



## Data-Driven Modelling of Soil Moisture Dynamics for Smart Irrigation Scheduling

Erion Bwambale <sup>a,b,c,\*</sup>, Felix K. Abagale <sup>a,b</sup>, Geophrey K. Anornu <sup>d</sup>

<sup>a</sup> West African Center for Water, Irrigation and Sustainable Agriculture (WACWISA), University for Development Studies, P. O. Box TL 1882, Tamale, Ghana

<sup>b</sup> Department of Agricultural Engineering, University for Development Studies, P. O. Box TL 1882, Tamale, Ghana

<sup>c</sup> Department of Agricultural and Biosystems Engineering, Makerere University, P. O. Box 7062, Kampala, Uganda

<sup>d</sup> Civil Engineering Department, Regional Water and Environmental Sanitation Center Kumasi (RWESCK), Kwame Nkrumah University of Sciences and Technology, Kumasi, Ghana



### ARTICLE INFO

Editor: Stephen Symons

**Keywords:**

Soil moisture dynamics  
State space modelling  
System identification  
Model predictive control

### ABSTRACT

In the face of increasing water scarcity and uncertainties of climate change, improving crop water use efficiency and productivity, while minimizing negative environmental impacts, is becoming crucial to meet the surging global food demand. Smart irrigation has a potential of improving water use efficiency in precision agriculture especially when efficient irrigation control strategies are adopted. Conventionally, irrigation systems rely on heuristic methods to schedule irrigation which either leads to over-irrigation or under-irrigation. This influences the crop physiological characteristics as well as the water use efficiency. To tackle this menace, model-based irrigation management has been overemphasized. A closed-loop irrigation control strategy relies on a mathematical model of the system for irrigation scheduling decisions. In this study, a data-driven approach was used to learn soil moisture dynamics from a drip irrigated tomato in an open field agricultural system. A total number of 9674 data samples were collected using an ATMOS41 weather station, TERROS 12 soil moisture sensor and a YFS-201 flow sensor for crop evapotranspiration and precipitation, soil moisture and irrigation volumes respectively. Data driven modelling was then performed using the system identification toolbox in a MATLAB environment. The model formulation was a multi-input single-output (MISO) system with reference evapotranspiration, irrigation and rainfall as inputs and soil moisture as the output. Different model structures including transfer functions, state space models, polynomial models and ARX models were evaluated. Model performance was evaluated using the mean square error (MSE), final prediction error (FPE) and estimated fit of the model approaches. Simulation results indicate that the soil moisture dynamics model provides a satisfactory approximation of the process dynamics with a state space model giving an estimated fit of 97.04 %, MSE and FPE of  $1.74 \times 10^{-7}$  and  $1.75 \times 10^{-7}$  respectively. This model will be used to design a model predictive controller for smart irrigation scheduling in open field environmental conditions.

### 1. Introduction

Water scarcity at a global scale is increasing, making it perplexing to achieve sustainable agricultural production [1–3]. Agricultural water scarcity is projected to increase in more than 80 % of the world's countries by the year 2050 [4]. This coupled with the increasing population, climate change impacts and competition from other sectors, puts food security at stake [5]. Irrigated agriculture is one of the major water users accounting for about 70 % of freshwater withdrawals [6],[7]. Irrigation accounts for 40 % of the world's food and fibre production and will therefore play a crucial role in feeding the 9.7 billion people by

2050 [8],[9]. Water withdrawal for irrigation is projected to hit 29000 km<sup>3</sup> by 2050 with most of the high demand in developing countries translating into 20 million hectares under irrigation [10]. The need to enhance water use efficiency has never been over emphasized throughout the research community and industry as it is presently.

Enhancing water use efficiency in irrigated agriculture necessitates effective use of available water resources to respond to real-time crop irrigation requirements. This requires monitoring water use and controlling how the irrigation water is applied. The goal of monitoring the crop environment in real-time is to control the root zone soil moisture between specific thresholds; field capacity and permanent wilting point

\* Corresponding author.

E-mail address: [erionbw209@uds.edu.gh](mailto:erionbw209@uds.edu.gh) (E. Bwambale).

[11]. On the other hand, irrigation control deals with applying what will keep the soil moisture between that threshold at the set time with precision.

Emerging technologies, such as wireless communications, remote sensing, machine learning, crop physiology and soil environment monitoring systems, the Internet of Things (IoT) and big data, have expounded the scope of irrigation monitoring and control [12–16]. These technologies facilitate the generation of large amounts of data which when processed and analyzed can be utilized for predictions. This has given rise to the paradigm shift from open loop irrigation control to model-based irrigation management [17,18]. In model-based irrigation management, data from soil, weather and plant makes an integral part of the predictive control strategy to improve crop water use efficiency while minimizing negative environmental consequences [19–22].

Over the years, the soil agro-hydrological model has been utilized to depict the soil moisture dynamics in the rootzone [23] as a basis for model based irrigation scheduling. The model characterizes the hydrological cycle between the soil, the atmosphere, and the crop [24]. Modelling soil moisture dynamics is very critical in irrigation automation. Unlike conventional irrigation systems that rely on open-loop irrigation strategies, closed-loop irrigation control for smart irrigation uses a plant model and the system's current state to predict future irrigation amounts [25–27]. Nonetheless, these models have been developed for controlled environment agricultural systems especially in developed countries like USA, China and Australia. About, 95 % of food consumed in developing countries is from open field agricultural systems where most of the water losses occur. Therefore, developing models on which irrigation scheduling decisions can be based in open field agriculture is of significant importance.

Several researchers have modelled soil moisture dynamics in an attempt to improve irrigation scheduling. For example, Saleem et al. [28] and Lozoya et al. [29] developed a soil moisture dynamics model for closed-loop irrigation. The authors assumed negligible precipitation and ignored the crop coefficient in the evapotranspiration model. The drawback then becomes its suitability in different agroclimatic zones where supplementary irrigation is necessary. Additionally, crops need different volumes of water per growth stage. Similarly, Mao et al. [30] used the agro-hydrological model to predict the future states of an agricultural system. The authors, however, rely on historical data for irrigation scheduling. This means real-time plant water needs are dependent on past data instead of real-time soil, weather, and plant system parameters. Recently, Abioye et al. [31] developed an Internet of Things (IoT) and data-driven modelling to model soil moisture dynamics in a controlled environment agricultural system. The authors developed a multi-input single-output (MISO) system with evapotranspiration and irrigation volumes as inputs and soil moisture as the output. Open field agricultural systems, however, depict a more dynamic nature of soil moisture with many factors contributing to changes in soil moisture compared to controlled environment agricultural systems. There is considerably a small body of studies that modelled soil moisture dynamics in open field agricultural systems using data-driven approaches. This is partly attributed to the costly monitoring equipment required to provide the data for all the parameters in the water balance equation. The major objective of this study is to therefore develop a soil moisture dynamics model in an open field agricultural system to aid model-based irrigation management for enhancement of water use efficiency in arid and semi-arid regions.

## 2. Materials and Methods

### 2.1. Study Area

The experiment was conducted at the University for Development Studies, Nyankpala campus at WACWISA experimental site located at  $9^{\circ}24'37''$  N and  $0^{\circ}58'53''$  E of the Equator as shown in Fig. 1. A drip irrigation experiment was designed on a 20\*18m plot under tomato

production spaced at 0.5m. The drip irrigation system consisted of 20mm HDPE laterals with online emitters with a flow rate of 2.1litres per hour. The experimental data was collected from September 21<sup>st</sup> to November 26, 2022, resulting in 9674 data sets for developing the soil moisture dynamics model collected in 10-mibute time-steps. Fig. 2 shows the experimental plot at WACWISA experimental field, University for Development Studies in Northern Ghana.

