Automatica 176 (2025) 112233

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Robust fault detection method based on interval neural networks optimized by ellipsoid bundles☆

Meng Zhou [a], Yinyue Zhang [a], Jing Wang [a], Tarek Raïssi [b,*], Vicenç Puig [c]

[a] School of Electrical and Control Engineering, North China University of Technology, 100144, Beijing, China
[b] Conservatoire National des Arts et Metiers (CNAM) Cedric-Laetitia, 292 Rue St-Martin, 75141, Paris Cedex 03, France
[c] Advanced Control Systems Research Group at Institut de Robòtica, CSIC-UPC, Universitat Politècnica de Catalunya-BarcelonaTech, Barcelona, Spain

ARTICLE INFO

Article history:
Received 27 January 2024
Received in revised form 7 November 2024
Accepted 19 January 2025
Available online 3 March 2025

Keywords:
Ellipsoid bundle
Data-driven
Interval estimation
Fault diagnosis
Wastewater treatment process

ABSTRACT

In this paper, a novel framework for interval prediction neural network model optimized by ellipsoid bundles is designed for data-driven systems with unknown but bounded noise and disturbances. Firstly, a point prediction neural network model is constructed with input and output data, the prediction error is assumed to be unknown but bounded. Then, a feasible set related to the output weights of the neural network model is described by ellipsoid bundles, which allows decreasing conservatism compared to other types of sets such as zonotopes or ellipsoids. To obtain a more precise description shape of the feasible set, an optimal iterative formula is theoretically derived by minimizing the Frobenius norm of ellipsoid bundles. Next, an outer bounding box is established to obtain the maximum and minimum weights that satisfy the feasible set, thereby providing the output prediction interval. This interval characterizes the systems' disturbance and can be used as the threshold when dealing with fault diagnosis applications. Finally, the effectiveness of the proposed interval neural network model is demonstrated by using faulty data from a wastewater treatment process. Simulation results verify that the proposed method can achieve accurate prediction outputs and fault diagnosis performance.

© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

In recent decades, with the rapid development of Internet of Things, big-data analysis, artificial intelligence technology, soft sensor modelling methods based on neural networks (NNs) have been widely investigated in various fields, due to their learning, adaptive, and nonlinear function approximation capabilities. Most commonly used neural network models mainly include backpropagation neural networks (BP), radial basis function neural networks (RBF), long short-term memory neural networks (LSTM), convolutional neural networks, recurrent neural network, just to name a few. NN approach has been applied for state monitoring, system identification and fault detection in multiple domains, such as wastewater treatment process (WWTP), industrial process control, environmental monitoring systems, smart grid control systems, among others (Jia et al., 2023; Sun & Ge, 2021; Wang et al., 2021).

It should be mentioned that most soft sensor models based on NNs are primarily utilized for point prediction tasks. However, in practice, due to the measurement noise, disturbances and modelling uncertainty, it is difficult to obtain accurate soft sensor models. Note that the deviation between the actual and the predicted values should be bounded in non-faulty scenarios if the obtained models are going to be used in fault detection. It is significant to analyse the effect of the unknown but bounded disturbances on the estimated parameters. To solve this problem, many scholars have investigated interval soft sensor models (Nourani et al., 2023).

For soft sensor models based on NNs, how to optimize the neural network weights is an important task in nonlinear fitting. Currently, uncertainty prediction models mainly involve probabilistic interval prediction based on error distributions and deterministic interval prediction based on set-membership estimation. For the first ones, the uncertainty is typically assumed to be random but with a known probability distribution, which may not be suitable when dealing with unknown deterministic behaviours (Combastel, 2015).

Instead of requiring probability distributions, set-membership estimation has gained significant attention since only considers

☆ This paper is funded by National Natural Science Foundation of China (62273007, 61973023), the Project of Cultivation for Young Top-Motch Talents of Beijing Municipal Institutions, China (BPHR202203032). The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Angelo Alessandri under the direction of Editor Thomas Parisini.

* Corresponding author.
E-mail addresses: zhoumeng@ncut.edu.cn (M. Zhou), yyzhang@mail.ncut.edu.cn (Y. Zhang), jwang@ncut.edu.cn (J. Wang), tarek.raissi@cnam.fr (T. Raïssi), vicenc.puig@upc.edu (V. Puig).

https://doi.org/10.1016/j.automatica.2025.112233
0005-1098/© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

that the uncertainties are bounded with a priori known bounds. There exist different set representations, e.g. ellipsoids, polytopes and zonotopes, that have been widely used for interval state and fault estimation in control systems. Due to their ability to achieve a good trade-off between estimation accuracy and computational complexity, zonotope-based methods have received considerable attention in the last years. In Li et al. (2021), a distributed set-membership filtering method is designed to deal with attacks signals and fading measurements based on zonotopes. Zhu et al. (2022) present an interval observer-based fault detection and isolation design method for T-S fuzzy systems based on zonotope analysis. Recently, a new set representation tool called ellipsoid bundle that combines the advantages of ellipsoids and zonotopes is proposed (Tang et al., 2022) and applied to linear parameters varying systems. The ellipsoid bundles can effectively describe sets with smooth boundaries and demonstrate the ability of zonotopes to represent sets with complex shapes.

However, very little attention has been paid on the application of set-membership estimation in neural networks. In Chai and Qiao (2014), a minimum trace ellipsoid set-membership estimation algorithm is proposed to estimate the weights of the output layer of RBF and applied to a wastewater treatment system. Since this method presents some conservatism, it has motivated the work presented in this paper. The main contributions of this paper are summarized as follows:

- This paper proposes a novel framework for an interval prediction neural network model optimized by means of ellipsoid bundles.
- To obtain a more precise description of the feasible set shape, a recursive interval parameter estimation approach is proposed using ellipsoid bundles. Besides, an optimal iterative formula of the ellipsoid bundle is derived by minimizing the Frobenius norm. It can generate tighter interval estimation and enhance the accuracy of fault diagnosis compared with some similar methods based on ellipsoids (Zhang et al., 2022) and zonotopes (Alamo et al., 2005).
- Based on the properties of ellipsoid bundles, an outer bounding box is established to obtain the maximum and minimum weights that satisfy the feasible set, allowing to determine the interval output prediction. This interval can be used as a fault detection threshold and applied to wastewater treatment fault diagnosis.

