Gaussian Process-based Methods for Process States Validation

Ibrahim Ahmed, Gyunyoung Heo*
Department of Nuclear Engineering, Kyung Hee University, 1732 Deogyeong-daero, Giheung-gu, Yongin-si, Gyeonggi-do 17104, Korea
**Corresponding author: gheo@khu.ac.kr*

1. Introduction

In industrial systems' operations such as avionics, fossil-fuel power plants, nuclear power plants (NPPs), and oil and gas installations, faults and failures can occur in sensors, equipment, and processes which can have impact on safety and on the performance of the system. Hence, monitoring and signal validation of those sensors installed for the safety parameters during operation is extremely important.

Even though a verse variety of model-based approaches to achieve condition monitoring purposes are available in literature [1], their application in large complex plants is sometimes infeasible and unrealistic because they require rigorous process models which are not usually available for complex systems. On the other hand, data-driven methods [2], an alternative to model-based approaches, use historical data measured and recorded during process operations to build an empirical model. Such methods include auto associative kernel regression (AAKR). Typically, data-driven models are developed under steady state plant conditions belonging to most of uptime, but it is significant to have condition monitoring during transient operations, considering the fact that most industrial systems' operations are time-varying. Transient operations are any non-steady state, time-varying signal conditions, such as the start-up, shutdown, and load following modes of the system.

The AAKR formulation of kernel regression (KR) model – a special and simple form of Gaussian process regression (GPR), has been successfully used in actual NPP steady-state operations for instrument channel calibration and condition monitoring [2], easy to use, and less computationally demanding. However, KR has limitations and seems inappropriate in transient operations because, it is based on the unilateral kernel and lack temporal information in that only the current query vector has effect on the model. Any previous information leading to the current query vector is completely ignored. Although this is acceptable and even preferable for many applications, it is not acceptable for transient operations, in which the previous information directly affects the next data point. In view of the above, the authors developed several methods which are simple, effective, easy to use, and capture both the spatial and temporal dependencies in the time-series data, for efficient implementation of process states validation and condition monitoring not only in steady-state operations but also in transient operations of industrial systems. These methods are summarized in this paper.

3. Gaussian Process-based Methods: An Overview

3.1 Gaussian process regression model

Gaussian process is a tool that has been successful in building surrogate models in which the GPR extends multivariate Gaussian distributions (MGD) to infinite dimensionality. It represents function by letting the data determine the structure of the model, less parametric, and can be formulated in terms of a kernel smoother, using an equivalent kernel or weight function. Hence, it can be viewed as equivalent to KR approaches prevalent in statistics, which have been adapted for use in on-line monitoring. Given some noisy observations of a dependent variable $\mathbf{y} \in \mathbb{R}^{M\times 1}$ at certain values of independent variable $\mathbf{x} \in \mathbb{R}^{M\times 1}$, the goal is to estimate the dependent variable $\hat{y}_q$ at a new value of x_q. Each observation y can be thought of as related to an underlying function $f(x)$ through a Gaussian noise model:

$$\mathrm{y} = f(x) + \mathcal{N}(0, \sigma_n^2) \tag{1}$$

where $f(x)$ is true function of y which is to be search for through regression, and $\mathcal{N}(0, \sigma_n^2)$ is a measurement noise distribution with mean of 0 and variance of σ_n^2. Since the key assumption in GP modeling is that, the data can be represented as a sample from a MGD, we have

$$\begin{bmatrix} y \\ y_q \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K & K_q^T \\ K_q & K_{qq} \end{bmatrix}\right) \tag{2}$$

where T indicates matrix transposition, and K, K_q and K_{qq} are $M \times M$ covariance matrix, $1 \times M$ covariance vector and 1×1 covariance vector, respectively, which are all computed from covariance kernel function:

$$k(x, x') = \sigma_f^2 exp\left[\frac{-d(x, x')^2}{2l^2}\right] \tag{3}$$

where $d(x, x')$ is the distance between observations, and l and σ_f^2 are bandwidth and variance of the true function, respectively, which are to be determined during training. Our interest is in the conditional probability, $p(y_q|\mathbf{y})$ that given the data, how likely is a certain estimation for y_q. From Eq. (2), it can be showed that this probability follows the following Gaussian distribution:

$$y_q|\mathbf{y} \sim \mathcal{N}\left(K_q K^{-1}\mathbf{y}, \quad K_{qq} - K_q K^{-1} K_q^T\right) \tag{4}$$

