ENGINEERING Information Technology & Electronic Engineering
https://www.academax.com/EITEE; https://www.engineering.org.cn
https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=11344065
E-mail: eng.itee@zju.edu.cn

Available on IEEE Xplore®

Research Article

https://doi.org/10.1631/ENG.ITEE.2026.0005

High-precision temperature prediction for atmospheric refractivity correction using Kalman spatiotemporal data fusion

Ziru LI[1,2,3], Zhaobin XU[1,2,3]✉, Tao ZHANG[1,2,3], Xinbo YUAN[1,2,3], Zhonghe JIN[1,2,3]

[1]Micro-Satellite Research Center, Zhejiang University, Hangzhou 310027, China
[2]Huanjiang Laboratory, Zhuji 311899, China
[3]Key Laboratory of Micro-Nano Satellite Research Zhejiang Province, Hangzhou 310027, China

Abstract: In absolute distance measurement and positioning applications, atmospheric refraction error is a critical factor limiting measurement accuracy. Temperature plays a dominant role in computing the atmospheric refractive index. However, accurately acquiring the temperature field along the ranging path in complex and dynamic outdoor environments remains challenging due to limited sensor deployment and environmental nonstationarity. We propose a spatiotemporal temperature data fusion method for atmospheric refraction correction, which integrates the strengths of the generalized regression neural network (GRNN) and Kriging interpolation within a Kalman filter. This method achieves dynamic prediction and high-accuracy reconstruction of temperature parameters. The proposed method is systematically validated through simulation analysis as well as indoor and kilometer-scale outdoor experimental measurements. The simulation results demonstrate that Kalman filter expanded fusion (KFEF) outperforms the traditional interpolation method radial basis function (RBF) and the state-of-the-art spatiotemporal interpolation and prediction methods spatiotemporal Kriging (STK) and Gaussian process (GP), in terms of both reconstruction accuracy and stability of the temperature field. Specifically, KFEF achieves a 61.54% reduction in root mean square error (RMSE) compared with RBF and reductions of 34.21% and 32.43% relative to STK and GP, respectively. This indicates its practical value for long-distance high-precision ranging engineering applications. Furthermore, the proposed spatiotemporal data fusion framework is highly general and scalable. It can also be applied to other temperature field prediction and reconstruction problems.

Key words: Temperature prediction; Kalman filter expanded fusion (KFEF); Atmospheric refraction correction; Absolute distance measurement; Generalized regression neural network (GRNN) optimization

1 Introduction

The absolute distance measurement accuracy and positioning accuracy are largely affected by the atmospheric refraction error (Meiners-Hagen et al., 2017; Ding et al., 2020; Bi et al., 2024). When electromagnetic waves such as lasers and microwaves propagate in the atmosphere, their refractive indices are dynamic and change with temperature, air pressure, humidity, etc. This leads to deviations in distance measurement results. The core of atmospheric refraction error correction is the acquisition of precise atmospheric refractivity along the electromagnetic wave propagation path. Existing studies indicate that within commonly used atmospheric refractivity models, such as the Edlén formula (Rüeger, 2002) and the Ciddor formula (Ciddor, 1996), meteorological parameters involved in refractive index computation exert markedly different levels of influence. Among these parameters, temperature plays a decisive role in determining the refractivity accuracy (Pisani et al., 2018).

In recent years, researchers have proposed various methods for acquiring or eliminating temperature parameters with high precision. The main temperature parameter elimination methods are the dual-wavelength method and the acoustic method. The former is not applicable to the microwave band (Tomberg et al., 2017; Guillory et al., 2024), whereas the latter is susceptible to industrial noise and is characterized by high system complexity and cost (Underwood et al., 2015; Pisani et al., 2018). In terms of high-precision temperature parameter acquisition, multiple sensors are typically deployed along the baseline to reconstruct the temperature meteorological field. The Physikalisch-Technische Bundesanstalt (PTB) in Germany and the National Institute of Metrology in China installed 60 temperature sensors outdoors, effectively correcting the refractive error (Pollinger et al., 2012; Liu XD et al., 2020). However, this approach is difficult to apply in mobile outdoor environments due to the large number of sensors.

✉ Zhaobin XU, zjuxzb@zju.edu.cn
Ziru LI, https://orcid.org/0009-0007-9038-5651
Zhaobin XU, https://orcid.org/0000-0003-3059-8974
Zhonghe JIN, https://orcid.org/0000-0002-2039-1390

CLC number: TP274

Received: Jan. 5, 2026; Revision accepted: Apr. 7, 2026;
Crosschecked: Apr. 15, 2026

In outdoor absolute distance measurements, a limited number of sensors are usually used to collect temperature data. Interpolation methods are then employed to estimate the temperature distribution along the propagation path. Previous studies have suggested various methods, including the equivalent area method (Chen Y et al., 2018) and linear interpolation combined with the radial basis function (RBF) (Gu et al., 2019). However, the equivalent area method is prone to large errors due to linear interpolation. RBF is relatively sensitive to noise in the input data, and its model performance largely depends on the availability of a sufficient number of training samples (Wang LC et al., 2016; Ghritlahre and Prasad, 2018). Generalized regression neural network (GRNN) can maintain a low mean-squared error even under small-sample and noisy conditions (Ghritlahre and Prasad, 2018; Li G et al., 2018). However, GRNN neglects the inherent spatial correlation and heterogeneity of the temperature meteorological field (Chen KQ et al., 2024).

Unlike other prediction methods, Kriging interpolation considers not only the distance between unknown and sampled points but also the spatial correlation structure between sampled points (Tobler, 1970). Kriging interpolation describes how the variance between data points changes with distance (Webster and Oliver, 2007). However, Kriging interpolation is not suitable for dynamic environments. On one hand, it is computationally complex, typically having a time complexity of $O(n^3)$, where n is the number of data points (Srinivasan et al., 2010). On the other hand, the accuracy of Kriging interpolation depends on the choice of the variogram model and its parameterization (Erdogan Erten et al., 2022).

Air temperature can be considered a realization of a spatiotemporal function. Therefore, performing interpolation only in the spatial or temporal domain is suboptimal (Li S et al., 2020). Spatiotemporal Kriging (STK) and Gaussian process (GP) are two state-of-the-art approaches for temperature field reconstruction in meteorological applications. STK extends the conventional Kriging framework by constructing spatiotemporal variograms; however, it remains fundamentally an interpolation-based method with limited capability to capture complex nonlinear dynamic variations and often results in oversmoothing effects (Varouchakis et al., 2019; Duff et al., 2025). GP, as a nonparametric Bayesian modeling approach, characterizes the correlation structure of spatiotemporal data through kernel functions and provides a solid theoretical foundation for temperature field modeling (Chauhan et al., 2024). Nevertheless, it tends to inadequately represent abrupt temperature changes (Cheng et al., 2025).

Beyond these two mainstream approaches, both the state transition matrix in state-space Kriging (Shi and Zhou, 2021; Cheng et al., 2026) and the filtering model parameters in methods that combine Kriging with the ensemble Kalman filter (EnKF) or unscented Kalman filter (UKF) (Katzfuss et al., 2020) are typically predefined and fixed. Consequently, these methods fail to fully exploit the advantages of machine learning models in small-sample prediction scenarios.

Deep learning methods rely heavily on large-scale training data, exhibit limited interpretability, and are sensitive to noise. These drawbacks make them difficult to deploy robustly in field environments characterized by sparse observations (Wang SZ et al., 2022; Chen KQ et al., 2024; Liu EB et al., 2025). Ensemble learning approaches, which integrate deep learning models with Kriging interpolation, also lack an effective theoretical framework for incorporating multivariate spatiotemporal dependencies and nonstationary trends. This limitation constrains their applicability and practical value (Janowicz et al., 2020; Chen KQ et al., 2024; Nourani et al., 2024; Liu EB et al., 2025). In summary, existing temperature field reconstruction methods still exhibit limitations in capturing spatiotemporal coupling, modeling nonlinear dynamics, and adapting to small-sample scenarios. This situation motivates the need for a unified framework that effectively integrates statistical modeling with data-driven approaches.

To improve the accuracy of temperature parameters, in this study, we propose a Kalman filter expanded fusion (KFEF) algorithm that integrates the advantages of GRNN prediction and Kriging interpolation. The main contributions of this paper are as follows:

1. A dynamic GRNN optimization algorithm based on simulated annealing (SA), named GRNN-SA, is proposed. Unlike existing methods based on particle swarm optimization (PSO), genetic algorithms (GA), etc., SA employs the prediction error as the energy function and accepts inferior solutions with a certain probability. This enables rapid and accurate identification of optimal GRNN parameters.

2. A data filtering system named Kalman filter fusion (KFF), composed of Kalman filter (KF), GRNN, and Kriging interpolation, is proposed to predict data. Moreover, an adaptive KF is proposed to address the divergence problem. This filter dynamically adjusts the state transition matrix and the measurement noise matrix using measurement data to adapt to complex outdoor environments.

3. To enhance the robustness of the system, a fixed-interval interpolation method is proposed. This method first employs GRNN-SA and Kriging interpolation to extend the data. Then, it treats the extended data as observations and processes them using KFF.