The paper is organized as follows. The preliminaries are introduced in Section 2. The interval prediction model construction is described in Section 3. The proposed interval estimation model is presented in Section 4. Simulation results are discussed in Section 5. Some conclusions are given in Section 6.

2. Preliminaries

According to Tang et al. (2022), some fundamental knowledge about ellipsoid bundles is given.

Definition 1. An n-order ellipsoid bundle $\xi(c, \Psi) \subset \mathbb{R}^m$ is denoted as the Minkowski sum of n ellipsoids and a vector $c \in \mathbb{R}^m$. The expression is given as follows:

$$\xi(c, \Psi) = c \oplus \bigoplus_{i=1}^{n} \mathcal{E}(0, H_i) \tag{1}$$

where $\mathcal{E}(0, H_i)$ is an ellipsoid with the expression of $\mathcal{E}(0, H_i) = \{\theta \in \mathbb{R}^m : \theta = 0 + H_i z, z \in \mathbb{R}^n, z^T z \leq 1\}$, $\Psi = \{H_1, H_2, \ldots, H_n\}$ is an ordered set of matrices of the ellipsoid bundle. $H_i \in \mathbb{R}^{m \times n}$, $i = 1, \ldots, n$ represents the generation matrix of ellipsoids. $\oplus$ denotes the Minkowski sum.

(a) 2-dimensional (b) 3-dimensional

Fig. 1. Ellipsoid bundles.

Fig. 1(a) shows a two-dimensional plane projection of the ellipsoid bundle and Fig. 1(b) shows an ellipsoid bundle in three-dimensional space.

Property 1. *Given two ellipsoid bundles $\xi(c_1, \Psi_1) \in \mathbb{R}^n$ and $\xi(c_2, \Psi_2) \in \mathbb{R}^n$, their Minkowski sum satisfies:*

$$\xi(c_1, \Psi_1) \oplus \xi(c_2, \Psi_2) = \xi(c_1 + c_2, [\Psi_1, \Psi_2]) \tag{2}$$

Property 2. *For the ellipsoid bundle $\xi(c, \Psi) \subset \mathbb{R}^m$ in (1) and a matrix $\varphi \in \mathbb{R}^{r \times m}$, the following equation holds:*

$$\varphi \xi(c, \Psi) = \xi(\varphi c, \Phi(\varphi, \Psi)) \tag{3}$$

where

$$\Phi(\varphi, \Psi) = [\mathscr{H}_1, \ldots, \mathscr{H}_n], \mathscr{H}_i = \varphi H_i \varphi^T \tag{4}$$

The Minkowski sum of ellipsoid bundles can be transformed into simple matrix operations, which is similar to zonotopes.

Property 3. *For the ellipsoid bundle $\xi(c, \Psi) \subset \mathbb{R}^m$ in (1), a new ellipsoid bundle $\xi(c, \tilde{\Psi})$ is generated by reordering $\xi(c, \Psi)$ in a decreasing trace of H_i, such that $\tilde{\Psi} = \{H_{\sigma(1)}, H_{\sigma(2)}, \ldots, H_{\sigma(n)}\}$ and $\sigma(i)$ is a permutation of $i = 1, \ldots, m$. Then given a fixed integer $s (1 < s < n)$, we can obtain an s-order ellipsoid bundle $\xi(c, \mathcal{R}_s(\Psi))$ such that $\xi(c, \Psi) \subset \xi(c, \mathcal{R}_s(\Psi))$ with:*

$$\mathcal{R}_s(\Psi) = \{H_{\sigma(1)}, \ldots, H_{\sigma(s-1)}, \tilde{H}\}, \tilde{H}\tilde{H}^T = \tilde{P} \tag{5}$$

$$\tilde{P} = \sum_{i=s}^{n} \sqrt{tr(H_{\sigma(i)})} \sum_{j=s}^{n} \frac{P_{\sigma(j)}}{\sqrt{tr(H_{\sigma(j)})}} \tag{6}$$

where $P_{\sigma(j)} = H_{\sigma(j)} H_{\sigma(j)}^T$ and $\tilde{H}$ can be obtained by a Cholesky decomposition if $\tilde{P}$ is positive definite, otherwise by a singular value decomposition.

Lemma 1. *(Tang et al., 2020) If $x_k \in \xi(c, \Psi) = c \oplus \bigoplus_{i=1}^{n} \mathcal{E}(0, H_{k,i})$ with $P_{k,i} = H_{k,i} H_{k,i}^T$, we can obtain an interval vector $[\underline{x}_k, \bar{x}_k]$ such that $\underline{x}_k \leq x_k \leq \bar{x}_k$ with:*

$$\begin{cases} \underline{x}_k = c - \widehat{b}_k \\ \bar{x}_k = c + \widehat{b}_k \end{cases} \tag{7}$$

where $\widehat{b}_k(l) = \sum_{i=1}^{n} \sqrt{P_{k,i}(l, l)}, l = 1, 2, \ldots, m$.

3. Problem description

A general soft-sensing model established by neural network models can be described as:

$$y_k = f(x_{1k}, x_{2k}, \ldots, x_{qk}) \tag{8}$$

where x_{ik}, $i = 1, \ldots, q$ represent q auxiliary variables that determine the output. k represents the kth time step, $f(\cdot)$ is an equivalent function of a soft-sensing model that approximates the nonlinear relationship between the auxiliary variables and the

output. $y_k \in \mathbb{R}^{1\times n_y}$ represents the output with n_y variables.