### 2.2. Agro-hydrological balance model

This model describes the soil moisture dynamics in an open field agricultural system. The agro-hydrologic balance model states that changes in rootzone soil moisture at a particular point in time is due to water inflows including rainfall, capillary rise, irrigation less the water outflows, deep percolation, water runoff and evapotranspiration as depicted in Fig. 3 [27]. The agro-hydrological model can be mathematically expressed as in Eq. (1) [28,32].

$$\dot{\theta}(t) = P(t) + C_r(t) + I(t) + L_i(t) - D_p(t) - R_0(t) - ET_c(t) - L_o(t) \quad (1)$$

Where;

$\dot{\theta}(t)$  is soil moisture storage,  $P(t)$  is effective rainfall,  $C_r(t)$  is capillary rise,  $I(t)$  is irrigation,  $L_i(t)$  is lateral inflow,  $D_p(t)$  is deep percolation,  $R_0(t)$  is runoff,  $ET_c(t)$  is crop evapotranspiration, and  $L_o(t)$  is lateral outflow

In most agricultural fields the lateral water inflow is negligible, or the total effect is always zero. Likewise, the capillary rise contribution in soils with a deep-water table is zero. The runoff in a well-designed smart irrigation system is taken as zero, and thus the change in rootzone soil moisture becomes as shown in Eq. (2).

$$\dot{\theta}(t) = P(t) + I(t) - D_p(t) - ET_c(t) \quad (2)$$

#### 2.2.1. Evapotranspiration

The crop evapotranspiration then is a product of reference evapotranspiration and the crop coefficient as depicted in Eq. (3) [33]. Allen et al. [33] defines the crop evapotranspiration under standard conditions, denoted as  $ET_c$ , as the evapotranspiration from disease-free, well-fertilized crops, grown in large fields, under optimum soil water conditions, and achieving full production under the given climatic conditions. The FAO Penman Monteith model for reference evapotranspiration was adopted for this study.

$$ET_c = ET_0 * K_c \quad (3)$$

Where;

$ET_c$  is the crop evapotranspiration( $\text{mm day}^{-1}$ ),  $ET_0$  is the reference evapotranspiration( $\text{mm day}^{-1}$ ) and  $K_c$  is the crop factor which depends on the growth stage of the plant.

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273}U_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34U_2)} \quad (4)$$

Where;

$ET_0$  is Reference evapotranspiration ( $\text{mm day}^{-1}$ ),  $\Delta$  is the slope vapour pressure curve ( $\text{kPa }^{\circ}\text{C}^{-1}$ ),  $R_n$  is net radiation at the crop surface ( $\text{MJ m}^{-2}\text{day}^{-1}$ ),  $G$  is Soil heat flux density ( $\text{MJ m}^{-2}\text{day}^{-1}$ ),  $\gamma$  is Psychometric constant ( $\text{kPa }^{\circ}\text{C}^{-1}$ ),  $T$  is Air temperature ( $^{\circ}\text{C}$ ),  $U_2$  is Wind speed ( $\text{ms}^{-1}$ ),  $e_s$  is Saturation vapour pressure (kpa),  $e_a$  is Actual vapour pressure (kpa)

An ATMOS 41 weather station was used to monitor the plant environment in real-time and the data was sent to a wireless Zentrac data logger and stored in the cloud. The automated weather station gives solar radiation, air temperature, vapour pressure, vapour pressure deficit, precipitation, and wind speed among other parameters. Net radiation at the crop surface  $R_n$ , Soil heat flux density  $G$ , slope vapour pressure curve and psychometric constant were derived from measured parameters using procedures described in Allen et al. [33].

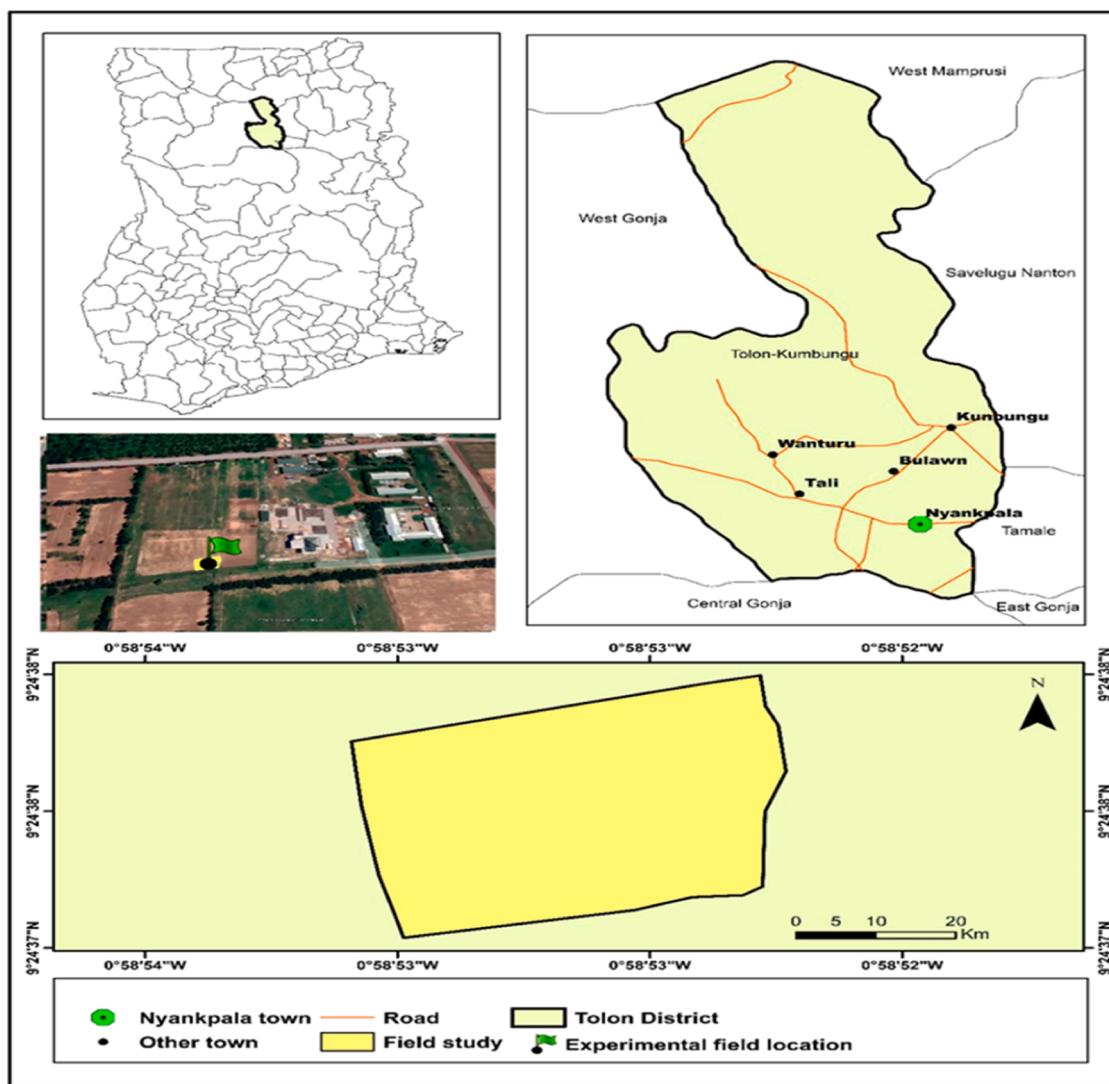


Fig. 1. Study Area.