2 Brief review of GRNN and Kriging

2.1 GRNN time prediction

GRNN is a method with high training speed. Its performance surpasses those of multivariate linear and nonlinear regression models (Ghritlahre and Prasad, 2018; Li G et al., 2018). GRNN is a special type of RBF neural network. The number of nodes in the hidden layer equals the number of training samples. The centers of RBF are determined by the samples. GRNN calculates the weighted average of the outputs based on the similarity between the input and training samples. Given an input sample, GRNN predicts the output as follows:

$$\hat{y}(\boldsymbol{x})=\frac{\sum_{i=1}^{N} y_i \exp\left(-\frac{\|\boldsymbol{x}-\boldsymbol{x}_i\|^2}{2\sigma^2}\right)}{\sum_{i=1}^{N} \exp\left(-\frac{\|\boldsymbol{x}-\boldsymbol{x}_i\|^2}{2\sigma^2}\right)}, \tag{1}$$

where σ is the smoothing factor and N is the number of training samples. It determines the weight distribution of samples in the prediction (Rooki, 2016). A smaller σ makes the model more sensitive to noise, leading to overfitting. A larger σ results in excessively smooth predictions and underfitting. The prediction results of GRNN are compared with those of RBF in Fig. 1. The results show that RBF can

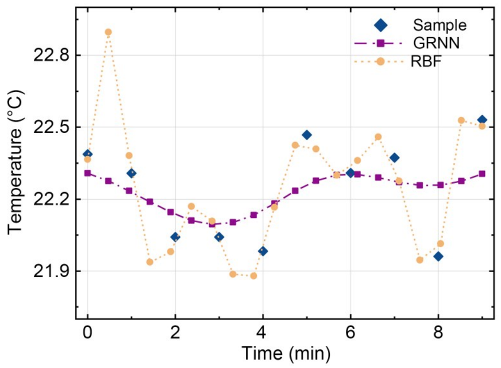


Fig. 1 Temperature prediction comparison among sample, GRNN, and RBF

model partial input data. In contrast, GRNN achieves superior overall prediction accuracy and effectively addresses nonlinear problems. However, the performance of GRNN is highly sensitive to the choice of the smoothing factor. Empirical settings and simple search strategies can trap the model in local optima, limiting its generalization capability.

PSO (Anggraeni et al., 2024), GA (Golpîra et al., 2011), and SA (Lee and Kim, 2020) are used to optimize the smoothing factor of GRNN. The parameter settings for the three methods are summarized in Table 1. Owing to the inherent stochasticity of PSO, GA, and SA, each algorithm is independently executed 30 times using different random seeds. The average root mean square error (RMSE) over the 30 runs is presented in Fig. 2. The PSO algorithm diverges in the early stage. GA shows poor stability in the later stage. RMSE fluctuates repeatedly during iterations 12–20. Overall, the SA algorithm demonstrates superior performance. It achieves the lowest RMSE. In Section 3, SA combined with temperature parameter characteristics is used to perform dynamic GRNN optimization.

2.2 Kriging spatial interpolation

The interpolation formula uses observational data to estimate parameters at unobserved locations (Tobler, 1970). To accomplish this task in complex spatial environments, an ideal interpolation model should account for the multiscale structural information embedded in spatial temperature data (Chen KQ et al., 2024). Kriging interpolation fully considers the multiscale structural information contained in temperature data. It is an optimal, unbiased interpolation method based on the spatial statistical theory.

Kriging uses a variogram to characterize the correlation between spatial variables and estimate values at unknown locations (Webster and Oliver, 2007). Suppose that the variable to be estimated at spatial location $\boldsymbol{x}_0$ is $\hat{Z}(\boldsymbol{x}_0)$. $\hat{Z}(\boldsymbol{x}_0)$ is expressed as a linear combination of n neighboring known sample points:

$$\hat{Z}(\boldsymbol{x}_0)=\sum_{i=1}^{n}\lambda_i Z(\boldsymbol{x}_i), \tag{2}$$

where λ_i represents the weight coefficient, satisfying the unbiasedness constraint $\sum_{i=1}^{n}\lambda_i=1$. The optimal weights are obtained by minimizing the prediction variance:

$$\sigma^2=\mathrm{Var}\,[\,\hat{Z}(\boldsymbol{x}_0)-Z(\boldsymbol{x}_0)\,]. \tag{3}$$

Table 1 Parameter settings for the three algorithms

Algorithm	Parameter	Value
SA	Initial temperature	26 °C
	Maximum number of iterations	100
	Cooling factor	0.95
PSO	Number of particles	20
	Maximum number of iterations	50
	Inertia weight w	0.5
	Learning factors c_1, c_2	2, 2
GA	Population size	20
	Maximum number of generations	50
	Crossover probability	0.8
	Mutation probability	0.05

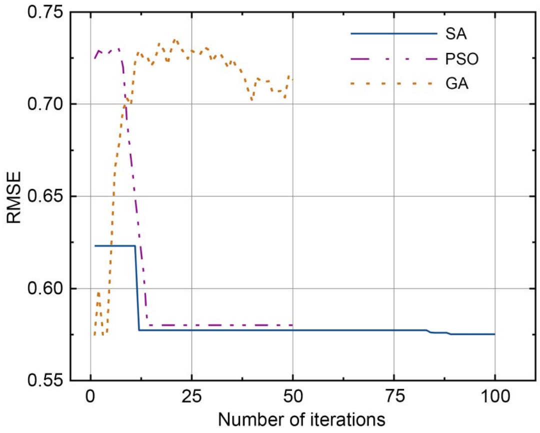


Fig. 2 Average RMSE convergence curves over 30 runs

The minimum prediction variance can also be expressed directly using the variogram:

$$y(h)=\frac{1}{2N(h)}\sum_{i=1}^{N(h)}[Z(\boldsymbol{x}_i)-Z(\boldsymbol{x}_i+h)]^2, \tag{4}$$

where h represents the spatial lag and $N(h)$ is the number of sample pairs separated by distance h. By modeling and fitting the variogram (e.g., spherical, exponential, or Gaussian model), the Kriging weights can be derived using a parametric form. Variogram models reflect the smoothness and scale of the spatial correlation of variables. Selecting the appropriate model significantly impacts the interpolation results (Miao and Wang, 2024). Fig. 3 shows the temperature interpolation results using different variogram models.

As shown in Fig. 3, different variogram models exhibit pronounced differences. These differences manifest in how they characterize the variation trends of the temperature field. The spherical model emphasizes global stationarity. The exponential model strikes a favorable balance between trend continuity and responsiveness to local variations. The Gaussian model, on the other hand, is more sensitive to local perturbations. In this study, the exponential variogram model is uniformly adopted as the core fitting model:

$$\gamma(h)=C_0+C_1\left(1-\mathrm{e}^{-\frac{h}{a}}\right), \tag{5}$$

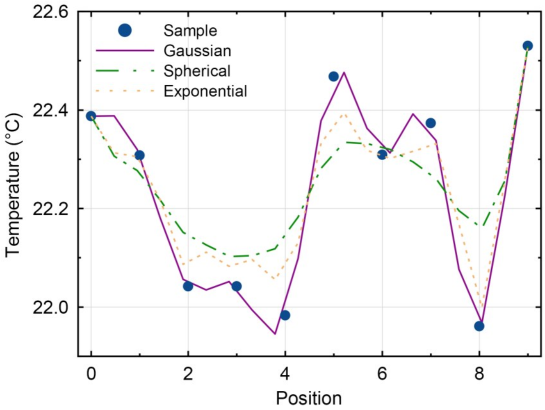


Fig. 3 Comparison of temperature interpolation results among Gaussian, spherical, and exponential variogram models

where C_0 denotes the nugget effect, C_1 is the partial sill, and a is the range. In this study, the empirical variogram is first computed from spatial temperature observations. Subsequently, a weighted least-squares (WLS) method is employed to fit the empirical variogram to the exponential variogram model. Finally, leave-one-out cross-validation (LOOCV) is used to optimize the model parameters, with the objective of minimizing the Kriging prediction variance to determine the final variogram parameters. To ensure the adaptability and generalization capability of the variogram model across different experimental scenarios, scenario-specific data sources and processing strategies are adopted for variogram fitting in simulations, indoor experiment, and kilometer-scale field experiment.

Compared with traditional interpolation methods, Kriging interpolation explicitly models spatial correlation through the variogram. This gives trend representation a clear statistical–physical meaning and demonstrates distinct advantages in terms of stability, interpretability, and trend consistency.

3 Methodology

As discussed above, GRNN assumes that the inputs are independent and identically distributed during the prediction process. In contrast, Kriging interpolation can model the relationships among input measurements, making it more rigorous and interpretable. We propose the KFEF algorithm to overcome the limitations of existing single prediction methods. This algorithm integrates the advantages of GRNN prediction and Kriging interpolation. The proposed KFEF framework is illustrated in Fig. 4.

To provide a clearer and more coherent description of the overall framework of the proposed method, we formulate GRNN-based prediction, Kriging-based spatial modeling, and KF within a unified state-space data fusion model.

First, within this unified framework, a dynamic GRNN-SA optimization algorithm based on the minimum RMSE is proposed. The algorithm has two components: outer-loop temperature update and inner-loop state update. The annealing temperature is updated iteratively in the outer loop using a geometric cooling strategy. In the inner loop, candidate parameters undergo multiple random perturbations at a fixed temperature. The acceptance of new solutions is determined according to the Metropolis criterion. The details of GRNN-SA are introduced in Section 3.1.

Second, a novel temperature prediction method named KFF is introduced. This method uses temporal and spatial correlations in temperature data to estimate temperature states within a KF. Moreover, an adaptive KF-based temperature estimation framework is constructed. This framework consists of a set of structurally consistent adaptive KFs and is theoretically grounded in Kriging variogram models to establish a spatial state model with statistical significance. Meanwhile, to avoid the uncertainties introduced by the empirically specified measurement noise covariance, we use GRNN to model errors in the temperature time series and construct the

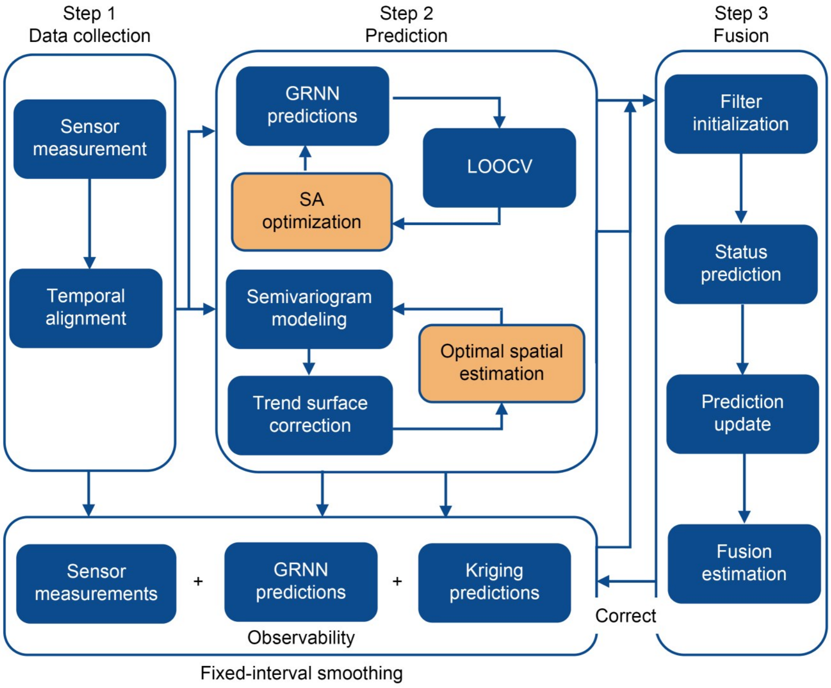


Fig. 4 Proposed temperature prediction system

measurement noise covariance matrix. The details of KFF are described in Section 3.2.

