Spatio-temporal field estimation using kriged kalman filter (KKF) with sparsity-enforcing sensor placement

Roy, Venkat; Simonetto, Andrea; Leus, Geert

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Spatio-Temporal Field Estimation Using Kriged Kalman Filter (KKF) with Sparsity-Enforcing Sensor Placement

Venkat Roy [1,*], Andrea Simonetto [2] and Geert Leus [3]

[1] NXP Semiconductors, High Tech Campus 46, 5656 AE Eindhoven, The Netherlands
[2] Optimisation and Control group, IBM Research Ireland, Dublin 15, Ireland; andrea.simonetto@ibm.com
[3] Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands; g.j.t.leus@tudelft.nl
* Correspondence: venkat.roy@nxp.com

Abstract: We propose a sensor placement method for spatio-temporal field estimation based on a kriged Kalman filter (KKF) using a network of static or mobile sensors. The developed framework dynamically designs the optimal constellation to place the sensors. We combine the estimation error (for the stationary as well as non-stationary component of the field) minimization problem with a sparsity-enforcing penalty to design the optimal sensor constellation in an economic manner. The developed sensor placement method can be directly used for a general class of covariance matrices (ill-conditioned or well-conditioned) modelling the spatial variability of the stationary component of the field, which acts as a correlated observation noise, while estimating the non-stationary component of the field. Finally, a KKF estimator is used to estimate the field using the measurements from the selected sensing locations. Numerical results are provided to exhibit the feasibility of the proposed dynamic sensor placement followed by the KKF estimation method.

1. Introduction

Tracking the spatio-temporal evolution of any field using a limited number of homogeneous/heterogeneous sensors with a desired accuracy is one of the most common applications of wireless sensor networks (WSNs) [1,2]. Different types of environmental, geophysical and biological processes exhibit complicated spatial as well as temporal variability. Spatial and temporal variability of a spatio-temporally stationary physical field can be modelled by its correlation over space and time [3]. If the field is non-stationary then a suitable dynamic model can be used to model the spatio-temporal evolution of the field [3]. If the field exhibits both a stationary and non-stationary behavior over space and time then the field can be dynamically monitored by the combination of kriging [3] and Kalman filtering, i.e., a kriged Kalman filter (KKF) [4] or space-time Kalman filter [5]. The key idea behind the KKF is the liaison of kriging [3] and Kalman filtering. The unknown physical field is modelled as a combination of a non-stationary (capturing the dynamics) and a stationary (capturing the low magnitude spatial effects) stochastic component. Assuming that the dynamics of the non-stationary component and the second-order statistics of the stationary component (e.g., covariance structure) are perfectly known, KKF jointly estimates both of these field components using the spatial observations at every time instant. The KKF paradigm has been used for different applications like wireless communications (e.g., spectrum sensing [6] and path delay estimation [7]) and field estimation [5]. From a practical perspective, KKF can also be used for air

pollution level forecasting applications [8]. Also, a distributed implementation of KKF can be used for environment monitoring using a robotic sensor network [9].

One of the important overheads of dynamic field estimation using a WSN is the lack of sufficient measurements at every time instant. This is related to the shortage of sensor life time, availability of bandwidth, and other resource-related economical constraints. In such scenarios, we need to efficiently place/move the available sensors into the most informative locations over space and time. One notable approach for sensor placement in Gaussian processes is based on exploiting the submodularity of the mutual information between the sensing locations [10]. Submodularity of the frame potential and related greedy algorithms for sensor placement are proposed in [11]. Dynamic sensor scheduling is a well-cultivated topic in the fields of signal processing as well as control theory [12,13]. Some recent notable contributions are [14–16]. Prior knowledge regarding the correlation of the field over space and time can be exploited in a multi-layer design of sensor networks [17]. Selecting the most informative sensing locations can be treated as a sensor selection problem, which can be formulated as a convex optimization problem [18]. This can be solved for linear [19] as well as non-linear measurement models [20]. Sparsity-promoting approaches for sensor placement are also exhibited in [15,21], where the placement algorithm is formulated using the alternating direction method of multipliers (ADMM). In [22], a generalized sparsity-enforcing and performance-constrained sensor placement method is proposed, where the field can be either stationary or non-stationary. The aforementioned method can be implemented for a single snapshot or multiple snapshot ahead sensor placement and for a general class of spatio-temporal covariance matrices, which can either be ill-conditioned or well-conditioned. Seminal contributions on the convex formalism of sensor selection (like [18]) assume that the measurement noise components are spatio-temporally uncorrelated. However, this can be an unrealistic assumption in some practical scenarios [23]. However, even in those scenarios, it has been shown that the sensor selection problem can be formulated as a convex optimization problem [16,24].

In this work, we develop a unified framework of sensor placement followed by a KKF estimator to dynamically monitor a physical field that exhibits both stationarity and non-stationarity over space and time. In the first step, we select the most informative locations to deploy/move the sensors and in the second step we estimate the field by using the measurements from those selected locations. The key contributions can be summarized as follows,

- The performance metrics to estimate the stationary as well as the non-stationary components of the field are represented in closed form as an explicit function of the sensor location selection vector.
- The aforementioned analytical formalism tackles two important issues in the sensor placement step. First, the developed method takes care of the fact that the estimation of the non-stationary component of the field involves the stationary component of the field as a spatially correlated observation noise. Second, the proposed method is applicable for a general class of spatial covariance matrices of the stationary component of the field, even when they are ill-conditioned or close to singular [25].
- The proposed sensor placement problem is formulated in a way that minimizes a cost function that involves the sum of the mean square error (MSE) of the stationary and the non-stationary component of the field as well as a spatial sparsity enforcing penalty. The overall optimization problem also satisfies a flexible resource constraint at every time instant.

One of the aspects that distinguishes the proposed sensor placement method from the prior works in sensor placement for environmental field estimation is the specific statistical nature of the unknown physical field, which yields an additive coupling of stationary and the non-stationary components. Secondly, we develop a unified framework for the efficient utilization of the spatio-temporal variability of the field in order to design an opportunistic sensor placement method using a convex approach. We develop a parsimonious sensor placement algorithm followed by a KKF estimator, which can be used to dynamically monitor a general class of environmental fields (based on the assumed process model and spatial statistics of the field components). However, the developed approach is

similar to [22] in terms of the primary measurement model, which is considered to be linear and underdetermined. We emphasize that the proposed technique is a model-based centralized sensor placement method, where we resort to the Bayesian estimation philosophy. We assume that the available prior statistical knowledge regarding the unknown physical field like spatial correlation information and dynamics are perfectly known a priori. It is also assumed that for the current centralized setup the communication range of the sensors are sufficient to communicate with the fusion center, which can be located inside/outside the given service area.

Notations: Matrices are in upper case bold while column vectors are in lower case bold. The notation $[\mathbf{X}]_{ij}$ is the (i, j)-th entry of the matrix $\mathbf{X}$, $[\mathbf{x}]_i$ is the i-th entry of the vector $\mathbf{x}$, and $\mathrm{tr}[\mathbf{X}]$ denotes the trace of $\mathbf{X}$, i.e., the sum of the diagonal elements of $\mathbf{X}$. The notation $\mathrm{supp}(\mathbf{x})$ is defined as the set of indices of the non-zero entries of $\mathbf{x}$, while $\mathrm{diag}(\mathbf{x})$ and $\mathrm{diag}(\mathbf{X})$ are the diagonal matrix with diagonal $\mathbf{x}$ and the main diagonal of the matrix $\mathbf{X}$, respectively. An identity matrix of size $N \times N$ is denoted by $\mathbf{I}_N$. The notation $(\cdot)^T$ is the transpose operator, $\hat{\mathbf{x}}$ is the estimate of $\mathbf{x}$, and $\|\mathbf{x}\|_p = (\sum_{i=0}^{N-1} |[\mathbf{x}]_i|^p)^{1/p}$ is the ℓ_p norm of $\mathbf{x}$. Vectors of all zeros and ones of length N are denoted by $\mathbf{0}_N$ and $\mathbf{1}_N$, respectively. An all zero matrix of size $N \times N$ is given by $\mathbf{0}_{N \times N}$. The set of symmetric matrices of size $N \times N$ and the set of symmetric positive-definite matrices of size $N \times N$ are denoted by $\mathbb{S}^N$ and $\mathbb{S}^N_{++}$, respectively.

2. Signal Modelling and Problem Formulation

2.1. Measurement Model

Let us denote the spatially continuous field by $u_t(\mathbf{x})$, at any discrete time index t and location $\mathbf{x} \in \mathbb{R}^2$. We assume that the entire service area of interest is *uniformly discretized* over N pixels, where we would like to estimate the field intensities. The field intensities of the N pixels at time t are represented by $\mathbf{u}_t \in \mathbb{R}^N$. It is assumed that the field intensity is the same everywhere within a pixel, and it can be represented by $[\mathbf{u}_t]_j = u_t(\mathbf{x}_j)$, where $\mathbf{x}_j \in \mathbb{R}^2$ is the centroid of the j-th pixel, with $j = 1, \ldots, N$. We consider a linear underdetermined measurement model

$$\mathbf{y}_t = \mathbf{C}_t \mathbf{u}_t + \mathbf{e}_t \tag{1}$$

$$= \mathbf{C}_t(\mathbf{v}_t + \mathbf{s}_t) + \mathbf{e}_t, \tag{2}$$

where $\mathbf{v}_t \in \mathbb{R}^N$ is the non-stationary component of the field and $\mathbf{s}_t \in \mathbb{R}^N$ is a stationary component of the field capturing the non-dynamic spatial effects. It is assumed that $\mathbf{v}_t$ and $\mathbf{s}_t$ are mutually uncorrelated.

At any time t, the measurements are given by $\mathbf{y}_t \in \mathbb{R}^{M_t}$ collected from M_t ($M_t < N$) sensing locations (pixels) of the entire service area. The time-varying sensing or measurement matrix $\mathbf{C}_t \in \{0, 1\}^{M_t \times N}$ selects M_t measurements from N field locations. The measurement matrix $\mathbf{C}_t$ is constructed by removing the zero rows of $\mathrm{diag}(\mathbf{w}_t)$, where $\mathbf{w}_t \in \{0, 1\}^N$ is a *sensor location selection* vector. If $[\mathbf{w}_t]_j = 1(0)$, then the j-th pixel is selected (not selected) for sensor placement. Based on this, at any time t, the number and the constellation of the selected sensing locations are given by $\mathbf{1}^T \mathbf{w}_t = M_t$ and $\mathrm{supp}(\mathbf{w}_t)$, respectively. Using the considered construction of $\mathbf{C}_t$, we have the relations

$$\mathbf{C}_t \mathbf{C}_t^T = \mathbf{I}_{M_t}; \qquad \mathbf{C}_t^T \mathbf{C}_t = \mathrm{diag}(\mathbf{w}_t). \tag{3}$$

Note that the type of measurement matrix used in this work is similar to an incidence matrix, which can be viewed as a flexible data collection method using heterogeneous sensing equipments. In practice, when different types of sensing modalities are used, we may not know the process by which any of the sensors gathers its measurement but only its recorded value is important. Also, we rigorously exploit the property of the structure of $\mathbf{C}_t$ mentioned in (3), later in this paper.

The error incurred by the measurement process is modelled through $\mathbf{e}_t$, which is uncorrelated with both $\mathbf{v}_t$ and $\mathbf{s}_t$, respectively. The spatio-temporally white measurement noise $\mathbf{e}_t$ is characterized by $\mathbf{e}_t \sim \mathcal{N}(\mathbf{0}_{M_t}, \sigma_e^2 \mathbf{I}_{M_t})$.

