Spatiotemporal Multitask Learning for 3-D Dynamic Field Modeling

Di Wang, Kaibo Liu, *Member, IEEE*, Xi Zhang, *Member, IEEE*, and Hui Wang

***Abstract*— 3-D dynamic field modeling using data acquired from sensor networks is typically complex due to the data sparsity and missing problem. In this article, we consider the ubiquitous missing data problem in current sensor networks and aim to take complete advantage of the existing sensor data for thermal field modeling. In the common scenario, data from the target network are not always obtainable, but data from other neighboring networks with homogeneous fields are accessible. Thus, a novel method that captures the information acquired from these neighboring networks is proposed. To achieve accurate thermal field estimation using limited sensor observations, we develop a mixed-effect model framework in which the dynamic field is decomposed into a mean profile and local variability. In particular, we establish a spatiotemporal field multitask learning (FML) approach to identify the spatiotemporal correlation by integrating a multitask Gaussian process (MGP) framework into an autoregressive (AR) model using neighboring data sources from homogeneous fields. Our proposed method is verified through a real case study of thermal field estimation during grain storage.**

***Note to Practitioners*—The proposed method aims to obtain an accurate estimation of a thermal field when certain sensor data are inaccessible. To better implement this method in practice, three things are noteworthy: First, the mean profile of the thermal field should be extracted using the thermodynamic model, so that the remaining data are able to follow a Gaussian process. Second, the FML approach considers neighboring data sources from homogeneous thermal fields to achieve an accurate estimation of the target thermal field. Thus, the target thermal field and other thermal fields should be under similar external conditions, e.g., environmental surroundings, geographical location, and field size. Third, the proposed method can not only process the data from grid-based sensor networks, but also can be extended to other topological structures of sensor networks for field estimation.**

***Index Terms*— Dynamic thermal field, field multitask learning (FML), spatiotemporal dependence.**

Manuscript received October 23, 2018; revised June 5, 2019; accepted September 1, 2019. Date of publication October 25, 2019; date of current version April 7, 2020. This article was recommended for publication by Associate Editor X. Li and Editor K. Saitou upon evaluation of the reviewers' comments. This work was supported in part by the National Science Foundation of China under Grant 71571003, Grant 71771004, and Grant 71690232, and in part by the National Science Foundation under Grant 1637772. *(Corresponding author: Xi Zhang.)*

D. Wang is with the Department of Industrial Engineering and Management, Peking University, Beijing 100871, China, and also with the Dr. Kaibo Liu's Lab, University of Wisconsin–Madison, Madison, WI 53706 USA (e-mail: di.wang@pku.edu.cn).

K. Liu is with the Department of Industrial and Systems Engineering, University of Wisconsin–Madison, Madison, WI 53706 USA (e-mail: kliu8@wisc.edu).

X. Zhang is with the Department of Industrial Engineering and Management, Peking University, Beijing 100871, China (e-mail: xi.zhang@pku.edu.cn).

H. Wang is with the Department of Industrial and Manufacturing Engineering, Florida State University, Tallahassee, FL 32306 USA (e-mail: hwang10@eng.famu.fsu.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASE.2019.2941736

NOMENCLATURE

c	Heat capacity of the grains.
I	Total number of test data at each time point in the target granary.
k	Number of iterations in the EM–LS algorithm.
m	Number of thermal fields.
M	Total number of thermal fields.
n	Number of distinct sensor locations of thermal fields.
N	Number of interested sites.
$\mathbb{R}$	Set of real numbers.
$\mathbf{s}$	Space position.
$\mathbf{S}$	Set of distinct sensor locations in $\{\mathbf{S}_t^m, t = t_0, \ldots, t_T, m = 1, \ldots, M\}$.
$\boldsymbol{\mathcal{S}}$	Set of interested sites.
t	Time point.
T	Total number of time points.
(x, y, z)	Space position in a Cartesian coordinate system.
δ	Range parameter that corresponds to the distance around a thermal field.
$\boldsymbol{\kappa}$	Kernel matrix between $\mathbf{S}$ and $\boldsymbol{\mathcal{S}}$ obtained by the kernel function.
π	Precision of the hyperprior distribution for $\boldsymbol{\mu}_t$.
ρ	Density of the grains.
τ	Precision of the hyperprior distribution for $\mathbf{C}_t$.
Ψ	Set of hyperprior parameters.
ϵ	Predetermined threshold.
$b^m(\mathbf{s}, t)$, $b_t^m(\mathbf{s})$	Bias of the thermal field m at location $\mathbf{s}$ and time t.
$b_{t-l}^m(\mathbf{s})$	Bias of the thermal field m at location $\mathbf{s}$ and time $t-1$.
$\hat{b}_{t,k}^m(\mathbf{s})$	Estimate of $b_t^m(\mathbf{s})$ in iteration k.
$\hat{b}_{t-l,k}^m(\mathbf{s})$	Estimate of $b_{t-l}^m(\mathbf{s})$ in iteration k.
$b'^m_{t,k}(\mathbf{s})$	Processed data of the thermal field m at location $\mathbf{s}$ and time t in iteration k.
$\mathbf{C}_t$	Covariance matrix of the Gaussian process $\boldsymbol{\alpha}_t^m$.
$\mathbf{C}_{t,k}$	Estimate of $\mathbf{C}_t$ in iteration k.
$\mathbf{C}_{t,k}^m$	Estimate of the covariance matrix in terms of $\boldsymbol{\alpha}_{t,k}^m$.
$\mathbf{C}_{wt}$	Covariance matrix of the Gaussian process $\mathbf{w}_t^m$.
$\mathbf{D}_t^m$	Accessible data set of thermal field m at time t that is composed of $\mathbf{S}_t^m$ and $\mathbf{g}_t^m$.
$g^m(\mathbf{s}, t)$	Deviance of the sensor data from the mean profiles of the thermal field m.

$\mathbf{g}_t^m$	Deviance vector of the sensor data from the mean profiles of the thermal field m.
n_t^m	Number of accessible sensor data of the thermal field m at time t.
$\mathbf{s}_i$	Space position with index i.
$\mathbf{S}_t^m$	Set of space positions of the thermal field m at time t.
t_0	Initial time point.
t_T	Last time point.
$\mathrm{tr}(\cdot)$	Trace of a matrix.
$u_{t,k}^m(\mathbf{s})$	Difference between $b'^m_{t,k}(\mathbf{s})$ and $\hat{w}^m_{t,k-1}(\mathbf{s})$.
$v_{t,k}^m(\mathbf{s})$	Deviance between the bias of the thermal field m at time $t b'^m_{t,k}(\mathbf{s})$ and those at previous L time points $\sum_{l=1}^{L} \hat{\beta}^m_{l,k} b'^m_{t-l,k}(\mathbf{s})$.
$\mathbf{v}_{t,k}^m$	Vector of $\{v_{t,k}^m(\mathbf{s})\}, \forall \mathbf{s} \in \mathbf{S}_t^m$.
$w_t^m(\mathbf{s})$	Variation of the local variability of the thermal field m at location $\mathbf{s}$ and time t.
$\hat{w}_{t,k}^m(\mathbf{s})$	Estimate of $w_t^m(\mathbf{s})$ in iteration k.
$\mathbf{w}_t^m$	$N \times 1$ vector in terms of $W_t^m(\mathbf{s})$
$y^m(\mathbf{s}, t)$	Temperature of the thermal field m at location $\mathbf{s}$ and time t.
$\hat{y}^m(\mathbf{s}, t)$	Estimate of $y^m(\mathbf{s}, t)$.
$\boldsymbol{\alpha}_t^m$	Vector of the weight parameters for thermal field m at time t.
$\hat{\boldsymbol{\alpha}}_{t,k}^m$	Estimate of $\boldsymbol{\alpha}_t^m$ in iteration k.
$\alpha_{t,i}^m$	ith element of $\boldsymbol{\alpha}_t^m$.
$\hat{\alpha}_{t,i,k}^m$	Estimate of $\alpha_{t,i}^m$ in iteration k.
β_l^m	lth parameter associated with the temporal effects from time $t-l$.
$\hat{\beta}_{l,k}^m$	Estimate of β_l^m in iteration k.
$\boldsymbol{\beta}^m$	Vector of the AR model parameters of the thermal field m.
$\Delta\boldsymbol{\beta}_k^m$	Change in $\boldsymbol{\beta}^m$ between time t and time $t-1$ in iteration k.
$\varepsilon^m(\mathbf{s}, t)$	Model error of the thermal field m at location $\mathbf{s}$ and time t.
$\kappa(\mathbf{s}_i, \mathbf{s})$	Kernel function in terms of locations $\mathbf{s}_i$ and $\mathbf{s}$.
$\boldsymbol{\kappa}_0$	Kernel matrix among $\mathbf{S}$.
$\boldsymbol{\kappa}_m$	Kernel matrix between $\mathbf{S}_t^m$ and $\mathbf{S}$ obtained by the kernel function.
$\lambda_x, \lambda_y, \lambda_z$	Thermal conductivity of the grains in the x, y, and z directions, respectively.
$\mu^m(\mathbf{s}, t)$	Mean function of the thermal field m at location $\mathbf{s}$ and time t.
$\boldsymbol{\mu}_t$	Mean vector of the Gaussian process $\boldsymbol{\alpha}_t^m$.
$\boldsymbol{\mu}_{t,k}$	Estimate of $\boldsymbol{\mu}_t$ in iteration k.
$\boldsymbol{\mu}_{wt}$	Mean vector of the Gaussian process $\mathbf{w}_t^m$.
σ_t^2	Variance of the normal distribution for $\varepsilon^m(\mathbf{s}, t)$.
$\sigma_{t,k}^2$	Estimate of σ_t^2 in iteration k.
Φ_t	Set of the MGP model parameters at time t.
$\hat{\Phi}_{t,k}$	Estimate of Φ_t in iteration k.

I. Introduction

DYNAMIC fields, such as electric, magnetic, and thermal fields, widely exist in various dynamic systems and are critical for the operation, design, and maintenance of a system. The appropriate modeling of such fields provides good opportunities to understand the system mechanism and obtains useful information for potential improvement of a system. Successful examples can be found in a variety of domains, including ecology [1], [2], healthcare [3], and meteorology [4].

A thermal field is a classic dynamic field that exists in a variety of engineering systems, and temperature is typically regarded as an important index to indicate system performance. For example, in a grain storage depot, the local temperature in barns can be affected by a number of external factors (e.g., ambient temperature and airiness condition) and internal factors (e.g., grain self-breath and mildew), thereby leading to considerable spatial and temporal temperature variations during grain storage. Modeling thermal distribution by relying only on expertise in thermodynamics does not work efficiently due to lack of understanding regarding the dynamics of grain activities. Consequently, obtaining a precise thermal distribution of grains in certain complex systems is critically important. This is because the change in the thermal distribution may result in grain quality deterioration and excessive energy consumption in warehouse ventilation. A similar problem is also often observed in the cooling systems of nuclear plants [5] and supercomputer centers [6].

