# Optimal Control for Precision Irrigation of a Large-Scale Plantation

Kassing, R.C.; De Schutter, B.H.K.; Abraham, E.

**DOI**
10.1029/2019WR026989

**Publication date**
2020

**Document Version**
Final published version

**Published in**
Water Resources Research

**Citation (APA)**
Kassing, R. C., De Schutter, B. H. K., & Abraham, E. (2020). Optimal Control for Precision Irrigation of a Large-Scale Plantation. *Water Resources Research*, *56*(10), 1-22. Article e2019WR026989. https://doi.org/10.1029/2019WR026989

# Optimal Control for Precision Irrigation of a Large-Scale Plantation

**R.C. Kassing[1*], B. De Schutter[2†], E. Abraham[1‡]**

[2]Delft Center for System and Control, Delft University of Technology, Mekelweg 2, 2628 CD, Delft, The Netherlands

[1]Water Management Department, Delft University of Technology, Stevinweg 1, 2628 CN, Delft, The Netherlands

**Key Points:**

- A two-level optimization method is proposed that is tailored to solve the optimal irrigation problem of a large-scale plantation
- AquaCrop-OS is used to model the agro-hydrological system using real data and to simulate a plantation consisting of multiple fields
- The proposed optimization method produces optimal irrigation schedules in periods of drought and maximizes water productivity

Corresponding author: Edo Abraham, E.Abraham@tudelft.nl



This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2019WR026989

**Abstract**

Distributing water optimally is a complex problem that many farmers face yearly, especially in times of drought. In this work, we propose optimization-based feedback control to improve crop yield and water productivity in agriculture irrigation for a plantation consisting of multiple fields. The interaction between soil, water, crop (sugarcane in this work), and the atmosphere is characterized by an agro-hydrological model using the crop water productivity modeling software AquaCrop-OS. To optimally distribute water over the fields, we propose a two-level optimal control approach. In this approach, the seasonal irrigation planner determines the optimal allocation of water over the fields for the entire growth season to maximize the crop yield, by considering an approximation of the crop productivity function. In addition, the model predictive controller takes care of the daily regulation of the soil moisture, respecting the water distribution decided on by the seasonal planner. To reduce the computational complexity of the daily controller, a mixed-logic dynamical model is identified based on the AquaCrop-OS model. This dynamical model incorporates saturation dynamics explicitly to improve model quality. To further improve performance, we create an evapotranspiration model by considering the expected development of the crop over the season using remote-sensing-based measurements of the canopy cover. The performance of the two-level approach is evaluated through a closed-loop simulation in AquaCrop-OS of a real sugarcane plantation in Mozambique. Our optimal control approach boosts water productivity by up to 30% compared to local heuristics and can respect water use constraints that arise in times of drought.

## 1 Introduction

Global water demands are projected to increase by 20% to 30% by 2050 compared to 2010 levels, due to population growth and emerging economies (Burek et al., 2016). Agriculture is the largest global consumer of the available water resources, accounting for 69% of annual water withdrawals (*Water Energy Nexus: Excerpt from the World Energy Outlook 2016*, 2016). Therefore, a pivotal step in addressing the alarming water-scarcity problem is improving irrigation efficiency in agriculture. Conventionally, irrigation operates under open-loop conditions, making use of heuristics or empirical data. These traditional approaches miss out on real-time feedback information from the fields, such as evapotranspiration rates and soil moisture measurements. As a result, open-loop irrigation methods are often unable to irrigate with precision, leading to overirrigation.

With increasing interest to achieve precision irrigation, Model Predictive Control (MPC) has been explored in the past decade as an optimal feedback control method. In MPC, the evolution of a system (e.g., the crop field) is predicted using a model, subject to the control inputs (e.g., amount and timing of irrigation). These control inputs are then optimized with respect to the desired trajectory of the system states that would optimize some performance (e.g., maximize crop yield or minimize water use). The control inputs corresponding to the first time step are implemented and for the next time step the optimization procedure is repeated using newly obtained measurements and state estimates. Park et al. (2009) used MPC to maintain the soil moisture and salt levels of a field below a desired threshold. McCarthy et al. (2014) divided a heterogeneous field in zones based on spatial and temporal differences, and used MPC to optimize irrigation. An MPC framework was developed in Saleem et al. (2013) using a water balance model for real-time irrigation scheduling to minimize the root zone soil moisture deficit and the amount of irrigation. Delgoda et al. (2016) identified a linear model based on the crop water productivity model AquaCrop using a water balance and used MPC to minimize the root zone soil moisture deficit, with a limit on the water supply. Mao et al. (2018) approximated a nonlinear agro-hydrological model with a linear parameter varying model and used MPC to maintain the soil moisture within a desired target zone.

The typical objective of the aforementioned studies is to maintain the soil moisture of a single field at a desired set-point or within a target zone using hourly or daily time steps. However, the aforementioned methods do not relate water stress to yield, which is the main performance indicator for any crop in the real setting. Furthermore, water use should also be considered over the entire growth season (a priori preferably because farmers need to be able to plan ahead), not just from day to day as water resources are often limited and droughts are becoming increasingly common. This issue has been addressed in feedforward approaches (e.g., Dudley et al. (1971); Protopapas and Georgakakos (1990); Wardlaw and Barnes (1999); Georgiou and Papamichail (2008)) that schedule irrigation over entire growth seasons. However, the main shortcoming of these later methods is that they miss out on real-time feedback information from the fields to help decide where and when to allocate the water. Recently, to optimize water allocation on both a daily and seasonal temporal scale, Nahar et al. (2019) have proposed a hierarchical feedback approach to maximize yield over an entire growth season. A top-level scheduler calculates a target soil moisture value for each remaining day in the growth season for a single field and a lower-level controller follows the target soil moisture. Then, the proposed closed-loop scheduling approach is compared to open-loop scheduling over a full season (by crop yield and water consumption). However, the authors were unable to show significant difference between the two approaches.

In this paper, unlike in the papers discussed above, we consider the more general problem of optimal soil moisture control of multiple fields (e.g., a plantation) for a full growth season. Here we consider allocating water optimally over the growth season with daily irrigation control for multiple fields, while considering water availability changes throughout the season. This adds complexity to the control problem, as operational constraints need to be included (such as a limited number of fields that can be irrigated in a day) and trade-offs need to be made between irrigation and crop yield of the different fields. Furthermore, the growth stages of the fields can be different, as often not all fields can be planted and harvested at the same time. We propose a novel two-level feedback control approach for the water distribution of an entire plantation, in order to improve the water use efficiency and yield. A process schematic of the proposed two-level control approach is depicted in Figure 1. The seasonal irrigation planner determines the optimal distribution of irrigation water over the fields and growth stages of the crop for the entire growth season, in order to maximize the crop yield under scarce water conditions. We could also maximize economic return from the water. However, in this work we consider a plantation that has the same crop on all fields and so maximizing yield is the same as maximizing economic return. The daily irrigation controller uses MPC to regulate a target soil moisture zone subject to the (seasonal) irrigation constraints set by the seasonal irrigation planner and operational constraints. In our framework, we apply MPC in a receding horizon implementation, where the input sequence that yields an optimal predicted output while simultaneously satisfying all the constraints is computed using predicted weather conditions. Receding horizon control accomplishes feedback against uncertain disturbances by applying only part of the computed input sequence (here for one day) and then updating the system states using new information (e.g., deviations in temperature and precipitation) and recomputing the next optimal sequence a day later. As the prediction horizon slides along, this process of state updates using measurements, prediction and control optimization is repeated. This process is generally inherently robust to a small amount of uncertainty (Mayne et al., 2000), and good closed-loop performance can be expected in this application if the water availability and plant water demand do not change drastically. To validate our method we model a real sugarcane plantation from Mozambique in AquaCrop-OS (Foster et al., 2017) (the Food and Agriculture Organization of the United Nations (FAO) standard for simulating the crop yield response when water is a key limiting factor in crop production) and evaluate the performance of our two-level optimization algorithm using closed-loop simulations in AquaCrop-OS.

This paper is organized as follows:

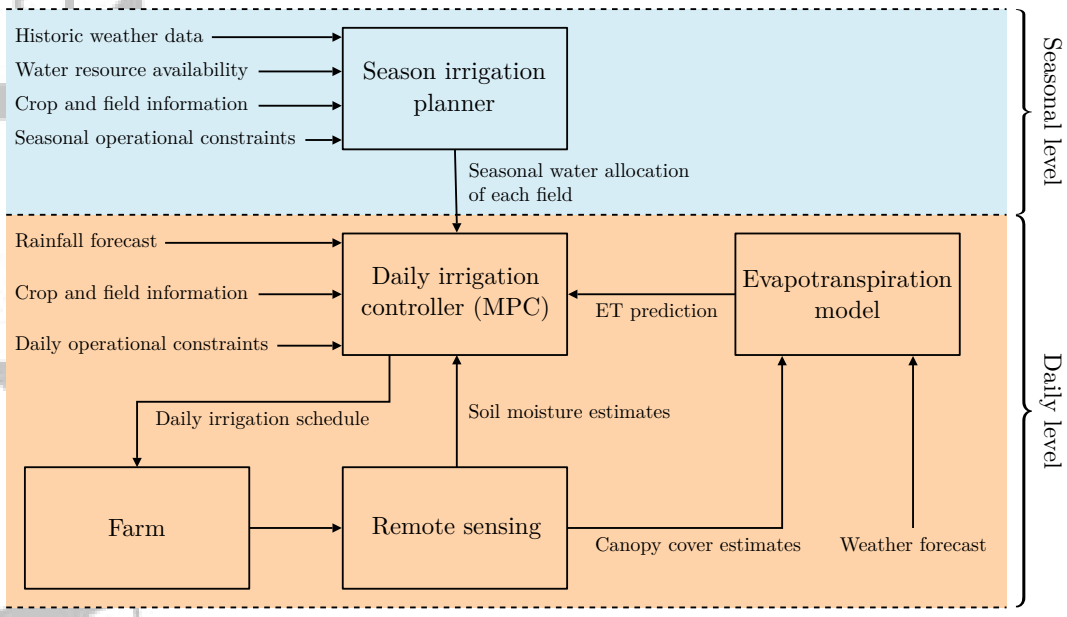


**Figure 1.** Process diagram for the proposed two-level optimization approach for precision irrigation. The seasonal irrigation planner first calculates the water amount that each field can use during each growth stage; the daily model predictive controller then regulates the soil moisture using irrigation control while respecting the water amount that the seasonal irrigation planner imposed. Note that if the water availability changes during the season the seasonal irrigation planner can reevaluate the water distribution.

