Computers and Electronics in Agriculture 214 (2023) 108282

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Agent-based sensor location strategy for smart irrigation of large crop fields

Jorge Lopez-Jimenez [a,b,*], Nicanor Quijano [a], Alain Vande Wouwer [b]

[a] Department of Electrical and Electronic Engineering, Universidad de los Andes, Bogotá, 111711, Colombia
[b] Systems, Estimation, Control and Optimization (SECO), University of Mons, Mons, 7000, Belgium

ARTICLE INFO

Keywords:
Sensor placement
State estimation
Distributed observer
Genetic algorithm
Smart farming

ABSTRACT

Efficient monitoring of large crop fields is important to ensure the optimal use of resources such as water in irrigation policies, but at the same time represents a challenge to determine the structure of the sensor network. A balance must be accomplished between the acquisition, operation, and maintenance costs of this sensor network with the amount of information that can be collected in real-time to support the optimal use of the resource, e.g., an optimal irrigation policy. In this study, a sensor location strategy is proposed based on an agent-based model (ABM) of the crop–soil system, a state estimation algorithm reconstructing non-measured variables, and an objective function balancing the convergence of the estimation technique and the costs of the sensor network. The ABM model describes the crop–soil dynamics and allows conveniently representing uneven landscapes where water exchanges take place between different portions of the land. Various state estimation techniques can be considered and an extended Kalman filter is implemented in the present study, whose error covariance matrix can be exploited to assess practical observability and observer convergence. Finally, an economic cost function combines the observability measure with the sensor costs in order to select an optimal or suboptimal sensor array. For validation purposes, a numerical simulation case study, corresponding to a rugged land located in Colombia, is used to test various scenarios including the variability of climatic inputs.

1. Introduction

The growth of the world population and the increasing demand for food production has put the focus on the improvement of agricultural processes. The problem of food production has two facets. On the one hand, farming procedures must involve proper knowledge of the soil and crop in the context of climate change. On the other hand, the overuse of limited resources such as water must be drastically reduced, considering that agriculture uses about 70% of freshwater worldwide.

Precision agriculture (PA) includes a set of tools and techniques to deal with the above-mentioned issues while foreseeing efficiency and sustainability. One of the open challenges in PA is the efficient management of resources such as water, fertilizer, or herbicides and the optimal deployment and use of sensors (Cobbenhagen et al., 2021). In Visalini et al. (2019), for instance, the authors propose an algorithm for sensor placement where information from low-resolution remote sensors are fused with proximal sensors connected to a wireless sensor network (WSN). However, the sensor cost and range, which are affected by weather and farming operation outages, are recurrent issues. Moreover, when a large set of sensors are deployed in a WSN to monitor a field, the costs associated with installation and operation do not decrease with the increasing number of sensors, independently of the technology and topology of the network (Thakur et al., 2019). Hence, the design of systems for monitoring large crops is conditioned by economic constraints and technical limitations. The economic aspect of sensor location to build efficient crop monitoring systems has been tackled by combining multiple data sources and statistical tools (López-Lozano and Baruth, 2019).

Besides the economic limitations, the prediction of the crop–soil system is critical in water management since the hydrological dynamics are significantly influenced by the landscape heterogeneity (Gao et al., 2018) and climate variability (Pelak et al., 2017). One of the possible approaches to address land variations is agent-based modeling (ABM). This concept has been successfully used to represent uneven landscapes, where agents correspond to homogeneous parts of soil (Lopez-Jimenez et al., 2021). The connections between agents are represented by a directed graph, where water exchanges occur between nodes. This framework reduces the complexity of solving the crop-environment model under climate variability compared to a 2- or 3-dimensional partial differential equation model.

The general approach to sensor location has been defined as an optimization problem where a sensor configuration achieves the minimum capital cost while meeting specific performance criteria. In Chmielewski

* Corresponding author at: Department of Electrical and Electronic Engineering, Universidad de los Andes, Bogotá, 111711, Colombia.
E-mail addresses: jorgelopez@uniandes.edu.co (J. Lopez-Jimenez), nquijano@uniandes.edu.co (N. Quijano), alain.vandewouwer@umons.ac.be (A. Vande Wouwer).

https://doi.org/10.1016/j.compag.2023.108282
Received 26 January 2023; Received in revised form 18 August 2023; Accepted 25 September 2023
Available online 14 October 2023

et al. (2002), the authors provide a theoretical framework to prove the feasibility of a distributed observer for nonlinear systems with precision and detectability constraints. Particularly for the case of nonlinear systems, observability analysis is a challenge, and indirect assessment methods are proposed. In Rodriguez et al. (2021), an empirical observability approach is presented based on the Gramian of measured variables. The estimation error is evaluated as the trace of the error covariance matrix. However, the underlying computation becomes challenging when the number of variables increases (Bwambale et al., 2022). To address this issue, several algorithmic solutions have been proposed. For instance, multi-rate discretization, reduction of model order (Paul et al., 2016), and the use of genetic algorithms (GA) to minimize the cost function (Abioye et al., 2020). Moreover, the use of GA in the design of irrigation policies has been reported in Perea et al. (2019, 2021).

In this study, we focus attention on sensor placement in large crop fields with land heterogeneity, climatic variability, and limited budget. A distributed estimator based on an ABM model and an extended Kalman filter (EKF) is implemented to reconstruct the non-measured state variables and to formulate an objective function balancing the EKF performance with the costs of the sensor network. This objective function is optimized using a genetic algorithm.

More specifically, the main contributions of this work are:

- the design of an agent-based EKF to estimate the state variables of a heterogeneous crop–soil system;
- the formulation of a sensor placement strategy based on the covariance matrix of the EKF and various terms accounting for the sensor costs, exploiting a genetic algorithm to deal with the problem dimensionality; and
- the application of the methodology to a realistic case study, corresponding to a crop–soil system located in the province of Samacá, in the department of Boyacá, Colombia, where the agricultural activity is carried out on rough terrain, and water has a high cost, highlighting the potential benefits of monitoring and cost savings.

This paper is organized as follows. In Section 2, the underlying methods are described, i.e., agent-based modeling, EKF estimation, and genetic algorithm. Section 3 deals with the design of the distributed state estimator. In Section 4, a case study is presented as a testbed and numerical results are discussed. Conclusions and future directions are drawn in the final section.

2. Methods

2.1. Agent-based modeling

Agent-based modeling (ABM) is a methodology used to simulate interactions between autonomous individuals. An agent is an entity located in a specific environment and capable of autonomous actions to meet some objectives (Siegfried, 2014). The complete definition includes identifiability, a set of non-uniform attributes, asynchronous interactions, and uncertainty related to parameters or states. Moreover, an agent has at least two levels of interaction: an upper layer to interact with other agents and a lower layer to solve internal processes. The first layer is concerned with interactions with other agents, and the second one is concerned with the environment. Global behavior emerges from the combination of individuals.

The ABM shown in this work is intended to predict the evolution of crop water content in soil and biomass by using local information from a limited set of sensors. Especially, the model considers environmental inputs from a meteorological station (either portable or local) located in close proximity to the cultivation area. Two types of agents are considered, e.g., crop–soil agents and irrigation agents. The first ones correspond to a portion of homogeneous soil with a surface of regular shape, while the second ones are designed for irrigation management. This model is particularly well suited to represent crop landscapes with rugged topography where the water exchanges give the interactions between crop–soil agents. The conceptual representation of the discretized land surface is shown in Fig. 1, where a sample agent is highlighted with the instrumentation, estimation, and control loop. The internal structure of crop–soil agents is based on a previous work focused on the description of the crop–soil dynamics (Lopez-Jimenez et al., 2021). The list of variables is summarized in Table 1.

Table 1
Notation.

N	Number of agents
$n = 1, \dots, N$	Index of agents
k	Time index
j	Index of environmental inputs
n_s	Number of states per agent
n_m	Number of sensed patches
n_e	Number of environmental inputs
$N_s = n_s N$	Total number of states
$N_m = n_s n_m$	Total number of measurements
x_1	Water content in soil
x_2	Cumulative temperature
x_3	Cumulative temperature until maturity to reach 50% radiation interception
x_4	Biomass
u	Irrigation
$u_e^{(1)}$	Rainfall
$u_e^{(2)}$	Reference evapotranspiration
$u_e^{(4)}$	Solar radiation
$\mathbf{x}$	Enhanced state vector
$\mathbf{F}$	Linear system matrix
$\mathbf{B}$	Input matrix
$\mathbf{C}$	Measurement matrix
$\mathbf{H}$	Matrix of direct effect of environmental inputs
$\mathbf{Q}$	Covariance matrix of system
$\mathbf{R}$	Covariance matrix of measurements
w_k	Uncertainties in the process
v_k	Uncertainty in measurements
λ	Weighting factor of cost function
$\gamma^{(n)}$	Normalized elevation

The $n = 1, \dots, N$, soil–crop agents are described by a set of n_s state variables $x_i^{(n)}, \quad i = 1, \dots, n_s$, which are influenced by the n_e climatic variables $u_e^{(j)}, \quad j = 1, \dots, n_e$, and by the irrigation inputs $u^{(n)}$. Part of the state variables can be measured by sensors with output $z_i^{(n)}$ under the influence of measurement errors (or noise) $v_i^{(n)}$. The irrigation agent acts as a controller to determine the water delivered to the crop–soil agents, and r represents the amount of water that is available. As instrumentation is limited, an estimator is required to provide the state estimates $\hat{x}_i^{(n)}$. A description of the variables related to the main blocks of Fig. 1 is presented next, where the estimator block is described in Section 2.2.

