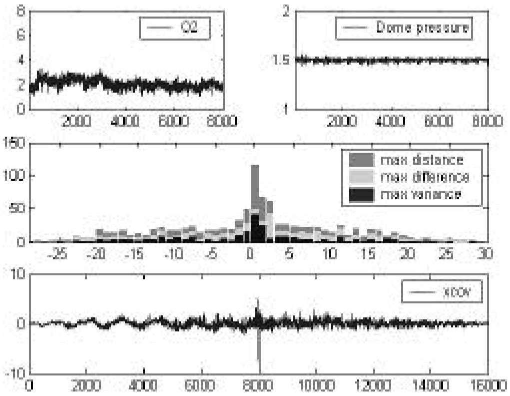

Figure 11. Two correlated steady state process signals, their MLPC correlated bar charts and Matlab's xcov function.

We will use a comparison between MLPC and the mathematical cross-covariance to illustrate the new method's capabilities. Here, the two functions are applied to two process signals from a test plant (the heating plant of Figure 4). The signals are the oxygen level inside the boiler and the pressure in the dome. They have been measured during two weeks in December 2002. First we apply the functions on a period of stable production, see Figure 11. One week later, there was a major disturbance in both signals due to a plant trip, see Figure 12. The disturbance has a significant impact on the cross-covariance function, but has only a modest effect on the MLPC method.

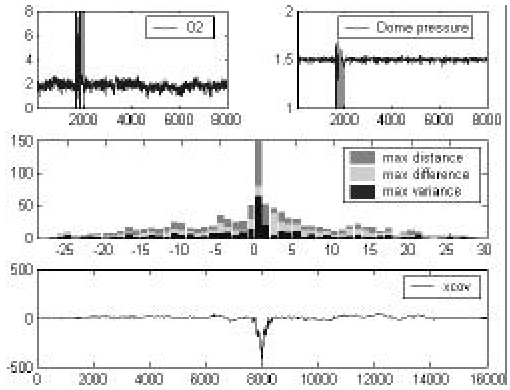

Figure 12. Two correlated process signals suffering a major disturbance, their MLPC correlated bar charts and Matlab's xcov function. Note the scale change (from Figure 11) in the cross-covariance plot. There is only a small change in the output of MLPC.

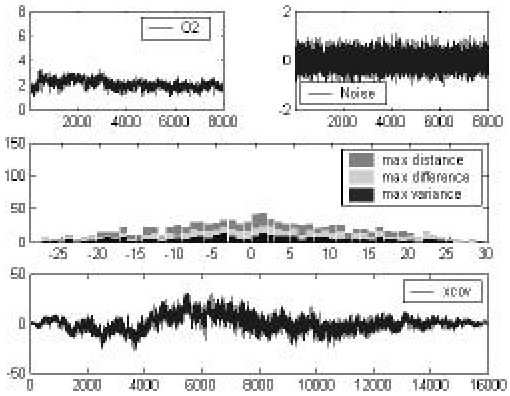

Figure 13. A process signal and normal distributed noise. The absence of peaks shows that the signals are uncorrelated.

It is important to test the method for uncorrelated cases too. In Figure 12 we use the oxygen level again but correlate it with normal distributed noise. The absence of a clear peak shows that the two signals are uncorrelated.

MLPC is robust against noise. In Figure 13, the two process signals have an additive noise with the same variance as the signal themselves. For industrial applications this is a high signal to noise ratio. Again, it can be seen that MLPC is quite robust in its detection of dependencies.