In Section 2, we model the: (1) agro-hydrological system using a simple water balance from literature which we expand on by including a novel mixed-logic saturation model, (2) evapotranspiration process by assimilating case study specific crop growth data and atmospheric parameters, (3) effect of water stress on crop growth using piecewise linear approximations of a crop yield model for sugarcane. We complete the section with an exposition of the complete optimization problem and our proposed two-level optimization method that is used to solve it. In Section 3, our proposed models are identified and validated on an AquaCrop-OS simulation model of a real sugarcane plantation in Mozambique. In Section 4, we evaluate the performance of our two-level optimization method by conducting three closed-loop simulations using AquaCrop-OS involving (1) the effect of decreasing water use on crop yield, (2) comparing the water productivity of our proposed approach with the plantation's approach of irrigation every 3-4 days to field capacity, (3) assessment of the performance of our proposed approach under water scarcity. Section 5 summarizes some conclusions and further research required to realize an implementation of the proposed approach in the real irrigation system.

## 2 Modeling and Optimal Control

The objective of precision irrigation is to maximize the yield and water efficiency by irrigating the fields with an adequate amount of water at the most suitable time instants. Therefore, the crop yield is maximized by minimizing the amount of water stress the crop is exposed to. This water stress occurs when not enough water is available in the root zone of the soil for the crop to absorb with its roots. Therefore, in order to minimize the water stress, a model of the agro-hydrological system is needed that characterizes the interaction between soil moisture, the crop and the atmosphere. Furthermore,

the effect of the water stress on the yield needs to be quantified per growth stage, such that irrigation is prioritized in the water sensitive stages. This is achieved by considering a simplification of the yield prediction function proposed in Raes et al. (2006). Finally, the irrigation of the fields is constrained by operational limitations, such as water availability or availability of human operators and/or machinery such as sprinklers.

## 2.1 Agro-hydrological modeling

Here we consider a plantation with multiple fields, where each field $j$ is characterized by a field-specific agro-hydrological dynamical system. A schematic of this agro-hydrological model for a single field is shown in Figure 2, together with water fluxes at the boundaries of the root zone. In the agro-hydrological system, water transportation takes place by means of rainfall, drainage, evaporation, transpiration, and irrigation. Here, soil homogeneity for each individual field is assumed, i.e. a field is modeled using a single soil layer, and a simplified water balance approach (M. Jensen et al., 1971; Delgoda et al., 2016) is used for irrigation control. This simplified water balance approach is selected over more complex modeling approaches like those of (Mao et al., 2018) for computational feasibility reasons, as the complexity of the scheduling problem will grow exponentially with the number of fields in the later. Furthermore, the system will operate using a daily time step, for which the dynamics are more easily captured in a simple water balance. A crop experiences water stress when the potential energy of the soil water limits the availability of water extraction by the plants roots, i.e. the Total Available Water (TAW) in Figure 2 drops below the water stress threshold. This water stress threshold depends on the crop type and the evaporation power of the atmosphere. The water stress is expressed by the Root Zone Depletion (RZD) (Allen et al., 1998), which is the combined amount of rain and irrigation needed to bring the soil moisture content of the root zone back to field capacity. A negative RZD indicates excess water, which will be drained away from the field over time. Let the variable $D$ represent the RZD of a field; then its dynamics can be given as (Allen et al., 1998):

$$D(k) = D(k-1) + E^*(k) - P_{\mathrm{e}}(k) - I(k) + G(k) + R(k), \tag{1}$$

where $D(k)$ is the RZD at the end of day $k$, $E^*(k)$ is the total crop evapotranspiration for the day, $P_{\mathrm{e}}(k)$ the effective rainfall (i.e. the rainfall reaching the soil after interception by the canopy), $I(k)$ the irrigation depth, $G(k)$ the deep percolation (drainage), $R(k)$ the runoff, and $k$ the time step in days. The groundwater tables of the fields we consider here are low; therefore, the contribution of capillary rise is considered negligible. Note that for notational convenience, we omit the field index $j$ from the equations here. Furthermore, the variables from Eq. (1) are all expressed in millimeters by normalizing over the surface area.

### *2.1.1 A mixed-logic saturation model*

The RZD is, as the name implies, dependent on the root zone depth ($Z_{\mathrm{r}}$) of the plant during its growth:

$$D(k) = Z_{\mathrm{r}}(k) \cdot (\theta_{\mathrm{FC}} - \theta(k)), \tag{2}$$

where $\theta_{\mathrm{FC}}$ is the volumetric soil water content at field capacity and $\theta(k)$ the soil water content at end of day $k$, expressed in $\mathrm{mm}^3/\mathrm{mm}^3$. Following heavy rain or irrigation, the RZD might exceed the field capacity ($D(k) < 0$). Depending on the hydrological parameters of the soil this excess water will drain over multiple days (Raes et al., 2006, 2018). However, for reasons of simplicity, we assume all fields are well drained and that the soil moisture content is back to field capacity at the end of the first day of free drainage. Note that for soils that are not well drained the deep percolating of the excess water may take longer than one day. However, we assume that for the majority of the days that the fields are operated the soils are unsaturated. Furthermore, in a closed-loop system with fre-

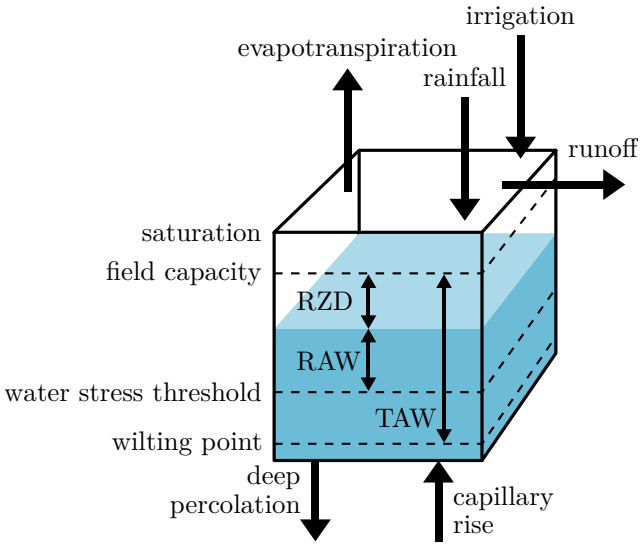


**Figure 2.** Basic water balance model (adapted from Allen et al. (1998, Figure 43)), RAW and TAW stand for readily and total available water to the plant, respectively. The RZD indicates the required water amount to bring the soil moisture content back to field capacity.

quent measurements the soil moisture estimation errors introduced by this saturation modeling error are small.

When the soil becomes saturated it cannot hold any more water and the remaining irrigation or rainfall can be considered as surface runoff. However, any excess water is assumed drained in one day, so saturation never occurs and $R(k) = 0, \forall k$. Consequently, we rewrite Eq. (1) to obtain the following agro-hydrological model:

$$D(k) = \begin{cases} 0 & \text{if } D(k-1) + E^*(k) - P_e(k) - I(k) \leq 0, \\ D(k-1) + E^*(k) - P_e(k) - I(k) & \text{otherwise.} \end{cases} \quad (3)$$

The conditionally switched depletion dynamics in Eq. (3) can be converted into a set of linear dynamic equations and linear inequalities involving real and integer variables called a Mixed-Logic Dynamical (MLD) model (Bemporad & Morari, 1999).

### *2.1.2 Remote sensing data based evapotranspiration model*

Evapotranspiration is the largest contributor to consumptive water use by crops in the agro-hydrological cycle. Therefore, it is the main driver for irrigation and it should be estimated accurately to achieve precision irrigation. Using the Pennman–Monteith method (Allen et al., 1998) or the Priestley–Taylor equation (Priestley & Taylor, 1972), a reference evapotranspiration can be calculated that can then be adjusted with growth stage dependent coefficients for each crop type to estimate the actual evapotranspiration of a crop. In recent years, the application of remote-sensing data for agricultural decision making has been researched extensively (e.g., Estes et al. (1978); Champagne et al. (2010); Mulla (2013); Martens et al. (2017)). A particularly promising remote-sensing based vegetation index is the Normalized Difference Vegetation Index (NDVI), which measures reflectance of the near-infrared and red range of the light spectrum. Vegetation can be particularly well measured using NDVI as plants contain varying amounts of chlorophyll that reflect the light in a distinct way. As a result, NDVI is strongly correlated with physiological processes like evapotranspiration (Trout et al., 2008; Kamble et al., 2013) and biomass (Goswami et al., 2015). In Zhang et al. (2015), evaporation for sugarcane

in Hawaii is estimated using NDVI measurements. The NDVI values are used to estimate the canopy cover using linear regression and subsequently the canopy cover is used to estimate evapotranspiration using another set of linear regression models. This motivates using NDVI directly to obtain estimates of the crop evapotranspiration.