Generally, neural network models can be designed to construct the prediction model $f(\cdot)$, which typically consists of input, hidden, and output layers. Specially, the connection from the hidden layer to the output layer can be represented as:

$$y_k = \varphi_k\theta_k + e_k \tag{9}$$

where $\theta_k \in \mathbb{R}^{m\times n_y}$ represents the weight matrix of the hidden layer, m is the number of neuron nodes in the hidden layer, $\varphi_k \in \mathbb{R}^{1\times m}$ is the feature matrix of the hidden layer output. $e_k \in \mathbb{R}^{1\times n_y}$ denotes the modelling error of the neural network, which is affected by the selection of the output weights of the neural network and the noise in the sample data.

Note that the output y_k is in regressor form in (9), i.e., linearly dependent with respect to the parameters θ_k. Since the existence of the modelling error, it is difficult to obtain the exact point prediction value y_k. According to Schweppe (1968), it is common to assume that the error vector is unknown but bounded. Hence, in this paper, the modelling error e_k is expressed as bounded in a determined set, which is related to the approximation ability of the neural network. The ith term of e_k has

$$|e_k^i| \le \gamma, \quad i = 1, \ldots, n_y \tag{10}$$

where γ is the upper bound of the modelling error of the neural network after the training process.

Combining (9) and (10), we can obtain:

$$(y_k - \varphi_k\theta_k)(y_k - \varphi_k\theta_k)^T \le \gamma^2 \tag{11}$$

As γ is known, for each pair of y_k and φ_k, the corresponding θ_k can be described as a set. In this paper, denote S_k the set of the hidden layer weights θ_k satisfying Eq. (11) as:

$$S_k = \{\theta_k : (y_k - \varphi_k\theta_k)(y_k - \varphi_k\theta_k)^T \le \gamma^2\} \tag{12}$$

From a geometrical point of view, S_k is a convex set. For each pair of parameters in sequence (y_k, φ_k), all elements in S_k are reasonable estimates of the actual parameter θ_k. That means the feasible set of the parameter θ_k can be obtained by measured output and known error bound. If (y_k, φ_k) is available at time instance k, then the parameter θ_k is described as (12). Finally, the feasible set of weights must contain the actual results of the weight at any time step k, hence θ_k should belong to the intersection of the corresponding sequence of S_k, namely:

$$\theta_k \in \Theta_k = \bigcap_{k=1}^{n} S_k \tag{13}$$

where Θ_k represents the feasible set of weight θ_k and n represents the number of sets S_k.

The input data of the neural network (9) should satisfy the persistently excitation condition well-known in the literature, see Ljung (1998), converting different operating points to obtain an accurate model. This condition also plays a crucial role in minimizing the size of the prediction intervals. Due to Θ_k is a monotone non-increasing sequence, if the training data is unlimited, the set Θ_k converges to a certain set. However, although the set Θ_k can be easily expressed, the analytical expression is not easy to find, hence the goal of this article is to seek an approximate analytical expression for Θ_k and obtain a feasible set of weights for neural network. The schematic of the proposed approach is shown in Fig. 2.

Remark 1: The weights of NNs can be first trained by gradient descent algorithm, then they can be further optimized by the proposed ellipsoid bundle estimation algorithm. The difference between the training weights of the traditional gradient descent algorithm and the set-membership estimation algorithm is that the latter can obtain the weights of the neural network and its upper and lower bounds, simultaneously, considering that the uncertainty of the soft-sensor prediction model can be described by ellipsoid bundles.

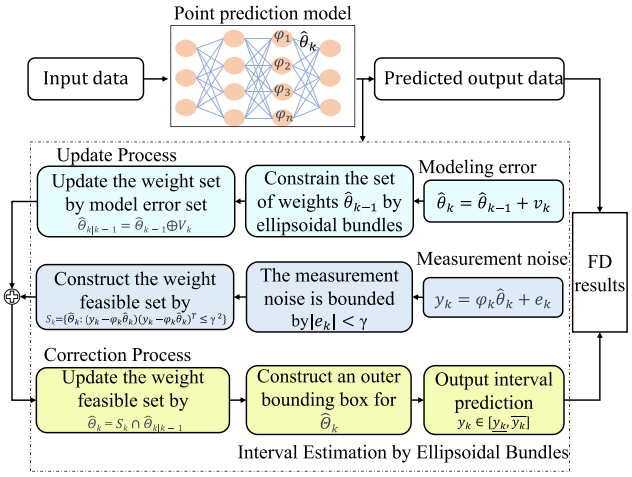

Fig. 2. Schematic of the proposed approach.

4. Interval neural network model construction

As discussed in Section 1, the ellipsoid bundle integrates the advantages of both ellipsoids and zonotopes, making it more suitable for representing sets with complex shapes. Moreover, it does not have limitations when describing weight sets with smooth boundaries (Tang et al., 2022). To further reduce the conservatism of the ellipsoid, this subsection aims to describe the feasible set of weights Θ_k by means of the ellipsoid bundle method.

Motivated by Kalman filters and observers, the estimation process can be divided into two parts: update process and correction process (Jose & Wen, 2009).

In this paper, the disturbance and modelling error are considered as bounded, hence $\widehat{\theta}_{k|k-1}$, $\widehat{\theta}_{k-1}$ and v_k are assumed to be restricted in terms of ellipsoid bundles as:

$$\widehat{\theta}_{k|k-1} \in \widehat{\Omega}_{k|k-1} = \xi(\widehat{\theta}_{k|k-1}, \widehat{\Psi}_{k|k-1}) \tag{14}$$

$$\widehat{\theta}_{k-1} \in \widehat{\Omega}_{k-1} = \xi(\widehat{\theta}_{k-1}, \widehat{\Psi}_{k-1}) \tag{15}$$

$$v_k \in V_k = \xi(0, Q) \tag{16}$$

where $\widehat{\Psi}_{k|k-1}$ is the shape and size matrix of $\widehat{\Omega}_{k|k-1}$, which is obtained by the Minkowski sum of $\widehat{\Omega}_{k-1}$ and V_k. Q represents the shape and size matrix of the error ellipsoid bundle V_k, which is generally set to the same value as $\widehat{\Psi}_0$ and does not change during the iterative update.