Finally, a fixed-interval interpolation method is proposed to enhance the prediction accuracy and expand the sensor data. The fixed-interval interpolation method has two components: interpolation and fusion processing. The data predicted by GRNN and Kriging are interpolated at equal intervals into the measurement data. The interpolated data are treated as observations, which are then processed using KFF. The processed data are regarded as the final prediction results. The details of the interpolation method are introduced in Section 3.3.

3.1 GRNN-SA optimization algorithm

To address GRNN's tendency to become trapped in local optima and its limited generalization capability, a dynamic GRNN-SA optimization algorithm is proposed. This algorithm introduces probabilistic distributions under different conditions to perform refined searches and achieve convergence in regions where superior solutions have been identified. Thus, it balances global exploration and local exploitation.

The dynamic GRNN-SA optimization algorithm treats the GRNN prediction error as the energy function of the SA algorithm. In the early stage, the initial temperature is high. According to the Metropolis criterion, inferior solutions are accepted with a high probability. This gives the system strong global search capabilities and the ability to escape local optima. During the search process, the algorithm explores the optimal solution in the solution space. It does so under the guidance of the current optimal region, updating the probability of accepting different solutions. The stages of the interaction process between SA and GRNN are illustrated in Fig. 5.

The RMSE of the validation set is defined as the energy function of the SA algorithm:

$$E(\sigma)=\frac{1}{M}\sum_{k=1}^{M}\left(y_k-\hat{y}(\boldsymbol{x}_k,\sigma)\right)^2, \tag{6}$$

where y_k denotes the dimensionality of the test point, $\hat{y}(\boldsymbol{x}_k,\sigma)$ represents the temperature predicted by GRNN, and M denotes the number of validation set samples. The one-dimensional Gaussian kernel function is defined as

$$a_i(\boldsymbol{x},\sigma)=\exp\left(-\frac{\|\boldsymbol{x}-\boldsymbol{x}_i\|^2}{2\sigma^2}\right),\quad i=1,2,\cdots,N, \tag{7}$$

where $\boldsymbol{x}_i$ represents the i^{th} input of the training sample. The normalized weight formulation is represented as

$$w_i(\boldsymbol{x},\sigma)=\frac{a_i(\boldsymbol{x},\sigma)}{\sum_{j=1}^{N}a_j(\boldsymbol{x},\sigma)}, \tag{8}$$

where j is the summation index over the training samples. The GRNN prediction is given as

$$\hat{y}(\boldsymbol{x},\sigma)=\sum_{i=1}^{M}w_i(\boldsymbol{x},\sigma)y^i. \tag{9}$$

To prevent overfitting and improve the prediction capability for unknown points, the LOOCV method is used to calculate the error. Hence, the energy function is expressed as

$$E_{\text{LOO}}(\sigma)=\frac{1}{m}\sum_{k=1}^{m}\left(y^k-\hat{y}_{-k}(\boldsymbol{x}^k,\sigma)\right)^2, \tag{10}$$

where $\hat{y}_{-k}(\boldsymbol{x}^k,\sigma)$ represents the GRNN prediction obtained after removing the k^{th} sample from the dataset. The optimal value of σ is defined as

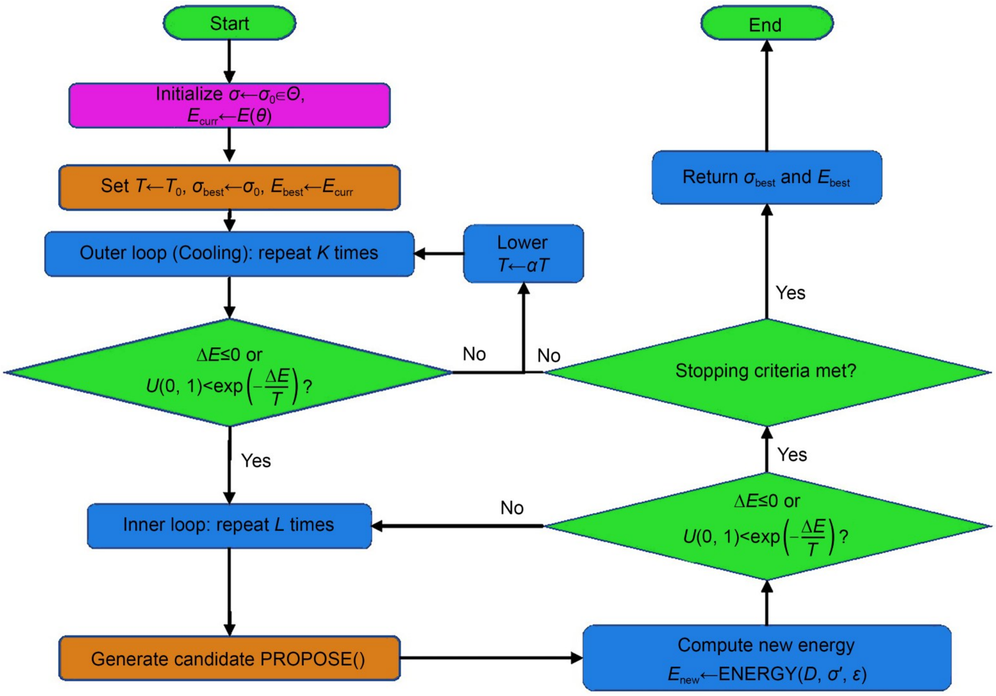


Fig. 5 GRNN-SA optimization process

$$\sigma^{*}=\arg\min_{\sigma\in\Theta} E(\sigma), \tag{11}$$

where Θ is the search range. According to the Metropolis criterion, a new state, σ', is randomly selected from the current state, σ^{*}. The energy difference is represented by $\Delta E=E(\sigma')-E(\sigma^{*})$, and the acceptance probability is expressed as follows:

$$P_{\text{accept}}=\begin{cases}1, & \Delta E\leq 0,\\ \exp\left(-\dfrac{\Delta E}{T_k}\right), & \Delta E>0,\end{cases} \tag{12}$$

where T_k is the temperature parameter at the k^{th} iteration in the SA process, which is determined by the current GRNN prediction. When the current RMSE exceeds the initial RMSE, there is still a certain probability of accepting a suboptimal solution. When the predicted output temperature does not meet the error requirements, it will be gradually reduced according to the geometric cooling strategy: $T_{k+1}=\alpha T_k$, $0<\alpha<1$. The detailed procedure is explained in Algorithm 1.

Algorithm 1 GRNN-SA algorithm

```
Input: dataset $D$, outer iteration number $K$, inner iteration number $L$, error threshold $E_t$, initial temperature $T_0$, and cooling coefficient $\alpha$ ($0<\alpha<1$).
Output: $\sigma_{best}$, $E_{best}$.
Initialize smoothing factors $\sigma\in\Theta$ and $E_{curr}$ according to Eq. (10), and
  set $\sigma_{best}\leftarrow\sigma_0$, $E_{best}\leftarrow E_{curr}$, $T\leftarrow T_0$;
for $i$=1 to $K$ do
    if $E_{best}\leq E_t$ then
        Break;
    end if
    for $j$=1 to $L$ do
        Generate candidate solution $\sigma'=\sigma\cdot 10^{\varepsilon}$, and add perturbation $\varepsilon$
          on a logarithmic scale;
        Calculate new energy $E_{new}=\mathrm{ENERGY}(D,\sigma',\varepsilon)$;
        Update $\Delta E=E_{new}-E_{curr}$;
        if $\Delta E\leq 0$ or $U(0,1)<\exp(-\Delta E/T)$ then
            Update $\sigma=\sigma'$, $E_{curr}=E_{new}$;
            if $E_{curr}<E_{best}$ then
                Update $\sigma_{best}=\sigma'$, $E_{best}=E_{curr}$;
            end if
        end if
    end for
    $T\leftarrow\alpha T$;
end for
Return $\sigma_{best}$, $E_{best}$;
```

3.2 KFF method

In small-scale temperature field modeling, the temperature field tends to reach a steady state, and a sensor fusion algorithm based on a linear KF is proposed. The data fusion process of the KF algorithm is divided into two stages: prediction stage and correction and update stage. In the prediction stage, the state estimate from the previous time step is used to predict the current state. The correction and update stage uses observations of the current state to correct the prediction obtained in the previous stage, yielding a relatively accurate state estimate at the current time, as shown in Fig. 6.

1. Prediction stage: KF uses the control input $\boldsymbol{u}_{k-1}$, the posterior estimate $\hat{\boldsymbol{x}}_{k-1}$ at time $k-1$, and its covariance matrix $\boldsymbol{P}_{k-1}$ to predict the value $\hat{\boldsymbol{x}}_k^-$ and the covariance $\boldsymbol{P}_k^-$ at the current time k as follows:

$$\begin{cases}\hat{\boldsymbol{x}}_k^-=\boldsymbol{A}\hat{\boldsymbol{x}}_{k-1}+\boldsymbol{B}\boldsymbol{u}_{k-1},\\ \boldsymbol{P}_k^-=\boldsymbol{A}\boldsymbol{P}_{k-1}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{Q},\end{cases} \tag{13}$$

where $\boldsymbol{A}$ denotes the state transition matrix, $\boldsymbol{Q}$ is the process noise covariance matrix, and $\boldsymbol{B}$ represents the control input matrix.