For scheduling irrigation, we propose a new approach to model evapotranspiration that uses NDVI data as a surrogate for canopy size (as directly measuring canopy size for a large-scale plantation is infeasible) to indicate the expected progress of crop growth in a typical growth season. The evapotranspiration is partitioned into (soil) evaporation and transpiration. The evaporation largely depends on the area of the soil covered by the canopy, i.e., when the canopy size increases a smaller fraction of the soil is exposed to the sun and the evaporation will decrease. Conversely, when the canopy size increases the transpiration will increase as a larger canopy will generally consume more water. Therefore, an NDVI-based canopy size curve estimate can be used as a measure of the increasing transpiration levels over the season and also the corresponding fraction of water lost as evaporation, which decreases over the season. Creating an estimate of this expected crop growth in a growth season requires an NDVI dataset, weather data, and planting and harvest dates from at least one past growth season of the fields. However, the growth of a crop depends heavily on the crop varieties, water stress, application of fertilizer, and soil type. Therefore, for a plantation, we assume that the group of fields used in the method grow the same crop with similar characteristics; this method can be repeated for each group of fields in a plantation to obtain accurate evapotranspiration estimates. In applications where small holder farmers are involved, this assumption may not be true and evapotranspiration models for each crop type could be aggregated.

The first step of the method is to model the effect of seasonal air temperature changes on the different growth seasons. By incorporating the evolution of temperature over time we correct for the effect of temperature on the crop development over the growing cycle. This is necessary because the planting and harvest dates describing each growth season can be different for each field, often due to logistical reasons, such as number of harvesting/planting machines available, human operator availability, and water availability. Using the planting date and air temperature data, the NDVI data points are linked to heat units, expressed in Growing Degree Days (GDD), to describe crop development (in $^{\circ}$C day$^{-1}$). In this approach, the time required to reach a particular growth stage is expressed in Cumulative Growing Degree Days (CGDD) instead of number of days. Here we use the method from (McMaster & Wilhelm, 1997), a popular approach from among many methods, where both maximum and minimum temperatures are bounded before calculating the CGDD:

$$\mathrm{GDD}(k) = T_{\mathrm{avg}}(k) - T_{\mathrm{base}}, \tag{4}$$

where the base temperature $T_{\mathrm{base}}$ is the lower bound below which crop development halts. The average temperature $T_{\mathrm{avg}}$ is calculated using:

$$T_{\mathrm{avg}}(k) = \frac{T^*_{\mathrm{max}}(k) + T^*_{\mathrm{min}}(k)}{2}, \tag{5}$$

$$T^*_{\mathrm{max}}(k) = \begin{cases} T_{\mathrm{upper}} & \text{if } T_{\mathrm{max}}(k) \geq T_{\mathrm{upper}}, \\ T_{\mathrm{base}} & \text{if } T_{\mathrm{max}}(k) \leq T_{\mathrm{base}}, \\ T_{\mathrm{max}}(k) & \text{otherwise}, \end{cases} \tag{6}$$

$$T^*_{\mathrm{min}}(k) = \begin{cases} T_{\mathrm{upper}} & \text{if } T_{\mathrm{min}}(k) \geq T_{\mathrm{upper}}, \\ T_{\mathrm{base}} & \text{if } T_{\mathrm{min}}(k) \leq T_{\mathrm{base}}, \\ T_{\mathrm{min}}(k) & \text{otherwise}, \end{cases} \tag{7}$$

where $T_{\mathrm{max}}(k)$ and $T_{\mathrm{min}}(k)$ are the maximum and minimum air temperature, respectively, and $T^*_{\mathrm{max}}(k)$ and $T^*_{\mathrm{min}}(k)$ are the maximum and minimum air temperature adjusted for the upper and lower thresholds that limit crop development. The upper threshold tem-

perature $T_{\text{upper}}$ indicates the upper bound on air temperature above which crop development no longer increases and $T_{\text{base}}$ is the temperature below which growth does not progress.

The second step, involves fitting a curve, $f_{\text{NDVI}}$, through the NDVI–CGDD points of all growth seasons of all fields that represents the expected evolution of the crop size over the season. Note that the type of curve appropriate to use depends on the quality and quantity of the data. The obtained curve describes the expected canopy size over the growth season as a function of the amount of 'warmth' that the crops have experienced since planting.

As a final step, evapotranspiration is split into an evaporation and a transpiration part. The evolution of both the transpiration and evaporation is described by the evolution of the canopy size captured by our NDVI-based curve. However, with an increase in the canopy size the transpiration increases and the evaporation decreases. Let $\mathring{k}$ (a new counter) denote the CGDD of the crop at the end of day $k$:

$$\mathring{k} = \sum_{\kappa=k_0}^{k} \text{GDD}(\kappa), \tag{8}$$

where $k_0$ is the planting day of the field. Then, the evolution of the evaporation, $f_{\text{evap}}$, can be described as a function of the curve $f_{\text{NDVI}}$:

$$f_{\text{evap}}(\mathring{k}) = f_{\text{NDVI}}^{\max} - f_{\text{NDVI}}(\mathring{k}). \tag{9}$$

where

$$f_{\text{NDVI}}^{\max} = \max_{\mathring{k}}(f_{\text{NDVI}}(\mathring{k})). \tag{10}$$

Consequently, the evaporation contribution to the evapotranspiration is assumed to be negligible when the canopy is fully grown. If this is not the case, a value higher than $f_{\text{transp}}^{\max}$ can be chosen to increase the value of $f_{\text{evap}}(\mathring{k})$ in full grown canopy conditions. For the transpiration we can directly use the NDVI-based growth curve. Finally, the curves describing the evolution of the transpiration and evaporation over the growth season are used to estimate the evapotranspiration ($\hat{E}$) of the crop over the growth season:

$$\hat{E}(\mathring{k}) = \underbrace{\alpha_1 f_{\text{NDVI}}(\mathring{k}) \cdot E_{\text{ref}}(\mathring{k})}_{\text{transpiration}} + \underbrace{\alpha_2 f_{\text{evap}}(\mathring{k}) E_{\text{ref}}(\mathring{k})}_{\text{evaporation}} + \alpha_3 \mathring{k}, \tag{11}$$

where $\alpha_1$,$\alpha_2$, and $\alpha_3$ are coefficients that are estimated through linear least-squares using the historic data of the reference evapotranspiration $E_{\text{ref}}(\mathring{k})$. Furthermore, the linear term $\alpha_3\mathring{k}$ corrects for the linear dynamics in evapotranspiration as a function of the season, which is not captured by the transpiration and evaporation curves. Note that temperature predictions are needed to estimate $\mathring{k}$, which then allows to predict future evapotranspiration values using Eq. (11). Furthermore, the reference evapotranspiration $E_{\text{ref}}(\mathring{k})$ is the same for all fields, as we assume the fields are located in the same geographically small (i.e. same climate) area.

## 2.2 Crop yield model

### *2.2.1 Defining water stress*

The amount of water available to the crop is the difference between the permanent wilting point (minimal amount of soil water required for the plant not to wilt) and the field capacity. However, when the soil water content drops below the water stress threshold (see Figure 2), the roots cannot uptake water quickly enough to respond to the water demand and the crop starts experiencing stress. Furthermore, the deeper the roots the bigger the soil column from which the plant can draw water. Therefore, the point

at which the crop starts experiencing water stress can be expressed as a function of the effective rooting depth:

$$p(\theta_{\mathrm{FC}} - \theta_{\mathrm{WP}})Z_{\mathrm{r}}(k), \tag{12}$$

where $p \in [0, 1]$ is the fraction of the total available water below which the crop experiences water stress, $\theta_{\mathrm{FC}}$ is the soil moisture content as a fraction of the field capacity, and $\theta_{\mathrm{WP}}$ is the soil moisture content as a fraction of the wilting point. The roots are initially established at a small depth when the crops are first planted and gradually grow to the maximum depth. However, if the soil becomes too wet ($D(k) < 0$) the crop also starts experiencing water stress, due to anaerobic conditions and water logging. Therefore, we define the target zone in which the crop experiences no water stress as

$$0 \leq D(k) \leq p(\theta_{\mathrm{FC}} - \theta_{\mathrm{WP}})Z_{\mathrm{r}}(k). \tag{13}$$

The lower bound is enforced by the assumed rapid drainage, that is $D(k) \geq 0\ \forall k$. To track the violation $D_{\mathrm{v}}(k)$ of the upper bound of the target zone, a performance constraint is introduced:

$$D_{\mathrm{v}}(k) = \begin{cases} 0 & \text{if } D(k-1) + E^*(k) - P_{\mathrm{e}}(k) - I(k) \leq 0, \\ D(k-1) + E^*(k) - P_{\mathrm{e}}(k) - I(k) - \epsilon(k) & \text{otherwise,} \end{cases} \tag{14}$$

where $\epsilon(k) \in [0, p(\theta_{\mathrm{FC}} - \theta_{\mathrm{WP}})Z_{\mathrm{r}}(k)]$ is a virtual state with the same bounds as $D(k)$ in Eq. (13). Note that when $D(k)$ does not violate the target zone, we have $\epsilon(k) = D(k)$, resulting in $D_{\mathrm{v}}(k) = 0$. This is true because we are solving a linear problem. Then, by penalizing the value of $D_{\mathrm{v}}(k)$, the violation of the upper bound of the target zone is minimized.

