Measurement 188 (2022) 110621

Contents lists available at ScienceDirect

Measurement

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

Bayesian prediction of tunnel convergence combining empirical model and relevance vector machine

Xiangyu Chang [a], Hao Wang [a,*], Yiming Zhang [a,*], Feiqiu Wang [b], Zhaozhong Li [a,b]

[a] Key Laboratory of C&PC Structures of Ministry of Education, Southeast University, Nanjing 211189, Jiangsu, China
[b] China Railway 24th Bureau Group Co., Ltd., Shanghai 200071, China

ARTICLE INFO

Keywords:
Tunnel engineering
Convergence prediction
Empirical model
Bayesian estimation
Relevance vector machine

ABSTRACT

Convergence monitoring of the tunnel is a direct and reliable way to reveal its status. Accurate prediction of convergence is essential to prevent safety hazards, such as rock collapse, project delay, and even geological disasters. Convergence prediction is usually carried out based on numerical methods (NMs) and empirical models (EMs). The accurate model parameters of NMs are difficult to estimate from limited field geological tests, which severely undermines its prediction accuracy. Although the EMs involve the advantages of computational simplicity and efficiency, their prediction capacity is relatively limited. This paper presents a probabilistic model that combines the EM, Bayesian estimation, and relevance vector machine (RVM) to predict convergence. Firstly, various EMs integrated with Bayesian estimation are established and prediction results are compared to select the EM with higher prediction accuracy. Prediction residuals of the EM are then modeled by the RVM to further improve the accuracy. A high-speed railway tunnel is utilized to demonstrate the effectiveness of the presented approach. The results show that the root-mean-squared error values of the combined probabilistic model are reduced by 92.6% and 95.8% compared with the EM for two data sets. Moreover, the comparison results show that the presented model exhibits higher prediction accuracy than backpropagation neural network and Gaussian process regression.

1. Introduction

High-speed railways serve as a significant component in providing convenient travel options and promoting economic development [1,2]. As a vital part of high-speed railway, the high-speed railway tunnel exceeds 5,500 km in length by the end of 2019 in China [3]. The size and the burial depth of tunnels are significantly increasing, especially for tunnel projects using the New Austrian Tunneling Method (NATM) [4]. This trend is inevitably accompanied by the increasing convergence deformation during construction. Severe convergence deformation could cause unexpected safety problems, such as cavern collapses, vault cracking, trapping, jamming of the boring machine, and even geological disasters [5]. Predicting the monitoring data of tunnel convergence at early stages is essential to identify the structural condition and timely adjust the construction design [6].

Generally, convergence prediction can be carried out through numerical methods (NMs) [7–9], and empirical models (EMs) [10–12]. NMs (such as FLAC 3D) predict tunnel convergence by simulating the actual tunnel construction and solving mechanical equations. The attractiveness of these techniques comes from some extruding advantages, including high efficiency, fast convergence, and low expense. Nevertheless, it is not easy to accurately estimate rock mass parameters in the finite element model from the limited field geological tests [13]. Alternatively, EMs [14–16] do not rely on the finite element model and directly predict the tunnel convergence using the measured data. Sulem et al. [15] developed a new EM for predicting the tunnel convergence and adapted the support according to the conditions encountered during the excavation. Kontogianni et al. [16] estimated the tunnel convergence based on the monitoring data of two tunnels. To guarantee the prediction accuracy, various optimization techniques are used to estimate parameters of EMs, including least squares, Levenberg-Marquardt, and trust-region [17,18]. Such optimization algorithms are deterministic methods that only provide a single value for EM parameters and fail to quantify uncertainties induced by smooth blasting, groundwater, and/or non-standard construction. Bayesian estimation has been extensively applied in rock mechanics and engineering [19]. It quantifies the uncertainties by offering the posterior distribution of the parameter based on the prior and the new information provided by successive convergence measurements [20,21]. Following the EM

* Corresponding authors at: Key Laboratory of C&PC Structures of Ministry of Education, School of Civil Engineering, Southeast University, Nanjing 211189, China.
E-mail addresses: wanghao1980@seu.edu.cn (H. Wang), zhangyiming1992@seu.edu.cn (Y. Zhang).

https://doi.org/10.1016/j.measurement.2021.110621
Received 20 March 2021; Received in revised form 2 October 2021; Accepted 15 December 2021
Available online 21 December 2021

Nomenclature

X	Characteristic length related to the distance of influence of the tunnel face
$C_{\infty x}$	Instantaneous closure obtained at an infinite rate of the face advance
T	Characteristic time depending on the ground
x	Distance from the measuring point to the tunnel face
t	Time for the excavation of the monitoring section
A_x	Coefficient that depends on the distance x
m	Ratio between A_x and $C_x(x)$
d_x	Time-dependent convergence at a stable level
U	Measured values of tunnel convergence
μ	Mean value
σ^2	Variance value
ψ	Location parameter
γ	Scale parameter
D	Monitoring data of the tunnel convergence
$\boldsymbol{\theta}$	Model parameters of EMOS
$\boldsymbol{\theta}^*$	Model parameters of EMOS with minimum RMSE value
$\boldsymbol{\theta}_0$	First sample of the Markov chain
ϕ	Prior knowledge of EMOS parameters
E	Predicted residuals of EMOS
p	Length of the observation window
$\boldsymbol{l}_k$	Input vector
$\boldsymbol{h}$	Target vector
ε_k	Random noise
$\boldsymbol{w}$	Weight vector
w_0	Bias
Φ	Design matrix
$\boldsymbol{\alpha}$	Hyperparameters
$\sum$	Covariance matrix
u	Kernel parameter of RBF kernel function
$\widehat{C}$	Prediction results of the EM
$\widehat{E}$	Residual prediction results of RVM
NATM	New Austrian Tunneling Method
NM	Numerical method
EM	Empirical model
EMOS	An EM of Sumel et al.
EMOK	An EM of Kontogianni et al.
GPR	Gaussian process regression
BPNN	Backpropagation neural network
GA	Genetic algorithm
BU	Bayesian updating method
RVM	Relevance vector machine
MSE	Mean square error
RMSE	Root-mean-squared error
MAPE	Mean absolute percentage error
NSE	Nash–Sutcliffe efficiency
ML	Machine learning
PDF	Probability density function
MCMC	Markov Chain Monte Carlo simulation
CI	Confidence interval

proposed by Sulem et al. [15], Feng et al. [13] combined the EM with a Bayesian updating approach to predict tunnel convergence. It was found that the developed approach exhibited better predictive performance than the deterministic methods. Fundamentally, although the EMs involve the advantages of computational simplicity and efficiency, its prediction accuracy may be undermined when the nonlinear and nonstationary relationships are dominant [22]. Thus, significant residuals may exist in the prediction results of EMs, especially for tunnel construction under the complicated geological environment [23].

Machine learning (ML) techniques, such as support vector machine (SVM) [24], Gaussian process regression (GPR) [25], and backpropagation neural network (BPNN) [26], have been widely used in predicting the tunnel convergence. Compared with simple regression analysis methods, ML techniques can provide high-order approximations for nonlinear problems [25]. However, some inherent drawbacks have also been observed, such as slow convergences, less generalized performances, arriving at local minima, and overfitting problems [27]. BPNN and GPR generally need a large number of training samples to obtain a satisfactory prediction accuracy [24,26]. SVM-based models typically exhibit superior generalization capacities in handling small sample problems since they adopt a structural risk minimization principle to minimize the training error [27]. But, the application of SVMs is inherently limited by Mercer's condition and the difficulty in selecting appropriate hyperparameters [24]. To overcome such issues, Tipping [28] proposed the relevance vector machine (RVM) algorithm based on the Bayesian theory and SVM. RVM is characterized by its two merits. Firstly, most of its parameters are automatically updated to the optimal value during the learning process [29]. Secondly, RVM proposes a sparse probabilistic framework without the limitation of kernel function form [24]. Due to these advantages, RVM generally outperforms SVM-based models concerning the prediction accuracy [27]. RVM has been applied in various fields such as fault diagnosis [29], traffic flow forecasting [30], and risk assessment [31]. However, the accuracy of predictive values generated by ML techniques is decreased in the long-term prediction of tunnel convergence [25,26]. In addition, long-term predictions of tunnel convergence are indeterminate, non-linear and nonstationary processes. Therefore, it is difficult to accurately predict tunnel convergence with a single ML method.

The machine learning-based residual learning approach is expected to provide an effective solution to further improve the prediction accuracy of the tunnel convergence [32]. It can use prediction residuals of EMs as input and models their internal relationship to modify the final prediction results [33]. Thus, in recognition of the modeling strength of RVM, it is applied to predict residuals of the EM and combined with the EM to improve the prediction accuracy for tunnel convergence.

In summary, the main contributions of this paper are three-fold. First, a probabilistic model combining the EM, Bayesian estimation, and RVM is presented to improve the prediction accuracy for tunnel convergence. More specifically, the EM with Bayesian estimation is applied to predict the tunnel convergence and produces residuals of the EM. The prediction residuals are modified by the RVM model. It retains the simplicity and efficiency of the EM and enables probabilistic prediction that offers the posterior distribution of the parameter. Second, to guarantee the modeling accuracy, different EMs with Bayesian estimation are fully compared to select the optimal one in terms of the prediction accuracy for tunnel convergence. Third, the feasibility of the proposed combined probabilistic model is validated using the convergence measurements collected from a high-speed railway tunnel. Results show that the proposed model exhibits higher accuracy and generalization performance than BPNN and GPR. The rest of this article is organized as follows. In Section 2, a new combined probabilistic model is presented to predict the tunnel convergence based on the EM, Bayesian estimation, and RVM. In Section 3, the effectiveness of the present approach is demonstrated by the monitoring data collected from a high-speed railway tunnel in construction. A comparison of eight prediction models is also performed. Finally, the concluding remarks of this study are presented in Section 4.

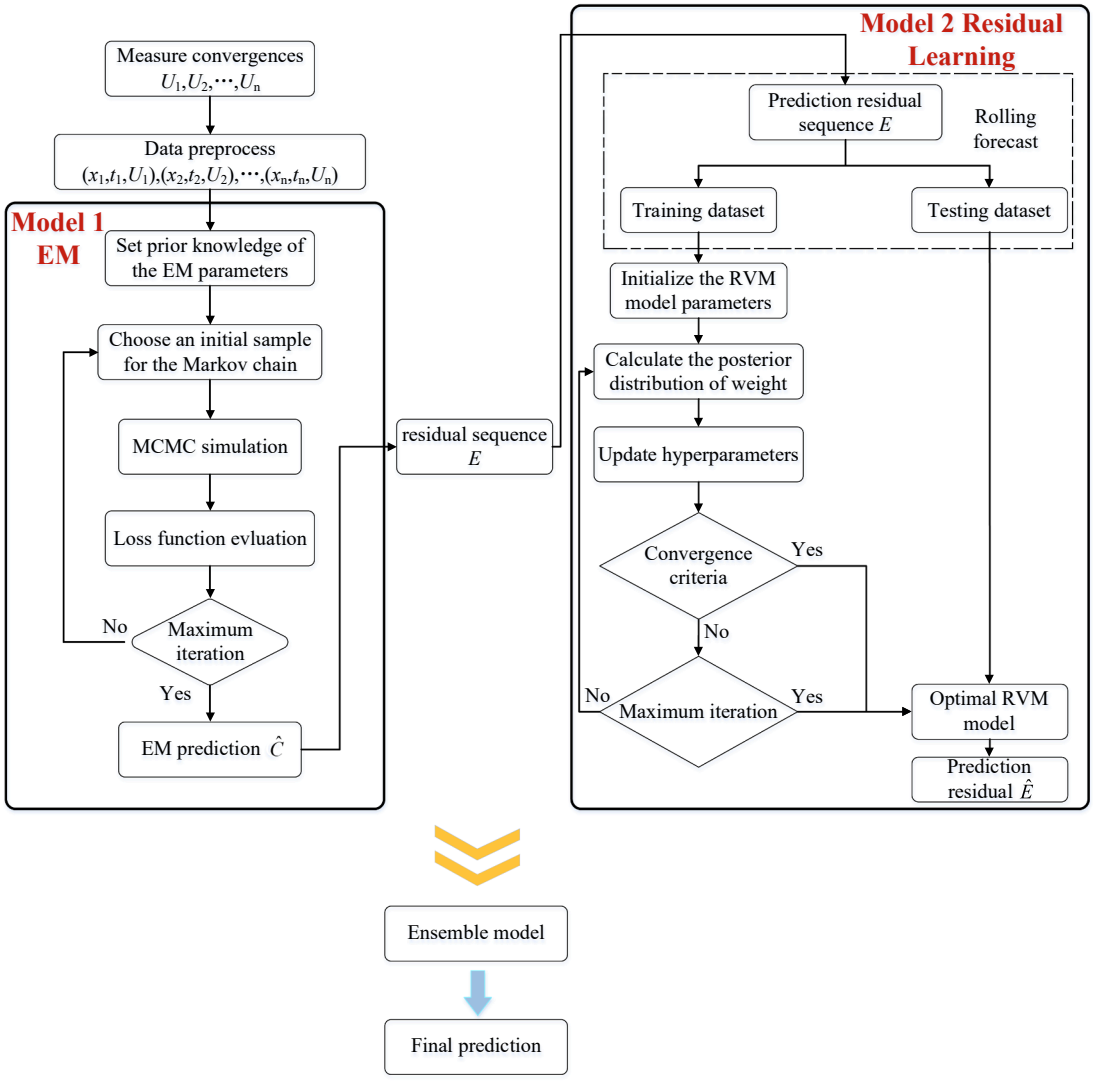

Fig. 1. The procedure of the combined probabilistic prediction model.