One of the main reasons for developing a new algorithm was the difficulty of making an automatic and reliable decision on whether two signals are correlated or not using the unbounded cross-correlation. With MLPC, the number of segments, N, and the number of samples in every segment, L, are known a priori. The probability that a feature will occur at a specific index, i, is

$$p(i) = \frac{1}{L}.$$

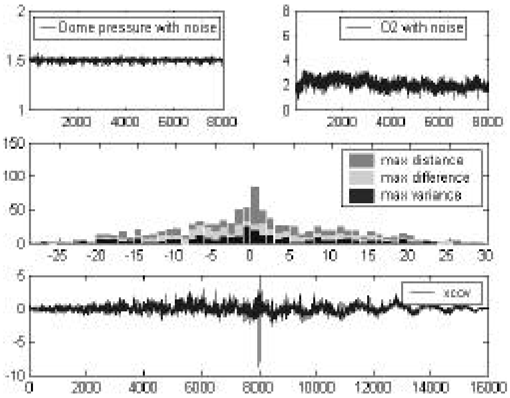

Figure 14. The two process signals with added noise. The noise only has a small effect on the MLPC.

In case that the signals are independent and uncorrelated, the probability of the time shift, t, between the indices in two corresponding segments is

$$p(i - j = t) = (L - |t|)\frac{1}{L^2}.$$

This probability can be used to set boundaries for when two signals are to be considered correlated. In Figure 13 and throughout the next section we will use

$$l(t) = 2N(L - |t|)\frac{1}{L^2}$$

as a limit.

The number of segments bounds the output value, and the bar charts can be merged using a fuzzy system that enable us to put different weights on the properties. The results can be presented in a compact way by selecting one signal and showing the degree of MLPC correlation for the other signals, see Figure 18.

The introduced method is already functional and ready for industrial applications. It is easy to append more functions detecting additional properties, for instance, changes in applied adaptive filter parameters or neural network weights.

7 Experimental Results

The new correlation method has been applied to process signals from a bio-fueled District Heating Plant, Flintrännan (FFC) in Malmö, Sweden. A part of the same process is also described by the MFM-model in Figure 4. We have focused on five signals from the model, the oxygen level, the carbon monoxide level, the dome pressure, and the incoming and outgoing water temperatures, see Figure 14. The signals show a large disturbance in the middle of the period, which affects all signals except the return temperature. The signals do not seem to have any other properties in common, but according to the MFM-model, there should be correlation between some of the signals.

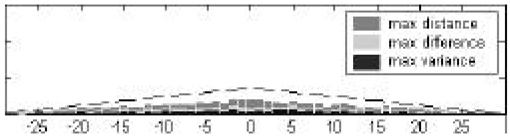

Figure 15. The correlation of two random (noise) signals with the threshold l(t) marked.

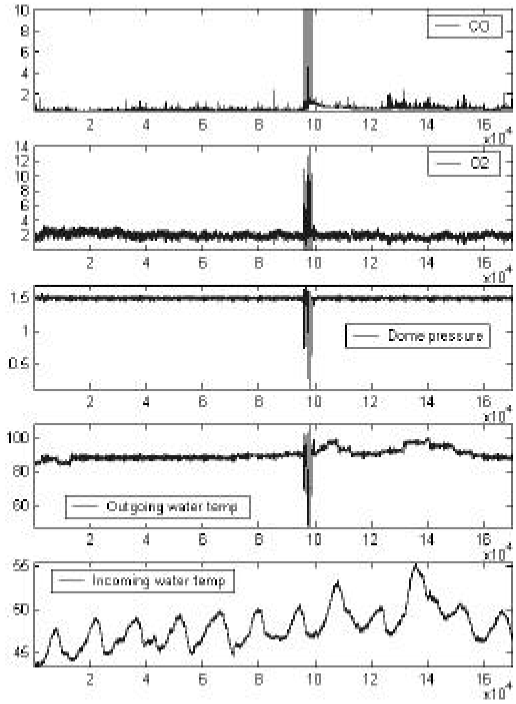

Figure 16. Five measured signals from the heating plant, the oxygen level, the carbon monoxide level, the dome pressure, and the incoming and outgoing water temperatures.

From the model we know that the following signal pairs should be correlated:

- The oxygen level and the carbon monoxide level.
- The oxygen level and the dome pressure.
- The dome pressure and the outgoing water temperature.

Furthermore, the dome pressure should only be slightly correlated with the incoming water temperature. The incoming water temperature should affect the outgoing temperature, but not the other way around. The reason for this is that there are many other contributors to the district heating system. Even if the plant is shut down, the district heating system will still work.

