Fault detection and precedent-free localization in numerically discretized thermal–fluid systems

K.P. Carpenter, D. Djurdjanovic *, A.K. da Silva

Department of Mechanical Engineering, The University of Texas at Austin, 204 E. Dean Keeton St. Stop C2200, Austin, TX 78712-1591, USA

A R T I C L E   I N F O

Keywords:
Anomaly detection
Anomaly localization
GSMMS
Channel flow

A B S T R A C T

This paper uses the Growing Structure Multiple Model System (GSMMS) method for fault detection and precedent-free localization of unwanted heating anomalies in two different configurations of channel flow systems operated under dynamic conditions: (i) straight channel and (ii) straight channel with an internal flow disruptor. Unlike commonly used fault detection methods, the newly proposed approach does not require prior information regarding the fault location, fault severity or data emitted in the presence of a fault to build the model of that fault and recognize it. The new detection mechanism is based only on the models of normal behavior for various portions of the monitored system. The obtained results indicate that the detection and localization of the unwanted heating element (i.e., heat source) can be achieved through distributed GSMMS-based anomaly detection, with multiple anomaly detectors monitoring different parts of each configuration. The results also suggest that fault detection and localization are strongly related to a system’s configuration and operational conditions.

C 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Thermal–fluid systems, such as heat exchangers and fuel cells, are host to a variety of potential problems, such as fouling, overheating, leakage, and general wear and deterioration (Calisto, Martins, & Afgan, 2008; Karlsson, Arriagada, & Genrup, 2008). Not only do these problems negatively affect a system’s performance, but their effects develop slowly and may go unnoticed until abrupt failure occurs. Additionally, besides the difficulty of detecting the presence of faults, its localization can also be challenging as its effects might propagate throughout the system.

In the last two decades, fault detection methods applied to thermal–fluid systems have mostly used data-driven models. For example, the use of neural networks for fault detection in thermal science has been increasing in the recent years due to their robustness in dealing with more complex phenomena (Yang, 2008). For instance, fault detection methods have been created for a steam turbine (Karlsson et al., 2008), furnace (Calisto et al., 2008), a solar water heater (Kalogirou, 2008), an internal combustion engine (Wu, Huang, Chang, & Shiao, 2010), and a heat exchanger (Garcia, 2012). In each method, the system was modeled using a neural network, and their modeling residuals were analyzed and compared with residuals of a known fault. This means that in the aforementioned studies, these fault detection methods were precedent-based methods (i.e., prior knowledge regarding the potential faults and their effects are required).

The major drawback of this approach is that such methods are limited to situations when faulty behavior data is readily available for training of a dynamic model (neural network or any other model) that is representative of that faulty behavior model or if a priori knowledge of the fault characteristics exists. Clearly, these constraints limit the applicability of precedent-based methods, because as system complexity increases, it becomes infeasible to anticipate all possible faults, at all possible locations, and for all possible working regimes of the system. Also, many existing fault detection methods are only capable of detecting a single fault occurring, whereas the development and occurrence of problems in multiple parts of a thermal–fluid system are not uncommon. Thus, a more sophisticated fault detection method would be desirable.

Based on the above, this paper uses a precedent-free localization technique based on the recently introduced Growing Multiple Model System (GSMMS) method for modeling system dynamics to detect and localize anomalies in a thermal–fluid system. The term “precedent-free” indicates that this approach requires only normal system behavior data to achieve localization of the source(s) of anomalous behavior. Previously, this approach has been applied to lumped parameter systems, such as an electronically controlled throttle system (Liu, Djurdjanovic, Marko, & Ni, 2009), an exhaust gas recirculation (EGR) system (Cholette & Djurdjanovic, 2009,
2.1. Overview of the GSMMS-based detection and localization method

Fig. 1 describes a brief overview of the approach. Training data, which is comprised of normal system inputs/outputs, is used by the GSMMS to create a dynamic model of the system. This system model can then be given a set of new inputs to predict the corresponding outputs of the system. The predictions are then compared with the actual system output, and these comparisons (residuals) are statistically analyzed to determine if the system is behaving normally or not.

The foundation of this method is modeling of normal system behavior using the GSMMS approach, which uses a growing self-organizing network (SON, (Kohonen, 1995)) to partition the operating space into local models. Using successive passes through a set of training data, as described in Section 2.2 (training data consists of inputs and outputs corresponding to normal system behavior), the SON adjusts itself during training to appropriately

![Diagram](image-url)

**Fig. 1.** Overview of the GSMMS-based approach to anomaly detection and localization.
partition the operating space, as described in detail in Section 2.3. Each region is then approximated using an analytically tractable linear model. Besides the local model tractability, another advantage of this “divide and conquer” approach to system modeling is the capability to locally monitor the residuals (the differences between the GSMMS-predicted outputs and the monitored, actual outputs) within each GSMMS sub-region.