2. Prediction of tunnel convergence combining the EM and RVM

2.1. Tunnel convergence EMs

Convergence measurements are crucial for obtaining the tunnel deformation, selecting the excavation method, and ensuring construction safety and quality [34]. Generally, the convergence is induced by two effects: (i) the stress variation associated with the advancing tunnel face; and (ii) the time-dependent rheological behavior of the rock mass [13]. Several EMs [15,35] have been proposed to predict tunnel convergence by considering these two effects. Two EMs with satisfactory prediction performance are discussed as follows [16].

The tunnel convergence due to face advance effect is determined by [15]:

$$C_x(x) = C_{\infty x}\left[1 - \left(\frac{X}{X+x}\right)^2\right] \tag{1}$$

where $\times$ represents the distance from the measuring point to the tunnel face; X is the characteristic length related to the distance of influence of the tunnel face; and $C_{\infty x}$ corresponds to instantaneous closure obtained at an infinite rate of the face advance, excluding time-dependent effect.

When the excavation is suspended or the measuring point is far away from the tunnel face, the tunnel convergence mainly depends on the time-dependent properties of the rock mass (e.g., creep deformation and progressive extension of the yield zone). The tunnel convergence caused by the time-dependent effect can be expressed as follows [18]:

$$C_{ts}(t) = A_x\left[1 - \left(\frac{T}{T+t}\right)^{0.3}\right] \tag{2}$$

where t denotes the time for the excavation of the monitoring section; T represents a characteristic time depending on the ground; and A_x is a coefficient that depends on the distance x. The ratio between A_x and $C_x(x)$ is equal to a constant m. If the tunnel sections have similar characteristics (e.g., rock mass properties, excavation method, and support), the corresponding m values should be of the same order of magnitude [13].

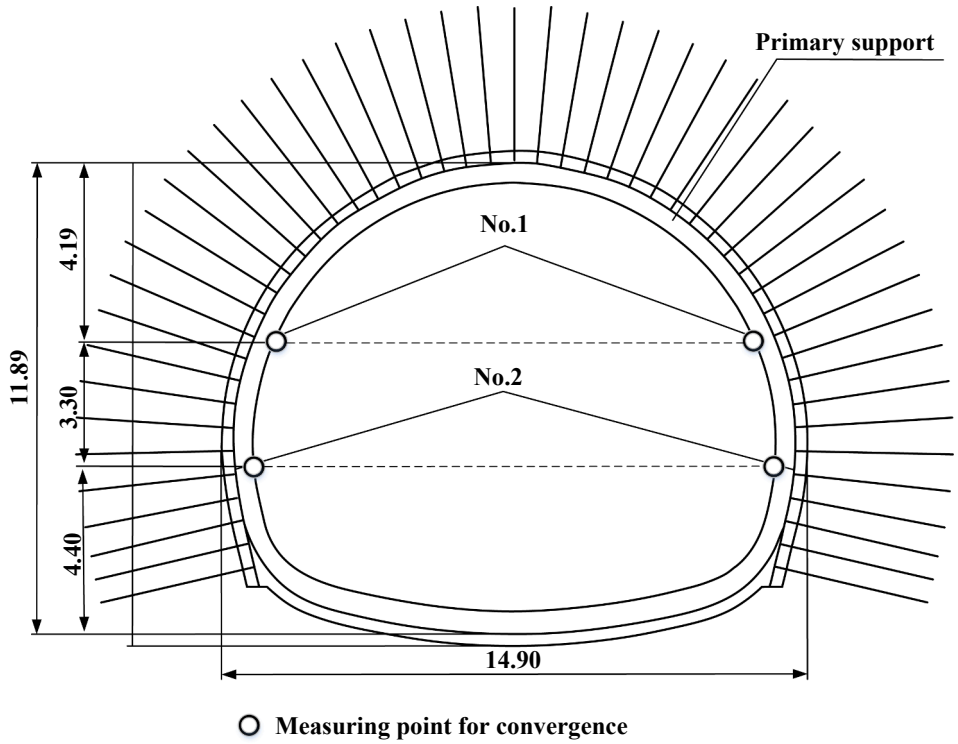

Fig. 2. A typical cross-section of the Yangshan tunnel (Unit: m).

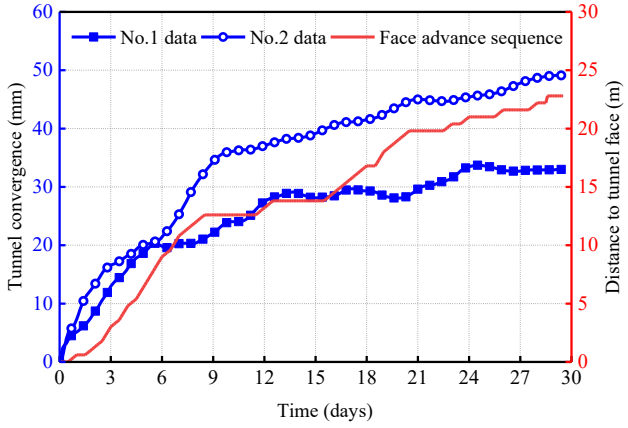

Fig. 3. Face advance sequence and convergence information for the Yangshan tunnel.

Table 1
The prior information of $C_{\infty x}$, m, and d_x.

Parameter	Mean	variance	Distribution
X	3	9	Normal
T	7	9	Normal

Table 2
The prior information of X and T.

Parameter	Range	Distribution
$C_{\infty x}$	0–60	Uniform
m	0–60	Uniform
d_x	0–60	Uniform

Assume the total convergence $C_1(x,\ t)$ is presented by the joint consideration of the face distance and the time effect. An EM of Sumel (EMOS) is defined as [15]:

$$C_1(x,\quad t) = C_{\infty x}\left[1-\left(\frac{X}{X+x}\right)^2\right]\left[1+m\left(1-\left(\frac{T}{T+t}\right)^{0.3}\right)\right] \tag{3}$$

For the time-dependent effect, Kontogianni et al. [16] also developed a model according to the Kelvin-Voigt creep model, which is expressed as:

$$C_{tk}(t) = d_x\left[1-exp(\frac{-t}{T})\right] \tag{4}$$

where d_x is the time-dependent convergence at a stable level (theoretically at the infinite time).

Kontogianni et al. [16] proposed the EM of Kontogianni (EMOK) and demonstrated that the total convergence is the sum of the tunnel face advance and the time-dependent effect. Given by Eq. (2) and Eq. (4), the EMOK is written as:

$$C_2(x,\quad t) = C_{\infty x}\left[1-\left(\frac{X}{X+x}\right)^2\right] + d_x\left[1-exp(\frac{-t}{T})\right] \tag{5}$$

2.2. Bayesian estimation of EM parameters

The EM parameters directly affect its prediction performance. Bayesian estimation approach is implemented to estimate EM parameters for considering the involved uncertainties. In this section, the parameters of the EMOS are taken as an example to illustrate the procedure of the Bayesian estimation.

2.2.1. Prior distribution

Let U_i (i = 0, 1, 2…, n) denote the measured value of tunnel convergence when the face is at distance × = x_i and for a time defined by $t = t_i$; and $C_1(x_i, t_i)$ denotes the tunnel convergence predicted by the EMOS at × = x_i and $t = t_i$. According to previous studies, the prior

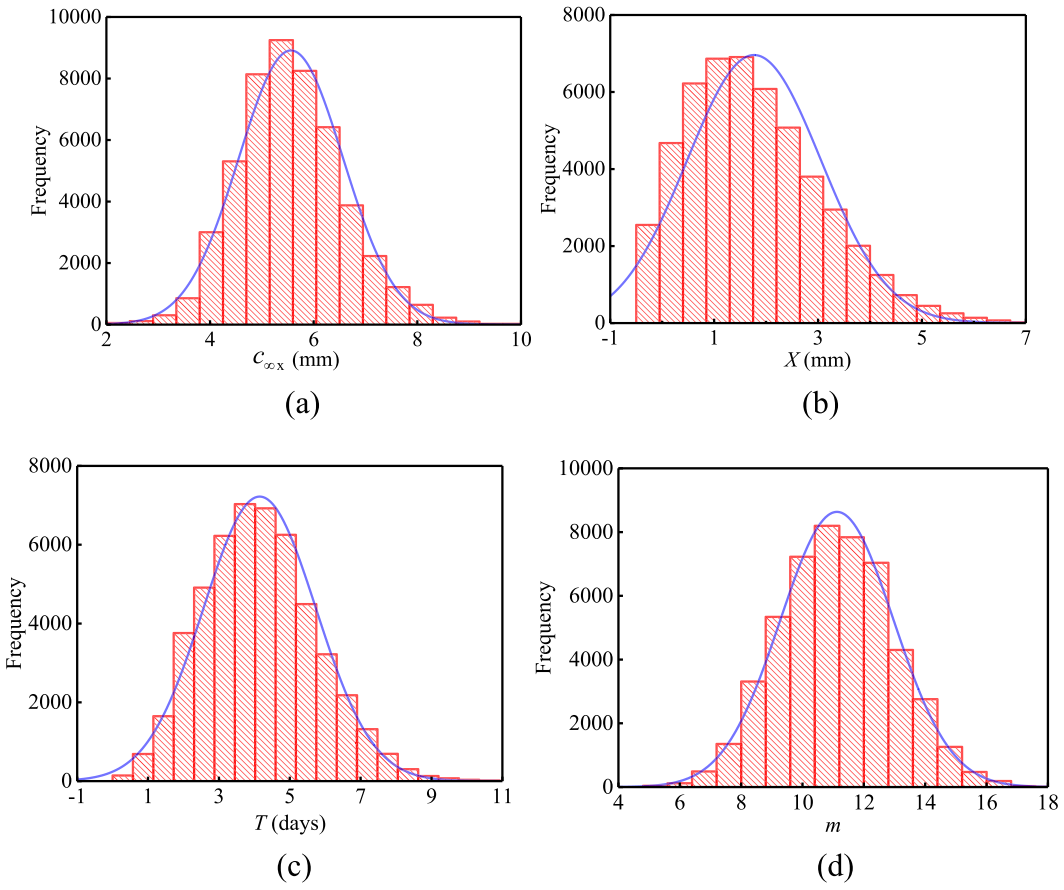

Fig. 4. Posterior parameter distributions of the EMOS used in MCMC (a) $C_{\infty x}$ (b) X (c) T (d) m.

distribution of $C_1(x_i, t_i)$ is assumed to be a normal random variable with the mean μ_0 and variance σ_0^2 [13,15,17]. Based on Kontogianni and Stiros [36], the variance σ_0 of good quality geodetic measurement methods ranges within 1 and 2 mm. Thus, σ_0 is assumed to be a half-Cauchy random variable with the location parameter $\psi = 0$ and scale parameter $\gamma = 1$. Then, the distribution of $C_1(x_i, t_i)$ is given by:

$$C_1(x_i, t_i) \sim N\left(\mu_0 = U_i, \sigma_0^2\right) \tag{6}$$

$$\sigma_0 \sim HalfCauchy(\psi, \gamma) \tag{7}$$

EM parameters are assumed to be mutually independent. The prior information of the parameters X and T are derived from the Frejus and Las Planas tunnel [15–17]. X and T are modeled by normal random variables. Note that the parameters $C_{\infty x}$ and m (d_x for EMOK) have no obvious regularity according to previous literature [15,16]. Thus, uniform distributions are used to reflect their prior knowledge instead of using the information directly.