2.1.1. Crop–soil agents

The dynamic behavior of the crop–soil agents is related to their internal mechanistic functions, and the coupling with other agents. The dynamical evolution is described by

$$\dot{\mathbf{x}}^{(n)} = f_n(\mathbf{x}, \mathbf{u_e}, \mathbf{u_c}, \Theta^{(n)}); \qquad n = 1, \dots, N, \tag{1}$$

where $f_n(\cdot)$ is a nonlinear function, $\mathbf{x}$ corresponds to the vector of state variables (of all the agents), $\mathbf{u_e}$ is the vector of environmental inputs, $\mathbf{u_c}$ is the vector of management inputs (i.e., irrigation), and $\Theta^{(n)}$ is the set of parameters related to the nth agent, including soil, crop, and management parameters.

In practice, the control actions will be determined based on information that is received at regular time instants k, and a discrete-time formulation of the model is more appropriate. In order to discretize the model, we use the usual sampling interval (1 day). This is motivated by two reasons: (i) the growing process at the crop scale is assumed to be affected only by the time course of temperature, which enables the use of thermal time (degree days above a base temperature) to compute the

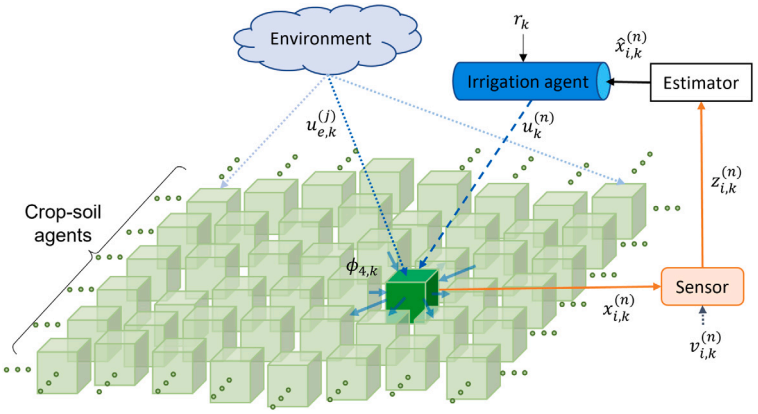

Fig. 1. Conceptual diagram of the ABM with $n = 1, \ldots, N$, crop–soil agents, i indexing the state variables of each agent, and j indexing the environmental inputs. A sample agent is highlighted to relate the variables and signals of the control loop. The subindex k corresponds to the time evolution index.

daily aging of the crop; and (ii) the environmental inputs are available daily. In a more detailed form, the discrete-time equations of the nth agent can be summarized as follows:

$$x_{1,k+1}^{(n)} = x_{1,k}^{(n)} - \phi_{1,k}^{(n)} - \phi_{2,k}^{(n)} - \phi_{3,k}^{(n)} + \phi_{4,k}^{(n)} + u_{e,k}^{(1)} + u_{c,k}^{(n)} \tag{2}$$

$$x_{2,k+1}^{(n)} = x_{2,k}^{(n)} + h_{1,k}^{(n)} \tag{3}$$

$$x_{3,k+1}^{(n)} = x_{3,k}^{(n)} + \theta_{11}^{(n)}(1 - h_{2,k}^{(n)}) + \theta_{12}^{(n)}(1 - h_{3,k}^{(n)}) \tag{4}$$

$$x_{4,k+1}^{(n)} = x_{4,k}^{(n)} + \theta_{13}^{(n)} h_{6,k}^{(n)} h_{7,k}^{(n)} h_{8,k}^{(n)} g_k^{(n)} u_{e,k}^{(4)}, \tag{5}$$

where the state variables are the water content of every land patch ($x_1^{(n)}$), the cumulative temperature ($x_2^{(n)}$), the cumulative temperature until maturity to reach 50% radiation interception due to leaf senescence ($x_3^{(n)}$), and the biomass ($x_4^{(n)}$). The fluxes related to the water balance in soil are the crop transpiration ($\phi_1^{(n)}$), the surface runoff ($\phi_2^{(n)}$), the deep drainage ($\phi_3^{(n)}$), and the flux coming from the neighbors located in any highest part of the terrain with respect to the agent ($\phi_4^{(n)}$). The independent environmental inputs include the precipitations ($u_e^{(1)}$) and the solar radiation ($u_e^{(4)}$). The management input is the amount of water irrigated in every patch ($u_c^{(n)}$). The functions $h^{(n)}$ are continuous or piecewise continuous functions of environmental inputs ($u_e^{(j)}$) used to compute the stress factors, and $g^{(n)}$ is the growth function. Parameters $\theta_{11}^{(n)}$, $\theta_{12}^{(n)}$, and $\theta_{13}^{(n)}$ represent the maximum daily reduction in $x_3^{(n)}$ due to heat stress, the maximum daily reduction in $x_3^{(n)}$ due to drought stress, and the radiation use efficiency, respectively. The full set of parameters $\Theta^{(n)}$ includes soil parameters (i.e., θ_1 to θ_6), thermal crop parameters (i.e., θ_7 to θ_{16}), a parameter to account for the influence of CO_2 on the radiation use efficiency (i.e., θ_{17}), and management parameters (i.e., θ_{18} to θ_{20}). In particular, the parameters related to crops can be retrieved from literature sources, the ones related to soil can be measured in situ, and for those related to crop management, the choice should follow conservative cropping practices. A complete list of functions, variables, and parameters is described in Lopez-Jimenez et al. (2021).

The crop transpiration $\phi_{1,k}$ is given by

$$\phi_{1,k} = \min(\theta_1(x_{1,k} - \theta_2\theta_5), u_{e,k}^{(2)}), \tag{6}$$

where θ_1 is the water uptake coefficient, θ_2 is the wilting point, θ_5 is the root-zone depth, and $u_{e,k}^{(2)}$ is the reference evapotranspiration.

The drought stress $h_{3,k}$ can be expressed as

$$h_{3,k} = 1 - \theta_{14} h_{4,k}, \tag{7}$$

where θ_{14} is the sensitivity factor to radiation-use efficiency, and

$$h_{4,k} = \begin{cases} 1 - \frac{\phi_{1,k}}{w_{2,k}}, & \phi_{1,k} < u_{e,k}^{(2)} \\ 0, & \phi_{1,k} \geq u_{e,k}^{(2)}. \end{cases}$$

The incoming flux $\phi_{4,k}$ is computed as the sum of the outflows $\phi_{out,k}$ from all the neighbors, which have a normalized elevation γ higher than the considered agent. The outflow of such an agent is given by

$$\phi_{out,k} = \begin{cases} \frac{(x_{1,k} - \theta_6\theta_5) + \phi_{2,k}}{N_r}, & x_{1,k} > \theta_6\theta_5 \\ 0, & x_{1,k} \leq \theta_6\theta_5, \end{cases} \tag{8}$$

where N_r is the number of receiving neighbors, and θ_6 is the field capacity. Finally, the growth function g_k^n is given by

$$g_k^n = \begin{cases} \frac{\theta_{19}}{1 + e^{-0.01(x_{2,k} - \theta_{20})}}, & x_{2,k} \leq \frac{\theta_{18}}{2} \\ \frac{\theta_{19}}{1 + e^{0.01(x_{2,k} + x_{3,k} - \theta_{18})}}, & x_{2,k} > \frac{\theta_{18}}{2}, \end{cases} \tag{9}$$

where θ_{18} is the cumulative temperature requirement from sowing to maturity, θ_{19} is the maximum fraction of radiation interception that a crop can reach, and θ_{20} is the cumulative temperature requirement for leaf area development to intercept 50% of radiation.

2.1.2. Irrigation agent

The objective of this agent is to deliver a water quota to every crop–soil agent on specific days during the crop season. This agent is intended to follow an optimal policy. However, as the focus of this work is on the estimator design, the irrigation agent is only considered to close the control loop. It can be turned off ($u_c^{(n)} = 0, \forall n \in N$) or provide a fixed quota ($u_c^{(n)} = c$, where c is a constant), following a traditional management policy.