A thermal field can be generally considered a 3-D spatiotemporal stochastic system that is represented by time-varying temperatures at each location. The accurate estimation of a thermal field remains a challenging task because capturing spatiotemporal dynamics in a 3-D thermal field involves a large number of uncertainties. Conventional field modeling techniques obtain the spatiotemporal dynamics of a thermal field by establishing thermodynamic models, which merely consider the aforementioned extrinsic factors and employ physical principles (such as thermal mechanism) to describe a field. For example, nonlinear heat transmission models [7], [8] were proposed to describe the thermal field in a cuboid-shaped granary. A hydromechanical model was developed to characterize the air flow and a dynamic thermal field [9]. An improved heat transmission model was proposed by considering the self-heating property of grains to improve the field estimation performance [10]. These methods successfully estimate thermal fields under ideal conditions, i.e., no other unexpected changes occur in the system. In practice, however, uncertainties caused by internal factors also exist, which limits these conventional methods, leading to a large discrepancy between the estimated and actual thermal fields. Therefore, to acquire an accurate thermal field through a new modeling technique is quite desirable for field estimation and system improvement.

The remainder of this article is organized as follows. Section II provides the literature review on dynamic field modeling by using sensor data. Section III introduces the methodology of thermal field modeling and elaborates two major components of the mixed-effect model, i.e., modeling the mean profile and local variability, using the grain storage example. Section IV presents a real case study of grain thermal fields to evaluate the performance of the proposed method.

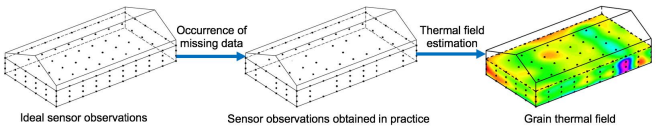


Fig. 1. Sensor observations from the grid-based sensor networks of a granary with missing data.

Section V provides conclusive remarks and suggestions for future research.

II. Literature Review

A large number of modeling techniques have emerged with the advancement of information and sensor technologies, which provide the infrastructural support to generate sensor observations for dynamic field modeling [11]–[13]. Data-driven approaches have been vastly developed to model the dynamic field based on sensor observations obtained from sensor networks [14], [15]. These approaches aim to characterize dynamic fields by considering spatiotemporal correlations using sensor observations. Pioneering studies have modeled the spatiotemporal correlations of dynamic fields by assuming that spatial and temporal components are independent of each other [16], [17]. For recent studies, Katzfuss and Cressie [18] proposed a spatiotemporal smoothing approach that independently represents spatial and temporal processes using remote-sensing data sets. Zheng *et al.* [4] modeled the spatiotemporal dependence of air quality by considering two independent classifiers, namely spatial and temporal classifiers. However, the model framework with independent spatial and temporal components cannot effectively capture the spatiotemporal interactions of dynamic fields.

To address this problem, studies have attempted to model spatiotemporal interactions by simultaneously considering spatial, temporal, and spatiotemporal correlations [19]. In recent years, kriging has elicited considerable attention in dynamic field modeling and has been widely used to model spatiotemporal interactions [20], [21]. However, the computational issue generally arises in the kriging method for large training data sets. Consequently, many researchers have made their efforts in modeling covariance functions to reduce computational complexity; these studies include sparse matrix algorithms [23], reduced rank techniques [24], and Gaussian random fields [25]. For example, Wang *et al.* [26] integrated a kriging model into a Gaussian Markov random field model by fully using grid-based sensor data, which reduced the computational burden and provided an acceptable performance in thermal distribution monitoring. The aforementioned methods require sufficient available sensor observations to achieve a balance between the field estimation accuracy and the computational burden. However, one of the major challenges in thermal field modeling is that limited sensor observations are available because only sparse sensors are deployed in sensor networks, given the high costs of sensors and limited budget (the distance between two adjacent sensors is set to be 5 m in a granary, as shown in Fig. 1). In addition, sensor observations may not be gathered appropriately into a single database for unexpected reasons, such as sensor aging, wireless communication failures, and data reading errors. Therefore, developing an effective method for spatiotemporal thermal field modeling when only a limited number of sensor observations are accessible is essential.

Fig. 2. Aerial view of a grain depot with multiple homogeneous granaries in the central part of China.

To solve this problem, researchers have adopted interpolation methods to fill in missing data using the existing sensor observations of thermal fields, including linear, spline [27], and Lagrange interpolations [28]. However, this strategy may cause a large bias and is inclined to lose the chance to capture local changes in a dynamic field, particularly when certain data from the target sensor network are missing while other neighboring networks with homogeneous fields are accessible. Hence, other researchers instead tend to borrow data from homogeneous fields to fill in the missing data for target dynamic fields. For example, missing thermal data in a granary are filled directly with sensor data from similar granaries. However, interpolation methods and the direct data filling strategy from other data pools lack substantial consideration of the dynamics of a thermal field and data uncertainties, e.g., varieties of data dependence may exist at different locations and time points.

Transfer learning provides an opportunity for thermal field modeling with limited sensor observations and for addressing the issue of data uncertainty; it aims to improve the learning of unobserved values in the target source by sharing knowledge or information from related data sources [29]–[32]. In recent years, transfer learning has been studied by numerous researchers and applied to various engineering domains, including WiFi localization [33], speech emotion recognition [34], and manufacturing shape deviation [35]. Multitask learning has emerged as one of the popular focuses for transfer learning problems. Compared with single-task learning (STL) [36], [37], which learns "knowledge" by using the training data in the target task, multitask learning is a machine learning framework that aims to improve the model performance of a target task by learning multiple similar-but-not-identical tasks (e.g., thermal field estimation) and then sharing the information of each task. Here, we use grain storage as an example, as shown in Fig. 2. Several columns of granaries with similar storage conditions (within the red boxes), including the

location, volume of granaries, stored grains, and environmental surroundings, are found in this grain depot. The thermal field estimation of a target granary can be obtained through the concept of multitask learning, which leverages information available from sensor observations collected from other homogeneous granaries.

Multitask learning has attracted considerable research interest within a decade. For example, Jin and Sun [38] combined multitask learning and neural networks to predict traffic flows. Huang *et al*. [39] proposed a deep architecture that combined a deep belief network with multitask learning for traffic flow prediction. Xu *et al*. [40] developed a least squares (LS) regularized regression algorithm for multi-task learning with Gaussian kernels. Unfortunately, these models fail to efficiently capture the spatiotemporal information in dynamic systems. Many researchers have also adopted multitask learning for complex spatial and temporal systems. Guo and Chen [41] applied multitask learning to human action recognition. Their method represents the spatiotemporal features of each human action independently but does not consider spatiotemporal correlations among human actions. Goncalves *et al*. [42] proposed a hierarchical multitask learning method for climate prediction. Their method can estimate the spatiotemporal correlation of climate, but is not applicable to high-dimensional spatiotemporal data using limited sensor observations, e.g., estimating a 3-D spatiotemporal thermal field. Shao *et al*. [43] developed a multitask learning model to estimate a 2-D-machined surface shape using limited sensor observations from homogeneous data sources. This model considers the spatial correlation of the surface shape and improves the modeling accuracy based on the sensor data for related surface shapes. The study of [43] mainly focused on static spatial surfaces, but failed in modeling spatiotemporal dynamic fields that vary across space and time.

In conclusion, the existing spatiotemporal models require sufficient observations in the spatial and temporal domains to estimate dynamic fields accurately. However, only sparse sensor observations in granaries are collected, and some sensor observations may be missing for unexpected reasons, thereby leading to poor field estimation results using these existing spatiotemporal models. Multitask learning provides an opportunity for thermal field modeling with limited sensor observations. However, the existing approaches for multitask learning have gaps in dynamic field modeling. First,multitask learning approaches have been widely used to model profiles or surfaces, but few studies have considered 3-D field modeling. Second, the existing multitask learning approaches are purely data driven. These purely data-driven approaches cannot achieve accurate field estimation results; useful engineering information must be considered in the modeling of 3-D dynamic fields. Third, the existing works mainly focus on the modeling of static systems. They mainly consider "borrowing" spatial information, but disregard the temporal correlation of a system. Few studies have addressed the issue of dynamic systems that vary over time using multitask learning.

To fill in the research gap, we propose a 3-D thermal field estimation method and model spatiotemporal dynamics of a thermal field using limited sensor observations from several homogeneous data sources. To achieve accurate thermal field estimation using limited sensor observations, we develop a mixed-effect model framework in which the dynamic field is decomposed into a mean profile and a local variability. We adopt a thermodynamic model to capture the mean profile of a thermal field by leveraging the physical mechanism of heat transmission. To fully utilize all the accessible sensor data, we develop a spatiotemporal field multitask learning (FML) approach to characterize local variability by considering missing data and data uncertainties. In particular, we establish the FML to capture spatiotemporal correlation by integrating a multitask Gaussian process (MGP) model into an autoregressive (AR) model with time series using neighboring data sources from homogeneous fields.

Compared with the existing spatiotemporal models, the proposed method addresses the challenges of 3-D thermal field estimation using spatiotemporal multitask learning in the following aspects. First, we propose an FML method to characterize the local variability of a 3-D thermal field by fully utilizing limited sensor observations from homogeneous data sources. Second, compared with purely data-driven methods, our proposed method leverages the physical mechanism of thermal conductivity and uses a thermodynamic model to capture the mean profile of a thermal field. We combine the thermodynamic model and data-driven methods to achieve accurate thermal field estimation. Third, the MGP and the time series method are simultaneously considered to effectively capture the spatiotemporal correlations of dynamic thermal fields.

III. Methodology

A mixed-effect model framework that incorporates thermal physics and information acquired from sensor networks using FML is proposed to obtain a dynamic field. Fig. 2 shows that the granaries are arranged using columns with identical environmental conditions. This situation naturally provides a good opportunity to capture the available sensor data in neighboring granaries to support the thermal field estimation of the target granary. Without loss of generality, we assume that M 3-D spatiotemporal thermal fields with similar conditions exist. As shown in Fig. 3, these M 3-D thermal fields present different missing data patterns. We regard the thermal field m as the target data source and estimate it by considering the sensor data in M thermal fields. For the thermal field $m (m = 1, \ldots, M)$, let $y^m(\mathbf{s}, t)$ be the response variable observed at location $\mathbf{s}$ and time t, where $\mathbf{s} \in \mathbb{R}^3$ and $t \in \mathbb{R}^+$. The underlying dynamic field can be represented as follows:

$$y^m(\mathbf{s},t) = \mu^m(\mathbf{s},t) + b^m(\mathbf{s},t) + \varepsilon^m(\mathbf{s},t) \tag{1}$$

where $\mu^m(\mathbf{s}, t)$ is the mean function of the thermal field m. The residual after the mean function is decomposed into two independent parts, namely the corresponding bias $b^m(\mathbf{s}, t)$, which captures the local variability of the thermal field, and Gaussian white noise $\varepsilon^m(\mathbf{s}, t)$, which captures model errors. Here, we assume that $\varepsilon^m(\mathbf{s}, t)$ is independent at each time point. For each $\varepsilon^m(\mathbf{s}, t), m = 1, \ldots, M$, $\varepsilon^m(\mathbf{s}, t)$ follows a normal distribution at time t with variance σ_t^2, that is $\varepsilon^m(\mathbf{s}, t) \sim N(0, \sigma_t^2)$.