The prediction process of the output layer weights of the neural network model can be expressed as:

$$\widehat{\Omega}_{k|k-1} = \widehat{\Omega}_{k-1} \oplus V_k \tag{17}$$

According to Property 1, we have:

$$\xi(\widehat{\theta}_{k|k-1}, \widehat{\Psi}_{k|k-1}) = \xi(\widehat{\theta}_{k-1}, \widehat{\Psi}_{k-1}) \oplus \xi(0, Q) \tag{18}$$

where

$$\widehat{\theta}_{k|k-1} = \widehat{\theta}_{k-1} \tag{19}$$

$$\widehat{\Psi}_{k|k-1} = [\widehat{\Psi}_{k-1}, Q] \tag{20}$$

As a result, the prediction ellipsoid bundle is obtained.

Next, the prediction set needs to be corrected by the measurement outputs. In order to describe the feasible set resulting from intersection with S_k, the constructed ellipsoid set $\widehat{\Omega}_{k|k-1}$ is used to determine the intersection with S_k, iteratively. Due to the irregular shape of the intersection between $\widehat{\Omega}_{k|k-1}$ and

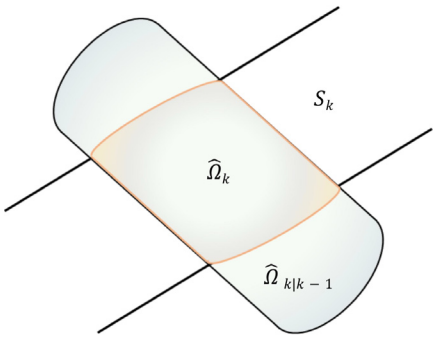

Fig. 3. Intersection between a ellipsoid bundle and a strip $\widehat{\Omega}_{k|k-1} \cap S_k \subseteq \widehat{\Omega}_k$.

S_k, which cannot be expressed by an analytical formula, a new bounding ellipsoid bundle $\widehat{\Omega}_k$ is constructed to indirectly describe this intersection. This iterative process, presented in Fig. 3, can be expressed as:

$$\widehat{\Omega}_k = \widehat{\Omega}_{k|k-1} \cap S_k \tag{21}$$

Then the following theorem provides an ellipsoid bundle $\widehat{\Omega}_k$ (parametrized by means of vector λ_k) that contains the intersection of an ellipsoid bundle $\widehat{\Omega}_{k|k-1}$ and a strip S_k. It is possible to choose the parameter vector λ_k in such a way that a size criterion for the obtained ellipsoid bundle is minimized.

Theorem 1. *Given the ellipsoid bundle* $\widehat{\Omega}_{k|k-1} = \{\theta_k : \theta_k = \widehat{\theta}_{k|k-1} + \widehat{\Psi}_{k|k-1} z\}$ *and the strip* $S_k = \{\theta_k \in \mathbb{R}^{m \times n_y} : (y_k - \varphi_k \theta_k)(y_k - \varphi_k \theta_k)^T \leq \gamma^2\}$, *the ellipsoid bundle that bounds* $\widehat{\Omega}_k$ *can be calculated iteratively as:*

$$\widehat{\Omega}_k = \widehat{\theta}_k \oplus \widehat{\Psi}_k z \tag{22}$$

where

$$\widehat{\theta}_k = \widehat{\theta}_{k|k-1} + \lambda_k (y_k - \varphi_k \widehat{\theta}_{k|k-1}) \tag{23}$$

$$\widehat{\Psi}_k = [(I - \lambda_k \varphi_k)\widehat{\Psi}_{k|k-1} \quad -\gamma \lambda_k] \tag{24}$$

with the vector $\lambda_k \in \mathbb{R}^{m \times 1}$ *determined by*

$$\lambda_k = \frac{\widehat{\Psi}_{k|k-1} \widehat{\Psi}_{k|k-1}^T \varphi_k^T}{\varphi_k \widehat{\Psi}_{k|k-1} \widehat{\Psi}_{k|k-1}^T \varphi_k^T + \gamma^2} \tag{25}$$

Proof According to the definition of ellipsoid bundles (Tang et al., 2020), adding and subtracting $\lambda_k \varphi_k \widehat{\Psi}_{k|k-1} z$ to $\widehat{\theta}_{k|k-1} + \widehat{\Psi}_{k|k-1} z$, we have:

$$\begin{aligned} \theta_k =& \widehat{\theta}_{k|k-1} + \lambda_k \varphi_k \widehat{\Psi}_{k|k-1} z + \widehat{\Psi}_{k|k-1} z - \lambda_k \varphi_k \widehat{\Psi}_{k|k-1} z \\ =& \widehat{\theta}_{k|k-1} + \lambda_k \varphi_k \widehat{\Psi}_{k|k-1} z + (I_m - \lambda_k \varphi_k) \widehat{\Psi}_{k|k-1} z \end{aligned} \tag{26}$$

where I_m denotes a m dimensional identity matrix. According to (11), there exists $\omega \in \beta^1$ such that $y_k - \varphi_k \theta_k = \gamma \omega$. Taking into account (2), $y_k - \varphi_k(\widehat{\theta}_{k|k-1} + \widehat{\Psi}_{k|k-1} z) = \gamma \omega$, we have

$$\varphi_k \widehat{\Psi}_{k|k-1} z = y_k - \varphi_k \widehat{\theta}_{k|k-1} - \gamma \omega \tag{27}$$

