Article

An Agent-Based Crop Model Framework for Heterogeneous Soils

Jorge Lopez-Jimenez [1,2], Nicanor Quijano [1] and Alain Vande Wouwer [2,*]

1   Department of Electrical and Electronic Engineering, Universidad de los Andes, Bogotá 111711, Colombia; jorgelopez@uniandes.edu.co (J.L.-J.); nquijano@uniandes.edu.co (N.Q.)
2   Systems, Estimation, Control, and Optimization (SECO), University of Mons, 7000 Mons, Belgium
*   Correspondence: alain.vandewouwer@umons.ac.be

Abstract: Climate change and the efficient use of freshwater for irrigation pose a challenge for sustainable agriculture. Traditionally, the prediction of agricultural production is carried out through crop-growth models and historical records of the climatic variables. However, one of the main flaws of these models is that they do not consider the variability of the soil throughout the cultivation area. In addition, with the availability of new information sources (i.e., aerial or satellite images) and low-cost meteorological stations, it is convenient that the models incorporate prediction capabilities to enhance the representation of production scenarios. In this work, an agent-based model (ABM) that considers the soil heterogeneity and water exchanges is proposed. Soil heterogeneity is associated to the combination of individual behaviours of uniform portions of land (agents), while water fluxes are related to the topography. Each agent is characterized by an individual dynamic model, which describes the local crop growth. Moreover, this model considers positive and negative effects of water level, i.e., drought and waterlogging, on the biomass production. The development of the global ABM is oriented to the future use of control strategies and optimal irrigation policies. The model is built bottom-up starting with the definition of agents, and the *Python* environment Mesa is chosen for the implementation. The validation is carried out using three topographic scenarios in Colombia. Results of potential production cases are discussed, and some practical recommendations on the implementation are presented.

Keywords: mathematical modelling; numerical simulation; agriculture; irrigation policy; state estimation; optimal control

1. Introduction

Increasing food production involves efficient use of resources including arable land [1]. However, not all the land available for agricultural activities is found in locations with smooth topography and freshwater availability for irrigation. This poses two challenges: the use of the land and the rational use of water. Land irrigation is a major concern, as it represents the largest consumption of freshwater all over the world [2]. Several attempts to improve the irrigation schemes have been proposed over the years, but a systemic understanding of water effects on crops is necessary and involves the use of software tools to achieve this purpose. As it takes a full season to assess an irrigation policy, research has been essentially carried out through numerical simulation. On the other hand, the land available for agriculture (e.g., wasteland), lacks useful information to forecast a food production scenario. In developing countries there are no detailed soil studies available. Then, a first approach is to assume the homogeneity of land based on a limited sample of the soil and to use traditional crop-growth models. Therefore, the crop-growth models assume homogeneous blocks of soil, which makes sense for small regions but is unrealistic at wide regional scale [3–5]. Improving these models is not straightforward due to the diverse methodologies and assumptions used to build them. This difficulty was evidenced in the first work of the Agricultural Model Intercomparison and Improvement Project (AgMIP), a global initiative aimed at comparing 27 main models and looking for possible multi-model

ensembles [6]. The results reported by Martre et al. [7] indicate that heterogeneity of land at regional scale is a current issue in agriculture.

Some crop simulation models consider soil heterogeneity only with respect to depth, by assuming that the dynamics of the movement of water in the soil only depends on the thickness of the different layers (e.g., AquaCrop [8], STICS [9] and APSIM [10]). However, soil variability across the surface has not been addressed by these models. To face these variations, this work proposes an agent-based model (ABM), that considers a mechanistic approach to crop growth and soil water mass balance. The model proposes a reduced set of equations for an agent to be scalable, and includes uncertainty by adding Gaussian noise in the states and the measurements. The model is fed with climatic variables, irrigation inputs and data from measurements of state variables. With these elements, this model starts from the global behaviour of the crop and allows interpreting the behaviour of smaller portions of land and its response to environmental variables and the amount of water in soil, therefore to propose or to assess irrigation policies.

The application of ABM is not new in agriculture. For instance, an agent structure has been proposed to represent the management decisions of farmers [11,12]. In relation to soil description, [13] presents an ABM framework to model the movement of water through a soil column. To the best of our knowledge, however, no integrated model has addressed surface land variability yet. Instead, large scale models are run in independent simulations and results are collected by a decision support software, e.g., DSSAT in [14]. The topographic surface land variability comes with additional water dynamics that are underestimated by the aforementioned models. To address this issue, a coupling between hydrology and crop growth models was reviewed by Siad et al. [15]. The main conclusions of their work are that coupling models are still at an early stage because data is scarce, and the assessment outputs are reported as regional average.

The main contribution of this work is to propose an agent-based model to represent the complex dynamical interactions between irrigation policy, heterogeneous soil, and crop growth. Each agent represents a portion of land and is associated with a dynamic model of crop growth. The model considers the negative effect of excess or lack of water through an original stress index. Another agent describes the irrigation system, and various irrigation strategies can be assessed in simulation taking constraints on water supply into account. The advantage of the ABM formulation is that it allows flexibility in the definition of the land topography and the consideration of uncertain parameters. In view of the limited (ground or aerial) instrumentation, the control structure involves an estimation scheme based on extended Kalman filtering. The irrigation strategy is based on the optimization of a performance index related to biomass production under irrigation constraints. The software application is illustrated with three topographic scenarios in Colombia. Indeed, 35% of the Colombian territory is available for agriculture but only 7% is currently cultivated. More than 50% of the cultivable area is located on mountainsides, in valleys and on uneven grounds (non-smooth topography and waterlogging in 42%, and 20% of cropping lands, respectively). Moreover, there are sustainable freshwater sources, but they need to be exploited rationally.

This paper is organized as follows. The next section introduces the agent-based model structure, including the estimation scheme and optimal irrigation policy, as well as the implementation in the Python environment Mesa. Three land topologies and four datasets corresponding to different rain patterns are introduced in Section 3, which are used to illustrate the predictive capability of the model and the effect of irrigation on soil stress and biomass production. Conclusions and perspectives are drawn in Section 4.

2. Model Structure and Implementation

Agent-based modelling (ABM) is a methodology used to simulate interactions between autonomous individuals [16]. An agent is a self-contained, self-directed, modular, and uniquely identifiable individual. The modularity requirement implies that an agent has a boundary. Moreover, every agent is an entity located in a specific environment and

capable of autonomous actions, in order to meet some objectives. The complete definition of an agent involves a set of states that varies over time, a set of non-uniform attributes, asynchronous interactions, and uncertainty related to parameters or states [17]. Furthermore, 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. In the first layer, the goal is to interact with other agents, and the second deals with the physical world.

The modelling process can be top-down or bottom-up. The bottom-up strategy is selected in this study because one of the modelling goals is to understand how the individual components affect the all-in-all system behaviour. The emergent system behaviour is the result of agent interactions [18]. A general structure (shown in Figure 1) of an agent-based model has typically three elements:

1. A set of agents, their attributes, and behaviours (lower layer).
2. A set of agent relationships and methods of interaction (i.e., interaction topology or upper layer).
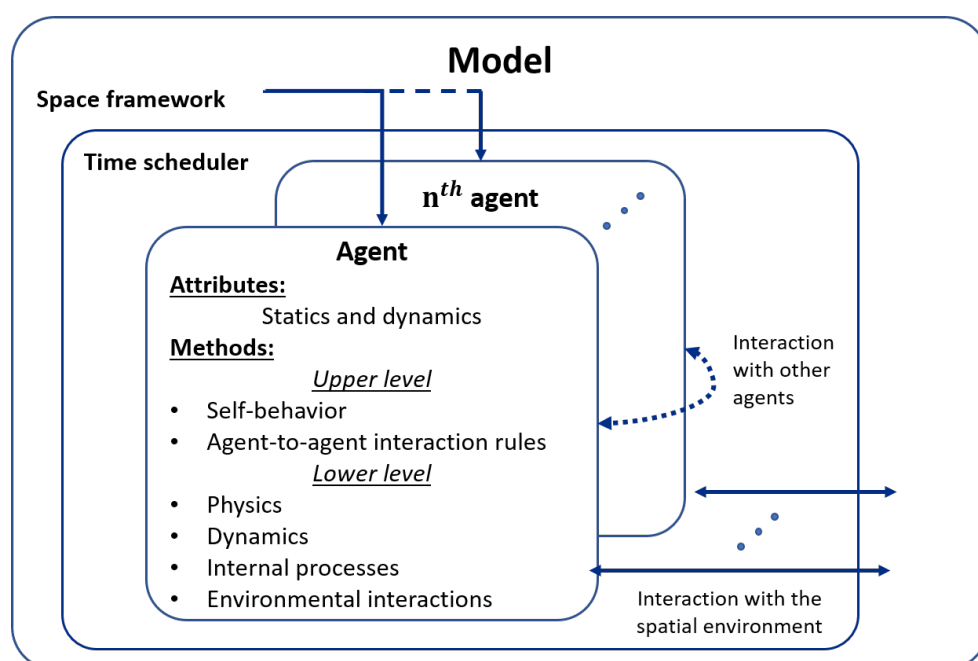3. The agent environment (i.e., spatial topology).

Figure 1. ABM conceptual structure.

For the particular case of this application, two types of agents are considered: (i) portions of homogeneous soil with a surface of regular shape, and (ii) irrigation equipment. The soil agents are associated with a dynamic model representing local crop growth as a function of water availability. The decision or upper level of all agents ensures the respect of the conditions for the exchange of water between neighbouring agents. Crop-soil agents require irrigation at regular time intervals and the irrigation agent answers this demand under the constraints of an irrigation quota.

The global model is intended to predict the evolution of the water content in soil using local information and assess the biomass produced by applying different irrigation schemes. The model structure is shown in Figure 2.

The soil-crop agents are described by a set of state variables $x^n(k)$, which are influenced by the climatic variables $w(k)$ and by the irrigation inputs $u^n(k)$. Part of the state variables can be measured by sensors with output $z^n(k)$ (ground sensors, aerial or satellite measurements) under the influence of measurement errors (or noise) $v^n(k)$. The irrigation agent acts as a controller to determine the water delivered to the crop-soil agents. The control action is determined based on information that it receives at regular time

intervals k regarding the amount of water available for irrigation of the entire field $r(k)$. As instrumentation is limited, an estimator is required which provides state estimates $\hat{x}^n(k)$.
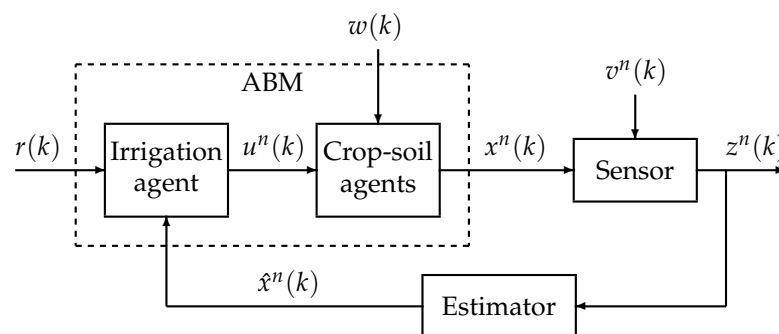

Figure 2. Block diagram of the model with $n = 1, ..., N$ crop-soil agents. All variables are described in Table A1 of Appendix A.

Each of the model blocks are described in the following sections and the compendium of the variables is presented in Table A1. Subsequently, the Python programming environment Mesa is described, as well as the methodology to define agents in their spatial and temporal environments.