2.2. Modelling of the Spatial Variability

The spatial effects of the field are modelled through a spatially colored yet temporally white discrete random process $\mathbf{s}_t \sim \mathcal{N}(\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s)$, where $\boldsymbol{\mu}_s \in \mathbb{R}^N$ is the mean and $\boldsymbol{\Sigma}_s \in \mathbb{S}_{++}^N$ is the spatial covariance matrix of $\mathbf{s}_t$. We assume that the process $\mathbf{s}_t$ is spatially second-order stationary as well as isotropic, which means that

$$\boldsymbol{\mu}_s = \mathbb{E}[\mathbf{s}_t] = \mu_s \mathbf{1}_N, \tag{4}$$

$$[\boldsymbol{\Sigma}_s]_{ij} = \mathbb{E}[(s_t(\mathbf{x}_i) - \mu_s)(s_t(\mathbf{x}_j) - \mu_s)] = f(\|\mathbf{x}_i - \mathbf{x}_j\|_2), \tag{5}$$

where $i, j = 1, \ldots, N$ [3]. Note that here we follow the same spatial discretization as mentioned in Section 2.1. There are several empirical as well as parametric model-based approaches to model the spatial covariance. In this work, we assume that the spatial covariance function is given by a simple squared exponential function:

$$f(\|\mathbf{x}_i - \mathbf{x}_j\|_2) = \sigma_s^2 \exp(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2^2}{\theta^2}), \tag{6}$$

where $\theta > 0$ is the parameter controlling the strength of the spatial correlation. The covariance function mentioned in (6) is plotted in Figure 1a for increasing values of the pairwise distance between the centroids of the pixels, i.e., $d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|_2$ and the parameter θ. Note that the aforementioned covariance function belongs to the family of Matérn covariance functions [3], which are widely used to model the spatial variability of a field in geostatistics and environmental sciences.

Using the squared exponential covariance function, the elements of the $N \times N$ spatial covariance matrix ($\boldsymbol{\Sigma}_s$) can be constructed by the Relation (5). Let us consider a service area with $N = 36$ pixels. The centroids of these 36 pixels, which are also the candidate locations for sensor deployment are shown in Figure 1b. These centroids are indexed from the top left to the bottom right. The elements of $\boldsymbol{\Sigma}_s$ are shown in Figure 1c. Note that based on the nature of the covariance Function (6), the spatial covariance matrix $\boldsymbol{\Sigma}_s$ is symmetric and based on the constellation of the candidate sensing locations (Figure 1b), $\boldsymbol{\Sigma}_s$ is also a block Toeplitz matrix. We assume that $\boldsymbol{\mu}_s$ and $\boldsymbol{\Sigma}_s$ are perfectly known a priori.

2.3. State Model

The spatio-temporal evolution of the non-stationary component of the field, i.e., $\mathbf{v}_t$, can be modelled by the following state model

$$\mathbf{v}_t = \mathbf{H}_t \mathbf{v}_{t-1} + \mathbf{q}_t. \tag{7}$$

Here, the time-varying state transition/propagator matrix is given by $\mathbf{H}_t \in \mathbb{R}^{N \times N}$. The process noise vector $\mathbf{q}_t$ is assumed to be characterized by $\mathbf{q}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_t)$. The elements of the state transition matrix $\mathbf{H}_t$ act as spatial redistribution weights for $\mathbf{v}_{t-1}$ for the temporal transition from $t-1$ to t [3]. Note that this spatial redistribution can be dependent on the temporal sampling interval. We model the elements of $\mathbf{H}_t$ by using a parameterized Gaussian kernel function

$$[\mathbf{H}_t]_{ij} = \nu \exp[-(\mathbf{x}_i - \mathbf{x}_j - \mathbf{a}_t^{ij})^T [\mathbf{D}_t^{ij}]^{-1} (\mathbf{x}_i - \mathbf{x}_j - \mathbf{a}_t^{ij})], \tag{8}$$

where $i, j = 1, \ldots, N$ and the spatio-temporally varying translation and dilation parameters are represented by $\mathbf{a}_t^{ij} \in \mathbb{R}^2$, and $\mathbf{D}_t^{ij} \in \mathbb{S}_{++}^2$, respectively. The scalar $\nu \in (0, 1)$ is a scaling parameter to avoid an explosive growth of $\mathbf{v}_t$, i.e., to keep the maximum eigenvalue of $\mathbf{H}_t$ below 1. The simplest

form of the state model (7) is defined by $\mathbf{H}_t \approx \nu \mathbf{I}_N$, which is similar to a Gaussian random walk model. This corresponds to the Gaussian kernel Function (8), with $\mathbf{D}_t = \zeta \mathbf{I}$ and $\zeta \ll 1$, and with $\mathbf{a}_t = \mathbf{0}$. In this work, we assume that the state transition matrix $\mathbf{H}_t$, whose elements are parameterized by $\{\mathbf{a}_t^{ij}\}$ and $\{\mathbf{D}_t^{ij}\}$ through the Function (8) is perfectly known a priori.

(a)

(b)

(c)
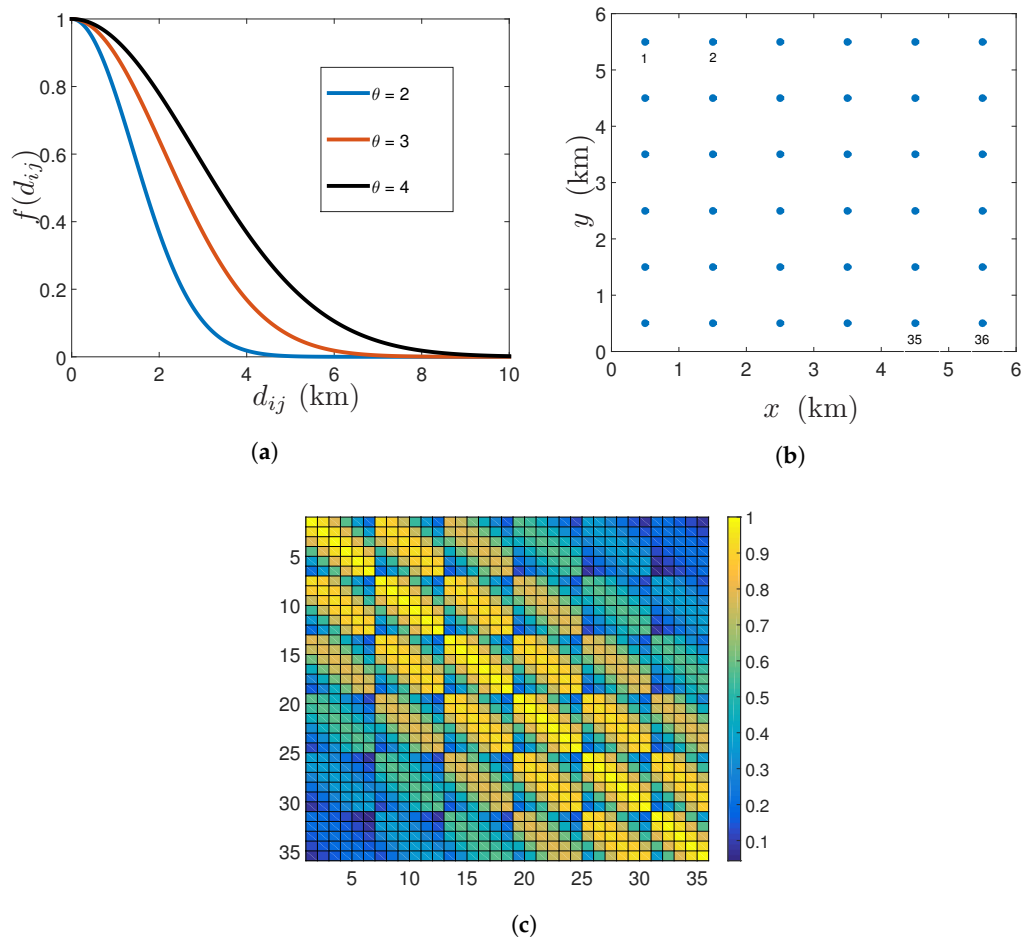
Figure 1. (**a**) Squared exponential covariance function for different values of θ (variance $\sigma_s^2 = 1$); (**b**) Service area with $N = 36$ candidate sensing locations; (**c**) Spatial covariance matrix.

From a practical point of view, the aforementioned Gaussian kernel-based spatio-temporal evolution can be used for rainfall prediction [26]. This modelling approach incorporates some physical properties of environmental fields, such as diffusion, advection etc. [26], so it can also be used for modelling the propagation of a general class of environmental fields (e.g., pollutants, aerosol movements).

2.4. Main Problem Statement

The main problem is to design an optimal sensor placement pattern, i.e., to design $\mathbf{w}_t$, whose support gives the optimal locations to deploy the sensors. At any t, the design goal is to minimize the estimation error for both the stationary and the non-stationary components of the field as well as to enforce sparsity in $\mathbf{w}_t$, i.e., to reduce the number of required sensing locations. If the estimation error of the stationary and non-stationary components of the field can be represented by a single performance metric $g(\mathbf{w}_t)$, the sensor placement problem can be represented by

$$\underset{\mathbf{w}_t \in \{0,1\}^N}{\arg\min} \quad g(\mathbf{w}_t) + \lambda_t \|\mathbf{w}_t\|_1 \tag{9a}$$

$$\text{s.t.} \quad K_t^{\min} \leq \mathbf{1}^T \mathbf{w}_t \leq K_t^{\max}. \tag{9b}$$

At any t, $K_t^{\min}$ and $K_t^{\max}$ denote the lower bound on the number of available sensors, and a given budget on the maximum number of available sensors, respectively. Sparsity is enforced through a sparsity-promoting penalty, i.e., an ℓ_1 norm of $\mathbf{w}_t$ in the second summand of the cost Function (9a) with a time-varying regularization parameter $\lambda_t > 0$ controlling the sparsity of the elements of $\mathbf{w}_t$. A detailed description regarding the structure of the objective function and the importance of the constraints in the optimization problem (9) are discussed in Section 3.1.

2.5. Simple KKF Estimator and Estimation Error Covariance

Using the measurement and state models of (1) and (7), respectively, the state estimate $\hat{\mathbf{u}}_t$, for $t = 1, 2, \ldots$, can be computed following the lines of a standard KKF [5,6]. First, a simple Kalman filter is used to track the dynamic component $\mathbf{v}_t$, where the stationary component $\mathbf{s}_t$ is interpreted as a noise term. In this case, the measurement model is given by

$$\breve{\mathbf{y}}_t = \mathbf{C}_t \mathbf{v}_t + \breve{\mathbf{e}}_t, \tag{10}$$

where $\breve{\mathbf{y}}_t = \mathbf{y}_t - \mathbf{C}_t \boldsymbol{\mu}_\mathrm{s}$, $\breve{\mathbf{e}}_t = \mathbf{C}_t \breve{\mathbf{s}}_t + \mathbf{e}_t$, and $\breve{\mathbf{s}}_t = \mathbf{s}_t - \boldsymbol{\mu}_\mathrm{s}$. Furthermore, $\breve{\mathbf{s}}_t \sim \mathcal{N}(\mathbf{0}_N, \boldsymbol{\Sigma}_\mathrm{s})$, and $\breve{\mathbf{e}}_t \sim \mathcal{N}(\mathbf{0}_{M_t}, \breve{\mathbf{R}}_t)$, with $\breve{\mathbf{R}}_t = \mathbf{C}_t \boldsymbol{\Sigma}_\mathrm{s} \mathbf{C}_t^T + \sigma_e^2 \mathbf{I}_{M_t}$. It can be easily shown that $\mathbf{v}_t$ and $\breve{\mathbf{e}}_t$ are mutually uncorrelated as it is already assumed in Section 2.1 that $\mathbf{v}_t$ is mutually uncorrelated with $\mathbf{s}_t$ and $\mathbf{e}_t$, respectively. Now, using the state model of (7) and the measurement model of (10), the non-stationary component $\mathbf{v}_t$ can be estimated following the lines of a simple Kalman filter [27]. In this case, the recursive state estimate at time t is given by

$$\hat{\mathbf{v}}_t = \mathbf{H}_t \hat{\mathbf{v}}_{t-1} + \mathbf{G}_t (\breve{\mathbf{y}}_t - \mathbf{C}_t \mathbf{H}_t \hat{\mathbf{v}}_{t-1}), \tag{11}$$

where the Kalman gain $\mathbf{G}_t$ can be expressed as

$$\begin{aligned} \mathbf{G}_t = [\mathbf{H}_t \mathbf{M}_{\mathbf{v}_{t-1}} \mathbf{H}_t^T + \mathbf{Q}_t] \mathbf{C}_t^T \times \\ [\breve{\mathbf{R}}_t + \mathbf{C}_t (\mathbf{H}_t \mathbf{M}_{\mathbf{v}_{t-1}} \mathbf{H}_t^T + \mathbf{Q}_t) \mathbf{C}_t^T]^{-1}. \end{aligned} \tag{12}$$