2. Correction and update stage: The Kalman gain $\boldsymbol{K}_k$ determines the weighting between the predicted information $\hat{\boldsymbol{x}}_k^-$ and the observed information z_k. Using $\boldsymbol{K}_k$, the state $\hat{\boldsymbol{x}}_k$ and the error covariance $\boldsymbol{P}_k$ are updated to perform correction as follows:

$$\begin{cases}\boldsymbol{K}_k=\boldsymbol{P}_k^-\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{P}_k^-\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R})^{-1},\\ \hat{\boldsymbol{x}}_k=\hat{\boldsymbol{x}}_k^-+\boldsymbol{K}_k(z_k-\boldsymbol{H}\hat{\boldsymbol{x}}_k^-),\\ \boldsymbol{P}_k=(\boldsymbol{I}-\boldsymbol{K}_k\boldsymbol{H})\boldsymbol{P}_k^-,\end{cases} \tag{14}$$

where $\boldsymbol{H}$ is the observation matrix that describes the mapping from the state space to the observation space, and $\boldsymbol{R}$ denotes the measurement noise covariance matrix.

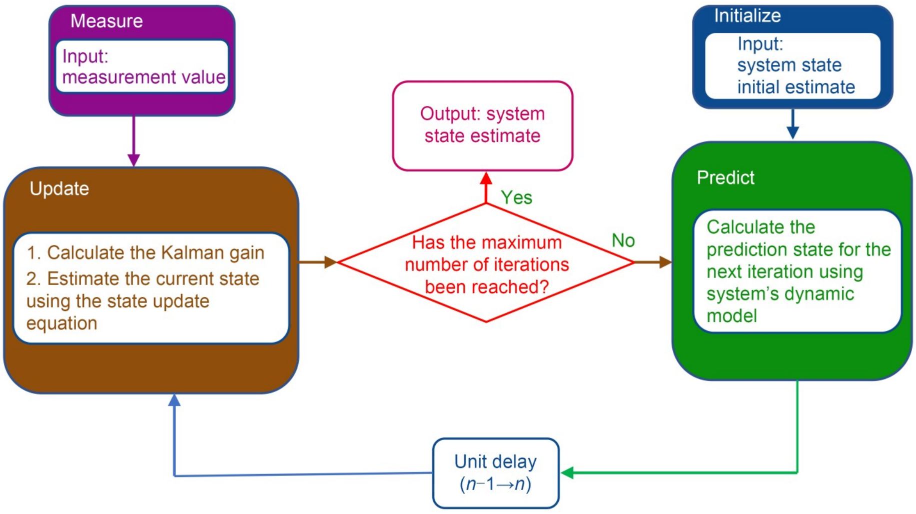


Fig. 6 Kalman filter principle

3.2.1 State transition matrix based on Kriging

According to the geostatistical theory, Kriging provides the best linear unbiased estimator (BLUE) under the assumption of second-order stationarity. The Kriging weights characterize the spatial dependence among neighboring points through the variogram model. Within the KF framework, the Kriging weight matrix is employed to construct the state transition matrix, enabling the filter to incorporate spatial correlations into the dynamic evolution of the temperature field. The set of spatial positions is denoted as $S=\{s_1, s_2, \cdots, s_n\}$, and the state vector is expressed as

$$\boldsymbol{x}_t=[T_t(s_1), T_t(s_2), \cdots, T_t(s_n)]^{\mathrm{T}}\in\mathbb{R}^n, \tag{15}$$

where s_n denotes the target estimation point and $T_t(s_n)$ represents the temperature at time t and s_n. The variogram γ is used to characterize the spatial correlation, expressed as

$$\gamma(h)=\frac{1}{2}\mathrm{Var}(T(s+h)-T(s)). \tag{16}$$

For a second-order stationary temperature field, the relationship between the variogram and the covariance is given by

$$\begin{aligned} C(h)&=\mathrm{Cov}(T(s), T(s+h)) \\ &=\sigma^2-\gamma(h), \end{aligned} \tag{17}$$

where σ^2 denotes the overall variance. Using the standard ordinary Kriging equations without an external trend term, m neighboring points s_j are employed to interpolate the unknown point s_0. For each neighborhood index $p=1, 2, \cdots, m$, the following condition is satisfied:

$$\sum_{j=1}^{m}\lambda_j\gamma(s_p-s_j)+\mu=\gamma(s_p-s_0), \tag{18}$$

where μ is the Lagrange multiplier and λ_j is the linear unbiased weight of Kriging satisfying the unbiasedness constraint:

$$\sum_{j=1}^{m}\lambda_j=1. \tag{19}$$

Combining Eqs. (17) and (18), the local variance matrix is defined, yielding

$$\Gamma_{pq}^{(i)}=\gamma(s_p-s_q)p, \quad q=1, 2, \cdots, m. \tag{20}$$

We define the vector $\boldsymbol{\gamma}_p^{(i)}\in\mathbb{R}^m$, yielding

$$\boldsymbol{\gamma}_p^{(i)}=\boldsymbol{\gamma}(\boldsymbol{s}_p-\boldsymbol{s}_0), \quad p=1, 2, \cdots, m, \tag{21}$$

where $\boldsymbol{\gamma}$ is a spatiotemporal vector-valued variogram with time scale, $\boldsymbol{s}_p$ represents the spatial coordinates of the p^{th} neighboring sampling point, and $\boldsymbol{\gamma}_p^{(i)}$ describes the variation characteristics from $\boldsymbol{s}_p$ to the unknown interpolation point $\boldsymbol{s}_0$ within the given time scale. Eqs. (21) and (20) differ distinctly in variable form and the integration of time scale.

Defining the vector $\mathbf{1}_m=(1, \cdots, 1)^{\mathrm{T}}\in\mathbb{R}^m$, the augmented matrix form of the linear system is expressed as

$$\begin{bmatrix}\boldsymbol{\Gamma}^{(i)} & \mathbf{1}_m \\ \mathbf{1}_m^{\mathrm{T}} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{\lambda}^{(i)} \\ \boldsymbol{\mu}\end{bmatrix}=\begin{bmatrix}\boldsymbol{\gamma}^{(i)} \\ \mathbf{1}\end{bmatrix}, \tag{22}$$

where the local weight $\boldsymbol{\lambda}^{(i)}$ is the Kriging weight at point i. For each temperature prediction target point $i=1, 2, \cdots, n$, the neighborhood $N(i)$ is selected, and $\boldsymbol{\lambda}^{(i)}$ is solved. The weights of each row are placed in the i^{th} row of matrix $\boldsymbol{\Lambda}\in\mathbb{R}^{n\times n}$, and each row represents the local Kriging weight vector:

$$\Lambda_{i,j}=\begin{cases}\lambda_j^{(i)}, & j\in N(i), \\ 0, & \text{otherwise.}\end{cases} \tag{23}$$

Following the above derivation and the autoregressive model of temperature, the state transition matrix is obtained:

$$\boldsymbol{A}=\alpha(\Delta t)\boldsymbol{I}+(1-\alpha(\Delta t))\boldsymbol{\Lambda}, \tag{24}$$

where $\alpha(\Delta t)=\mathrm{e}^{-\Delta t/\tau}$, τ is the time scale, and Δt is the step size.

Assumptions on the spatiotemporal characteristics of the temperature field are introduced to construct a state transition matrix based on Kriging weights. First, the spatial temperature field is assumed to satisfy second-order stationarity within a local region. Second, the temperature evolution at each spatial location is assumed to be primarily influenced by neighboring locations rather than distant points. Third, over short time intervals, the temporal evolution of temperature is assumed to exhibit weak temporal autocorrelation, implying relatively smooth variations. In this study, the spatial temperature field is primarily dominated by spatial propagation with weak short-term autoregression, yielding

$$\boldsymbol{A}=\boldsymbol{\Lambda}. \tag{25}$$

The matrix $\boldsymbol{Q}$ is modeled using the Kriging variance and its associated covariance structure. Specifically, the diagonal elements are expressed as

$$Q(i,i)=\sigma_{\mathrm{K}}^2(s_i)=\gamma(0)-\sum_{j=1}^{n}\lambda_j(s_i)\gamma\left(h_{j,s_i}\right)+\mu(s_i), \tag{26}$$

where $\gamma(\cdot)$ denotes the variogram of the temperature field defined in Eq. (16), n represents the number of neighboring sampling points used for Kriging interpolation, $\lambda_j(s_i)$ denotes the optimal unbiased Kriging weights at location s_i obtained from Eqs. (17)–(19), h_{j,s_i} is the spatial lag distance between the j^{th} sampling point and the target location, and $\mu(s_i)$ is the Lagrange multiplier obtained from solving the Kriging system. The off-diagonal elements can be expressed as

$$\begin{aligned} Q(i,l)=\mathrm{Cov}\left[\sigma_{\mathrm{K}}^2(s_i), \sigma_{\mathrm{K}}^2(s_l)\right]&=\gamma(h_{i,l})-\sum_{j=1}^{n}\lambda_j(s_i)\gamma(h_{j,s_i}) \\ &-\sum_{j=1}^{n}\lambda_j(s_l)\gamma(h_{i,s_i})+\mu(s_i)+\mu(s_l). \end{aligned} \tag{27}$$

The treatment of strongly nonstationary atmospheric fields is widely acknowledged as a fundamental challenge in both geostatistics and atmospheric modeling. Absolute distance measurements focus on scenarios in which the temperature field approximately satisfies local stationarity (Wu et al., 2017; Xu et al., 2020). In this study, the experimental data are likewise collected during periods with relatively stable meteorological conditions and are obtained by averaging multiple measurements, thereby satisfying the assumption of second-order stationarity. This ensures the applicability and effectiveness of the proposed modeling approach.