The result of the new correlation method is shown in Figure 15. It is clear that the dome pressure is strongly correlated to the outgoing water temperature. All the three properties (distance, difference, and variance) have high values in the center, (time delay = 0). The oxygen level and the carbon monoxide level also seem to be correlated. However, the max difference property is not concentrated to the center. Instead it has its highest value with a one-minute delay, with no contribution at all in the center. The correlation between the oxygen level and the dome pressure is barely noticeable. Again, the max difference property is delayed, this time with two minutes. As we suspected, the correlation between the dome pressure and the incoming water temperature is marginal. The center value comes mostly from the max variance property.

There seems to be a strong center max variance property value in all the signals, even from signals that should not be correlated. This probably comes from small but global disturbances in the process, that is, process noise.

The new correlation method can find correlation in process signals, which are not detectable with the ordinary cross-covariance function. With the use of our method, we can validate the model, that is, point out inconsistencies with real measurement values.

The presented method has already been used in model validation. When the MFM model of the district heating plant, Figure 4, was first developed, the authors assumed that the temperature in the local heating network would depend on the temperature in the heat exchanger. This is a natural assumption, because the plant is used to provide thermal energy to the network.

However, the local network is supported by several heating plants and well controlled, and variations in temperature in a single plant do not propagate into the network. In fact, limited variations cannot even be observed in the network temperature. Furthermore, the flow of this thermal energy described by the model is a combination of the water temperature and water flow, which is regulated with a shunt-valve. It is therefore not enough to compare the two temperatures without taking the water-flows in consideration. We first realized these facts when we applied MLPC and failed to discover any correlation. After further analysis, we realized that there is in fact no dependency, and the model was updated. Thus, the model in Figure 4 has been validated by our method and we have already discovered and corrected one model fault.

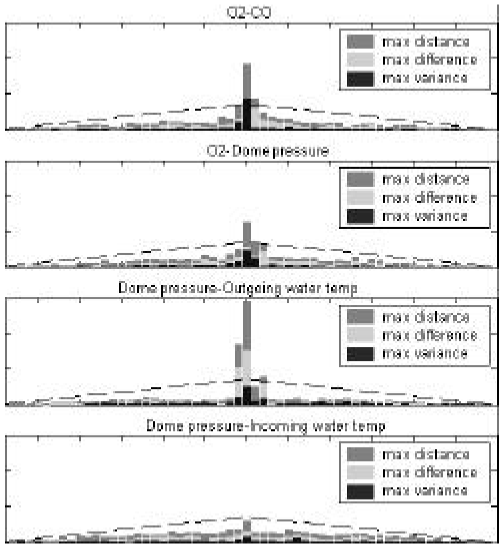

Figure 17. The results of correlating the different signals from the heating plant.

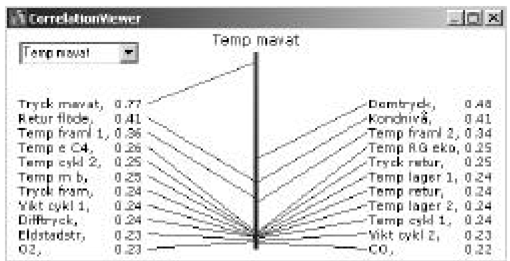

Figure 18. A compact way of showing the degree of MLPC correlation between one signal and the rest of the signals of the plant. The feed water temperature is strongly correlated to the feed water pressure but less dependent on other variables.

8 Properties of the MLPC Method

MLPC is an algorithm that can detect causal connections in industrial process. The algorithm is more sensitive to small changes than, for instance, the standard cross-correlation function, and it is at the same time more robust against noise. It is therefore suitable to use as a tool when trying to detect dependencies between real process signals. The simple structure and the merging of detected properties by a fuzzy system make it easy to enhance the MLPC with other signal properties.

There is a need for an efficient way of validating models. Using MLPC to create a dependency model from the process signals and compare its consistency with an MFM model is straightforward and will be further evaluated.