2.1. Crop-Soil Agents

The lower level behaviour of every crop-soil agent is related to its internal mechanistic function, and the information coming from other agents. The dynamical evolution of the nth agent is given by

$$\dot{x}^n = \phi_n(x^n, w, u^n, \Theta^n) \qquad n = 1, ..., N, \tag{1}$$

where $\phi_n(\cdot)$ is a nonlinear function, x^n corresponds to the vector of state variables, w the vector of environmental inputs, u^n the vector of management inputs (i.e., irrigation), and Θ^n is the set of parameters related to the nth agent, including soil, crop (e.g., type of cultivar), and management parameters. These parameters can be either constant or time varying. In general, they are known according to a probability distribution, e.g., the soil parameters are stochastic variables accounting for the soil variability.

In this study, the dynamical model underlying the crop-soil agents is built up as a combination of elements of three models. A water mass balance proposed by Ritchie [19,20] (CERES-Wheat), the drought index equation proposed by Woli et al. [21] (ARID model), a biomass growth model presented by Zhao et al. [22] (SIMPLE model), and an original waterlogging stress proposed by the authors (the excess of water reduces the crop transpiration and therefore, the biomass production). This set of equations allows computing the biomass produced by every agent as a response to the environmental inputs and the water content in the corresponding patch of soil. The model is developed in discrete-time using a sampling time of 1 day. This is motivated by two reasons: (i) the growth process at the crop scale is assumed to be affected only by the time course of temperature, which enables the use of thermal time (with units degree days, which represents the cumulative time integral of temperature) to compute the daily aging of the crop; and (ii) the environmental inputs are available daily. Then, the discrete-time equations for every agent $n = 1, ..., N$, can be summarized as follows:

$$x_1^n(k+1) = x_1^n(k) - f_1^n(k) - f_2^n(k) - f_3^n(k) + f_4^n(k) + w_1(k) + u^n(k) \tag{2}$$
$$x_2^n(k+1) = x_2^n(k) + h_1^n(k) \tag{3}$$
$$x_3^n(k+1) = x_3^n(k) + \theta_{11}^n(1 - h_2^n(k)) + \theta_{12}^n(1 - h_3^n(k)) \tag{4}$$
$$x_4^n(k+1) = x_4^n(k) + \theta_{13}^n h_6^n(k) h_7^n(k) h_8^n(k) g^n(k) w_4(k), \tag{5}$$

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 (f_1^n), the surface runoff (f_2^n), the deep drainage (f_3^n), and the flux coming from a neighbour (f_4^n). The independent inputs include the precipitations (w_1) and the solar radiation (w_4). The management input is the amount of water irrigated in every patch (u^n). The functions $h^n(k)$ are continuous or piecewise continuous functions of environmental inputs (w_j) used to compute the stress factors, and $g^n(k)$ is the growth function. Parameters θ_{11}^n, θ_{12}^n, and θ_{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. A complete list of variables and parameters is given in Tables A1 and A2, and the representation of an agent is shown in Figure 3.
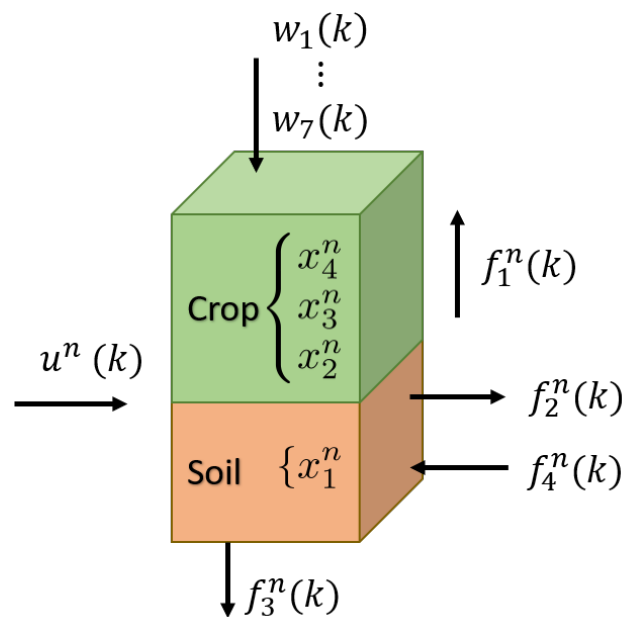

Figure 3. Crop-soil agent structure. All variables are described in Table A1.

Hereafter, the superscript n of all internal functions ($f^n(k)$ and $h^n(k)$), state variables, and parameters is omitted to simplify the notation, keeping in mind that each equation is specific to each agent. The flux components $f^n(k)$ related to Equation (2) are computed following the sequence proposed by [23]. The crop transpiration $f_1(k)$ is given by

$$f_1(k) = min(\theta_1(x_1(k) - \theta_2\theta_5), w_2(k)), \tag{6}$$

where θ_1 is the water uptake coefficient, θ_2 is the wilting point, θ_5 is the root-zone depth, and w_2 is the reference evapotranspiration. The surface runoff $f_2(k)$ is computed as

$$f_2(k) = \begin{cases} \frac{(w_1(k) - \theta_3)^2}{(w_1(k) + 4\theta_3)}, & w_1(k) > \theta_3 \\ 0, & w_1(k) \le \theta_3, \end{cases} \tag{7}$$

where θ_3 is the initial abstraction. The deep drainage $f_3(k)$ is estimated as

$$f_3(k) = \begin{cases} \theta_4(x_1(k) + w_1(k) - f_2(k) - \theta_6\theta_5), & x_1(k) + w_1(k) - f_2(k) > \theta_6\theta_5 \\ 0, & x_1(k) + w_1(k) - f_2(k) \le \theta_6\theta_5, \end{cases} \tag{8}$$

where θ_4 is the drainage coefficient, and θ_6 is the field capacity. The incoming flux $f_4(k)$ is computed as the sum of outflows $f_{out}(k)$ from all the neighbours, which have a normalized elevation γ higher than the considered agent. The outflow of such an agent is given by

$$f_{out}(k) = \begin{cases} \frac{(x_1(k) - \theta_6\theta_5 + f_2(k))}{N_r}, & x_1(k) > \theta_6\theta_5 \\ 0, & x_1(k) \leq \theta_6\theta_5, \end{cases} \tag{9}$$

where N_r is the number of receiving neighbours. The function $h_1(k)$ in Equation (3) represents the daily mean temperature added to the state variable x_2, i.e.,

$$h_1(k) = \begin{cases} w_3(k) - \theta_7, & w_3(k) > \theta_7 \\ 0, & w_3(k) \leq \theta_7, \end{cases} \tag{10}$$

where $w_3(k)$ is the mean air temperature and θ_7 is the base temperature for phenology development and growth. The cumulative temperature required to reach 50% of radiation interception during canopy senescence x_3 is increased by heat stress (i.e., faster canopy senescence) and drought stress. The heat stress $h_2(k)$ is computed as follows

$$h_2(k) = \begin{cases} 1, & w_6(k) \leq \theta_9 \\ 1 - \frac{w_6(k) - \theta_9}{\theta_{10} - \theta_9}, & \theta_9 < w_6(k) \leq \theta_{10} \\ 0, & w_6(k) > \theta_{10}, \end{cases} \tag{11}$$

where θ_9 is the threshold temperature to start accelerating senescence from heat stress, θ_{10} is the extreme temperature threshold when radiation use efficiency (RUE) becomes 0 due to heat stress, and $w_6(k)$ is the maximum daily temperature. The drought stress $h_3(k)$ is built based on the ARID index [21],

$$h_3(k) = 1 - \theta_{14}h_4(k), \tag{12}$$

where θ_{14} is the sensitivity factor to RUE, and

$$h_4(k) = \begin{cases} 1 - \frac{f_1(k)}{w_2(k)}, & f_1(k) < w_2(k) \\ 0, & f_1(k) \geq w_2(k). \end{cases} \tag{13}$$

One of the novel features of the model is to consider the negative effect of water excess. Hence, a waterlogging stress factor $h_5(k)$ is proposed in Equation (14), i.e.,

$$h_5(k) = \begin{cases} 1, & S_t(k) \leq \theta_{15} \\ 1 - \frac{S_t(k) - \theta_{15}}{\theta_{16} - \theta_{15}}, & \theta_{15} < S_t(k) \leq \theta_{16} \\ 0, & S_t(k) > \theta_{16}, \end{cases} \tag{14}$$

where θ_{15} is the stress time lower limit, and θ_{16} is the stress time upper limit. These limits are specific for every crop, and the values can be retrieved from literature. This index is built based on the anaerobic root stress reported by [24,25]. The rationale behind this index is that the negative effect is proportional to the amount of days $(S_t(k))$ under anaerobic stress. The limits are taken from the literature. The stress shape is also suggested by Steduto et al. [8]. The counting of days under waterlogging stress is computed as

$$S_t(k) = \begin{cases} S_t(k) + 1, & x_1(k) > \theta_6\theta_5 \\ 0, & x_1(k) \leq \theta_6\theta_5. \end{cases} \tag{15}$$

The combined negative effect of water excess, lack, and excessive heat is summarized into $h_6(k)$ using a minimum function due to the rationality that the crop never suffers simultaneously of more than one stress related to transpiration, in other words,

$$h_6(k) = min(h_2(k), h_3(k), h_5(k)). \tag{16}$$

The impact of low temperatures on biomass growth rate is calculated as follows

$$h_7(k) = \begin{cases} 0, & w_3(k) < \theta_7 \\ \frac{w_3(k) - \theta_7}{\theta_8 - \theta_7}, & \theta_7 \le w_3(k) < \theta_8 \\ 1, & w_3(k) \ge \theta_8, \end{cases} \tag{17}$$

where θ_8 is the optimal temperature for biomass growth. Then, the CO_2 impact in biomass growth rate $h_8(k)$ is estimated as

$$h_8(k) = \begin{cases} 1 + \theta_{17}(w_7(k) - 350), & 350 \le w_7(k) < 700 \\ 1 + 350\theta_{17}, & w_7(k) \ge 700, \end{cases} \tag{18}$$

where θ_{17} is the relative increase in RUE due to a higher CO_2 concentration, and $w_7(k)$ is the atmospheric CO_2 concentration. Finally, the growth function $g^n(k)$ is estimated as

$$g^n(k) = \begin{cases} \frac{\theta_{19}}{1 + e^{-0.01(x_2(k) - \theta_{20})}}, & x_2(k) \le \frac{\theta_{18}}{2} \\ \frac{\theta_{19}}{1 + e^{0.01(x_2(k) + x_3(k) - \theta_{18})}}, & x_2(k) > \frac{\theta_{18}}{2}, \end{cases} \tag{19}$$

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.

2.2. Irrigation Agent

This agent is the only one that interacts with all crop-soil agents. Its mission is to distribute water for irrigation, either following an arbitrary policy determined by some heuristics or an optimal policy resulting from the formulation of a performance index and the solution of an optimal control problem. There is an array of optimal control policies that could be considered including crop yield, water consumption, and constraints on water availability or distribution. Here, a very simple problem is considered at first, where the performance index corresponds to the average daily biomass, subject to the constraint of maximum available water on a daily basis W_{max}, i.e.,

$$\max_{u^n(k)} \quad \frac{1}{N} \sum_{n=1}^{N} \hat{x}_4^n(k) \tag{20}$$

$$\text{s.t.} \quad 0 \le u^n(k) \le W_{max}, \tag{21}$$

where N is the number of agents, $\hat{x}_4^n(k)$ is the biomass estimate and $u^n(k)$ is the irrigation for the crop-soil agent n at day k, respectively.

After the definition of the model agents, it is necessary to examine the flow of information from the sensor.