Abnormal behavior is then recognized by statistically evaluating the local residuals. Abnormal behavior is quantitatively assessed by computing the overlap between the probability density function (PDF) of the current, monitored residuals and the PDF of the normal residuals obtained during training for each region (the overlap is represented by the “confidence value” or CV). A global CV can then be created from the CVs from all the GSMMS regions, as described in Section 2.4. A CV value at or near 1 indicates normal behavior, while a CV less than 1 indicates the current system behavior is deviating from normal behavior. Together, the GSMMS model of normal behavior of the region and the CV based on the interpretation of modeling residuals enable detection of abnormalities in system behavior and thus, can be termed an anomaly detector. Localization of the abnormal behavior can then be done using multiple anomaly detectors, where each anomaly detector each monitors a pertinent sub-system, as discussed in Section 2.5. The method is outlined in Fig. 2.

2.2. GSMMS modeling approach

The GSMMS uses a multiple model structure, where each local model domain is defined by the self-organizing network (SON, Kohonen, 1995) induced Voronoi tessellation of the state-space of the model (Liu et al., 2009). The weight vectors $\xi_m, m = 1, 2, \ldots, M$ of a SON define a Voronoi tessellation:

$$V_m = \{x : \|x - \xi_m\| < \|x - \xi_j\|, \forall m \neq j\}$$

(1)

In this way, the operating space is divided into sub-regions, $V_m$, of “similar” input–output patterns. Within each of the $M$ regions, local models are assumed to be of the linear form

$$F_m(s(k)) = a_m^T s(k) + b_m$$

(2)

where $a_m$ and $b_m$ denote the vectors of parameters of the local model $m$ and

$$s(k) = [y^T(k), \ldots, y^T(k - n_y + 1), u^T(k - n_u), \ldots, u^T(k - n_u - n_b + 1)]^T$$

(3)

---

**Fig. 2.** Overview on how a system is modeled, how an anomaly is detected, and how an anomaly is localized.
is the state vector, where \( y(k) = [y_1(k), \ldots, y_p(k)]^T \) is a vector of \( p \) outputs and \( u(k) = [u_1(k), \ldots, u_q(k)]^T \) is a vector of \( q \) inputs of the system, \( n_d \) is the time delay between when the input reaches the output, and \( n_a \) and \( n_b \) are respectively the autoregressive and external input orders of the local model. Following Johansen and Foss (1995), the global model is then defined as

\[
y(k + 1) = \sum_{k=1}^{M} v_m(s(k))F_m(s(k))
\]

(4)

where \( \nu(s(k)) \) describes how the local models are interpolated into the global model. Following Liu (2007), \( \nu(s(k)) \) is a simple gating function

\[
v_m(s(k)) = \begin{cases} 1 & s(k) \in V_m \\ 0 & \text{otherwise} \end{cases}
\]

(5)

that says that each local model \( F_m(s(k)) \) is only valid in region \( m \).

2.3. Training the GSMMs

The training process for the GSMMs yields the structural parameters of the model and determines the valid regions of each model and the local model parameters of each region. The structure of the model is determined by the weight vectors of the SON underlying the model. The Voronoi tessellation induced by the weight vectors partitions the operation space into regions of similar input–output patterns. The weight vectors are obtained through unsupervised clustering of input/output vectors, \( s(k) \) (state-vectors), in the training set.

Essentially, training consists of successive passes through the training data corresponding to the normal behavior of the system, which updates and adjusts the SON weight vectors (and thus the resulting state-space partition induced by the corresponding Voronoi tessellation), as well as the local model parameters within each Voronoi region. Unsupervised clustering of the SON weight vectors, \( \zeta_m, m = 1, 2, \ldots, M \), is accomplished via recursive adjustments.

\[
\zeta_m(k + 1) = \zeta_m(k) + \zeta_m(k)h(k, d_m(b(k)))(s_m - \zeta_m(k))
\]

(6)

where \( k \) is the index of the training item \( s(k), \) and \( s_m \) is the sample mean of the training vectors for which \( b(k) \) is the Best Matching Unit (BMU). For each training sample, a BMU, \( b(k), \) is defined as

\[
b(k) = \arg \min_{m} \|s(k) - \zeta_m\|
\]

(7)