3.2.2 Measurement model

The KFF framework employs a multisource parallel observation architecture comprising three independent observation branches: physical sensor observation branch, GRNN-based temporal prediction branch, and Kriging-based spatial interpolation branch. The covariance matrix of the GRNN-SA prediction noise is obtained using the statistical error sampling method. The GRNN method naturally accommodates nonlinear models and is used for high-dimensional data prediction. Let the temperature time series be

$$y(t),\quad t=1,2,\cdots,N. \tag{28}$$

Generate the prediction values in the training set:

$$\hat{y}(t)=f_{\mathrm{GRNN}}(y(1{:}t-1)). \tag{29}$$

Define the prediction residual of temperature data as

$$e(t)=y(t)-\hat{y}(t). \tag{30}$$

The output of GRNN is expressed in the form of kernel regression as

$$\hat{y}(t)=\sum_{i=1}^{N}w_i(t)\,y(i), \tag{31}$$

where $w_i(t)=\dfrac{\exp\left(-\dfrac{[y(t-1)-y(i-1)]^2}{2\sigma^2}\right)}{\sum\limits_{i=1}^{N}\exp\left(-\dfrac{[y(t-1)-y(i-1)]^2}{2\sigma^2}\right)}$ and $\sum\limits_{i=1}^{N}w_i(t)=1$.

Therefore, the prediction error of GRNN consists of the weight error of neighboring samples introduced by kernel regression, the propagation of observation noise through kernel weighting, and the structural approximation error of GRNN. KF requires the measurement noise matrix satisfy the following relationship:

$$\boldsymbol{R}=E[\boldsymbol{e}(t)\boldsymbol{e}(t)^{\mathrm{T}}]. \tag{32}$$

Under the GRNN sample framework, the measurement noise is estimated as

$$\hat{\boldsymbol{R}}=\frac{1}{m-1}\sum_{i=1}^{m}(\boldsymbol{e}_i-\bar{\boldsymbol{e}})(\boldsymbol{e}_i-\bar{\boldsymbol{e}})^{\mathrm{T}}, \tag{33}$$

where $\bar{\boldsymbol{e}}=\dfrac{1}{m}\sum\limits_{i=1}^{m}\boldsymbol{e}_i$. For a multisensor temperature prediction system, the complete measurement noise covariance matrix is given by

$$\hat{\boldsymbol{R}}=\begin{bmatrix}\mathrm{Var}(\boldsymbol{e}_1) & \mathrm{Cov}(\boldsymbol{e}_1,\boldsymbol{e}_2) & \cdots & \mathrm{Cov}(\boldsymbol{e}_1,\boldsymbol{e}_m)\\ \mathrm{Cov}(\boldsymbol{e}_2,\boldsymbol{e}_1) & \mathrm{Var}(\boldsymbol{e}_2) & \cdots & \mathrm{Cov}(\boldsymbol{e}_2,\boldsymbol{e}_m)\\ \vdots & \vdots & & \vdots\\ \mathrm{Cov}(\boldsymbol{e}_m,\boldsymbol{e}_1) & \mathrm{Cov}(\boldsymbol{e}_m,\boldsymbol{e}_2) & \cdots & \mathrm{Var}(\boldsymbol{e}_m)\end{bmatrix}_{m\times m}. \tag{34}$$

3.3 Fixed-interval interpolation method

In accordance with the aforementioned unified spatiotemporal state-space model, the precision of state estimation is constrained when the number of sensor observation points is limited. To improve the accuracy of temperature measurement and enrich sensor data, a fixed-interval interpolation method is proposed. The fixed-interval interpolation is illustrated in Fig. 7. Here, Z_{m} denotes the measured temperature value, Z_{g} represents the GRNN-SA prediction result, and Z_{k} denotes the Kriging interpolation result. The value Z_{g} is predicted using Z_{m} as the prior, while Z_{k} is obtained by performing Kriging interpolation on Z_{m} and Z_{g}. Within the meteorological interval L, Z_{g} and Z_{k} are inserted into L at fixed intervals and unified into a local coordinate system aligned with the principal directions, forming the column vector $[Z_{\mathrm{m}}, Z_{\mathrm{k}}, Z_{\mathrm{g}}, Z_{\mathrm{k}}, Z_{\mathrm{m}}, \cdots]$.

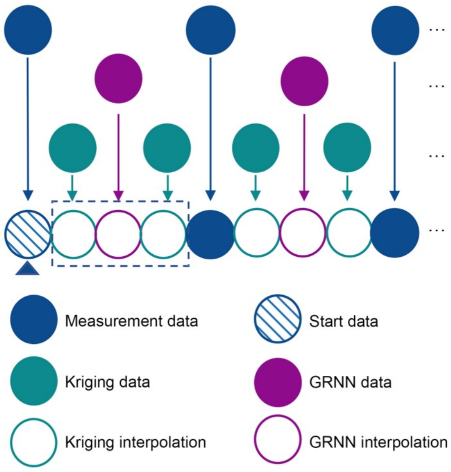


Fig. 7 Fixed interpolation model

By integrating the GRNN prediction, Kriging-based spatial interpolation, and KF processes into a unified spatiotemporal state-space model, the state estimation update equation is formulated as

$$\boldsymbol{x}_k=\boldsymbol{A}_{\mathrm{K}}\boldsymbol{x}_{k-1}+\boldsymbol{K}(\boldsymbol{Z}_t-\boldsymbol{H}\boldsymbol{A}\boldsymbol{x}_{k-1}). \tag{35}$$

The Kalman gain and covariance propagation equations are given by

$$\boldsymbol{K}=\boldsymbol{P}_{k|k-1}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{P}_{k|k-1}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}_{\mathrm{GRNN}})^{-1}, \tag{36}$$

$$\boldsymbol{P}_{k|k-1}=\boldsymbol{A}_{\mathrm{K}}\boldsymbol{P}_{k-1}\boldsymbol{A}_{\mathrm{K}}^{\mathrm{T}}+\boldsymbol{Q}. \tag{37}$$

In Eqs. (35)–(37), $\boldsymbol{Z}_t$ is the observation vector, i.e., $[Z_{\mathrm{m}}, Z_{\mathrm{k}}, Z_{\mathrm{g}}, Z_{\mathrm{k}}, Z_{\mathrm{m}}, \cdots]$, $\boldsymbol{A}_{\mathrm{K}}$ denotes the state transition matrix constructed from Kriging interpolation, $\boldsymbol{Q}$ represents the process noise derived from the Kriging variogram, $\boldsymbol{R}_{\mathrm{GRNN}}$ denotes the observation noise associated with the GRNN prediction, and $\boldsymbol{K}$ is the Kalman-computed fusion gain.

Therefore, the proposed KFEF model is simplified into two components: interpolation and fusion processing. First, the measured temperature values are obtained and expanded by performing fixed-interval interpolation based on GRNN-SA predictions introduced in Section 3.1 and Kriging interpolation. Second, fusion filtering is carried out using the KFF introduced in Section 3.2. After KFF processing, the final prediction results are produced which constitute the output of the proposed KFEF model.

4 Experiment and results

4.1 Experimental setup

To evaluate the performance of the proposed method, simulation, indoor experiment, and outdoor experiment are conducted. First, a temperature field simulation is established in Python. To meet the requirements of long-distance measurement, the simulation measurement length is set to L=1000 m. Terrain undulation is considered the main factor affecting temperature variation, and a one-dimensional temperature field random walk model is constructed along the measurement path. It is assumed that the temperature exhibits small random variations between adjacent spatial sampling points and satisfies the following relationship:

$$T(x_{k+1})=T(x_k)+\frac{6}{L}(\tan 1)\cdot\xi_k, \tag{38}$$

where $\xi_k\in\{-1, 0, 1\}$ represents the random perturbation caused by terrain undulation.

Second, atmospheric parameter acquisition experiments are conducted multiple times using a sensor array system at indoor and outdoor testing sites in Hangzhou, China. Additionally, the method for evaluating the prediction accuracy in the experiments is critically important. In this experiment, the traditional RBF interpolation method, GRNN-SA prediction, and ordinary Kriging interpolation are used as baselines to evaluate the interpolation performance. To provide a comprehensive evaluation, we further consider two representative spatiotemporal modeling approaches as baseline methods: STK and GP. Three evaluation metrics are introduced to assess the accuracy of the interpolation results: RMSE, mean absolute error (MAE), and mean absolute percentage error (MAPE). The calculation formulae are provided as follows:

$$\text{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}^{(i)}-y^{(i)})^2}, \tag{39}$$

$$\text{MAE}=\frac{1}{N}\sum_{i=1}^{N}|\hat{y}^{(i)}-y^{(i)}|, \tag{40}$$

$$\text{MAPE}=\frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}^{(i)}-y^{(i)}}{y^{(i)}}\right|\times 100\%, \tag{41}$$

where $y^{(i)}$ and $\hat{y}^{(i)}$ represent the observation and estimation values at position s_i, respectively. In both indoor and outdoor experiments, a compact sensor array is deployed at equal intervals along the measurement path to collect temperature distributions. The number of sensors remains constant, and the density of sensors varies at different distances during outdoor experiments. All sensors are installed at a height of approximately 1.5 m above the ground to mitigate the influence of ground thermal radiation. The sampling frequency is set to 1 Hz, with a temperature measurement accuracy of ±0.1 ℃. To ensure reproducibility, the main algorithmic parameters used in the experiments are summarized in Table 2.

4.2 Analysis of simulation results

In this study, we use the Monte Carlo method to perform 1000 random simulations based on the physical evolution of the temperature field and take the average, systematically comparing the interpolation and prediction performances of different methods. The results indicate that KFF outperforms all baseline methods in overall interpolation accuracy. Compared with RBF, KFF achieves an approximately 44% improvement in RMSE, significantly enhancing the upper limit of temperature field reconstruction accuracy and validating the effectiveness of the fusion framework in complex meteorological parameter estimation problems.

Furthermore, KFF is compared with two representative methods in Fig. 8: GRNN-SA and Kriging interpolation. The results show that KFF improves RMSE by approximately 21.8% and 14% compared to GRNN-SA and Kriging, respectively.