2.3. Sensor

The sensing technology considered in this work is 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. For every agent $n = 1, ..., N$, only $x_1^n(k)$, $x_2^n(k)$, and $x_4^n(k)$ are measurable accordingly to [26–28]. The sensor equation is

$$z^n(k) = H^n x^n(k) + v^n(k), \tag{22}$$

where $z^n(k)$ is the vector of sensor outputs, H^n is the measurement matrix, and $v^n(k)$ is a vector that represents the measurements errors. In this preliminary study, the measurement noise is assumed to follow a normal distribution with covariance matrix $R^n(k)$, but other distributions could be considered as well. All measurements are assumed uncorrelated. For the case of $x_4^n(k)$, the measurement is a fraction of the crop biomass (i.e., above ground

biomass) [5]. With these signals, the next step is to infer the unmeasurable state variables using state estimation techniques.

2.4. Estimator

There are two reasons to incorporate an estimator in the model. First, not all state variables are measurable, and second, it is necessary to incorporate sensor data to adjust the model predictive capabilities. In view of the assumed Gaussian distribution of the noise, a classical solution is provided by an extended Kalman filter (EKF) in discrete-time form [29]. In the framework of the ABM model, the filter is deployed in a decentralized way by associating to each agent a simplified version of the filter where the coupling between agents is partly neglected.

The discrete-time state space representation of a crop-soil agent (Equation (1)) becomes

$$x^n(k+1) = \phi^n(x^n(k), w(k), u^n(k), \Theta^n) + \omega^n(k), \tag{23}$$

with the consideration of a zero-mean Gaussian process noise $\omega^n(k)$ with covariance matrix $Q^n(k)$. The state space equation is linearised along the estimated trajectory and the Jacobian matrix is obtained by partial differentiation

$$\Phi(k) = \left.\frac{\partial \phi^n(\cdot)}{\partial x^n}\right|_{x^n = \hat{x}^n(k)}, \tag{24}$$

where the coupling between agents is neglected so as to achieve a simple local evaluation (the form of $\Phi(k)$ is detailed in Appendix B).

The recursive algorithm of the filter begins with the initialisation with the initial state $\hat{x}^n(0)$, and the error covariance matrix $P^n(0)$. Hereafter, the n super index is omitted for simplicity and prior estimates are denoted with an index $^-$. The algorithm is divided into a prediction (or time update) step:

$$\hat{x}^-(k) = \phi(\hat{x}(k-1), w_j(k), u^n(k), \Theta^n). \tag{25}$$

$$P^-(k) = \Phi(k-1)P(k-1)\Phi^\top(k-1) + Q(k-1). \tag{26}$$

and a correction (or measurement update) step using the correction gain $K(k)$:

$$K(k) = P^-(k)H^\top(k)[H(k)P^-(k)H^\top(k) + R(k)]^{-1}, \tag{27}$$

$$\hat{x}(k) = \hat{x}^-(k) + K(k)[z(k) - H(k)\hat{x}^-(k)], \tag{28}$$

$$P(k) = [I - K(k)H(k)]P^-(k). \tag{29}$$

Every agent daily iterates the sequence of Equations (25) to (29). The considerations about the selection and computation of matrices $Q(k)$, $R(k)$ and $P(k)$, follow the recommendations of Reif et al. [30] to guarantee that $\|\Phi(k)\|$ is bounded and $\Phi(k)$ is nonsingular. These recommendations include a positive upper and lower bound for $P(k)$, and a positive lower bound for $Q(k)$ and $R(k)$ (which are positive definite matrix defining the process and measurement noise covariances).

As the exchange flux f_4 is not considered in the Jacobian evaluation, the estimator may be affected particularly regarding the estimation of the water content $\hat{x}_1$. An ad-hoc procedure is therefore suggested, which consists in "retuning" the variance of the process noise on the first state variable as follows

$$Q_m(1,1)(k) = Q(1,1)(k) + \alpha V[w_1(k)], \tag{30}$$

where α is a weight factor and $V[w_1(k)]$ is the variance of rainfall. There is indeed a direct relation between the rainfall pattern and the intensity of the exchange flux f_4. This

modification allows modulating the uncertainty related to the lack of exchange flux in the model linearisation. An example of tuning of the EKF is given in Appendix B, which is not meant to be optimal, but gives reasonable results.

2.5. Agent Spatial Environment

The construction of the ABM requires the selection of the crop and geographic location, e.g., species, cultivar, soil properties. Based on aerial images of the field, image processing is performed, e.g., color to grayscale, binarization, edge detection and region extraction [31,32]. This processing yields a 2-D map, which can be discretized using a spatial grid. The grid size is established based on the properties of the image and the knowledge about the field, e.g., information of the previous crop season. Then, an assignment of the N agents is carried out. The minimum ground area for parameter calibration is equal to 1 m^2, which is selected as the lower limit of agent size. Then, the maximum number of agents is the ratio between the minimum ground area and the total crop area. Once a label is assigned to an agent, the algorithm extracts extra information from the image, e.g., the normalized elevation parameter γ for every agent [33,34]. Subsequently, a graph is created, as well as the adjacency matrix and the matrix of neighbours. As the number of agents increases, the graph can be used to decompose the cropping region into smaller portions while keeping the relationships between neighbours.

2.6. Model Implementation

The survey paper [35] presents a comprehensive comparative literature review of the state-of-the-art software development in agent-based modelling. Mesa, a modular agent-based modelling framework in Python [36], appears as an efficient tool to develop the crop-soil model proposed in this study. It allows users to create agent-based models using built-in core components, such as spatial grids and agent schedulers, or customized implementations, to visualize them using a browser-based interface, and to analyse the results using Python data analysis tools. The modules are grouped into three categories:

1. **Modeling:** Modules used to build the models themselves: a model and agent classes, a scheduler to determine the sequence in which the agents act, and space for them to move around.
2. **Analysis:** Tools to collect data generated from the model, or to run it multiple times with different parameter values.
3. **Visualization:** Classes to create and launch an interactive model visualization, using a server with a JavaScript interface.

Once the image processing routine is executed and data is stored in text files, the model is built. This model is split into three files. The first file contains the model definition, where the irrigation agent is incorporated as a "method" (an object-oriented programming procedure) of the model, as there is no physical location of this agent (see Algorithm A2 of Appendix C). The second file contains the crop-soil agent definition and methods (see Algorithm A1 of Appendix C). Finally, the last file contains the web visualization and the plotting routine for the state variables. The whole model is intended to advance daily ($k = 1d$). However, every agent can behave asynchronously by setting its own *self_time* variable.

3. Case Studies

In order to assess the functionality of the model, three different scenarios with heterogeneous soil are shown in Figure 4. These examples only consider the water supply problem, but no nutrient stress. The sample fields are located in the region of Boyacá (Colombia) between 2400 and 2800 m.a.s.l. (meters above sea level), where there is no in situ meteorological station. Instead, the data comes from a meteorological station located 3 km away, in Villa Carmen (latitude = 5.51°, and longitude = −73.50°). The weather information is organized in four data sets between March 2017 and February 2019, considering

that the sowing months are March and September and the harvest months are August and February. These datasets expose different rain patterns:

- First dataset (March to August 2017), strong peak of rain in the middle of the season and 483 mm of rainfall.
- Second dataset (September 2017 to February 2018), strong peak of rain in the first half of the season and 244 mm of rainfall.
- Third dataset (March to August 2018), persistent rain in the second half of the season and 323 mm of rainfall.
- Fourth dataset (September 2018 to February 2019), persistent rain in the first half of the season and 325 mm of rainfall.

For all scenarios, the crop is wheat of the same cultivar (i.e., yecora rojo) and the land size is 1600 m^2. The study area is a square shaped land, and no incoming and outgoing fluxes beyond this area are considered. After image processing, $N = 16$ agents are chosen, and the parameters Θ^n are adjusted to match the regional reported values. All parameters are listed in Table A2.
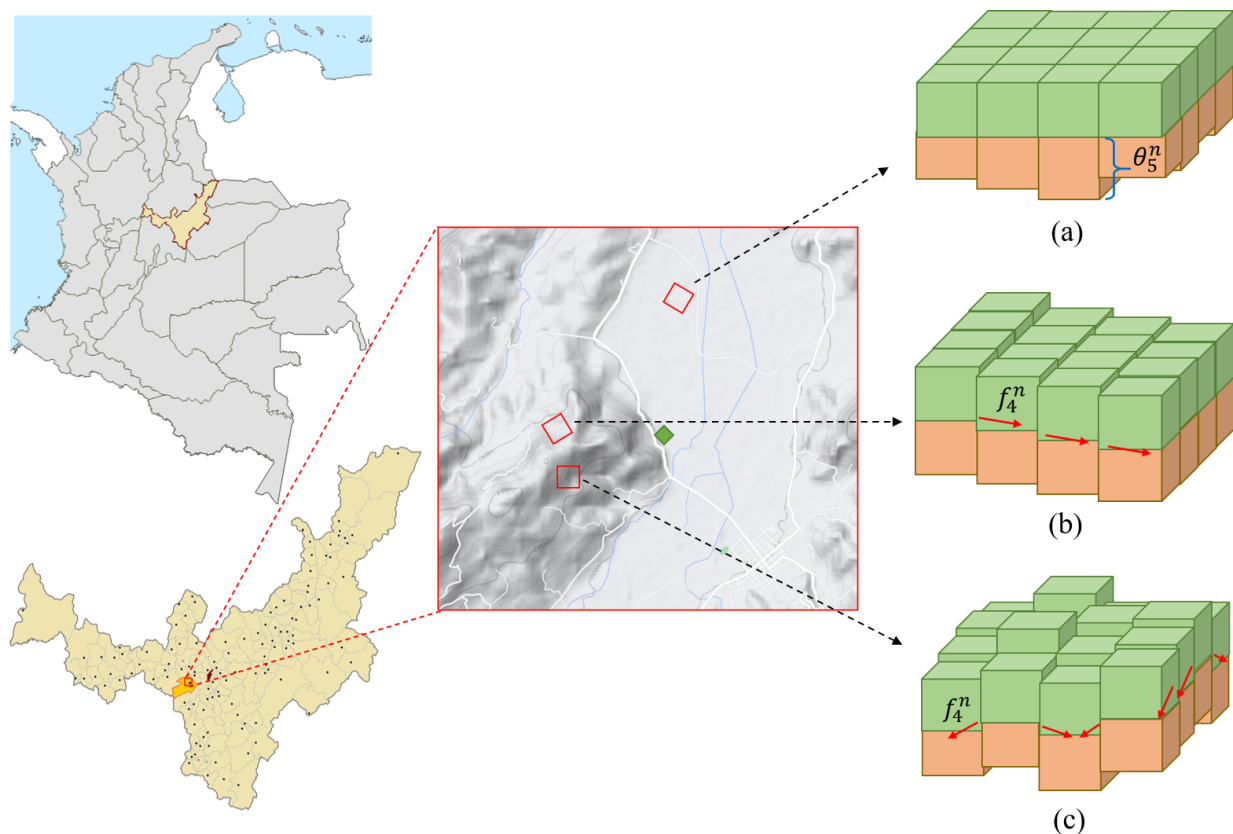

Figure 4. Simulation scenarios and their location in Samacá province of Boyacá region. (**a**) Flat land. (**b**) Slope land. (**c**) Rugged land. θ_5^n and f_4^n are the root-zone depth and incoming flux for every agent, respectively. The green diamond shows the location of Villa Carmen.

The first scenario is a flat terrain (i.e., slope less than 3%) in which the root-zone depth θ_5^n changes according to a uniform probability distribution as suggested in Woli et al. [37]. This parameter was selected after evaluating the impact that the parameters θ_1 to θ_6 have on the state variable x_1, following the methodology proposed by Woli et al. [37]. This study shows that the root-zone depth parameter has a 42% influence on the output based on the Fourier amplitude sensitivity test. This scenario is evaluated during the two sowing periods each year. The second scenario considers a sloping ground (i.e., slope around 20%) with the same climatic data as in the previous scenario. Finally, the third scenario is

concerned with a rugged ground that allows evaluating different interactions between the agents due to the topography. The soil is mostly loamy and sandy-loamy in all scenarios.