Weather data used in the calculation of reference evapotranspiration was collected at 10-minute intervals and stored on the Zetracloud. Using the FAO Penman Monteith Equation, a 10-minute reference evapotranspiration was determined for the 9674 samples.

#### 2.2.2. Soil Moisture

Soil moisture measurements were conducted by using a Terros 12 volumetric water content sensor from meter group. The sensor was located at a depth of 30 cm in order to measure the soil moisture at the effective rootzone depth of tomato crop. The crop grows over a sandy loam type of soil with a field capacity and permanent wilting point of 24.5 and 12.6, respectively and an infiltration rate of 12 l/h.

#### 2.2.3. Irrigation

Similarly, in order to estimate the irrigation volume, the IoT based flowmeter produces a digital output pulse which has been counted by the IoT Expresso board to calculate the water flow rate during irrigation. The water flowrate is calibrated in Eq. (5) and flow volume in Eq. (6).

$$R = \frac{N * 60(\text{Pulse per minute})}{M(\text{Pulse per litre})} \quad (5)$$

$$\text{Irrigation volume} = \sum \text{Flow rate} * \text{Time - Interval} \quad (6)$$

#### 2.3. Model Structure

##### 2.3.1. Transfer functions

For a dynamical system that does not change over time, the transfer function provides a useful representation. [34]. Inspection or basic algebraic manipulations of the differential equations that characterize the systems can yield the transfer function. Even infinite-dimensional systems driven by partial differential equations may be understood with the help of transfer functions. Experiments on a system can be used to establish its transfer function. Eq. 7 depicts the mathematical formulation of a transfer function [35].

$$G(s) = C[sI - A]^{-1} - B + D \quad (7)$$

Where

$A, B, C$  and  $D$  are matrices

##### 2.3.2. The Box Jenkins Model

The Box Jenkins model uses a rational polynomial function to give totally separate parameterization for the dynamics and noise. Eqs. 8a and b represent the Box Jenkins model [36].

$$y(t) = x^{-d} \frac{B(x^{-1})}{F(x^{-1})} u(t) + \frac{C(x^{-1})}{D(x^{-1})} e(t) \quad (8a)$$

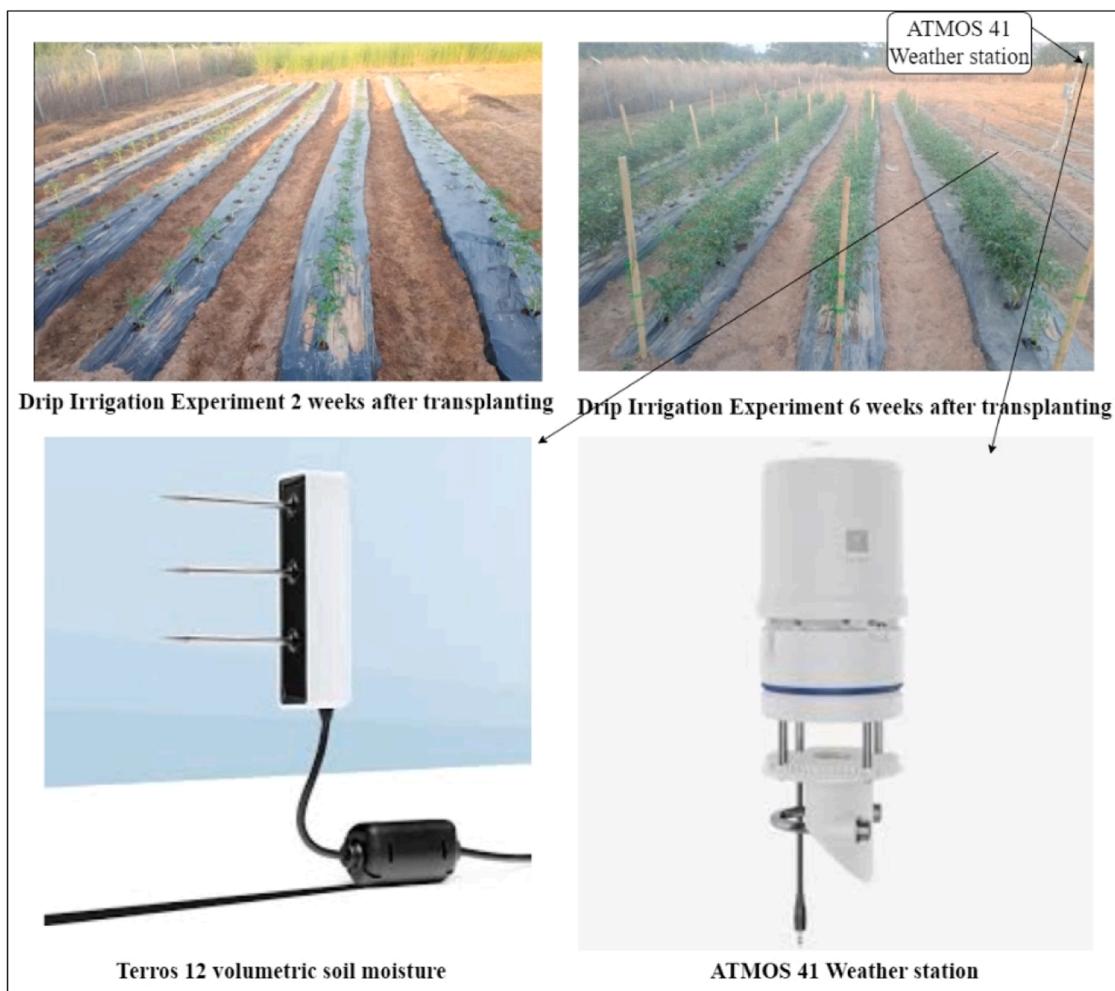


Fig. 2. Layout of drip irrigation experiment with monitoring sensors.

Where

$$D(x^{-1}) = 1 + \sum_{k=1}^{n_a} D x^{-k} \quad (8b)$$

Water content in volume at time t is denoted by  $y(t)$ , input variables are denoted by  $u(t)$ , and estimate error is denoted by  $e(t)$ .  $A(x)$ ,  $B(x)$ ,  $C$

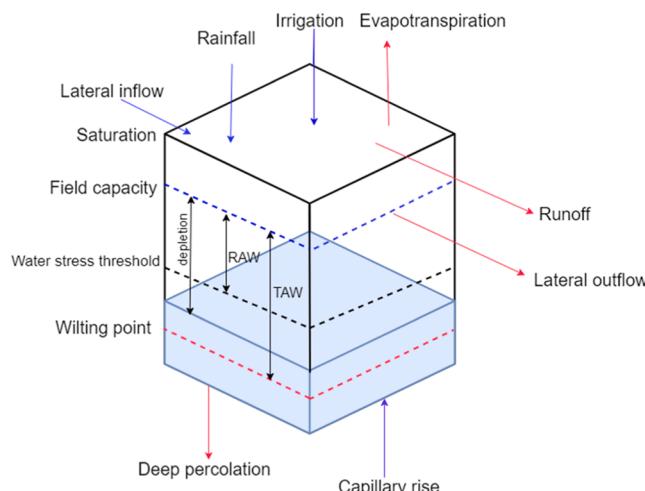


Fig. 3. Agro-hydrological system.

$(x)$ ,  $D(x)$ , and  $F(x)$  are the polynomials that determine the model's output, inputs, and estimate errors, respectively.