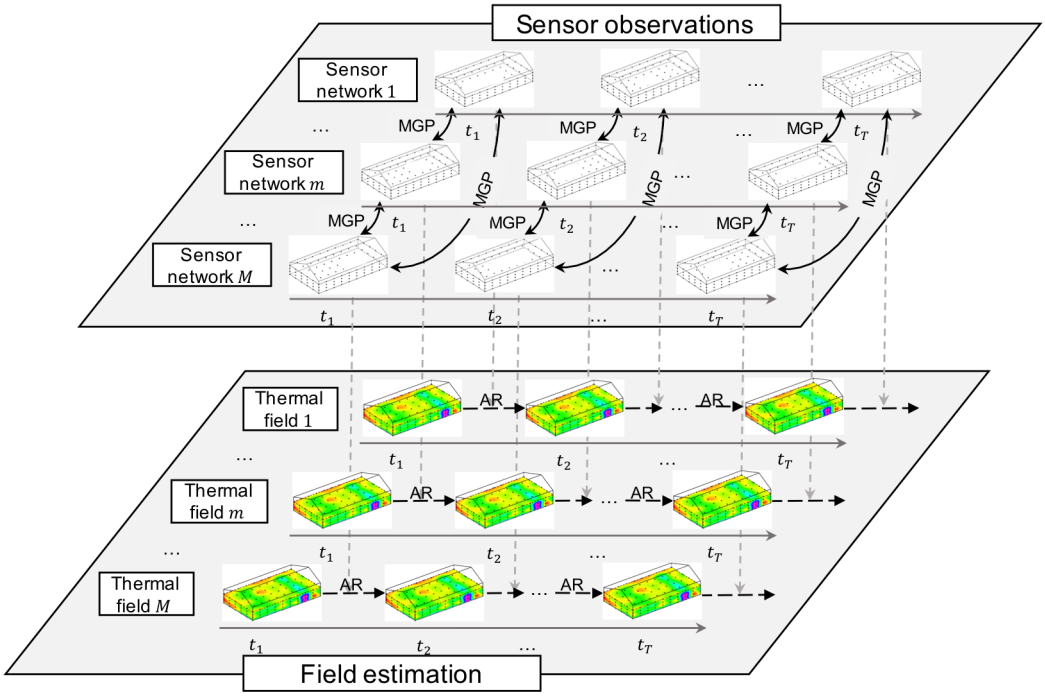


Fig. 3. Main concept of research methodology. Dynamic field estimation for 3-D thermal fields by integrating the MGP model and the AR model based on sparse sensor data with missing data problems from homogeneous sensor networks.

A. Modeling Mean Profile

We adopt a thermodynamic model as a mean function to model mean profiles of thermal fields following thermal physics knowledge. The thermodynamic model has been widely used to estimate thermal fields under ideal conditions. In particular, we employ a 3-D unsteady heat transfer function as the thermodynamic model to characterize the spatiotemporal mean profiles. This function is a second-order differential equation, which describes the heat distribution (or temperature variation) in a given space over time. For the thermal field m, the location $\mathbf{s}$ in (1) is denoted (x, y, z) in a Cartesian coordinate system, and the 3-D unsteady heat transfer function is given as follows:

$$\rho c \frac{\partial \mu^m(x,y,z,t)}{\partial t} = \lambda_x \frac{\partial^2 \mu^m(x,y,z,t)}{\partial x^2} + \lambda_y \frac{\partial^2 \mu^m(x,y,z,t)}{\partial y^2} + \lambda_z \frac{\partial^2 \mu^m(x,y,z,t)}{\partial z^2} \quad (2)$$

where $\mu^m(x, y, z, t)$ denotes the temperature response of the thermal field m. The parameter set $\boldsymbol{\theta} = \rho, c, \lambda_x, \lambda_y, \lambda_z$ is determined based on the properties of the target thermal field. For stored grains, ρ denotes the density of grains, c denotes the heat capacity of grains, and $\lambda_x, \lambda_y, \lambda_z$ denote the thermal conductivity of grains in the x, y, and z directions, respectively.

Given the temperature at time t_0 as the initial values and the temperature at the walls of the granary as the boundary conditions, we use a finite difference method to solve (2). Notably, the boundary conditions change over time and are affected by environmental factors, which in turn significantly affect the temperature of granary walls and further affect the temperature of inner grains (for detailed analysis regarding these effects, please refer to [8]).

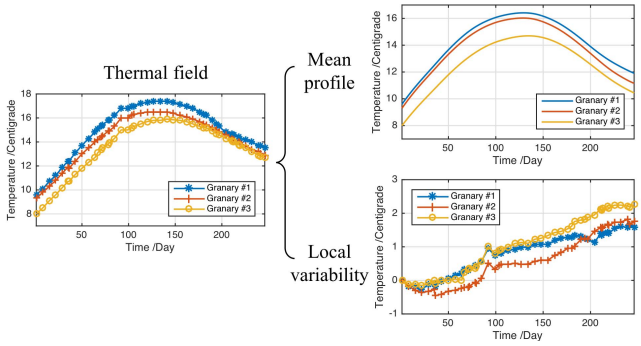


Fig. 4. Anatomy of a thermal field. (A thermal field is decomposed into the mean profile and the local variability.)

B. Modeling Local Variability

The local variability of a thermal field is affected by various latent intrinsic factors and cannot be easily characterized by an existing known deterministic model. Two variables at adjacent locations tend to exhibit a strong correlation when the distance between their locations is small in a thermal field. Thermal fields with similar conditions share common spatiotemporal structures. Notably, only the deviance of the limited sensor data from the mean profiles of a thermal field $g^m(\mathbf{s}, t) = y^m(\mathbf{s}, t) - \mu^m(\mathbf{s}, t)$ is known. To effectively estimate the local variability of a thermal field, spatiotemporal correlations should be captured by simultaneously considering multiple sensor observations from homogeneous thermal

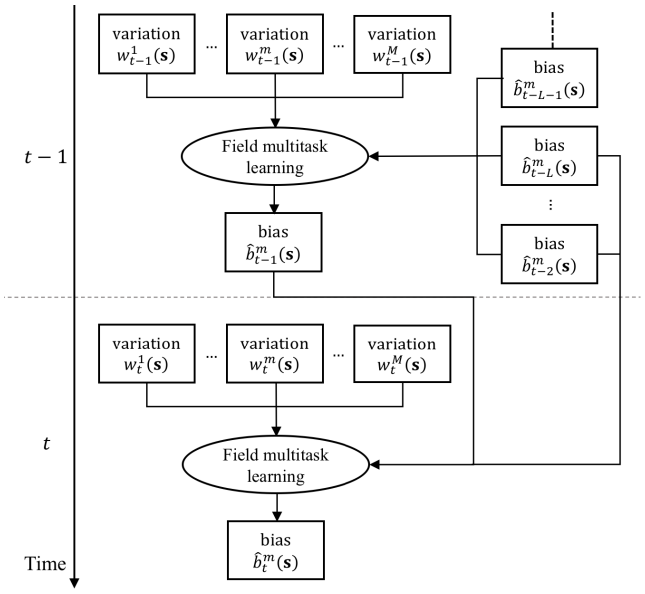


Fig. 5. Procedure for modeling the local variability of the thermal field m based on FML.

fields. Fig. 4 presents a demonstrative example of thermal sensor observations at the same location in three adjacent granaries with similar storage conditions. The curves of the mean profile are obtained using the thermodynamic model and the local variabilities marked with dots are the values of the deviance of the sensor observations from the mean profile. The local variability in each granary varies and exhibits considerable spatial and temporal correlations.

We propose an FML method to describe the local variability, which integrates the MGP model into an AR model to characterize spatiotemporal correlation by considering temperature from neighboring fields at previous L time points. For the thermal field m

$$b_t^m(\mathbf{s}) = \sum_{l=1}^{L} \beta_l^m b_{t-l}^m(\mathbf{s}) + w_t^m(\mathbf{s}) \tag{3}$$

where the bias $b_t^m(\mathbf{s})$ denotes the local variability at location $\mathbf{s}$ and time t of the thermal field m, which is a simplified notation of $b^m(\mathbf{s}, t)$. β_l^m is the lth parameter associated with the temporal effects from the $t-l$ time point. $w_t^m(\mathbf{s})$ corresponds to the spatial effects that are used to model the variation in location $\mathbf{s}$ at time t. We consider $w_t^m(\mathbf{s})$ as a Gaussian process and use an MGP model to characterize the spatial correlation between the thermal field m and other $M-1$ thermal fields at time t using multiple sensor observations from these M homogeneous thermal fields.

Fig. 5 shows the procedure of the FML method. The bias $b_t^m(\mathbf{s})$ is composed of a weighted combination of the bias from its previous L time points and the variation term of the local variability at time t. As the flowchart shows, at time $t-1$, we consider the bias from time $t-L-1$ to time $t-2$ as well as the variation terms at time $t-1$ of all the M homogeneous fields and use the proposed FML model to estimate the bias at time $t-1$. Similarly, at time t, we consider the bias from time $t-L$ to time $t-1$, as well as the variation terms at time t of all the M homogeneous fields, and use FML to estimate the bias at time t. We can see that the FML model considers data from homogeneous fields and also characterizes spatiotemporal correlations.

1) MGP for $w_t^m(\mathbf{s})$: For the thermal field m, recall that we can only observe the deviance between the limited sensor data and the thermodynamic model. This condition poses a considerable challenge to accurately estimate the local temperature variation if nearby sensors are inaccessible. To address this issue, we consider the estimation of M related functions in terms of local variability based on the observed deviance of M similar thermal fields, which substantially increases the size of the available data set.

At time $t(t = t_0, \ldots, t_T)$, we represent the accessible data for the thermal field m as $\mathbf{D}_t^m = \{\mathbf{S}_t^m, \mathbf{g}_t^m\}$, where $\mathbf{S}_t^m$ denotes the location and $\mathbf{g}_t^m$ denotes the observed deviance. Each thermal field processes accessible sensor data with different sites. Thus, a total of n distinct sensor locations exist in $\{\mathbf{S}_t^m, t = t_0, \ldots, t_T, m = 1, \ldots, M\}$ with $\max\{n_t^m, t = t_0, \ldots, t_T, m = 1, \ldots, M\} \le n \le \sum_{m=1}^{M} \sum_{t=t_0}^{t_T} n_t^m$, where n_t^m indicates the number of accessible sensor data for the thermal field m at time t. We denote $\mathbf{S}$ as the set of distinct sensor locations in $\{\mathbf{S}_t^m, t = t_0, \ldots, t_T, m = 1, \ldots, M\}$.