The function, \( h(k, d_m(b(k))) \), is the neighborhood function, which describes how each vector is updated using training samples in neighboring regions and defined as

\[
h(k, d_m(b(k))) = \exp \left( \frac{-d_m(b(k))^2}{2\sigma^2(k)} \right)
\]

(8)

The neighborhood function shrinks with increasing distance away from the BMU and with increasing passes through the data. The width parameter, \( \sigma^2(k) \), defines the effective range of the weighting function and decreases to zero as \( k \to \infty \) to achieve convergence and global ordering of the SON (Kohonen, 1995). In other words, at the beginning of the training, each training sample has an effect on the parameters of a wide area of local models, but as \( k \to \infty \), the affected area of local models narrows because the further away a given region \( m \) is from the BMU, the less significant the effects are of the current observation on the parameter estimates of the model in region \( m \). Following Cholette and Djurdjanovic (2009), in this paper, we adopt \( \sigma(k) = \frac{1}{k} \). Finally, the term, \( \text{dis}(m(b(k))) \), denotes the shortest topological distance between the node \( m \) and the BMU and is found using the Breadth-first procedure (Sedgewick, 1995).

The penalty term, \( \zeta_m(k) \) in Eq. (6) helps achieve a more accurate model by balancing the effects of visiting frequencies and model errors across different regions (Liu et al., 2009). If a region is not frequently visited, the region could be poorly approximated and in need of more local models. Furthermore, if modeling errors in a given region are high, that region may need to be redefined and additional SON models should move toward it. The distance-based weight vector updating Eq. (6) already ensures that the state space is partitioned according to visitation frequencies (more nodes in more frequently visited regions). Conversely, the penalty term is described by

\[
\zeta_m(k) = \frac{e_m(k)}{\sum_{m=1}^{M} e_m(k)}
\]

(9)

where \( e_m(k) \) is the root mean squared (RMS) modeling errors in the \( m \)th region. Therefore, the weight vectors will tend to move towards regions with higher modeling errors, resulting in a finer partition in those areas.

The local model parameters in region \( m \) are determined by minimizing the sum of the weighted squared output errors in each region using

\[
j_m(\theta_m) = \frac{1}{K} \sum_{k=1}^{K} w_m(s(i))|y(i) - \hat{y}_m(s(i))|^2
\]

(10)

where \( \theta_m \) denotes the model parameters for the \( m \)th region, \( y(i) \) is the training output at time \( i \), and \( \hat{y}_m(s(i)) \) is the predicted output of model \( m \) at time \( i \). The weighting function, \( w_m(s(i)) \), determines the effect of sample \( i \) on the neighboring regional models estimates, and is defined as

\[
w_m(s(i)) = \exp \left( \frac{-\text{dis}(m, b(k))^2}{2\sigma^2(k)} \right), m = 1, 2, \ldots, M
\]

(11)

In this way, each training sample affects all the local models, with this effect lessening as the distance from the model corresponding to the BMU grows.

The self-organizing network is allowed to "grow" after a predetermined number of passes through the training data by inserting a new node half-way between the two nodes corresponding to the two poorest performing models. This growing-gas-like (Martinetz, Berkovich, & Schulten, 1993) mechanism allows the underlying model structure to grow and adapt to the data. Training ends when one of the following two stopping criteria was met, (1) the total RMS error was below a pre-determined tolerance or (2) the number of nodes exceeded a pre-determined number.

2.4. Analysis of the residuals

Once training is finished, the statistical characteristics of the modeling residuals during normal behavior are known, where the residuals are defined as the differences between the actual system output and the GSMMs-predicted output. If any anomaly occurs in the system, i.e., if the system dynamics changes in any way, the modeling residuals behavior in at least some GSMMs regions changes as well. Thus, to detect an anomaly, one can compare the characteristics of the training residuals with that of the current residuals, and abnormal behavior can be indicated when differences are detected.

The operating regions within the GSMMs all have different levels of approximation accuracy. If the system inputs change and drive the system to a different operation region, the modeling residuals could also change. Thus, the modeling residuals can change for reasons other than an anomaly. The "divide and conquer" framework of the GSMMs models is able to work around these potential false alarms because residual interpretation can be done based on the simpler, regional residuals.
Following Liu (2007), system deviation inside each region $m$ is quantified using the concept of regional confidence values (CVs), defined as

$$CV(m, k) = \frac{|f_m(e) \cdot g_m(e, k)|}{|f_m(e)||g_m(e, k)|}$$  \hspace{1cm} (12)

where $f_m(e)$ is the probability density function (PDF) of the modeling residuals displayed during normal behavior, and $g_m(e, k)$ is the PDF of the residuals corresponding to the current behavior at time $k$. $||$ denotes the inner product between two products and $||$ denotes the norm of the function.