### 2.3.3. The State Space Model

In statistics, a state-space model is a kind of probabilistic graphical model that graphically expresses the probabilistic dependency between a latent state variable and an observed measurement. [37]. It represents the dynamics of an  $N^{\text{th}}$  order system as a first-order differential equation in an  $N$ -vector, which is called the state [38]. The important attribute of the state vector  $x(t)$  in the state space formulation is that it entirely specifies the system at a time  $t$  [39]. Future states are determined solely by the current state,  $x(t)$ , and any inputs,  $u(t)$ , at time  $t$  and beyond. The current state  $x(t)$  summarizes all previous states and the full input history; it thus encompasses all of the system's memory. Eqs. 9a and 9b can be used to represent state space models.

$$x(t + Ts) = A x(t) + B u(t) + ke(t) \quad (9a)$$

$$y(t) = C x(t) + D u(t) + e(t) \quad (9b)$$

Where,

$X \in R^N$  is the state vector at time  $n$ ,  $u$  is a  $p \times 1$  vector of inputs,  $A$  is an  $n \times n$  state transition matrix,  $B$  is an  $n \times r$  input coefficient matrix,  $y$  is a column vector of the output variables  $y(t)$ ,  $C$  is an  $m \times n$  matrix of the constant-coefficient  $c_{ij}$  that weight of state variables and  $D$  is an  $m \times n$  matrix of the constant-coefficient  $d_{ij}$  weight of the system inputs.

### 2.3.4. The Output Error Model

When only white measurement noise is affecting the process output, the Output Error model structure is used as represented in Eqs. 10a and b [40].

$$G(x^{-1}) = x^{-d} \frac{B(x^{-1})}{F(x^{-1})} \text{ and } H(x^{-1}) = 1 \quad (10a)$$

This model can be expressed as

$$y(t) = x^{-d} \frac{B(x^{-1})}{F(x^{-1})} u(t) + e(t) \quad (10b)$$

### 2.3.5. The Autoregressive with External Input Model (ARX) Model

The ARX model is derived from Eq. 11 [41].

$$G(x^{-1}) = x^{-d} \frac{B(x^{-1})}{A(x^{-1})} \text{ and } H(x^{-1}) = \frac{1}{A(x^{-1})} \quad (11)$$

Where  $A(x^{-1}) = 1 + \sum_{k=1}^{n_a} a_k x^{-k}$  and  $B(x^{-1}) = \sum_{k=0}^{n_b} b_k x^{-k}$ . These are polynomial matrices; d is the delay of the system. The ARX model is therefore simplified using equation 12 as shown.

$$A(x^{-1})y(t) = x^{-d}B(x^{-1})u(t) + C(x^{-1})e(t) \quad (12)$$

### 2.3.6. The Autoregressive Moving Average with Extra Input Model

This autoregressive moving average with external input model (ARMAX) is in broader structure compared to the ARX as shown in Eq. 13.

$$G(x^{-1}) = x^{-d} \frac{B(x^{-1})}{A(x^{-1})} \text{ and } H(x^{-1}) = \frac{C(x^{-1})}{A(x^{-1})} \quad (13)$$

Where  $C(x^{-1}) = 1 + \sum_{k=1}^{n_c} c_k x^{-k}$ . This is a polynomial where the noise term is explicitly modelled and expressed as shown in Eq. 14.

$$A(x^{-1})y(t) = x^{-d}B(x^{-1})u(t) + C(x^{-1})e(t) \quad (14)$$

## 2.4. System Identification in MATLAB Environment

System identification toolbox in MATLAB was used to model the soil moisture dynamics starting with a lower order, increasing and decreasing the order of the polynomial and the variable delay. Fig. 4 shows the interface of the system identification toolbox where various

model structures were used. A total of 9674 samples collected at 600s time steps were used in the modelling with 70 % of the data as estimation and 30 % as validation data. Additionally, pre-processing of the experimental data was performed to get rid of the trends, outliers, and means. as suggested by [42].

### 2.5. Model Evaluation

With the help of the lowest order of the system's dynamics, the models constructed inside the MATLAB system identification toolbox are able to predict the behaviour of the system's output and graphically represent the relationship between the system's input and output variables [31]. Best fit, final prediction error (FPE), and mean square error (MSE) are three statistical measures that may be used to evaluate the efficacy of the developed models [31].

#### 2.5.1. Final Prediction Error

Model quality is quantified by the absolute value of the prediction error obtained during validation on an independent data set. According to Abdul-Rahim et al. [43] the accuracy of the derived model is dependent on how small the FPE is. This is expressed in Eq. 15.

$$FPE = \det \left( \frac{1}{N} \sum_1^N e(t, \hat{\theta}_N)(e(t, \hat{\theta}_N))^N \right) \left( \frac{1 + \frac{d}{N}}{1 - \frac{d}{N}} \right) \quad (15)$$

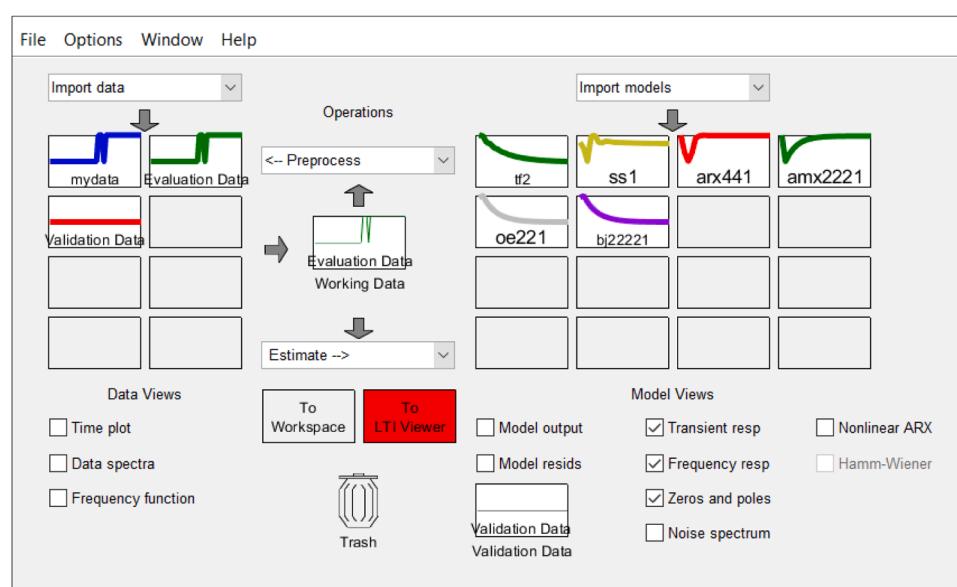
Where;

Estimation data is denoted by N, and the associated prediction errors are represented by  $e(t)$ , a vector of dimensions  $(n, -1)$ . The estimated parameters are denoted by  $\hat{\theta}_N$ , while the total number of parameters is denoted by d

#### 2.5.2. Mean Square Error

The mean square error (MSE) is a statistical metric used to evaluate prediction quality in the same unit of the variable. The MSE score is near to zero, indicating that the model prediction is correct. The MSE is defined by Eq. 16:

$$E = \frac{1}{n} \sum_1^n (y_i - \hat{y}_i)^2 \quad (16)$$



**Fig. 4.** MATLAB System identification toolbox.

Where;

$y_i$  and  $\hat{y}_i$  are observed and predicted values at time ( $i = 1, 2, 3, \dots, n$ ) and  $n$  is the data point.