We assume $\mathbf{w}_t^m = \{w_t^m(\mathbf{s}), \mathbf{s} \in \boldsymbol{\mathcal{S}} \subset \mathbb{R}^3\}$, is the $N \times 1$ vector of the variation term, with $m = 1, \ldots, M$, where $\boldsymbol{\mathcal{S}}$ denotes a set of interested sites among thermal fields and N denotes the number of interested sites, to be a Gaussian process $\mathbf{w}_t^m \sim N(\boldsymbol{\mu}_{wt}, \mathbf{C}_{wt})$, which is characterized by a mean vector $\boldsymbol{\mu}_{wt}$ and a covariance matrix $\mathbf{C}_{wt}$. To capture the similarities of local variability among granaries, we assume that $\mathbf{w}_t^m$ shares a common structure by modeling the parameters of the Gaussian process $\boldsymbol{\mu}_{wt} = \boldsymbol{\kappa}^T \boldsymbol{\mu}_t$ and $\mathbf{C}_{wt} = \boldsymbol{\kappa}^T \mathbf{C}_t \boldsymbol{\kappa}$, where $\boldsymbol{\kappa}$ denotes the $n \times N$ kernel matrix obtained by the kernel function. For $\forall \mathbf{w}_t^m$ with $m = 1, \ldots, M$, there exists a unique $\boldsymbol{\alpha}_t^m$, such that $\mathbf{w}_t^m = \boldsymbol{\kappa}^T \boldsymbol{\alpha}_t^m$. Here, $\boldsymbol{\alpha}_t^m$ denotes a vector of the weight parameter for thermal field m and $\boldsymbol{\alpha}_t^m \sim N(\boldsymbol{\mu}_t, \mathbf{C}_t)$, where $\boldsymbol{\mu}_t$ denotes the mean vector of $\boldsymbol{\alpha}_t^m$ and $\mathbf{C}_t$ denotes the covariance matrix of $\boldsymbol{\alpha}_t^m$. To obtain the maximum likelihood estimates of $\boldsymbol{\mu}_t$ and $\mathbf{C}_t$, we describe the hyperprior distribution of $\boldsymbol{\mu}_t$ and $\mathbf{C}_t$ via the normal-inverse Wishart distribution, which is the conjugate prior for the multivariate Gaussian distribution

$$\{\boldsymbol{\mu}_t, \mathbf{C}_t\} \sim N(\boldsymbol{\mu}_t | 0, \frac{1}{\pi} \mathbf{C}_t) IW(\mathbf{C}_t | \tau, \boldsymbol{\kappa}^{-1})$$

where $\boldsymbol{\mu}_t$ is specified by a prior mean 0 and a covariance matrix $\mathbf{C}_t$ with precision π and $\mathbf{C}_t$ is specified by the kernel matrix $\boldsymbol{\kappa}$ with precision τ. Thus, the MGP framework is described in detail as follows.

Step 1: We initiate $\boldsymbol{\mu}_t$ and $\mathbf{C}_t$ by following the normal-inverse Wishart distribution.

Step 2: For each thermal field $m, m = 1, \ldots, M$, we obtain $\boldsymbol{\alpha}_t^m$ by $\boldsymbol{\alpha}_t^m \sim N(\boldsymbol{\mu}_t, \mathbf{C}_t)$.

Step 3: Given $\mathbf{s}_i \in \mathbf{S}, \mathbf{s} \in \boldsymbol{\mathcal{S}} \subset \mathbb{R}^3$, we obtain $w_t^m(\mathbf{s}) = \sum_{i=1}^{n} \alpha_{t,i}^m \kappa(\mathbf{s}_i, \mathbf{s})$, where $\alpha_{t,i}^m$ denotes the ith element of $\boldsymbol{\alpha}_t^m$ and $\kappa(\mathbf{s}_i, \mathbf{s})$ denotes the kernel function in terms of locations $\mathbf{s}_i$ and $\mathbf{s}$.

Various types of kernels can be used as candidates for the MGP model. To effectively capture the spatial dynamics of

a thermal field, we select a spatial covariance kernel that describes the spatial correlation on the basis of the Euclidean distance between two different locations of the thermal field. The spatial covariance kernel is extensively used in grain thermal fields to characterize the spatial correlation patterns of local variability among granaries as follows:

$$\kappa(\mathbf{s}_i, \mathbf{s}) = \exp\left(-\frac{\|\mathbf{s}_i - \mathbf{s}\|^2}{\delta^2}\right) \tag{4}$$

where $\mathbf{s}_i$ and $\mathbf{s}$ denote two different locations and δ denotes the range parameter that corresponds to the distance around a thermal field. The range parameter can be determined on the basis of the maximum length of a thermal field.

This MGP model based on $\boldsymbol{\mu}_t$ and $\mathbf{C}_t$ has two advantages. First, the MGP model is convenient for parameter estimation. We develop an iterative EM–LS algorithm that integrates an expectation–maximum (EM) algorithm into the LS method and iteratively estimates parameters. The detailed information is expressed in Section III-B2. Second, we realize $\mathbf{w}_t^m$ by a spatial covariance kernel $\boldsymbol{\kappa}$, which can characterize the nonlinear spatial correlation patterns in a thermal field.

Algorithm 1 Iterative EM–LS Algorithm for Parameter Estimation

Input: deviance data of homogeneous fields
Output: estimated parameters in AR model $\hat{\boldsymbol{\beta}}^m$, with $m = 1, \ldots, M$; estimated parameters in MGP model $\hat{\Phi}_t = \left\{\hat{\boldsymbol{\mu}}_t, \hat{\mathbf{C}}_t, \hat{\sigma}_t^2\right\}$ and $\hat{\boldsymbol{\alpha}}_t^m$ for $t = t_0, \ldots, t_T$, with $m = 1, \ldots, M$.

1: **Procedure:**
2: Set $\hat{w}_{t,0}^m(\mathbf{s}) = 0$ and $k = 1$.
3: Use Gaussian process to fill up missing values and obtain $\{b'^m_{t,k}(\mathbf{s})\}, t = t_0, \ldots, t_T$.
4: **Do**
5: AR model: obtain $u_{t,k}^m(\mathbf{s}) = b'^m_{t,k}(\mathbf{s}) - \hat{w}_{t,k-1}^m(\mathbf{s})$, and apply the least squares method to estimate $\hat{\boldsymbol{\beta}}_k^m$ using $\{u_{t,k}^m(\mathbf{s}), b'^m_{t-1,k}(\mathbf{s}), \ldots, b'^m_{t-L,k}(\mathbf{s})\}$.
6: MGP model: obtain $v_{t,k}^m(\mathbf{s}) = b'^m_{t,k}(\mathbf{s}) - \sum_{l=1}^L \hat{\beta}_{l,k}^m b'^m_{t-l,k}(\mathbf{s})$ for $t = t_0, \ldots, t_T$, and apply the EM algorithm to estimate $\hat{\Phi}_{t,k}$ and $\hat{\boldsymbol{\alpha}}_{t,k}^m$ using $\{v_{t,k}^m(\mathbf{s})\}$.
7: Set $k = k + 1$.
8: Field estimation: fill up missing values using $\hat{\boldsymbol{\beta}}_{k-1}^m$ and $\hat{\boldsymbol{\alpha}}_{t,k-1}^m$ and obtain$\{b'^m_{t,k}(\mathbf{s})\}$, $t = t_0, \ldots, t_T$.
9: **Stop** the loop until $\Delta\boldsymbol{\beta}_k^m = \left\|\hat{\boldsymbol{\beta}}_k^m - \hat{\boldsymbol{\beta}}_{k-1}^m\right\|$ converges to 0.
10: Set $\hat{\boldsymbol{\beta}}^m = \hat{\boldsymbol{\beta}}_k^m$, $\hat{\Phi}_t = \hat{\Phi}_{t,k}$, and $\hat{\boldsymbol{\alpha}}_t^m = \hat{\boldsymbol{\alpha}}_{t,k}^m$ for $t = t_0, \ldots, t_T$, with $m = 1, \ldots, M$.

2) Parameter Estimation for the FML Model: In the proposed FML model, we assume that the parameters of all M bias functions in (3) share a common mean and a common covariance matrix of the Gaussian process, because the considered dynamic fields have a nearly identical environment and same operating conditions (e.g., same type of grains stored in granaries). Nevertheless, the associated parameters from different thermal fields vary. We capture the dependence among bias functions based on the corresponding sensor data in all accessible sensor networks.

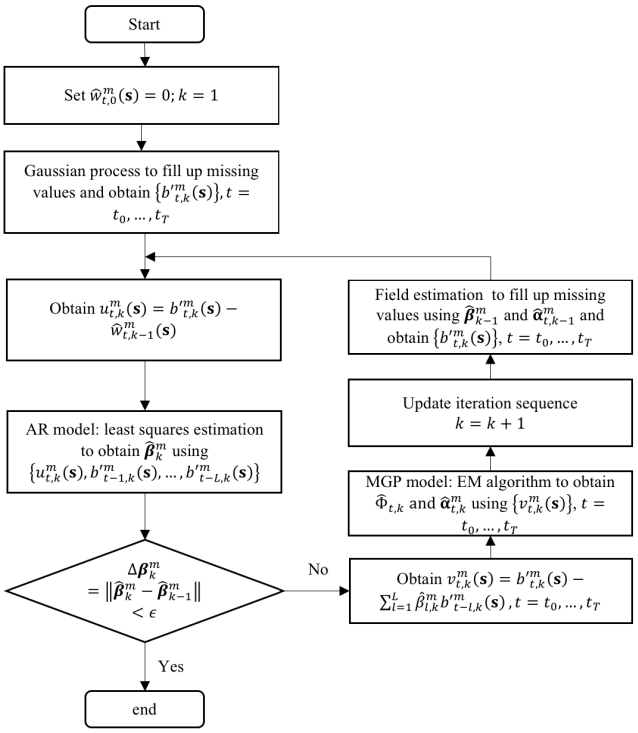


Fig. 6. Iterative EM–LS algorithm for parameter estimation.

The proposed FML model contains a set of parameters: hyperprior parameters $\Psi = \{\pi, \tau, \delta\}$; parameters of the MGP model $\Phi_t = \{\boldsymbol{\mu}_t, \mathbf{C}_t, \sigma_t^2\}, t = t_0, \ldots, t_T$; and parameters of the AR model $\boldsymbol{\beta}^m = \{\beta_1^m, \beta_2^m, \ldots, \beta_L^m\}, m = 1, \ldots, M$. Hyperprior parameters Ψ can be determined via expert knowledge of thermal fields, and small π and τ values generally provide good estimation performance [43]. Meanwhile, δ can be determined based on the maximum length of a thermal field. In addition, two types of parameters, namely$\{\Phi_t\}_{t=t_0}^{t_T}$ and $\{\boldsymbol{\beta}^1, \ldots, \boldsymbol{\beta}^M\}$ should be estimated. The simultaneous estimation of these parameters is challenging because the changes in one parameter may affect the estimation of the other. This phenomenon has motivated us to consider the EM algorithm and the LS method simultaneously. In particular, we develop an iterative EM–LS algorithm that integrates an EM algorithm into the LS method and iteratively estimates the two types of parameters, i.e., $\{\Phi_t\}_{t=t_0}^{t_T}$ and $\{\boldsymbol{\beta}^1, \ldots, \boldsymbol{\beta}^M\}$.

The flowchart in Fig. 6 and the pseudocodes illustrate the iterative algorithm procedure for parameter estimation. The deviance of the sensor data from the mean profiles of thermal field $m\,\{\mathbf{g}_t^m\}_{t=t_0}^{t_T}$ is known. At iteration $k = 1$, where k denotes the index of iteration, we employ a Gaussian process model [26] to fill in the missing data in the grid-based sensor networks using the deviance data $\{\mathbf{g}_t^m\}_{t=t_0}^{t_T}$, given that there are missing data on the sensor locations. The processed data set is denoted as $\{b'^m_{t,k}(\mathbf{s})\}, \forall \mathbf{s} \in \mathbf{S}, t = t_0, \ldots, t_T$.