The MSE matrix of the estimate $\hat{\mathbf{v}}_t$ at time t is given by $\mathbb{E}[(\mathbf{v}_t - \hat{\mathbf{v}}_t)(\mathbf{v}_t - \hat{\mathbf{v}}_t)^T] = \mathbf{M}_{\mathbf{v}_t}$, which is related to the MSE matrix of the estimate at time $t - 1$, i.e., $\mathbf{M}_{\mathbf{v}_{t-1}}$, by the recursive relation [27]

$$\mathbf{M}_{\mathbf{v}_t} = [(\mathbf{H}_t \mathbf{M}_{\mathbf{v}_{t-1}} \mathbf{H}_t^T + \mathbf{Q}_t)^{-1} + \mathbf{C}_t^T \breve{\mathbf{R}}_t^{-1} \mathbf{C}_t]^{-1}. \tag{13}$$

In the next stage, the estimate of $\hat{\mathbf{v}}_t$ in (11) is used to compute the stationary component $\mathbf{s}_t$ using kriging. In spatial statistics, the intensity of an environmental field in an unknown location can be interpolated using a variogram model [3,4]. For a spatially stationary field, this variogram can be expressed as a covariance [3]. In this way kriging can be viewed as a simple linear minimum mean square error (LMMSE) interpolator [27]. The linear model is given by $\mathbf{y}_t - \mathbf{C}_t \hat{\mathbf{v}}_t = \mathbf{C}_t \mathbf{s}_t + \mathbf{e}_t$ and the related estimator has the form

$$\hat{\mathbf{s}}_t = \boldsymbol{\mu}_\mathrm{s} + \boldsymbol{\Sigma}_\mathrm{s} \mathbf{C}_t^T (\mathbf{C}_t \boldsymbol{\Sigma}_\mathrm{s} \mathbf{C}_t^T + \sigma_e^2 \mathbf{I}_M)^{-1} (\mathbf{y}_t - \mathbf{C}_t \hat{\mathbf{v}}_t - \mathbf{C}_t \boldsymbol{\mu}_\mathrm{s}), \tag{14}$$

where we use the prior information $\mathbf{s}_t \sim \mathcal{N}(\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s)$. The MSE matrix [27] of the estimate $\hat{\mathbf{s}}_t$, i.e., $\mathbf{M}_{\mathbf{s}_t}$ is given by

$$\mathbf{M}_{\mathbf{s}_t} = [\boldsymbol{\Sigma}_\mathrm{s}^{-1} + \sigma_e^{-2} \mathbf{C}_t^T \mathbf{C}_t]^{-1}. \tag{15}$$

Finally, the overall field estimate at time t is given by $\hat{\mathbf{u}}_t = \hat{\mathbf{v}}_t + \hat{\mathbf{s}}_t$.

2.6. Performance Metrics as a Function of $\mathbf{w}_t$

In this section, we express the MSE matrices, i.e., $\mathbf{M}_{\mathbf{v}_t}$ and $\mathbf{M}_{\mathbf{s}_t}$ as functions of $\mathbf{w}_t$. First of all, we mention some facts regarding the structure of the error covariance matrices presented in the Expressions (13) and (15).

It should be noted that the measurement noise in (10) is correlated over space through the off-diagonal elements of $\check{\mathbf{R}}_t$. Due to this fact, sensor selection approaches using the standard convex framework like [18,20,22], i.e., designing a $\mathbf{w}_t$ by directly optimizing the Expression (13) is difficult due to the presence of the off-diagonal elements of $\check{\mathbf{R}}_t$. It should also be noted that the expression of $\check{\mathbf{R}}_t$ is a function of the measurement matrix $\mathbf{C}_t$ and thus the selection vector $\mathbf{w}_t$. However, we do not encounter this issue in the performance metric to estimate the stationary component $\mathbf{s}_t$, i.e., (15), as the measurement noise $\mathbf{e}_t$ is assumed to be spatially uncorrelated in this case.

In the expression of $\mathbf{M}_{\mathbf{s}_t}$, i.e., (15), we assume that $\boldsymbol{\Sigma}_s$ is well-conditioned, i.e., accurately invertible. However, this may not be the case in some scenarios. The condition number of $\boldsymbol{\Sigma}_s$ strongly depends on the correlation of the field, spatial sampling distance, grid size etc. [25]. The variation of the condition number of $\boldsymbol{\Sigma}_s$ with different values of θ for both $N = 36$ and $N = 144$ is plotted in Figure 2a. It is seen that for a higher resolution or a strong spatial correlation, the spatial covariance matrix becomes increasingly ill-conditioned and thus close to singular. In such circumstances, we cannot compute the estimation error covariance matrix $\mathbf{M}_{\mathbf{s}_t}$ using the Expression (15). In that case, $\mathbf{M}_{\mathbf{s}_t}$ can be computed using the alternate expression of (15) given by

$$
\begin{aligned}
\mathbf{M}_{\mathbf{s}_t} &= [\boldsymbol{\Sigma}_s^{-1} + \sigma_e^{-2}\mathbf{C}_t^T\mathbf{C}_t]^{-1} \\
&= \boldsymbol{\Sigma}_s - \boldsymbol{\Sigma}_s\mathbf{C}_t^T(\mathbf{C}_t\boldsymbol{\Sigma}_s\mathbf{C}_t^T + \sigma_e^2\mathbf{I}_{M_t})^{-1}\mathbf{C}_t\boldsymbol{\Sigma}_s,
\end{aligned}
\tag{16}
$$

which is obtained using the matrix inversion lemma (MIL). It should be noted that the alternative expression of the MSE can be used to compute the MSE (without inverting $\boldsymbol{\Sigma}_s$), but it is difficult to express it as an explicit function of $\mathbf{w}_t$.
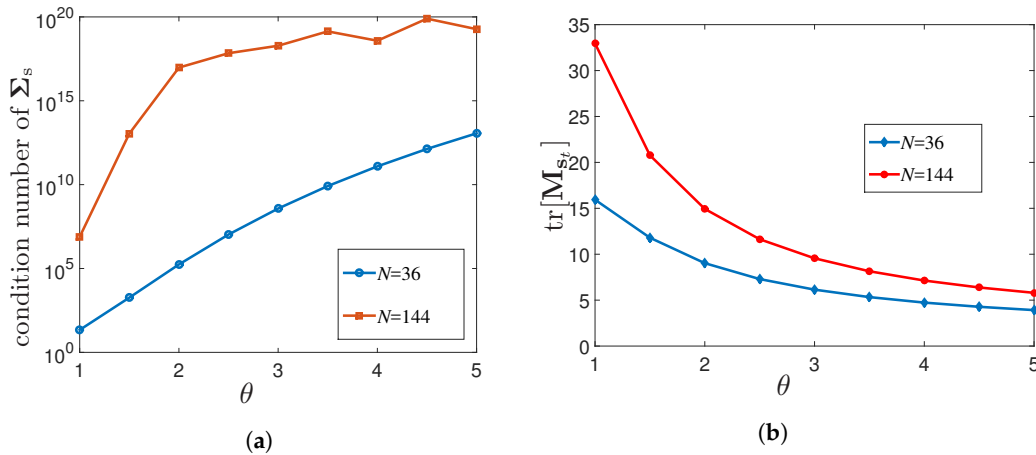
Figure 2. (**a**) Plot of the condition number of $\boldsymbol{\Sigma}_s$ vs. θ with different number of candidate sensing locations (N); (**b**) MSE of the estimate of $\mathbf{s}_t$ vs. θ for different numbers of candidate sensing locations (N); $M_t = N$; $\sigma_e^2 = 1$; The spatial resolution is increased by representing one pixel of Figure 1a by 4 pixels.

In Figure 2b, we plot $\text{tr}[\mathbf{M}_{\mathbf{s}_t}]$ for the best case, i.e., the MSE with all the pixel centroids equipped with sensors ($\mathbf{w}_t = \mathbf{1}_N$ or $\mathbf{C}_t = \mathbf{I}_N$) for different values of θ, and for two different spatial resolutions ($N = 36$ and $N = 144$) of the same 6×6 square km service area. It is seen that $\text{tr}[\mathbf{M}_{\mathbf{s}_t}]$ decreases as the strength of the correlation is increased by increasing θ.

To circumvent the effect of ill-conditioning as well as to handle the correlated measurement noise in the expression of $\mathbf{M}_{\mathbf{v}_t}$, we propose the following approach. We start by introducing the substitution

$$\boldsymbol{\Sigma}_{\mathrm{sr}} = \boldsymbol{\Sigma}_{\mathrm{s}} + \alpha \mathbf{I}, \tag{17}$$

where $\boldsymbol{\Sigma}_{\mathrm{sr}}$ is a well-conditioned matrix and $\alpha \in \mathbb{R}$. More specifically, the range of α is considered to be $0 < \alpha < \sigma_e^2$. Substituting $\boldsymbol{\Sigma}_{\mathrm{s}} = \boldsymbol{\Sigma}_{\mathrm{sr}} - \alpha \mathbf{I}$, we can represent the measurement error covariance matrix of (10) as, $\check{\mathbf{R}}_t = \mathbf{C}_t \boldsymbol{\Sigma}_{\mathrm{sr}} \mathbf{C}_t^T + \zeta \mathbf{I}_{M_t}$, where $\zeta = \sigma_e^2 - \alpha$ and where we used $\mathbf{C}_t \mathbf{C}_t^T = \mathbf{I}_{M_t}$. Substituting $\check{\mathbf{R}}_t = \mathbf{C}_t \boldsymbol{\Sigma}_{\mathrm{sr}} \mathbf{C}_t^T + \zeta \mathbf{I}_{M_t}$ in (13), the MSE matrix for the estimate of the non-stationary component is given by

$$\mathbf{M}_{\mathbf{v}_t} = [(\mathbf{H}_t \mathbf{M}_{\mathbf{v}_{t-1}} \mathbf{H}_t^T + \mathbf{Q}_t)^{-1} + \\ \mathbf{C}_t^T (\mathbf{C}_t \boldsymbol{\Sigma}_{\mathrm{sr}} \mathbf{C}_t^T + \zeta \mathbf{I}_{M_t})^{-1} \mathbf{C}_t]^{-1}. \tag{18}$$

Using the MIL, we have the following matrix identity

$$(\boldsymbol{\Sigma}_{\mathrm{sr}}^{-1} + \mathbf{C}_t^T (\zeta \mathbf{I}_{M_t})^{-1} \mathbf{C}_t)^{-1} = \boldsymbol{\Sigma}_{\mathrm{sr}} \\ - \boldsymbol{\Sigma}_{\mathrm{sr}} \mathbf{C}_t^T (\mathbf{C}_t \boldsymbol{\Sigma}_{\mathrm{sr}} \mathbf{C}_t^T + \zeta \mathbf{I}_{M_t})^{-1} \mathbf{C}_t \boldsymbol{\Sigma}_{\mathrm{sr}}, \tag{19}$$

from which we can derive

$$\mathbf{C}_t^T (\mathbf{C}_t \boldsymbol{\Sigma}_{\mathrm{sr}} \mathbf{C}_t^T + \zeta \mathbf{I}_{M_t})^{-1} \mathbf{C}_t = \boldsymbol{\Sigma}_{\mathrm{sr}}^{-1} \\ - \boldsymbol{\Sigma}_{\mathrm{sr}}^{-1} (\boldsymbol{\Sigma}_{\mathrm{sr}}^{-1} + \zeta^{-1} \mathrm{diag}(\mathbf{w}_t))^{-1} \boldsymbol{\Sigma}_{\mathrm{sr}}^{-1}, \tag{20}$$

where we used $\mathbf{C}_t^T \mathbf{C}_t = \mathrm{diag}(\mathbf{w}_t)$. Substituting (20) in (18) we obtain the following expression for $\mathbf{M}_{\mathbf{v}_t}$.