The regional confidence value, $CV(m, k)$ can be seen as describing the normalized area of the overlap of the PDFs in that region. If $CV(m, k) = 1$, the current residual PDF matches with the residual PDF obtained during training, which indicates normal behavior. The PDF $f_m(e)$ was approximated using Gaussian Mixture Models due to their universal approximation capability (McLachlan & Peel, 2000), and $g_m(e, k)$ was calculated by updating $f_m(e)$ recursively during operation (Zivkovic & van der Heijden, 2004). To simplify monitoring, a single, global CV is created as the geometric mean of the regional CVs.

2.5. Fault localization

Using only normal behavior data, fault localization can be achieved using the paradigm of distributed anomaly detection, where a set of anomaly detectors (ADs) are used to monitor pertinent sub-systems. Thus, if the source of the anomalous behavior occurs in one of the sub-systems, only the CVs from the affected sub-systems will indicate abnormal behavior, whereas the CVs from the unaffected sub-systems will indicate normal behavior. Thus, fault localization can be achieved.

3. Numerical simulations

Two systems were modeled: (i) a simple channel flow with purely convective heat transfer, and (ii) a channel flow with a solid flow disruptor with conductive and convective heat transfer, where both systems were previously studied numerically from a heat transfer perspective in references da Silva, Lorente, and Bejan (2004) and Young and Vafai (1998), respectively. These two configurations and their respective geometrical variables are shown in Fig. 3. The purpose of having two configurations was to observe the effect of the complexity of the geometry of the configuration on the effectiveness of the detection and localization approach.

![Fig. 3. Representation of the 2-D channel flow numerical domain: (a) straight channel (adapted from da Silva et al. (2004)) (b) straight channel with a flow disruptor (adapted from Young and Vafai (1998)).](image)

To validate the simulations for both configurations, the present study relied on the geometric features and mathematical formulations of both systems as reported in references da Silva et al. (2004) and Young and Vafai (1998), which will be briefly mentioned below as detailed information can be found in the original references. For both configurations, a two-dimensional channel with laminar flow conditions and constant properties were assumed.

The mass, momentum, and energy equations, shown below, were solved within the fluid part of the numerical domain for both configurations:

$$\nabla \cdot \mathbf{v} = 0$$  \hspace{1cm} (13)

$$\rho \frac{D\mathbf{v}}{Dt} = -\nabla p + \mu \nabla^2 \mathbf{v}$$  \hspace{1cm} (14)

$$\rho c_p \frac{DT}{Dt} = k \nabla^2 T$$  \hspace{1cm} (15)

Note that differently from the mathematical model presented in references da Silva et al. (2004) and Young and Vafai (1998), the set of equations above is time dependent due to the need to create anomalies, here represented by the imposition of transient values for a heat flux dissipated in the heat source of each of the domains – the anomalies imposed to the system will be detailed in Section 4. Also, for the solid conducting domain present in Fig. 3b, the 2-D transient conduction equations was simultaneously solved along with the equations above. The interface between the fluid and solid domains was coupled by equating the heat fluxes between these two domains.

For convenience, all equations and relevant parameters were normalized with their respective set of non-dimensional variables, however, for conciseness, we only report the ones that are discussed in this study, which can be seen in Table 1.

As for the boundary conditions, while both configurations used a non-slip condition inside the channel, the configuration shown in Fig. 3a had an extended outlet domain with slip ($L_{out} = 0.5 L$). Also, both configurations used zero gradients (i.e., hydrodynamic and thermal) at the exit in the flow direction. As for the inlet, the configuration shown in Fig. 3a has a uniform velocity imposed, and the one in Fig. 3b has a fully developed parabolic profile. Furthermore, all internal walls of both configurations were considered adiabatic, except for the heat source area when this is active. Finally, in the original reference studies da Silva et al. (2004) and Young and Vafai (1998), the inlet temperature remained constant during the numerical calculations; in this study, however, the inlet temperature was also ramped up and down as way to test a dynamic behavior, which will be discussed in the following section.