For this case study, aerial images are only considered to configure each of the scenarios and establish the initial conditions of the parameters and variables. Subsequently, the daily information from the sensors is generated by the model (i.e., Equations (2) to (5)) with the addition of Gaussian noise with zero mean and standard deviation in the range described by [26,27], and [28] for the variables $x_1(k)$, $x_2(k)$ and $x_4(k)$, respectively.

3.1. First Scenario

In this scenario, there is no exchange of water between agents. Instead, the root depth zone of every agent θ_5 is chosen from a uniform distribution $U(300, 600)$. The assignment of agents is presented in Figure 5. The value of $x_1^n(0)$ is at field capacity for all agents. To assess the irrigation effect on every agent, the model runs first with no irrigation and then with the irrigation policy. For illustration purposes, only agent 1 is considered. All state variables are displayed in Figure 6 as all of them could be considered to make management decisions, i.e., x_1 for water use, x_2 to control the phenological development of the crop, x_3 to evaluate the effect of stress factors in crop transpiration, and x_4 to evaluate production. The estimator is working satisfactorily with a convergence time of about 30 days. x_3, which is the only unmeasured variable, is estimated with a slight offset. The error of estimation of the biomass x_4, which is used in our performance index to determine the irrigation policy, is ranging between 0.5 and 5% in the several simulation tests, see Table 1.
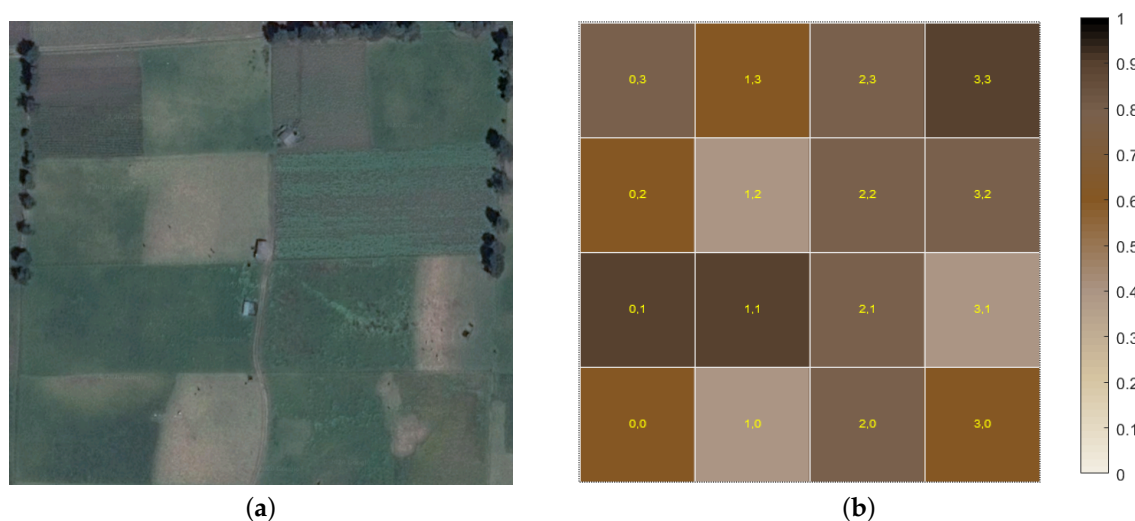

Figure 5. First scenario. (**a**) Aerial view of the land. (**b**) Grid representation after image processing. Here, the size of the side of every patch is 100 m and the coordinates identifier is depicted in the center of patches. The color of every patch indicates the normalized amount of water at time k, the darker means high water content and the lighter indicate dryness.

Figure 7 shows the normalized soil water content with respect to the field capacity without and with irrigation. In the latter case, the water level is around field capacity. Moreover, the water stress indexes (i.e., drought and waterlogging) indicate that only if the field capacity is exceeded for more than 3 days the waterlogging stress is triggered. After irrigation, the drought stress decreases and the field becomes prone to waterlogging stress when the precipitations increase. The positive effect of irrigation is seen in the increment of the production of biomass. Table 1, shows that the water quantity used for irrigation is lower than the total rainfall except in the situation of dataset 2 where it is 132 mm higher.
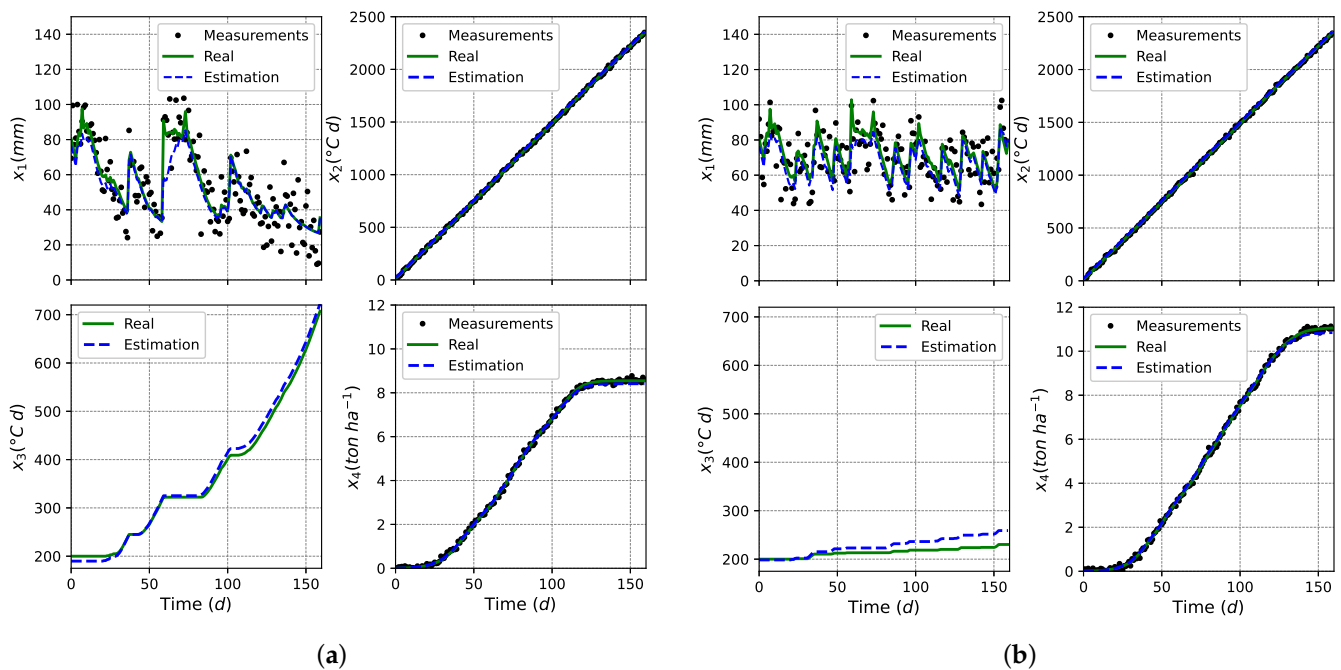
Figure 6. State variables for sample agent 1 under (**a**) no irrigation, and (**b**) irrigation policy for the first weather dataset.

Table 1. Summary of water effect on biomass production for representative agents in each scenario.

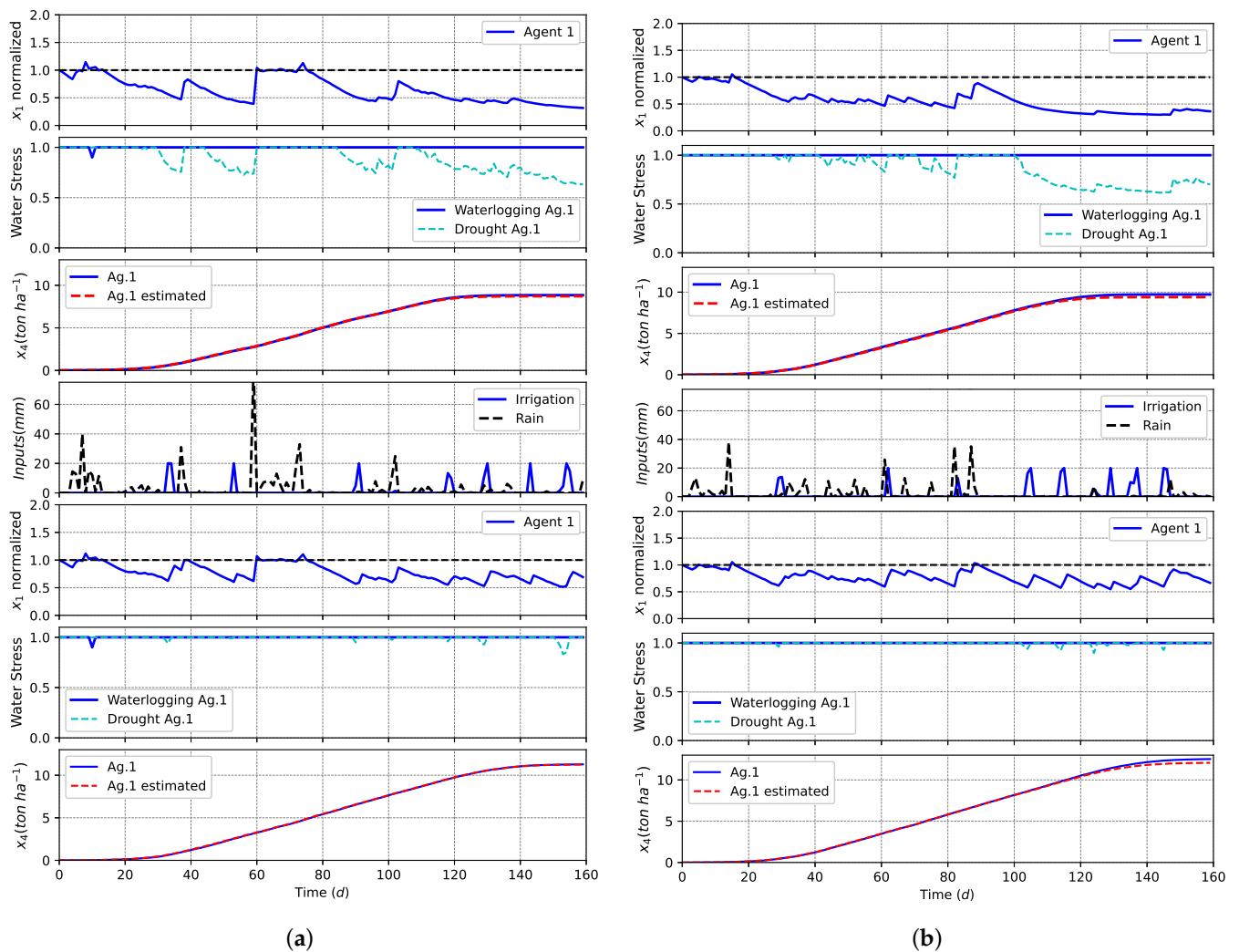| | | | No Irrigation | | | | Irrigation | | | | |
| | | | Water Stresses | | | | Water Stresses | | | | |
Scenario	Dataset	Agent	Drought (d)	Waterlogging (d)	x_4 (ton/ha)	$\hat{x}_4$ (ton/ha)	Drought (d)	Waterlogging (d)	x_4 (ton/ha)	$\hat{x}_4$ (ton/ha)	Irrigation (mm)
1	1	1	97	1	8.7	8.5	5	1	11.4	11.3	188
	2	1	108	0	8.2	8.1	3	0	12.4	12.0	356
	3	1	109	1	7.0	6.7	8	3	11.3	11.4	282
	4	1	99	0	9.6	9.3	8	0	12.6	12.4	324
2	1	1	89	1	9.2	9.0	7	8	11.2	11.3	324
		16	56	37	8.2	8.3	8	55	8.0	8.2	215
	2	1	108	0	8.2	8.1	40	1	11.6	9.7	291
		16	109	0	8.2	8.0	32	1	11.8	11.1	325
	3	1	107	0	7.1	6.7	6	2	11.4	11.3	299
		16	107	0	7.1	6.7	8	2	11.4	11.3	275
	4	1	93	0	9.8	9.5	7	0	12.6	12.5	303
		16	70	4	10.2	10.2	11	19	11.6	12.0	200
3	1	2	96	1	8.7	8.6	5	4	11.3	11.4	227
		10	59	38	8.0	7.8	33	58	7.9	7.1	145
	2	2	108	0	8.3	8.0	11	0	12.0	11.8	336
		10	111	0	8.1	8.0	7	0	12.1	12.1	329
	3	2	116	1	6.7	6.4	3	2	11.4	11.2	301
		10	115	1	6.8	6.4	4	2	11.4	11.2	259
	4	2	93	0	9.8	9.4	13	0	12.6	12.5	318
		10	69	4	10.2	10.3	53	22	10.0	6.9	109

Figure 7. Water effect on agent 1 in a flat ground (first scenario) and (**a**) dataset 1, (**b**) dataset 4. In each case, from top to bottom: soil water content normalized under no irrigation, water stresses under no irrigation, biomass under no irrigation, water inputs, soil water normalized content after irrigation, water stresses under irrigation, biomass under irrigation.