2.2.2. Bayesian estimation

Let $D = \{((x_1, t_1), U_1), ((x_2, t_2), U_2), \cdots, ((x_n, t_n), U_n)\}$ denote the monitoring data of the tunnel convergence; and EMOS parameters in Eq. (3) are defined as $\boldsymbol{\theta} = [X, T, C_{\infty x}, m]$. In the Bayesian framework, prior distribution quantifies the prior knowledge, and the likelihood function determines the information provided by monitoring data D [20]. Thus, EMOS parameters in $\boldsymbol{\theta}$ are estimated by the combination of the prior knowledge and the likelihood function. The posterior distribution of $\boldsymbol{\theta}$ is given by [37]:

$$p(\mu_0, \mu_i, \sigma_0, \sigma_i | D) = Z_1 p(D | \mu_0, \mu_i, \sigma_0, \sigma_i) p(\mu_i, \sigma_i) \tag{8}$$

where μ_i and $\sigma_i^2 (i = 1, 2, 3, 4)$ are mean and variance of EMOS parameters in $\boldsymbol{\theta}$, respectively; $Z_I = \left(\iint p(D|\mu_0, \mu_i, \sigma_0, \sigma_i) p(\mu_i, \sigma_i) d\mu_i d\sigma_i\right)^{-1}$ is a normalizing constant; $p(D|\mu_0, \mu_i, \sigma_0, \sigma_i) = Pr\left[\left(C_j(x_i, t_i) = U_i\right) | \mu_0, \mu_i, \sigma_0, \sigma_i\right]$ is the likelihood function that quantifies the information supplied by the data D [37]; and $p(\mu_i, \sigma_i)$ is the prior distribution of $\boldsymbol{\theta}$.

Computing the posterior distributions of EMOS parameters in $\boldsymbol{\theta}$ are challenging in Eq. (8) because they are not common probability density functions (PDFs). The Markov Chain Monte Carlo (MCMC) simulation method has been used in geotechnical engineering to simulate sequence samples of random variables from prior knowledge [38]. The basic idea of MCMC is to extract values from an approximate distribution and then modify these values to better approximate the target posterior distribution [39]. The PDFs of EMOS parameters in $\boldsymbol{\theta}$ are obtained using numerical samples generated from the MCMC simulation [40]. An initial sample of the MCMC simulation is randomly selected from the prior knowledge of $\boldsymbol{\theta}$ and the next sample simulates only based on the current sample [13]. After a large number of samples of $\boldsymbol{\theta}$ are obtained, statistical analyses are calculated on the samples of $\boldsymbol{\theta}$ to evaluate their statistics (e.g., mean and standard deviation) and probability distribution (e.g., PDF and cumulative distribution function).

2.2.3. Loss function

To reduce the influence of the initial sample, the MCMC simulation is repeated 100 times. The optimal posterior distribution of $\boldsymbol{\theta}^*$ is obtained

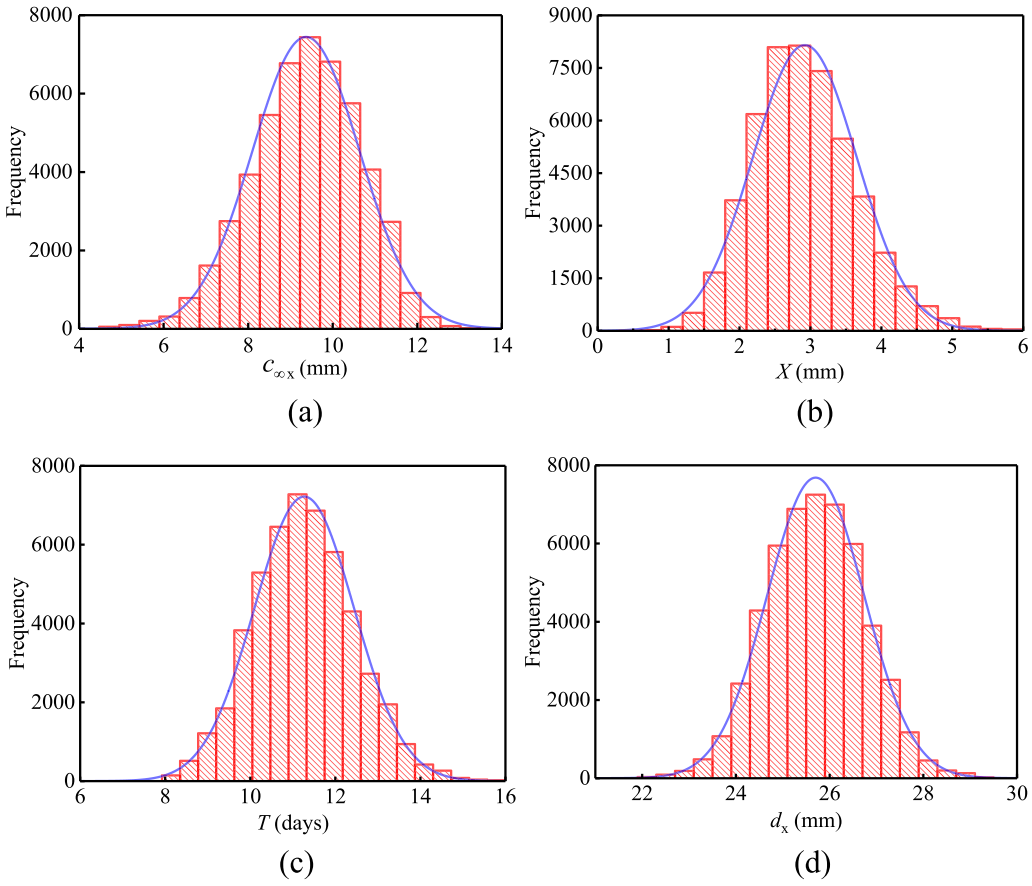

Fig. 5. Posterior parameter distributions of the EMOK used in MCMC (a) $C_{\infty x}$ (b) X (c) T (d) d_x.

Table 3
Summary of the estimated mean and variance of model parameters.

EMs		EMOS				EMOK			
Data		$C_{\infty x}$	X	T	m	$C_{\infty x}$	X	T	d_x
No.1	Mean	5.560	1.769	4.152	11.123	9.368	2.922	11.271	25.701
	Variance	1.015	1.665	2.524	3.416	1.622	0.539	1.379	1.080
No.2	Mean	5.722	0.618	6.136	18.852	3.240	1.903	9.846	46.984
	Variance	0.264	0.095	0.363	2.468	0.005	0.007	0.002	0.005

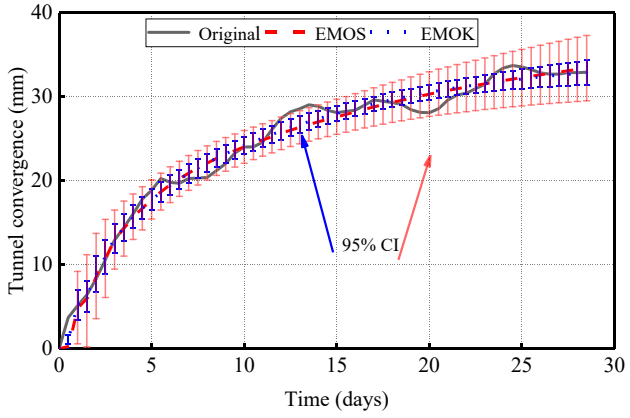

Fig. 6. Prediction results of two EMs for the No.1 data set (95% confidence interval).

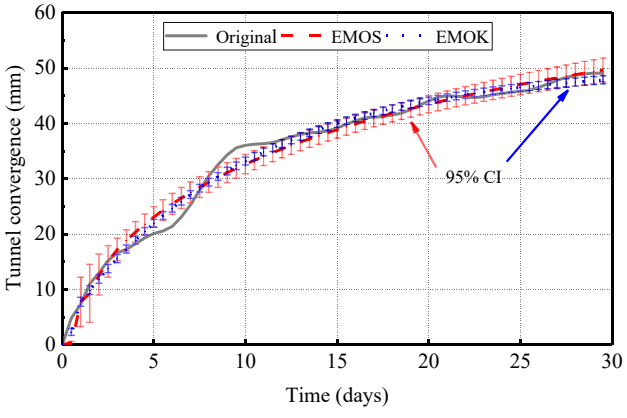

Fig. 7. Prediction results of two EMs for the No.2 data set (95% confidence interval).

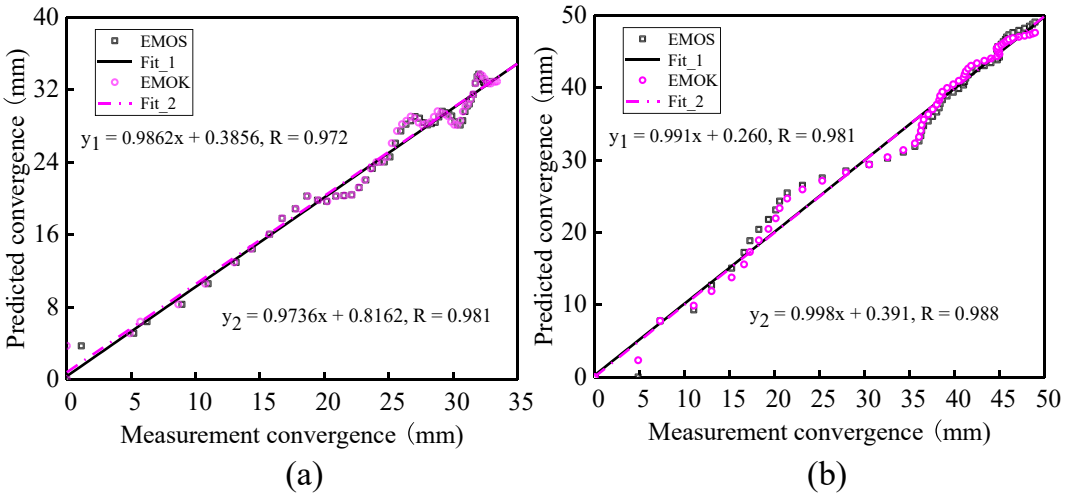

Fig. 8. Regression plots of original and predicted convergence (a) No.1 data set (b) No.2 data set.