2.1.3. Sensor configuration

The sensing technology considered in this work could be either direct or remote. The direct methods include ground sensors, and remote sensing (RS) is mostly based on images. Images have the advantage of capturing the whole cropping area and can be calibrated with punctual measurements of climatic variables either by a meteorological station in situ or by ground sensors (Wu et al., 2022). However, the daily availability of images is not guaranteed; instead, ground sensors are continuously providing data, but the trade-off is that measurements are restricted to the location of the sensors (Visalini et al., 2019). For every agent, $n = 1 \ldots, N$, it is assumed that n_s state variables can be measured or estimated based on local information. Therefore, in the distributed observer context, when an agent is selected as the location for sensing, all state variables are available at every time k either by sensor fusion or direct RS. The sensor equation for every agent is

$$\mathbf{z}_k^{(n)} = \mathbf{C}^{(n)}\mathbf{x}_k^{(n)} + \mathbf{H}^{(n)}\mathbf{u}_{e,k}^{(n)} + \mathbf{v}_k^{(n)}, \tag{10}$$

where $\mathbf{z}_k^{(n)}$ is the vector of sensor outputs, $\mathbf{C}^{(n)}$ is the measurement matrix, $\mathbf{H}^{(n)}$ is the matrix associated with the direct effect of environmental

inputs, such as rainfall and air temperature, on outputs such as water in the soil, and $\mathbf{v}_k^{(n)}$ is a vector that represents the measurements errors. In this work, the measurement noise is assumed to follow a normal distribution with a covariance matrix $\mathbf{R}_k^{(n)}$. All measurements are assumed uncorrelated. In practice, the measurement of water in soil $x_{1,k}^{(n)}$ is directly influenced by rainfall, the measurement of $x_{4,k}^{(n)}$, i.e., aboveground biomass, is directly influenced by solar radiation, while the two other states, i.e., cumulative temperature $x_{2,k}^{(n)}$ and cumulative temperature until maturity $x_{3,k}^{(n)}$, are influenced by the air temperature. For the latter, environmental inputs have an indirect effect since they change the stress factors.

Since the meteorological conditions have a significant impact on the soil–crop behavior (Rossello et al., 2019), $\mathbf{u}_{e,k}$ is assumed to be collected daily for the full area of the field.

2.2. State estimation approach

As mentioned before, the agents are sharing information through the water exchanges. These exchanges are driven by the location of every agent concerning their neighbors accordingly to the topography of the terrain and the soil properties. A directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is proposed to track the information inside the model. There, $\mathcal{V}$ represents a set of N nodes (one per every agent) and $\mathcal{E}$ is given by the incoming flux from neighbors $\phi_{out,k}^{(n)}$. Notice that $\phi_4^{(n)} = \sum \phi_{out,k}^{(n)}$ in Eq. (2).

To estimate the states of all agents using a limited number of sensing locations, a custom version of the extended Kalman filter (EKF) is proposed. This version is adapted from the Kalman filter algorithm presented in Jiang et al. (2017). The discretization and time update are assumed to be one day for practical purposes. Considering the representation of the entire field given by an assembled nonlinear model as

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{u}_{\mathbf{e},k}, \mathbf{u}_{c,k}, \Theta), \tag{11}$$

where $\mathbf{x} = [x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, x_4^{(1)}, \dots, x_1^{(N)}, x_2^{(N)}, x_3^{(N)}, x_4^{(N)}]^\top$, Eq. (11) can be linearized along the estimated trajectory, and to this end, the Jacobian matrix is obtained by partial differentiation, i.e.,

$$\mathbf{F} = \begin{bmatrix} F^{(1)} & \cdots & [0] & \dots & [0] \\ \vdots & \ddots & \vdots & & \vdots \\ [0] & \cdots & F^{(n)} & \cdots & [0] \\ \vdots & & \vdots & \ddots & \vdots \\ [0] & \cdots & [0] & \cdots & F^{(N)} \end{bmatrix}.$$

where $F^{(n)}$ represents the Jacobian matrix per crop–soil agent.

Including the measurement equation, the field system becomes

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k, \tag{12}$$

$$\mathbf{z}_k = \mathbf{C}\mathbf{x}_k + \mathbf{H}\mathbf{u}_k + \mathbf{v}_k, \tag{13}$$

where $\mathbf{x} \in \mathbb{R}^{N_s}$ is the global vector of states, $\mathbf{u} \in \mathbb{R}^{n_e}$ represents the vector of manipulated and environmental inputs, $\mathbf{z} \in \mathbb{R}^{N_m}$ is the vector of measured variables, $\mathbf{F} \in \mathbb{R}^{N_s \times N_s}$ and $\mathbf{B} \in \mathbb{R}^{N_s \times n_e}$ are system matrices, $\mathbf{C} \in \mathbb{R}^{N_m \times N_s}$ is the measurement matrix, and $\mathbf{H} \in \mathbb{R}^{N_s \times n_e}$ is the matrix of direct effect of environmental inputs on outputs. $\mathbf{w}_k$ represents perturbations and unmodeled processes dynamics and is assumed to be Gaussian with zero mean and covariance matrix $\mathbf{Q}_p \in \mathbb{R}^{N_s \times N_s}$. Measurements are susceptible to noise $\mathbf{v}_k$, which is also assumed to be Gaussian with zero mean and covariance matrix $\mathbf{R} \in \mathbb{R}^{N_m \times N_m}$. $\mathbf{w}_k$ and $\mathbf{v}_k$ are assumed to be uncorrelated. Notice that $\mathbf{B}\mathbf{u}_k = \mathbf{B}_c\mathbf{u}_{c,k} + \mathbf{B}_e\mathbf{u}_{e,k}$ encompasses the management and environmental inputs.

The recursive algorithm of the filter begins with the initialization of the state $\hat{\mathbf{x}}_0$, and the error covariance matrix $\mathbf{P}_0$ given by

$$E[\mathbf{x}_0] = \hat{\mathbf{x}}_0; \quad \mathbf{P}_0 = E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^\top].$$

The algorithm starts with a prediction (or time update) step:

$$\hat{\mathbf{x}}_k^- = f(\hat{\mathbf{x}}_{k-1}, \mathbf{w}_k, \mathbf{u}_k^{(n)}, \Theta^{(n)}) \tag{14}$$

$$\mathbf{P}_k^- = \mathbf{F}_{k-1}\mathbf{P}\mathbf{F}_{k-1}^\top + \mathbf{Q}_{k-1}, \tag{15}$$

where the super index $^-$ denotes the prior estimates and $\mathbf{Q}$ is the adjusted system covariance matrix as described later. Then, a correction (or measurement update) step is applied using,

$$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{C}_k^\top[\mathbf{C}_k\mathbf{P}_k^-\mathbf{C}_k^\top + \mathbf{R}_k]^{-1}, \tag{16}$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k[\mathbf{z}_k - \mathbf{C}_k\hat{\mathbf{x}}_k^-], \tag{17}$$

$$\mathbf{P}_k = [\mathbf{I} - \mathbf{K}_k\mathbf{C}_k]\mathbf{P}_k^-. \tag{18}$$

Due to the assembled Jacobian of the state space, $\mathbf{F}$ is independent of the incoming fluxes $\phi_4^{(n)}$, and the links between agents are lost. Therefore, to compensate for such loss of connectivity, we define

$$\mathbf{Q} = \mathbf{Q}_p + \mathbf{Q}_g. \tag{19}$$

$\mathbf{Q}_p$ is a diagonal matrix of the variances of the environmental inputs, which are the most relevant for each one of the state variables, and $\mathbf{Q}_g$ is the weighted adjacency matrix of the graph $\mathcal{G}$, which indirectly takes the missing fluxes $\phi_4^{(n)}$ into account. The weighting factor for every element in $\mathbf{Q}_g$ is the maximum reported value in the previous crop season for the selected location. These values are mapped in the $\mathbf{Q}$ matrix as $f(\phi_4)$. Therefore, the adjusted covariance matrix is given by

$$\mathbf{Q} = \begin{bmatrix} \begin{bmatrix} V[u_e^1] & 0 & 0 & 0 \\ 0 & V[u_e^3] & 0 & 0 \\ 0 & 0 & V[u_e^6] & 0 \\ 0 & 0 & 0 & V[u_e^4] \end{bmatrix} & \cdots & [f(\phi_4)] & \cdots & [f(\phi_4)] \\ \vdots & \ddots & \vdots & & \vdots \\ [f(\phi_4)] & \cdots & \begin{bmatrix} V[u_e^1] & 0 & 0 & 0 \\ 0 & V[u_e^3] & 0 & 0 \\ 0 & 0 & V[u_e^6] & 0 \\ 0 & 0 & 0 & V[u_e^4] \end{bmatrix} & \cdots & [f(\phi_4)] \\ \vdots & & \vdots & \ddots & \vdots \\ [f(\phi_4)] & \cdots & [f(\phi_4)] & \cdots & \begin{bmatrix} V[u_e^1] & 0 & 0 & 0 \\ 0 & V[u_e^3] & 0 & 0 \\ 0 & 0 & V[u_e^6] & 0 \\ 0 & 0 & 0 & V[u_e^4] \end{bmatrix} \end{bmatrix}$$

where $V[u_e^1]$, $V[u_e^3]$, $V[u_e^6]$, and $V[u_e^4]$ are the variances of the rainfall, air temperature, maximum daily temperature, and daily solar radiation, respectively.