Substituting $\varphi_k \widehat{\Psi}_{k|k-1} z$ into Eq. (26):

$$\begin{cases} \theta_k = & \widehat{\theta}_{k|k-1} + y_k - \lambda_k(\varphi_k \widehat{\theta}_{k|k-1} + \gamma \omega) \\ & + (I - \lambda_k \varphi_k)\widehat{\Psi}_{k|k-1} z \\ = & \widehat{\theta}_{k|k-1} + \lambda_k (y_k - \varphi_k \widehat{\theta}_{k|k-1}) - \lambda_k \gamma \omega \\ & + (I - \lambda_k \varphi_k)\widehat{\Psi}_{k|k-1} z \\ = & \widehat{\theta}_{k|k-1} + \lambda_k (y_k - \varphi_k \widehat{\theta}_{k|k-1}) \\ & + [(I - \lambda_k \varphi_k)\widehat{\Psi}_{k|k-1} \quad -\lambda_k \gamma][z \quad \omega]^T \\ \theta_k = & \widehat{\theta}_k + \widehat{\Psi}_k \beta^{n+1} \end{cases} \tag{28}$$

where

$$\begin{cases} \widehat{\theta}_k & = \widehat{\theta}_{k|k-1} + \lambda_k (y_k - \varphi_k \widehat{\theta}_{k|k-1}) \\ \widehat{\Psi}_k & = [(I - \lambda_k \varphi_k)\widehat{\Psi}_{k|k-1} \quad -\gamma \lambda_k] \end{cases} \tag{29}$$

To minimize the intersection between the enclosing ellipsoid bundle $\widehat{\Omega}_{k|k-1}$ and S_k, the optimal parameters λ_k need to be adjusted. The Frobenius norm of $\widehat{\Psi}_k$ can be utilized to measure the size of the intersection between $\widehat{\Omega}_{k|k-1}$ and S_k. The optimal λ_k is obtained by minimizing such a norm.

According to the above derivation, given $\widehat{\Psi}_k = [(I - \lambda_k \varphi_k) \widehat{\Psi}_{k|k-1} \quad -\lambda_k \gamma]$, let $A = [\widehat{\Psi}_{k|k-1} \quad \mathbf{0}_{m \times 1}]$ and $B^T = [-\varphi_k \widehat{\Psi}_{k|k-1} \quad -\gamma]$, $\widehat{\Psi}_k$ can be represented as:

$$\widehat{\Psi}_k = A + \lambda_k B^T \tag{30}$$

where $\mathbf{0}_{m \times 1}$ denotes an $m \times 1$ dimensional zero matrix. The Frobenius norm of $\widehat{\Psi}_k$ is :

$$\begin{aligned} \|\widehat{\Psi}_k\|_F^2 &= \|A + \lambda_k B^T\|_F^2 \\ &= tr((A^T + B\lambda_k^T)(A + \lambda_k B^T)) \\ &= tr(A^T A) + tr(B\lambda_k^T A) + tr(A^T \lambda_k B^T) \\ &\quad + tr(B\lambda_k^T \lambda_k B^T) \\ &= 2\lambda_k^T AB + (B^T B)\lambda_k^T \lambda_k + tr(A^T A) \end{aligned} \tag{31}$$

where $tr(A^T A)$ is a constant value. The minimum of $\|\widehat{\Psi}_k\|_F^2$ is achieved at

$$\lambda_k = \frac{-AB}{B^T B} = \frac{\widehat{\Psi}_{k|k-1} \widehat{\Psi}_{k|k-1}^T \varphi_k^T}{\varphi_k \widehat{\Psi}_{k|k-1} \widehat{\Psi}_{k|k-1}^T \varphi_k^T + \gamma^2} \tag{32}$$

Remark 2: Using the Frobenius norm, the size of the zonotope can be minimized at each sample time by an analytic expression (Alamo et al., 2005), being more convenient and easy to be implemented. Hence in this paper, a recursive formula for the parameters of the ellipsoid bundles is derived based on the Frobenius norm method, resulting in a theoretically minimal set of ellipsoid bundles.

After obtaining the theoretically smallest ellipsoid bundle set $\xi(\widehat{\theta}_k, \widehat{\Psi}_k)$, based on the properties of ellipsoid bundles and Lemma 1, the interval estimates of the outputs can be obtained.

Theorem 2. *Considering (9), if the interval set of weight parameters* $\xi(\widehat{\theta}_k, \widehat{\Psi}_k)$ *is obtained, then the prediction interval of output* y_k *can be calculated by:*

$$\begin{cases} \underline{y}_k = \varphi_k \underline{\theta}_k - \mathscr{R}(\gamma) \\ \overline{y}_k = \varphi_k \overline{\theta}_k + \mathscr{R}(\gamma) \end{cases} \tag{33}$$

where $\mathscr{R}(\gamma) \in \mathbb{R}^{1 \times n_y}$ *consists of a vector of repeat copies of* γ.

Proof: *If* $\theta_k \in \xi(\widehat{\theta}_k, \widehat{\Psi}_k) = \widehat{\theta}_k \oplus \bigoplus_{i=1}^n \mathcal{E}\left(0, \widehat{H}_{k,i}\right)$ *with* $\widehat{P}_{k,i} = \widehat{H}_{k,i} \widehat{H}_{k,i}^T$, *according to Lemma 1, an interval vector* $[\underline{\theta}_k, \overline{\theta}_k]$ *can be calculated, such that* $\underline{\theta}_k < \widehat{\theta}_k < \overline{\theta}_k$ *satisfying*

$$\begin{cases} \underline{\theta}_k = \widehat{\theta}_k - \widehat{b}_k \\ \overline{\theta}_k = \widehat{\theta}_k + \widehat{b}_k \end{cases} \tag{34}$$

with

$$\widehat{b}_k(l) = \sum_{i=1}^n \sqrt{\widehat{P}_{k,i}(l,l)}, \; l = 1, 2, \ldots, m \tag{35}$$

where $\widehat{H}_{k,i}$ *represents the ith matrix in the set of matrices* $\widehat{\Psi}_k$.

Then, combining (9) and (10), the model output interval can be expressed as $y_k \in [\underline{y}_k, \overline{y}_k]$ *in (33).*

Note that (35) differs from directly summing the absolute values of the generator matrix $\widehat{\Psi}_k$ in zonotope, resulting in a more tight interval. Furthermore, the generated interval can be utilized as a threshold for system monitoring and fault detection. Under normal operating conditions, the system output remains within the generated interval. However, in the event of a fault, the output values will exceed this interval. The fault detection logic is as follows:

$$\begin{cases} y_k \in [\underline{y}_k, \quad \overline{y}_k], & \text{No fault} \\ y_k \notin [\underline{y}_k, \quad \overline{y}_k], & \text{Fault occurs} \end{cases} \tag{36}$$

Table 1
Water quality parameter auxiliary variable of BSM1.