$$\mathbf{M}_{\mathbf{v}_t} = [(\mathbf{H}_t \mathbf{M}_{\mathbf{v}_{t-1}} \mathbf{H}_t^T + \mathbf{Q}_t)^{-1} + \boldsymbol{\Sigma}_{\mathrm{sr}}^{-1} \\ - \boldsymbol{\Sigma}_{\mathrm{sr}}^{-1} (\boldsymbol{\Sigma}_{\mathrm{sr}}^{-1} + \zeta^{-1} \mathrm{diag}(\mathbf{w}_t))^{-1} \boldsymbol{\Sigma}_{\mathrm{sr}}^{-1}]^{-1}. \tag{21}$$

Next, substituting $\boldsymbol{\Sigma}_{\mathrm{s}} = \boldsymbol{\Sigma}_{\mathrm{sr}} - \alpha \mathbf{I}$ in the inverse of the right most term of (16) and using $\mathbf{C}_t \mathbf{C}_t^T = \mathbf{I}_{M_t}$, we obtain

$$\mathbf{M}_{\mathbf{s}_t} = \boldsymbol{\Sigma}_{\mathrm{s}} - \boldsymbol{\Sigma}_{\mathrm{s}} \mathbf{C}_t^T (\mathbf{C}_t \boldsymbol{\Sigma}_{\mathrm{sr}} \mathbf{C}_t^T + \zeta \mathbf{I}_{M_t})^{-1} \mathbf{C}_t \boldsymbol{\Sigma}_{\mathrm{s}}. \tag{22}$$

Substituting the identity (20) into (22), we obtain the following expression of $\mathbf{M}_{\mathbf{s}_t}$.

$$\mathbf{M}_{\mathbf{s}_t} = \boldsymbol{\Sigma}_{\mathrm{s}} - \boldsymbol{\Sigma}_{\mathrm{s}} \boldsymbol{\Sigma}_{\mathrm{sr}}^{-1} \boldsymbol{\Sigma}_{\mathrm{s}} \\ + \boldsymbol{\Sigma}_{\mathrm{s}} \boldsymbol{\Sigma}_{\mathrm{sr}}^{-1} (\boldsymbol{\Sigma}_{\mathrm{sr}}^{-1} + \zeta^{-1} \mathrm{diag}(\mathbf{w}_t))^{-1} \boldsymbol{\Sigma}_{\mathrm{sr}}^{-1} \boldsymbol{\Sigma}_{\mathrm{s}}. \tag{23}$$

Note that, the expression of (23) avoids the inversion of an ill-conditioned $\boldsymbol{\Sigma}_{\mathrm{s}}$. Here, we only need to invert the well-conditioned $\boldsymbol{\Sigma}_{\mathrm{sr}}$.

In this work, we consider the overall MSE as a performance metric for sensor placement, i.e., $g(\mathbf{w}_t)$ as mentioned in (9a). This is given by

$$g(\mathbf{w}_t) = \mathrm{tr}(\mathbf{M}_{\mathbf{v}_t}) + \mathrm{tr}(\mathbf{M}_{\mathbf{s}_t}) \\ = \mathrm{tr}[\mathbf{X} - \mathbf{F}[\mathbf{F} + \zeta^{-1} \mathrm{diag}(\mathbf{w}_t)]^{-1} \mathbf{F}]^{-1} + \mathrm{tr}[\mathbf{Y}] \\ + \mathrm{tr}[\mathbf{Z}^T [\mathbf{F} + \zeta^{-1} \mathrm{diag}(\mathbf{w}_t)]^{-1} \mathbf{Z}], \tag{24}$$

where $\mathbf{X} = (\mathbf{H}_t \mathbf{M}_{\mathbf{v}_{t-1}} \mathbf{H}_t^T + \mathbf{Q}_t)^{-1} + \boldsymbol{\Sigma}_{\mathrm{sr}}^{-1}$, $\mathbf{F} = \boldsymbol{\Sigma}_{\mathrm{sr}}^{-1}$, $\mathbf{Y} = \boldsymbol{\Sigma}_{\mathrm{s}} - \boldsymbol{\Sigma}_{\mathrm{s}} \boldsymbol{\Sigma}_{\mathrm{sr}}^{-1} \boldsymbol{\Sigma}_{\mathrm{s}}$, and $\mathbf{Z} = \boldsymbol{\Sigma}_{\mathrm{sr}}^{-1} \boldsymbol{\Sigma}_{\mathrm{s}}$. Note that the matrices $\mathbf{X}$, $\mathbf{F}$, $\mathbf{Y}$, and $\mathbf{Z}$ are all independent of $\mathbf{w}_t$. To model $\boldsymbol{\Sigma}_{\mathrm{sr}}$ and $\mathbf{F} + \zeta^{-1} \mathrm{diag}(\mathbf{w}_t)$ as positive definite matrices we need $0 < \alpha < \sigma_e^2$.

The performance metric derived in (24) incorporates the MSE matrices of the estimates of the non-stationary ($\mathbf{v}_t$) as well as the stationary ($\mathbf{s}_t$) component of the field, as explicit functions of the

sensor location selection vector $\mathbf{w}_t$. Note that a formulation similar to (23), for the computation of the MSE matrix as a function of $\mathbf{w}_t$ is proposed in [22], where the field is considered to be either purely stationary or non-stationary.

3. KKF with Sensor Placement

In this section, we relax and reformulate the proposed sensor placement problem (9) as a semidefinite programming (SDP) problem. Then we present the overall KKF estimator followed by the sensor placement to dynamically monitor the field using only the measurements from the selected sensing locations.

3.1. Sensor Placement Problem as an SDP

Solving for the best subset of sensing locations is a combinatorially complex problem. However, it can be relaxed to a convex problem [18–20]. As discussed in Section 2.1, the sensor location selection vector $\mathbf{w}_t \in \{0,1\}^N$ acts as a weighting vector for all the N candidate pixels. Following the main optimization problem, i.e., (9), a sparsity-enforcing, low estimation error, and resource-constrained design of $\mathbf{w}_t$ can be obtained by solving

$$\underset{\mathbf{w}_t \in [0,1]^N}{\arg\min} \quad g(\mathbf{w}_t) + \lambda_t \|\mathbf{w}_t\|_1 \tag{25a}$$

$$\text{s.t.} \quad K_t^{\min} \leq \mathbf{1}^T \mathbf{w}_t \leq K_t^{\max}, \tag{25b}$$

where the expression of $g(\mathbf{w}_t)$ is given by (24). Here, we have relaxed the non-convex Boolean constraint $\mathbf{w}_t \in \{0,1\}^N$ of (9) to a convex box constraint $\mathbf{w}_t \in [0,1]^N$. The resource constraint of (25b) is affine and thus convex. Some comments regarding the formulation of the proposed sensor placement problem of (25) are presented next.

- First of all, let us consider the non-convex version of the optimization problem of (25) with $\lambda_t = 0$. This is given as

$$\underset{\mathbf{w}_t \in \{0,1\}^N}{\arg\min} \quad g(\mathbf{w}_t) \tag{26a}$$

$$\text{s.t.} \quad K_t^{\min} \leq \mathbf{1}^T \mathbf{w}_t \leq K_t^{\max}. \tag{26b}$$

  In this case, the MSE cost will be minimum, i.e., the best estimation performance is achieved, when we select the maximum number of available candidate locations or in other words, when $\mathbf{1}^T \mathbf{w}_t = K_t^{\max}$. Then, there is no way to reduce the number of selected locations below $K_t^{\max}$ and the constraint $\mathbf{1}^T \mathbf{w}_t \geq K_t^{\min}$ becomes redundant. In the aforementioned case, it is difficult to reduce the number of selected sensing locations below $K_t^{\max}$.

- Notice that, dropping the resource constraint (25b) and increasing λ_t will reduce the number of selected sensing locations. However, there is no explicit relation between λ_t and $\mathbf{1}^T \mathbf{w}_t$, i.e., it is difficult to directly control the resource allocation (i.e., $K_t^{\max}$) through λ_t.

- We mention that the proposed formulation of (25) is not a direct MSE minimization problem but it attains a specific MSE along with enforcing sparsity in spatial sensor location selection through the second summand of (25a). The sparsity enforcement is lower bounded by the minimum number of sensing locations to be selected at any t, i.e., $K_t^{\min}$. It should be noted that for an arbitrary selection of λ_t, the minimum number of selected sensing locations will always be $K_t^{\min}$.

- Lastly, it should be noted that a sparsity-enforcing design of $\mathbf{w}_t$ can be achieved by retaining only the second summand of the objective function of (25a) and using a separate performance constraint given as $g(\mathbf{w}_t) \leq \gamma_{t,\text{MSE}}$ [20,22]. The desired performance threshold $\gamma_{t,\text{MSE}}$ can be time-varying or independent of t based on the application. However, in many practical scenarios, it could be difficult to set the performance threshold $\gamma_{t,\text{MSE}}$ a priori for every t.

Based on the aforementioned arguments, we advocate the proposed design approach (25) that lowers the MSE along with enforcing sparsity in sensor placement satisfying a flexible resource allocation constraint.

After solving (25), we obtain $\hat{\mathbf{w}}_t \in [0,1]^N$ which can be converted to a Boolean selection vector $\mathbf{w}_t \in \{0,1\}^N$. This can be performed by either deterministic or stochastic rounding procedures as discussed below.

- The simplest approach could be to set the non-zero entries of $\hat{\mathbf{w}}_t$ to 1. However, there can be a huge difference between the magnitudes of any two non-zero elements in $\hat{\mathbf{w}}_t$. Considering the fact that the indices of the high magnitude (close to 1) elements of $\hat{\mathbf{w}}_t$ signify a more informative sensing location, $\hat{\mathbf{w}}_t$ can be sorted in ascending order of magnitude [18] and a selection threshold (γ) can be selected based on the magnitudes of the elements of the sorted $\hat{\mathbf{w}}_t$. The entries of the Boolean selection vector can be computed as $[\mathbf{w}_t]_j = 1$ if $[\hat{\mathbf{w}}_t]_j \geq \gamma$ else $[\mathbf{w}_t]_j = 0$, for $j = 1, \dots, N$.
- Another approach could be a stochastic approach, where every entry of $\hat{\mathbf{w}}_t$ is assumed to be the probability that this sensing location is selected at time t. Based on this, multiple random realizations of $\mathbf{w}_t \in \{0,1\}^N$ are generated, where the probability that $[\mathbf{w}_t]_j = 1$ is given by $[\hat{\mathbf{w}}_t]_j$, for $j = 1, \dots, N$. Then the realization that satisfies the constraints and minimizes the estimation error, i.e., $g(\mathbf{w}_t)$ is selected [20].

Let us now transform the optimization problem of (25) into an SDP. From the expression of (24), it is clear that minimizing $g(\mathbf{w}_t)$ w.r.t. $\mathbf{w}_t$ is equivalent to minimizing the expression $\text{tr}[\mathbf{X} - \mathbf{F}[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)]^{-1}\mathbf{F}]^{-1} + \text{tr}[\mathbf{Z}^T[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)]^{-1}\mathbf{Z}]$ as the matrix $\text{tr}[\mathbf{Y}]$ is independent of $\mathbf{w}_t$. In the first step, we represent the optimization problem of (25) in an epigraph form ([28], p. 134), ([18], Equations (25) and (26)) which is given by

$$\underset{\mathbf{w}_t \in [0,1]^N, \mathbf{V} \in \mathbb{S}^N, \mathbf{B} \in \mathbb{S}^N}{\arg\min} \quad \text{tr}[\mathbf{V}] + \text{tr}[\mathbf{B}] + \lambda_t \|\mathbf{w}_t\|_1 \tag{27a}$$

$$\text{s.t.} \quad \mathbf{V} \succeq [\mathbf{X} - \mathbf{F}[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)]^{-1}\mathbf{F}]^{-1}, \tag{27b}$$

$$\mathbf{B} \succeq \mathbf{Z}^T[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)]^{-1}\mathbf{Z}, \tag{27c}$$

$$K_t^{\min} \leq \mathbf{1}^T\mathbf{w}_t \leq K_t^{\max}, \tag{27d}$$

where we introduce the auxiliary variables $\mathbf{V} \in \mathbb{S}^N$ and $\mathbf{B} \in \mathbb{S}^N$. We notice that the epigraph form (27) is well-posed since by choosing $0 < \alpha < \sigma_e^2$ in (17) the matrix $[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)]$ is always positive definite and symmetric. In addition, also the matrix $[\mathbf{X} - \mathbf{F}[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)]^{-1}\mathbf{F}]$ is also positive definite by construction as derived in (18)–(21).