Based on the procedure to generate the matrices related to measurements proposed by Kadu et al. (2008), matrices $\mathbf{C}_k$ and $\mathbf{R}_k$ are calculated as follows:

$$\tilde{C}_k = \begin{bmatrix} \delta_{1,k} & . & 0 & . & 0 \\ . & . & . & . & . \\ 0 & . & \delta_{i,k} & . & 0 \\ . & . & . & . & . \\ 0 & . & 0 & . & \delta_{n_m,k} \end{bmatrix} \begin{bmatrix} c_{11} & . & c_{1j} & . & c_{1N} \\ . & . & . & . & . \\ c_{i1} & . & c_{ij} & . & c_{iN} \\ . & . & . & . & . \\ c_{n_m1} & . & c_{n_mj} & . & c_{n_mN} \end{bmatrix}$$

$$\tilde{R}_k = \begin{bmatrix} \delta_{1,k} & 0 & . & . & 0 \\ . & . & . & . & . \\ 0 & . & \delta_{i,k} & . & 0 \\ . & . & . & . & . \\ 0 & 0 & . & . & \delta_{n_sn_m,k} \end{bmatrix} \begin{bmatrix} \sigma_1^2 & . & 0 & . & 0 \\ . & . & . & . & . \\ 0 & . & \sigma_i^2 & . & 0 \\ . & . & . & . & . \\ 0 & . & 0 & . & \sigma_{n_sn_m}^2 \end{bmatrix}$$

where c_{ij} is the constant that relates the j_{th} state to the i_{th} measurement, and σ_i^2 is the variance associated with the i_{th} sensor. The variable $\delta_{i,k} = 1$ if the sensor location is active, and $\delta_{i,k} = 0$ otherwise. As the location of patches can change over time due to the search for the best quality of estimation, an operator *shape* removes zero rows from its

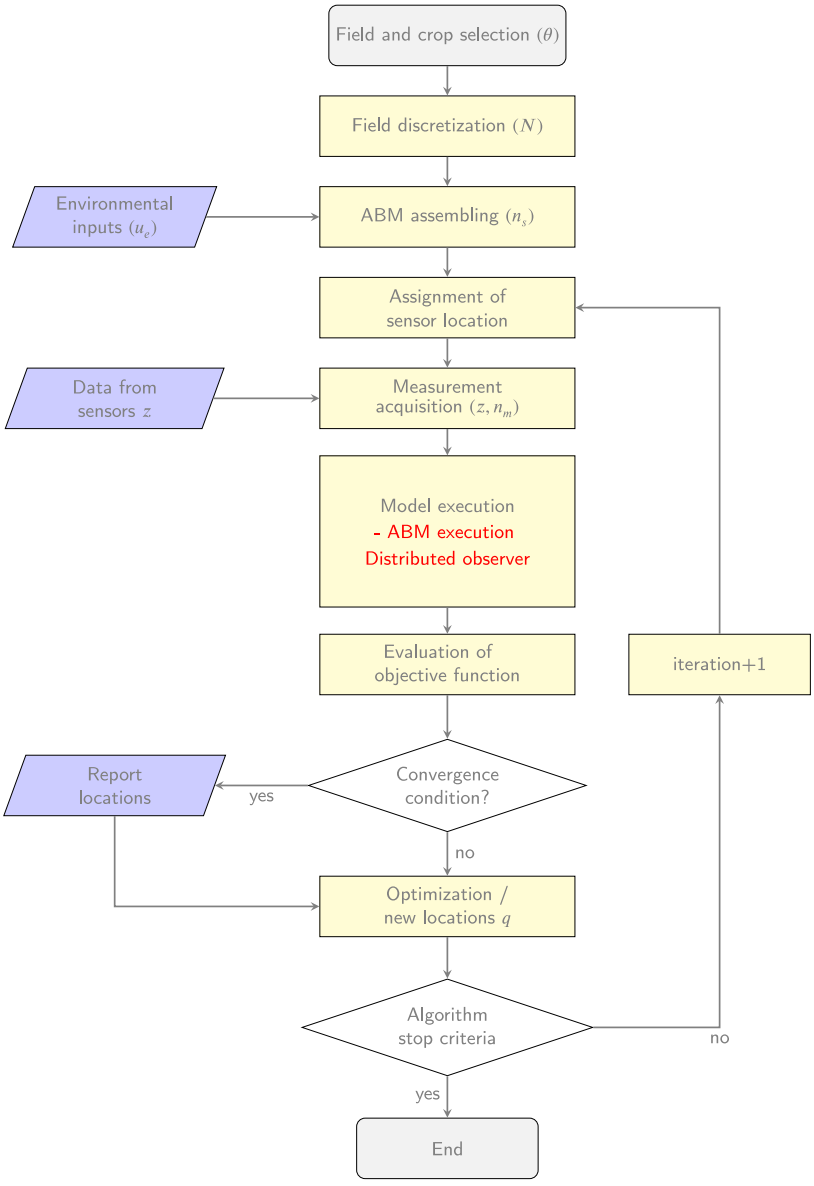

Fig. 2. Flowchart of the algorithm for the location of sensors.

argument, ensuring the appropriate size for $\mathbf{C}_k$ and $\mathbf{R}_k$, The operator over $\tilde{C}_k$ and $\tilde{R}_k$ is defined as,

$$\mathbf{C}_k = shape\{\tilde{C}_k\} \tag{20}$$

$$\mathbf{R}_k = shape\{\tilde{R}_k\}. \tag{21}$$

Notice that the four state variables of each agent are measured when a sensor location is active. On the contrary, if the site is not active, the four states of the corresponding agent must be estimated.

3. Sensor location and distributed observer

The proposed approach follows the flow diagram shown in Fig. 2. It begins with the choice of the field and crop. The selected terrain is represented by a 3D model, where the normalized elevation corresponds to a difference between the highest and the lowest altitude. This representation is built by applying image processing algorithms and a digital elevation model (DEM) (Mukherjee et al., 2013). Afterward, the crop choice and soil type determine the values of vector Θ.

The discretization of the field has the objective of determining the number of agents N that properly represent the changes in the soil dynamics due to water fluxes. For this purpose, we use the error of estimation of biomass. This error is computed as the difference between the production with the minimum patch size (1 m^2) and the production with a bigger patch size. In the resultant grid, every cell becomes a crop–soil agent. The ABM is assembled as described in Section 2, where the dynamics of every patch are linked to the n_s state variables.

To assemble and run the ABM, it is necessary to incorporate the environmental inputs and define what variables could be measured and which ones should be estimated. Subsequently, the number of sensors (n_m) and their distribution in the field should be selected. These locations are used in the sensor equation.

The model is executed after collecting data from sensors and their incorporation into the ABM. This covers the execution of the EKF described in the previous section. Later, an optimization procedure is performed to find the best set of locations with the outcome of prior steps.

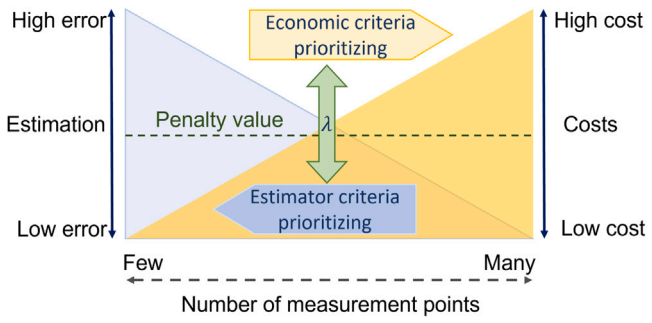

Fig. 3. Graphical representation of the compromises in the objective function.

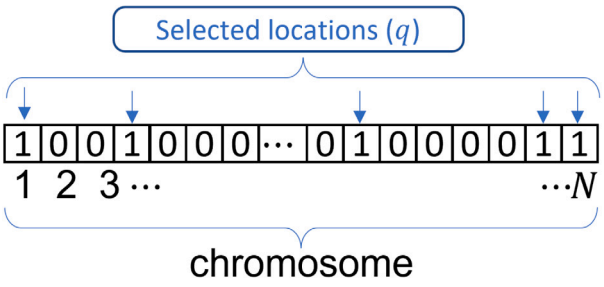

Fig. 4. Codification of sensor locations into a chromosome. The number of ones in the chromosome is n_m.

3.1. Decision criteria

To decide where to locate the sensors, two aspects are considered: the quality of the information obtained with these sensors, and their installation and operation costs. Fig. 3 shows a graphical description of the decision criteria based on the number of sensing points and their location. The left axis shows the quality of estimation related to the estimation error, which decreases as the number of sensors increases. On the right axis, the cost of sensors is depicted, and it is directly proportional to the amount of devices needed. The two components are inspired by the general procedure of a multiobjective optimization proposed by Kadu et al. (2008). These two elements are combined in a single cost function. However, the units and scales are not the same, and therefore, additional compensation elements are added. Moreover, a penalization factor λ is included to prioritize one objective over the other.