All simulations were performed in COMSOL Multiphysics® (Comsol Multiphysics, version 3.5a, User’s Guide) using the direct linear solver, PARDISO. In all the time-dependent simulations, a time step of 0.01 was used. The relative error was set as $10^{-8}$, and the absolute error was set as $10^{-6}$. As expected, the present numerical results were validated against the results of references da Silva et al. (2004) and Young and Vafai (1998). The GSMSMS code was linked to the COMSOL Multiphysics® numerical solutions in Matlab® (MATLAB, 2008). Several different procedures were imple-

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Relevant dimensionless variables.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fig. 3a (Section 4)</strong> (da Silva et al., 2004)</td>
<td><strong>Fig. 3b (Section 5)</strong> (Young &amp; Vafai, 1998)</td>
</tr>
<tr>
<td>Length scale</td>
<td>$L$</td>
</tr>
<tr>
<td>Time scale ($\tau$)</td>
<td>$U_a/L$</td>
</tr>
<tr>
<td>Dimensionless temperature ($\gamma^*$)</td>
<td>$(T - T_0)/</td>
</tr>
<tr>
<td>Reynolds number (Re)</td>
<td>$\rho U_a L / \mu$</td>
</tr>
</tbody>
</table>
mented to ensure the correctness of the present calculations for both configurations. In general, the procedure involved a mesh density study followed by a direct comparison of the present results and the original results published in reference da Silva et al. (2004) and Young and Vafai (1998). For the configuration shown in Fig. 3a, the global conductance (defined in reference da Silva et al., 2004) as \( Q / (k(T_{\text{in}} - T_{\text{w}})) \) was used as the comparison parameter, and for the configuration of Fig. 3b, the Nusselt number was used as the reference.

For the first configuration (Fig. 3a), three variations of the numerical code were implemented and compared with the steady-state results of reference da Silva et al. (2004) where \( D = 0.3 \) L. First, a steady-state code, then a transient code that was allowed to reach steady-state and finally, a transient process where the inlet temperature was oscillated at a very high frequency – note that this last variation aimed to verify the numerical accuracy of the oscillatory inlet boundary condition because it will be used later in the fault description sections (Section 4). For all the aforementioned cases, a mesh of 11,006 elements produced results that agreed within 8% to the reference values in Ref. (da Silva et al., 2004) when \( Re_1 = 1000 \) and \( Pr = 0.7 \) at \( S_{3a} = 0.1, 0.3, 0.5, 0.7, \) and 0.9.

The validation of the code for the configuration with an internal heat source (Fig. 3b) was performed with a steady-state code and a transient with an oscillatory inlet temperature. In this case, the Reynolds number based on the channel wall spacing and the Prandtl number were set to 200 and 0.72, respectively. The length and width (a and b from Fig. 3b) of the obstacle were both \( 0.25D \) in Fig. 3b, \( (Young & Vafai, 1998) \) The relative thermal conductance between the solid protuberance and the fluid was set to 10, and a dimensionless heat flux equal to the unit was imposed at the bottom surface of the solid block, similarly to reference Young and Vafai (1998). For both scenarios above, the results agreed within 10% of the values from (Young & Vafai, 1998), except at locations extremely close to the two top corners (at distances of 0.002 from each corner) of the obstacle, which agreed within 36% using a mesh of 227,475 elements.

4. Results for a straight channel

This section will begin by describing the anomalies that were simulated for the channel flow system shown in Fig. 3a. Several simulations were performed, where the anomaly was positioned at different locations along the channel wall, and different sensor configurations (the number of inputs, positions of inputs, and position of the output) were used to evaluate the capabilities and limitations of the anomaly detection and localization method. Only one type of anomaly was simulated in this study: a heat flux through a small portion of the channel wall that linearly increased with time, \( 0 \leq Q \leq 1 \) from \( 50 \leq \tau_{3a} \leq 300 \). This heat flux was simulated at three positions along the bottom wall. The anomalies corresponding to the three positions will be referred to as Anomalies 1, 2, and 3.

Anomalies 1, 2, and 3 each had a width \( S_{3a} = 0.3 \) that was 10% of the channel's length \( L \) and were located at \( S_{3a} = 0.3, 0.9, \) and 0.9, respectively (see Fig. 3a). When referring to anomaly simulations, for example, Anomaly 1 will refer to a simulation where normal behavior was first simulated (i.e., no heat dissipated by the heater), followed by abnormal behavior simulated (starting at \( \tau_{3a} = 50 \)) at position \( S_{3a} = 0.3 \). The same convention will be used for Anomaly 2 and Anomaly 3 at their positions, \( S_{3a} = 0.3 \) and 0.9, respectively.

For this study, the inputs/outputs orders and the time delay (see Section 2) were chosen based on the combination that resulted in the lowest root mean square (RMS) error using testing data, i.e., an iterative process was used that systematically varied the orders and the time delay, and the RMS errors were computed and compared. Normal system behavior was defined by a constant Reynolds number (\( Re_1 = 100 \)) with the inlet temperature changing at random times for random intervals for a period anywhere from \( \tau_{3a} = 4 \) to \( 7 \) and ranging between \( 0.9 \) to \( 1 \). The reason for these differences is that in the cases of Anomalies 1 and 2, the anomaly's effects have not dissipated much unnoticed. Therefore, it was expected that the anomaly, modeled active with an initial dimensionless heat flux of zero, linearly increasing to 1 at \( \tau_{3a} = 300 \).