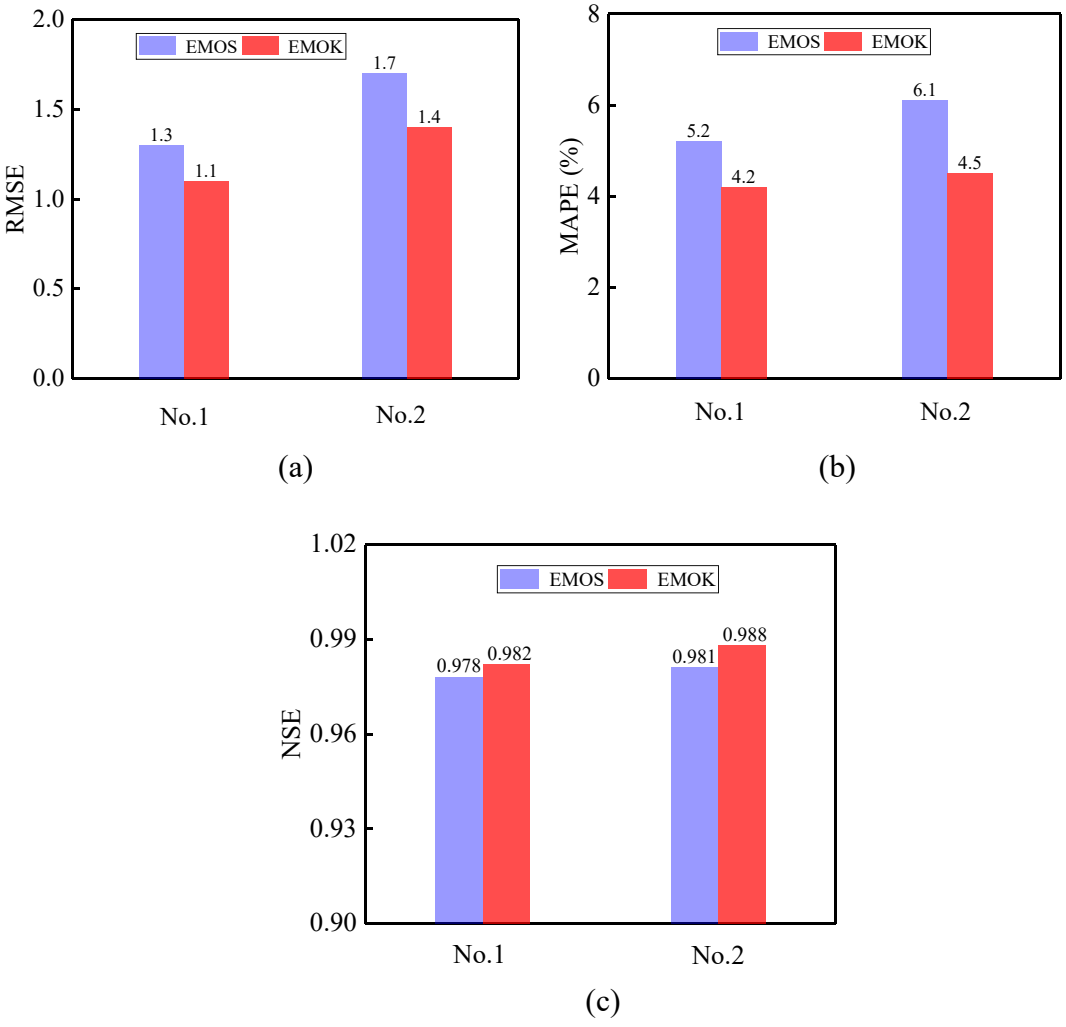

Fig. 9. RMSE, MAPE, and NSE of prediction results of two EMs for two data sets (a) RMSE (b) MAPE (c) NSE.

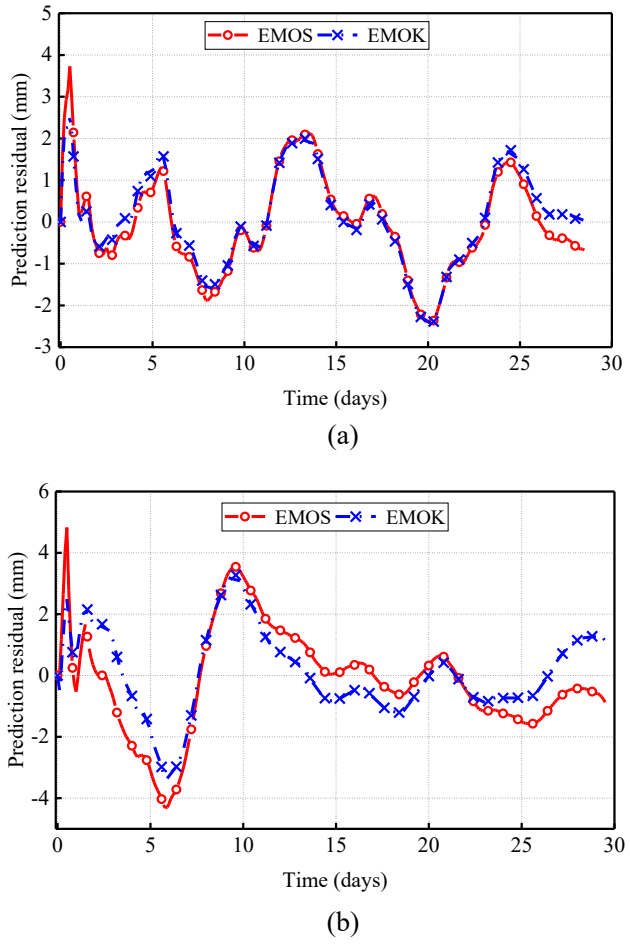

Fig. 10. EMs predicted residuals (a) No.1 data set (b) No.2 data set.

by minimizing the objective function $f(\boldsymbol{\theta})$, which is drawn from the MCMC simulation and can be expressed as follows:

$$\boldsymbol{\theta}^* = argminf(\boldsymbol{\theta}), \quad \boldsymbol{\theta} \in \varphi \tag{9}$$

where ϕ represents the prior knowledge of EMOS parameters; and $f(\boldsymbol{\theta})$ is the root-mean-squared error (RMSE) of prediction results to evaluate the performance of EMOS.

According to the law of large numbers, when an event is repeated in large numbers, the final frequency is infinitely close to the event probability [41]. Thus, the model parameters in $\boldsymbol{\theta}$ can be approximated by the posterior distribution of $\boldsymbol{\theta}^*$ with mean and variance. The EMOS then is developed by the mean values of the posterior distribution of $\boldsymbol{\theta}^*$, and the confidence interval (CI) of the EMOS is evaluated by their variance values. EM with the Bayesian estimation method is summarized in Algorithm 1.

Algorithm 1. EM with the Bayesian estimation

Step I: Initialize the first sample $\boldsymbol{\theta}_0$ that draws from the prior information as the first state of the Markov chain.
Step II: Generate samples of $\boldsymbol{\theta}$ using the MCMC simulation. The next sample in the Markov chain only relies on the current sample.
Step III: Estimate the mean, variance, and the probability distribution of $\boldsymbol{\theta}$. Calculate the validation error.
Step IV: Update the optimal posterior distribution of $\boldsymbol{\theta}^*$ in terms of the RMSE of prediction results.
Step V: Repeat steps II, III, and IV until the maximum number of iterations is reached.

2.3. RVM for residual learning model

Although prediction results of EMs are obtained, significant residual sequences $E = \{e_i\},\ i = 1, 2, \cdots, n$ may exist in the prediction results of the EM. RVM is a sparse probability model proposed by Tipping within a Bayesian framework [28]. It is used here for correcting the prediction results of the EM.

2.3.1. RVM formulation

Training samples are obtained by rolling forecast (the length of the observation window is p). The first 70% of data is selected as the training data, and the remaining 30% of data is used for testing the residual learning model. The training samples are set as $\{\boldsymbol{l}_k, h_k\}_{k=1}^{N}$ with N sets of training samples, where $\boldsymbol{l}_k = [e_{k+1}, e_{k+2}, \cdots, e_{k+p}]$ is the input vector and $h_k = e_{k+p+1}$ is the output value. Assume that the output value is function $f(\boldsymbol{l}_k)$ with additional noise ε_k [28]:

$$h_k = f(\boldsymbol{l}_k) + \varepsilon_k \tag{10}$$

where $\varepsilon_k \sim N(0, \sigma^2)$ is random noise. The common decision function of the RVM model is [24]:

$$f(\boldsymbol{l}_k) = \sum_{k=1}^{N} w_k \kappa(\boldsymbol{l}, \boldsymbol{l}_k) + w_0 \tag{11}$$

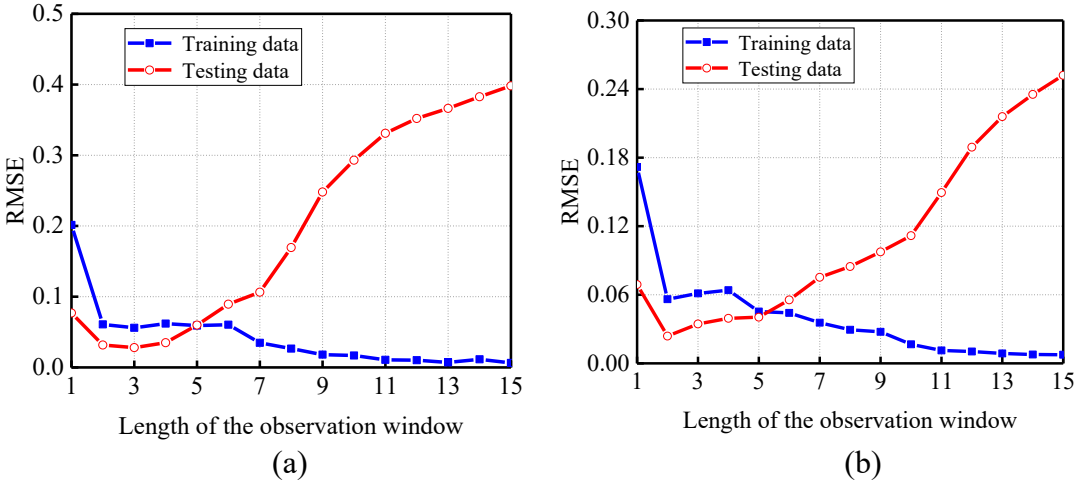

Fig. 11. Comparison of the different lengths of the observation window for the No.1 data set (a) EMOS-RVM (b) EMOK-RVM.

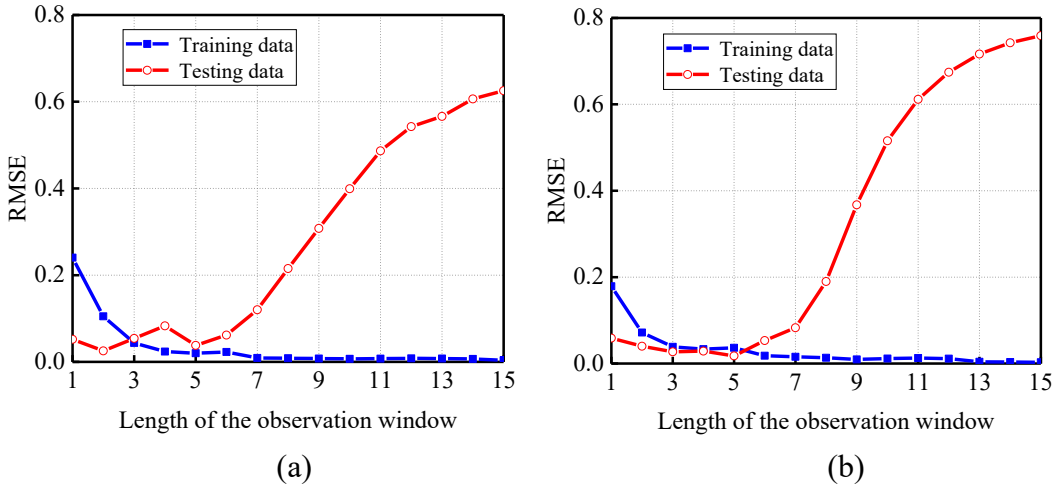

Fig. 12. Comparison of the different lengths of the observation window for the No.2 data set (a) EMOS-RVM (b) EMOK-RVM.

Table 4
Optimal solutions of hyperparameters of EMOS-RVM.

Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value
α_1	0.0744	α_{23}	0.0002	α_{115}	0.0043	α_{197}	0.0019
α_2	0.1369	α_{26}	0.0002	α_{118}	0.0062	α_{202}	1.7650
α_3	0.3618	α_{31}	0.0015	α_{124}	0.0005	α_{207}	0.0397
α_4	0.8292	α_{53}	0.0312	α_{129}	0.0004	α_{208}	0.0011
α_5	2.4139	α_{54}	0.0061	α_{131}	0.0023	α_{217}	0.0186
α_6	74.3832	α_{57}	0.0108	α_{140}	0.0353	α_{228}	0.0007
α_7	69.6772	α_{58}	0.0058	α_{143}	0.0092	α_{230}	0.0004
α_8	7.3082	α_{76}	0.0008	α_{165}	0.0012	σ^2	240.4619
α_9	0.0256	α_{85}	0.0004	α_{172}	0.0003	–	–
α_{13}	0.0012	α_{95}	0.0001	α_{183}	0.0030	–	–
α_{17}	0.0036	α_{112}	0.0006	α_{186}	0.0014	–	–

where $\kappa(\boldsymbol{l}, \boldsymbol{l}_k)$ represents the kernel function of RVM; $\boldsymbol{l} = [\boldsymbol{l}_1, \boldsymbol{l}_2, \cdots, \boldsymbol{l}_N]$ is the input of RVM; w_k is the weight of the kernel function; and w_0 is the bias.

Assume h_k as an independent random variable, and the likelihood function of the training samples is expressed as [30]:

$$p(\boldsymbol{h}|\boldsymbol{w}, \sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} exp\left\{ -\frac{1}{2\sigma^2}\|\boldsymbol{h} - \Phi\boldsymbol{w}\|^2 \right\} \tag{12}$$

where $\boldsymbol{h} = [h_1, \cdots, h_N]^T$ is the target values; $\boldsymbol{w} = [w_1, \cdots, w_N]^T$ is the weight vector; $\Phi = [\boldsymbol{\phi}(\boldsymbol{l}_1), \boldsymbol{\phi}(\boldsymbol{l}_2), \cdots, \boldsymbol{\phi}(\boldsymbol{l}_N)]^T$ represents the "design" matrix whose columns comprise the complete N sets of "basis vectors"; and $\boldsymbol{\phi}(\boldsymbol{l}_k) = [1, \kappa(\boldsymbol{l}_k, \boldsymbol{l}_1), \kappa(\boldsymbol{l}_k, \boldsymbol{l}_2), \cdots, \kappa(\boldsymbol{l}_k, \boldsymbol{l}_N)]$.

Of note is the development of independent hyperparameters vector $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \cdots, \alpha_N)$, and each respectively controls the prior about its associated weight. The prior distribution of $\boldsymbol{w}$ is defined as [24]:

$$p(\boldsymbol{w}|\boldsymbol{\alpha}) = \prod_{k=0}^{N} N\left(w_k|0, \alpha_k^{-1}\right) \tag{13}$$

According to the Bayesian theory, assume that σ^2 is subject to the normal distribution. The predictive distribution with the corresponding target h_* can be represented as [28]:

$$p(h_*|\boldsymbol{h}) = \int p\left(h_*|\boldsymbol{w}, \boldsymbol{\alpha}, \sigma^2\right) p\left(\boldsymbol{w}, \boldsymbol{\alpha}, \sigma^2|\boldsymbol{h}\right) d\boldsymbol{w} d\boldsymbol{\alpha} d\sigma^2 \tag{14}$$

The posterior parameter distribution based on the training samples is obtained by Eqs. (13) and (14) within the Bayesian framework [31]:

Table 5
Optimal solutions of hyperparameters of EMOK-RVM.

Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value
α_1	0.1539	α_{23}	0.0003	α_{115}	0.0198	α_{197}	0.0017
α_2	0.0420	α_{26}	0.0004	α_{118}	0.0030	α_{202}	0.0261
α_3	0.2126	α_{31}	0.6966	α_{124}	0.0023	α_{207}	0.0452
α_4	2.1266	α_{53}	0.0018	α_{129}	0.0006	α_{208}	0.0071
α_5	5.2247	α_{54}	0.0013	α_{131}	0.0001	α_{217}	0.0015
α_6	38.8243	α_{57}	0.0010	α_{140}	0.0006	α_{228}	0.0012
α_7	0.0007	α_{58}	0.0006	α_{143}	0.0016	α_{230}	0.0027
α_8	0.0182	α_{76}	0.0096	α_{165}	0.0330	σ^2	213.2945

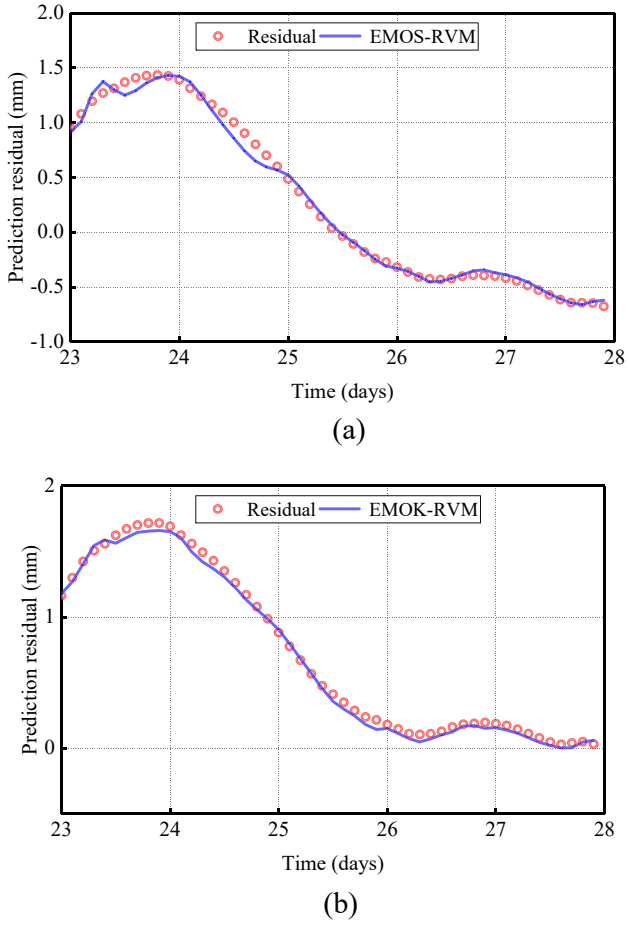

Fig. 13. Residual prediction results of two models for the No.1 data set (a) EMOS-RVM (b) EMOK-RVM.

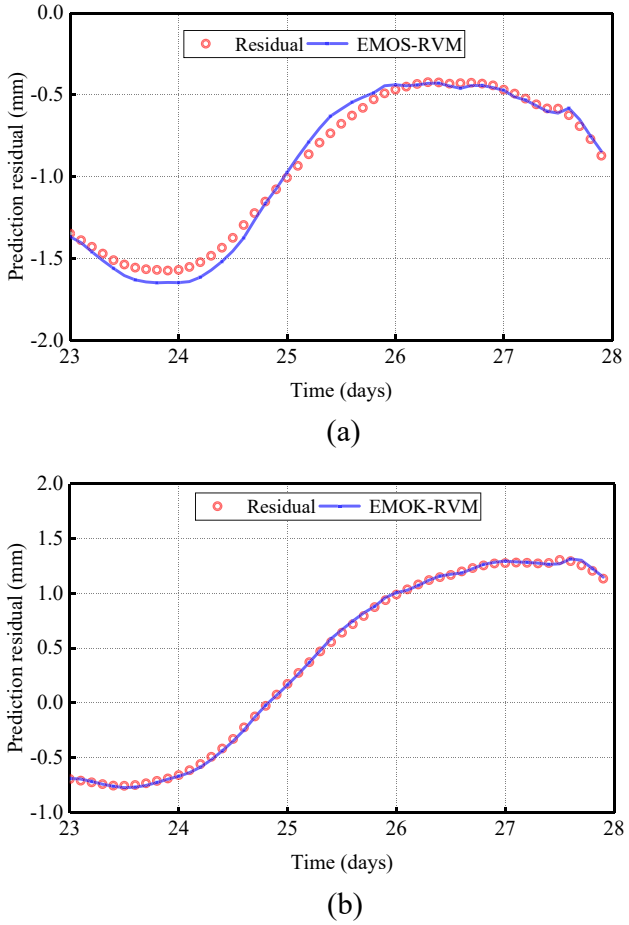

Fig. 14. Residual prediction results of two models for the No.2 data set (a) EMOS-RVM (b) EMOK-RVM.

$$p(\boldsymbol{w}|\boldsymbol{h},\boldsymbol{\alpha},\sigma^2) = \frac{p(\boldsymbol{h}|\boldsymbol{w},\sigma^2)p(\boldsymbol{w}|\boldsymbol{\alpha})}{p(\boldsymbol{h}|\boldsymbol{\alpha},\sigma^2)} = (2\pi)^{-\frac{N+1}{2}}\left|\sum\right|^{-\frac{1}{2}} exp\left(-\frac{1}{2}(\boldsymbol{w}-\boldsymbol{\mu})^T\sum^{-1}(\boldsymbol{w}-\boldsymbol{\mu})\right) \tag{15}$$

where the covariance matrix $\sum$ is $(\boldsymbol{B}+\sigma^{-2}\Phi^T\Phi)^{-1}$; the mean matrix $\boldsymbol{\mu}$ is $\sigma^{-2}\sum\Phi^T\boldsymbol{h}$; and $\boldsymbol{B}$ is defined as diag $(\alpha_1,\cdots,\alpha_N)$.

The hyperparameter α_k can be computed by maximizing the term $p(\boldsymbol{h}|\boldsymbol{\alpha},\sigma^2)$, then [29]:

$$p(\boldsymbol{h}|\boldsymbol{\alpha},\sigma^2) = (2\pi)^{-\frac{N}{2}}\left|\sigma^2\mathbf{I}+\Phi\boldsymbol{B}^{-1}\Phi^T\right|^{-\frac{1}{2}} exp\left(-\frac{1}{2}\boldsymbol{h}^T(\sigma^2\mathbf{I}+\Phi\boldsymbol{B}^{-1}\Phi^T)^{-1}\boldsymbol{h}\right) \tag{16}$$

Partial differentiation Eq. (16) of $\boldsymbol{\alpha}$ and σ^2, equating to 0 respectively [24]. Then the updated hyperparameter $\boldsymbol{\alpha}$ and variance σ^2 can be obtained:

$$\begin{cases} \alpha_k^{new} = \dfrac{\gamma_k}{\mu_k^2} \\ (\sigma^2)^{new} = \dfrac{\|\boldsymbol{h}-\Phi\boldsymbol{\mu}\|^2}{N-\sum_{i=0}^{N}\gamma_k} \end{cases} \tag{17}$$

where $\gamma_k = 1-\alpha_k\sum_{kk}$ and $\sum_{kk}$ is the k-th diagonal element of the posterior weight covariance. The learning algorithm updates the posterior mean $\boldsymbol{\mu}$ and the covariance matrix $\sum$ by calculating the α_k^{new} and $(\sigma^2)^{new}$ until the maximum number of iterations or the convergence criterion is satisfied [28].

Since the deformation of the surrounding rock of the tunnel is nonlinear, the RBF kernel function suitable for high order nonlinear problem is adopted as follows [25]:

$$\kappa(\boldsymbol{l},\boldsymbol{l}_k) = exp\left(-\frac{\|\boldsymbol{l}-\boldsymbol{l}_k\|^2}{u^2}\right) \tag{18}$$

where u is the kernel parameter.

2.3.2. Residual learning model

The time series residuals E are divided into training and testing data sets. The learning steps of the RVM residual learning model are displayed in Algorithm 2.

Algorithm 2. The RVM residual learning model.

Step I: Initialize the RVM model parameters, including the variance σ^2, the maximum iterations number, convergence criterion, and hyperparameter α_i.
Step II: Calculate the posterior distribution mean $\boldsymbol{\mu}$ and covariance $\sum$ of weight.
Step III: Update σ^2 and α_i based on Eq. (17).
Step IV: Repeat steps II and III until the maximum number of iterations or the convergence criterion is reached.
Step V: Use the trained model to predict the testing samples to verify the model performance.

2.4. Evaluation criteria

Many measures of forecast accuracy have been proposed in the past [42]. They can be divided into three categories: (1) scale-dependent measures; (2) percentage error-based measures; and (3) correlation-based measures.