#### *2.2.2 Modeling the effect of water stress on yield*

The effect of violating the target soil water content, i.e. water stress, on the yield depends on both the reduction in the evapotranspiration (i.e. the level of water deficit) and the growth stage. As in the literature and standard Food and Agriculture Organization of the United Nations (FAO) approaches (Steduto et al., 2012), we model the response of yield to water stress using sensitivity indices $\lambda_\ell$ for each growth stage $\ell$. These sensitivity indices, together with the potential water use $W_{\mathrm{p},\ell}$ and actual water use $W_{\mathrm{a},\ell}$ specify the effect of water stress on yield decline in each growth stage (Doorenbos & Kassam, 1979):

$$1 - \left(\frac{Y_{\mathrm{a}}}{Y_{\mathrm{p}}}\right)_\ell = \lambda_\ell \left(1 - \frac{W_{\mathrm{a},\ell}}{W_{\mathrm{p},\ell}}\right), \tag{15}$$

where $\left(\frac{Y_{\mathrm{a}}}{Y_{\mathrm{p}}}\right)_\ell$ is the relative yield, the left-hand side is the relative decrease in yield, and $\left(1 - \frac{W_{\mathrm{a},\ell}}{W_{\mathrm{p},\ell}}\right)$ denotes the fraction of reduction in water available to the crop in growth stage $\ell$ compared to the potential water use, i.e. is the total crop water use without stress or other limiting factors. This $W_{\mathrm{p},\ell}$ value can be estimated using the evapotranspiration estimate derived in Section 2.1.2:

$$\hat{W}_{\mathrm{p},\ell} = \sum_{k \in \mathcal{K}_\ell} \hat{E}(\mathring{k}), \tag{16}$$

where $\mathcal{K}_\ell$ is a subset of days in the growth season during which the crop is in growth stage $\ell \in \{1, \ldots, N_{\mathrm{gs}}\}$, where $N_{\mathrm{gs}}$ denotes the number of growth stages. Assuming optimal agronomic practice, the actual water use is equal to the potential water use when the soil moisture is in the target zone. However, when the tolerable depletion exceeds the threshold ($D_{\mathrm{v}}(k) > 0$) the crop experiences water stress and the actual water use is diminished:

$$W_{\mathrm{a},\ell} = \sum_{k \in \mathcal{K}_\ell} E^*(\mathring{k}). \tag{17}$$

The effect of RZD on $E^*(\mathring{k})$ can be modeled by introducing a water stress factor $K_\mathrm{s}(\mathring{k})$ (Allen et al., 1998):

$$E^*(\mathring{k}) = K_\mathrm{s}(\mathring{k})\hat{E}(\mathring{k}), \tag{18}$$

where

$$K_\mathrm{s}(\mathring{k}) = \begin{cases} 1 - \frac{D_\mathrm{v}(\mathring{k})}{(1-p)\cdot(\theta_\mathrm{FC}-\theta_\mathrm{WP})Z_\mathrm{r}(\mathring{k})} & \text{if } 0 < D_\mathrm{v}(\mathring{k}) \leq (1-p)\cdot(\theta_\mathrm{FC}-\theta_\mathrm{WP})Z_\mathrm{r}(\mathring{k}), \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

We assume that under feedback control the soil moisture content will not be allowed to drop below the wilting point; therefore, the water stress factor can be reduced to:

$$K_\mathrm{s}(\mathring{k}) = 1 - \frac{D_v(\mathring{k})}{(1-p)\cdot(\theta_\mathrm{FC}-\theta_\mathrm{WP})Z_\mathrm{r}(\mathring{k})}. \tag{20}$$

Next, by substituting Eqs. (16)–(20) in Eq. (15) we can express the yield as a function of the root zone water dynamics:

$$1 - \left(\frac{Y_\mathrm{a}}{Y_\mathrm{p}}\right)_\ell = \lambda_\ell \frac{\sum_{k\in\mathcal{K}_\ell}\left(\frac{D_v(\mathring{k})}{(1-p)\cdot(\theta_\mathrm{FC}-\theta_\mathrm{WP})Z_\mathrm{r}(\mathring{k})}\hat{E}(\mathring{k})\right)}{\sum_{k\in\mathcal{K}_\ell}\hat{E}(\mathring{k})}. \tag{21}$$

Over the growth season, the sequence of relative yield decreases in each stage are multiplied to compute the overall yield decrease (J. Jensen, 1968; Hanks, 1974). This however results in a high-order (i.e. order equal to number of growth stages) nonlinear function with multiplicative terms between all variables. Therefore, the overall yield is often approximated using a first-order (i.e. additive) compounding function (Stewart et al., 1977; Bras & Cordova, 1981), which we use for the MPC controller as the multiplicative approach is not computationally feasible. Although the additive approach is used to formulate our yield maximizing objective functions, the performance of the MPC controllers are tested using the multiplicative formula in closed-loop simulations with AquaCrop-OS. Note that for the seasonal irrigation planner, the multiplicative compounding function is approximated by a piecewise function, see Section 2.3. Consequently, the performance criterion of the MPC controller is expressed as:

$$\left(1 - \frac{Y_\mathrm{a}}{Y_\mathrm{p}}\right) \approx \sum_{\ell=1}^{N_\ell}\left(1 - \left(\frac{Y_\mathrm{a}}{Y_\mathrm{p}}\right)_\ell\right) = \sum_{\ell=1}^{N_\ell}\lambda_\ell \frac{\sum_{k\in\mathcal{K}_\ell}\left(\frac{D_v(\mathring{k})}{(1-p)\cdot(\theta_\mathrm{FC}-\theta_\mathrm{WP})Z_\mathrm{r}(\mathring{k})}\hat{E}(\mathring{k})\right)}{\sum_{k\in\mathcal{K}_\ell}\hat{E}(\mathring{k})}. \tag{22}$$

The only variable in Eq. (22) that we can manipulate is $D_\mathrm{v}(k)$ by controlling the amount of irrigation. The other variables can all be calculated or predicted before each optimization iteration of the MPC problem. Therefore, the water stress cost ($J_\mathrm{ws}$) to minimize in order to maximize the yield can be expressed as:

$$J_\mathrm{ws} = \sum_{\ell=1}^{N_\ell}\left(\frac{\lambda_\ell}{\sum_{k\in\mathcal{K}_\ell}\hat{E}(\mathring{k})}\sum_{k\in\mathcal{K}_\ell} w_\ell(\mathring{k})D_\mathrm{v}(\mathring{k})\right), \tag{23}$$

where $w_\ell(\mathring{k})$ is the weight on water stress in growth stage $\ell$:

$$w_\ell(\mathring{k}) = \frac{\hat{E}(\mathring{k})}{(1-p)\cdot(\theta_\mathrm{FC}-\theta_\mathrm{WP})Z_\mathrm{r}(\mathring{k})}. \tag{24}$$

The evapotranspiration of each growth stage can be predicted a priori for the entire growth season using historic data, predictions of the temperature, and Eq. (11). Furthermore, estimates of the crop sensitivity indices can be obtained from literature (see Allen et al. (1998)) or can be estimated using a crop model analysis. Moreover, estimates of the root depth can be obtained from field measurements or literature. Therefore, the weights $w_\ell$ can be calculated a priori for the current growth season. Moreover, note that the framework presented in this work is flexible in choosing sensitivity indices: any number of growth stages can be defined.

The crop yield will be maximized over a finite prediction horizon with length $N_\mathrm{p}$ using a rolling horizon approach (Camacho & Bordons, 1995), i.e. MPC. For computational reasons, the prediction horizon will likely not cover the entire growth season. Therefore, we rewrite Eq. (23) to:

$$J_\mathrm{ws} = \sum_{\ell=1}^{N_\ell} \Big( \frac{\lambda_\ell}{\sum_{k\in\mathcal{K}_\ell} \hat{E}(\mathring{k})} \sum_{k\in\mathcal{K}_\ell^\mathrm{p}} w_\ell(\mathring{k}) D_\mathrm{v}(\mathring{k}) \Big), \tag{25}$$

where $\mathcal{K}_\ell^\mathrm{p}$ is a subset of days in the prediction horizon during which the crop is in growth stage $\ell \in \{1, \ldots, N_\mathrm{gs}\}$.

### 2.3 Optimal irrigation allocation over growth stages

Anticipating droughts, local water authorities often limit the amount of water that each farmer can use for irrigation in the upcoming growth season or year. However, how to allocate this water over the fields and the growth stages to maximize profit or yield is non-trivial. The allocation does not only depend on maximization of the yield, but also costs such as the fixed costs of planting (or not planting), costs of fertilizer use, and costs of required machinery for irrigation. However, all these costs can simply be included as objectives in the optimization problem, emphasizing the generality of our modeling approach. For the purpose of this work, we only consider the allocation of the available water over all fields to maximize the yield, such that no crops die. This optimal water allocation optimization is performed by the seasonal irrigation planner, before the growth season starts, as the management of the plantation will have to plan ahead the allocation of other resources such as machines and labor. Although the water availability throughout the season could change, in this work we assume that the water availability is known before and will not change (as it is agreed upon in the yearly contract with the local water authority). However, in our formulation we keep track of how much water was applied in past growth stages of each of the fields and their current progress in the season. Therefore, the seasonal irrigation planner can be used to redistribute the water throughout the season given the new water availability constraints. To model the effect of water deficit and stress on yield as accurately as possible, the multiplicative compounding function is used to plan water allocation for a growth season.