3.2. Second Scenario

With a sloping ground, the water flux $f_4^n(k) \neq 0$ from the most elevated side of the terrain as represented by the arrows in the graph (see Figure 8). To illustrate this situation, agents 1 and 16 are selected, which are in the highest and lowest part of the ground, respectively, and also are at maximum distance from one another.

The trajectories of Agent 1 and 16 and their state estimates are shown in Figure 9 for state variables x_1, and x_3. The performance of the state estimator is satisfactory for agent 1, but the estimation of the water content x_1 of agent 16 suffers from transient errors particularly in periods of strong rainfall, where the water content is underestimated. The tuning parameter α can be used to partly alleviate this situation, but is not enough to guarantee a good estimation in all cases. After irrigation, state x_3 decreases as a result of the reduction in the drought stress and the increase of the crop transpiration.
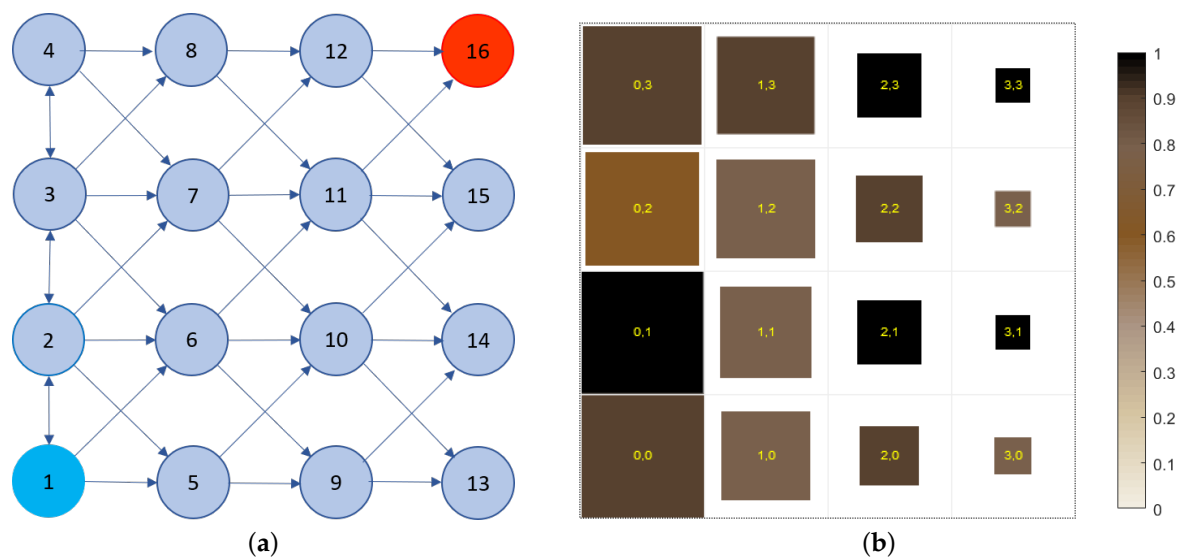
Figure 8. Second scenario: (**a**) Possible water flux exchanges (**b**) Grid representation after image processing. The width of each square represents the normalized elevation of the portion of land with respect to the highest point in the grid. The maximum elevation considered is around 10 m. The color of every patch indicates the normalized amount of water at time k, the darker means high water content and the lighter indicate dryness.

Figure 9. State variables x_1, and x_3 and their estimates for agents 1 and 16 under (**a**) no irrigation, and (**b**) irrigation policy for the first dataset.

Figure 10 shows the water effect on biomass production for the first and fourth climatic datasets.

The waterlogging stress during the first 30 days of growth has no significant impact on the final biomass production. However, after this period the irrigation increases the waterlogging stress of agent 16 under strong rainfall, and delays the biomass production. The biomass production x_4 increases by 22, and 29% for agent 1 but decreases by 2.4, and 14% for agent 16 for datasets 1 and 4, respectively (see Table 1).
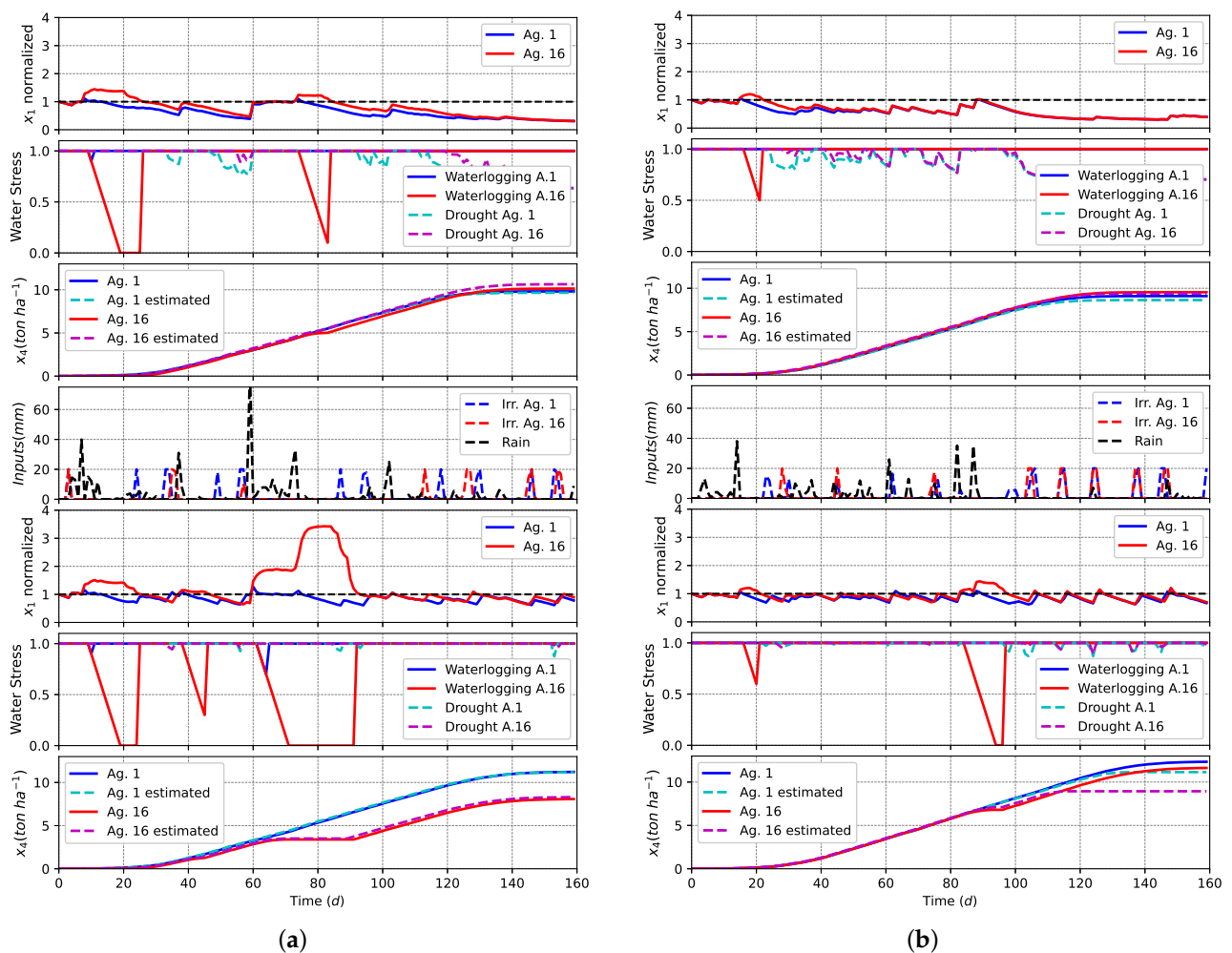
Figure 10. Water effect on agent 1 and 16 in a sloppy land (scenario 2) with (**a**) dataset 1 and (**b**) dataset 4. In each case, from top to bottom: soil water content normalized under no irrigation, water stresses under no irrigation, biomass under no irrigation, water inputs, soil water normalized content after irrigation, water stresses under irrigation, biomass under irrigation.

3.3. Third Scenario

In this scenario there are several interactions between agents (Figure 11). Agent 2 represents the highest point in the field and agent 10 corresponds to the lowest location.

The estimation of the state variables x_1 and x_3 with and without irrigation for the first meteorological dataset is shown in Figure 12. Again, it can be observed, that the water content of agent 10, which is at a low point and therefore receives more water, is underestimated under stronger precipitations. The estimation of x_3 with and without irrigation is subject to a slight offset. The reduction of x_3 by irrigation is related to the decrease in the drought stress and therefore an increase of crop transpiration.

Figure 13 shows the water effect on biomass production for datasets 1 and 4. For the first dataset it is noticeable that the gusts of rain generate a waterlogging stress even for the agents in the highest points of the terrain (e.g., agent 2). This effect is enhanced by irrigation. Moreover if the waterlogging stress is triggered in the growth stage, the biomass production is affected negatively. The positive effect of irrigation on the biomass production is about 30% for agent 2 and between 1.2 to 2.5% for agent 10 for datasets 1 and 4 (see Table 1).