As seen in Fig. 4, the CV plot for Anomaly 3 dropped the most after the initiation of the heat dissipation, clearly indicating that abnormal behavior was occurring; this result is expected since the output is relatively close to the anomaly. However, a detection lag is observed for the CVs with the same input/output pair but with Anomalies 1 and 2 simulated because the anomaly was positioned further upstream from the output – note that the same inlet temperature conditions and the same linearly increasing heat flux were used for all three anomalies, i.e., the only difference is the position of the anomaly. For instance, for Anomaly 3, the overall CV drop was approximately 0.7 at \( \tau_{3a} = 300 \), whereas the overall CV drop for Anomalies 1 and 2 were 0.95 and 0.9, respectively. The reason for these differences is that in the cases of Anomalies 1 and 2, the anomaly's effects have dissipated by the time they have reached the temperature sensor at the end of the channel (at the output sensor). Conversely, in the case of Anomaly 3, the anomaly's effects have not dissipated much
because the anomaly is located very near the output, and thus, the temperature is much higher (i.e., the observed temperature and the GSMMS-predicted temperature are very different). Finally, the above results suggest that since anomalous effects dissipate, the anomaly detection sensitivity will be different depending on where the anomaly is located relative to the position of the sensors. Thus, a more refined approach to sensor configuration placement is needed that can better detect a potential fault located anywhere in the system.

Next, rather than using only a single input, as done previously, multiple sensors were distributed throughout the system and multiple inputs were used to observe if the detection sensitivity increased. Table 2 displays the $x_C^*$ positions for the inputs that were tested. The input/output configurations tested are also graphically shown in Fig. 5d. For each of the four input configurations, three different anomalies were tested. The positions of the anomalies are the same as the ones used in Fig. 4.

Table 2  
Dimensionless location $x_C^*$ of the inputs tested (note that $y_C^* = 0.05$ for all inputs).  

<table>
<thead>
<tr>
<th>Number of inputs</th>
<th>Input locations ($x_C^*$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0 0.3 0.7</td>
</tr>
<tr>
<td>5</td>
<td>0 0.2 0.4 0.6 0.8</td>
</tr>
<tr>
<td>10</td>
<td>0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9</td>
</tr>
</tbody>
</table>

As seen in Fig. 5a and b, the detection sensitivity for Anomalies 1 and 2 were not affected as the number of inputs increased, i.e., adding more inputs did not increase the overall CV drop. In fact, for both Anomalies 1 and 2, when the anomaly was placed at $S_{x_3}=0$ and 0.3, respectively, the CV dropped slightly more for configurations with fewer inputs (i.e., there was more sensitivity with fewer inputs). The reason for less detection sensitivity with more inputs is likely because the GSMMS model included more inputs that were affected by the anomaly and thus, ended up in the “unusual” (not well-trained) SON regions where the local models are not as reliable (i.e., regions where the variances of the modeling errors and parameter estimates are too high). Alternatively, with fewer inputs affected by the anomaly, the GSMMS remains in the “usual” (well-trained) SON regions, where local models are more reliable, and therefore more sensitive to anomalies. In contrast, the results for Anomaly 3 indicate that generally, the greater the number of inputs, the more the CV will drop. However, a limit does seem to exist, i.e., as more inputs were added, there was less and less of a change in the overall CV drop. The reason why we see this pattern with Anomaly 3 and not with Anomalies 1 and 2 is that, unlike the cases of Anomalies 1 and 2, the bulk of the inputs are unaffected by Anomaly 3 and thus, the GSMMS is able to gain additional information from the additional inputs without jumping into SON regions where local GSMMS models are unreliable.

From the aforementioned results, it is clear that the relative distance between the output and the anomaly is important. Only if...
the output is close to the anomaly will more inputs benefit fault detection sensitivity. However, if an anomaly's effects have dissipated, the anomaly can go undetected, and the use of more inputs will not significantly enhance the sensitivity. Clearly, since the objective of this fault detection method is to identify unforeseen faults, a sensor configuration that does rely on the output sensor being near a potential anomaly is needed.