Let $\mathcal{F}$ denote the set of fields and let $\hat{W}_{\mathrm{p},\ell,j}$ be the estimate of the potential water use in growth stage $\ell$ of field $j \in \mathcal{F}$. Then, the total yield $Y_\mathrm{tot}$ of a plantation for a growth season is described by the summation of crop yields computed by the multiplicative compounding function:

$$Y_\mathrm{tot} = \sum_{j\in\mathcal{F}} A_j Y_\mathrm{p} \prod_{\ell=1}^{N_\ell} \Big( 1 - \lambda_\ell \Big( 1 - \frac{W_{\mathrm{a},\ell,j}}{\hat{W}_{\mathrm{p},\ell,j}} \Big) \Big), \tag{26}$$

where $A_j$ is the surface area of field $j$, $Y_\mathrm{p}$ the potential yield per area (which is assumed the same for each field), and $W_{\mathrm{a},\ell,j}$ is the actual water use of field $j$ in growth stage $\ell$. Furthermore, the total water use over the growth season is

$$W_\mathrm{tot} = \sum_{j\in\mathcal{F}} A_j \sum_{\ell=1}^{N_\ell} W_{\mathrm{a},\ell,j}. \tag{27}$$

Then, if the plantation is assigned an amount of water $W_\mathrm{max}$ for the growth season, the optimal water allocation problem can be defined as

$$\max_{W_{\mathrm{a},\ell,j}} \quad Y_\mathrm{tot} \tag{28a}$$

$$\text{s.t.} \quad W_\mathrm{tot} \leq W_\mathrm{max} \tag{28b}$$

$$W_{\mathrm{a},\ell,j} \geq T_\ell \hat{W}_{\mathrm{p},\ell,j}, \tag{28c}$$

where $T_\ell \in [0, 1]$ is the minimum relative water use of each growth stage below which the crop suffers terminal moisture stress and dies ($Y_\mathrm{p} = 0$). Note that the optimization problem is nonlinear and non-convex, due to the multiplicative compounding function. Furthermore, the constraint $W_\mathrm{max} \geq \sum_{j \in \mathcal{F}} A_j \sum_{\ell=1}^{N_\ell} T_\ell \hat{W}_{\mathrm{p},\ell,j}$ needs to be satisfied for the problem to be feasible.

A crop kite is a tool that visualizes the space of possible yields and actual water uses when considering the effect the temporal distribution of water use has on the yield (Smilovic et al., 2016). The solutions to the optimization problem from Eq. (28) define the upper bound of the crop kite, see Figure 3. The crop is essentially a convex set that indicates how much effect the temporal water use distribution can have on the yield. For example, when only 70% of the potential water use is available the resulting yield can be anywhere from 22% to 55% of the potential yield, depending on how the water is allocated across the growth stages in the growth season.

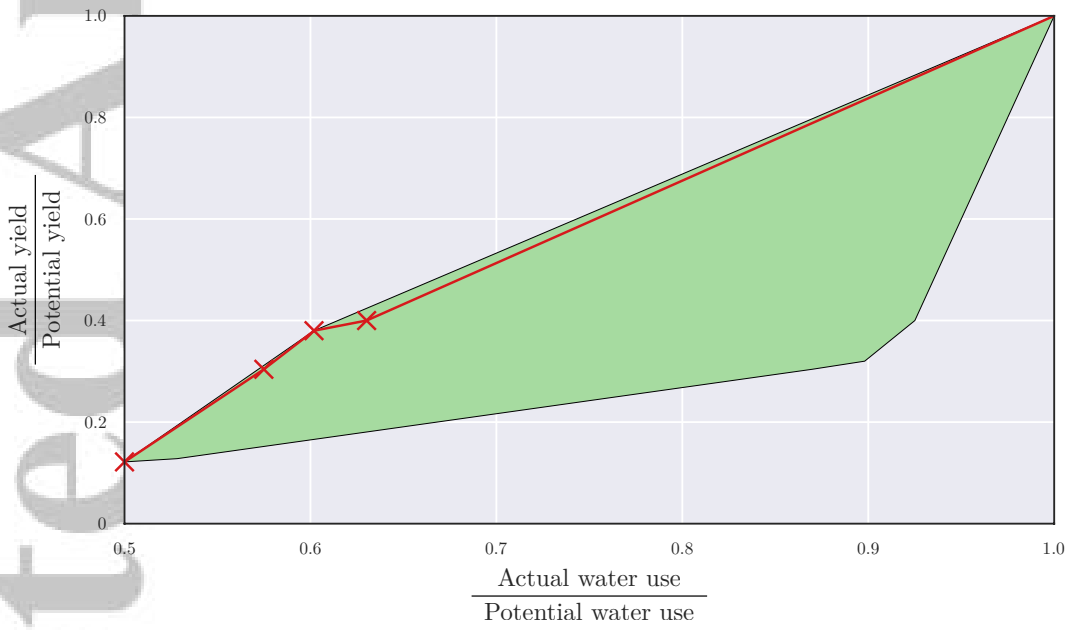


**Figure 3.** In green the crop kite of sugarcane, using crop sensitivity values from Table 1 and $T_\ell = 0.5$, $\forall \ell$. The red crosses and lines represent a piecewise affine approximation of the upper bound of the crop kite using Eq. (32).

The optimization problem from Eq. (28) is non-convex and the difficulty of solving it scales exponentially with the number of fields. Therefore, we approximate it to render our seasonal irrigation planner computationally feasible. First, consider the alternative problem formulation in which we have to allocate water to one growth stage $\ell$ until $W_{\mathrm{a},\ell} = W_{\mathrm{p},\ell}$, then where to allocate the remaining water to another growth stage, and so on. This alternative optimization problem is a scheduling problem: *What is the best order of growth stages to allocate the available water to in order to maximize the yield*? To answer this question, consider the contribution values $c_\ell$ of each growth stage $\ell$ per water unit to the yield:

$$c_\ell = \frac{\lambda_\ell \big(1 - T_\ell\big)}{\hat{W}_{\mathrm{p},\ell} - T_\ell \hat{W}_{\mathrm{p},\ell}} = \frac{\lambda_\ell}{\hat{W}_{\mathrm{p},\ell}}. \tag{29}$$

The ranking of the $c_\ell$ values determines the solution to our alternative problem: each growth stage receives the minimum required amount of irrigation for the crop not to die (feasibility requirement for a solution to this problem), then the growth stage with high-

est contribution $c_\ell$ receives water up to potential water use (or however much water is available), then the remaining water is allocated to the growth stage with second highest contribution value, and so on. Let $\bar{\mathcal{L}}_i$ be the set of growth stages with a contribution value $c_i$ or higher and $\underline{\mathcal{L}}_i$ be the set of growth stages with a contribution value lower than $c_i$, where $i \in \{1, \dots, N_{\text{gs}}\}$. Then, the upper bound of the crop kite is approximated by a set $\mathcal{P}$ of points (one for each growth stage). Each of the points $P(W_i, Y_i) \in \mathcal{P}$ consists of a relative yield coordinate ($Y_i$) and a relative water use coordinate ($W_i$):

$$Y_i = \prod_{\ell \in \underline{\mathcal{L}}_i} \Big(1 - \lambda_\ell (1 - T_\ell)\Big), \tag{30}$$

$$W_i = \frac{\sum_{\ell \in \underline{\mathcal{L}}_i} T_\ell \hat{W}_{\text{p},\ell} + \sum_{\ell \in \bar{\mathcal{L}}_i} \hat{W}_{\text{p},\ell}}{\sum_{\ell=1}^{N_{\text{gs}}} \hat{W}_{\text{p},\ell}}. \tag{31}$$

These points can be calculated a priori for each growth season and are connected with straight lines (as we increase water use for one growth stage at a time) to form a piecewise affine approximation of the upper bound of the crop kite:

$$\frac{Y_{\text{a}}}{Y_{\text{p}}} \approx \begin{cases} \frac{\frac{W_{\text{a}}}{\hat{W}_{\text{p}}} - W_1}{W_2 - W_1} \cdot (Y_2 - Y_1) + Y_1 & \text{if } W_1 \le \frac{W_{\text{a}}}{\hat{W}_{\text{p}}} < W_2, \\ \frac{\frac{W_{\text{a}}}{\hat{W}_{\text{p}}} - W_2}{W_3 - W_2} \cdot (Y_3 - Y_2) + Y_2 & \text{if } W_2 \le \frac{W_{\text{a}}}{\hat{W}_{\text{p}}} < W_3, \\ \vdots & \vdots \\ \frac{\frac{W_{\text{a}}}{\hat{W}_{\text{p}}} - W_{N_{\text{gs}}}}{1 - W_{N_{\text{gs}}}} \cdot (1 - Y_{N_{\text{gs}}}) + Y_{N_{\text{gs}}} & \text{if } W_{N_{\text{gs}}} \le \frac{W_{\text{a}}}{\hat{W}_{\text{p}}} \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{32}$$

This piecewise affine approximation of the upper bound of the crop kite can be calculated for all fields using the potential water use predictions. Furthermore, by using binary variables to model the switching behavior, Eq. (32) can be converted to an MLD system following the method of Bemporad and Morari (1999). Then, denote by $\hat{Y}_{\text{a},j} = A_j \left(\frac{Y_{\text{a}}}{Y_{\text{p}}}\right)_j$ the expected yield of field $j$ using the MLD reformulation of Eq. (32). Consequently, the nonlinear, non-convex optimization problem from Eq. (28) can be reduced to an Mixed-Integer Linear Programming (MILP) problem:

$$\max_{W_{\text{a},j}} \quad \sum_{j \in \mathcal{F}} \hat{Y}_{\text{a},j} \tag{33a}$$

$$\text{s.t.} \quad \sum_{j \in \mathcal{F}} A_j W_{\text{a},j} \le W_{\max} \tag{33b}$$

$$W_{\text{a},j} \ge \sum_{\ell=1}^{N_{\text{gs}}} T_\ell \hat{W}_{\text{p},\ell}, \tag{33c}$$

which can be efficiently solved using state-of-the-art MILP solvers such as CPLEX or Gurobi. The solution can then be used to constrain the water use of each growth stage for each field, in order to maximize the total yield with the available water.