The epigraph form (27) is not a strictly convex program, in the sense that there are multiple $\mathbf{V}$ and $\mathbf{B}$ matrices that achieve the minimal cost value. This is due to the inequality constraints of (27b) and (27c). At optimality, the eigenvalues of $\mathbf{V}$ and $\mathbf{B}$ must be equivalent to their lower bounds in (27b) and (27c). Hence, an optimizer of the problem is $\mathbf{V} = [\mathbf{X} - \mathbf{F}[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)]^{-1}\mathbf{F}]^{-1}$ and $\mathbf{B} = \mathbf{Z}^T[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)]^{-1}\mathbf{Z}$.

We proceed by simplifying the constraint (27b). Let us introduce another auxiliary variable $\mathbf{A} \in \mathbb{S}^N$ and substitute (27b) with two constraints

$$\mathbf{V} \succeq [\mathbf{X} - \mathbf{A}]^{-1}, \tag{28}$$

$$\mathbf{A} \succeq \mathbf{F}[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)]^{-1}\mathbf{F}. \tag{29}$$

With this in place, the optimization Problem (27) can be formulated as

$$\arg\min_{\mathbf{w}_t \in [0,1]^N, \mathbf{V}, \mathbf{A}, \mathbf{B} \in \mathbb{S}^N} \quad \text{tr}[\mathbf{V}] + \text{tr}[\mathbf{B}] + \lambda_t \|\mathbf{w}_t\|_1 \tag{30a}$$

$$\text{s.t.} \quad \mathbf{V} \succeq [\mathbf{X} - \mathbf{A}]^{-1}, \tag{30b}$$

$$\mathbf{A} \succeq \mathbf{F}[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)]^{-1}\mathbf{F}, \tag{30c}$$

$$\mathbf{B} \succeq \mathbf{Z}^T[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)]^{-1}\mathbf{Z}, \tag{30d}$$

$$K_t^{\min} \leq \mathbf{1}^T\mathbf{w}_t \leq K_t^{\max}, \tag{30e}$$

It can be claimed that the optimization problem (30) is equivalent to (27) given that it yields a decision variable $\mathbf{w}_t$ with the same optimal cost of (27) . To prove this, let us choose an arbitrary $\mathbf{w}_t$ say $\bar{\mathbf{w}}$. For a fixed yet arbitrary $\bar{\mathbf{w}}$ verifying (30e), the optimization problem (30) minimizes both $\mathbf{V}$ and $\mathbf{B}$. This means that due to (30b) it minimizes also $\mathbf{A}$: in fact, as $\mathbf{V} \succeq [\mathbf{X} - \mathbf{A}]^{-1}$ the lower bound for $\mathbf{V}$ is minimal if the positive definite matrix $[\mathbf{X} - \mathbf{A}]$ is maximal, that is $\mathbf{A}$ is minimal. Therefore, $\mathbf{A}$ must attain its lower bound. As mentioned earlier, there are multiple optimizers, yet one is $\mathbf{A} = \mathbf{F}[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)]^{-1}\mathbf{F}$. In addition, $\mathbf{V} = [\mathbf{X} - \mathbf{A}]^{-1} = [\mathbf{X} - \mathbf{F}[\mathbf{F} + \zeta^{-1}\text{diag}(\bar{\mathbf{w}})]^{-1}\mathbf{F}]^{-1}$ at optimality, as well. The same reasoning holds also for $\mathbf{B}$, which at optimality is $\mathbf{B} = \mathbf{Z}^T[\mathbf{F} + \zeta^{-1}\text{diag}(\bar{\mathbf{w}})]^{-1}\mathbf{Z}$. Since this reasoning is valid for any feasible $\bar{\mathbf{w}}$, it is also valid for an optimal one and therefore the equivalence claim follows. It should be noted that a similar argument was also presented in [24], where only the issue of correlated measurement noise is considered.

Using the Schur complement lemma the constraints (30b) and (30c) can be equivalently represented by the linear matrix inequalities (LMI) :

$$\begin{bmatrix} \mathbf{X} - \mathbf{A} & \mathbf{I} \\ \mathbf{I} & \mathbf{V} \end{bmatrix} \succeq \mathbf{0} \tag{31a}$$

$$\begin{bmatrix} \mathbf{A} & \mathbf{F} \\ \mathbf{F} & \mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t) \end{bmatrix} \succeq \mathbf{0} \tag{31b}$$

The constraint (30c) can be equivalently represented by an LMI using the Schur complement [28]. In other words, using the fact that $[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)] \succ \mathbf{0}$, we obtain

$$\begin{bmatrix} \mathbf{B} & \mathbf{Z}^T \\ \mathbf{Z} & \mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t) \end{bmatrix} \succeq \mathbf{0}. \tag{32}$$

Finally, an SDP representation of the overall optimization problem of (27) can be expressed as

$$\arg\min_{\mathbf{w}_t \in [0,1]^N, \mathbf{A}, \mathbf{B}, \mathbf{V} \in \mathbb{S}^N} \quad \text{tr}[\mathbf{V}] + \text{tr}[\mathbf{B}] + \lambda_t \|\mathbf{w}_t\|_1, \tag{33a}$$

$$\text{s.t.} \quad \text{LMIs in } (31a), (31b), (32) \tag{33b}$$

$$K_t^{\min} \leq \mathbf{1}^T\mathbf{w}_t \leq K_t^{\max} \tag{33c}$$

The solution of the aforementioned SDP is a selection vector $\hat{\mathbf{w}}_t \in [0,1]^N$.

3.2. Spatial Sensor Placement for Stationary Field Estimation

Let us consider the effect of the stationary component of the field $\mathbf{s}_t$ for any t. In this case, we consider that $\mathbf{v}_t = \mathbf{0}$. In this case, the measurement model of (1) is given by $\mathbf{y}_t = \mathbf{C}_t\mathbf{s}_t + \mathbf{e}_t$. Exploiting the prior information regarding $\mathbf{s}_t$, i.e., $\mathbf{s}_t \sim \mathcal{N}(\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s)$ an LMMSE estimator of $\mathbf{s}_t$ can be presented by $\hat{\mathbf{s}}_t = \boldsymbol{\mu}_s + \boldsymbol{\Sigma}_s\mathbf{C}_t^T(\mathbf{C}_t\boldsymbol{\Sigma}_s\mathbf{C}_t^T + \sigma_e^2\mathbf{I}_{M_t})^{-1}(\mathbf{y}_t - \mathbf{C}_t\boldsymbol{\mu}_s)$. The performance of the aforementioned estimator is given by the MSE matrix $\mathbf{M}_{\mathbf{s}_t} = [\boldsymbol{\Sigma}_s^{-1} + \sigma_e^{-2}\mathbf{C}_t^T\mathbf{C}_t]^{-1} = \boldsymbol{\Sigma}_s - \boldsymbol{\Sigma}_s\mathbf{C}_t^T(\mathbf{C}_t\boldsymbol{\Sigma}_s\mathbf{C}_t^T + \sigma_e^2\mathbf{I}_{M_t})^{-1}\mathbf{C}_t\boldsymbol{\Sigma}_s$. Considering the fact that $\boldsymbol{\Sigma}_s$ can be ill-conditioned, following the formulation of (23), the expression of $\mathbf{M}_{\mathbf{s}_t}$ can be expressed as a function of $\mathbf{w}_t$ as

$$\mathbf{M}_{\mathbf{s}_t} = \mathbf{Y} + \mathbf{Z}^T[\mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t)]^{-1}\mathbf{Z}, \tag{34}$$

where $\mathbf{Y} = \boldsymbol{\Sigma}_s - \boldsymbol{\Sigma}_s\boldsymbol{\Sigma}_{sr}^{-1}\boldsymbol{\Sigma}_s$, $\mathbf{Z} = \boldsymbol{\Sigma}_{sr}^{-1}\boldsymbol{\Sigma}_s$, and $\mathbf{F} = \boldsymbol{\Sigma}_{sr}^{-1}$. Note that the matrices $\mathbf{F}$, $\mathbf{Y}$, and $\mathbf{Z}$ are all independent of $\mathbf{w}_t$. Considering $g(\mathbf{w}_t) = \text{tr}[\mathbf{M}_{\mathbf{s}_t}]$ and following the same SDP formulation of Section 3.1, the proposed sensor placement problem of (9) can be represented as

$$\underset{\mathbf{w}_t \in [0,1]^N, \mathbf{B} \in \mathbb{S}^N}{\arg\min} \quad \text{tr}[\mathbf{B}] + \lambda_t\|\mathbf{w}_t\|_1, \tag{35a}$$

$$\text{s.t.} \quad \begin{bmatrix} \mathbf{B} & \mathbf{Z}^T \\ \mathbf{Z} & \mathbf{F} + \zeta^{-1}\text{diag}(\mathbf{w}_t) \end{bmatrix} \succeq \mathbf{0}, \tag{35b}$$

$$K_t^{\min} \le \mathbf{1}^T\mathbf{w}_t \le K_t^{\max}. \tag{35c}$$

The optimization problem of (35) gives the spatial sensor placement pattern for any snapshot t, when the field is stationary over space. However, if the field is also temporally stationary then the sensor placement problem of (35) can be extended to blocks of multiple snapshots. In this case, the performance metric can be computed using the same approach as [22]. In the simulation section, we show the effects of spatial correlation on sensor placement.

3.3. Sparsity-Enhancing Iterative Design

In order to eschew the effect of the magnitude dependencies of the elements of $\hat{\mathbf{w}}_t$, we individually weigh each element of $\mathbf{w}_t$. In this case, we consider a vector form for the regularization parameter : $\boldsymbol{\lambda}_t \in \mathbb{R}^N$. The weight associated to the each element of $\mathbf{w}_t$ is the corresponding element of $\boldsymbol{\lambda}_t \in \mathbb{R}^N$. We iteratively refine the weighting vector $\boldsymbol{\lambda}_t$ in the ℓ_1 minimization term of the problem (33) [29]. Using this approach, higher weights are applied on the smaller elements of $\mathbf{w}_t$ to push them towards 0 and the magnitudes of the larger elements are maintained by applying a smaller weight. In this way, a sparser solution can be obtained compared to the standard sparsity-promoting method. The iterative algorithm can be summarized as

- **Initialize** $i = 0$, weight vector $\boldsymbol{\lambda}_t^0 = \mathbf{1}_N$, ϵ, and maximum number of iterations I.
- **for** $i = 0, \dots, I$

$$\hat{\mathbf{w}}_t^i = \underset{\mathbf{w}_t \in [0,1]^N, \mathbf{A}, \mathbf{B}, \mathbf{V} \in \mathbb{S}^N}{\arg\min} \quad \text{tr}[\mathbf{V}] + \text{tr}[\mathbf{B}] + (\boldsymbol{\lambda}_t^i)^T\mathbf{w}_t, \tag{36a}$$

$$\text{s.t.} \quad \text{LMIs in } (31a), (31b), (32) \tag{36b}$$

$$K_t^{\min} \le \mathbf{1}^T\mathbf{w}_t \le K_t^{\max} \tag{36c}$$

- $[\boldsymbol{\lambda}_t^{i+1}]_j = \frac{1}{\epsilon + [\hat{\mathbf{w}}_t^i]_j}$, for every $j = 1, \dots, N$
- **end**;
- **set** $\hat{\mathbf{w}}_t = \hat{\mathbf{w}}_t^I$.

After solving the above algorithm, we still obtain $\hat{\mathbf{w}}_t \in [0,1]^N$. We convert this to a Boolean selection vector $\mathbf{w}_t \in \{0,1\}^N$ using a deterministic/stochastic rounding method as mentioned in Section 3.1.