### 2.5.3. Estimated Fit

The estimated fit (%) is a measure of the correlation between the measured output  $y_i$  and the predicted output  $\hat{y}_i$ . This is computed as shown in Eq. 17.

$$\text{Estimated Fit} = \left( 1 - \frac{\sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\sqrt{\sum_{i=1}^n (y_i - \frac{1}{N} \sum_{i=1}^n y_i)^2}} \right) \times 100 \quad (17)$$

## 3. Results and Discussions

### 3.1. Daily $ET_0$ and Solar Radiation

Fig. 5 presents the relationship between reference evapotranspiration  $ET_0$ , and Solar Radiation  $Rs$ , during the experimental data collection. It is observed that peak solar radiation received in the afternoon which translates into peak values of  $ET_0$ . The maximum  $ET_0$  recorded during the study period is  $4.8 \text{ mm day}^{-1}$  while the maximum solar radiation was  $1357.9 \text{ W m}^{-2}$ . Reference evapotranspiration and solar radiation are crucial for effective irrigation management.

### 3.2. Reference evapotranspiration, water content and precipitation

The soil moisture content at the start of the experiment was  $0.34 \text{ m}^3/\text{m}^3$  as depicted in Fig. 6. During this time precipitation values were high given that the rainy season was coming to an end. As the precipitation reduced, the soil moisture content reduced to field capacity where it was being maintained through irrigation. The dynamics may be seen altering because of plant water absorption and environmental (weather) effects that induce water loss from the soil, resulting in a changing variation in

the volumetric water content. As evapotranspiration causes water loss, the volumetric water content of the soil changes, indicating that more or less water is required to be provided for irrigation. As a result, these variables have a direct influence on the amount of water applied to the plant.

### 3.3. Statistical Performance of the Models

From the system identification simulation results in MATLAB presented in Table 1, it is evident that state space model performed better than other model structures (Transfer Functions, ARMAX, Box Jenkins, ARX and Output Error model). The state space mode yielded an estimated fit of 97.04 % being the highest amongst all the model structures evaluated. Additionally, in terms of the mean square error and final prediction error, the state space model performed better with the least values of the statistical criterion. The Output Error model on the other hand performed poorly amongst all the model structures. It could not mimic the soil moisture dynamics in the field and hence was not chosen for a model predictive controller design. The step ahead prediction results presented in Table 2 also support the state space model's ability to represent the dynamic nature of the soil-water agricultural system. With a 5-step ahead prediction it gives an estimated fit of 98.92 % making it the best choice for MPC design. The chosen model estimated the parameters as shown in Eqs. 15 and 17 in discrete-time state space model. Fig. 7 shows the 5-step ahead prediction of the system.

The discrete time state space model of the system are presented as Eqs. 18 a-g.

$$x(t+Ts) = A x(t) + B u(t) + K e(t) \quad (18a)$$

$$y(t) = C x(t) + D u(t) + e(t) \quad (18b)$$

$$A = \begin{matrix} . & x1 & x2 \\ x1 & 1 & -0.0008092 \\ x2 & -0.01197 & 0.9167 \end{matrix} \quad (18c)$$

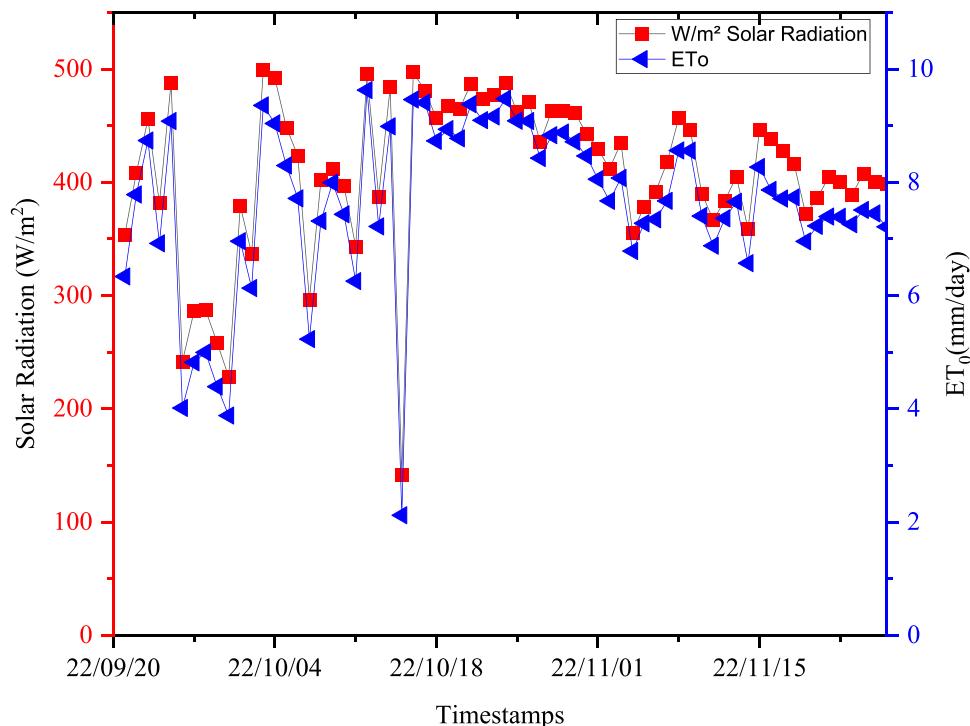
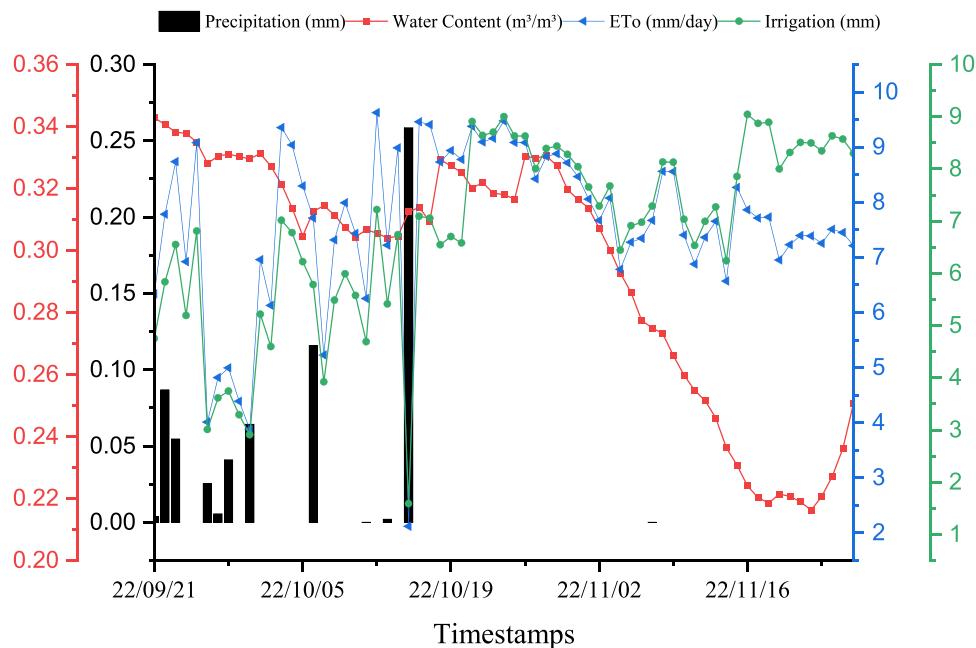


Fig. 5. A graph of Solar radiation vs  $ET_0$  during the experimental period.