Step 1: Estimating parameters in the AR model

To estimate the parameters of the AR model, we first obtain the estimates of $\{\hat{w}^m_{t,k}(\mathbf{s})\}$ using the current estimated parameters of the MGP model for the thermal field m ($m = 1, \ldots, M$), where we initiate $\hat{w}^m_{t,0}(\mathbf{s}) = 0$. The calculation of $\hat{w}^m_{t,k}(\mathbf{s})$ at $k = 1, 2, \ldots, K$ is introduced in Step 2. Then, we obtain the data set $\{u^m_{t,k}(\mathbf{s})\}$ by calculating the difference between $\{b'^m_{t,k}(\mathbf{s})\}$ and $\{\hat{w}^m_{t,k-1}(\mathbf{s})\}$ using the following equation:

$$u^m_{t,k}(\mathbf{s}) = b'^m_{t,k}(\mathbf{s}) - \hat{w}^m_{t,k-1}(\mathbf{s}). \tag{5}$$

We establish an AR model $d^m_{t,k}(\mathbf{s}) = \sum_{l=1}^{L} \beta^m_{l,k} b^m_{t-l,k}(\mathbf{s})$ and its corresponding data set $\{u^m_{t,k}(\mathbf{s}), b'^m_{t-1,k}(\mathbf{s}), \ldots, b'^m_{t-L,k}(\mathbf{s})\}, \forall \mathbf{s} \in \mathbf{S}, t = t_L, \ldots, t_T$, and we use the LS method to estimate the parameters of the AR model and obtain the estimator $\hat{\boldsymbol{\beta}}^m_k$.

Step 2: Estimating parameters in the MGP model

To estimate the parameters of MGP, we first obtain the data set that corresponds to $w^m_t(\mathbf{s}), m = 1, \ldots, M$ by calculating the deviance between the bias at time t and those at previous L time points, that is,

$$v^m_{t,k}(\mathbf{s}) = b'^m_{t,k}(\mathbf{s}) - \sum_{l=1}^{L} \hat{\beta}^m_{l,k} b'^m_{t-l,k}(\mathbf{s}). \tag{6}$$

Then, we apply the EM algorithm to estimate $\Phi_{t,k} = \{\boldsymbol{\mu}_{t,k}, \mathbf{C}_{t,k}, \sigma^2_{t,k}\}$ and $\boldsymbol{\alpha}^m_{t,k}$ using the obtained data set of M thermal fields $\{v^1_{t,k}(\mathbf{s}), \ldots, v^M_{t,k}(\mathbf{s})\}, \forall \mathbf{s} \in \mathbf{S}^m_t, t = t_0, \ldots, t_T$, and to acquire the estimated parameters in MGP, i.e., $\hat{\Phi}_{t,k}$ and $\hat{\alpha}^m_{t,k}$.

E-step: The expectation of $\boldsymbol{\alpha}^m_{t,k}$ and the covariance matrix $\mathbf{C}^m_{t,k}$ are estimated using current $\Phi_{t,k}$ for $m = 1, \ldots, M$

$$\boldsymbol{\alpha}^m_{t,k} = \left(\frac{1}{\sigma^2_{t,k}} \boldsymbol{\kappa}^T_m \boldsymbol{\kappa}_m + \mathbf{C}^{-1}_{t,k}\right)^{-1} \left(\frac{1}{\sigma^2_{t,k}} \boldsymbol{\kappa}^T_m \mathbf{v}^m_{t,k} + \mathbf{C}^{-1}_{t,k} \boldsymbol{\mu}_{t,k}\right)$$

$$\mathbf{C}^m_{t,k} = \left(\frac{1}{\sigma^2_{t,k}} \boldsymbol{\kappa}^T_m \boldsymbol{\kappa}_m + \mathbf{C}^{-1}_{t,k}\right)^{-1}$$

where $\boldsymbol{\kappa}_m$ denotes the $n^m_t \times n$ kernel matrix between $\mathbf{S}$ and $\mathbf{S}^m_t$ obtained using the kernel function in (4) and $\mathbf{v}^m_{t,k}$ denotes the vector of $\{v^m_{t,k}(\mathbf{s})\}, \forall \mathbf{s} \in \mathbf{S}^m_t$.

M-step: $\boldsymbol{\mu}_{t,k}$, $\mathbf{C}_{t,k}$, and $\sigma^2_{t,k}$ are optimized based on the last E-step; thus, we obtain the updated $\boldsymbol{\mu}_{t,k}$, $\mathbf{C}_{t,k}$, and $\sigma^2_{t,k}$ as follows:

$$\boldsymbol{\mu}_{t,k} = \frac{1}{\pi + M} \sum_{m=1}^{M} \boldsymbol{\alpha}^m_{t,k}$$

$$\mathbf{C}_{t,k} = \frac{1}{\tau + M} \left(\pi \boldsymbol{\mu}_{t,k} \boldsymbol{\mu}^T_{t,k} + \tau \boldsymbol{\kappa}^{-1}_0 + \sum_{m=1}^{M} \mathbf{C}^m_{t,k} + \sum_{m=1}^{M} \left(\boldsymbol{\alpha}^m_{t,k} - \boldsymbol{\mu}_{t,k}\right)\left(\boldsymbol{\alpha}^m_{t,k} - \boldsymbol{\mu}_{t,k}\right)^T \right)$$

$$\sigma^2_{t,k} = \frac{1}{\sum_{m=1}^{M} n^m_t} \sum_{m=1}^{M} \left\| \mathbf{v}^m_{t,k} - \boldsymbol{\kappa}^T_m \boldsymbol{\alpha}^m_{t,k} \right\|^2 + \mathrm{tr}\left(\boldsymbol{\kappa}_m \mathbf{C}^m_{t,k} \boldsymbol{\kappa}^T_m\right)$$

where $\boldsymbol{\kappa}_0$ denotes the $n \times n$ kernel matrix among $\mathbf{S}$ and $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. After the EM algorithm is implemented, we obtain $\hat{\Phi}_{t,k} = \{\hat{\boldsymbol{\mu}}_{t,k}, \hat{\mathbf{C}}_{t,k}, \hat{\sigma}^2_{t,k}\}$ and $\hat{\alpha}^m_{t,k}$, $t = t_0, \ldots, t_T$. The estimated inductive functions in terms of $w^m_t(\mathbf{s})$ are as follows:

$$\hat{w}^m_{t,k}(\mathbf{s}) = \sum_{i=1}^{n} \hat{\alpha}^m_{t,i,k} \kappa(\mathbf{s}_i, \mathbf{s}) \tag{7}$$

where $\hat{\alpha}^m_{t,i,k}$ denotes the ith element of $\hat{\alpha}^m_{t,k}$.

We update the iteration sequence $k = k + 1$ after both types of parameters at iteration k have been estimated. Since missing data exist in the target field, we use the estimated parameters in the last iteration to fill in the missing values in the grid-based sensor networks using the proposed FML model as follows:

$$\hat{b}^m_{t,k}(\mathbf{s}) = \sum_{l=1}^{L} \hat{\beta}^m_{l,k-1} \hat{b}^m_{t-l,k}(\mathbf{s}) + \sum_{i=1}^{n} \hat{\alpha}^m_{t,i,k-1} \kappa(\mathbf{s}_i, \mathbf{s}) \tag{8}$$

and obtain the processed data set $\{b'^m_{t,k}(\mathbf{s})\}, \forall \mathbf{s} \in \mathbf{S}, t = t_0, \ldots, t_T$. Then, we repeat Steps 1 and 2 to estimate the two types of parameters until convergence is achieved. These parameters interweave with each other; therefore, selecting one type of parameter for the convergence test is reasonable. We adopt the convergence criterion to check whether the parameters of the AR model $\boldsymbol{\beta}^m$ converge at $m = 1, \ldots, M$. Convergence is considered achieved when the changes in $\boldsymbol{\beta}^m$ between two consecutive iterations are within a predetermined threshold ϵ

$$\Delta \boldsymbol{\beta}^m_k = \| \boldsymbol{\beta}^m_k - \boldsymbol{\beta}^m_{k-1} \| < \epsilon, \quad \forall m = 1, \ldots, M. \tag{9}$$

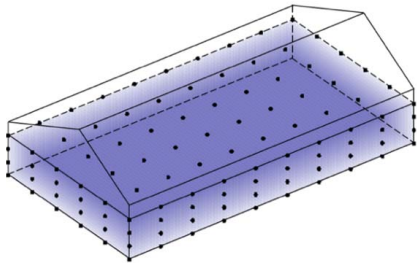


Fig. 7. Illustration of a traditional cuboid granary (the blue cloud represents grain, whereas the black points represent sensor locations).

TABLE I
PHYSICAL PROPERTIES OF GRAIN IN THE GRANARY

Material	Density ρ	Thermal conductivity in x, y, and z directions λ_x, λ_y, and λ_z	Specific heat capacity c
Mixed wheat	750 kg/m^2	0.0834, 0.0833, 0.2257 W/(m · K)	2000 J/(kg · K)

C. Spatiotemporal Field Estimation

The thermal field $m (m = 1, \ldots, M)$ can finally be obtained with the mean function and local variability using sensor observations from M homogeneous fields based on the estimated parameters in the FML model

$$\hat{y}^m(\mathbf{s}, t) = \mu^m(\mathbf{s}, t) + \sum_{l=1}^{L} \hat{\beta}^m_l b^m_{t-l}(\mathbf{s}) + \sum_{i=1}^{n} \hat{\alpha}^m_{t,i} \kappa(\mathbf{s}_i, \mathbf{s}) \tag{10}$$

TABLE II
RMSEs OF THE PROPOSED MODEL WITH DIFFERENT ORDERS FOR GROUP A

Granary	#1		#2		#3		Model performance
Order	$\boldsymbol{\beta}^m$	RMSE	$\boldsymbol{\beta}^m$	RMSE	$\boldsymbol{\beta}^m$	RMSE	Average RMSE
1	{0.9986}	0.0949	{0.9990}	0.1210	{0.9990}	0.0918	0.1026
2	{0.9841, 0.0162}	0.0926	{0.9846, 0.0158}	0.1196	{0.9874, 0.0128}	0.0842	0.0982
3	{0.9811, 0.0105, 0.0080}	0.0939	{0.9894, 0.0068, 0.0042}	0.1207	{0.9779, 0.0136, 0.0090}	0.0905	0.1017

Note: the number in parentheses is the standard deviation of RMSEs under 100 replications

where $\hat{y}^m(\mathbf{s}, t)$ denotes the estimate given a new location $\mathbf{s} \in \boldsymbol{\mathcal{S}} \subset \mathbb{R}^3$ and time t; $b^m_{t-l}(\mathbf{s}) = \hat{y}^m(\mathbf{s}, t-l) - \mu^m(\mathbf{s}, t-l)$; $\hat{\beta}^m_l$ and $\hat{\alpha}^m_{t,i}$ denote the estimated parameters of the AR model and the MGP model, respectively.