Thus, it is very clear from Eq. (4) that the best estimate of y_q is the mean of the distribution. Hence, the GPR estimation is given as

$$\hat{y}_q = K_q K^{-1}\mathbf{y} \tag{5}$$

If $\mathbf{w} = K_q K^{-1}$ is defined as vector of weights, i.e., $\mathbf{w} \in \mathbb{R}^{1\times M}$, Eq. (5) can be written as

$$\hat{y}_q = \mathbf{wy} \tag{6}$$

This implies that the GPR is a weighted sum of the historical observations [3], which is equivalent to KR as shown in the next section.

3.2 Kernel regression model

KR is the process of estimating a parameter's value by calculating a weighted average of the historical observations. It is generally and more compactly represented by the Nadaraya-Watson estimator as [3]

$$\hat{y}_q = \frac{\sum_{i=1}^{M} k_{qi}(x_i, x_q) y_i}{\sum_{i=1}^{M} k_{qi}(x_i, x_q)} \tag{7}$$

where K_q $1 \times M$ Kernel vector whose elements are calculated from kernel function with bandwidth *h*:

$$k(x_i, x_q) = exp\left[\frac{-d(x_i, x_q)^2}{2h^2}\right] \tag{8}$$

Equation (7) can be re-written as

$$\hat{y}_q = \sum_{i=1}^{M} \frac{k_{qi}(x_i, x_q)}{\sum_{i=1}^{M} k_{qi}(x_i, x_q)} y_i = \sum_{i=1}^{M} w_i y_i = \mathbf{wy} \tag{9}$$

$\mathbf{w} \in \mathbb{R}^{1\times M}$ is a vector of weights whose elements $w_i = k_{qi}/\sum_{i=1}^{M} k_{qi}$. It can be seen that, the main different between KR and GPR is the $M \times M$ covariance matrix *K* whose its diagonal elements is σ_f^2 and off-diagonal elements are nearly zeros. However, if the normalizing factor, $\sum_{i=1}^{M} k_{qi}$ in Eq. (9) can be thought of as an approximate value of σ_f^2, Eqs. (9) and (6) can be assumed to be equivalent and KR can be said to be a simpler and less computational form of GPR.

Several formation of KR can be used, such as AAKR, inferential KR, and hetero-associative KR. The AAKR formulation is the most suitable for process monitoring. Despite the simplicity of KR and its successful applications in steady-state equipment condition assessments, it is obvious from its mathematical modelling that it lacks temporal information, which is crucial in transient operations. The major shortcoming of KR in those conditions is that each predicted signal variable from the query input vector are the same in most cases: the estimated takes the average of those signal variables when the training set contains similar training patterns for all of those variables. This problem has been clearly depicted in Fig. 1. In Fig. 1, to estimate the variable y given x_1 and x_2, the predictor vector at the point A which occurs at time t_4 is identical to the predictor vector at point B which occurred at t_8. In this case, the corresponding values of y's, $y(t_1)$ and $y(t_2)$ at the respective data points would not be estimated correctly by KR. This is because the distances computed at both points are the same which, in turn, will result in KR inability to differentiate between the two data vectors.

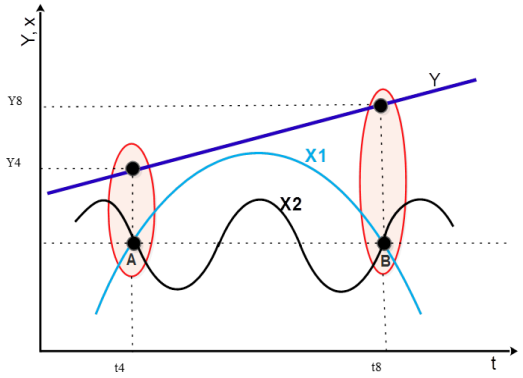


Fig. 1. Problem definition (estimation of y given x1 and x2)

Therefore, a new, related, and yet simple model is needed to capture both the spatial and temporal dependencies in the data for efficient implementation of process monitoring in transient operations. This led to the development of several methods that will be briefly discussed in the next section.