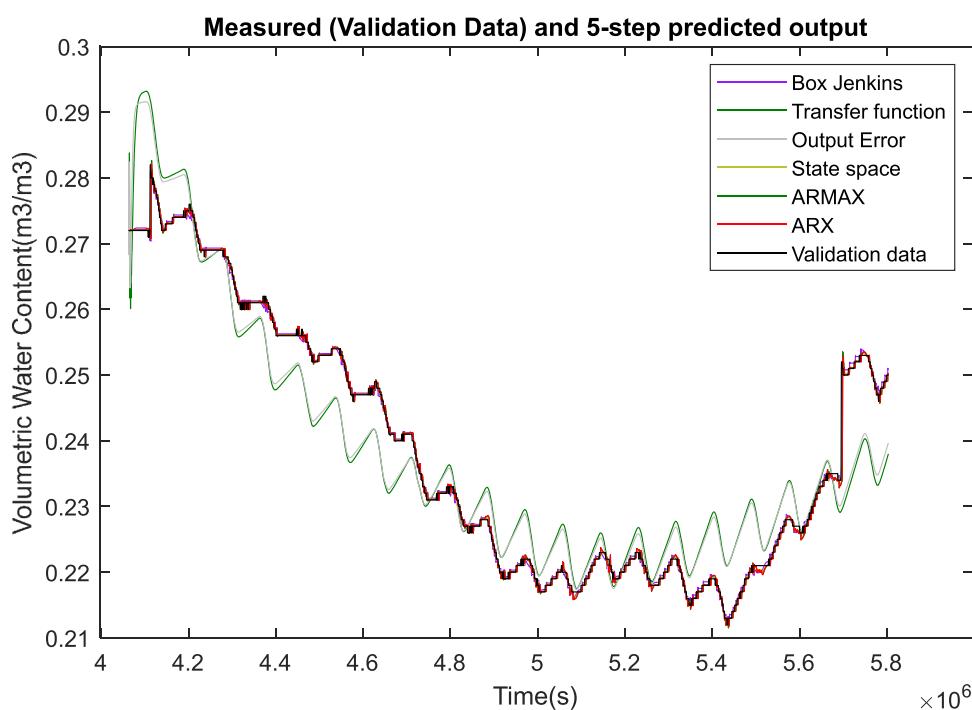
**Fig. 6.** A graph of  $ET_0$ , water content, Irrigation and Precipitation during the experimental period.

**Table 1**  
Statistical model performance of the system.

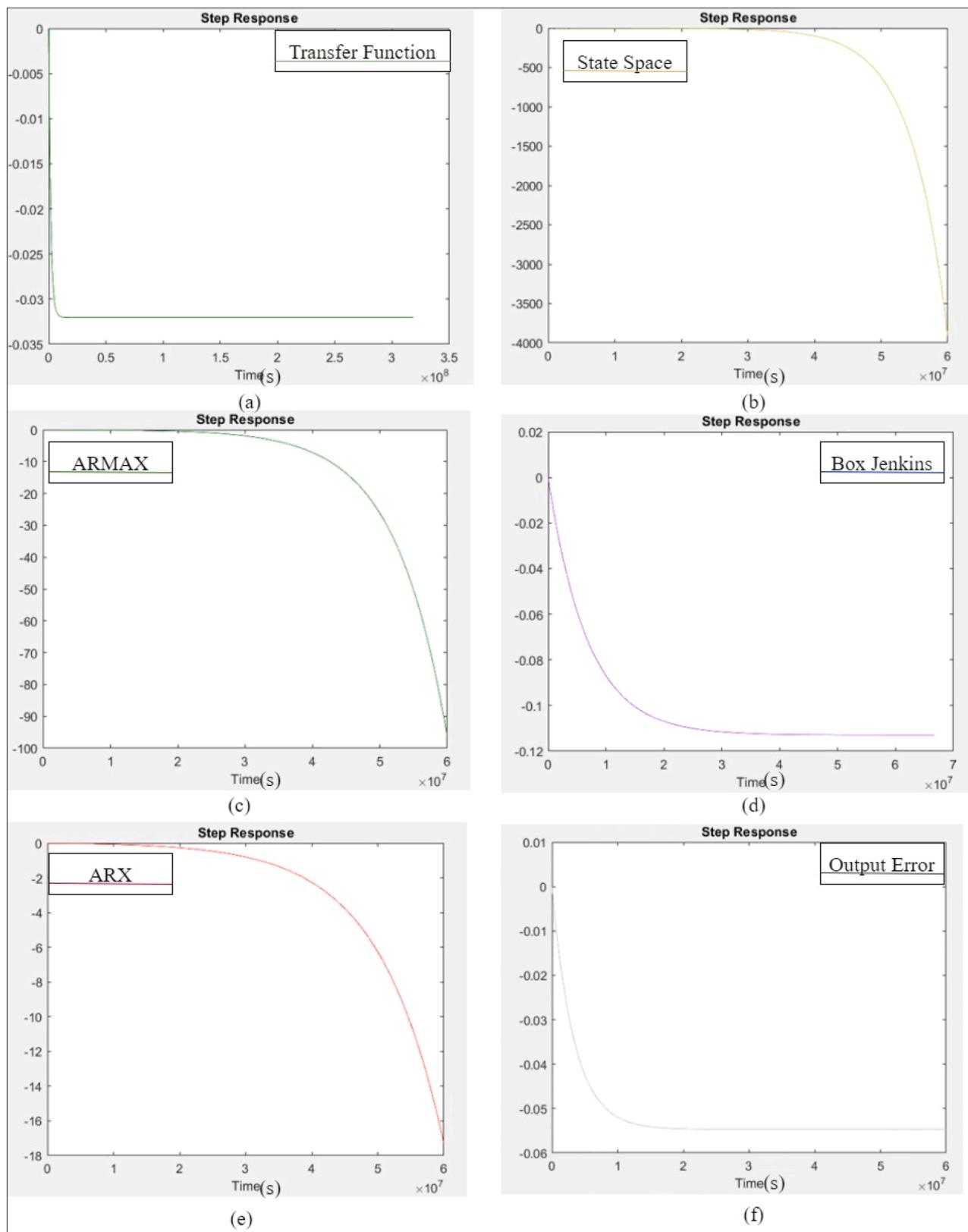
Model	FPE	MSE	Estimate Fit (%)
Transfer Functions	1.41e-4	1.40e-4	16.15
State Space Model	1.75e-7	1.74e-7	97.04
ARMAX	1.76e-7	1.76e-7	97.03
Box Jenkins	1.85e-7	1.79e-7	97.00
ARX	1.76e-7	1.75e-7	97.04
Output Error	1.44e-7	1.43e-7	15.29

**Table 2**  
Step-ahead predictions of the system.

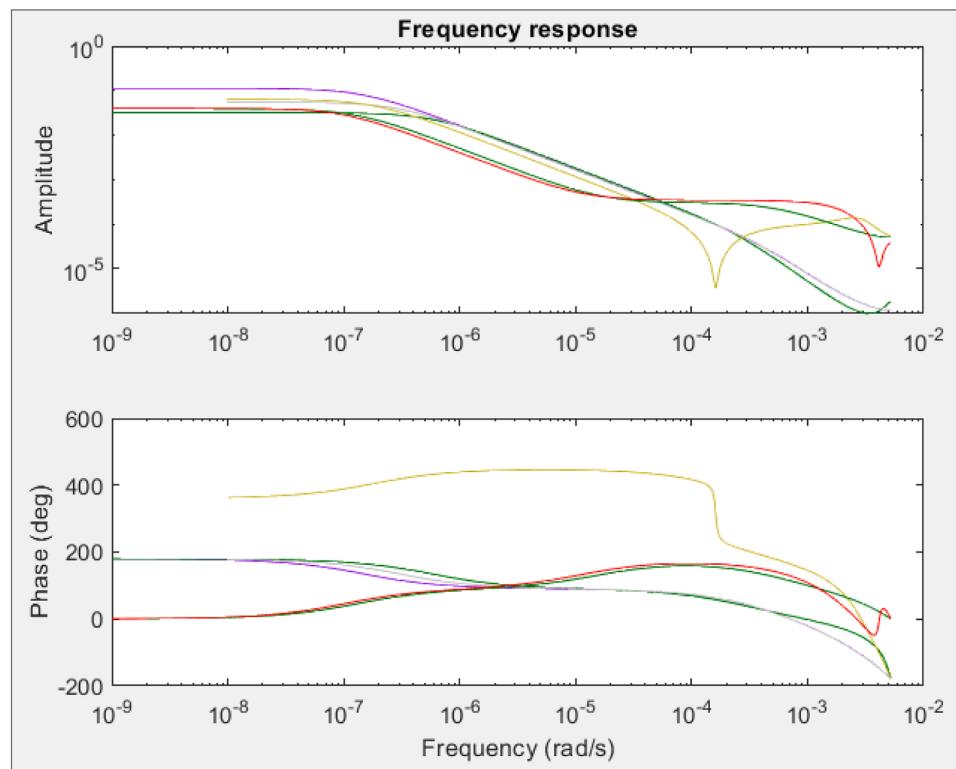
Model	5 step ahead prediction (%)	3 step ahead prediction (%)	1 step ahead prediction (%)
Transfer functions	62.48	62.48	62.48
State Space Model	98.92	95.98	97.64
ARMAX	94.82	95.92	97.63
Box Jenkins	94.92	95.89	97.63
ARX	94.82	95.92	97.62
Output Error	64.73	64.73	64.73