IV. CASE STUDY

Monitoring grain quality during storage is necessary to reduce losses. A grain thermal field should be estimated based on grain thermal sensor data for grain quality monitoring. Therefore, we conducted a real case study on grain thermal field estimation to test the performance of our proposed model. We selected three adjacent traditional granaries that store the same type of grains in a national grain depot located in central China. The thermal sensor data for these granaries were collected synchronously from thermal sensor networks from July 2, 2012 to March 4, 2013. Fig. 7 presents an illustration of a traditional cuboidal granary. As shown in Fig. 7, 240 evenly spaced temperature sensors are distributed in the granaries, which have a volume of 46 m (length), 26 m (width), and 6 m (height). The sensors are deployed at the initial positions of 0.5, 0.5, and 0.3 m in the x, y, and z directions, respectively, and located at 5-m intervals in the x and y directions and 1.8-m intervals in the z-direction. We obtained 11 280 samples of $10 \times 6 \times 4$ meshgrid-based sensor data at 47 time points from each selected granary.

Missing sensor data commonly exist in grain storage sensor networks [44]. Generally, once the ratio of missing sensor data is beyond 30%, grain quality monitoring performance will greatly decline, leading to a large number of grain losses. In a real grain storage scenario, the range of missing data ratio is approximately from 20% to 40%. Therefore, in order to be consistent with the real scenario, for each granary, we randomly selected a number of sensor data as training data and set the rest as test data, which are assumed missing. We set the ratios of the training data to the test data as 7/3 (Group A), 6/4 (Group B), and 5/5 (Group C) in the case study to evaluate the performance of our proposed method. Specifically, we set the ratio of training data to test data as 5/5 to test the effectiveness of our proposed model in some extreme cases with a large ratio of missing data. Furthermore, we repeated the procedure by randomly selecting the training and test data 100 times to evaluate the proposed method.

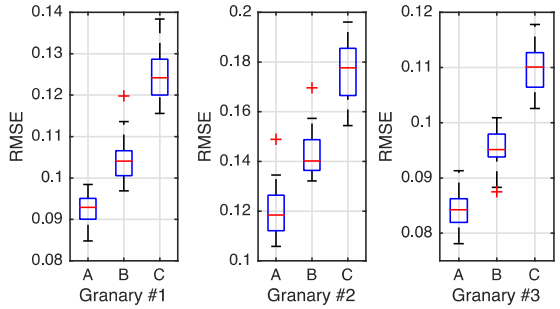


Fig. 8. Model performance of the proposed method under 100 replications with group A (30% testing), group B (40% testing), and group C (50% testing).

We considered the root-mean-square errors (RMSEs) between the real data values and the estimated values of granary $m (m = 1, \ldots, M)$, defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{I(T+1)} \sum_{t=t_0}^{t_T} \sum_{i=1}^{I} (y^m(\mathbf{s}_i, t) - \hat{y}^m(\mathbf{s}_i, t))^2} \quad (11)$$

where $y^m(\mathbf{s}_i, t)$ is the observation in the test data set of granary m and $\hat{y}^m(\mathbf{s}_i, t)$ is the estimated value of granary m at location $\mathbf{s}_i$ and time t.

For the mean function, the physical properties of stored grains with regard to the static parameters in the thermodynamic model are given in Table I. Environmental conditions considerably affect the temperature of granary walls (e.g., the seasonal variation range of temperature reaches 25 °C) and further affect the inner grain temperature. Accordingly, we set the temperature of the granary walls as the boundary condition to determine the environmental effects on the grain thermal field. We acquired the sensor data for the granary walls collected by the sensors on the walls and adopted a Gaussian process model [21] for interpolation using the sensors mounted on the granary walls to obtain the boundary condition of the thermodynamic model. Afterward, we adopted the thermodynamic model to obtain the global profile of each granary. We calculated the RMSEs using the thermodynamic model for three granaries which are 0.7468, 1.0521, and 0.9938, thereby indicating that the thermodynamic model can mainly capture global trends in grain thermal fields.

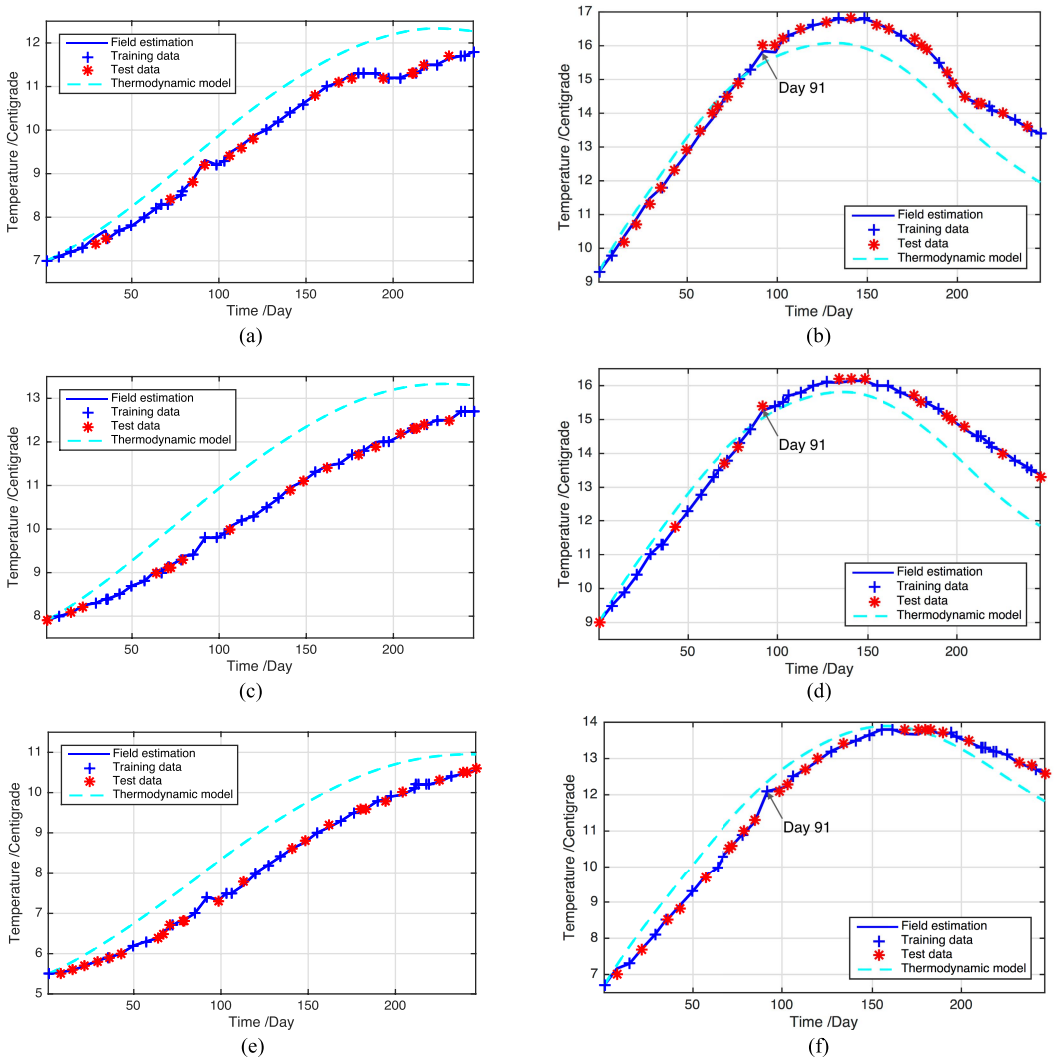


Fig. 9. Examples of grain thermal profiles in the three granaries estimated using the proposed method and the thermodynamic model. Coordinates (a) Granary #1 (10.5, 30.5, 2.1) m. (b) Granary #1 (15.5, 35.5, 3.9) m. (c) Granary #2 (10.5, 30.5, 2.1) m. (d) Granary #2 (15.5, 35.5, 3.9) m. (e) Granary #3 (10.5, 30.5, 2.1) m. (f) Granary #3 (15.5, 35.5, 3.9) m.

For local variability, the hyperprior parameter Ψ was determined via expert knowledge of thermal fields. Small π and τ values generally provide a good estimation performance [43]. Without loss of generality, we set $\pi = 1$ and $\tau = 0$. δ was determined on the basis of the maximum length of the grain thermal field. The parameters of the MGP model $\Phi_t = \{\boldsymbol{\mu}_t, \mathbf{C}_t, \sigma_t^2\}, t = t_0, \ldots, t_T$ and the parameters in the AR model $\boldsymbol{\beta}^m = \{\beta_1^m, \beta_2^m, \ldots, \beta_L^m\}, m = 1, \ldots, M$ were estimated using the proposed iterative EM–LS algorithm in (5)–(9). After parameter estimation, we estimated the grain thermal fields using (10) and calculated the RMSEs to validate the model performance.

Then, we introduced the determination of the order L in the AR model. In grain thermal field estimation, on one hand, we tended to set the order as a medium value because the grain temperature during storage changes slowly. On the other hand, we determined the order L using a best-fit strategy, which compares the model performances (RMSEs) with different values of the order and selected the proper order L with a minimum average RMSE value. We used the data in group A as an example. The estimated $\boldsymbol{\beta}^m$ and RMSEs for the three granaries are listed in Table II if the order of the proposed model is set to $L = 1, 2,$ and 3. We compared the average RMSEs of the three granaries to select a proper value of order L. Table II shows that the proposed model with order $L = 2$ has the best model performance, i.e., the minimum average RMSE value.

Fig. 8 illustrates the performance of our proposed field estimation method under 100 replications. The RMSEs of the discrepancy between the test data and the estimated values by our proposed model for the three granaries are as follows: 0.0926, 0.1196, and 0.0842 for Group A; 0.1048, 0.1431, and 0.0951 for Group B; and 0.1253, 0.1759, and 0.1096 for Group C. Compared with the performance when only the thermodynamic model is used, our proposed method not only characterizes the global profile of thermal fields but also captures their local variability using sensor observations, and ultimately achieves a higher field estimation than the thermodynamic

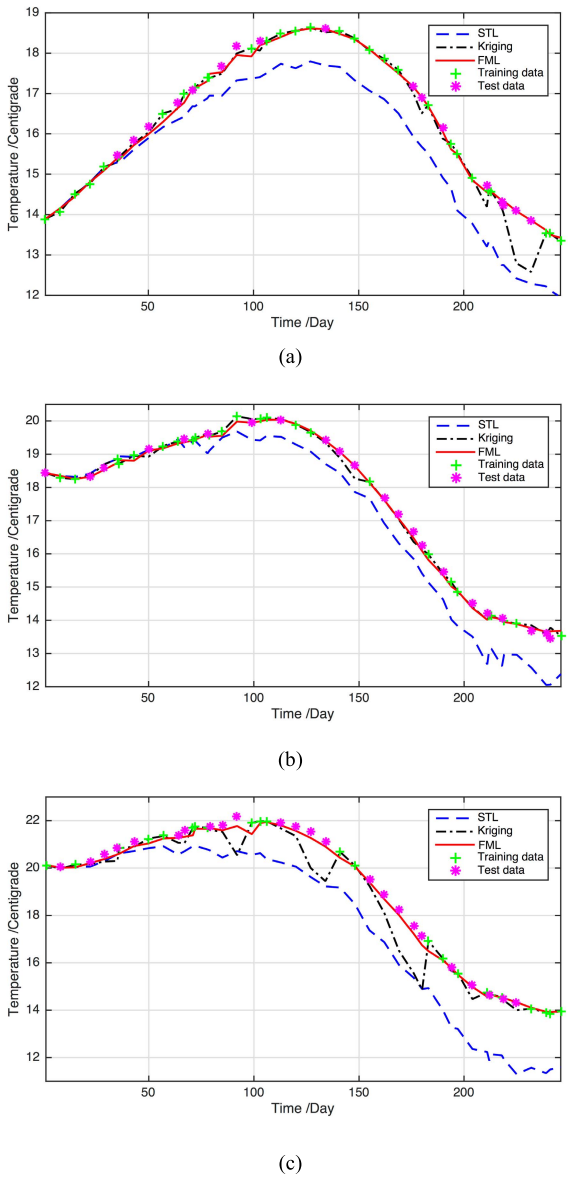


Fig. 10. Estimated grain temperature profiles in the three granaries using comparative methods. [Coordinate (10.5, 10.5, 3.9)] (a) Granary #1. (b) Granary #2. (c) Granary #3.