Auxiliary Variable	Unit
Influent chemical oxygen demand (I_{COD})	g/m^3
Effluent chemical oxygen demand (E_{COD})	g/m^3
Influent total nitrogen (I_{TN})	g/m^3
Effluent total nitrogen (E_{TN})	g/m^3
Influent biochemical oxygen demand (I_{BOD})	g/m^3
Influent suspended solids (I_{SS})	g/m^3
Effluent suspended solids (E_{SS})	g/m^3

Table 2
Faults in BSM1.

Fault.No	Fault Type	Fault Time
Fault 1	Sensor abrupt fault in BOD (increased by 2 g/m^3)	10–12 (day)
Fault 2	Toxic impact fault ($\mu H = 8, v = 100$)	10–14 (day)
Fault 3	Sludge expansion fault ($\mu H = 8, BH = 10, v = 100$)	10–14 (day)

Table 3
Parameter setting.

Model	ϵ	γ	$\hat{\Psi}_0$	Q
BP-EB	50	0.03	$0.02I_{10}$	$0.02I_{10}$
RBF-EB	5	0.03	$0.2I_{10}$	$0.2I_{10}$
LSTM-EB	20	0.03	$0.05I_{10}$	$0.05I_{10}$
KELM-EB	200	0.03	$0.005I_{10}$	$0.005I_{10}$

Algorithm 1 shows the overall pseudocode of the proposed method.

Algorithm 1 Fault detection based on ellipsoid bundles

Input: Initialize $\widehat{\Psi}_0$, $\widehat{\theta}_0$, γ, s and n.
Output: Fault detection results.
1: **for** $k = 1$ to n **do**
2: Update $\widehat{\theta}_{k-1}$ and $\widehat{\Psi}_{k-1}$ to $\widehat{\theta}_{k|k-1}$ and $\widehat{\Psi}_{k|k-1}$ by (19) and (20).
3: Calculate the prediction error by $y_k - \varphi\widehat{\theta}_{k|k-1}$.
4: Correct $\widehat{\theta}_{k|k-1}$ to $\widehat{\theta}_k$, $\widehat{\Psi}_{k|k-1}$ to $\widehat{\Psi}_k$ and update λ_k by (23)-(25).
5: Reorder the dimensions of $\widehat{\Psi}_k$ by (5) and (6).
6: Set $k \leftarrow k + 1$.
7: **if** $k = n$ **then**
8: Obtain $\widehat{\theta}_n$ and $\widehat{\Psi}_n$, construct an outer bounding box for the ellipsoid bundle, and obtain $\overline{\theta}_k$ and $\underline{\theta}_k$ by (34)-(35).
9: Calculate $[\underline{y}_k \quad \overline{y}_k]$ by (33).
10: Determine whether a fault occurs by(36).
11: **else**
12: Go back to Step 2.
13: **end**
14: **end for**

5. Simulation results

In this section, simulation results in a WWTP based on Benchmark Simulation Model 1 (BSM1) (Pons et al., 2021) are presented to demonstrate the effectiveness of the proposed strategy.

First, a 14-day simulation of BSM1 is conducted with constant flow and composition. During this time, data are collected every 15 min. Seven crucial variables are selected for monitoring, which include 4 input variables and 3 output variables related to water quality as shown in Table 1. Besides, the following three types of faults in WWTP are considered and defined in Table 2.

The weights of some typical neural networks such as BP, RBF, LSTM and KELM are trained based on the gradient descent method. The parameter settings of the model are shown in Table 3. These models are trained with the first 1535 data samples, and tested with the remaining 481 samples. Then the trained weights are optimized by the set-membership methods. In this paper, ellipsoids and zonotopes are used for comparison. Besides, the shape matrix of ellipsoid bundles and zonotopes are both reduced to 20 dimensions.

To evaluate the generated prediction interval and fault diagnosis performance, the PI coverage probability (PICP), PI normalized averaged width (PINAW), root mean squared error (RMSE), fault diagnosis rate (FDR), and fault alarm rate (FAR) (Ding et al., 2019; Quan et al., 2014) are used. The expressions are given as:

$$\text{PICP} = \frac{1}{n}\sum_{k=1}^{n}\pi_k \tag{37}$$

$$\text{PINAW} = \frac{1}{nR}\sum_{k=1}^{n}(\overline{y}_k - \underline{y}_k) \tag{38}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(y_{k,true} - \varphi_k\widehat{\theta}_k - \gamma)^2} \tag{39}$$

$$FDR = (\frac{FD}{TD})\times 100\% \tag{40}$$

$$FAR = (\frac{FA}{TA})\times 100\% \tag{41}$$

where $y_{k,true}$ and $\varphi_k\hat{\theta}_k$ represent the simulated value and the predicted result of effluent indicator, respectively. If the actual value is within the range $[\underline{y}_k, \overline{y}_k]$, then $\pi_k = 1$. Otherwise, $\pi_k = 0$. If all the true values are included in the prediction interval, PICP= 100%. Hence, the larger the PICP, the better the monitoring performance. R is the range of the underlying target defined as the difference between its minimum and maximum values. PINAW is used to measure the width of the generated interval, which indicates that the narrower the interval, the higher the information and value. *FD* is the times that the system indicates a fault occurs when there is no fault, *TD* is the sum of true alarms and false alarms. *FA* is the times that the system correctly identifies a fault, *TA* is the sum of correct and missed detections.

As shown in Table 4 and Fig. 4, compared with other methods such as in Chai and Qiao (2014) (RBF-MTOBE), in Jose and Wen (2009) (BP-MTOBE) and in Zhou et al. (2023) (LSTM-MTOBE), the proposed method has significant improvements both in prediction accuracy and interval performance. Additionally, obtaining a 100% confidence interval using kernel density estimation (KDE) method (Stief & Baranowski, 2021) is challenging and carries the risk of fault detection. This also demonstrates that the proposed method can generate more tight intervals, achieving more accurate fault detection results.