Next, rather than changing the number of inputs and their position, only the position of the output will be changed. Fig. 6 shows the pattern of CVs from distributed GSMMS anomaly detectors for Anomalies 1, 2, and 3. Anomaly detectors were formed based on GSMMS models using two inputs from the beginning of the channel, both located at $x_i^* = 0$, but with different heights ($y_i^* = 0.05$ and $0.25$) and an output placed downstream of the input at several different locations. The individual anomaly detectors (corresponding to individual CV plots) differ according to the position of the output, where the output position was moved downstream. There were ten output positions spaced evenly from the inlet to the outlet, where the larger position number indicates a position further downstream. The output positions used for each of the three subplots along with their respective CV plots are shown in Fig. 6.

**Fig. 6a** shows the results for Anomaly 1, which is placed at the beginning of the channel ($x_i^* = 0$). Clearly, anomalous behavior is indicated by the large overall CV drop at positions 2 and 3. As the output is moved further downstream, the overall CV dropped less and less. This result agrees with the previous results in the sense that as the distance between the output and the anomaly becomes larger, the anomalous effects dissipate and the corresponding CVs drop less. Nevertheless, unlike what we had in the previous section where only a single anomaly detector was used (a single CV profile), the pattern of CVs created by the series of anomaly detectors in Fig. 6a clearly indicates that the anomaly's point of origin is between the inlet and position 2 or 3, which is where the anomaly is indeed located.

Similarly, Anomalies 2 and 3 can be detected and localized by interpreting the CV patterns shown in **Fig. 6b** and **c**, respectively. For Anomaly 2, as shown in **Fig. 6b**, normal behavior is indicated at position 2. At position 4, the CV drops slightly, indicating anomalous behavior, which is expected because position 4 and the anomaly's location overlap slightly. At position 6, the CV experienced the largest drop, and as the output is positioned further downstream, the overall CV drop decreases. Thus, an anomaly is detected, and
from the pattern of CV plots, it can be localized between positions 4 and 6. Similar analysis of the CV patterns in Fig. 6c results in localizing Anomaly 3.

In addition, the same approach can be used to detect and localize multiple anomalies occurring simultaneously in the system, which has been a limitation of past fault detection methods. Fig. 7 shows the CV plots when both Anomalies 1 and 2 occur simultaneously. As expected, the CV drops at positions 2 and 3, which represents a similar behavior as in Fig. 6a. However, at position 5, the CV drops even more than the previous position’s CV indicating that, despite the dissipative effects, an additional anomaly is occurring.

5. Results for a straight channel with a flow disruptor

This section describes the distributed anomaly detection approach applied to a more complex system. The system is characterized by a channel flow with a sharp flow disruptor, which greatly complicates the system dynamics (Young & Vafai, 1998). Similarly to the faults considered in the previous section, faults or defects are mimicked by introducing one anomaly at three possible locations. Anomalies 1 and 2 are located in the bottom wall, and Anomaly 3 is located in the top wall. In each of these locations, a patch length of $S_{3b} = 0.1$ is subjected to a linearly increasing heat flux occurring in the time period from zero to one ($0 \leq Q \leq 1$) over 200 dimensionless time steps. For each anomaly, the heat flux begins at $\tau_{3b} = 100$. The three anomalies were positioned as shown in Fig. 8, i.e., Anomalies 1, 2, and 3 were at $S_{3b} = 1.7, 2.35,$ and $2.1$, respectively.

In all simulations for this configuration, normal system behavior was characterized by $Re_{y_0} = 200$ and $Pr = 0.72$. The inlet temperature ranged from $0 \geq T_{in} \geq 1$ and changed at random times for random intervals, between $\tau_{3b} = 4 - 6$. Temperature measurements were sampled every $\tau = 0.1$, and 0.5% noise was added to each measurement. The sensors were located a distance of $y_0 = 0.05$ away from the walls/obstacle, spaced at a distance of $x_0 = 0.1$ apart along the channel walls and spaced at a distance of $x_0 = 0.025$ along the obstacle. The sensor configuration is shown in Fig. 8. For each sub-system, the sensor directly upstream of the output was used as the only input. Also, as was done in Section 4, the orders of each model were chosen based on the lowest RMS.