To illustrate the performance of the piecewise approximation, the upper bound is approximated for the crop kite depicted in Figure 3, see the red curve and crosses. Because of the non-linearity of the multiplicative compounding function the contribution values actually change as water is allocated to growth stages. However, in the approximation we assumed fixed contribution values (i.e. average contribution per unit of water). This assumption introduces errors when the contribution values are similar, as then the available water should be divided between the competing growth stages for optimality. This is actually the case for the third and fourth growth stages of the example in Figure 3. Here, the contribution values of the third and fourth growth stage are 0.2026 and 0.2208, respectively. However, the approximation is still good in this case, as the estimation error is small. As a future improvement the water could be split between growth

stages that have similar contribution values to reduce the error, for example: if the contribution values are within a tolerance of 0.05 the remaining water is divided half-half between the two growth stages.

### 2.4 Operational constraints

The feasibility of the calculated daily irrigation schedules depends on the local operational constraints of the plantation. These constraints can be divided in two groups:

- Resource constraints: These constraints indicate a limitation in the availability of certain resources required for irrigation, e.g., availability of human operators or sprinklers. Another example that often constrains the irrigation of the fields, is the maximum amount of water that can be used for irrigation over a growth season or year. This amount is often decided on by local water authorities or the government in anticipation of droughts.
- Hydraulic infrastructure constraints: These constraints define the bounds of the daily amount of irrigation. The pumps, channels, and gates conveying water to the fields have a limited capacity. Note that the hydraulic dynamics of water conveyance and the operating of pumps, valves, and gates is not part of this work and will require further research.

Any type of resource and hydraulic infrastructure constraints can be included in the framework presented in our work. However, for illustrational purposes, we only consider two resource constraints and one hydraulic infrastructure constraint. The resource constraints are the maximum number of fields that can be irrigated in a day and the maximum irrigation water use of the entire season calculated using the approach from Section 2.3. The hydraulic infrastructure constraint defines an upper bound on the daily amount of irrigation that is determined by the maximum gate flow through the inlet to the fields.

#### *2.4.1 Maximum number of fields scheduled for irrigation*

Depending on the water delivery capacity of the hydraulic infrastructure and the daily availability of the aforementioned resources a certain maximum number of fields can be irrigated in a working day. Let $\delta_j(k)$ be the binary decision variable that indicates whether to irrigate a plot $j \in \mathcal{F}$ at day $k$. Moreover, let $I_j^*(k)$ be the desired amount of irrigation. Then, the following mixed-integer linear inequalities (Bemporad & Morari, 1999) will satisfy $[\delta_j(k) = 0] \Rightarrow [I_j(k) = 0]$, $[\delta_j(k) = 1] \Rightarrow [I_j(k) = I_j^*(k)]$:

$$I_j(k) \leq M_j \delta_j(k) \tag{34a}$$
$$I_j(k) \geq m_j \delta_j(k) \tag{34b}$$
$$I_j(k) \leq I_j^*(k) - m_j(1 - \delta_j(k)) \tag{34c}$$
$$I_j(k) \geq I_j^*(k) - M_j(1 - \delta_j(k)), \tag{34d}$$

where

$$M_j \triangleq \max_{k \in \{1, N_\mathrm{p}\}} I_j^*(k) \tag{35}$$
$$m_j \triangleq \min_{k \in \{1, N_\mathrm{p}\}} I_j^*(k). \tag{36}$$

It is assumed that the number of fields that can be irrigated on day $k$ of the prediction horizon, $N_\mathrm{f}(k)$, is known a priori at each optimization iteration. Thus, to respect the maximum number of fields to irrigate the constraint

$$\sum_{j \in \mathcal{F}} \delta_j(k) \leq N_\mathrm{f}(k), \quad \forall k \in \{1, \ldots, N_\mathrm{p}\}, \tag{37}$$

is sufficient.

#### *2.4.2 Restricting water use over growth seasons*

The seasonal irrigation planner solves the optimization problem from Eq. (33) to determine the maximum water use $W_{\mathrm{a},\ell,j}$ for each growth stage $\ell$ for each field $j$. This amount can then be used to constrain the water use for each growth stage of each field. However, the prediction horizon may not extend to the entire growth stage. In that case, the daily irrigation controller cannot oversee the consequences of using too much water at the start of the growth season. Therefore, the following constraint is used for the water use of field $j$ in growth stage $\ell$ in the prediction horizon:

$$\sum_{k \in \mathcal{K}^{\mathrm{p}}_{\ell,j}} \Big( I_j(k) \Big) \leq \Big( W_{\mathrm{a},\ell,j} - W_{\mathrm{prev},\ell,j} \Big) \cdot \frac{\left| \mathcal{K}^{\mathrm{p}}_{\ell,j} \right|}{N^{\mathrm{remain}}_{\ell,j}}, \tag{38}$$

where $W_{\mathrm{prev},\ell,j}$ is the amount of water that has previously been applied to the field in growth stage $\ell$, $\mathcal{K}^{\mathrm{p}}_{\ell,j}$ is the set of days that field $j$ is in growth stage $\ell$ in the prediction horizon, and $N^{\mathrm{remain}}_{\ell,j}$ is the number of remaining days that field $j$ will be in growth stage $\ell$. At the start of each optimization iteration the value of $W_{\mathrm{prev},\ell,j}$ is updated with the irrigation from the previous day(s). In case there is a lot of precipitation at the start of the growth stage, the amount irrigated will be small and $W_{\mathrm{prev},\ell,j}$ will remain relatively small. As a result, more water is available for the rest of the season. Furthermore, if the water availability changes in the season, the seasonal irrigation planner can optimize Eq. (33) again given the new situation and the previously applied water amounts, to obtain a new water distribution schedule for the future growth stages of all fields.

### 2.5 Operational Optimization problem formulation

In the daily irrigation control, minimizing the water stress cost defined in Eq. (25) results in a minimization of the violation $D_{\mathrm{v}}(k)$ for all fields over the prediction horizon. As a result, Eq. (14) can be rewritten into:

$$D_{\mathrm{v}}(k) \geq D(k-1) + E^*(k) - P_{\mathrm{e}}(k) - I(k) - \epsilon(k) \tag{39}$$
$$D_{\mathrm{v}}(k) \geq 0. \tag{40}$$

As the optimization problem is linear, the minimization of $D_{\mathrm{v}}(k)$ implicitly also minimizes $D(k)$; consequently, Eq. (3) can be rewritten as:

$$D(k) \geq D(k-1) + E^*(k) - P_{\mathrm{e}}(k) - I(k), \tag{41}$$
$$D(k) \geq 0. \tag{42}$$

Reformulation of the two state update equations (Eq. (3) and Eq. (14)) into four linear inequalities eliminates the need to model their switching behavior using binary variables, reducing the computational burden significantly. However, the optimization problem is

still an MILP problem, due to the operational irrigation constraints:

$$\min_{\mathbf{I}(k)} J = (1-\beta)J_{\text{ws}} + \beta J_{\text{irr}}, \tag{43a}$$

$$\text{s.t.} \quad J_{\text{ws}} = \sum_{j\in\mathcal{F}}\sum_{\ell=1}^{N_\ell}\Big(\frac{\lambda_\ell}{\sum_{k\in\mathcal{K}_{\ell,j}}\hat{\text{ET}}(\mathring{k})}\sum_{k\in\mathcal{K}_{\ell,N_\text{p},j}} w_{\ell,j}(\mathring{k})D_{\text{v,j}}(\mathring{k})\Big), \tag{43b}$$

$$J_{\text{irr}} = \sum_{j\in\mathcal{F}}\sum_{k=1}^{N_\text{p}}\gamma(k)I_j(k), \tag{43c}$$

$$\mathbf{D}(k) \geq \mathbf{D}(k-1) + \mathbf{E}^*(k) - \mathbf{P}_\text{e}(k) - \mathbf{I}(k),\ \forall k \in \{1, N_\text{p}\}, \tag{43d}$$

$$\mathbf{D}_\text{v}(k) \geq \mathbf{D}(k-1) + \mathbf{E}^*(k) - \mathbf{P}_\text{e}(k) - \mathbf{I}(k) - \boldsymbol{\epsilon}(k),\ \forall k \in \{1,\ldots,N_\text{p}\}, \tag{43e}$$

$$0 \leq \boldsymbol{\epsilon}(k) \leq p(\boldsymbol{\theta}_\text{FC} - \boldsymbol{\theta}_\text{WP})\mathbf{Z}_\text{r}(k),\ \forall k \in \{1,\ldots,N_\text{p}\}, \tag{43f}$$

$$\mathbf{D}(k) \geq 0,\ \forall k \in \{1,\ldots,N_\text{p}\}, \tag{43g}$$

$$\mathbf{D}_\text{v}(k) \geq 0,\ \forall k \in \{1,\ldots,N_\text{p}\}, \tag{43h}$$

the operational irrigation constraints in Eqs. (34)–(37),
and the water availability constraint from Eq. (38),

where $\beta \in [0,1]$ is a weight factor indicating the relative importance of water use minimization with respect to water stress minimization and $\gamma(k) > 0$ is a weight factor indicating the relative importance of irrigating earlier or later in the prediction horizon. Furthermore, $D_{\text{v},j}(k)$ is the RZD upper bound violation on field $j$. Note that the bold symbols (e.g., $\mathbf{D}(k)$) denote vectors containing the respective variables for each field in the plantation. The formulated MILP problem can be efficiently solved using state-of-the-art MILP solvers such as CPLEX or Gurobi.