4. The Developed Methods

In this section, the methods developed to alleviate those issues in KR are summarized.

4.1 Kernel regression using derivatives

Here, the conventional KR is modified through the development of time-dependent transformation equation which is derived from the Taylor series expansion [4].

$$x(t+h) = x(t) + h\frac{dx(t)}{dt} + \frac{h^2}{2!}\frac{d^2x(t)}{dt^2} + \cdots \tag{10}$$

From the Taylor series of Eq. (10), the function at $x(t_2)$ can be approximated as

$$x(t_2) = x(t_1) + (t_2 - t_1)\frac{dx}{dt} \tag{11}$$

with $\eta = (t_2 - t_1)$, which is the time step interval. Equation (11) can then be rearranged as

$$x(t_i) - t_i\frac{dx}{dt} = x(t_{i-1}) - t_{i-1}\frac{dx}{dt}. \tag{12}$$

We then substituted

$$\psi_i = t_i\frac{dx}{dt_i} \tag{13}$$

with assumption that

$$\begin{cases} \dfrac{dx}{dt_1} \neq \dfrac{dx}{dt_2}\ but \\ \left.\dfrac{dx}{dt}\right|_{t=t_1} \approx \dfrac{dx}{dt_1}, and \\ \left.\dfrac{dx}{dt}\right|_{t=t_2} \approx \dfrac{dx}{dt_2} \end{cases} \tag{14}$$

Thus, our proposed transformation dynamic equation is given by

$$\psi_i \cong \psi_{i-1} + (x_i - x_{i-1}) \tag{15}$$

where ψ_i is the resulting transformed value of the current data point x_i and dx/dt_{i-1} is the first derivative at time t_{i-1} of immediate past data point x_{i-1}. It can be observed that, the transformed output ψ_i has the same

unit as that of the original value x_i. In this case, the information from previous input vectors is incorporated into the KR through the gradient, the time, and the difference between the previous and the current input vector. This in fact will gives a more details representation of the estimations compare to that of the conventional KR that ignored any information leading to the current data point. The derivative at each data point can be approximated using backward finite difference derivative approximation.

Both historical/memory data and the query input data should be transformed to a new space defined by Eq. (15). The transformed data can then be used to compute the similarity and kernel weights. By re-writing Eq. (9), the estimation of the query input can be performed as follows:

$$\hat{y}_q = \sum_{i=1}^{M} \frac{k_{qi}(\psi_i, \psi_q)}{\sum_{i=1}^{M} k_{qi}(\psi_i, \psi_q)} y_i \tag{16}$$

The transformed signal will be shifted to a new space difference from its original signal. With this transformation, the data points at A and B in Fig. (1) could be distinguished and the correct estimates from Eq. (16) can be made.

4.2 Kernel regression aggregating bilateral directions

Having developed the modified KR presented in Section 4.1 and found to have resolved the issues, we discovered that, during on-line implementation, the query time input (Eq. (13)) monotonically increases and becomes indefinite. Although this is known for a particular historical data set within the specific period of time at which the data is collected, $\{0 \le t_q \le T\}$, it is virtually impossible to collect historical data that covers the operational lifespan of large industrial components. In this regards, we propose a novel model that we call weighted-distance Auto Associative Bilateral Kernel Regression (AABKR) [5, 6].

Given a sequence of historical *p*-dimensional time-series data $\mathbf{X} \in \mathbb{R}^{M\times p}$, with *M* observation sequence vectors, and $x_{i,j}$ represents the *i*th observation of the *j*th variable. For every on-line query observation, $\mathbf{x}_q^* \in \mathbb{R}^{1\times p}$, the prediction of the signal is performed according to the formulation

$$\hat{x}_{q,j}^* = \frac{\sum_{i=1}^{M}(k_i^f \circledast k_i^t)x_{i,j}}{\sum_{i=1}^{M}(k_i^f \circledast k_i^t)} \tag{17}$$

where $k_i^f \circledast k_i^t = k_i^{ab}$ is an adaptive bilateral kernel evaluated at $\mathbf{x}_i$. The symbol, $\circledast$ represents the bilateral kernel combination operator that combined the feature and temporal kernels together, formulated as