3.4. KKF Algorithm with Sensor Placement

The informative M_t locations to deploy/move the sensors at any t is denoted by $\text{supp}(\hat{\mathbf{w}}_t)$, where $\mathbf{1}^T\hat{\mathbf{w}}_t = M_t$. The noisy measurements collected from the aforementioned M_t locations are stored in $\mathbf{y}_t$. The sensing matrix $\mathbf{C}_t$ is constructed by removing the all-zero rows of $\text{diag}(\hat{\mathbf{w}}_t)$ at every t. This measurement matrix is used for the estimation of the non-stationary and the stationary components by (11) and (14), respectively. Then the overall field estimate at time t is computed by $\hat{\mathbf{u}}_t = \hat{\mathbf{v}}_t + \hat{\mathbf{s}}_t$. Note that the estimation steps, i.e., (11) and (14) do not require the computation of the

inverse of $\boldsymbol{\Sigma}_s$. The error covariance of the non-stationary component can be updated by (13), which also does not require the inverse of $\boldsymbol{\Sigma}_s$. At every t, the overall estimation performance can be computed by the expression of (24). The best case performance, i.e., the performance with all the locations selected can also be computed by the expression of (24) by using $\mathbf{w}_t = \mathbf{1}_N$.

In many practical environmental fields (such as rainfall), the field is generally non-negative. To achieve a non-negative estimate at every t, the estimates of the stationary and non-stationary components can be projected onto the non-negative orthant, i.e., the negative values are set to zero. This is obtained by adopting

$$\hat{\mathbf{u}}_t = [\hat{\mathbf{v}}_t + \hat{\mathbf{s}}_t]_+. \tag{37}$$

However, in this case, the performance metrics $\mathrm{tr}[\mathbf{M}_{\mathbf{v}_t}]$ and $\mathrm{tr}[\mathbf{M}_{\mathbf{s}_t}]$ are only the approximations. The overall sensor placement followed by a KKF algorithm is presented in Algorithm 1.

Algorithm 1 Sensor placement followed by a KKF estimator

1: **Initialize:** $t = 0$, $\hat{\mathbf{v}}_t$, $\mathbf{M}_{\hat{\mathbf{v}}_t}$, $\lambda_t^0 = \mathbf{1}_N$.

2: **Given:** $\{\mathbf{H}_t\}_{t=1}^T$, $\mathbf{Q}_t = \mathbf{Q}$, $\boldsymbol{\Sigma}_s$, σ_e^2, I, ϵ, $K_t^{\max}$, $K_t^{\min}$.

3: **for** $t = 1, \ldots, T$

4:     **Prediction using the state model.**

5:     **Sensor placement:** iterative sparsity-enhancing design of $\mathbf{w}_t \in \{0,1\}^N$ (Section 3.3).

6:     **Correction:** estimation of $\mathbf{v}_t$ and $\mathbf{s}_t$ using the measurements from the selected sensing locations.

7:     **Overall KKF estimate:** $\hat{\mathbf{u}}_t = [\hat{\mathbf{v}}_t + \hat{\mathbf{s}}_t]_+$.

8:     **Covariance update.**

9: **end for**

---

4. Simulation Results

In this section, we perform some numerical experiments to exhibit the practicality of the developed sparsity-enforcing sensor placement followed by the KKF estimation method. We select a service area of 6×6 square km with 1 square km spatial resolution as illustrated in Figure 3. The spatial distribution of the non-stationary component at time $t = 0$, i.e., $\mathbf{v}_0$, is generated by the following exponential source-field model

$$[\mathbf{v}_0]_j = \sum_{k=1}^K s_k \exp(-d_k \|\mathbf{x}_j - \boldsymbol{\rho}_k\|_2), \quad j = 1, \ldots, N, \tag{38}$$

where K is the number of field-generating points/sources. The parameters $\boldsymbol{\rho}_k$, s_k, and d_k are the location, amplitude, and the spatial decaying factor of the k-th source at time $t = 0$. Based on this function, we generate the non-stationary component of the field at time $t = 0$, i.e., $\mathbf{v}_0 \in \mathbb{R}^N$ using the parameters $K = 1$, $\boldsymbol{\rho}_1 = [1.5, 1.5]^T$, $s_1 = 2$, and $d_1 = 1$. The spatial distribution of $\mathbf{v}_0$ in the specified service area is shown in Figure 3.

The state model of the non-stationary component $\mathbf{v}_t$ is modelled by (7). The state transition matrix is modelled by the Gaussian kernel function given by (8). For the sake of simplicity, we consider a spatially invariant translation parameter and spatio-temporally invariant dilation parameters given as $\mathbf{a}_t^{ij} = \mathbf{a}_t$ and $\mathbf{D}_t^{ij} = \mathbf{D}$, respectively, for $i, j = 1, \ldots, N$. The elements of the state transition matrix are given by

$$[\mathbf{H}_t]_{ij} = \nu \exp[-(\mathbf{x}_i - \mathbf{x}_j - \mathbf{a}_t)^T \mathbf{D}^{-1}(\mathbf{x}_i - \mathbf{x}_j - \mathbf{a}_t)]. \tag{39}$$

The spatio-temporal evolution of the true value of the field, i.e., $\mathbf{u}_t = \mathbf{v}_t + \mathbf{s}_t$ is generated in the following two ways.
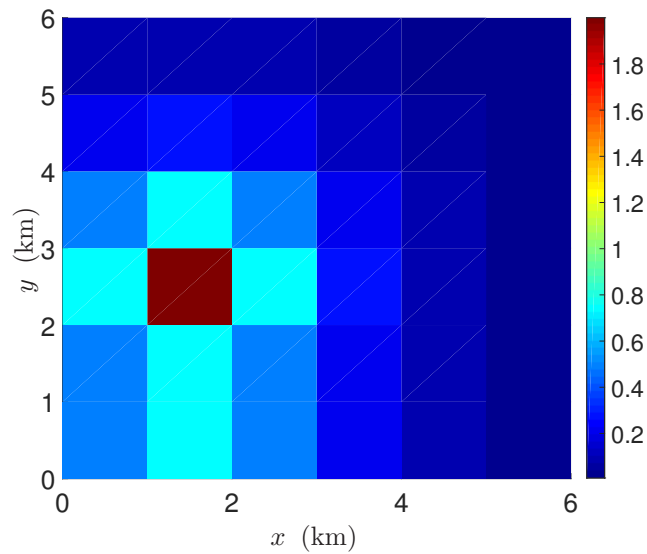
Figure 3. Field distribution at $t = 0$ with a single source: $K = 1$, $\boldsymbol{\rho}_1 = [1.5, 1.5]^T$, $s_1 = 2$, $d_1 = 1$.

In the **first** case, we consider a pure advective process, i.e., we select a very low dilation parameter given by $\mathbf{D} = 10^{-4}\mathbf{I}_2$ for all $t = 1, \ldots, 8$ and $\nu = 0.8$. It is assumed that the temporal resolution is 1 min. The translation vectors, i.e., $\mathbf{a}_t$, are assumed to be changing every t as $[1, 1]^T$, $[-1, -1]^T$, $[1, 1]^T$, $[0, 0]^T$, $[1, 1]^T$, $[-1, -1]^T$, $[0, 1]^T$, and $[-1, -1]^T$. It is assumed that at $t = 0$, $\mathbf{v}_t$ is generated by the source as shown in Figure 3. The different states of $\mathbf{v}_t$ for $t = 1, \ldots, 8$ are generated by the state model of (7). The spatially colored yet temporally uncorrelated process noise is characterized by $\mathbf{q}_t \sim \mathcal{N}(\mathbf{0}_N, \mathbf{Q})$, where $[\mathbf{Q}]_{ij} = 10^{-4}\exp(-\|\mathbf{x}_i - \mathbf{x}_j\|_2)$. The stationary component $\mathbf{s}_t$ is modelled by $\mathbf{s}_t \sim \mathcal{N}(\mathbf{1}_N, \boldsymbol{\Sigma}_s)$. The parameters of the squared exponential covariance function of (6) are given by $\sigma_s^2 = 0.001$ and $\theta = 1$. Note that increasing the value of θ, the field becomes spatially more correlated and the condition number of $\boldsymbol{\Sigma}_s$ increases. However, as mentioned earlier, our proposed formulation, i.e., both the selection and the estimation, does not involve the inversion of $\boldsymbol{\Sigma}_s$. A highly spatially correlated $\mathbf{s}_t$ is considered in the next case. For the first case, the true field $\mathbf{u}_t = \mathbf{v}_t + \mathbf{s}_t$ for $t = 1, \ldots, 8$ can be simulated as shown in Figure 4.
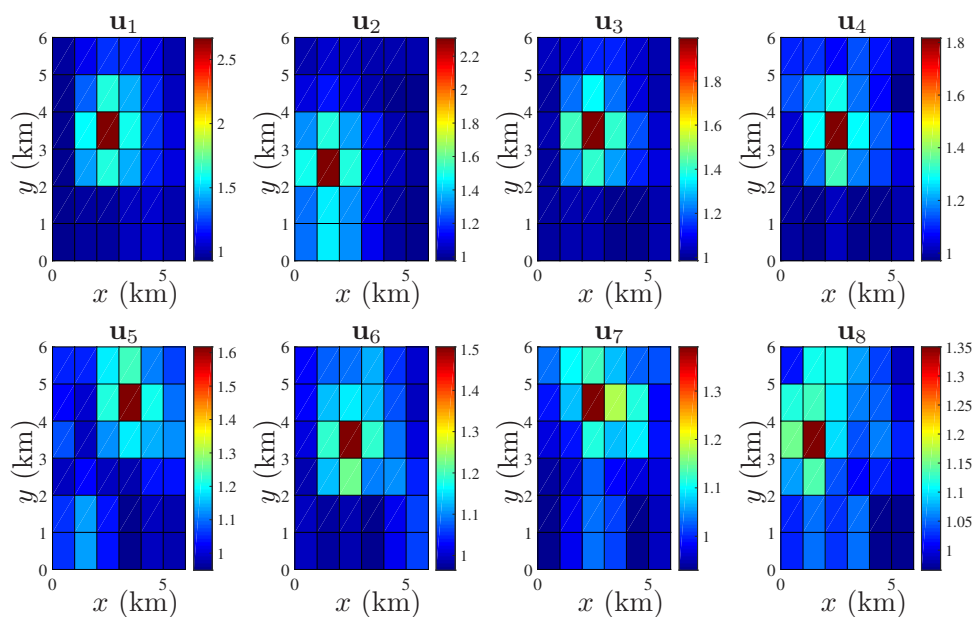
Figure 4. Spatio-temporal evolution of $\mathbf{u}_t$ in a 6×6 square km area; spatial resolution: 1×1 square km; time varying $\mathbf{H}_t$ for $t = 1, \ldots, 8$; strength of spatial correlation: $\theta = 1$.