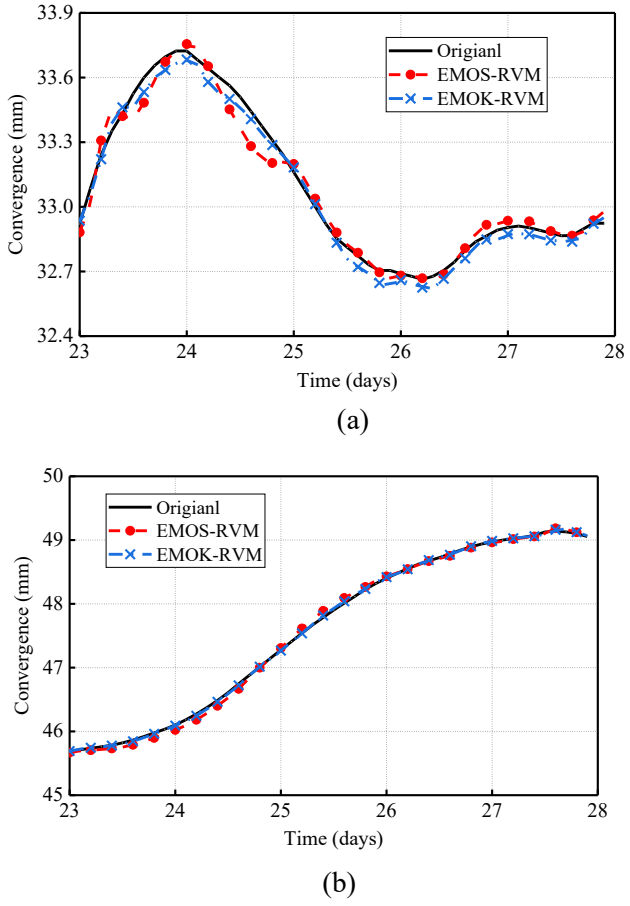

Fig. 15. Prediction results of EMOS-RVM and EMOK-RVM for two data (a) No.1 data set (b) No.2 data set.

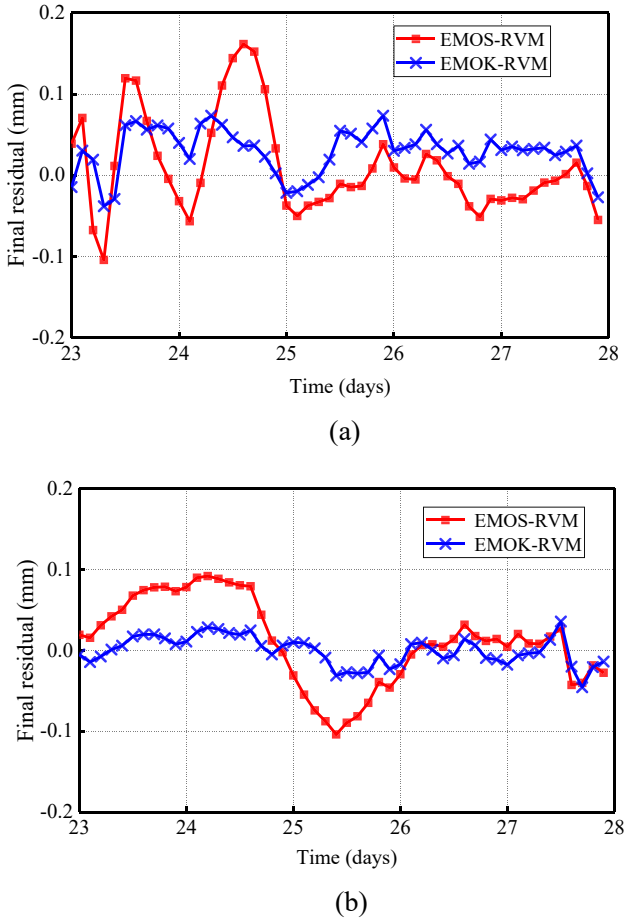

Fig. 16. Final prediction residuals of EMOS-RVM and EMOK-RVM for two data sets (a) No.1 data set (b) No.2 data set.

Firstly, for the scale-dependent measures, the commonly used scale-dependent measures are based on the absolute error or squared errors, such as mean square error (MSE) and RMSE. The scale of these measures depends on the scale of the data. They are useful when comparing different methods applied to the same set of data. Often, the RMSE is preferred to the MSE as it is on the same scale as the data [43].

Secondly, percentage errors have the advantage of being scale-independent, and thus can be used to compare forecast performance across different data sets. The commonly used measures include mean absolute percentage error (MAPE), median absolute percentage error, and root mean square percentage error. However, these measures put a heavier penalty on positive errors than on negative errors [44].

Thirdly, correlation-based measures are adopted to measure the correlation between the predicted and measured values. The commonly used measures are Pearson correlation coefficient, Nash–Sutcliffe efficiency (NSE), and Spearman rank correlation coefficient [45]. For instance, NSE (large values better) theoretically varies on the range $[-\infty, 1]$. It is easily understood that NSE values less than zero are undesirable since NSE $= 0$ implies that the prediction results give MSE performance no better than the mean of the observed system output.

In this study, RMSE, MAPE, and NSE are used to evaluate prediction performance. They are described as follows:

$$RMSE = \sqrt{\frac{1}{r}\sum_{i=1}^{r}(y_i - \hat{y})^2} \tag{19}$$

$$MAPE = \frac{100\%}{r}\sum_{i=1}^{r}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{20}$$

$$NSE = 1 - \frac{\sum_{i=1}^{r}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{r}(y_i - \bar{y}_i)} \tag{21}$$

where y_i and $\hat{y}_i$ stand for the actual and estimated values, respectively; $\bar{y}_i$ is the mean value of the actual values; and r is the number of the observations.

2.5. Probabilistic prediction model of tunnel convergence

A combined probabilistic model based on the EM, Bayesian estimation, and RVM is presented to predict tunnel convergence. The EM is applied to predict the tunnel convergence and produce residuals of the EM. The prediction residuals of EM are modified by the RVM model. The framework of the presented model consists of six steps as follows (Fig. 1):

Step I: Determine the prior information of the EM parameters in $\boldsymbol{\theta}$. Meanwhile, initialize the first sample of the Markov chain.

Step II: Generate samples of $\boldsymbol{\theta}$ using the MCMC simulation. Estimate and update the optimal posterior distribution of $\boldsymbol{\theta}^*$ in terms of the RMSE value of prediction results of the EM.

Step III: Establish the EM with minimum RMSE value to predict the tunnel convergence; and calculate the predicted residuals E according to the prediction results $\widehat{C}$ of the EM.

Step IV: Obtain the training and testing samples based on the predicted residuals E, and initialize the RVM model parameters.

Step V: Train the RVM model with the training samples and update RVM parameters σ^2 and α_i based on Eq. (17).

Step VI: Predict the testing samples to verify the model performance

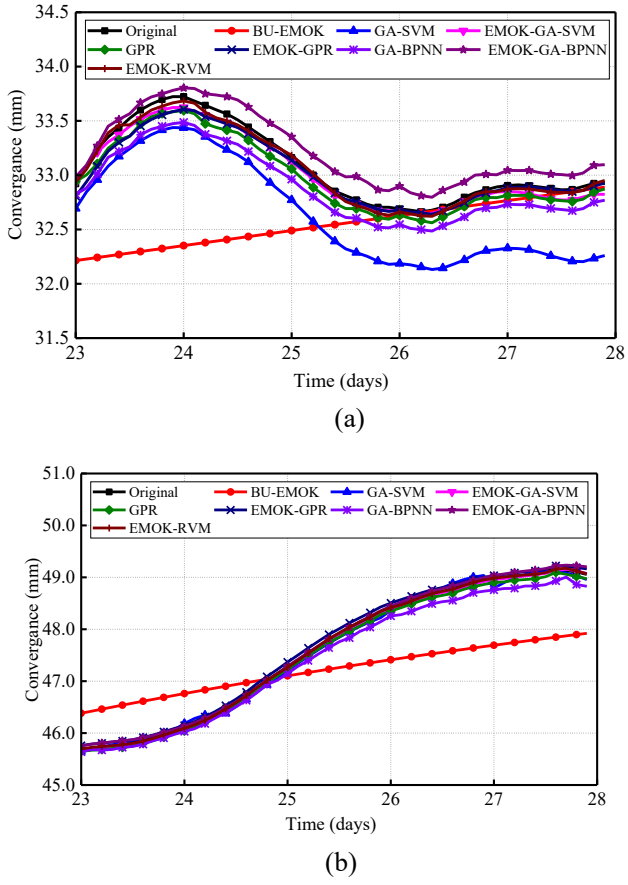

Fig. 17. Prediction results of convergence by eight models (a) No.1 data set (b) No.2 data set.

of RVM and obtain the residual prediction results $\widehat{E}$; and the prediction results $\widehat{C}$ of the EM are modified by residual prediction results $\widehat{E}$ as the final prediction results.

3. Case study

3.1. Project description and monitoring data

The Yangshan high-speed railway tunnel with a length of 850 m and a maximum overburden depth of 120 m is employed to illustrate the proposed approach. The excavation of this tunnel started in 2019 using NATM. The tunnel is supported by radial rock bolts and concrete lining, and the shallow section is approximately 100 m with maximum overburden of 16 m. The rock mass classification of the shallow section is V grade with broken and poor stability. The convergence is the distance between two points of the tunnel wall at the same height, which directly reflects the tunnel deformation. The typical cross-section is shown in Fig. 2, in which No.1 and No.2 represent the convergence measuring points.

To ensure construction safety and quality, tunnel convergences were recorded by an automatic monitoring system over 30 days. Without loss of generality, two measuring points from the shallow tunnel section are chosen for analysis. Fig. 3 shows their time-dependent curves and face advance sequence. It is observed that the time-dependent curves of the tunnel convergence increase and fluctuate over time after the excavation and the advance of the tunnel face.

3.2. Tunnel convergence EM

Several EMs [15,16] have been proposed to predict time-dependent tunnel convergence. It was reported that the EMs appropriately described the relationship between rock mass stress redistribution and tunnel convergence [17]. The EM parameters directly affect the prediction performance. Therefore, Bayesian estimation is employed to analyze the convergence data from the Yangshan tunnel.

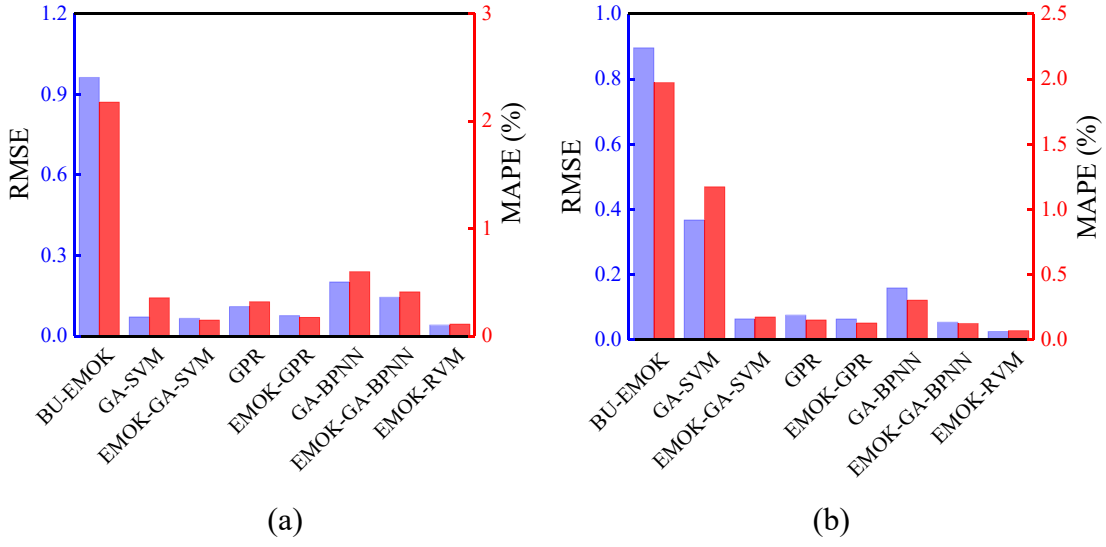

Fig. 18. Comparison results of prediction convergence in terms of RMSE and MAPE by eight models (a) No.1 data set (b) No.2 data set.