## 3 Model Identification and Validation

The control approach proposed in our work will be evaluated on a simulation of a specific furrow-irrigated area of the Xinavane Sugar Estates (Tongaat Hulett) in Xinavane, Mozambique (see Figure 4). The sugarcane plantation is competing with the wa-

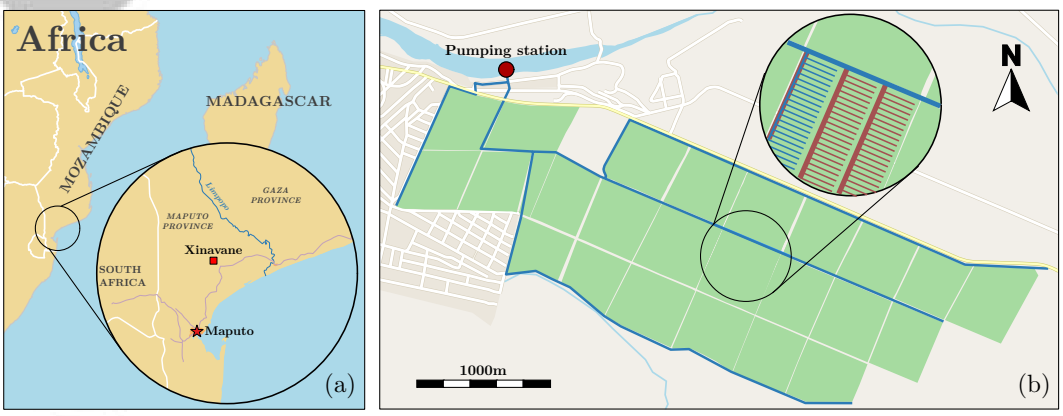


**Figure 4.** In (a) the geographical location of the study area is depicted. In (b) the layout of the fields subject in this work is depicted. A pumping station pumps water out of the river into a system of concrete canals from which the water is conveyed into earth canals by opening gates. From the earth canals (that are adjacent to the fields) the water is siphoned into the furrows of the fields.

ter needs of a rapidly urbanizing population downstream, increasing agricultural and industrial water uses under the stresses of climate change (Gelcer et al., 2018). Furthermore, storing water is not possible and the large-scale plantations consist of multiple clusters of fields, which each have their own (downstream) pumping stations that draw water form the same river; therefore, Tongaat Hulett has the objective to use as little possible as water while maximizing yield. The selected area has a size of approximately 420 ha divided over 24 plot served by a network of open canals. Each plot consists of 2-4 fields, adding up to 79 fields in the selected area. There is an upstream gate and canal present for each of the fields, so they can be irrigated separately. The fields are currently irrigated every 3-4 days by human operators, the net irrigation requirement is approximately 1650 mm per year, and the average precipitation in the area is approximately 650 mm per year. Through a network of open canals each field is provided water by an upstream gate with a maximum capacity of 60 liters per second. This maximum flow to the field together with the 8 hour working day of the operators, constrain the maximum amount of irrigation that can be applied to each plot. Each of the fields considered in this work is modeled as an agro-hydrological system in AquaCrop-OS, using soil and crop parameters that are representative of the in-field situation in Xinavane (and that are taken the same for each field). Note that lateral dynamics are not included in the modeling approach: so fields have no effect on adjacent fields. Furthermore, the rainfall and temperature is assumed equivalent for each field, as the fields are located in the same geographically small (same weather) area. Note that the application efficiency of furrow irrigation is incorporated in the identification process of the water balance model used in the daily irrigation controller; therefore, our method is not restricted to a specific infield irrigation strategy as we can account for the different application efficiencies in the MPC model. The complete set of parameters of the soil, the crop, and the field management settings used for modeling the fields can be found in the supplementary material.

### 3.1 Data for agro-hydrological modeling

For each field in the case study area, the recorded planting and harvest dates from 2013-2017 were used. Furthermore, from conversations with local operators and management the local irrigation practice can be summarized as follows:

- After initially planting the sugarcane (or shoots in case of ratooning), the (dry) field is irrigated to field capacity to promote the germination, this initial irrigation cycle can take up to 3 days.
- Next, over the course of the growth season the field is irrigated to field capacity with an irrigation frequency of once every 4 days.
- Finally, when the sugarcane has matured, the field is harvested and after approximately one week the field is replanted and the next growth season commences.

Unfortunately, accurate and sufficiently detailed data on local irrigation records and soil moisture measurements were not available. Therefore, using the aforementioned irrigation practice, real historic growth season weather data, and the calibrated soil and crop parameters (e.g., canopy decline/growth coefficients, soil evaporation coefficients and soil composition, see the supplementary material), the growth seasons are *recreated* using simulations in AquaCrop-OS. The resulting data set of 316 growth seasons spread over 5 years and 79 fields can then be used for identification of the proposed model. This data set consists of evapotranspiration data, canopy cover data, soil moisture data, rooting depth data, and yield data.

### 3.2 Identifying and validating the evapotranspiration model

The evapotranspiration model (see Eq. (11)) uses an estimate of the size of the canopy cover as a function of thermal time (in the form of an NDVI-CGDD curve, see Eqs. (9)

and (11)) to fit the coefficients $\alpha_1$, $\alpha_2$ and $\alpha_3$ using the evapotranspiration dataset generated in AquaCrop-OS. However, AquaCrop-OS does not generate NDVI data; instead canopy cover data is directly used to estimate the size of the canopy cover as a function of the thermal time. The average evolution of the canopy cover over thermal time is estimated using a random half of the generated canopy cover data set. During each growth season the evolution of the canopy cover will be slightly different, due to different atmospheric conditions. To obtain a smooth curve that best describes the average canopy cover development, we use the following procedure:

- Divide the canopy cover data points into $N_g$ groups, based on evenly spaced sections of the thermal time. This number of groups depends on the density of the data: the more data points the more groups can be used.
- For each group calculate the average thermal time and the average size of the canopy cover; this is a new data point. Therefore, the number of groups should be chosen such that the distance between the new data points is minimized, while still obtaining a smooth curve.
- Finally, draw a curve through the new averaged data points. We use a Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) (Fritsch & Butland, 1984) that interpolates the data locally using a cubic function. Note that quadratic or linear interpolation can also be used. In this case, cubic interpolation was chosen as a compromise between number of coefficients and accuracy of the fit.

The resulting size of the canopy cover when using $N_g = 100$ is depicted in Figure 5. In Figure 5 (a), the raw datapoints of the size of the canopy cover as a function of thermal time are depicted, that are obtained by simulating local practice of irrigating to full capacity every 4 days. Note the enlarged box illustrating the (tight) spread of the data. In Figure 5 (b) the resulting PCHIP curve is plotted that connects the decimated data points. The curve is smooth without any sudden jumps.

Next, the canopy cover PCHIP approximation, temperature data, and reference evapotranspiration data from the identification half of the data set are use to estimate the model coefficients from Eq. (11) using linear least-squares. The model coefficients are determined as $[\alpha_1,\alpha_2,\alpha_3] = [-3.2677{\cdot}10^{-4}, 1.1247, 0.8918]$ with an Root Mean Square Error (RMSE) of 0.55 mm and a Variance Accounted For (VAF) of 93.58%. Furthermore, in Figure 5 (c) the spread of the modeling error as a function of the thermal time for the validation set is depicted. Here, the modeling error (ME) is defined as:

$$\mathrm{ME}(k) = \frac{\hat{E}(k) - E^*(k)}{E^*(k)} \cdot 100\%. \tag{44}$$

The evapotranspiration model performs best when the canopy cover reaches its maximum size. Conversely, when the size of the canopy cover is low and when the crop starts to age the contribution of the evaporation to the overall evapotranspiration is high. The evaporation dynamics are not well captured as wetting events are not included in the modeling. However, the model is still able to provide estimates for the evapotranspiration within a 25% error margin for over 75% of the growth season. Furthermore, sugarcane is relatively insensitive to water stress in the last growth stage (Doorenbos & Kassam, 1979; Robertson & Donaldson, 1998); therefore, modeling errors in this growth stage will likely not have a large influence on the controller's performance.