The optimization problem can be formulated as follows:

$$\min_{q} \quad \sum_{k=1}^{T}[trace\{P_k\}] + \lambda \left[\sum_{i=1}^{n_m} Fc_i + \sum_{k=1}^{T}\sum_{i=1}^{n_m} \delta_{ik} O_i\right] \tag{22a}$$

$$\text{s.t.} \quad q \in M \tag{22b}$$

$$n_m < N, \tag{22c}$$

where, the decision variable $q = \{m_1, m_2, \dots, m_{n_m}\}$ is one of the possible locations vector picked from $M = \{q_1, q_2, \dots, q_{2^N}\}$, which is the vector of all possible combinations of sensor locations on the field. Notice that q is a set of integers. The term $\sum_{k=1}^{T}[trace\{P_k\}]$ provides a scalar measure of estimation quality. A large value of this term indicates poor quality of estimate or large uncertainty in the state estimation. The second term on the right side of Eq. (22a) stands for the scalar measure of sensing cost, which is calculated as the sum of the installation cost and the operating cost. Fc_i and O_i represent the installation cost and operating cost for the i_{th} sensor, respectively. The variable $\delta_{ik} \in \{0, 1\}$ is defined as in previous section. λ is a weighting factor for balance profitability, scaling, and matching units. Considering the features of the optimization problem as the combinatorial multiobjective cost function with integer decision variables, a genetic algorithm (GA) is proposed to solve the placement problem. The steps and considerations are discussed next.

3.2. Genetic optimization algorithm

Genetic algorithms (GA) are search and optimization techniques based on genetics and natural selection principles (Katoch et al., 2021). GA is built on a set of individuals and different types of rules or genetic operators in the population. The iterative procedure of GA is as follows. First, a population of individuals with a particular structure is created. In this step, the key is the encoding of information in the structure of the individual (i.e., the definition of chromosomes). Second, a set of operators introduce changes in the population (i.e., selection, crossover, mutation). Third, the fitness of the population members is evaluated and the best ranked are selected keeping constant the number of individuals through the iterations. This sequence of steps is repeated until the algorithm meets a stopping criterion (e.g., the maximum number of iterations is reached, the fitness value does not change after a finite number of iterations, or the change in fitness value between consecutive iterations is below a threshold).

The GA is used in this work for its advantages in dealing with non-convex cost functions, the minimal information required to start the search, and the enhanced space exploration. However, this solution methodology can have limitations in terms of its computational speed, issues related to convergence, and the appropriate selection of initial values. Moreover, there is no rigorous guarantee that the solution reaches optimal values.

As mentioned before, the initial step in implementing the GA is the definition of the population and encoding information. Every individual in the population has a unique chromosome. The codification of the information in such chromosomes provides the possible location of sensors. Thus, every chromosome has N binary positions of length. When the particular location is active, the value of this position in the chromosome is one and zero on the contrary. This specific combination of ones and zeros in a candidate chromosome is what we call q, as shown in Fig. 4.

This study employs three distinct genetic operators, namely elitism, crossover, and mutation, as outlined by Konak et al. (2006). The initial value represents a positive integer that indicates the number of individuals in the present generation that are ensured to live and reproduce in the subsequent generation. The second parameter is the proportion of chromosomes contributed by each selected parent in the population, excluding elite offspring, which are utilized by the crossover function to generate novel individuals. The mutation is a crucial component of the genetic algorithm, as it involves introducing minor random alterations to individuals within the population, hence giving rise to mutation offspring. This operator facilitates the introduction of genetic diversity, hence expanding the search field for the genetic algorithm. In this study, the setup of these operators is achieved using MATLAB.

The computational complexity of solving the optimization problem is directly linked to the available sensing locations and the size of the population. The number of computations is related to the scale of the possible locations by 2^N, and it is related to the size of the chromosome. However, the number of individuals in the population is a fixed number. As the assessment of the genetic algorithm to generate new feasible locations is directly related to the computations between members of the population and the population does not change along the solving process, the effect on the complexity is merely arithmetic. For this work, a population of 50 individuals was used in all tests. This size is determined by multiple trials pursuing a balance between computational burden and changes in the population in every iteration of the algorithm.

3.3. Practical observability

As the sensor number and location are investigated, the question of system observability is crucial. A system is said to be observable if the state vector can be estimated using the available output measurements over a finite time period. Observability is a structural property of linear systems but also depends on the input signals for nonlinear systems. Several methods and dedicated software have been proposed for the observability analysis of nonlinear systems (see for instance Díaz-Seoane et al., 2023; Stigter and Joubert, 2021) but the piecewise continuity and nonlinearity of the ABM model, combined with its distributed nature, make the application of these methods difficult. To alleviate this difficulty, a practical observability criterion is proposed instead.

Structural observability tests, in addition to being difficult to apply to our ABM model, also have the disadvantage of providing a yes or no answer, without giving an insight into the degree of observability of the system. Such an insight is however provided by the examination of the error covariance matrix of the Kalman filter. The idea of analyzing the properties of this matrix traces back to the work of Ham and Brown (1983) which highlights the interpretation of the eigenvalues and eigenvectors. In the present study, we make use of the trace of the covariance matrix of the estimation error produced by the Kalman filter. The decreasing value of the trace over time ensures the consistency of the estimator in the spirit of the minimum variance achieved by Kalman filtering. Alternative quantifiers could have been the determinant or the condition number of the covariance matrix.

The minimal set of sensors required to monitor the field while keeping the consistency of the filter over time is also related to the minimal set of paths from a source node to a sink node in the directed graph $\mathcal{G}$ of the water flows (as illustrated in the following application example, see Fig. 8), which provides a physical interpretation of the results. The source and sink nodes are located in the terrain's highest and lowest parts, respectively. Indeed, the flow of information in the system is related to the flow of water on the land surface. Therefore, graph connectivity is a condition to achieve practical observability. If the water flows do not exist, there is no flow of information, and as a consequence, it is not possible to reconstruct the states of the unmeasured patch.

4. Case study

To illustrate the performance of the sensor location strategy, a rugged test field is selected in the province of Samacá in the department of Boyacá, Colombia. The size of the land is 1600 m^2, and the soil is primarily loamy with a thickness of the root zone ($\theta_5^{(n)}$) in the range of 300 mm to 600 mm. A 3D representation of the field is shown in Fig. 5, where the normalized elevation corresponds to a difference between the highest and the lowest altitude (i.e., a difference of about 60 m). This representation is built by applying image processing algorithms and a digital elevation model (DEM) (Mukherjee et al., 2013).

After selecting the field location, the next step is field discretization, which consists of a partition based on a grid. The size of the grid determines the number of agents. This grid is deployed over the ground surface, and the normalized elevation of the center of every square patch is taken as the corresponding agent elevation.

Several simulations are run to define the number of agents, considering the minimal size of grid patches of 1 × 1 m as the reference. This minimal size of patches is intended to provide the best possible representation of the system dynamics related to biomass production. In this paper, after ten simulations with and without irrigation, the average of the errors is calculated for each N. These results are shown in Fig. 6. For the sample square field of 400 × 400 m, a selection of 16 patches ($N = 16$) generates an error in biomass estimation of about 10%. For $N = 81$, the error is about 5%, and for $N = 144$, the error decreases to 3%, which is acceptable.

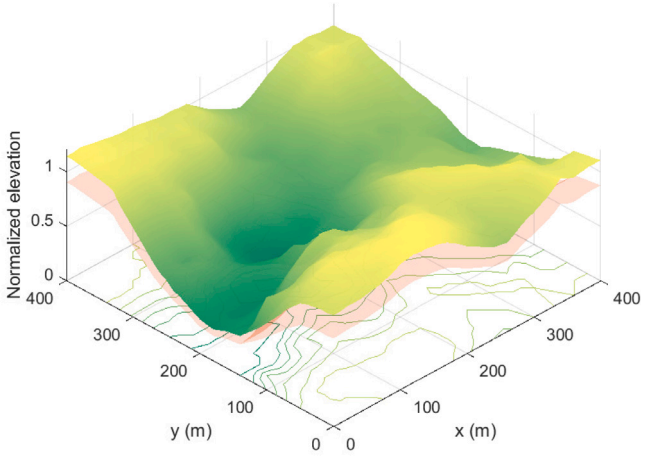

Fig. 5. Digital elevation representation of the test field. The grid in red over the surface indicates the size of agents and their projection to the x, y plane. The number in the corners notes the sequence of the numbering of agents.

Accepting an error in the estimation of about 3%, the ABM with $N = 144$ agents is assembled and executed after collecting their corresponding relative elevations. The agents' tags and the graph with water links related to the relative elevation are shown in Figs. 7 and 8, respectively. The climatic variables to feed the model are retrieved from a meteorological station located at 4 km Northeast for the full 2018 year.

To run the model, the complementary assumptions are, that no in- or out-flux are considered beyond the borders of the field. The environmental inputs are available daily and the values of the vector of parameters $\Theta = \{\theta_j\}$ with $j = 1 \dots 20$, are taken from Lopez-Jimenez et al. (2021). When an agent is selected as a sensing location all state variables are measurable according to the assumptions given in Section 2.1.3 and $n_m < N$.

4.1. Numerical results

The sensor location algorithm is tested in two different cases under no irrigation. In the first case, the optimizer runs under no constraints of the number of sensors, and λ is changing from 0.0001 to 100. The economic part of the cost function is reduced to a scalar value proportional to the number of sensors. The objective here is to set the value of λ that provides a reasonable trade-off between quality and cost. With a fixed λ, the second case solves the optimization problem for a set number of sensors and assesses the performance with the complete cost function.