$$k_i^{ab} = k_i^f \circledast k_i^t = \begin{cases} k_i^f * k_i^t, & s \le i \le M \;\&\; i \ne \varepsilon \\ \dfrac{(k_i^f + k_i^t)}{2}, & i = \varepsilon \end{cases} \quad ; \quad i\epsilon[s, M] \tag{18}$$

where *s* is a moving window length for the require data points use to compute backward-difference derivative approximation, ε is a time position index (required for the temporal distance calculation which is determined on-line using derivative-based comparator [5,6]) of the nearest vector in memory data vectors to the query on-line vector observation, k_i^f and k_i^t are feature kernel and temporal kernel, respectively. The two kernels are calculated as follows:

$$k_i^f = exp\left(\frac{-{d_i}^2}{2h_f^2}\right) \tag{19}$$

$$k_i^t = exp\left(\frac{-{\delta_i}^2}{2h_t^2}\right) \tag{20}$$

where h_f is a kernel bandwidth for feature preservation, which controls how much the nearby memory feature vector is weighted, h_t is the bandwidth for the time-domain preservation, which can serve as noise rejection and controls how much the nearby times in the memory vectors are weighted, $d_i = d_i(\boldsymbol{X}_i, \boldsymbol{x}_q^*) = \left\|\boldsymbol{X}_i - \boldsymbol{x}_q^*\right\|_1$ is a feature distance, and δ_i is temporal distance, which accounted for the variation in time at which the query vector is observed [5]. It can be seen in Eq. (17) that both the spatial and temporal information are captured in the model during prediction, and thus will provide a better signal estimation, particularly during normal process transient operations than that of KR.

The proposed weighted-distance AABKR has been applied to a start-up transient operation of a pressurized water reactor (PWR) NPP. Due to unavailability of real data from the plant, we used a real-time simulator data collected from the compact nuclear simulator (CNS) during heating from the cool-down mode (start-up operation) as a normal start-up transient operation for building the model. Six sensors' process variables from the reactor coolant system (RCS) were selected for monitoring during this operation: S1 (cold leg temperature), S2 (core exit temperature), S3 (hot leg temperature), S4 (safety injection flow) S5 (residual heat removal flow) and S6 (sub-cooling margin temperature). The data consist of 1000 observations sequentially collected at constant time intervals of 1s. Because these data are fault-free, we simulated abnormal conditions on them by adding fault to a particular sensor data at a time. To do this, we conducted a thousand-run Monte Carlo simulation experiment on this dataset, of which a run consist of 1000 observations of faulty dataset. In order to simulate a realistic scenario for the fault detection capability of the proposed model, the fault magnitude was a random number sampled from a bimodal uniform distribution, $U([-10,-2] \cup [2,10])$, and added to a variable. Also, the sensor variable to be in fault at a particular time step during a single run was uniformly random among the six possible sensor variables. At the end of each run, true alarm rate (TAR): the rate of fault detection in only a sensor that actually has the fault without fault being detected in other fault-free sensors, is calculated. Fig. 2 shows the distribution of the TARs of the thousand-run Monte Carlo experiment for AAKR (Fig. 2(a)) and the

weighted-distance AABKR (Fig. 2(b)). The means of these distributions are 61.4% and 99.8% for AAKR and weighted-distance AABKR, respectively. Thus, for single-sensor faults, the proposed model, on average, had a significantly higher TAR than AAKR, and can be used to validate the sensors' status during operation.

4.3 Bilateral kernel regression using dynamic time warping

The proposed method discussed in Section 4.2, having been successfully applied to the start-up transient operation, was tested for its generality by applying it to a typical steady-state operation. It was discovered from the result that its performance suffered significantly from the spillover effects (faults are detected in process sensors different from those actually affected by the faults). This was due to wrong identification of time position index, ε using derivative-based comparator, as the derivative calculated from the typical steady-state process is nearly constant, where the process change in time is almost negligible. To this end, a novel approach based on dynamic time warping (DTW) for the identification of time position index has been developed for efficient implementation of AABKR applicable not only in transient process operations but also in steady-state operations [3]. The basic idea in using DTW instead of derivatives is that, the query input of time window length s will be compared with the sequence of time-series memory data to find the optimal alignment between the two time sequences dynamically. The time location within the memory time-series data where the optimal alignment is located will have the smallest DTW distance compare to other locations. With this, the correct identification of time position index can be obtained. Another importance modification is that, by extending the adaptive approach of combining the two kernels, a robust bilateral kernel evaluation algorithm has been proposed, which dynamically compensates for faulty sensor inputs, resulting in a more robust model with less spillover during monitoring.