Figure 11. Third scenario: (**a**) Possible water flux exchanges (**b**) Grid representation after image processing. The width of each square represents the normalized elevation of the portion of land with respect to the highest point in the grid. The maximum elevation considered is around 10 m. The color of every patch indicates the normalized amount of water at time k, the darker means high water content and the lighter indicate dryness.
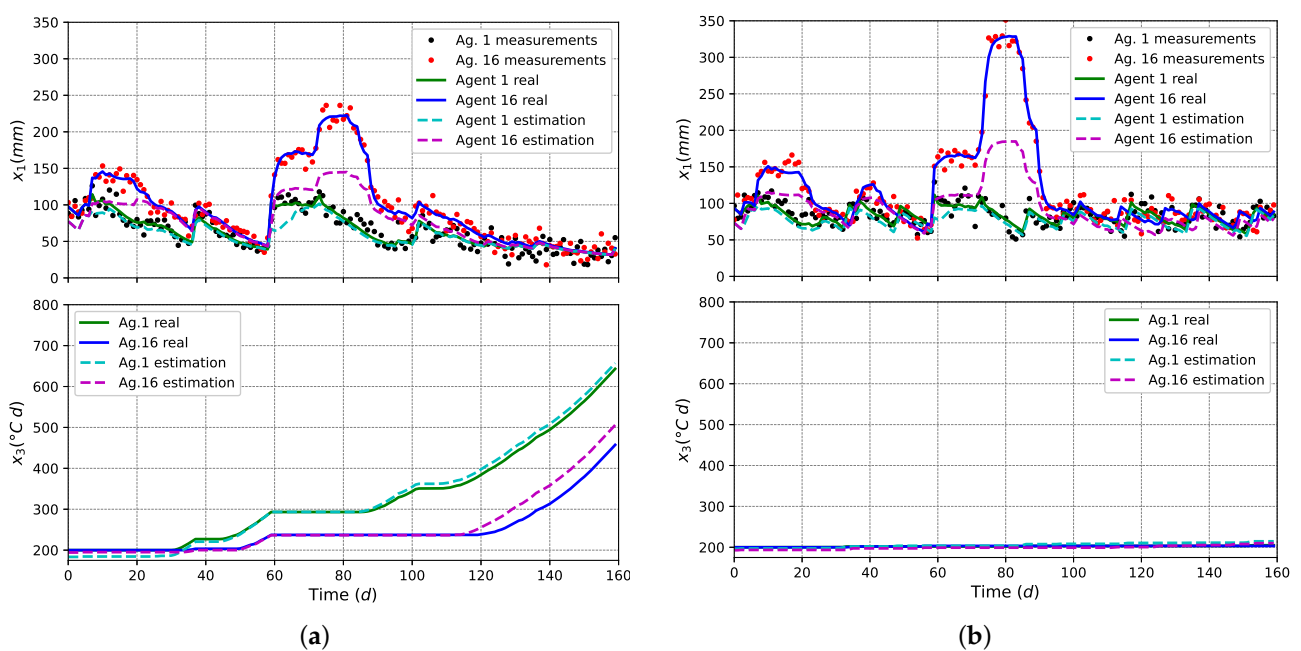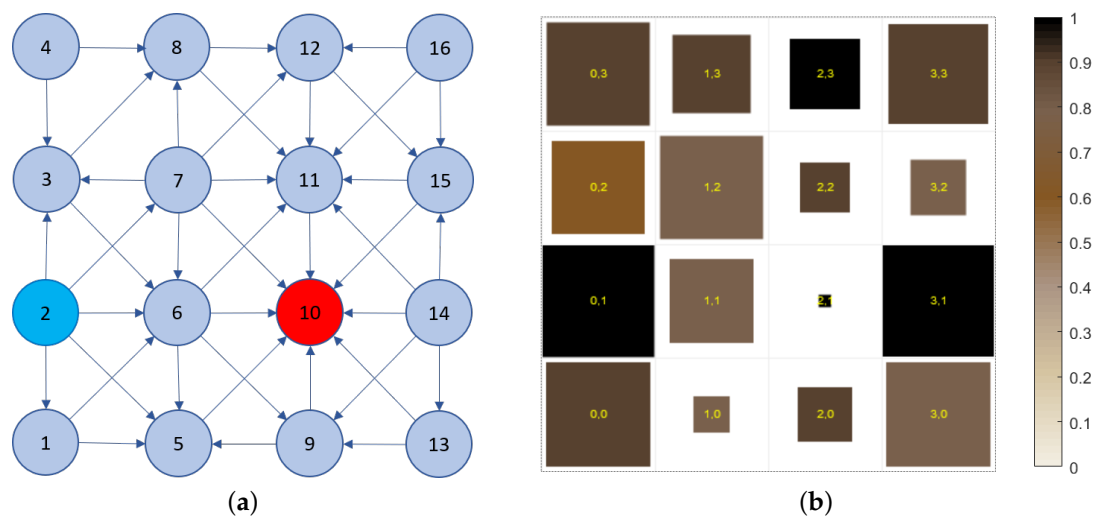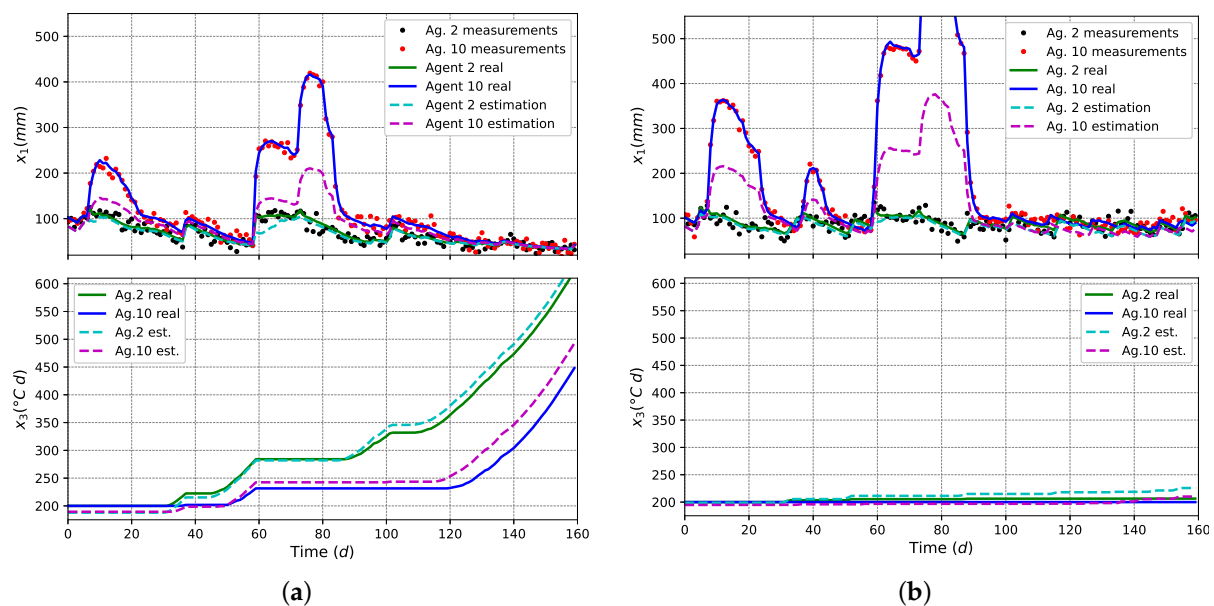

Figure 12. Rugged Land (scenario 3)—state variables x_1 and x_3 and their estimates for agents 2 and 10, under (**a**) no irrigation, and (**b**) irrigation policy for the first weather dataset.

In order to compare the performance of the irrigation policy in the various cases, a consume ratio is proposed as

$$CR = \frac{\sum\limits_{m=1}^{M} \sum\limits_{n=1}^{N} u^n}{M.N.W_{\max}},$$ (31)

where $W_{max} = 20$ mm, $N = 16$ and $M = 160$ (the harvest day is $t_f = 160$ d).

Table 2 presents a global view of the biomass production and water consumption for irrigation. The latter represents 6 to 11% of the maximum available water in all cases. The positive effect of irrigation is evidenced by a global rise of the production ranging between 10 to 62% depending on the case (land topology and weather). Obviously, the total biomass maximization does not necessarily increase the individual agent biomasses as

shown in Table 1, where for scenarios 2 and 3 the agents that are located at low points (i.e., agents 16 and 10) suffer from waterlogging instead of drought. The estimator is performing reasonably well, with usual errors in a range of 5%, but with transient discrepancies up to 20% for low points of uneven lands and strong precipitations.
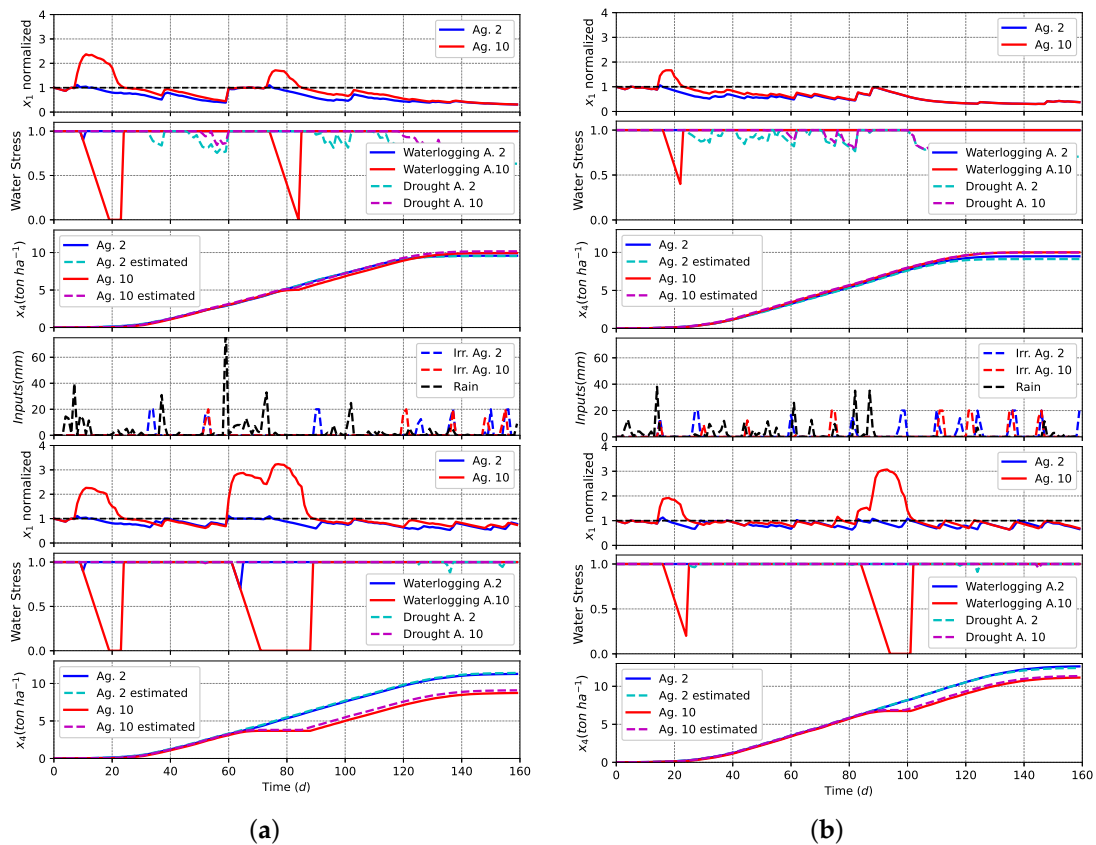

Figure 13. Water effect on agents 2 and 10 in a rugged land (third scenario) with (**a**) dataset 1 and (**b**) dataset 4. In each case, from top to bottom: soil water content normalized under no irrigation, water stresses under no irrigation, biomass under no irrigation, water inputs, soil water normalized content after irrigation, water stresses under irrigation, biomass under irrigation.

Table 2. Total biomass and irrigation for the whole cropping area (16 ha).

Scenario	Dataset	No Irrigation		Irrigation		Total Irrigation (mm)	Consume Ratio
		x_4 (ton)	$\hat{x}_4$ (ton)	x_4 (ton)	$\hat{x}_4$ (ton)		
1	1	145.0	142.2	180.9	181.5	3369	6.7%
	2	128.4	125.1	191.2	187.3	5704	11.3%
	3	113.6	107.6	181.5	181.6	4622	9.1%
	4	152.1	146.6	189.5	166.0	4165	8.2%
2	1	139.5	133.9	153.6	154.0	3587	7.1%
	2	130.0	126.9	186.9	177.0	5439	10.8%
	3	111.2	105.3	182.0	180.8	4416	8.7%
	4	159.8	157.6	190.2	177.3	3596	7.1%
3	1	137.8	138.0	165.4	168.6	3417	6.8%
	2	128.8	124.6	192.7	191.6	5439	10.8%
	3	108.4	102.9	182.0	180.7	4324	8.6%
	4	156.6	152.7	181.7	150.1	3034	6.0%

4. Conclusions and Perspectives

An agent-based model of an agricultural system is presented, which can be used to assess different irrigation schemes and production scenarios. The dynamic crop model underlying each agent is built based on existing models, but includes a novel waterlogging stress factor, proposed to consider the negative effect of water excess in soil. The ABM also includes an estimator of the unmeasured variables in the form of an extended Kalman filter, and a simple optimal irrigation strategy based on the average daily biomass. This framework is flexible and scalable, which allows the user to add new agents (or cropping areas) with diverse crops and different soil properties. Furthermore, crop-soil agents can be resized to catch the main soil changes across the cropping land based on soil properties and topographic variations. The ABM model is coded in the Python environment Mesa which is suitable for this purpose. However, there is the need to develop complementary modules to enhance the model functionality, e.g., GIS integration.