Table 6
Comparison results of prediction convergence in terms of NSE by eight models.

model		BU-EMOK	GA-SVM	EMOK-GA-SVM	GPR	EMOK-GPR	GA-BPNN	EMOK-GA-BPNN	EMOK-RVM
data	No.1	−6.6705	−0.8099	0.9660	0.901	0.9524	0.666	0.8287	0.9865
	No.2	0.4881	0.9968	0.9989	0.9963	0.9974	0.9838	0.9981	0.9998

Model parameters of the EMs in Eqs. (3) and (5) are listed in Tables 1 and 2. MCMC simulation is performed to generate 50,000 samples of model parameters in $\boldsymbol{\theta}$ with the prior information in Tables 1 and 2. Conventional statistical analyses are performed on the samples to determine the statistic and posterior distribution of $\boldsymbol{\theta}$ [20]. Also, to obtain the posterior distribution of $\boldsymbol{\theta}$ with the minimum RMSE, MCMC simulation is performed 100 times.

For brevity, only the PDFs and curve fittings of EM parameters for the No.1 data set are displayed in Figs. 4 and 5. It is found that the normal distribution is the good-fitting model for EM parameters. The good agreement shows that the presented approach provides reasonable estimations of two EMs. Thus, the PDFs of EM parameters can be used as a reference for convergence prediction of adjacent monitoring sections.

To quantify simulation capability, Table 3 summarizes the estimation of the mean and variance of the two EMs parameters. It is observed that estimations of the X and T have slight change for the same EM, which is consistent with the conclusion presented by Sulem et al. [15]. Moreover, the variances of the EMOK's parameter are much smaller than the EMOS's variances, and thus the EMOK with the MCMC simulation has better convergence.

The prediction results of the two EMs with minimum RMSE are calculated and presented in Figs. 6 and 7. The Bayesian estimation has an advantage of performing probabilistic predictions from the PDFs of the model parameters. CI refers to the estimation interval of overall parameters based on simulated samples, which is calculated at predetermined levels of significance [46]. The 95% CI of prediction results is also shown in Figs. 6 and 7.

As shown in Figs. 6 and 7, for two data sets, the prediction results exhibit a similar trend compared to the measured values, indicating that the presented approach has good prediction performance. The narrow 95% CI of the EMOK's forecasting results reveals its higher reliable performance. Fig. 8 shows regression plots of the original data set versus the forecasted one. The prediction residuals mainly exist in large values for the No.1 data set and the median value for the No.2 data set. By combining this with the convergence data shown in Fig. 3, it is easily observed that the prediction residuals mainly occur at the fluctuation of measured data, which illustrates EMs perform poorly when predicting the fluctuating trend.

The RMSE, MAPE, and NSE of two EMs are calculated and shown in Fig. 9. It is found that the RMSE and MAPE values of EMOK are smaller than those of EMOS for the two data sets. It is also shown that the NSE values of EMOK are higher than those of EMOS for the two data sets. This can be explained by the fact that local monitoring errors can influence the forecasting curve of EMOS in any direction [15]. As a result, the EMOK owns higher prediction accuracy compared with the EMOS.

Fig. 10 illustrates the predicted residuals for two EMs. It is observed that the prediction results have obvious residuals, indicating that the EMs cannot fully adapt to the convergence prediction in a complex construction environment. According to the above analysis results, the EMOK with the MCMC simulation exhibits good accuracy for tunnel convergence prediction. However, significant residuals between the original and predicted convergences are found and may affect the assessment accuracy of the tunnel status. In the next section, the RVM residual learning model will be discussed.

3.3. Residual learning model

RVM residual learning model is presented in a fully probabilistic framework based on the Bayesian theory. RVM includes a prior over the model weights governed by a set of hyperparameters, whose most probable values are iteratively estimated in the learning process [28]. The rolling forecast is applied to obtain training and testing data. According to the original data characteristics, the size of the observation window increases from 1 to 15. Based on Algorithm 2, EMOS-RVM and EMOK-RVM residual learning models are established based on the prediction results of the two EMs. Model performances of the observation window with different lengths are presented in Figs. 11 and 12.

As shown in Fig. 11, for the training data, the RMSE results for both EMOS-RVM and EMOK-RVM become smaller as the length of the observation window increases. For the testing data, the minimum RMSE values for these two models are 0.0598 and 0.0404, respectively. Thus, the EMOK-RVM shows higher accuracy in comparison with the EMOS-RVM. The length of the observation window corresponding to the minimum RMSE is 5, which can be used in the convergence prediction of adjacent monitoring sections of the Yangshan high-speed railway tunnel. Tables 4 and 5 show the optimal solutions of two hyperparameters $\boldsymbol{\alpha}$ and σ^2 (all other $\alpha_i = 0$), which can be estimated according to Eq. (17). Similar results can be found for the No.2 data set. Figs. 13 and 14 show the residual prediction results of EMOS-RVM and EMOK-RVM for testing data.

3.4. Prediction results and comparison

The forecasting results of EMOS-RVM and EMOK-RVM are calculated and presented in Fig. 15. Compared with the EMs, the prediction accuracy is significantly improved after the residual correction. More specifically, for the two data sets, the forecasting results show a similar trend with the measurements, particularly in predicting the fluctuating trend. It can be concluded that RVM can adapt to different EMs and the RVM residual learning model exhibits satisfactory prediction performance.

The predicted residuals of the EMOS-RVM and EMOK-RVM are shown in Fig. 16 to compare the model performance. It is found that the residuals of prediction results are significantly reduced in the two data points, which indicates that EMOS-RVM and EMOK-RVM have satisfactory prediction accuracy. This can be explained by the fact that the relationship of convergence monitoring data is highly nonlinear and thus could not be modeled accurately by the EMs. It is also seen from Fig. 16 that the final residuals of EMOK-RVM are overall smaller than those of EMOS-RVM. It can be concluded that the EMOK-RVM model exhibits superiority in predicting convergence over the EMOS-RVM. Compared with EMOK, the RMSE values of EMOK-RVM are reduced by 92.6% and 95.8% for two data sets respectively.

To verify the performance of the EMOK-RVM model, the following models will be used for the comparative purpose: (1) an EMOK with Bayesian updating (BU-EMOK); (2) an SVM with genetic algorithm (GA) optimization; (3) an EMOK-SVM with the MCMC simulation; (4) a GPR; (5) an EMOK-GPR with the MCMC simulation; (6) a GA-BPNN; and (7) an EMOK-GA-BPNN with the MCMC simulation. BU-EMOK considers various sources of uncertainties such as model uncertainty and model parameters uncertainty, which is used to predict tunnel convergence [13]. SVM is a popular machine learning technique for regression analysis, which is used to predict tunnel convergence [6]. GA is used to optimize the SVM's hyperparameters. GPR is a nonparametric probabilistic modeling approach to generalize nonlinear and complex functions in a dataset [43]. Hence, the approach has been used in various studies for solving different problems in tunnel projects [25]. BPNN is a forward feedback network, which is also the commonly used neural network and is especially used in tunnel displacement prediction [26]. All of these methods use the same learning settings as proposed in this paper. The prediction results for these methods are provided in Fig. 17. To quantify prediction capability, RMSE and MAPE are calculated and shown in Fig. 18, and NSE is also calculated and shown in Table 6.

As shown in Figs. 17 and 18, BU-EMOK has the poorest prediction results with the highest RMSE and MAPE. Also, it is observed that the NSE value of BU-EMOK is the lowest in Table 6. This can be explained by the fact that the highly nonlinear relationship of tunnel convergence could not be modeled accurately by the EMOK. It is easily observed that both EMOK-GA-SVM and EMOK-RVM have a better performance than GA-SVM. This is because SVM is inherently limited by Mercer's condition and the difficulty of the hyperparameter selection. It is also seen that the GPR and GA-BPNN show higher RMSE and MAPE values and lower NSE values than EMOK-RVM, which indicates that these two

models have lower prediction accuracies than EMOK-RVM. This proves that, for complex nonlinear problems, GPR and BPNN usually require more modeling samples to improve the prediction accuracy [32,47]. According to Figs. 17 and 18, compared with GPR, EMOK-GPR has a significant improvement on the prediction accuracy, which indicates the combined model based on EMOK and GPR has higher prediction precision. Similar comparison results can be found when GA-BPNN and EMOK-GA-BPNN are used to predict tunnel convergence. Moreover, it is easily observed that EMOK-RVM has a better performance than the other seven models for both two data sets. This is because RVM proposes a sparse probabilistic framework and exhibits superior generalization capacities for small sample data sets. In summary, EMOK-RVM is suitable for modeling the nonlinear relationship of tunnel convergence.

4. Conclusion

In this paper, a combined probabilistic model based on the empirical model (EM), Bayesian estimation, and relevance vector machine (RVM) has been developed to predict tunnel convergence. Monitoring data from the automatic monitoring system of a high-speed railway tunnel for one month is employed to demonstrate the feasibility of the presented approach. The following conclusions are obtained:

(1) Two EMs with Bayesian estimation are established to predict the tunnel convergence by considering the tunnel face advance and the time-dependent effect. Specifically, the parameters of the EMs are simulated by the Markov Chain Monte Carlo (MCMC) simulation. The probability density functions of these two model parameters show a good convergence and can be used as a reference for convergence prediction of adjacent monitoring sections. The EM of Kontogianni (EMOK) offers more accurate prediction results and lower uncertainty compared with the EM of Sumel (EMOS). Nevertheless, significant prediction residuals exist in predicting the fluctuating trend of monitoring data.
(2) The prediction residuals of the EMs are corrected by the RVM residual learning model. After selecting the appropriate observation window, the relationship among prediction residuals is properly developed by RVM and the prediction accuracy is significantly improved. The EMOK-RVM exhibits better performance compared with the EMOS-RVM.
(3) To demonstrate the performance of the EMOK-RVM, it is compared with EMOK with Bayesian updating (BU-EMOK), SVM with genetic algorithm (GA) optimization, EMOK-GA-SVM, Gaussian process regression (GPR), EMOK-GPR, back-propagation neural network (BPNN) with genetic algorithm optimization, and EMOK-GA-BPNN. The EMOK-RVM shows higher prediction accuracy than these seven prediction models. This is because the EMOK-RVM proposes a sparse probabilistic framework and exhibits superior generalization capacities for small sample data sets.

CRediT authorship contribution statement

Xiangyu Chang: Conceptualization, Methodology, Software, Validation, Writing – original draft. **Hao Wang:** Supervision, Funding acquisition, Investigation, Writing – review & editing. **Yiming Zhang:** Methodology, Supervision, Writing – review & editing. **Feiqiu Wang:** Formal analysis. **Zhaozhong Li:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to gratefully acknowledge the supports from the National Natural Science Foundation of China (Grant No. 51978155), the Jiangsu Provincial Key Research and Development Program (Grant No. BE2018120), the National Ten Thousand Talent Program for Young Top-notch Talents (Grant No. W03070080), and the Research and Development Program of Ministry of Housing and Urban-Rural Development (Grant No. K20200249).