The correlation method can also be used in other applications. In previous papers, the authors have described methods for reducing the number of faulty alarms from an industrial plant by analyzing signal history and moving the alarm limits and adding filters, a so called alarm cleanup, [1, 2]. MLPC could be used to guide the selection of new candidate signals for cleanup. If one signal is giving off many alarms, while a correlated one gives no alarms, this is an indication that at least one of the signals has wrongly set alarm limits.

9 Related Work

The main contributions to MFM have been made by Morten Lind and his group. Lind [25, 28] describes the basics of MFM, while [26] contains an early suggestion for a diagnostic system. Lind has also treated real-time diagnosis [27], and design of operator interfaces [24].

MFM has also been used in nuclear safety research, [3, 7, 14], in operator interfaces for fault diagnosis [8], for constructing COGSYS diagnostic systems [37], for fault diagnosis in process industry [38], and in intelligent man-machine systems for nuclear plants [29].

Larsson has used MFM in monitoring and diagnosis for intensive-care units [20-22].

MFM is the main basis for the startup company GoalArt, which works with diagnostics for various branches of industry, for example nuclear power generation [23].

MFM can be compared to other modeling and diagnosis methodologies, such as rule-based expert systems, fuzzy logic, qualitative physics based on Reiter's algorithm [11, 12, 36], classical statistical methods, methods from control theory [9], and neural networks. In comparison to expert systems and fuzzy logic, MFM imposes a deep model structure of means and ends, as opposed to a shallow rule-based representation. It differs from qualitative physics in that it explicitly represents goals and functions, avoids general logic, and is computationally more efficient, while qualitative physics has been geared towards diagnosis of electrical circuits, a task which MFM is not very well adapted for. MFM differs from statistical and control theory methods in that it uses discrete and more abstract representations, and thus is useful on a higher level of decision and diagnosis. For example, control theory methods are usually aimed at fault detection on control loop level, while MFM is aimed at diagnostic reasoning on a plant-wide level. Finally, MFM differs strongly from neural networks in that it explicitly represents human knowledge using linguistic concepts, and that the model construction relies almost completely on available human knowledge and not on automatic generalization of test cases.

10 Conclusions

In order to enable practically useful alarm handling and diagnostic reasoning, a good modeling technology is needed. This article has described how multilevel flow models can be extended to handle differences in construction and behavior, that is, variable causal properties. We present a general solution, which allows efficient modeling of systems with varying causality without a large multiplication of language objects. The proposed solution also allows varying causality depending on time and operating state.

The article also presents a new correlation-style method for analyzing groups of signals to find dependencies. The aim of the presented method is to give a measure of how well a model and an industrial process match. In this way, the model of the process can be validated and kept correct. With a validated model, the method will also be able to detect faults in the process, that is, to provide fault detection. Finally, the method provides process understanding for designers, operators, and service technicians.

11 Acknowledgements

The authors would like to thank Morten Lind, who invented MFM and have supported our efforts from the beginning, and Johannes Petersen for his contributions on varying causality in MFM connections. We would also like to thank the staff at GoalArt, Antonio Calzada, David Sjölander, Jan Tuszynski, and Anu Uus for providing inspiration, help, and a nice mental environment in our daily lives. Furthermore, we are grateful to Professor Gustaf Olsson for his excellent ideas and encouragement. Special thanks go to Ann-Britt Östberg and Tomas Erlandsson, Carl Bro

Energikonsult AB, Johnny Pettersson, Sydkraft Värme Syd AB, Johan Söderbom, and Lennart Gårdman, Vattenfall Utveckling AB, for their support.

Part of this work has been sponsored by VINNOVA, the Swedish Agency for Innovation Systems under the project number AIS-33, and by Värmeforsk, the Swedish Thermal Engineering Research Institute under the project number P4-312.

References

[1] J. Ahnlund, T. Bergquist, and M. Råberg, "Larmsanering med Generella Metoder," Värmeforskrapport P4-211, Värmeforsk, Stockholm, Sweden, September, 2003.