5. Conclusions

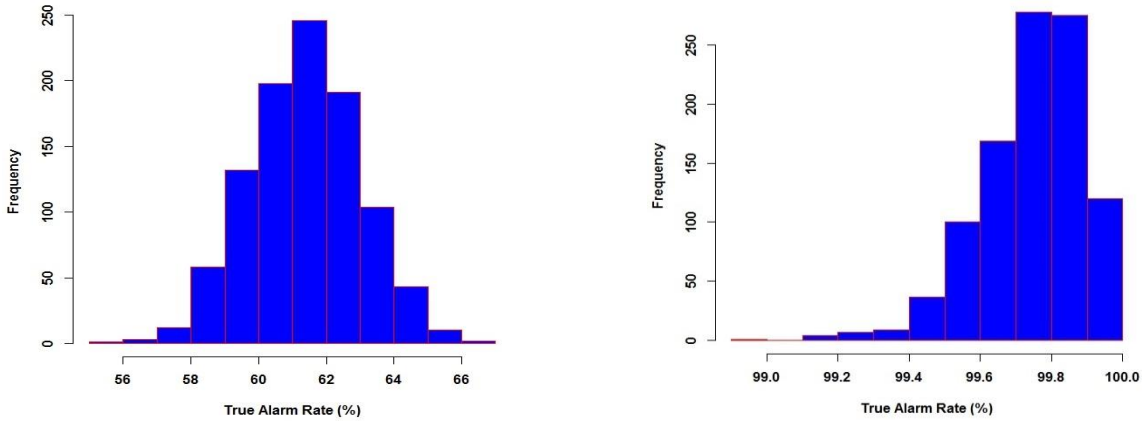
This paper discussed several methods developed for signal validation and condition monitoring during transient operations, in which several attempts have been successfully carried out and been applied to simulation dataset for transients of a PWR NPP. It turned out that, the conventional GPR can be adjusted for various purposes, and in steady-state as well as transient conditions is maintained higher. Furthermore, the accuracy, robustness, and computational time need to be improved and verified.

Acknowledgements

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No. 20163010140550); and by the National Research Foundation of Korea(NRF) grant funded by the Korean Government (MSIP: Ministry of Science, ICT and Future Planning) (No. 2017M2B2B1072806).

REFERENCES

[1] J. Chen and R. J. Patton, Robust model-based fault diagnosis for dynamic systems, Kluwer Academic Publishers, 1999.
[2] G. Heo, Condition Monitoring Using Empirical Models: Technical Review and Prospects for Nuclear Applications, Nuclear Engineering Technology, Vol. 40, p. 49–68, 2008.
[3] I. Ahmed, Gaussian Process and Deep learning-based Methods for Time-series States Validation in Industrial Systems, Ph.D. Thesis, Kyung Hee University, Republic of Korea, 2019.
[4] G. Heo, I. Ahmed, and G. Ha, "Apparatus and Method for Authenticating Time-Varying Signal in Online via Kernel Regression Model", Korea Patent, 10-1967524, Korea, 2019.
[5] I. Ahmed, G. Heo, and E. Zio, On-line process monitoring during transient operations using weighted distance Auto Associative Bilateral Kernel Regression, ISA Transactions, In Press, https://doi.org/10.1016/j.isatra.2019.02.010, 2019.
[6] G. Heo, I. Ahmed, and G. Ha, "Apparatus and Method for Online Signal Data Validation via Auto Associative Bilateral Kernel Regression", Korea Patent Application, 10-2018-0170632, Korea, 2018.



(a) AAKR (b) Weighted-distance AABKR

Fig. 2. Histograms of the TARs (true alarm rates) for a thousand-run Monte Carlo experiment