As shown in Fig. 8, the three methods exhibit significant differences in local regions. In region A (where temperature changes are drastic and local fluctuations are pronounced), the GRNN-SA method, by leveraging historical time-series information, effectively captures short-term dynamic temperature variations. In contrast, Kriging interpolation, due to sparse spatial samples, responds inadequately to rapidly changing temperature gradients, exhibiting a certain smoothing bias. In region B (where the overall trend is clear and spatial continuity is strong), the Kriging method demonstrates good capability in capturing the overall temperature distribution trend but cannot accurately reflect local details. The GRNN-SA

Table 2 Parameter settings for the three methods

Method	Parameter	Value/Description
KFEF	GRNN smoothing factor search range	[0.01, 10]
	Kalman filter sampling interval	1 s
	Number of Kriging neighbors	5
	Ordinary variogram model	Exponential
	SA initial temperature	26 ℃
	Cooling factor	0.95
STK	Time variogram model	Exponential
	Space variogram model	Exponential
GP	Kernel function type	Separable spatiotemporal squared exponential
	Hyperparameter optimization method	Maximum likelihood estimation
	Number of neighboring samples	20
	Initial noise variance	0.01

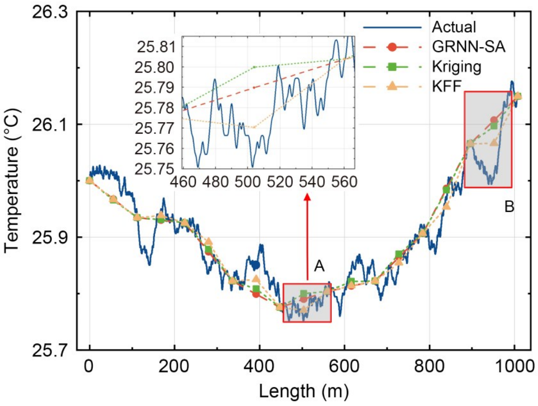


Fig. 8 Comparison of temperature prediction among GRNN-SA, Kriging, and KFF in simulations

method exhibits some errors in the spatial consistency of temperature. The KFF method, by integrating the prior information from Kriging and GRNN in regions A and B, achieves better results than both GRNN-SA and Kriging predictions.

Unfortunately, simulations show that when the number of temperature observation points is limited, the accuracy of KFF is relatively low. To analyze the effect of the number of temperature observation points on the fusion accuracy of KFF, numerical simulations are conducted under different observation densities. Specifically, 10, 20, 30, 40, and 50 temperature observation points are set, and the prediction results are quantitatively evaluated using RMSE, MAE, and MAPE (Table 3).

Table 3 Comparison of the three evaluation metrics between KFF and KFEF at different numbers of points in simulations

Number of points	RMSE		MAE		MAPE	
	KFF	KFEF	KFF	KFEF	KFF	KFEF
10	0.045	0.026	0.034	0.023	0.13	0.11
20	0.036	0.025	0.032	0.022	0.12	0.10
30	0.031	0.025	0.027	0.021	0.11	0.09
40	0.025	0.024	0.021	0.020	0.09	0.09
50	0.024	0.023	0.020	0.020	0.08	0.08

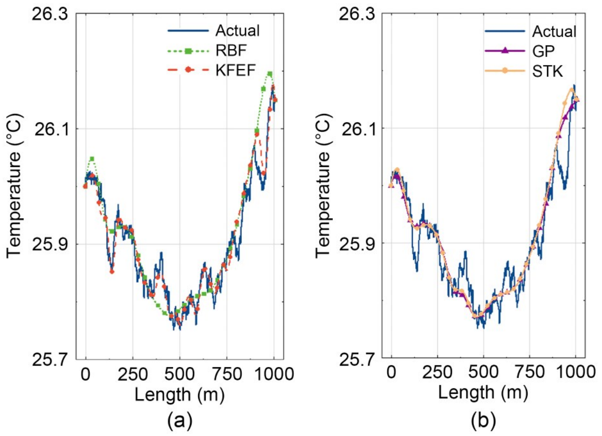
As shown in Table 3, the prediction accuracy of KFF is closely related to the number of observation points. When the number of observation points is small, for example, 10 or 20, the fusion results exhibit relatively large errors. This indicates that under limited observation information, the state estimation is insufficiently constrained, leading to a decrease in fusion accuracy.

As the number of observation points increases, KFF prediction performance shows a clear improvement trend. When the number of observation points increases from 10 to 50, RMSE decreases by approximately 47%, and both MAE and MAPE exhibit a consistent monotonic decline. Moreover, when the number of observation points reaches 40 and 50, the differences in the error metrics are small, and the prediction accuracy stabilizes.

However, in practical applications, the number of observation points is limited by sensor deployment costs and environmental conditions, making it difficult to meet the requirements for high-density sampling. To address the performance limitations of KFF under sparse observation conditions, in this study we introduce a fixed-interval interpolation method and construct the KFEF approach. Under identical observation density conditions, the quantitative comparison results with KFEF are also presented in Table 3. The results indicate that the KFEF scheme consistently achieves higher accuracy than KFF across all observation densities, with the performance gain becoming more pronounced as the observations become sparser. Specifically, when the number of observation points is 10, KFEF achieves an RMSE improvement of 42% relative to KFF.

The performance advantages of KFEF in dynamic temperature data modeling and prediction are even more pronounced. Fig. 9 shows the fitted temperature field based on the prediction data. Fig. 9a shows that the KFEF predicted temperature field closely matches the actual temperature field in both overall trends and local fluctuation features, effectively tracking rapid temperature changes. Fig. 9b shows that although STK accounts for both spatial and temporal correlations, its performance is constrained by sparse observations, resulting in an insufficient response to dynamically varying temperature gradients and a pronounced smoothing effect. While GP achieves lower overall errors compared with STK, it remains inadequate in capturing abrupt changes, leading to relatively large errors at discontinuity points.

Fig. 10 illustrates the error distribution between the prediction results of KFEF, STK, and GP and the actual temperature field. The



(a)

(b)

Fig. 9 Temperature prediction in simulations: (a) RBF and KFEF; (b) GP and STK

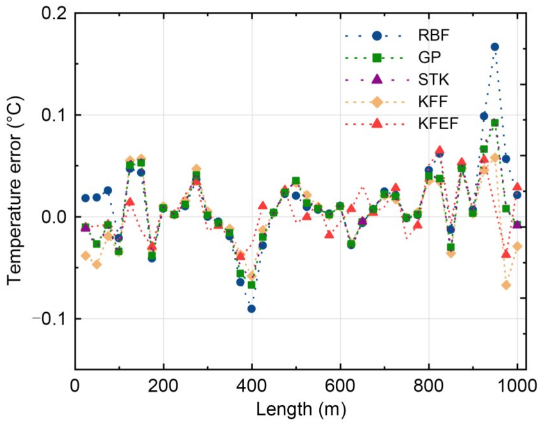


Fig. 10 Distribution of temperature errors for the five methods in simulations

KFEF prediction errors are generally small and are evenly distributed, with no significant systematic bias or localized error concentration. The error of GP increases markedly in regions with sharp temperature gradients, whereas the error of STK exhibits a relatively uniform overall distribution. The maximum absolute errors of both methods are significantly higher than those of KFEF.

These observations are fully consistent with the underlying theoretical framework: The spatiotemporal variogram fitting in STK is highly sensitive to sample density. In small-sample scenarios, the lack of sufficient data-driven support leads to overly smooth interpolation. The kernel function in GP is difficult to optimize globally in small-sample scenarios, resulting in local prediction fluctuations. In contrast, the proposed KFEF method recursively integrates spatial covariance information and temporal prediction capability within a KF framework, effectively suppressing the local measurement noise and reducing the spatial interpolation bias.

To quantitatively evaluate the overall performance of different methods in temperature field reconstruction, we conduct a comparative analysis of RBF, KFF, GP, STK, and KFEF methods. RMSE, MAE, and MAPE metrics between the interpolation results and the actual temperature data are summarized in Table 4. The results indicate that KFEF outperforms both RBF and KFF across all error metrics, and the performance improvements of GP, STK, KFF, and KFEF relative to RBF are shown in Table 5. Compared with STK and GP, both of which account for spatiotemporal correlations, KFF achieves modest reductions in RMSE of 5.2% and 2.7%, respectively, thereby validating the effectiveness of the KFF framework in integrating spatiotemporal information. With the introduction of fixed-interval interpolation, the KFEF method further enlarges the RMSE reductions to 34.21% and 32.43% relative to STK and GP, respectively, highlighting the substantial improvement in spatiotemporal modeling accuracy achieved through data augmentation strategies under sparse observation conditions.

Table 4 Three error evaluation metrics between the interpolation results of five methods and the actual temperature in simulations

Method	RMSE	MAE	MAPE
RBF	0.065	0.044	0.17
STK	0.038	0.035	0.14
GP	0.037	0.034	0.13
KFF	0.036	0.032	0.12
KFEF	0.025	0.021	0.09

Table 5 Reduction ratios of the four methods relative to RBF in terms of the three error evaluation metrics

Method	Reduction ratio (%)		
	RMSE	MAE	MAPE
STK	41.54	20.45	17.65
GP	43.08	22.73	23.53
KFF	44.05	28.83	28.75
KFEF	61.54	52.03	47.06

4.3 Indoor results

To further verify the accuracy of the proposed KFEF method, an indoor validation experiment is conducted. Atmospheric parameters are collected using a sensor array, ranging measurements are performed with a self-developed microwave ranging system (Zhao et al., 2013), and a high-precision laser ranging system is employed as a reference for result validation; the experimental setup is shown in Fig. 11.

Taking a propagation distance of 30 m as a representative case, the KFEF method is compared with RBF. As shown in Fig. 12, the fitting results of KFEF are closer to the actual temperature curve, whereas RBF performs a local fit based on the measured temperatures and exhibits noticeable deviations from the actual temperature.