6. Conclusion

In this paper, a novel interval neural network model framework is proposed in which the weights are optimized by ellipsoid bundles. The proposed method is suitable for any kind of neural network models that the relationship between the hidden layer

Table 4
Fault Detection results of different methods.

Model		Fault 1					Fault 2					Fault 3				
		PICP	PINAW	RMSE	FDR	FAR	PICP	PINAW	RMSE	FDR	FAR	PICP	PINAW	RMSE	FDR	FAR
BP	MTOBE	100%	0.3832	0.0775	89.3%	0%	100%	0.3777	0.0812	100%	0%	100%	0.2011	0.0865	100%	0%
	Zonotope	100%	0.2158	0.0625	94.6%	0%	100%	0.2419	0.0512	100%	0%	100%	0.2019	0.0462	100%	0%
	EB	100%	0.1758	0.0511	98.5%	0%	100%	0.1786	0.0462	100%	0%	100%	0.1957	0.0454	100%	0%
	KDE	99.8%	0.1612	0.0785	100%	0.1%	99.8%	0.1791	0.0792	100%	0.1%	99.7%	0.1747	0.0821	100%	0.1%
RBF	MTOBE	100%	0.3765	0.0738	86.9%	0%	100%	0.3772	0.0749	100%	0%	100%	0.2020	0.0736	100%	0%
	Zonotope	100%	0.2623	0.0651	87.3%	0%	100%	0.2767	0.0643	100%	0%	100%	0.2021	0.0643	100%	0%
	EB	100%	0.1826	0.0613	98.5%	0%	100%	0.1753	0.0642	100%	0%	100%	0.1787	0.0613	100%	%
	KDE	99.6%	0.1721	0.0749	100%	0.1%	99.7%	0.1752	0.0751	100%	0.1%	99.9%	0.1757	0.0760	100%	0.1%
LSTM	MTOBE	100%	0.3647	0.0623	88.6%	0%	100%	0.3778	0.0489	100%	0%	100%	0.1896	0.0587	100%	0%
	Zonotope	100%	0.2323	0.0500	89.7%	0%	100%	0.2689	0.0545	100%	0%	100%	0.3654	0.0509	100%	0%
	EB	100%	0.1852	0.0448	95.5%	0%	100%	0.1736	0.0523	100%	0%	100%	0.1819	0.0462	100%	0%
	KDE	99.8%	0.1786	0.0598	100%	0.1%	99.7%	0.1786	0.0556	100%	0.1%	99.7%	0.1897	0.0551	100%	0.1%
KELM	MTOBE	100%	0.4152	0.0118	86.9%	0%	100%	0.3998	0.0112	100%	0%	100%	0.1703	0.0155	100%	0%
	Zonotope	100%	0.4295	0.0118	86.9%	0%	100%	0.4801	0.0114	100%	0%	100%	0.1667	0.0156	100%	0%
	EB	100%	0.2470	0.0112	94.0%	0%	100%	0.2327	0.0113	100%	0%	100%	0.1651	0.0156	100%	0%
	KDE	99.8%	0.2992	0.0118	100%	0.1%	99.7%	0.2595	0.0118	100%	0.1%	99.8%	0.1627	0.0156	100%	0.1%

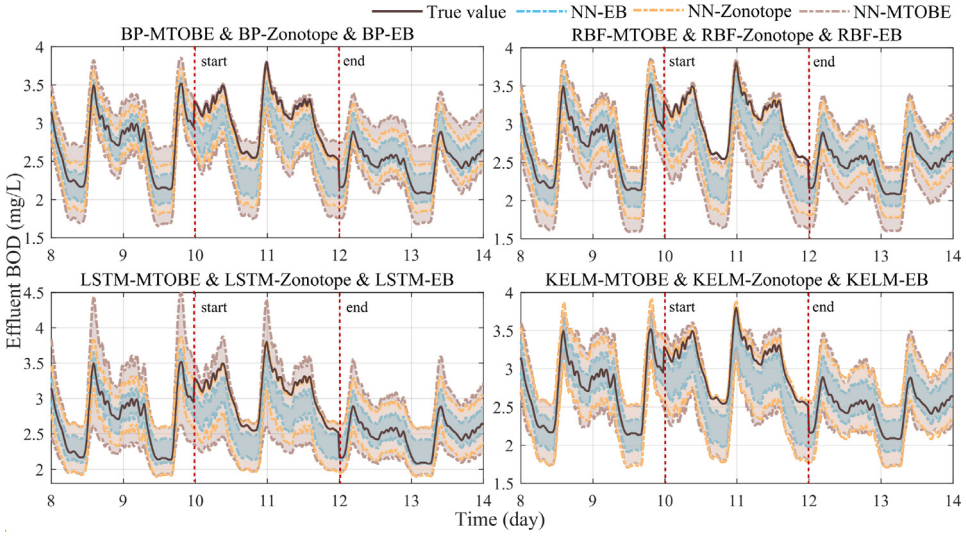

Fig. 4. Comparative fault detection results for Fault 1.

and the output layer satisfies linear characteristic. This paper takes BP, RBF, LSTM, KELM as examples and applied them to fault detection for a wastewater treatment process. Simulation results demonstrate that the proposed method can enhance higher prediction accuracy and fault detection performance compared with other methods.

References

Alamo, T., Bravo, J. M., & Camacho, E. F. (2005). Guaranteed state estimation by zonotopes. *Automatica*, *41*(6), 1035–1043.

Chai, W., & Qiao, J. F. (2014). Passive robust fault detection using RBF neural modeling based on set membership identification. *Engineering Applications of Artificial Intelligence*, *28*, 1–12.

Combastel, C. (2015). Zonotopes and Kalman observers: Gain optimality under distinct uncertainty paradigms and robust convergence. *Automatica*, *55*, 265–273.