In the **second** case, we consider $\mathbf{D} = \mathbf{I}_2$ for all $t = 1, \ldots, 8$ and the translation parameters are fixed as $\mathbf{a}_t = [0.4, 0.4]^T$ for $t = 1, \ldots, 4$, and no translation for the last 4 snapshots, i.e., $\mathbf{a}_t = [0, 0]^T$ for $t = 5, \ldots, 8$. The state of $\mathbf{v}_t$ at $t = 0$ is the same as before. The scaling parameter is given by $\nu = 0.35$. The process noise $\mathbf{q}_t$ is the same as before. In this case, we assume that the stationary component $\mathbf{s}_t$ is spatially more correlated than the last case. The parameters of the covariance Function (6) are taken as $\sigma_s^2 = 0.01$ and $\theta = 4$, which generates an ill-conditioned $\mathbf{\Sigma}_s$ (Figure 2a). Using these, the true field, i.e., $\mathbf{u}_t = \mathbf{v}_t + \mathbf{s}_t$ for $t = 1, \ldots, 8$ is shown in Figure 5.
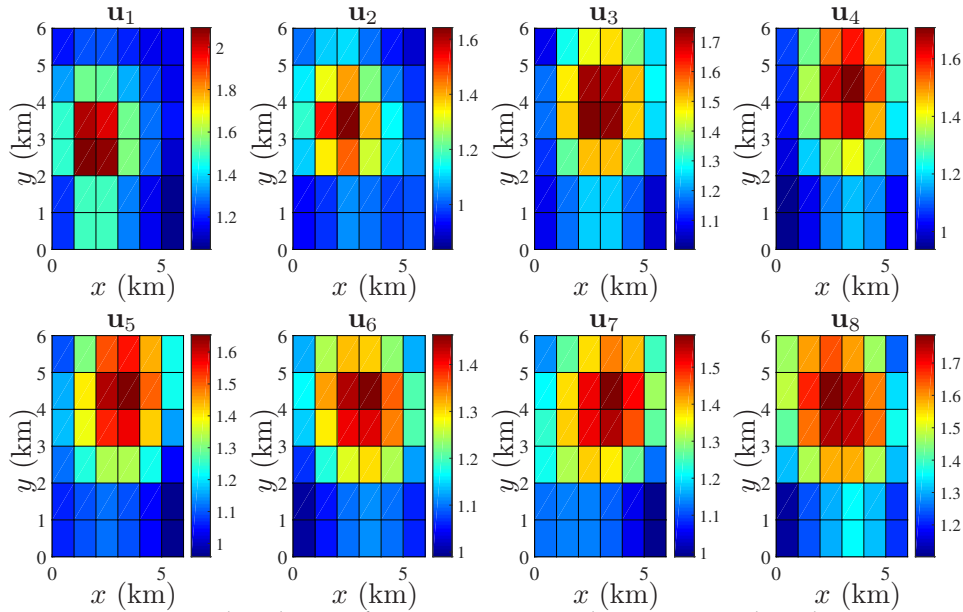
Figure 5. Spatio-temporal evolution of $\mathbf{u}_t$ in a 6×6 square km area; spatial resolution: 1×1 square km; time varying $\mathbf{H}_t$ for $t = 1, \ldots, 8$; strength of spatial correlation: $\theta = 4$.

4.1. Sensor Placement Followed by Field Estimation Using KKF

We select the optimal sensing locations and use them to estimate the field for $t = 1, \ldots, 8$ snapshots for the two different scenarios of the spatio-temporal evolution of the field, as mentioned in the previous section. We use the same service area shown in Figure 1b, where the centroids of the $N = 36$ pixels are the candidate sensing locations. We assume that the measurement noise variance is given by $\sigma_e^2 = 0.001$. We solve the optimization problem of (36) with the parameters $I = 2$ and $\epsilon = 10^{-6}$. The weighting vectors are initialized as $\boldsymbol{\lambda}_0 = \mathbf{1}_N$. The resource constraints are given as $K_t^{\max} = 30$ and $K_t^{\min} = 25$ for all t. To extract the Boolean solution $\mathbf{w}_t \in \{0, 1\}^N$ from $\hat{\mathbf{w}}_t \in [0, 1]^N$, we adopt the randomized rounding method. We use the software CVX [30] (parser CVX, solver SeDuMi [31]) to solve the SDP problem (36).

Following the above simulation setup, the selected sensing locations for the first and the second scenario are shown in Figure 6a,b respectively for the 8 snapshots. The indices of the pixel midpoints are the same as in Figure 1b (vertical axis). The main observations from the selected locations are listed below.

- First of all, it is clearly seen that the selected sensing locations depend on the dynamics. Note that Figure 6a gives the optimal placement pattern, when $\mathbf{H}_t$ is changing every t (different $\mathbf{a}_t$ on every t). Figure 6b shows the optimal sensing locations when we have the same $\mathbf{H}_t$ for $t = 1, ..., 4$ ($\mathbf{a}_t = [0.4, 0.4]^T$) and another $\mathbf{H}_t$ for $t = 5, ..., 8$ ($\mathbf{a}_t = [0, 0]^T$).
- When $\mathbf{H}_t$ is changing every t, i.e., the spatio-temporal evolution of the field is guided by the time-varying spatial translation parameter $\mathbf{a}_t$ (Figure 1b), the optimal selection pattern also depends upon this translation (Figure 6a).

- In the second scenario, we have assumed a very low and fixed translation, i.e., $\mathbf{a}_t = [0.4, 0.4]^T$ for the first 4 snapshots and $\mathbf{a}_t = [0, 0]^T$, i.e., no translation, for the last 4 snapshots (Figure 5). It is seen that almost the same set of sensors are selected in the last 4 snapshots of Figure 6b. In general, when $\mathbf{H}_t$ is not changing with time, the estimation error of the non-stationary component reaches a steady state after a number of snapshots and the same set of sensors are selected every t.

The overall estimation performance using the measurements from the selected locations of Figure 6a,b is shown in Figure 7a,b, respectively. In these figures, we exhibit the pixel-wise comparison of the estimates for $T = 8$ snapshots, i.e., the estimation performance of $36 \times 8 = 288$ pixels. We initialize the KKF iterations with $\hat{\mathbf{v}}_t = \mathbf{1}_N$ and $\mathbf{M}_{\mathbf{v}_t} = 0.001\mathbf{I}_N$ at $t = 0$.
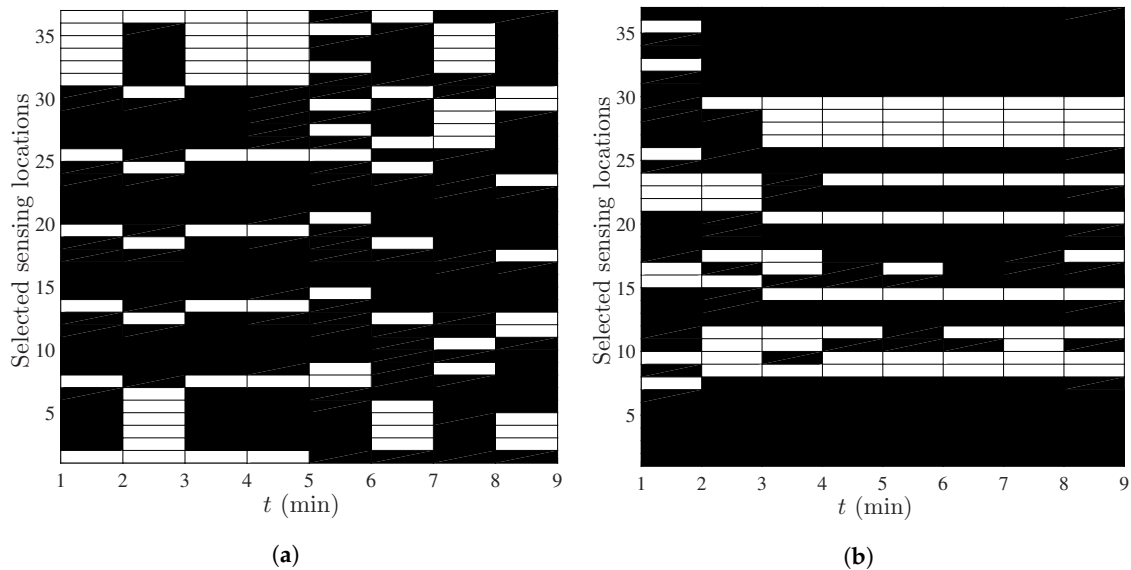

Figure 6. (**a**) Selected sensing locations to estimate the field with the first scenario of the true value; (**b**) Selected sensing locations to estimate the field with the second scenario of the true value.
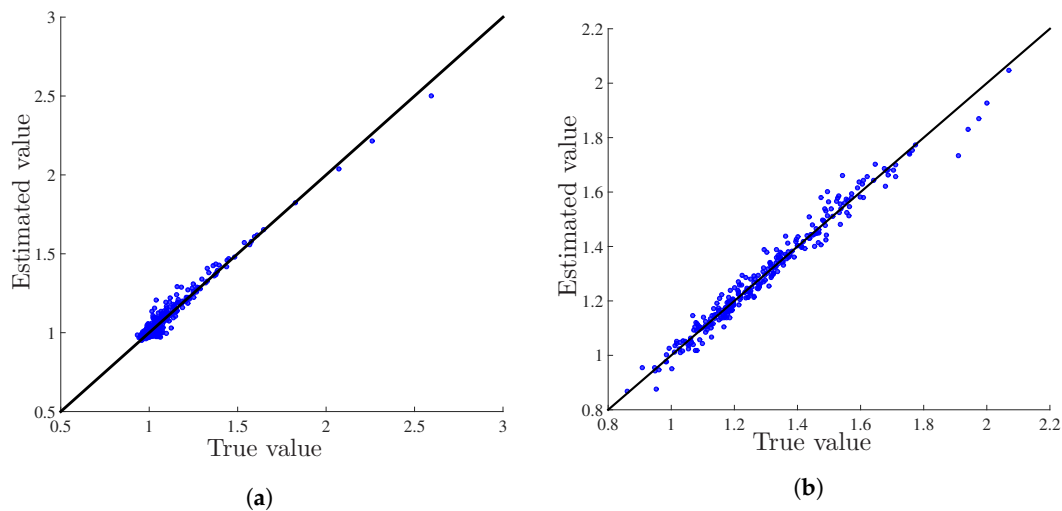
Figure 7. (**a**) Comparison of the KKF estimate ($\hat{\mathbf{u}}_t$) with the true value ($\mathbf{u}_t$) (Figure 4) with the measurements from the selected locations shown in Figure 6a; (**b**) Comparison of the KKF estimate ($\hat{\mathbf{u}}_t$) with the true value ($\mathbf{u}_t$) (Figure 5) with the measurements from the selected locations shown in Figure 6b.

4.2. Performance Analysis

We compare the estimation performance of the developed sensor placement method by comparing the performance metric, i.e., $g(\hat{\mathbf{w}}_t) = \text{tr}[\mathbf{M}_{\mathbf{s}_t}] + \text{tr}[\mathbf{M}_{\mathbf{v}_t}]$ with a random sensor placement (with the same M_t, i.e., $\|\hat{\mathbf{w}}_t\|_0 = M_t$ as for the developed method) and with the best case performance (i.e., $M_t = N$ or $\mathbf{w}_t = \mathbf{1}_N$). For the random placement, we generate 100 different realizations of $\mathbf{w}_t \in \{0,1\}^N$ at every t with the same M_t as for the proposed method. Then $g(\mathbf{w}_t)$ is computed for every $\mathbf{w}_t$ and their average is considered. Similarly, we compute the best case performance, i.e., $g(\mathbf{1}_N)$ for every t and in this case $\mathbf{M}_{\mathbf{v}_t}$ is updated with $\mathbf{w}_t = \mathbf{1}_N$. We use the same set of parameters as mentioned in the first case of Section 4.1. Only the resource allocation constraint is simplified as $\mathbf{1}^T \mathbf{w}_t = 15$, i.e., we fix that only 15 sensing locations will be selected every t. The performance comparison is shown in Figure 8. It is seen that the proposed approach slightly outperforms the random placement. However, the random placement of sensors does not optimize any performance criterion.
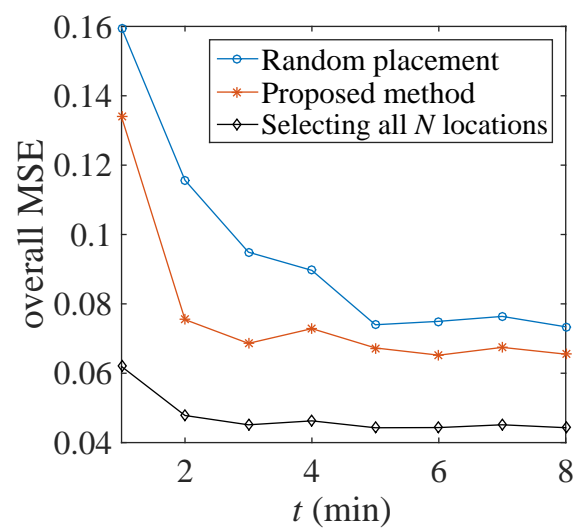

Figure 8. Comparison of the performance metric for the proposed method, random placement and the best case (only 15 sensors are selected on every t).