**Fig. 7.** 5-step ahead prediction of the system.



**Fig. 8.** Step response of the system: (a) Transfer function, (b) State space model, (c) ARMAX model, (d) Box Jenkins model, (e) ARX model, and (f) Output error model.



**Fig. 9.** Frequency Response Curve.

$$\begin{matrix} & u_1 & u_2 & u_3 \\ B = \begin{matrix} x_1 \\ x_2 \end{matrix} & -7.977e-07 & 7.204e-07 & -1.457e-05 \\ & -0.0001524 & 9.12e-05 & 0.001182 \end{matrix} \quad (18d)$$

$$C = \begin{matrix} & x_1 & x_2 \\ y_1 & -22.49 & 0.0009123 \end{matrix} \quad (18e)$$

$$D = \begin{matrix} & u_1 & u_2 & u_3 \\ x_1 & 0 & 0 & 0 \end{matrix} \quad (18f)$$

$$K = \begin{matrix} & y_1 \\ x_1 & -0.0463 \\ x_2 & -0.005392 \end{matrix} \quad (18g)$$

### 3.4. Step Response of Each Model

**Fig. 8** represents the step response of the model structures evaluated in the MATLAB environment. The model structures are graded using the lowest order of the dynamic response. The first-order models' output was found to be steady, with no overshoot or undershoot but a sluggish response time. There is no lag and a faster rising time compared to other models, even the Box Jenkins model with an offset. The frequency response curve, shown in **Fig. 9**, reveals that the dynamics of all model structures are identical over a wide range of frequencies. The purpose of this data-driven modelling was to create the forecasting models integral to creating a model-based controller for the physical system. The results show that despite the nonlinearity of the inputs, the output is quite steady. Despite experimental validation, the prediction model created in this study has some margin of error.

The step and frequency response of the models can provide valuable information for practical agricultural purposes. For example, the step response can help to understand how quickly the system responds to changes in input (e.g., irrigation) and to identify any delays or oscillations in the response. The frequency response can help to identify the natural frequencies of the system and to design controllers that can regulate the system at different frequencies.

### 4. Conclusion

Improving water use efficiency in precision agriculture is of paramount importance. Model-based irrigation management is one of the approaches through which irrigation scheduling can be responsive in real time especially for open field agricultural systems. This study demonstrates a data-driven approach where data from monitoring sensors in a drip irrigated tomato experiment is used to develop a soil moisture dynamics model using system identification. The resulting model is a state space model for great performance in terms of statistical evaluation criteria. In comparison to other model structures, it gives an estimation of fit of 97.04%, mean square error of  $1.74 \times 10^{-7}$  and final prediction error of  $1.75 \times 10^{-7}$ . Future research will involve utilizing the state space model to develop a model predictive controller which will later be deployed on a micro-controller in an open field agricultural system.

### Funding Information

The Institute de Recherché pour le Développement (IRD), AFD, and the West African Center for Water, Irrigation, and Sustainable Agriculture (WACWISA), University for Development Studies, Ghana, financed this research.

### CRediT authorship contribution statement

**Erion Bwambale:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft. **Felix K. Abagale:** Writing – review & editing, Visualization, Supervision, Project administration. **Geophrey K. Anornu:** Writing – review & editing, Supervision.

### 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.

## Acknowledgement

The West African Centre for Water, Irrigation, and Sustainable Agriculture (WACWISA), the Regional Water and Sanitation centre Kumasi, and the University for Development Studies are all deserving of our gratitude for their roles in making this study possible through their funding and provision of facilities, respectively. Likewise, we appreciate the work of the unnamed reviewers who made it possible for this study to be published.