For practical purposes, the cost function in Eq. (22a) can be rewritten as

$$J = J_1 + \lambda J_2, \tag{23}$$

where J_1 represents the quality of estimation given by the trace of P, and J_2 stands for the cost of measurements. In the first case, i.e., when the economic part is only related to the cost of sensors, the cost structure becomes $J_2 = n_m F_c$, where F_c is a fixed value, independently of the location. For the second case, the cost is as Eq. (22a). The values for the execution of the genetic algorithm are presented in Table 2.

4.1.1. Case 1

In this case, the aim is to evaluate the influence of λ in the cost function J based on the trade-off between the quality of estimation and the affordable cost of monitoring. Therefore, the GA solves the optimization problem under no constraints. In Fig. 9, the execution changing λ from 0.0001 to 100 is shown. A value of $\lambda \in (0.01, 0.1)$

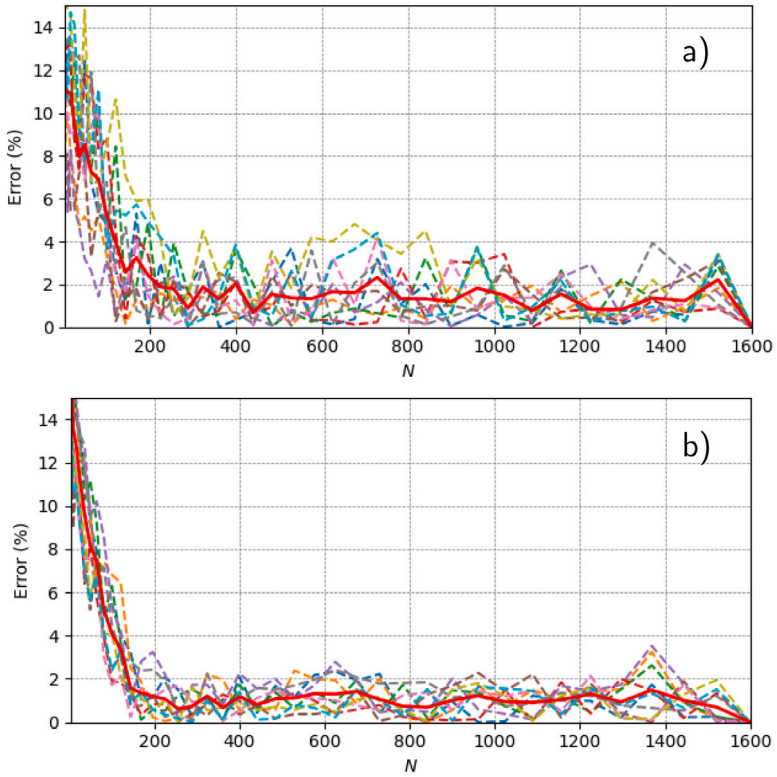

Fig. 6. Error of biomass prediction based on an average discretization of a field (a) under no irrigation; and (b) with a fixed schema of irrigation. The average of ten trials is plotted in red.

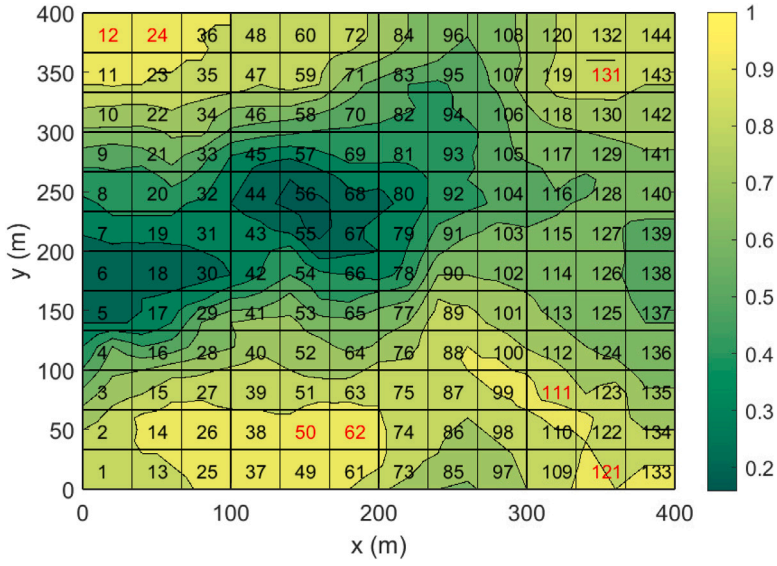

Fig. 7. The test field from above, where the color bar indicates the relative altitude. The agents with no incoming fluxes are labeled in red (See Fig. 8).

is finally selected as it provides a balance between quality and cost. At the top of the graph, the number of sensing locations are displayed with respect to the corresponding value of λ in the horizontal axis.

After assessing the impact of λ in the cost function, it is relevant to know the location of sensors in the field. Fig. 10 shows the location of agents when $n_m = 14$ and $n_m = 40$. The distribution of agents when the priority is the cost reveals clusters of sensors across the field. At the same time, the areas where the focus is the quality of estimation indicate a preference for locations in the highest part of the terrain and broad covering. Therefore, it is suggested to locate sensors in the parts of the terrain with fewer water interactions.

4.1.2. Case 2

In this case, the GA performance is assessed with the constraint of the number of sensors in the field, and the goal of the optimization problem is to focus on finding the best location for a given number of sensors. The advantage is the consideration of a limited sensor budget. Here, J_2 is as defined in Eq. (22a).

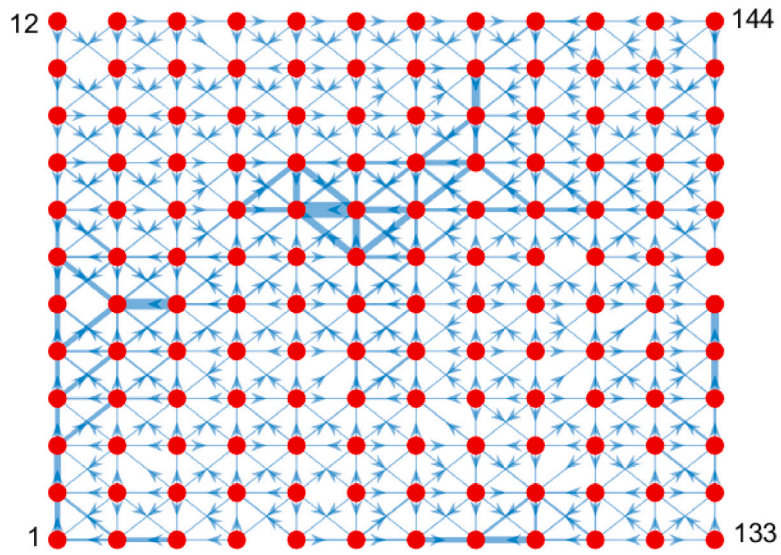

Fig. 8. Digraph of the case study. The arrows in the edges denote the direction of fluxes, while the thickness reveals the relative amount of water for a complete crop cycle.

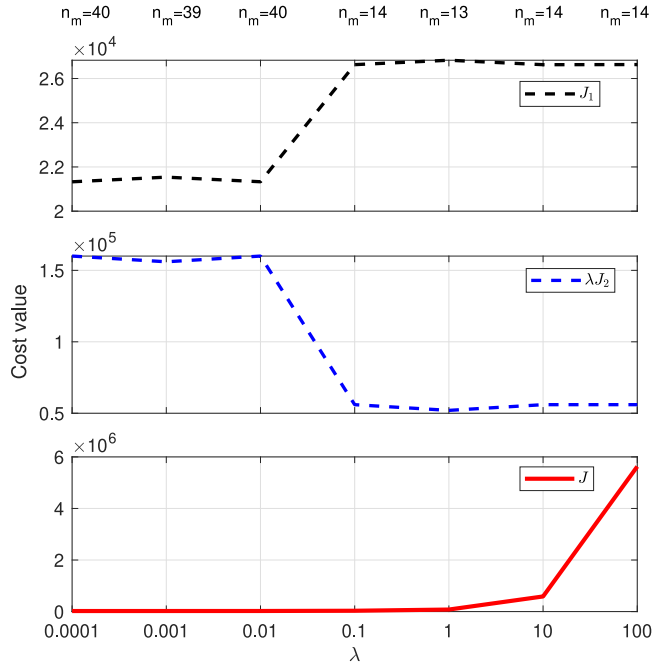

Fig. 9. Evaluation of unrestricted cost function J for $\lambda \in [0.0001, 100]$, where the fix cost is given by $F_c = \$4000$.