[2] T. Bergquist, J. Ahnlund, and J. E. Larsson, "Alarm Reduction in Industrial Process Control," Proceedings of the ETFA 2003, Lisbon, Portugal, September, 2003

[3] T. Businaro, A. Di Lorenzo, G. B. Meo, M. I. Rabbani, and E. Rubino, "An Application of MFM Method for Nuclear Plant State Identification," Proceedings of the Halden Programmer's Group Meeting on Computerized Man-Machine Communication, Gothenburg, 1985.

[4] F. Dahlstrand, "Alarm Analysis with Fuzzy Logic and Multilevel Flow Models", Proceedings of the 18th Annual International Conference of the British Computer Society Special Group on Expert Systems, ES98, Cambridge, England, pp.173–188, 1998.

[5] F. Dahlstrand, *Methods for Alarm Reduction with Multilevel Flow Models of Industrial Processes,* Licentiate thesis, Department of Information Technology, Lund University, Lund, Sweden, 2000.

[6] F. Dahlstrand, "Consequence Analysis Theory for Alarm Analysis," *Knowledge-Based Systems,* vol. 15, no. 1-2, pp. 27-36, 2002.

[7] M. K. De, J. A. Rumancik, A. J. Impink, and J. R. Easter, "A Functional Design Approach to PWR Safety," Proceedings of the International Meeting on Thermal Nuclear Reactor Safety, Chicago, Illinois, 1982.

[8] K. D. Duncan and N. Prætorius, "Flow Displays Representing Complex Plant for Diagnosis and Process Control," Proceedings of the 2nd European Meeting on Cognitive Science Approaches to Process Control, Siena, 1989.

[9] P. M. Frank, "Analytical and Qualitative Model-Based Fault Diagnosis ? A Survey and Some New Results," *European Journal of Control,* vol. 2, pp. 6–28, 1996.

[10] J. B. Fussell, "A Review of Fault Tree Analysis with Emphasis on Limitations," Proceedings of the 6th Triennial World Congress of the International Federation of Automatic Control, Pittsburgh, Pennsylvania, 1975.

[11] R. Greiner, B. A. Smith, and R. W. Wilkerson, "A Correction to the Algorithm in Reiter's Theory of Diagnosis," *Artificial Intelligence,* vol. 41, pp. 79–88, 1989.

[12] W. Hamscher, L. Console, and J. de Kleer, (Eds.), Readings in Model-Based Diagnosis, Morgan Kaufmann, San Mateo, California, 1992.

[13] D. Ingström, "MFM Modeling and Alarm Analysis of the Barsebäck Nuclear Power Plant," Master's thesis, Department of Information Technology, Lund Institute of Technology, Lund, 1998.

[14] K. Inoue and E. J. Henley, "Computer-Aided Reliability and Safety Analysis of Complex Systems," Proceedings of the 6th Triennial World Congress of the International Federation of Automatic Control, Pittsburgh, Pennsylvania, 1975.

[15] J. E. Larsson, *Knowledge-Based Methods for Control Systems,* Doctor's thesis, TFRT–1040, Department of Automatic Control, Lund Institute of Technology, Lund, 1992.

[16] J. E. Larsson, "Diagnostic Reasoning Strategies for Means-End Models," *Automatica,* vol. 30, no. 5, pp. 775–787, 1994.

[17] J. E. Larsson, "Diagnosis Based on Explicit Means-End Models," *Artificial Intelligence,* vol. 80, no. 1, pp. 29–93, 1996.

[18] J. E. Larsson, "Diagnostic Reasoning Based on Means-End Models: Experiences and Future Prospects," *Knowledge-Based Systems,* vol. 15, no. 1-2, pp. 103-110, 2002.

[19] J. E. Larsson and F. Dahlstrand, "New Algorithms for MFM Alarm Analysis," invited paper, Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, San Diego, California, 1998.

[20] J. E. Larsson and B. Hayes-Roth, "Guardian: An Intelligent Autonomous Agent for Medical Monitoring and Diagnosis," *IEEE Intelligent Systems,* vol. 13, no. 1, pp. 58–64, 1998.