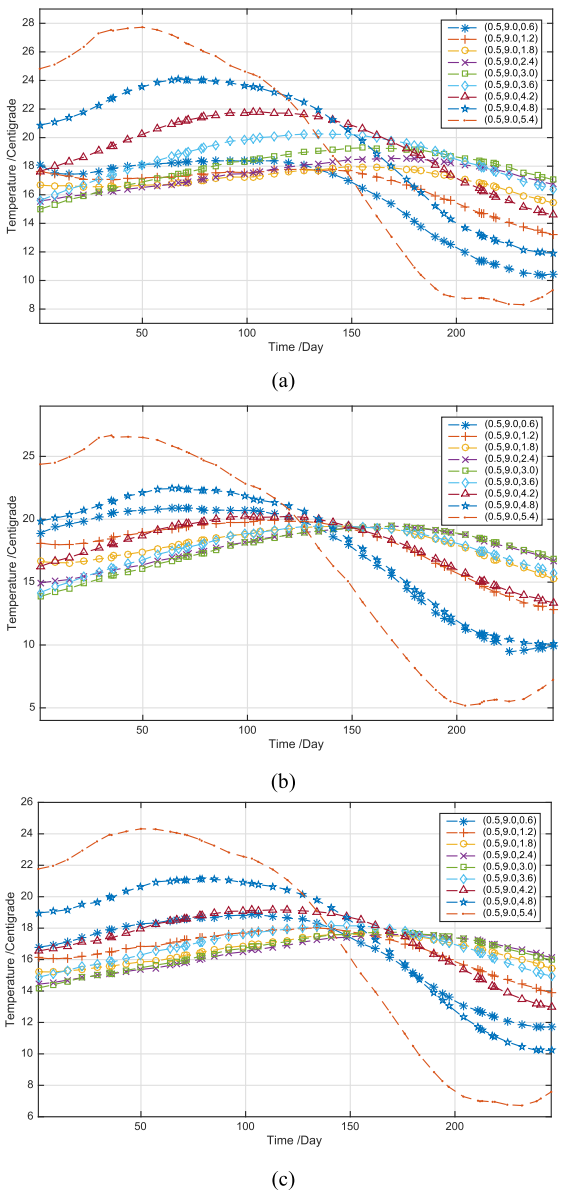


Fig. 11. Estimated grain temperature profiles at different locations in the three granaries. (a) Granary #1. (b) Granary #2. (c) Granary #3.

model. Fig. 9 presents examples of the estimated thermal fields using our proposed model v.s. using only the thermodynamic model based on the three granaries. The thermodynamic model can capture the global profile of the grain thermal field in each granary. Although the sensor observations of each granary are limited, our proposed model can accurately characterize the spatiotemporal correlations of local variability by sharing sensor observations from homogenous granaries. For example, we accurately estimate the local variability of Granaries #1 and #2 at day 91 by sharing the corresponding sensor observations from Granary #3, as shown in Fig. 9(b), (d), and (f).

To further validate the performance of our proposed method, we compared the FML model with two alternative models: an STL model and a kriging model. Compared with our proposed method, these two models also employed the mixed-effect model framework and used the thermodynamic model as the mean function, but they characterized the local variability only using sensor data collected from the corresponding granary. The STL model assumed that grain temperature data at different time points were independent of one another. It characterized the spatial correlation of the local variability at each time point. Meanwhile, the kriging model was used to characterize the spatiotemporal correlation of local variability. We considered a covariance function in terms of space and time and applied the kriging model to describe the local variability of the grain thermal field. We implemented these models 100 times and calculated their average RMSEs (Table III). The FML model outperforms the STL and kriging models. The average computation time of field estimation by STL, Kriging, and FML is 156.67, 387.69, and 769.26 s, respectively. Since the grain temperature changes slowly during storage, to reduce the data storage and processing costs, the sampling frequency

TABLE III
COMPARISON OF MODEL PERFORMANCE USING THREE MODELS

Model	Group of test data	RMSEs (Granary #1)	RMSEs (Granary #2)	RMSEs (Granary #3)
STL	A	0.6911 (0.0135)	0.9995 (0.0248)	0.7803 (0.0181)
	B	0.6945 (0.0128)	1.0114 (0.0156)	0.7845 (0.0162)
	C	0.7013 (0.0063)	1.0190 (0.0118)	0.7936 (0.0115)
Kriging	A	0.2040 (0.0175)	0.3619 (0.0333)	0.2237 (0.0205)
	B	0.2433 (0.0145)	0.4514 (0.0326)	0.2814 (0.0203)
	C	0.2855 (0.0143)	0.5299 (0.0270)	0.3437 (0.0204)
FML	A	0.0926 (0.0034)	0.1196 (0.0092)	0.0842 (0.0036)
	B	0.1048 (0.0054)	0.1431 (0.0098)	0.0951 (0.0040)
	C	0.1253 (0.0064)	0.1759 (0.0134)	0.1096 (0.0042)

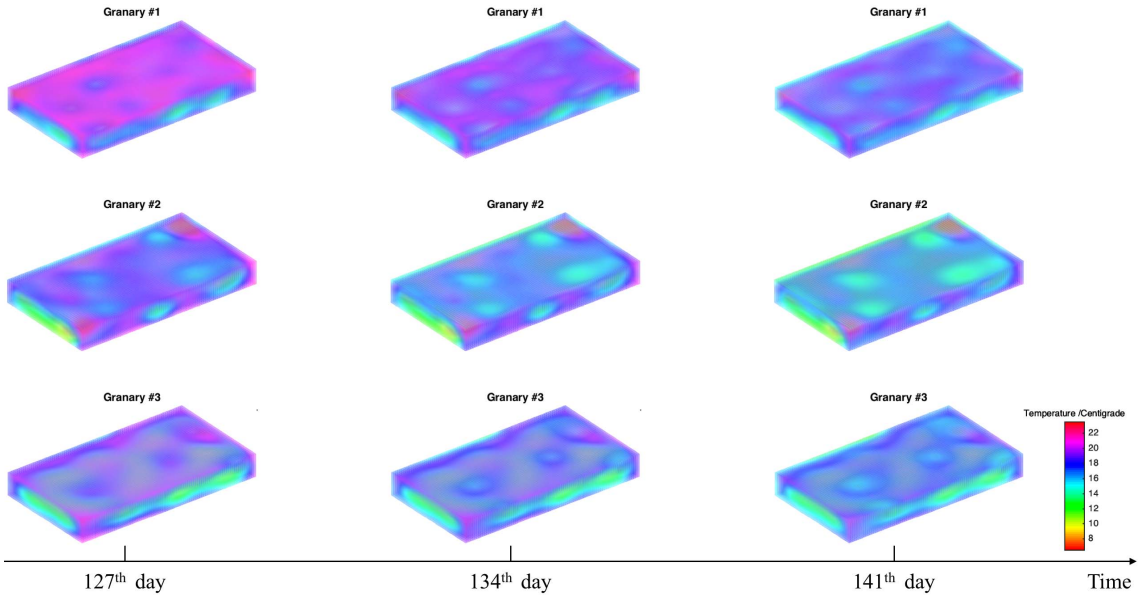


Fig. 12. Estimated grain thermal fields at consecutive time points in the three granaries.

of the grain temperature sensor data is comparatively low and the computational time of these three models is acceptable under this scenario. Fig. 10 presents examples of thermal fields estimated by the three models using the data setup of Group B. The STL model cannot accurately characterize the profiles of the thermal field since it only considers the spatial correlation of the grain thermal field at each time point with only limited sensor observations. The kriging model efficiently captures local variability when sensor observations in the target thermal field are sufficient. However, when sensor observations are sparse, the kriging model fails to achieve an accurate result. The FML model captures trend and local variability more accurately than the two alternative models.

We estimated the thermal fields of the three granaries using our proposed method. Fig. 11 shows the estimated grain temperature profiles at different locations in the three granaries. The grain temperature profiles at the same locations in the three granaries exhibit similar variation patterns over time. The grain temperature at the center of the granaries varies more slowly and has a smaller fluctuation range than the grain temperature near the boundary of the granaries. Fig. 12 shows the estimated thermal field at three consecutive time points in the three granaries. The thermal fields exhibit similar spatial correlation patterns and evolve over time with similar trends because the granaries share the same environmental conditions. These findings provide a new quantitative understanding of temperature field evaluation in granaries, which is of practical significance for monitoring the quality of grains and reducing grain loss during storage.

V. CONCLUSION

Thermal field estimation plays an essential role in a number of industrial domains by providing useful information for system improvement. One of the major challenges in thermal field

estimation is the limited availability of sensor observations, because only sparse sensors are engaged in sensor networks to collect data. Moreover, missing data arise due to sensor malfunction and data loss during transmissions. An efficient spatiotemporal approach for estimating the thermal field under such a scenario should be developed. In this article, we propose a dynamic thermal field estimation method using limited sensor observations from neighboring homogeneous sources. By fully considering limited sensor observations, we adopt a thermodynamic model to capture the trend of a thermal field and propose an FML method that integrates an MGP model into an AR model to characterize the spatiotemporal dynamics of local variability.

In our future work, given an accurate thermal field, we will focus on establishing an effective online strategy for simultaneously monitoring multiple thermal fields using limited sensor observations by fully considering their common features. We will also focus on developing a field prediction approach and modeling the spatiotemporal dynamics of a thermal field in the upcoming time points using limited sensor observations from several homogeneous data sources.

REFERENCES

[1] P. F. J. Lermusiaux, "Uncertainty estimation and prediction for interdisciplinary ocean dynamics," *J. Comput. Phys.*, vol. 217, no. 1, pp. 176–199, Sep. 2006.

[2] G. Moser and S. B. Serpico, "Automatic parameter optimization for support vector regression for land and sea surface temperature estimation from remote sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 909–921, Mar. 2009.

[3] Y. Chen and H. Yang, "Sparse modeling and recursive prediction of space–time dynamics in stochastic sensor networks," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 1, pp. 215–226, Jan. 2016.