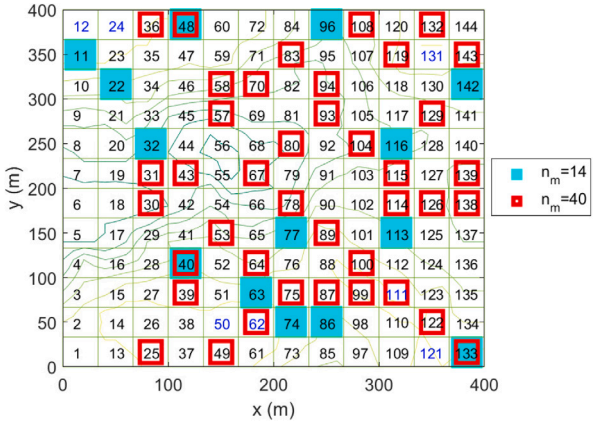

Fig. 10. Sensor placement distribution over the field for $\lambda = 0.01$ and $\lambda = 0.1$, which corresponds to $n_m = 40$ and $n_m = 14$, respectively. In both cases, the algorithm is run for 4000 iterations.

Table 2
Parameters of the GA and the EKF.

Parameter	Value
Population size	50
Length of chromosomes	$N = 144$
Crossover fraction	0.8
Elite count	2.5
Function tolerance	1e−06
Maximum generations	100
Mutation parameter	0.01
k	1 day
t_f	160
$\sigma^2_{x_1}$	0.16 (mm^2)
$\sigma^2_{x_2}$	0.01 (°C^2)
$\sigma^2_{x_3}$	0.01 (°C^2)
$\sigma^2_{x_4}$	0.1225 (ton/ha^2)
F_c	$2000
C_o	$100 per day
O_i	$\gamma^{(n)} \times C_o$

The location results obtained for a fixed number of sensors are shown in Fig. 11. For $n_m = 48$ (representing a third of the possible locations), coverage of the entire field is evidenced by an even distribution of sensing locations. When the number of sensors decreases to 12, the sensors shift to the highest parts of the terrain. The same situation happens when only six possible locations are considered. This indicates spots where there is less information flow, which makes sense given that there is no inflow to these areas except what enters through rain and irrigation.

Additionally, two extreme cases are analyzed with one and two sensors that support the previous conclusion. When the number of sensors is less than 6, the consistency of the estimator is compromised. The value of 6 agents is the limit from which the estimate is unreliable. It is essential to highlight that the good results obtained with $n_m = 6$ rely on the fact that, during more than 80% of the crop cycle, the connectivity of the graph (given by surface water flows $\phi_4^{(n)}$) between agents is preserved.

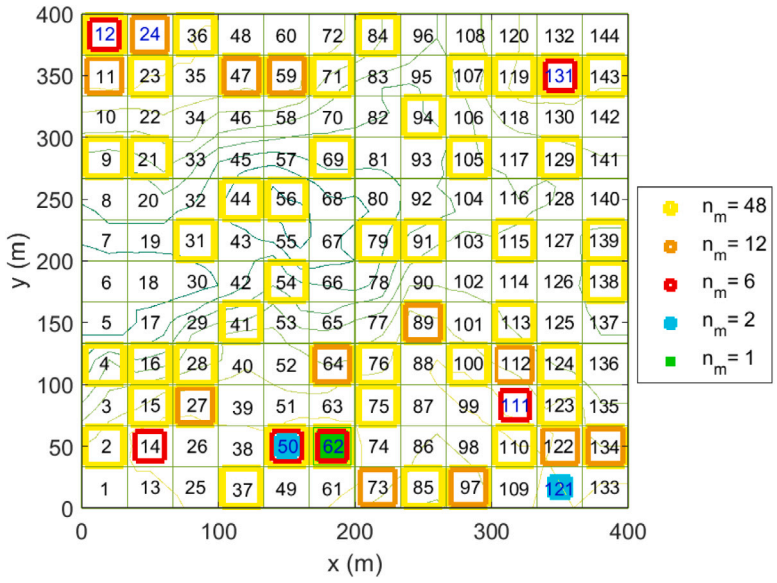

Fig. 11. Optimal location of a given number of sensors with $\lambda = 10$.

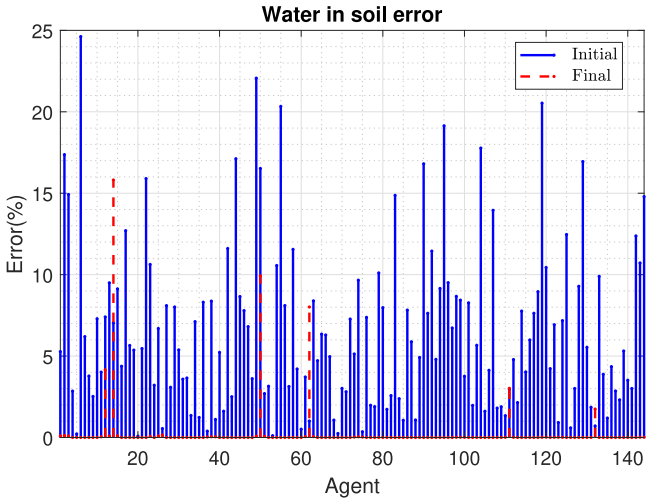

Fig. 12. Comparison between initial and final error in water in the soil for the 144 agents.

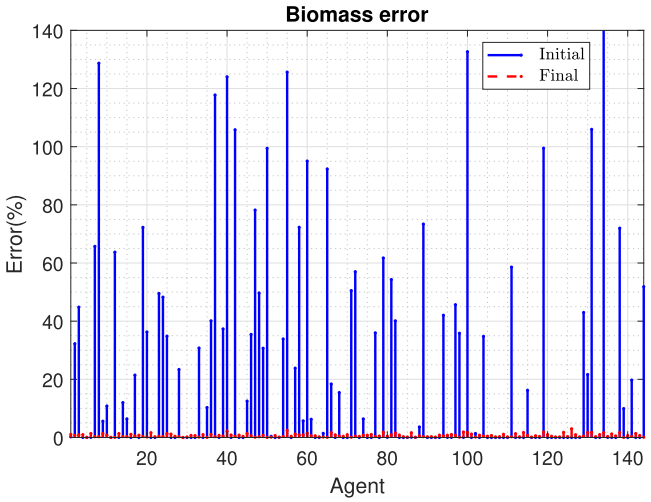

Fig. 13. Comparison between initial and final error in biomass for the 144 agents.

To corroborate this, Fig. 12 shows the initial error given by the difference between the initial condition of the water in the soil for each of the 144 agents and the estimation value and compared with the final error computed as the difference between the simulation of the model and the estimation. The initial conditions for all agents are in the range of 25%, which makes sense for the water-holding capacity of the soil. The only agents where the final error is more significant are those where the sensors are located since the error is calculated concerning the last measurement, which is not necessarily close to the estimated value.

Fig. 13 shows the initial and final biomass error produced for all the agents, evidencing a similar conclusion as in the case of water in the soil. Here, the initial error is set in the span of 140%. To illustrate the temporal evolution of the estimation, a random set of four sensing locations is chosen for water in soil and biomass (see Figs. 14 and 15). Notice that for the sensing location of agent 12, both the amount of water in the soil and the biomass are estimated around the value of the measurements.

The comparative results for both cases suggest that when priority is given to the economic objective, the sensing points are located in the highest parts of the terrain. On the other hand, if relevancy is given to the estimation quality, the algorithm distributes the sensing locations over the field. As the constraint is imposed on the possible sensing sites, the algorithm keeps selecting areas on the highest part of the terrain where no water exchanges and water is going outside, carrying information to agents in the lower parts. In this sense, the algorithm keeps most of the global information. Notice that for extreme cases as a couple of sensors, the locations are skew to high positions, but the algorithm's consistency is no longer preserved.

The performance of the genetic algorithm as an optimization method can be seen in Fig. 16. The generations correspond to each complete iteration of the algorithm, which means an entire cycle of generating a population and applying the genetic operators such as the selection, crossover, and mutation over such population. Additionally, since there are integer variables, the cost function is referred to as the penalty value.

The results indicate that the efficiency of the algorithm decreases after 20 iterations; after that, the improvement in penalty value does

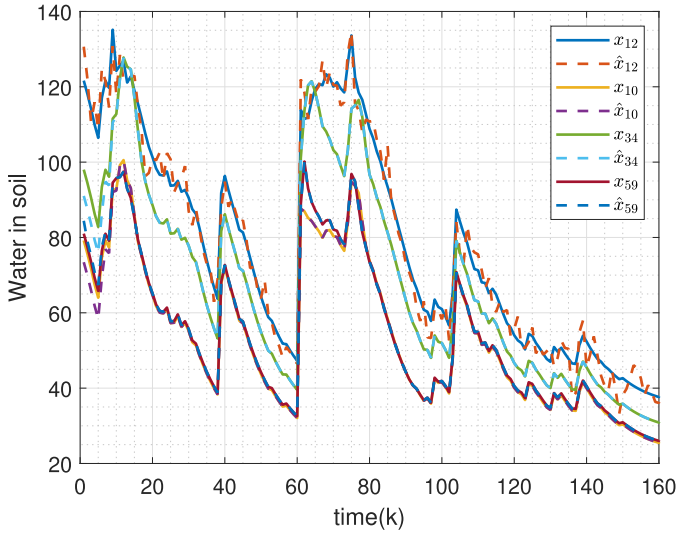

Fig. 14. Water in soil evolution over time for four random locations.