Ding, S. X., Li, L., & Krüger, M. (2019). Application of randomized algorithms to assessment and design of observer-based fault detection systems. *Automatica*, *107*, 175–182.

Jia, M. W., Xu, D. Y., Yang, T., Liu, Y., & Yao, Y. (2023). Graph convolutional network soft sensor for process quality prediction. *Journal of Process Control*, *123*, 12–25.

Jose, D. J. R., & Wen, Y. (2009). Neural networks training with optimal bounded ellipsoid algorithm. *Neural Computing and Applications*, *18*, 623–631.

Li, X., Wei, G. L., & Wang, L. C. (2021). Distributed set-membership filtering for discrete-time systems subject to denial-of-service attacks and fading measurements: A zonotopic approach. *Information Sciences*, *547*, 49–67.

Ljung, L. (1998). *System identification: Theory for the user second edition*. United States of America: Pearson Education.

Nourani, V., Shahidi, Z., & Dini, M. (2023). Estimation of prediction intervals for uncertainty assessment of artificial neural network based wastewater treatment plant effluent modeling. *Journal of Water Process Engineering*, *55*, Article 104145.

Pons, M. N., Alex, J., & Benedetti, L. (2021). Benchmark simulation model no. 1 (BSM1). Wastewater Modelling. Retrieved from https://wwtmodels.pubpub.org/pub/web0fxkd.

Quan, H., Srinivasan, D., & Khosravi, A. (2014). Particle swarm optimization for construction of neural network-based prediction intervals. *Neurocomputing*, *127*, 172–180.

Schweppe, F. (1968). Recursive state estimation: Unknown but bounded errors and system inputs. *IEEE Transactions on Automatic Control*, *13*, 22–28.

Stief, A., & Baranowski, J. (2021). Fault diagnosis using interpolated kernel density estimate. *Measurement*, *176*, Article 109230.

Sun, Q. Q., & Ge, Z. Q. (2021). A survey on deep learning for data-driven soft sensors. *IEEE Transactions on Industrial Informatics*, *17*, 5853–5866.

Tang, W. T., Zhang, Q. H., Wang, Z. H., & Shen, Y. (2020). Ellipsoid bundle and its application to set-membership estimation. *IFAC-PapersOnLine*, *53*, 13688–13693.

Tang, W. T., Zhang, Q. H., Wang, Z. H., & Shen, Y. (2022). Set-membership estimation based on ellipsoid bundles for discrete-time LPV descriptor systems. *Automatica*, *145*, Article 110580.
Wang, G. M., Jia, Q. S., Zhou, M. C., Bi, J., Qiao, J. F., & Abusorrah, A. (2021). Artificial neural networks for water quality soft-sensing in wastewater treatment: A review. *Artificial Intelligence Review*, *55*, 565–587.
Zhang, W. H., Wang, Z. H., Raissi, T., & Shen, Y. (2022). Ellipsoid-based interval estimation for Lipschitz nonlinear systems. *IEEE Transactions on Automatic Control*, *67*, 6802–6809.
Zhou, M., Zhang, Y. Y., Wang, J., & Xue, T. L. (2023). LSTM-OBE based interval prediction of effluent BOD for wastewater treatment. *IFAC-PapersOnLine*, *56*(2), 8488–8493.
Zhu, F. L., Tang, Y. Y., & Wang, Z. H. (2022). Interval-observer-based fault detection and isolation design for T-S fuzzy system based on zonotope analysis. *IEEE Transactions on Fuzzy Systems*, *30*, 945–955.

Meng Zhou received her B.S. degree from Harbin Engineering University, Harbin, China, in 2011 and her M.S. degree and Ph.D. degree from Harbin Institute of Technology, Harbin, China, in 2013 and 2018, respectively. From 2018 to 2020, she was a Post-doctor with the College of Information Science & Technology, Beijing University of Chemical Technology, Beijing. She is currently a Professor in School of Electrical and Control Engineering, North China University of Technology, Beijing. Her interests include fault detection and fault tolerant control, multi-agent cooperation control.

Yinyue Zhang received his B.S. degree from North China University of Technology in 2021 and his M.S degree from North China University of Technology in 2024, respectively. His main research is fault detection for wastewater treatment process.

Jing Wang received the B.S. degree in industry automation, and the Ph.D. degree in control theory and control engineering from Northeastern University, Shenyang, China, in 1994 and 1998, respectively. She is a Professor with the College of Information Science & Technology, Beijing University of Chemical Technology, Beijing, China. She was a Visiting Professor with the University of Delaware, Newark, DE, USA, in 2014. Her interests are in data-driven modelling, optimization and control for complex industrial process.

Tarek Raïssi (Senior Member, IEEE) received the engineering degree in electrical engineering from Ecole Nationale d'Ingenieurs de Tunis, Tunis, Tunisia, in 2000, the M.Sc. degree in automatic control from Ecole Centrale deLille, Villeneuve-d'Ascq, France, in 2001, and the Ph.D. degree in automatic control from the University of Paris XII, Creteil, France, in 2004. He was accreditation to Supervise/Conduct Research (HDR) in 2012 from the University of Bordeaux 1, Talence, France, where, from 2005 to 2011, he was an Associate Professor. He is currently a Full Professor with the Conservatoire National des Arts et Metiers, Paris, France. His research interests include fault detection and isolation, nonlinear systems estimation, and robust control. Dr. Raissi is a Member of the IFAC Technical Committee Modeling, Identification, and Signal Processing.

Vicenç Puig holds a Ph.D. in Automatic Control, Vision and Robotics and is the leader of the research group Advanced Control Systems (SAC) at the Polytechnic University of Catalonia. He has important scientific contributions in the areas of fault diagnosis and fault tolerant control using interval models. He participated in more than 20 international and national research projects in the last decade. He led many private contracts, and published more than 80 articles in JCR journals and more than 350 in international conference/workshop proceedings. Prof. Puig supervised over 20 Ph.D. theses and over 50 MA/BA theses.