Future research entails the continued validation of the proposed model, for instance following the methodology of [38–41], including the use of global sensitivity analysis to assess the influence of the key parameters and possibly simplify the model structure in order to ease parameter estimation. Parameter identifiability is a critical aspect to develop a reliable prediction model, based on the collection of experimental data.

In addition, the consideration of other estimation schemes should be the subject of further investigation. Indeed, the current study was limited to a simple decentralized extended Kalman filter scheme, which shows poor performance for uneven grounds where water can accumulate at low points. Future developments therefore entail the use of centralized estimation schemes (i.e., encompassing all the agents), and the consideration of alternative estimation methods in order to avoid model linearisation and to handle non-Gaussian noise distributions as well as measurement correlations, which might be more appropriate in the extraction of information from image analysis.

Finally, the definition of the irrigation agent might be refined to describe the situation where irrigation is not available everywhere and/or not in the same supply. Different performance indexes might also be investigated, including weather forecast, in the framework of model predictive control. After validation and additional tests, an updated version of the ODD protocol [42] will be made available to the readers upon request.

Author Contributions: Conceptualization, N.Q. and A.V.W.; Methodology, N.Q. and A.V.W.; Investigation, J.L.-J., N.Q. and A.V.W.; Formal analysis, J.L.-J.; Software, J.L.-J.; Validation, J.L.-J., Funding acquisition, N.Q. and A.V.W.; Project administration, N.Q. and A.V.W.; Supervision, N.Q. and A.V.W.; Writing—original draft, J.L.-J.; Writing—review & editing, N.Q. and A.V.W. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study is available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Variables and Parameters

Table A1. Description of variables for every agent and units.

Name	Description	Units
x_1	Water content in soil	mm
x_2	Cumulative mean temperature	°C d
x_3	Cumulative temperature required from maturity to 50% of radiation interception during canopy senescence	°C d
x_4	Cumulative biomass	ton ha^{-1}
w_1	Daily rainfall	mm
w_2	Reference evapotranspiration	mm
w_3	Mean daily temperature	°C
w_4	Daily solar radiation	MJ m^{-2}
w_5	Daily minimum temperature	°C
w_6	Daily maximum temperature	°C
w_7	Atmospheric CO_2 concentration	ppm
f_1	Crop transpiration	mm d^{-1}
f_2	Surface runoff	mm d^{-1}
f_3	Deep drainage	mm d^{-1}
f_4	Incoming flux	mm d^{-1}
h_1	Daily mean temperature	°C
h_2	Heat stress	-
h_3	Drought stress	-
h_4	ARID index	-
h_5	Waterlogging stress	-
h_6	Combined stress on transpiration	-
h_7	Low temperature stress	-
h_8	CO_2 growth factor	-
g	Growth function	g m^{-2}

Table A2. Parameters and constants. The values in the fourth column related to management for the case study are adjusted to match the regional production values reported by the Colombian Ministry of Agriculture in its website (https://www.agronet.gov.co/estadistica/Paginas/home.aspx?cod=1).

Parameter	Description	Units	Value	Reference
θ_1	Water uptake coefficient (alpha)	-	0.096	[21]
θ_2	Wilting point	-	0.06	[21]
θ_3	Initial abstraction	mm	27.4	[21]
θ_4	Drainage coefficient (Beta)	-	0.55	[19]
θ_5	root-zone depth	mm	300–600	[37]
θ_6	Field capacity	-	0.21	[21]
θ_7	Base temperature for phenology development and growth	°C	0	[22]
θ_8	Optimal temperature for biomass growth	°C	15	[22]
θ_9	Threshold temperature to start accelerating senescence from heat stress	°C	34	[22]
θ_{10}	The extreme temperature threshold when RUE becomes 0 due to heat stress	°C	45	[22]
θ_{11}	The maximum daily reduction in x_3 due to heat stress	°C d	100	[22]
θ_{12}	The maximum daily reduction in x_3 due to drought stress	°C d	25	[22]
θ_{13}	Radiation use efficiency (above ground only and without respiration)	g MJ^{-1}	1.24	[22]
θ_{14}	Sensitivity of RUE to drought stress (ARID index).	-	0.4	[43]
θ_{15}	Stress time lower limit	d	3	[43]
θ_{16}	Stress time upper limit	d	13	[43]
θ_{17}	Relative increase in RUE per ppm elevated CO2 above 350 ppm.	-	0.08	[22]
θ_{18}	Cumulative temperature requirement from sowing to maturity.	°C d	2200	[22]
θ_{19}	Maximum fraction of radiation interception the crop can reach	-	0.95	[22]
θ_{20}	Cumulative temperature requirement for leaf area development to intercept 50% of radiation	°C d	480	[22]

Appendix B. Jacobian and EKF Tuning

The Jacobian is computed as follows

$$\Phi(k) = \begin{bmatrix} \frac{\partial \phi_1}{\partial x_1} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{\partial \phi_3}{\partial x_1} & 0 & 1 & 0 \\ \frac{\partial \phi_4}{\partial x_1} & \frac{\partial \phi_4}{\partial x_2} & \frac{\partial \phi_4}{\partial x_3} & 1 \end{bmatrix}\Bigg|_{x^n = \hat{x}^n(k)}, \tag{A1}$$

where

$$\frac{\partial \phi_1}{\partial x_1} = 1 - \frac{\partial f_1(k)}{\partial x_1(k)} - \frac{\partial f_3(k)}{\partial x_1(k)}, \tag{A2}$$

$$\frac{\partial f_1(k)}{\partial x_1(k)} = \begin{cases} \theta_1, & \theta_1(\hat{x}_1(k) - \theta_2\theta_5) \le w_2(k) \\ 0, & otherwise, \end{cases} \tag{A3}$$

$$\frac{\partial f_3(k)}{\partial x_1(k)} = \begin{cases} \theta_4, & \hat{x}_1(k) - w_1(k) - f_2(k) > \theta_5\theta_6 \\ 0, & otherwise, \end{cases} \tag{A4}$$

$$\frac{\partial \phi_3}{\partial x_1} = \begin{cases} -\frac{\theta_1\theta_{12}\theta_{14}}{w_2(k)}, & f_1(k) < w_2(k) \\ 0, & otherwise, \end{cases} \tag{A5}$$

$$\frac{\partial \phi_4}{\partial x_1} = \begin{cases} \theta_1\theta_{13}\theta_{14}h - 7(k)h_8(k)g(k)w_4(k), & f_1(k) < w_2(k) \\ 0, & otherwise, \end{cases} \tag{A6}$$

$$\frac{\partial \phi_4}{\partial x_2} = \begin{cases} \frac{0.01[e^{(-0.01(\hat{x}_2(k)-\theta_{20}))}]\theta_{19}}{[1+e^{(-0.01(\hat{x}_2-\theta_{20}))}]^2}, & \hat{x}_2(k) \le \frac{\theta_{18}}{2} \\ -\frac{0.01[e^{(-0.01(\hat{x}_2(k)+\hat{x}_3(k)-\theta_{18}))}]\theta_{19}}{[1+e^{(-0.01(\hat{x}_2+\hat{x}_3(k)-\theta_{18}))}]^2}, & \hat{x}_2(k) > \frac{\theta_{18}}{2}, \end{cases} \tag{A7}$$

$$\frac{\partial \phi_4}{\partial x_3} = \begin{cases} 0, & \hat{x}_2(k) \le \frac{\theta_{18}}{2} \\ -\frac{0.01[e^{(-0.01(\hat{x}_2(k)+\hat{x}_3(k)-\theta_{18}))}]\theta_{19}}{[1+e^{(-0.01(\hat{x}_2+\hat{x}_3(k)-\theta_{18}))}]^2}, & \hat{x}_2(k) > \frac{\theta_{18}}{2}, \end{cases} \tag{A8}$$

The selection of the parameters of the extended Kalman filter has not been fine tuned, and the following values are given as examples:

$$R = \begin{bmatrix} 0.16 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.1225 \end{bmatrix} \tag{A9}$$

Table A3. Suggested tuning for $Q(k)$, where $\gamma_n = \frac{altitude\ of\ agent\ n(m.a.s.l)}{altitude\ of\ the\ higuest\ point\ of\ the\ terrain\ (m.a.s.l)}$ and $N_{r,n}$ is the number of agents connected to the agent n with $\gamma_j < \gamma_n$ for $j = 1, ..., N$ and $n \neq j$. The physical units are R_{11} (mm²), R_{22} (°C²), R_{33} ((ton/ha)²), Q_{11} (mm²), Q_{22} (°C²), Q_{33} (°C²) and Q_{44} ((ton/ha)²).

Scenario	$Q(k)$	α
1	$0.1I$	1
2	$0.4I$	$0.1(1 + N_{r,n})$
3	$0.7I$	$0.05(1 - \gamma_n)$

Appendix C. ABM Algorithms

Algorithm A1: Crop-soil agent Python structure.

    `#Import modules from mesa`
    import: Agent
    `#Import other modules`
    import: numpy, random
    `#Define agent class`

1   **class CropSoil(Agent):**
2   spatial inputs $\leftarrow$ data from images
3        $w_j \leftarrow$ meteorological data
4   Initialize attributes (position, id, *self_time*, and Θ^n)
5   `#Define methods`
6   **def get:** irrigation   # *Method to collect data from the irrigation agent*
7   **def get:** neighborhood   # *Method to retrieve neighbors from the graph*
8   **def get:** measurements   # *Method to collect data from sensors*
9   collect measurements by Equation (22)
10   **def get:** Jacobian
11   compute the jacobian of $\phi(k)$ by Equation (24)
12   **def get:** Extended Kalman filter   # *Method to collect the estimation of states*
13   project $\hat{x}^-(k)$ by Equation (25)
14   propagate $P^-(k)$ by Equation (26) with $Q_m(k)$ as in Equation (30)
15   compute the Kalman gain matrix by Equation (27)
16   state estimate update by Equation (28)
17   error covariance update by Equation (29)
18   **def step** (u^n, w_j):
19   update neighborhood (get: neighborhood)
20   compute x_1 by Equation (2) from Equations 6,7, and 8
21   compute f_4 from neighbors' data (x_1 and elevation) by Equation (9)
22   compute x_2 by Equation (3) from Equation (10)
23   `#compute stress factors`
24   Heat stress by Equation (11)   # *High temperature*
25   Temperature stress by Equation (17)   # *Low temperature*
26   Drought stress by Equation (12) from Equation (13)   # *Lack of water*
27   Waterlogging stress by Equation (14) from Equation (15) # *Excess of water*
28   compute x_3 by Equation (4)
29   compute CO_2 effect by Equation (18)
30   compute x_4 by Equation (5) from Equations 16 and 19
31   get measurements (get: measurements)
32   compute $\Phi(\cdot)$ (get: Jacobian)
33   execute EKF (get: Extended Kalman filter)
34   get irrigation (get: irrigation)
35          $k = k + 1$