[21] J. E. Larsson, B. Hayes-Roth, and D. M. Gaba, "Goals and Functions of the Human Body: An MFM Model for Fault Diagnosis," *IEEE Transactions on Systems, Man, and Cybernetics,* vol. 27, no. 6, pp. 758–765, 1997.

[22] J. E. Larsson, B. Hayes-Roth, D. M. Gaba, and B. E. Smith, "Evaluation of a Medical Diagnosis System Using Simulator Test Scenarios," *Artificial Intelligence in Medicine,* vol. 11, pp. 119–140, 1997.

[23] J. E. Larsson, "Avoiding Human Error," Proceedings of the International Conference on Control and Instrumentation in Nuclear Installations, Bristol, England, 2000.

[24] M. Lind, "Human-Machine Interface for Diagnosis Based on Multilevel Flow Modeling," Proceedings of the 2nd European Meeting on Cognitive Science Approaches to Process Control, Siena, 1989.

[25] M. Lind, "Representing Goals and Functions of Complex Systems — An Introduction to Multilevel Flow Modeling," Technical report, 90–D–38, Institute of Automatic Control Systems, Technical University of Denmark, Lyngby, 1990.

[26] M. Lind, "Abstractions Version 1.0 — Descriptions of Classes and Their Use," Technical report, 90–D–380, Institute of Automatic Control Systems, Technical University of Denmark, Lyngby, 1990.

[27] M. Lind, "An Architecture for Real-Time MFM Diagnosis," Technical report, Institute of Automatic Control Systems, Technical University of Denmark, Lyngby, 1990.

[28] M. Lind, "Modeling Goals and Functions of Complex Industrial Plants," *Applied Artificial Intelligence,* vol. 8, no. 2, pp. 259–283, 1994.

[29] K. Monta, J, Takizawa, Y. Hattori, T. Hayashi, N. Sato, J. Itoh, A. Sakuma, and E. Yoshikawa, "An Intelligent Man-Machine System for BWR Nuclear Power Plants," Proceedings of AI91 — Frontiers in Innovative Computing for the Nuclear Industry, Jackson, Wyoming, 1991.

[30] J. Petersen, *Knowledge-Based Support for Situation Assessment in Human Supervisory Control,* Ph.D. thesis, Department of Automation, Technical University of Denmark, Lyngby, Denmark, 2000.

[31] B. Öhman, "Failure Mode Analysis Using Multilevel Flow Models," Proceedings of the 5th European Control Conference, Karlsruhe, Germany, 1999.

[32] B. Öhman, "Code Generation for Alarm Analysis with Multilevel Flow Models," Proceedings of the Second International Symposium on Engineering of Intelligent Systems, EIS '2000, Paisley, Scotland, 2000.

[33] B. Öhman, "Alarm Analysis on Large Systems Using Multilevel Flow Models," Proceedings of the IFAC Symposium on Artificial Intelligence in Real-Time Control, Budapest, AIRTC–2000, Hungary, 2000.

[34] B. Öhman, *Real-Time Diagnosis of Industrial Processes Using Multilevel Flow Models,* Licentiate thesis, Department of Information Technology, Lund University, Lund, Sweden, 2001.

[35] B. Öhman, "Discrete Sensor Validation with Multilevel Flow Models," *IEEE Intelligent Systems,* vol. 17, no. 3, pp. 55-61, 2002.

[36] R. Reiter, "A Theory of Diagnosis from First Principles," *Artificial Intelligence,* vol. 32, pp. 732–737, 1987.

[37] J. M. A. Sassen, *Design Issues of Human Operator Support Systems,* Doctor's thesis, Faculty of Mechanical Engineering and Marine Technology, Laboratory for Measurement and Control, Delft University of Technology, Delft, 1993.

[38] J. Å. Walseth, *Diagnostic Reasoning in Continuous Systems,* Doctor's thesis, ITK-rapport 1993: 164–W, Division of Engineering Cybernetics, Norwegian Institute of Technology, Trondheim, 1993.