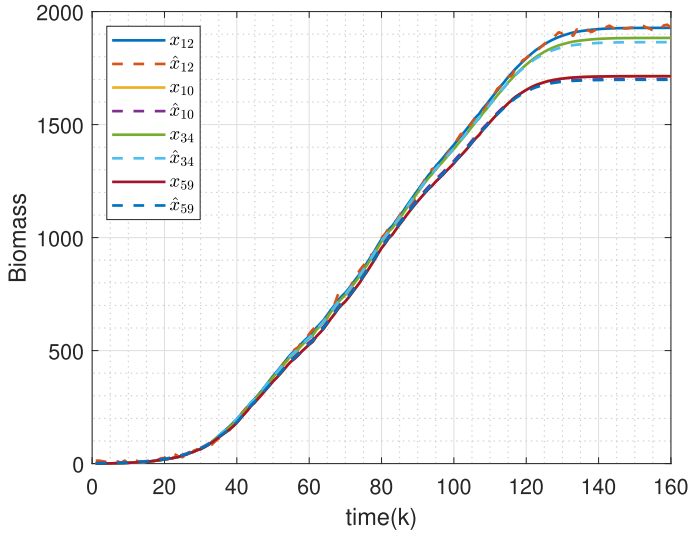

Fig. 15. Biomass evolution over time for four random locations.

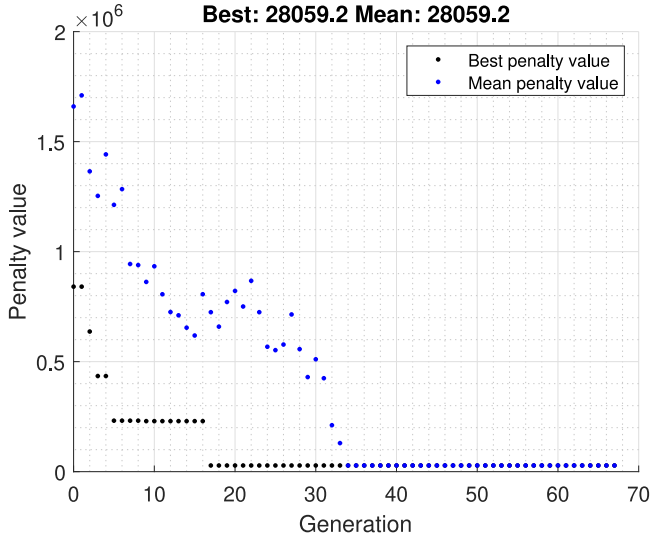

Fig. 16. Evolution of the performance of the GA towards convergence without constraints. The algorithm converges after 67 iterations.

not change significantly. This behavior suggests two possible causes. The first one may be that it is at a local minimum, or the second is that the evolution of the population requires adjustments in the genetic operators. The resulting population can be set as the initial population for a new cycle of algorithm execution.

5. Conclusions and perspectives

In this work, a successful framework to address the problem of sensor location has been presented. It encompasses an ABM for terrain interpretation, an EKF to estimate the states of the areas not sensed, a cost function that combines a factor of quality of estimation given by the trace of the error covariance matrix and accounting for the cost of measurements, and a strategy to solve the sensor placement problem by a GA. The proposed solution is stable over time if the fluxes between agents exist. Therefore the connectivity of the graph that represents the terrain's topography must be guaranteed.

Future work will deal with detectability under low rainfall conditions leading to changes in the estimation technique and the optimization problem.

Declaration of competing interest

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

Data availability

Data will be made available on request.

References

Abioye, E.A., Abidin, M.S.Z., Mahmud, M.S.A., Buyamin, S., Ishak, M.H.I., Abd Rahman, M.K.I., Otuoze, A.O., Onotu, P., Ramli, M.S.A., 2020. A review on monitoring and advanced control strategies for precision irrigation. Comput. Electron. Agric. 173, 105441.

Bwambale, E., Abagale, F.K., Anornu, G.K., 2022. Smart irrigation monitoring and control strategies for improving water use efficiency in precision agriculture: A review. Agricult. Water Manag. 260, 107324.

Chmielewski, D.J., Palmer, T., Manousiouthakis, V., 2002. On the theory of optimal sensor placement. AIChE J. 48 (5), 1001–1012.

Cobbenhagen, A., Antunes, D., van de Molengraft, M., Heemels, W., 2021. Opportunities for control engineering in arable precision agriculture. Annu. Rev. Control 51, 47–55.

Díaz-Seoane, S., Rey Barreiro, X., Villaverde, A.F., 2023. STRIKE-GOLDD 4.0: user-friendly, efficient analysis of structural identifiability and observability. Bioinformatics 39 (1), btac748.

Gao, H., Sabo, J.L., Chen, X., Liu, Z., Yang, Z., Ren, Z., Liu, M., 2018. Landscape heterogeneity and hydrological processes: a review of landscape-based hydrological models. Landsc. Ecol. 33 (9), 1461–1480.

Ham, F.M., Brown, R.G., 1983. Observability, eigenvalues, and Kalman filtering. IEEE Trans. Aerosp. Electron. Syst. (2), 269–273.

Jiang, Y., Huang, Y., Xue, W., Fang, H., 2017. On designing consistent extended Kalman filter. J. Syst. Sci. Complex. 30 (4), 751–764.

Kadu, S.C., Bhushan, M., Gudi, R., 2008. Optimal sensor network design for multirate systems. J. Process Control 18 (6), 594–609.

Katoch, S., Chauhan, S.S., Kumar, V., 2021. A review on genetic algorithm: past, present, and future. Multimedia Tools Appl. 80 (5), 8091–8126.

Konak, A., Coit, D.W., Smith, A.E., 2006. Multi-objective optimization using genetic algorithms: A tutorial. Reliab. Eng. Syst. Saf. 91 (9), 992–1007.

Lopez-Jimenez, J., Quijano, N., Vande Wouwer, A., 2021. An agent-based crop model framework for heterogeneous soils. Agronomy 11 (1), 85.

López-Lozano, R., Baruth, B., 2019. An evaluation framework to build a cost-efficient crop monitoring system. Experiences from the extension of the European crop monitoring system. Agric. Syst. 168, 231–246.

Mukherjee, S., Joshi, P.K., Mukherjee, S., Ghosh, A., Garg, R., Mukhopadhyay, A., 2013. Evaluation of vertical accuracy of open source Digital Elevation Model (DEM). Int. J. Appl. Earth Obs. Geoinf. 21, 205–217.

Paul, P., Bhattacharyya, D., Turton, R., Zitney, S.E., 2016. Dynamic model-based sensor network design algorithm for system efficiency maximization. Comput. Chem. Eng. 89, 27–40.

Pelak, N., Revelli, R., Porporato, A., 2017. A dynamical systems framework for crop models: Toward optimal fertilization and irrigation strategies under climatic variability. Ecol. Model. 365, 80–92.

Perea, R.G., Ballesteros, R., Ortega, J.F., Moreno, M.Á., 2021. Water and energy demand forecasting in large-scale water distribution networks for irrigation using open data and machine learning algorithms. Comput. Electron. Agric. 188, 106327.

Perea, R.G., Poyato, E.C., Montesinos, P., Díaz, J.R., 2019. Prediction of irrigation event occurrence at farm level using optimal decision trees. Comput. Electron. Agric. 157, 173–180.

Rodriguez, L.P., Tupaz, J.A., Sánchez, M.C., 2021. Sensor location for nonlinear state estimation. J. Process Control 100, 11–19.

Rossello, N.B., Carpio, R.F., Gasparri, A., Garone, E., 2019. A novel observer-based architecture for water management in large-scale (Hazelnut) Orchards. IFAC-PapersOnLine 52 (30), 62–69.

Siegfried, R., 2014. Modeling and Simulation of Complex Systems: A Framework for Efficient Agent-Based Modeling and Simulation. Springer.

Stigter, J., Joubert, D., 2021. Computing measures of identifiability, observability, and controllability for a dynamic system model with the StrucID app. IFAC-PapersOnLine 54 (7), 138–143.

Thakur, D., Kumar, Y., Kumar, A., Singh, P.K., 2019. Applicability of wireless sensor networks in precision agriculture: A review. Wirel. Pers. Commun. 107 (1), 471–512.

Visalini, K., Subathra, B., Srinivasan, S., Palmieri, G., Bekiroglu, K., Thiyaku, S., 2019. Sensor placement algorithm with range constraints for precision agriculture. IEEE Aerosp. Electron. Syst. Mag. 34 (6), 4–15.

Wu, B., Zhang, M., Zeng, H., Tian, F., Potgieter, A.B., Qin, X., Yan, N., Chang, S., Zhao, Y., Dong, Q., Boken, V., Plotnikov, D., Guo, H., Wu, F., Zhao, H., Deronde, B., Tits, L., Loupian, E., 2022. Challenges and opportunities in remote sensing-based crop monitoring: a review. Natl. Sci. Rev. 10 (4), nwac290. http://dx.doi.org/10.1093/nsr/nwac290, arXiv:https://academic.oup.com/nsr/article-pdf/10/4/nwac290/49583809/nwac290.pdf.