Algorithm A2: Model class structure.

```
    #Import modules from mesa
    import: Model, DataCollector, space (Grid), and time (RandomActivation)
    #Import other modules
    import: numpy, random, scipy
    #Define model class
1   class CropSoilModel(Model) :
2       Load data from the field (image processing info)
3       Initialize attributes (simulation_time, N)
4       #Set up model objects
5       Schedule = RandomActivation
6       Space = Grid
7       Define variables to collect (x_1, x_2, x_3, x_4, u)
8       #Create agents
9       From input data create N agents on their corresponding site in the grid
10      #Define methods
11      def get: state variables
12          collect data from sensors
13          execute the estimator
14      def get: Irrigation agent
15          compute Equation (20)
16      def step (self):
17          execute schedule step
18          execute irrigation agent
19          collect data (get: state variables)
20          simulation_time = simulation_time + 1
```

References

1. Godfray, H.C.J.; Beddington, J.R.; Crute, I.R.; Haddad, L.; Lawrence, D.; Muir, J.F.; Pretty, J.; Robinson, S.; Thomas, S.M.; Toulmin, C. Food Security: The Challenge of Feeding 9 Billion People. *Science* **2010**, *327*, 812–818. [CrossRef] [PubMed]

2. Rosegrant, M.W.; Ringler, C.; Zhu, T. Water for Agriculture: Maintaining Food Security under Growing Scarcity. *Annu. Rev. Environ. Resour.* **2009**, *34*, 205–222. [CrossRef]

3. Jones, J.W.; Antle, J.M.; Basso, B.; Boote, K.J.; Conant, R.T.; Foster, I.; Godfray, H.C.J.; Herrero, M.; Howitt, R.E.; Janssen, S.; et al. Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. *Agric. Syst.* **2017**, *155*, 269–288. [CrossRef] [PubMed]

4. Jin, X.; Kumar, L.; Li, Z.; Feng, H.; Xu, X.; Yang, G.; Wang, J. A review of data assimilation of remote sensing and crop models. *Eur. J. Agron.* **2018**, *92*, 141–152. [CrossRef]

5. Huang, J.; Gómez-Dans, J.L.; Huang, H.; Ma, H.; Wu, Q.; Lewis, P.E.; Liang, S.; Chen, Z.; Xue, J.H.; Wu, Y.; et al. Assimilation of remote sensing into crop growth models: Current status and perspectives. *Agric. For. Meteorol.* **2019**, *276–277*, 107609. [CrossRef]

6. Rosenzweig, C.; Jones, J.W.; Hatfield, J.L.; Ruane, A.C.; Boote, K.J.; Thorburn, P.; Antle, J.M.; Nelson, G.C.; Porter, C.; Janssen, S.; et al. The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. *Agric. For. Meteorol.* **2013**, *170*, 166–182. [CrossRef]

7. Martre, P.; Wallach, D.; Asseng, S.; Ewert, F.; Jones, J.W.; Rötter, R.P.; Boote, K.J.; Ruane, A.C.; Thorburn, P.J.; Cammarano, D.; et al. Multimodel ensembles of wheat growth: many models are better than one. *Glob. Chang. Biol.* **2015**, *21*, 911–925. [CrossRef]

8. Steduto, P.; Hsiao, T.C.; Raes, D.; Fereres, E. AquaCrop—The FAO Crop Model to Simulate Yield Response to Water: I. Concepts and Underlying Principles. *Agron. J.* **2009**, *101*, 426–437. [CrossRef]

9. Brisson, N.; Gary, C.; Justes, E.; Roche, R.; Mary, B.; Ripoche, D.; Zimmer, D.; Sierra, J.; Bertuzzi, P.; Burger, P.; et al. An overview of the crop model stics. *Eur. J. Agron.* **2003**, *18*, 309–332. [CrossRef]

10. Holzworth, D.P.; Huth, N.I.; deVoil, P.G.; Zurcher, E.J.; Herrmann, N.I.; McLean, G.; Chenu, K.; van Oosterom, E.J.; Snow, V.; Murphy, C.; et al. APSIM—Evolution towards a new generation of agricultural systems simulation. *Environ. Model. Softw.* **2014**, *62*, 327–350. [CrossRef]

11. Schreinemachers, P.; Berger, T. An agent-based simulation model of human–environment interactions in agricultural systems. *Environ. Model. Softw.* **2011**, *26*, 845–859. [CrossRef]

12. Murray-Rust, D.; Robinson, D.T.; Guillem, E.; Karali, E.; Rounsevell, M. An open framework for agent based modelling of agricultural land use change. *Environ. Model. Softw.* **2014**, *61*, 19–38. [CrossRef]

13. Mewes, B.; Schumann, A.H. IPA (v1): a framework for agent-based modelling of soil water movement. *Geosci. Model Dev.* **2018**, *11*, 2175. [CrossRef]

14. Thorp, K.R.; DeJonge, K.C.; Kaleita, A.L.; Batchelor, W.D.; Paz, J.O. Methodology for the use of DSSAT models for precision agriculture decision support. *Comput. Electron. Agric.* **2008**, *64*, 276–285. [CrossRef]

15. Siad, S.M.; Iacobellis, V.; Zdruli, P.; Gioia, A.; Stavi, I.; Hoogenboom, G. A review of coupled hydrologic and crop growth models. *Agric. Water Manag.* **2019**, *224*, 105746. [CrossRef]

16. Siegfried, R. *Modeling and Simulation of Complex Systems: A Framework for Efficient Agent-Based Modeling and Simulation*; Springer: Riemerling, Germany, 2014; p. 233. [CrossRef]

17. Macal, C.M.; North, M.J. Tutorial on agent-based modeling and simulation. In Proceedings of the Winter Simulation Conference, Orlando, FL, USA, 4 December 2005; p. 14. [CrossRef]

18. Macal, C.M. Everything you need to know about agent-based modelling and simulation. *J. Simul.* **2016**, *10*, 144–156. [CrossRef]

19. Ritchie, J.T. A User-Orientated Model of the Soil Water Balance in Wheat. In *Wheat Growth and Modelling*; Springer: Boston, MA, USA, 1985; pp. 293–305.

20. Jiang, Z.; Chen, Z.; Zhou, Q.; Ren, J. Global sensitivity analysis of CERES-Wheat model parameters. *Trans. Chin. Soc. Agric. Eng.* **2011**, *27*, 236–242.

21. Woli, P.; Jones, J.W.; Ingram, K.T.; Fraisse, C.W. Agricultural Reference Index for Drought (ARID). *Agron. J.* **2012**, *104*, 287–300. [CrossRef]

22. Zhao, C.; Liu, B.; Xiao, L.; Hoogenboom, G.; Boote, K.J.; Kassie, B.T.; Pavan, W.; Shelia, V.; Kim, K.S.; Hernandez-Ochoa, I.M.; et al. A SIMPLE crop model. *Eur. J. Agron.* **2019**, *104*, 97–106. [CrossRef]

23. Wallach, D.; Makowski, D.; Jones, J.W.; Brun, F. *Working with Dynamic Crop Models: Methods, Tools and Examples for Agriculture and Environment*; Academic Press: Cambridge, MA, USA, 2018.

24. Parent, C.; Capelli, N.; Berger, A.; Crèvecoeur, M.; Dat, J.F. An overview of plant responses to soil waterlogging. *Plant Stress* **2008**, *2*, 20–27.

25. Guang, C.; Xiugui, W.; Yu, L.; Wenbing, L. Effect of Water Logging Stress on Cotton Leaf Area Index and Yield. *Procedia Eng.* **2012**, *28*, 202–209. [CrossRef]

26. Neale, C.M.U.; Geli, H.M.E.; Kustas, W.P.; Alfieri, J.G.; Gowda, P.H.; Evett, S.R.; Prueger, J.H.; Hipps, L.E.; Dulaney, W.P.; Chávez, J.L.; et al. Soil water content estimation using a remote sensing based hybrid evapotranspiration modeling approach. *Adv. Water Resour.* **2012**, *50*, 152–161. [CrossRef]

27. Khanal, S.; Fulton, J.; Shearer, S. An overview of current and potential applications of thermal remote sensing in precision agriculture. *Comput. Electron. Agric.* **2017**, *139*, 22–32. [CrossRef]

28. Campos, I.; González-Gómez, L.; Villodre, J.; González-Piqueras, J.; Suyker, A.E.; Calera, A. Remote sensing-based crop biomass with water or light-driven crop growth models in wheat commercial fields. *Field Crops Res.* **2018**, *216*, 175–188. [CrossRef]

29. Haykin, S. *Kalman Filtering and Neural Networks*; John Wiley & Sons: Hoboken, NJ, USA, 2004; Volume 47.

30. Reif, K.; Gunther, S.; Yaz, E.; Unbehauen, R. Stochastic stability of the discrete-time extended Kalman filter. *IEEE Trans. Autom. Control* **1999**, *44*, 714–728. [CrossRef]

31. Dorafshan, S.; Thomas, R.J.; Maguire, M. Benchmarking image processing algorithms for unmanned aerial system-assisted crack detection in concrete structures. *Infrastructures* **2019**, *4*, 19. [CrossRef]

32. Li, X.; McCarty, G.W. Application of topographic analyses for mapping spatial patterns of soil properties. In *Geospatial Analyses of Earth Observation (EO) Data*; IntechOpen: London, UK, 2019. [CrossRef]

33. Hutchinson, M.F.; Gallant, J.C. Digital elevation models and representation of terrain shape. In *Terrain Analysis: Principles and Applications*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2000; pp. 29–50.

34. Dandois, J.P.; Ellis, E.C. Remote sensing of vegetation structure using computer vision. *Remote Sens.* **2010**, *2*, 1157–1176. [CrossRef]

35. Abar, S.; Theodoropoulos, G.K.; Lemarinier, P.; O'Hare, G.M.P. Agent Based Modelling and Simulation tools: A review of the state-of-art software. *Comput. Sci. Rev.* **2017**, *24*, 13–33. [CrossRef]

36. Masad, D.; Kazil, J. MESA: An agent-based modeling framework. In Proceedings of the 14th PYTHON in Science Conference, Austin, TX, USA, 6–12 July 2015; pp. 53–60.

37. Woli, P.; Jones, J.W.; Ingram, K.T. Assessing the Agricultural Reference Index for Drought (ARID) Using Uncertainty and Sensitivity Analyses. *Agron. J.* **2013**, *105*, 150–160. [CrossRef]

38. Pullum, L.L.; Cui, X. *Techniques and Issues in Agent-Based Modeling Validation*; Technique Report; Oak Ridge National Laboratory (ORNL): Oak Ridge, TN, USA, 2012; pp. 5–6.

39. Lee, J.S.; Filatova, T.; Ligmann-Zielinska, A.; Hassani-Mahmooei, B.; Stonedahl, F.; Lorscheid, I.; Voinov, A.; Polhill, G.; Sun, Z.; Parker, D.C. The complexities of agent-based modeling output analysis. *J. Artif. Soc. Soc. Simul.* **2015**, *18*. [CrossRef]

40. Kang, J.Y.; Aldstadt, J. Using multiple scale spatio-temporal patterns for validating spatially explicit agent-based models. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 193–213. [CrossRef]

41. Grimm, V.; Revilla, E.; Berger, U.; Jeltsch, F.; Mooij, W.M.; Railsback, S.F.; Thulke, H.H.; Weiner, J.; Wiegand, T.; DeAngelis, D.L. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science* **2005**, *310*, 987–991. [CrossRef] [PubMed]

42. Grimm, V.; Railsback, S.F.; Vincenot, C.E.; Berger, U.; Gallagher, C.; DeAngelis, D.L.; Edmonds, B.; Ge, J.; Giske, J.; Groeneveld, J.; et al. The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism. *J. Artif. Soc. Soc. Simul.* **2020**, *23*. [CrossRef]

43. Li, P.; Chen, J.; Wu, P. Agronomic Characteristics and Grain Yield of 30 Spring Wheat Genotypes under Drought Stress and Nonstress Conditions. *Agron. J.* **2011**, *103*, 1619–1628. [CrossRef]