The error distributions of different methods are shown in Fig. 13. The RBF method demonstrates the most significant deviations and fluctuations, as it does not consider the spatial correlations or temporal evolution of the temperature field. The incorporation of spatial and temporal factors in both GP and STK results in a substantial reduction in error compared to RBF; nevertheless, their errors persist

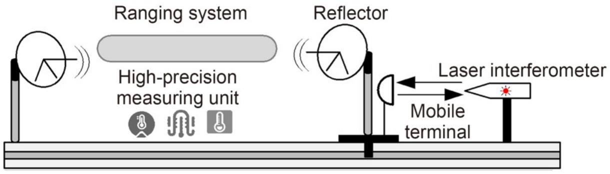


Fig. 11 Indoor validation platform

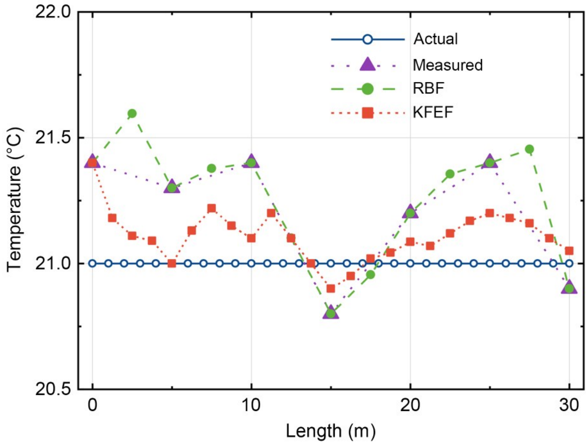


Fig. 12 Temperature prediction comparison between RBF and KFEF in the indoor experiment

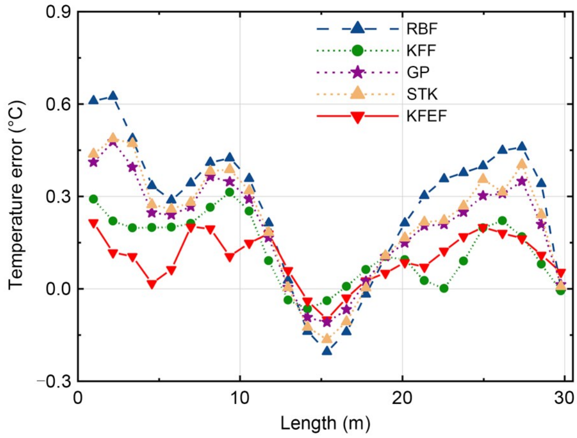


Fig. 13 Distribution of temperature errors for the five methods in the indoor experiment

at a slightly elevated level compared to those of KFF. By integrating prior information from both temporal and spatial domains, KFF achieves recursive error suppression and effectively prevents error accumulation. The KFEF method attains the lowest error and the smallest fluctuation amplitude. The fixed-interval interpolation strategy augments the observational information, mitigating the degradation in state estimation caused by sparse observations, while KFF fusion performs optimal estimation over the augmented multi-source data. These results are consistent with the theoretical findings from the simulations.

The quantitative error metrics are summarized in Table 6, showing that KFEF outperforms the comparative methods across all error indicators. Compared with RBF, RMSE decreases from 0.3504 to 0.1471, representing an approximate reduction of 58%; relative to KFF, KFEF also achieves a significant improvement in accuracy, thereby validating the effectiveness of data augmentation.

Table 6 Three error evaluation metrics between the interpolation results of five methods and the actual temperature in the indoor experiment

Method	RMSE	MAE	MAPE
RBF	0.350	0.312	1.49
STK	0.195	0.164	0.73
GP	0.193	0.154	0.72
KFF	0.190	0.148	0.70
KFEF	0.147	0.121	0.57

Ranging accuracy is used to validate KFEF, and the empirical relationship between the atmospheric refractive index and temperature is given as follows (Rüeger, 2002):

$$n-1=77.6890\frac{P}{T_{\mathrm{a}}}-6.3938\frac{e_{\mathrm{w}}}{T_{\mathrm{a}}}+375\,463\frac{e_{\mathrm{w}}}{T_{\mathrm{a}}^{2}}, \tag{42}$$

where P is the air pressure, T_{a} is the air temperature, and e_{w} is the water vapor partial pressure. The relationship between the actual distance and the microwave-measured distance is expressed as follows:

$$l_{\mathrm{re}}=l_{\mathrm{mr}}/n_{\mathrm{a}}+\delta_{\mathrm{mr}}, \tag{43}$$

where l_{re} denotes the actual distance measured using a high-precision laser ranging system in this experiment, l_{mr} represents the microwave-measured distance, n_{a} is the atmospheric refractive index along the measurement path, and δ_{mr} denotes the uncertainty of microwave ranging. By substituting the corrected temperatures listed in Table 6 into the distance correction formula, the distance deviations of the ranging system before and after correction are summarized in Table 7.

Table 7 Atmospheric refraction correction test results in the indoor experiment

Distance (m)	Deviation	
	Before correction (mm)	After correction (μm)
10	3.09	5.25
20	6.13	3.09
30	9.11	7.61

The results indicate that after refractive index correction, the ranging error is reduced from the millimeter level to the micrometer level, demonstrating that temperature prediction accuracy has a direct and significant impact on microwave ranging accuracy.

4.4 Outdoor results

A kilometer-scale baseline experiment is conducted in Hangzhou, China. The atmospheric meteorological parameters along the path are continuously collected and combined with a precision microwave ranging system to model and correct atmospheric refraction errors. The experimental path is gradually extended, with the maximum ranging distance reaching approximately 1000 m.

Due to the limited number of temperature sensors deployed along the field ranging path, it is difficult to obtain a continuous ground-truth temperature field. Therefore, the temperature field must be estimated or reconstructed from sparse observations. In the outdoor experiment, we adopt the standard deviation (STD), which is a commonly used metric in metrology, as the evaluation criterion for temperature field reconstruction (Wu et al., 2017). Specifically, the mean value of the estimated temperatures along the path is first computed, and the temperature STD is then calculated with respect to this mean value to characterize the degree of spatial variation along the path. The path temperature STDs obtained by different methods are presented in Table 8.

Table 8 Comparison of temperature prediction STD among four methods in the outdoor experiment

Distance (m)	Temperature STD (℃)			
	RBF	GP	STK	KFEF
24	0.19	0.12	0.13	0.10
48	0.28	0.21	0.22	0.15
72	0.31	0.23	0.24	0.17
168	0.38	0.29	0.31	0.20
312	0.40	0.34	0.35	0.22
432	0.43	0.37	0.38	0.24
600	0.45	0.39	0.40	0.25
792	0.49	0.36	0.37	0.26
1008	0.51	0.45	0.48	0.27

As shown in Table 8, at short distances, such as 24, 48, and 72 m, where the sensors are relatively densely deployed, the error reduction achieved by KFEF compared with GP and STK is limited; consequently, the performance advantage of KFEF is not significant. However, as the distance increases without a corresponding increase in the number of sensors, the sensor distribution becomes relatively sparse, and the STD achieved by KFEF is significantly lower than those of GP and STK. Specifically, at 24 m, the temperature STD of KFEF is 0.10, representing a 47.37% reduction compared with that of RBF and reductions of 16.67% and 23.08% compared with those of GP and STK, respectively. At 1008 m, the temperature STD of KFEF is 0.27, corresponding to a 47.06% reduction relative to that of RBF and reductions of 40.00% and 43.75% compared with those of GP and STK, respectively. From the above analysis and experimental validation, it can be concluded that under sparse sensor deployment in dynamic outdoor environments, KFEF is still capable of accurately capturing the spatial variability of the temperature field. Although the error is slightly higher than that in stable indoor environments, it remains within acceptable limits for current engineering applications.

Furthermore, a comparative analysis between the corrected ranging results and the reference data with known accuracy shows that the maximum difference is less than 1 mm. This result further validates the effectiveness and engineering value of the proposed atmospheric refraction correction method for long-distance microwave ranging applications.

5 Conclusions

In this paper, we propose a temperature field construction method based on KF-integrated GRNN prediction and Kriging interpolation for data fusion, which is suitable for constructing atmospheric refraction coefficients and various temperature prediction problems. First, a dynamic GRNN-SA optimization algorithm based on minimizing RMSE is proposed, using SA to quickly find the optimal parameters of GRNN. Second, a temperature data filtering system composed of KF, GRNN-SA, and Kriging interpolation is proposed to adapt to dynamic field environments. Finally, to improve the prediction accuracy, a fixed-interval interpolation method is introduced to further expand the temperature dataset. The proposed algorithm continuously corrects temperature measurement data and enriches the dataset using all available data knowledge, reducing measurement errors from temperature sensors and enhancing the accuracy of temperature field predictions. The performance of the proposed method has been validated through simulation data as well as indoor and outdoor evaluations. The results demonstrate that the KFEF-based temperature field construction method can achieve stable and accurate predictions of meteorological temperature data.

While the proposed method exhibits favorable performance in experimental settings, its applicability is contingent on specific conditions. First, the proposed method is founded on the second-order stationarity assumption inherent in ordinary Kriging. In actual atmospheric environments, strong convective activity or localized heat source disturbances may induce pronounced nonstationary characteristics in the temperature field, thereby affecting the accuracy of spatial correlation modeling. The modeling of nonstationary spatial fields has long been recognized as a challenging problem in both geostatistics and atmospheric environmental modeling. In practical applications of atmospheric refraction correction, data acquisition and experiments are typically conducted during periods characterized by relatively stable meteorological conditions. Consequently, this assumption is reasonable and acceptable within the experimental scenarios considered in this study.

Future research will focus on temperature field modeling under complex nonstationary atmospheric conditions, for example, by incorporating universal Kriging models, adaptive variogram estimation methods, and deep learning-based spatiotemporal modeling frameworks, to further enhance the adaptability and robustness of the proposed approach in complex environments.

Author contributions

Ziru LI designed the research, processed the data, and drafted the paper. Zhaobin XU revised the paper. Tao ZHANG and Xinbo YUAN helped organize the paper. Zhonghe JIN finalized the paper.