[4] Y. Zheng, F. Liu, and H.-P. Hsieh, "U-Air: When urban air quality inference meets big data," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Min.*, Aug. 2013, pp. 1436–1444.

[5] J. Y. Kim and D. H. Park, "Thermal analysis and statistical evaluation of EPR used in nuclear power plants," in *Proc. IEEE Electr. Insul. Conf.*, Jun. 2015, pp. 5–8.

[6] C. Conficoni, A. Bartolini, A. Tilli, C. Cavazzoni, and L. Benini, "Integrated energy-aware management of supercomputer hybrid cooling systems," *IEEE Trans. Ind. Informat.*, vol. 12, no. 4, pp. 1299–1311, Aug. 2016.

[7] C. Jia, D.-W. Sun, and C. Cao, "Finite element prediction of transient temperature distribution in a grain storage bin," *J. Agricult. Eng. Res.*, vol. 76, no. 4, pp. 323–330, Aug. 2000.

[8] D. Wang and X. Zhang, "A prediction method for interior temperature of grain storage via dynamics models: A simulation study," in *Proc. IEEE Int. Conf. Autom. Sci. Eng.*, Aug. 2015, pp. 1477–1483.

[9] O. A. Khatchatourian and F. A. De Oliveira, "Mathematical modelling of airflow and thermal state in large aerated grain storage," *Biosyst. Eng.*, vol. 95, no. 2, pp. 159–169, Oct. 2006.

[10] R. Rusinek and R. Kobyłka, "Experimental study and discrete element method modeling of temperature distributions in rapeseed stored in a model bin," *J. Stored Prod. Res.*, vol. 59, pp. 254–259, Oct. 2014.

[11] J. Li and S. Huang, "Regression-based process monitoring with consideration of measurement errors," *IIE Trans.*, vol. 42, no. 2, pp. 146–160, Nov. 2009.

[12] D. Wang, K. Liu, and X. Zhang, "Spatio temporal thermal field modeling using partial differential equations with time-varying parameters," *IEEE Trans. Autom. Sci. Eng.*, to be published.

[13] A. Wang, X. Xian, F. Tsung, and K. Liu, "A spatial-adaptive sampling procedure for online monitoring of big data streams," *J. Qual. Technol.*, vol. 50, no. 4, pp. 329–343, Oct. 2018.

[14] T. Feng, X. Qian, K. Liu, and S. Huang, "Dynamic inspection of latent variables in state-space systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 3, pp. 1232–1243, Jul. 2019.

[15] T. Liu, "Parallel reinforcement learning: A framework and case study," *IEEE/CAA J. Automatica Sinica*, vol. 5, no. 4, pp. 827–835, Jul. 2018.

[16] Q. Huang, "Physics-driven Bayesian hierarchical modeling of the nanowire growth process at each scale," *IIE Trans.*, vol. 43, no. 1, pp. 1–11, Dec. 2010.

[17] R. Soundararajan and A. C. Bovik, "Video quality assessment by reduced reference spatio-temporal entropic differencing," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 4, pp. 684–694, Apr. 2013.

[18] M. Katzfuss and N. Cressie, "Spatio-temporal smoothing and EM estimation for massive remote-sensing data sets," *J. Time Ser. Anal.*, vol. 32, no. 4, pp. 430–446, 2011.

[19] X. Liu, K. Yeo, Y. Hwang, J. Snigh, and J. Kalagnanam, "A statistical modeling approach for air quality data based on physical dispersion processes and its application to ozone modeling," *Ann. Appl. Stat.*, vol. 10, no. 2, pp. 756–785, Jan. 2016.

[20] M. L. Stein, *Interpolation of Spatial Data: Some Theory for Kriging*. New York, NY, USA: Springer, 1999.

[21] C. E. Rasmussen, "Gaussian processes in machine learning," in *Advanced Lectures on Machine Learning*. Berlin, Germany: Springer, 2004, pp. 63–71.

[22] H. Jiang, X. Deng, V. López, and H. F. Hamann, "Online updating of computer model output using real-time sensor data," *Technometrics*, vol. 58, no. 4, pp. 472–482, Oct. 2016.

[23] N. Cressie and G. Johannesson, "Fixed rank Kriging for very large spatial data sets," *J. Roy. Statist. Soc. B (Statist. Methodol.)*, vol. 70, no. 1, pp. 209–226, Feb. 2008.

[24] B. Zhang, H. Sang, and J. Z. Huang, "Full-scale approximations of spatio-temporal covariance models for large datasets," *Stat. Sin.*, vol. 25, no. 1, pp. 99–114, Jan. 2015.

[25] L. Xu and Q. Huang, "Modeling the interactions among neighboring nanostructures for local feature characterization and defect detection," *IEEE Trans. Autom. Sci. Eng.*, vol. 9, no. 4, pp. 745–754, Oct. 2012.

[26] D. Wang, K. Liu, and X. Zhang, "Modeling of a three-dimensional dynamic thermal field under grid-based sensor networks in grain storage," *IISE Trans.*, vol. 51, no. 5, pp. 531–546, May 2019.

[27] R. J. Aliaga, "Real-time estimation of zero crossings of sampled signals for timing using cubic spline interpolation," *IEEE Trans. Nucl. Sci.*, vol. 64, no. 8, pp. 2414–2422, Aug. 2017.

[28] Z. Cheng and J. W. Mark, "Channel estimation by modulated Lagrange interpolation," *Signal Process.*, vol. 90, no. 9, pp. 2749–2759, Sep. 2010.

[29] P. Zhao, S. C. H. Hoi, J. Wang, and B. Li, "Online transfer learning," *Artif. Intell.*, vol. 216, pp. 76–102, Nov. 2014.

[30] Y. Lin, S. Huang, G. E. Simon and, and S. Liu, "Analysis of depression trajectory patterns using collaborative learning," *Math. Biosci.*, vol. 282, pp. 191–203, Dec. 2016.

[31] X. Liu, G. Wang, Z. Cai, and H. Zhang, "Bagging based ensemble transfer learning," *J. Ambient Intell. Human. Comput.*, vol. 7, no. 1, pp. 29–36, Feb. 2016.

[32] M. Bussas, C. Sawade, N. Kühn, T. Scheffer, and N. Landwehr, "Varying-coefficient models for geospatial transfer learning," *Mach. Learn.*, vol. 106, nos. 9–10, pp. 1419–1440, Oct. 2017.

[33] S. J. Pan, V. W. Zheng, Q. Yang, and D. H. Hu, "Transfer learning for WiFi-based indoor localization," in *Proc. AAAI Workshop Tech. R.*, Jan. 2008, pp. 43–48.

[34] P. Song, Y. Jin, M. Xin, and L. Zhao, "Speech emotion recognition using transfer learning," *IEICE Trans. Inf. Syst.*, vol. 97, no. 9, pp. 2530–2532, Sep. 2014.

[35] L. Cheng, F. Tsung, and A. Wang, "A statistical transfer learning perspective for modeling shape deviations in additive manufacturing," *IEEE Robot. Autom. Lett.*, vol. 2, no. 4, pp. 1988–1993, Oct. 2017.

[36] C. E. Magnuson, J. B. Shea, and J. T. Fairbrother, "Effects of repeated retention tests on learning a single timing task," *Res. Quart. Exerc. Sport*, vol. 75, no. 1, pp. 39–46, Mar. 2004.

[37] D. Maslovat, R. Chua, T. D. Lee, and L. M. Franks, "Contextual interference: Single task versus multi-task learning," *Motor Control*, vol. 8, no. 2, pp. 213–233, Apr. 2004.

[38] F. Jin and S. Sun, "Neural network multitask learning for traffic flow forecasting," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jun. 2008, pp. 1897–1901.

[39] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: deep belief networks with multitask learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2191–2201, Oct. 2014.

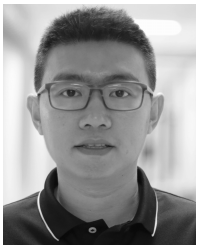
[40] Y.-L. Xu, X.-X. Li, D.-R. Chen, and H.-X. Li, "Learning rates of regularized regression with multiple Gaussian kernels for multi-task learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 11, pp. 5408–5418, Nov. 2018.

[41] W. Guo and G. Chen, "Human action recognition via multi-task learning base on spatial-temporal feature," *Inf. Sci.*, vol. 320, pp. 418–428, Nov. 2015.

[42] A. R. Goncalves, A. Banerjee, and F. J. Von Zuben, "Spatial projection of multiple climate variables using hierarchical multitask learning," in *Proc. 31st AAAI Conf. Artif. Intell.*, Jan. 2017, pp. 4509–4515.
[43] C. Shao, J. Ren, H. Wang, J. Jin, and S. Hu, "Improving machined surface shape prediction by integrating multi-task learning with cutting force variation modeling," *J. Manuf. Sci. Eng.*, vol. 139, no. 1, Jan. 2017, Art. no. 011014.
[44] Sina Finance, *35 Million Tons of Grain Losses Per Year in China*. Accessed: Jun. 2015. [Online]. Available: http://finance.sina.com.cn/money/future/20150616/083422442730.shtml

Di Wang received the B.S. degree in industrial engineering from Nankai University, Tianjin, China, in 2015. She is currently pursuing the Ph.D. degree with the Department of Industrial Engineering and Management, Peking University, Beijing, China.

Her research interests include statistically modeling of spatiotemporal data and complex systems.



Kaibo Liu (M'14) received the B.S. degree in industrial engineering and engineering management from Hong Kong University of Science and Technology, Hong Kong, in 2009, and the M.S. degree in statistics and the Ph.D. degree in industrial engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2011 and 2013, respectively.

He is currently an Associate Professor with the Department of Industrial and Systems Engineering, University of Wisconsin–Madison, Madison, WI, USA, where he is also the Associate Director of the UW–Madison IoT Systems Research Center. His research interests include system informatics, big data analytics, and data fusion for process modeling, monitoring, diagnosis, prognostics, and decision making.

Dr. Liu is a member of ASQ, INFORMS, SME, and IISE.

Xi Zhang (M'13) received the B.S. degrees in mechanical engineering and business administration from Shanghai Jiaotong University, Shanghai, China, in 2006, and the Ph.D. degree in industrial and management systems engineering from the University of South Florida, Tampa, FL, USA, in 2010.

He is currently an Associate Professor with the Department of Industrial Engineering and Management, Peking University, Beijing, China. His research interests include physics-based engineering data integration and analytics for process monitoring, diagnosis, control and optimization in complex engineering, and service systems.

Dr. Zhang is a member of ASQ, INFORMS, and IISE.

Hui Wang received the B.S. degree in mechanical engineering from Shanghai Jiaotong University, Shanghai, China, in 2001, the M.S. degree in mechanical engineering from the University of Michigan, Ann Arbor, MI, USA, in 2003, and the Ph.D. degree in industrial engineering from the University of South Florida, Tampa, FL, USA, in 2007.

He is currently an Assistant Professor with the Department of Industrial and Manufacturing Engineering, Florida State University, Tallahassee, FL, USA. His research interests include manufacturing system design, automation, and process control by integrating applied statistics, image processing, optimization, and control theory with engineering knowledge with broad applications including automotive, energy system, semiconductor, and nanomanufacturing.

Dr. Wang is a member of ASME, IISE, and INFORMS.