## References

- [1] D.D. Chiarelli, et al., Competition for water induced by transnational land acquisitions for agriculture, *Nat. Commun.* 13 (1) (2022) 1–9, <https://doi.org/10.1038/s41467-022-28077-2>.
- [2] F. Dolan, J. Lamontagne, R. Link, M. Hejazi, P. Reed, J. Edmonds, Evaluating the economic impact of water scarcity in a changing world, *Nat. Commun.* 12 (1) (2021) 1–10, <https://doi.org/10.1038/s41467-021-22194-0>.
- [3] B.D. Richter, et al., Water scarcity and fish imperilment driven by beef production, *Nat. Sustain.* 3 (4) (2020) 319–328, <https://doi.org/10.1038/s41893-020-0483-z>.
- [4] X. Liu, W. Liu, Q. Tang, B. Liu, Y. Wada, H. Yang, Global Agricultural Water Scarcity Assessment Incorporating Blue and Green Water Availability Under Future Climate Change, *Earth's Futur* 10 (4) (2022), <https://doi.org/10.1029/2021EF002567>.
- [5] FAO, IFAD, UNICEF, WFP, and WHO, *The State of Food Security and Nutrition in the World 2022*, 10, FAO, Rome, 2022.
- [6] ICID, *Agricultural Water Management For Sustainable Rural Development: Annual report 2020-21*, 2020.
- [7] United Nations, “The United Nations World Water Development Report 2021: Valuing Water,” Paris, 2021. doi: 10.4324/9781315627250-3.
- [8] FAO, *The State of Food and Agriculture 2021. Making agrifood Systems More Resilient to Shocks and Stresses*, 2021.
- [9] T.P. Higginbottom, R. Adhikari, R. Dimova, S. Redicker, T. Foster, Performance of large-scale irrigation projects in sub-Saharan Africa, *Nat. Sustain.* (2021), <https://doi.org/10.1038/s41893-020-00670-7>.
- [10] N. Ungureanu, V. Vlăduț, G. Voicu, Water scarcity and wastewater reuse in crop irrigation, *Sustain* 12 (21) (2020) 1–19, <https://doi.org/10.3390/su12219055>.
- [11] D.E. Eisenhauer, D.L. Martin, *Irrigation Systems Management*, American Society of Agricultural Engineers, 2021.
- [12] J.D. Gil, M. Munoz, L. Roca, F. Rodriguez, M. Berenguel, An IoT based control system for a solar membrane distillation plant used for greenhouse irrigation, Glob. IoT Summit, GIOTS 2019 - Proc. (2019), <https://doi.org/10.1109/GIOTS.2019.8766370>.
- [13] H. Benyezza, M. Bouhedda, S. Rebouh, Zoning irrigation smart system based on fuzzy control technology and IoT for water and energy saving, *J. Clean. Prod.* 302 (2021), 127001, <https://doi.org/10.1016/j.jclepro.2021.127001>.
- [14] F.J. Montáns, F. Chinesta, R. Gómez-Bombarelli, J.N. Kutz, Data-driven modeling and learning in science and engineering, *Comptes Rendus - Mec* 347 (11) (2019) 845–855, <https://doi.org/10.1016/j.crme.2019.11.009>.
- [15] J. Berberich, J. Köhler, M.A. Müller, F. Allgöwer, Data-driven model predictive control: closed-loop guarantees and experimental results, *Autom* 69 (7) (2021) 608–618, <https://doi.org/10.1515/auto-2021-0024>.
- [16] G. Cáceres, P. Millán, M. Pereira, D. Lozano, Smart farm irrigation: model predictive control for economic optimal irrigation in agriculture, *Agronomy* 11 (9) (2021) 1–18, <https://doi.org/10.3390/agronomy11091810>.
- [17] J. Wanyama, E. Bwambale, Precision Water Management, *Encycl. Smart Agric. Technol.* (2023) 1–8, [https://doi.org/10.1007/978-3-030-89123-7\\_213-1](https://doi.org/10.1007/978-3-030-89123-7_213-1).
- [18] E. Bwambale, F.K. Abagale, *Smart Irrigation Monitoring and Control. Encyclopedia of Smart Agriculture Technologies*, Springer International Publishing, Cham, 2023, pp. 1–7.
- [19] E. Bwambale, F.K. Abagale, G.K. Anornu, Data-driven model predictive control for precision irrigation management, *Smart Agric. Technol.* 3 (2023), 100074, <https://doi.org/10.1016/J.ATECH.2022.100074>.
- [20] W.-H. Chen, et al., Data-driven robust model predictive control framework for stem water potential regulation and irrigation in water management, *Control Eng. Pract.* 113 (November 2020) (2021), 104841, <https://doi.org/10.1016/j.conengprac.2021.104841>.
- [21] J. Nahar, S. Liu, Y. Mao, J. Liu, S.L. Shah, Closed-Loop Scheduling and Control for Precision Irrigation †, *Ind. Eng. Chem. Res.* 58 (26) (2019) 11485–11497, <https://doi.org/10.1021/acs.iecr.8b06184>.
- [22] C. Shang, W.H. Chen, A.D. Stroock, F. You, Robust Model Predictive Control of Irrigation Systems with Active Uncertainty Learning and Data Analytics, *IEEE Trans. Control Syst. Technol.* 28 (4) (2020) 1493–1504, <https://doi.org/10.1109/TCST.2019.2916753>.
- [23] L.S. Pereira, P. Paredes, N. Jovanovic, Soil water balance models for determining crop water and irrigation requirements and irrigation scheduling focusing on the FAO56 method and the dual Kc approach, *Agric. Water Manag.* 241 (June) (2020), 106357, <https://doi.org/10.1016/j.agwat.2020.106357>.
- [24] Y. Mao, S. Liu, J. Nahar, J. Liu, F. Ding, Soil moisture regulation of agro-hydrological systems using zone model predictive control, *Comput. Electron. Agric.* 154 (March) (2018) 239–247, <https://doi.org/10.1016/j.compag.2018.09.011>.
- [25] B.T. Agyeman, S.R. Sahoo, J. Liu, S.L. Shah, LSTM-based model predictive control with discrete actuators for irrigation scheduling, *IFAC-PapersOnLine* 55 (7) (2022) 334–339, <https://doi.org/10.1016/j.ifacol.2022.07.466>.
- [26] E. Bwambale, F.K. Abagale, G.K. Anornu, Smart irrigation monitoring and control strategies for improving water use efficiency in precision agriculture: a review, *Agric. Water Manag.* 260 (107324) (2022) 1–12, <https://doi.org/10.1016/j.agwat.2021.107324>.
- [27] S.R. Sahoo, J. Liu, Adaptive Model Reduction and State Estimation of Agro-hydrological Systems, *Comput. Electron. Agric.* 195 (March) (2022), 106825, <https://doi.org/10.1016/j.compag.2022.106825>.
- [28] S.K. Saleem, et al., Model Predictive Control for Real-Time Irrigation Scheduling, *IFAC Proc. Vol.* 46 (18) (2013) 299–304, <https://doi.org/10.3182/20130828-2-SF-3019.00062>.
- [29] C. Lozoya, et al., Model predictive control for closed-loop irrigation, *IFAC Proc. Vol.* 19 (January) (2014) 4429–4434, <https://doi.org/10.3182/20140824-6-za-1003.02067>.
- [30] Y. Mao, S. Liu, J. Nahar, J. Liu, F. Ding, Regulation of soil moisture using zone model predictive control, *IFAC-PapersOnLine* 51 (18) (2018) 762–767, <https://doi.org/10.1016/j.ifacol.2018.09.271>.
- [31] E.A. Abioye, et al., IoT-based monitoring and data-driven modelling of drip irrigation system for mustard leaf cultivation experiment, *Inf. Process. Agric.* 8 (2) (2021) 270–283, <https://doi.org/10.1016/j.inpa.2020.05.004>.
- [32] C. Lozoya, C. Mendoza, A. Aguilar, A. Román, R. Castelló, Sensor-Based Model Driven Control Strategy for Precision Irrigation, *J. Sensors* 2016 (2016), <https://doi.org/10.1155/2016/9784071>.
- [33] R.G. Allen, L.S. Pereira, D. Raes, M. Smith, *Crop Evapotranspiration. Guidelines for Computing Crop Water Requirements*. FAO Irrigation and Drainage Paper 56, Food and Agricultural Organization of the United Nations, Rome, Italy, 1998.
- [34] L. Wang, *Model Predictive Control System Design and Implementation Using MATLAB. Advances in Industrial Control*, Springer, 2009, p. 403.
- [35] A.C. Kemp, R.J. Telford, Transfer functions. *Handb. Sea-Level Res.*, 2015, pp. 470–499, <https://doi.org/10.1002/9781118452547.ch31>.
- [36] G. Box, G.M. Jenkins, G.C. Reinsel, Ljung Greta, *Time Series Analysis: Forecasting and Control, 5th Edition* | Wiley, Wiley, 2017.
- [37] Z. Chen, E. Brown, State space model, *Scholarpedia* 8 (3) (2013) 30868, <https://doi.org/10.4249/scholarpedia.30868>.
- [38] J. How, E. Frazzoli, *Feedback Control Systems. Fall 2010*. Massachusetts Institute of Technology: MIT OpenCourseWare, 2010.
- [39] O.J. Smith, MUS420 Introduction to Linear State Space Models. Introduction to Linear State Space Models, 2019, <https://ccrma.stanford.edu/~jos/StateSpace/StateSpace.pdf>, accessed May 27, 2021.
- [40] E. Tohme, *Initialization of Output Error Identification Algorithms*, University of balamand, 2006.
- [41] M. Horner, S.N. Pakzad, N.S. Gulgec, Parameter Estimation of Autoregressive-Exogenous and Autoregressive Models Subject to Missing Data Using Expectation Maximization, *Front. Built Environ.* 5 (September) (2019) 1–18, <https://doi.org/10.3389/fbuil.2019.00109>.
- [42] L. Ljung, *The Control Systems Handbook*, CRC Press, 2018.
- [43] H. Abdul Rahim, F. Ibrahim, M.N. Taib, System identification of nonlinear autoregressive models in monitoring dengue infection, *Int. J. Smart Sens. Intell. Syst.* 3 (4) (2010) 783–806, <https://doi.org/10.21307/ijssis-2017-421>.