Fig. 9 shows the result for Anomaly 1, when the anomaly is upstream of the flow disruptor – for clarity, only a portion of the wall is shown. The position number shown in the title of each CV plot refers to the output position. The corresponding input is always the previous position number. From Fig. 9, abnormal behavior is clearly indicated at position 18, where the overall CV dropped to approximately 0.73. Also, by the end of the simulation, the CV at positions 17, 19, and 20 dropped approximately 0.05, 0.05, and 0.02, respectively. However, localization is less clear. Anomalous behavior can be isolated between positions 16 and 21 since both the CVs at those locations indicate normal behavior. Anomaly 1 was located between positions 18 and 19, but the first CV to drop was located at position 17, which is slightly upstream of the anomaly. Surprisingly, the largest drop occurred at position 18, which is above the anomaly. This CV pattern differs from the results in Section 4, where the first CV to drop was either the output above the upstream corner (that would be the CV at position in Fig. 9) or one of the two outputs downstream of the anomaly (that would be positions 19 or 20 in Fig. 9). The reason for this behavior can be explained by realizing that no recirculation exists at position 17 ($x_0 = 1.6$ and $y_0 = 0.05$), suggesting low neighboring velocities. Therefore, it is reasonable to assume that the heat flux is propagating outwards in a relatively symmetric pattern in all directions such that the anomalous effects can affect the upstream sensor at position 17, and thus reduce the CV. Consequently, it is difficult to precisely localize the anomaly upstream of the obstacle without knowing how the particular anomaly propagates outwards into the system.

Fig. 10 shows the CV plots for Anomaly 2, which is downstream of the obstacle. The CVs along the obstacle wall drop slightly (positions 47–55). The first CV to substantially drop is at position 56, whereas the largest overall CV drop was at position 58. One
can note that the CVs at positions 57, 58, and 59 all approximately dropped the same amount, which is different from the previous results. Before, only a single, large CV drop was observed. The relatively large CV drop at the three adjacent positions suggests the existence of an anomaly whose size is potentially larger than what has been seen. This behavior can be explained by the recirculation of fluid around the flow disruptor.
pattern that is formed immediately downstream of the flow disruptor. In this corner, the fluid is rotating in the clockwise direction, which “spreads” the heat dissipated by the heater through the recirculation flow, substantially affecting four sensors. In other words, one can expect relatively high temperature gradients around the vicinity of the heater. Thus, upstream of the flow disruptor, the normal behavior dynamics are similar to that of the anomaly (presence of temperature gradients), and downstream of the obstacle, the normal behavior dynamics are different from that of the anomaly (very small temperature gradients). Therefore, inputs into the GSMMS models downstream of the obstacle in the vicinity of the anomaly are affected by the anomaly and thus, end up in the “unusual” SON regions, which deteriorate their performances.

Finally, it is important to discuss the CV results for Anomaly 3, where the anomaly is above the obstacle. Since the dynamics above the obstacle are similar to simple channel flow, the CV plot pattern was similar to the results from channel flow without the obstacle, as shown in Fig. 11. The results for Anomaly 3 show that the CV plot at position 22 dropped first, and the CV plot at position 23 exhibited the largest drop. Both positions 24 and 25 only dropped slightly. Thus, anomalous behavior was detected, and its origin can be determined to be between positions 21 and 23.

6. Conclusions

In this paper, a precedent-free fault detection and localization method based on the GSMMS method was applied to two thermal flow systems. The newly proposed approach does not require a model of the underlying faulty behavior to achieve detection and localization of faults at different positions within a system. To detect an abnormality, the method compares patterns of the residual differences between the outputs of the physical system (in this case, the two thermal–fluid systems modeled) and those predicted by the GSMMS models. The localization of the sources of anomalies is accomplished using the recently introduced distributed anomaly detection paradigm, where locations of anomalies are inferred through interpretation of the results from multiple distributed anomaly detectors. Our numerical studies show that the use of a single anomaly detector was inadequate even for the simplest of system. The reason is that in a thermal–fluid system, the effects of the anomaly dissipate due to the fluid flow, diminishing deviations from normal behavior at positions downstream from the anomaly and thus lessening the detection sensitivity. Conceivably, there would be a physical distance at which the effects of the anomaly will have completely dissipated and thus, despite the presence of the anomaly upstream of the output sensor, there would be no deviation in normal behavior at an output sensor position far away from the anomaly. Distribution of anomaly detectors prevents “blind spots” in the system by creating many input(s)/output combinations, ensuring that if an anomaly exists, an output sensor of one of the GSMMS models will be near the anomaly. With distributed anomaly detectors, both detection and localization of the anomaly was achieved. Localization was accomplished by analyzing the pattern of CVs, where the largest overall CV drop indicates a position very near the point of origin of the anomaly. However, the conclusions drawn from the analysis of the CV pattern should be done with care as different normal behavior dynamics will result in a different CV pattern for a given anomaly.

Despite the successful transition from a simple to a more complex system, the distributed anomaly detectors approach is limited by the number and location of sensors that can be physically and economically placed within the system. Thus, future work should focus on optimizing the number of sensors and the area monitored by the anomaly detectors. This sensor optimization could be based on expert knowledge, i.e., more anomaly detectors should be placed at the most probable fault locations.

References