References

[1] N.I. Ivanov, I.S. Boiko, A.E. Shashurin, The problem of high-speed railway noise prediction and reduction, Procedia Eng.. 189 (2017) 539–546, https://doi.org/10.1016/j.proeng.2017.05.086.
[2] Y.M. Zhang, H. Wang, J.X. Mao, F.Q. Wang, S.T. Hu, X.X. Zhao, Monitoring-based assessment of the construction influence of benoto pile on adjacent high-speed railway bridge: case study, J. Perform. Constr. Facil 33 (1) (2019) 04018106, https://doi.org/10.1061/(ASCE)CF.1943-5509.0001258.
[3] S. Tian, J. Gong, Statistics of railway tunnels in China as of end of 2019, Tunnel Construct. 40 (2020) 292.
[4] S. Nsubuga, M. Tsakiri, V. Georgiannou, A smart decision tool for the prediction of tunnel crown displacements, Appl. Geomatics 13 (S1) (2021) 77–91, https://doi.org/10.1007/s12518-020-00304-9.
[5] H. Zarei, K. Ahangari, M. Ghaemi, A. Khalili, A convergence criterion for water conveyance tunnels, Innovative Infrastruct. Solut. 2 (1) (2017), https://doi.org/10.1007/s41062-017-0098-z.
[6] S. Mahdevari, H. Shirzad Haghighat, S.R. Torabi, A dynamically approach based on SVM algorithm for prediction of tunnel convergence during excavation, Tunn. Undergr. Space Technol. 38 (2013) 59–68, https://doi.org/10.1016/j.tust.2013.05.002.
[7] D. Debernardi, G. Barla, New viscoplastic model for design analysis of tunnels in squeezing conditions, Rock Mech. Rock Eng. 42 (2) (2009) 259–288, https://doi.org/10.1007/s00603-009-0174-6.
[8] S. Nadimi, K. Shahriar, M. Sharifzadeh, P. Moarefvand, Triaxial creep tests and back analysis of time-dependent behavior of Siah Bisheh cavern by 3-dimensional distinct element method, Tunn. Undergr. Space Technol. 26 (1) (2011) 155–162, https://doi.org/10.1016/j.tust.2010.09.002.
[9] M. Sharifzadeh, A. Tarifard, M.A. Moridi, Time-dependent behavior of tunnel lining in weak rock mass based on displacement back analysis method, Tunn. Undergr. Space Technol. 38 (2013) 348–356, https://doi.org/10.1016/j.tust.2013.07.014.
[10] D. Sterpi, G. Gioda, Visco-plastic behaviour around advancing tunnels in squeezing rock, Rock Mech. Rock Eng. 42 (2) (2007) 319–339, https://doi.org/10.1007/s00603-007-0137-8.
[11] A. Fahimifar, F.M. Tehrani, A. Hedayat, A. Vakilzadeh, Analytical solution for the excavation of circular tunnels in a visco-elastic Burger's material under hydrostatic stress field, Tunn. Undergr. Space Technol. 25 (4) (2010) 297–304, https://doi.org/10.1016/j.tust.2010.01.002.
[12] P. Nomikos, R. Rahmannejad, A. Sofianos, Supported axisymmetric tunnels within linear viscoelastic burgers rocks, Rock Mech. Rock Eng. 44 (5) (2011) 553–564, https://doi.org/10.1007/s00603-011-0159-0.
[13] X. Feng, R. Jimenez, P. Zeng, S. Senent, Prediction of time-dependent tunnel convergences using a Bayesian updating approach, Tunn. Undergr. Space Technol. 94 (2019), https://doi.org/10.1016/j.tust.2019. 103118.
[14] S. Sakurai, Approximate time-dependent analysis of tunnel support structure considering progress of tunnel face, Int. J. Numer. Anal. Meth. Geomech. 2 (2) (1978) 159–175, https://doi.org/10.1002/nag.1610020205.
[15] J. Sulem, M. Panet, A. Guenot, Closure analysis in deep tunnels, Int. J. Rock Mech. Min. Sci. Geomech. Abst. 24 (3) (1987) 145–154, https://doi.org/10.1016/0148-9062(87)90522-5.
[16] V. Kontogianni, P. Psimoulis, S. Stiros, What is the contribution of time-dependent deformation in tunnel convergence? Eng. Geol. 82 (4) (2006) 264–267, https://doi.org/10.1016/j.enggeo.2005.11.001.
[17] T.M. Vu, J. Sulem, D. Subrin, N. Monin, J. Lascols, Anisotropic closure in squeezing rocks: the example of saint-martin-la-porte access gallery, Rock Mech. Rock Eng. 46 (2) (2012) 231–246, https://doi.org/10.1007/s00603-012-0320-4.
[18] E. Asadollahpour, R. Rahmannejad, A. Asghari, A. Abdollahipour, Back analysis of closure parameters of Panet equation and Burger's model of Babolak water tunnel conveyance, Int. J. Rock Mech. Min. Sci. 68 (2014) 159–166, https://doi.org/10.1016/j.ijrmms.2014.02.017.
[19] H. Lu, E. Kim, M. Gutierrez, Monte carlo simulation (MCS)-based uncertainty analysis of rock mass quality Q in underground construction, Tunn. Undergr. Space Technol. 94 (2019) 103089, https://doi.org/10.1016/j.tust.2019.103089.
[20] A.E. Aladejare, Y. Wang, Estimation of rock mass deformation modulus using indirect information from multiple sources, Tunn. Undergr. Space Technol. 85 (2019) 76–83, https://doi.org/10.1016/j.tust.2018.11.047.
[21] Y.M. Zhang, H. Wang, H.P. Wan, J.X. Mao, Y.C. Xu, Anomaly detection of structural health monitoring data using the maximum likelihood estimation-based Bayesian dynamic linear model, Struct. Health Monitor. 20 (6) (2021) 2936–2952, https://doi.org/10.1177/1475921720977020.

[22] R. Jimenez, D. Recio, A linear classifier for probabilistic prediction of squeezing conditions in Himalayan tunnels, Eng. Geol. 121 (3–4) (2011) 101–109, https://doi.org/10.1016/j.enggeo.2011.05.006.
[23] Ö. Satici, T. Topal, Prediction of tunnel wall convergences for NATM tunnels which are excavated in weak-to-fair-quality rock masses using decision-tree technique and rock mass strength parameters, SN Appl. Sci. 2 (4) (2020), https://doi.org/10.1007/s42452-020-2311-5.
[24] J. Ding, M. Wang, Z. Ping, D. Fu, V.S. Vassiliadis, An integrated method based on relevance vector machine for short-term load forecasting, Eur. J. Oper. Res. 287 (2) (2020) 497–510, https://doi.org/10.1016/j.ejor.2020.04.007.
[25] T.K. Mehdi, B. Sarshari, Predicting convergence rate of namaklan twin tunnels using machine learning methods, Arab. J. Sci. Eng. 45 (5) (2019) 3761–3780, https://doi.org/10.1007/s13369-019-04239-1.
[26] J.B. Fei, Z.Z. Wu, X.H. Sun, D. Su, X.H. Bao, Research on tunnel engineering monitoring technology based on Bpnn neural network and mars machine learning regression algorithm, Neural Comput. Appl. 33 (1) (2020) 239–255, https://doi.org/10.1007/s00521-020-04988-3.
[27] F. Wang, H. Lu, B. Gou, X. Han, Q. Zhang, Y. Qin, Modeling of shield-ground interaction using an adaptive relevance vector machine, Appl. Math. Model. 40 (9–10) (2016) 5171–5182, https://doi.org/10.1016/j.apm.2015.09.016.
[28] M.E. Tipping, Sparse Bayesian learning and the relevance vector machine, J. Mach. Learn. Res. 1 (3) (2001) 211–244.
[29] L. Wei, Z. Qian, H. Zareipour, Wind turbine pitch system condition monitoring and fault detection based on optimized relevance vector machine regression, IEEE Trans. Sustain. Energy 11 (4) (2020) 2326–2336.
[30] J. Lou, Z. Shen, Q. Shen, W. Hu, Z. Chen, Improving traffic flow forecasting with relevance vector machine and a randomized controlled statistical testing, Soft Comput. 24 (8) (2018) 5485–5497, https://doi.org/10.1007/s00500-018-03693-7.
[31] X. Ma, Y. Xue, C. Bai, H. Liu, Y. Yu, Prediction model for deformation risk grade of the soft rock tunnel based on GRA – Extension, IOP Conf. Ser.: Earth Environ. Sci. 440 (5) (2020) 052057, https://doi.org/10.1088/1755-1315/440/5/052057.
[32] R. Guo, Z. Liu, Y. Wei, Remaining useful life prediction for the air turbine starter based on empirical mode decomposition and relevance vector machine, Trans. Inst. Meas. Control 42 (13) (2020) 2578–2588, https://doi.org/10.1177/0142331220932651.
[33] W. Li, W. Sun, Y. Zhao, Z. Yuan, Y. Liu, Deep image compression with residual learning, Appl. Sci. 10 (11) (2020), https://doi.org/10.3390/app 10114023.
[34] M.J. Kavvadas, Monitoring ground deformation in tunnelling: current practice in transportation tunnels, Eng. Geol. 79 (1–2) (2005) 93–113, https://doi.org/10.1016/j.enggeo.2004.10.011.
[35] H. Tran-Manh, J. Sulem, D. Subrin, Progressive degradation of rock properties and time-dependent behavior of deep tunnels, Acta Geotech. 11 (3) (2016) 693–711, https://doi.org/10.1177/0142331220932651.
[36] V.A. Kontogianni, S.C. Stiros, Predictions and observations of convergence in shallow tunnels: case histories in Greece, Eng. Geol. 63 (3–4) (2002) 333–345, https://doi.org/10.1016/j.enggeo.2005.11.001.
[37] X.Y. Li, L.M. Zhang, S.H. Jiang, Updating performance of high rock slopes by combining incremental time-series monitoring data and three-dimensional numerical analysis, Int. J. Rock Mech. Min. Sci. 83 (2016) 252–261, https://doi.org/10.1016/j.ijrmms.2014.09.011.
[38] J. Zhang, W.H. Tang, L.M. Zhang, H.W. Huang, Characterising geotechnical model uncertainty by hybrid Markov Chain Monte Carlo simulation, Comput. Geotech. 43 (2012) 26–36, https://doi.org/10.1016/j.comp geo.2012.02.002.
[39] C. Hsein Juang, Z. Luo, S. Atamturktur, H. Huang, Bayesian updating of soil parameters for braced excavations using field observations, J. Geotech. Geoenviron. Eng. 139 (3) (2013) 395–406.
[40] Y. Wang, T. Zhao, Z. Cao, Site-specific probability distribution of geotechnical properties, Comput. Geotech. 70 (2015) 159–168, https://doi.org/10.1016/j.compgeo.2015.08.002.
[41] D.M. Chibisov, Bernoulli's Law of large numbers and the strong law of large numbers, Theory Prob. Appl. 60 (2) (2016) 318–319, https://doi.org/10.1137/s0040585x97t987696.
[42] R.J. Hyndman, A.B. Koehler, Another look at measures of forecast accuracy, Int. J. Forecast. 22 (4) (2006) 679–688, https://doi.org/10.1016/j.ijforecast.2006.03.001.
[43] K. Tamer, A. Mohamed, K. Sopian, M. Mahmoud, Assessment of artificial neural networks for hourly solar radiation prediction, Int. J. Photoenergy 2012 (2012) 1–7, https://doi.org/10.1155/2012/946890.
[44] K. Tamer, A. Mohamed, M. Mahmoud, K. Sopian, A new approach for meteorological variables prediction at kuala lumpur, malaysia, using artificial neural networks: application for sizing and maintaining photovoltaic systems, J. Sol. Energy Eng. 134 (2) (2012), https://doi.org/10.1115/1.4005754.
[45] G.H. Vijai, H. Kling, On typical range, sensitivity, and normalization of mean squared rrror and nash-sutcliffe rfficiency type metrics, Water Resour. Res. 47 (10) (2011), https://doi.org/10.1029/2011wr010962.
[46] Y.M. Zhang, H. Wang, J.X. Mao, Z.D. Xu, Y.F. Zhang, Probabilistic framework with Bayesian optimization for predicting typhoon-induced dynamic responses of a long-span bridge, J. Struct. Eng. 147 (1) (2021) 04020297, https://doi.org/10.1061/(ASCE)ST.1943-541X.0002881.
[47] H.M. Lu, J.S. Chen, W.C. Liao, Nonparametric regression via variance-adjusted gradient boosting gaussian process regression, IEEE Trans. Knowl. Data Eng. 33 (6) (2021) 2669–2679, https://doi.org/10.1109/tkde.2019.2953728.