4.3. Spatial Sensor Placement for Stationary Field Estimation

In this section, we show the effects of different spatial correlation patterns on sensor placement assuming the field is purely stationary. We solve the optimization problem of (35), for two different spatial covariance matrices ($\mathbf{\Sigma}_s$). In the first case, we consider that $\mathbf{\Sigma}_s$ is generated by the squared exponential Function (6) with $\theta = 2$ and $\sigma_s^2 = 0.01$ (Figure 9a). In the second case, we consider a randomly generated $\mathbf{\Sigma}_s$ (Figure 9b). The resource allocation constraint is the same as before, i.e., $K_t^{\min} = 25$, and $K_t^{\max} = 30$. We solve the optimization problem of (35), with the iterative approach of (36) with the same parameters as mentioned in the previous section. The selected locations (marked by black squares where the blue dots are the candidate locations as shown in Figure 1b) to deploy sensors are shown in Figure 10a,b for the spatial covariance matrices shown in Figure 9a,b, respectively.

First of all, it is observed that the spatial distribution of the optimal sensing locations depends upon the correlation pattern of the field. It is seen that when $\mathbf{\Sigma}_s$ is generated by a squared exponential covariance (stationary) function then the optimal sensor placement pattern is more or less symmetrically and uniformly distributed over the entire service area. However, for a random spatial covariance matrix the optimal sensing locations do not follow any specific pattern.

(a)

(b)
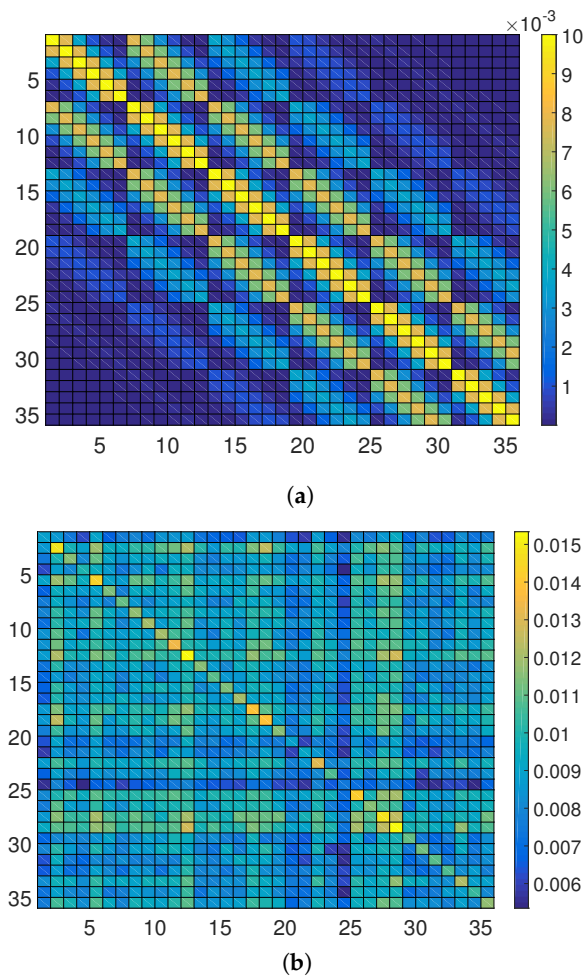
Figure 9. (a) Spatial covariance matrix generated by the squared exponential function ($\sigma_s^2 = 0.01$, $\theta = 2$); (b) Randomly generated spatial covariance matrix.

(a)

(b)
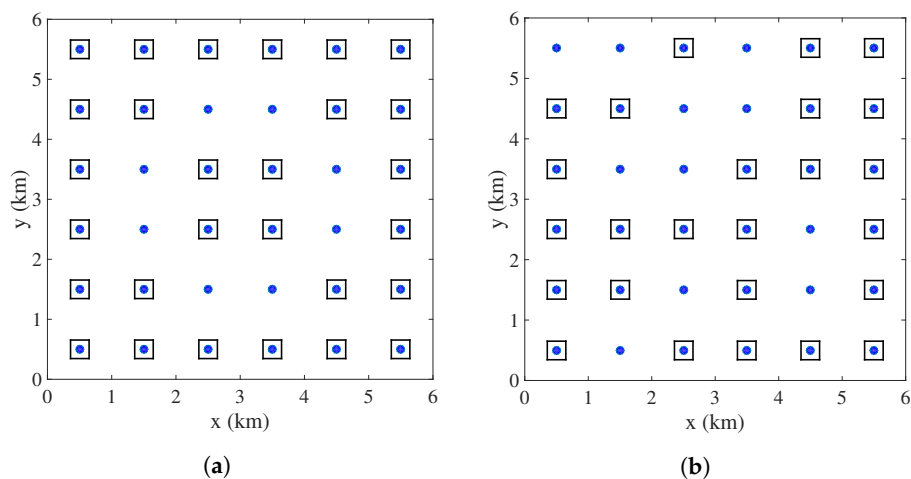
Figure 10. (a) Sensor placement pattern for the $\mathbf{\Sigma}_s$ as shown in Figure 9a; (b) Sensor placement pattern for the $\mathbf{\Sigma}_s$ as shown in Figure 9b.

5. Conclusions and Future Work

In this work, we have developed a sparsity-enforcing sensor placement followed by a field estimation technique using a KKF. The proposed methodology selects the most informative sensing

locations over space and time in a specified service area of interest. Along with minimizing the estimation error, the developed method also economizes the sensor placement (in terms of resources) at every temporal interval. The salient features of the proposed method include handling a general class of spatial covariance matrices and tackling correlated measurement noise. Numerical analysis shows the feasibility of the method. The effects of the dynamics and spatial correlation of the field in spatio-temporal sensor placement are discussed with numerical experiments.

In this work, we have considered the fact that the prior knowledge regarding the spatial variability and the dynamics are perfectly known a priori. In that case, the performance of a clairvoyant Kalman setup with Gaussian measurement and process noise is optimal. However, in many practical scenarios, the aforementioned spatio-temporal prior information may not be accurate and we require more information regarding the unknown field in the estimation step. Future research is envisioned to incorporate the effects of model imperfections in the developed method. Another future research area could be using distributed algorithms to apply the developed method for large scale sensor network applications. It will also be interesting to tailor the recent progress in time-varying optimization [32] to solve the SDPs in a tracking fashion, rather than at optimality at each sampling time.

References

1. Oliveira, L.; Rodrigues, J. Wireless sensor networks: A survey on environmental monitoring. *J. Commun.* **2011**, *6*, 143–151. [CrossRef]
2. Hart, J.K.; Martinez, K. Environmental sensor networks: A revolution in the earth system science? *Earth-Sci. Rev.* **2006**, *78*, 177–191. [CrossRef]
3. Cressie, N.; Wikle, K. *Statistics for Spatio-Temporal Data*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
4. Mardia, K.V.; Goodall, C.; Redfern, E.J.; Alonso, F.J. The kriged Kalman filter. *Test* **1998**, *7*, 217–282. [CrossRef]
5. Wikle, C.K.; Cressie, N. A dimension-reduced approach to space-time Kalman filtering. *Biometrika* **1999**, *86*, 815–829. [CrossRef]
6. Kim, S.J.; Anese, E.D.; Giannakis, G.B. Cooperative spectrum sensing for cognitive radios using kriged Kalman filtering. *IEEE J. Sel. Top. Signal Process.* **2011**, *5*, 24–36. [CrossRef]
7. Rajawat, K.; Dall'Anese, E.; Giannakis, G. Dynamic network delay cartography. *IEEE Trans. Inf. Theory* **2014**, *60*, 2910–2920. [CrossRef]
8. Sahu, S.K.; Mardia, K.V. A Bayesian kriged Kalman model for short term forecasting of air pollution levels. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **2005**, *54*, 223–244. [CrossRef]
9. Cortes, J. Distributed Kriged Kalman filter for spatial estimation. *IEEE Trans. Autom. Control* **2009**, *54*, 2816–2827. [CrossRef]
10. Krause, A.; Singh, A.; Guestrin, C. Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. *J. Mach. Learn. Res.* **2008**, *9*, 235–284.
11. Ranieri, J.; Chebira, A.; Vetterli, M. Near-optimal sensor placement for linear inverse problems. *IEEE Trans. Signal Process.* **2014**, *62*, 1135–1146. [CrossRef]
12. Carmi, A. Sensor scheduling via compressed sensing. In Proceedings of the 13th Conference on Information Fusion (FUSION), Edinburgh, UK, 26–29 July 2010.
13. Fu, Y.; Ling, Q.; Tian, Z. Distributed sensor allocation for multi-target tracking in wireless sensor networks. *IEEE Trans. Aerosp. Electron. Syst.* **2012**, *48*, 3538–3553. [CrossRef]
14. Patan, M. *Optimal Sensor Networks Scheduling in Identification of Distributed Parameter Systems*; Springer Science & Business Media: Berlin, Germay, 2012; Volume 425.
15. Liu, S.; Fardad, M.; Masazade, E.; Varshney, P.K. Optimal periodic sensor scheduling in networks of dynamical systems. *IEEE Trans. Signal Process.* **2014**, *62*, 3055–3068. [CrossRef]

16. Mo, Y.; Ambrosino, R.; Sinopoli, B. Sensor selection strategies for state estimation in energy constrained wireless sensor networks. *Automatica* **2011**, *47*, 1330–1338. [CrossRef]

17. Akyildiz, I.F.; Vuran, M.C.; Akan, O.B. On exploiting spatial and temporal correlation in wireless sensor networks. In Proceedings of the WiOpt, Cambridge, UK, March 2004; Volume 4, pp. 71–80.

18. Joshi, S.; Boyd, S. Sensor selection via convex optimization. *IEEE Trans. Signal Process.* **2009**, *57*, 451–462. [CrossRef]

19. Jamali-Rad, H.; Simonetto, A.; Ma, X.; Leus, G. Distributed sparsity-aware sensor selection. *IEEE Trans. Signal Process.* **2015**, *63*, 5951–5964. [CrossRef]

20. Chepuri, S.P.; Leus, G. Sparsity-promoting sensor selection for non-linear measurement models. *IEEE Trans. Signal Process.* **2015**, *63*, 684–698. [CrossRef]

21. Liu, S.; Cao, N.; Varshney, P.K. Sensor placement for field estimation via poisson disk sampling. In Proceedings of the 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Washington, DC, USA, 7–9 December 2016; pp. 520–524.

22. Roy, V.; Simonetto, A.; Leus, G. Spatio-temporal sensor management for environmental field estimation. *Signal Process.* **2016**, *128*, 369–381. [CrossRef]

23. Jindal, A.; Psounis, K. Modeling spatially correlated data in sensor networks. *ACM Trans. Sens. Netw.* **2006**, *2*, 466–499. [CrossRef]

24. Liu, S.; Chepuri, S.P.; Fardad, M.; Maşazade, E.; Leus, G.; Varshney, P.K. Sensor selection for estimation with correlated measurement noise. *IEEE Trans. Signal Process.* **2016**, *64*, 3509–3522. [CrossRef]

25. Ababou, R.; Bagtzoglou, A.C.; Wood, E.F. On the condition number of covariance matrices in kriging, estimation, and simulation of random fields. *Math. Geol.* **1994**, *26*, 99–133. [CrossRef]

26. Sigrist, F.; Künsch, H.R.; Stahel, W.A. A dynamic nonstationary spatio-temporal model for short term prediction of precipitation. *Ann. Appl. Stat.* **2012**, *6*, 1452–1477 [CrossRef]

27. Kay, S.M. *Fundamentals of Statistical Signal Processing: Estimation Theory*; PTR Prentice-Hall: Englewood Cliffs, NJ, USA, 1993.

28. Boyd, S.; Vandenberghe, S. *Convex Optimization*; Cambridge University Press: Cambridge, UK, 2009.

29. Candes, E.; Wakin, M.; Boyd, S. Enhancing sparsity by reweighted ℓ_1 minimization. *J. Fourier Anal. Appl.* **2008**, *14*, 877–905. [CrossRef]

30. Grant, M.; Boyd, S.; Ye, Y. *CVX, Matlab Software for Disciplined Convex Programming*; CVX Research, Inc.: Austin, TX, USA, 2008.

31. Sturm, J.F. Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones. *Optim. Methods Softw.* **1999**, *11*, 625–653. [CrossRef]

32. Simonetto, A.; Dall'Anese, E. Prediction-correction algorithms for time-varying constrained optimization. *IEEE Trans. Signal Process.* **2017**, *65*, 5481–5494. [CrossRef]