Conflict of interest

Zhonghe JIN is an editorial board member of *ENGINEERING Information Technology & Electronic Engineering*, and he was not involved with the peer review process of this paper. All the authors declare that they have no conflict of interest.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration on the use of generative AI tools

During the preparation of this work, the authors used ChatGPT to improve language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

Anggraeni W, Yuniarno EM, Rachmadi RF, et al., 2024. A hybrid EMD-GRNN-PSO in intermittent time-series data for dengue fever forecasting. *Expert Syst Appl*, 237: 121438. https://doi.org/10.1016/j.eswa.2023.121438

Bi TF, Li XL, Chen WB, 2024. LiDAR saturated waveform compensation-based real-time ranging method for traffic sign detection. *IEEE Trans Instrum Meas*, 73:8504011. https://doi.org/10.1109/tim.2024.3398092

Chauhan MS, Ojeda-Tuz M, Catarelli RA, et al., 2024. On active learning for Gaussian process-based global sensitivity analysis. *Reliab Eng Syst Saf*, 245:109945. https://doi.org/10.1016/j.ress.2024.109945

Chen KQ, Liu EB, Deng M, et al., 2024. DKNN: deep Kriging neural network for interpretable geospatial interpolation. *Int J Geogr Inform Sci*, 38(8):1486-1530. https://doi.org/10.1080/13658816.2024.2347316

Chen Y, Li JS, Miao DJ, et al., 2018. Development of outdoor baseline environment parameters measurement system based on sensor array. *Acta Metrol Sin*, 39(4): 455-460 (in Chinese). https://doi.org/10.3969/j.issn.1000-1158.2018.04.02

Cheng K, Papaioannou I, Lyu M, et al., 2025. State space Kriging model for emulating complex nonlinear dynamical systems under stochastic excitation. *Comput Methods Appl Mech Eng*, 442:117987. https://doi.org/10.1016/j.cma.2025.117987

Cheng K, Papaioannou I, Lyu M, et al., 2026. State space Kriging model for emulating nonlinear stochastic dynamical systems with parameter uncertainty. *Mech Syst Signal Process*, 243:113691. https://doi.org/10.1016/j.ymssp.2025.113691

Ciddor PE, 1996. Refractive index of air: new equations for the visible and near infrared. *Appl Opt*, 35(9):1566. https://doi.org/10.1364/ao.35.001566

Ding YP, Huang GW, Hu JB, et al., 2020. Indoor target tracking using dual-frequency continuous-wave radar based on the range-only measurements. *IEEE Trans Instrum Meas*, 69(8):5385-5394. https://doi.org/10.1109/tim.2019.2959424

Duff CM, Crawford J, Ip RHL, et al., 2025. Using spacetime geostatistical analysis to improve precipitation isoscape interpolation in Australia. *J Hydrol*, 650:132502. https://doi.org/10.1016/j.jhydrol.2024.132502

Erdogan Erten G, Yavuz M, Deutsch CV, 2022. Combination of machine learning and Kriging for spatial estimation of geological attributes. *Nat Resour Res*, 31(1):191-213. https://doi.org/10.1007/s11053-021-10003-w

Ghritlahre HK, Prasad RK, 2018. Investigation of thermal performance of unidirectional flow porous bed solar air heater using MLP, GRNN, and RBF models of ANN technique. *Therm Sci Eng Prog*, 6:226-235. https://doi.org/10.1016/j.tsep.2018.04.006

Golpîra H, Bevrani H, Golpîra H, 2011. Application of GA optimization for automatic generation control design in an interconnected power system. *Energy Conv Manag*, 52(5):2247-2255. https://doi.org/10.1016/j.enconman.2011.01.010

Gu YY, Jiang LX, Wang L, et al., 2019. Comparative study on thermometry methods of outdoor baseline. *Geom Spat Inform Technol*, 42(12):27-30 (in Chinese). https://doi.org/10.3969/j.issn.1672-5867.2019.12.008

Guillory J, Truong D, Wallerand JP, et al., 2024. A sub-millimetre two-wavelength EDM that compensates the air refractive index: uncertainty and measurements up to 5 km. *Meas Sci Technol*, 35(2):025024. https://doi.org/10.1088/1361-6501/ad0a22

Janowicz K, Gao S, McKenzie G, et al., 2020. GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *Int J Geogr Inform Sci*, 34(4):625-636. https://doi.org/10.1080/13658816.2019.1684500

Katzfuss M, Stroud JR, Wikle CK, 2020. Ensemble Kalman methods for high-dimensional hierarchical dynamic space-time models. *J Am Stat Assoc*, 115(530):866-885.

https://doi.org/10.1080/01621459.2019.1592753

Lee S, Kim SB, 2020. Parallel simulated annealing with a greedy algorithm for Bayesian network structure learning. *IEEE Trans Knowl Data Eng*, 32(6):1157-1166. https://doi.org/10.1109/tkde.2019.2899096

Li G, Liu ZY, Li JG, et al., 2018. Application of general regression neural network to model a novel integrated fluidized bed gasifier. *Int J Hydr Energy*, 43(11):5512-5521. https://doi.org/10.1016/j.ijhydene.2018.01.130

Li S, Griffith DA, Shu H, 2020. Temperature prediction based on a space–time regression-Kriging model. *J Appl Stat*, 47(7):1168-1190. https://doi.org/10.1080/02664763.2019.1671962

Liu EB, Chen KQ, Deng M, 2025. ST-KrigingNet: deep spatiotemporal Kriging neural network for multivariate and nonstationary spatiotemporal interpolation. *Int J Geogr Inform Sci*, 40(6):1807-1839. https://doi.org/10.1080/13658816.2025.2577860

Liu XD, Miao DJ, Zhang JY, et al., 2020. Development of environmental parameters automatic measurement system for 1.2 km standard baseline. *Acta Metrol Sin*, 41(8): 897-902 (in Chinese). https://doi.org/10.3969/j.issn.1000-1158.2020.08.01

Meiners-Hagen K, Meyer T, Mildner J, et al., 2017. SI-traceable absolute distance measurement over more than 800 meters with sub-nanometer interferometry by two-color inline refractivity compensation. *Appl Phys Lett*, 111(19):191104. https://doi.org/10.1063/1.5000569

Miao C, Wang Y, 2024. Interpolation of non-stationary geo-data using Kriging with sparse representation of covariance function. *Comput Geotech*, 169:106183. https://doi.org/10.1016/j.compgeo.2024.106183

Nourani V, Ghaffari A, Behfar N, et al., 2024. Spatiotemporal assessment of groundwater quality and quantity using geostatistical and ensemble artificial intelligence tools. *J Environ Manag*, 355:120495. https://doi.org/10.1016/j.jenvman.2024.120495

Pisani M, Astrua M, Zucco M, 2018. An acoustic thermometer for air refractive index estimation in long distance interferometric measurements. *Metrologia*, 55(1):67-74. https://doi.org/10.1088/1681-7575/aa9a7a

Pollinger F, Meyer T, Beyer J, et al., 2012. The upgraded PTB 600 m baseline: a high-accuracy reference for the calibration and the development of long distance measurement devices. *Meas Sci Technol*, 23(9):094018. https://doi.org/10.1088/0957-0233/23/9/094018

Rooki R, 2016. Application of general regression neural network (GRNN) for indirect measuring pressure loss of Herschel–Bulkley drilling fluids in oil drilling. *Measurement*, 85:184-191. https://doi.org/10.1016/j.measurement.2016.02.037

Rüeger JM, 2002. Refractive index formulae for radio waves. Proc FIG XXII Int Congress, p.1-13.

Shi ZC, Zhou XG, 2021. Spatio-temporal dynamic fields estimating and modeling of missing points in data sets using a flexible state-space model. *Appl Sci*, 11(19): 9050. https://doi.org/10.3390/app11199050

Srinivasan BV, Duraiswami R, Murtugudde R, 2010. Efficient Kriging for real-time spatio-temporal interpolation. Proc 20th Conf on Probability and Statistics in the Atmospheric Sciences, p.228-235.

Tobler WR, 1970. A computer movie simulating urban growth in the Detroit region. *Econ Geogr*, 46(S1):234-240. https://doi.org/10.2307/143141

Tomberg T, Fordell T, Jokela J, et al., 2017. Spectroscopic thermometry for long-distance surveying. *Appl Opt*, 56(2):239-246. https://doi.org/10.1364/ao.56.000239

Underwood R, Gardiner T, Finlayson A, et al., 2015. A combined non-contact acoustic thermometer and infrared hygrometer for atmospheric measurements. *Meteor Appl*, 22(S1):830-835. https://doi.org/10.1002/met.1513

Varouchakis EA, Theodoridou PG, Karatzas GP, 2019. Spatiotemporal geostatistical modeling of groundwater levels under a Bayesian framework using means of physical background. *J Hydrol*, 575:487-498. https://doi.org/10.1016/j.jhydrol.2019.05.055

Wang LC, Kisi O, Zounemat-Kermani M, et al., 2016. Solar radiation prediction using different techniques: model evaluation and comparison. *Renew Sustain Energy Rev*, 61:384-397. https://doi.org/10.1016/j.rser.2016.04.024

Wang SZ, Cao JN, Yu PS, 2022. Deep learning for spatio-temporal data mining: a survey. *IEEE Trans Knowl Data Eng*, 34(8):3681-3700. https://doi.org/10.1109/tkde.2020.3025580

Webster R, Oliver MA, 2007. Geostatistics for Environmental Scientists. John Wiley & Sons, Chichester, UK. https://doi.org/10.1002/9780470517277

Wu HZ, Zhao T, Wang ZY, et al., 2017. Long distance measurement up to 1.2 km by electro-optic dual-comb interferometry. *Appl Phys Lett*, 111(25):251901. https://doi.org/10.1063/1.4999537

Xu XY, Zhang ZQ, Zhang HY, et al., 2020. Long distance measurement by dynamic optical frequency comb. *Opt Expr*, 28(4):4398-4411. https://doi.org/10.1364/oe.381504

Zhao MC, Wang CH, Jin ZH, 2013. Design and derivation of the dual transponder carrier ranging system. *J Zhejiang Univ-Sci C (Comput & Electron)*, 14(5):383-394. https://doi.org/10.1631/jzus.C1200266