We note that the spread of the canopy cover data points is rather small and unrealistic in Figure 5a. To illustrate that the proposed method also works in the real-world, we have also constructed an NDVI–CGDD curve using real NDVI data for the Xinavane area, which is shown in the supporting material for this manuscript.

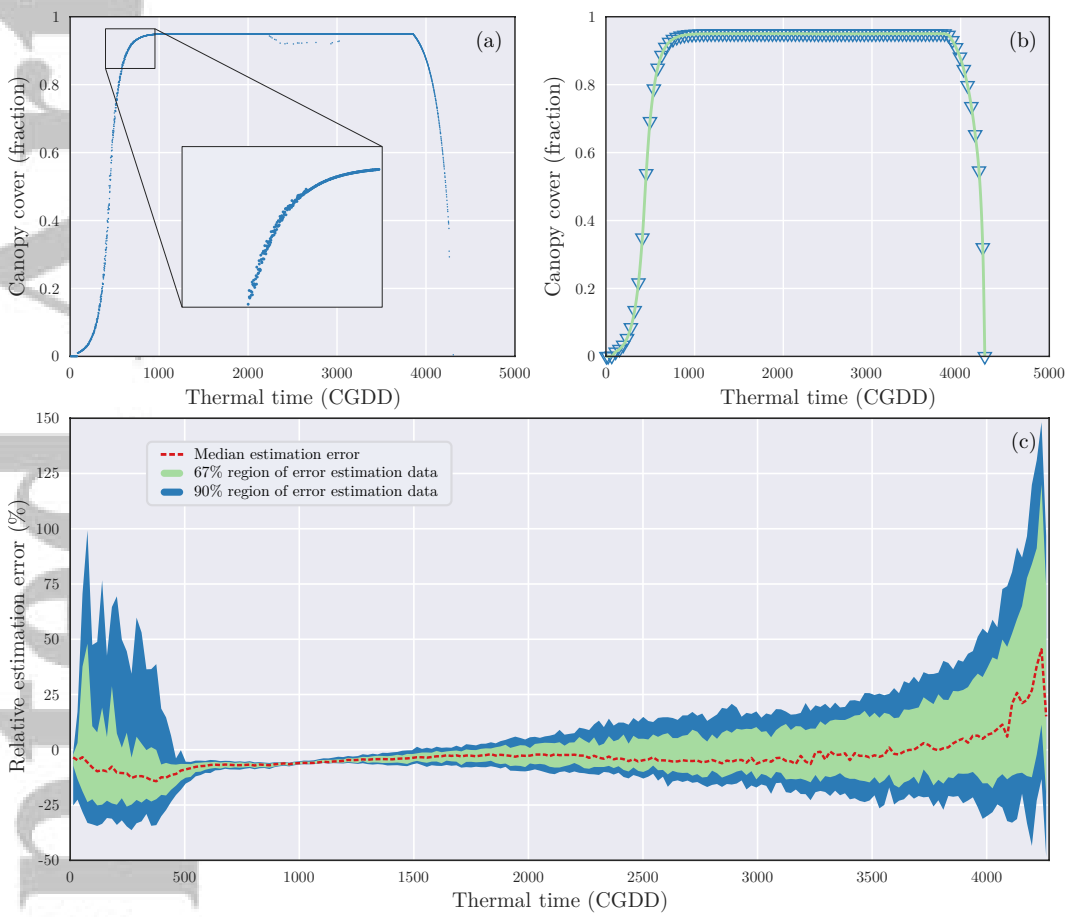


**Figure 5.** In (a) the canopy cover data points are depicted. In (b) the decimated data points (triangles) and the resulting PCHIP curve are shown. In (c) the modeling errors of the evapotranspiration model are depicted as a function of the thermal time.

## 4 Simulation Results and Discussion

To evaluate the performance of the proposed feedback control scheme, we conduct closed-loop simulations with the complex crop-water productivity model AquaCrop-OS. Although our control model uses a simple water balance (see Eq. (1)) with only a non-zero drainage term when the soil becomes saturated (see Eq. (3)), the AquaCrop-OS model does include more complex dynamics; for example filtration is modeled using spatial discretization of the soil layer(s) (Raes et al., 2009). The sufficiency of a simple control model is shown in the closed-loop performance of three complex system simulations on AquaCrop-OS in Sections 4.2-4.4. In the first simulation, we investigate the effect of decreasing the water availability on the yield, when using the proposed control scheme. In the second simulation, a comparison is made between the performance of the local irrigation approach and the proposed MPC approach in terms of water productivity and yield. In the last simulation, the water saving capabilities of the MPC approach are compared with that of the local irrigation approach by constraining the amount of irrigation for each growth season.

### 4.1 Simulation settings

The goal of the simulations is to investigate the maximum (theoretical) benefit that can be achieved by using an optimal control approach. Therefore, no stochastics are added to the simulations and perfect root zone soil moisture measurements for all fields are assumed to be available at the start of each day. In the preseason, the soil moisture measurements that the daily irrigation controller receives are at an depth of 30 cm, the depth at which the roots initially establish after planting. Furthermore, a perfect forecast of the evapotranspiration, temperature, rainfall, and planting and harvest dates is assumed to be available to the daily irrigation controller. Conversely, the controller imitating the local irrigation practice has information on the exact amount of irrigation required to irrigate to field capacity and applies this to the field, provided it does not violate the maximum amount of irrigation. To that end, each field is provided water by an upstream gate with a maximum capacity of 60 liters/second. This maximum flow to the field together with the 8 hour working day of the operators, constrain the maximum amount of irrigation that can be applied to each field. The sensitivity indices and growth stage definitions that the two-level approach uses are listed in Table 1. The sensitivity indices are derived from (Moutonnet, 2002), the growth stage lengths in number of days are retrieved from the provided example files for sugarcane in AquaCrop-OS, and the thermal times are calculated using Eq. (4) with a base temperature and maximum temperature of 12° and 32° Celsius, respectively, and historic data from Xinavane.

**Table 1.** Growth stage definitions used for simulations.

| Growth stage | Length (days) | Length (CGDD) | $\lambda_\ell$ |
|---|---|---|---|
| Establishment | 7 | 82 | 0.4[a] |
| Tillering | 53 | 618 | 1.2[a] |
| Yield formation | 270 | 3151 | 1.2[a] |
| Ripening | 35 | 408 | 0.1[b] |

[a] These values are taken from Moutonnet (2002).

[b] This value is taken from Doorenbos and Kassam (1979).

All simulations start on the 30$^{\text{th}}$ of May 2013, 7 days before the first field is planted and end on the 31st of December 2017. In this period, all fields will have had four full growth seasons. The relative cost of irrigation to water stress is set to $\beta = 10^{-7}$, a value small enough to ensure the controller irrigates whenever necessary to avoid violating the

RZD upper bound and big enough to provide an incentive to not waste water. Moreover, irrigation should be delayed if it is not directly required to avoid violating the defined soil moisture bounds. To provide this information to the daily MPC controller, we penalize irrigation at the end of the prediction with $\gamma(N_\mathrm{p}) = 1$ and with $\gamma(1) = 2$ (see Eq. (43c)) at the beginning of the horizon and linearly interpolate the cost between the two. Finally, the fraction of the available water at which the crop starts to experience stress is set to $p = 25\%$.

An important remark is that the daily MPC controller can irrigate in the fields in the preseason, if necessary. However, no costs are included in the optimization problem to penalize soil moisture deficits in the preseason; therefore, if the controller decides to irrigate a field in the preseason it is only in anticipation of a soil moisture threshold violation after the field has been planted. Furthermore, as mentioned above irrigation is delayed unless necessary to avoid violating the RZD bounds. Therefore, the daily irrigation controller will only irrigate in the preseason if the field is too dry before planting or when the daily irrigation controller is constrained by an irrigation frequency (which is the case in simulations 2 and 3 in Section 4.3 and Section 4.4, respectively). Note that if the daily irrigation controller applies irrigation in the preseason the used water is added to the total water use in the upcoming season in all simulations.

### 4.2 Simulation 1: Effect of decreasing water availability on yield

The effect of water stress on yield (summarized in yield sensitivity indices) is often based on open-loop predictions and regression methods. Therefore, in this first simulation we investigate the effects of restricting water use on closed-loop results. First, we establish the potential water use when using full (optimal) irrigation: each of the fields can be irrigated daily, only limited by the maximum amount of water that can be provided by the upstream gate and the length of a working day. Ideally, the prediction horizon covers the whole growth season; however limited by the available memory on the computer used for simulations we set the length of the prediction horizon to $N_\mathrm{p} = 100$ days (as large as possible). As the irrigation frequency is not bounded for this simulation, the constraints from Eq. (34) are relaxed (removing all integer variables from the optimization problem) and the problem is reduced to a continuous linear programming problem. The results indicate, for each field, how much water is required in each growth stage for maximum yield. Then, for a fraction of this potential water use we calculate the optimal water allocation for each growth season of each individual field using Eq. (32). These amounts are then used to constrain the water use in each growth stage of each field. The result is an approximation of the upper bound of the crop kite (from Figure 3), see the beeswarm plot in Figure 6 (a). The yield declines progressively as the water is increasingly restricted. In Figure 6 (b), the rainfall and evapotranspiration data is depicted for each of the seasons. The low amount of rainfall in season 3 explains the rapid yield decline when reducing actual water use: the decrease in amount of irrigation is a larger proportion of the total seasonal water in dry seasons than rainy seasons. The reason that season 1 maintains a higher yield under water stress than the other seasons is that the soil moisture is initialized at field capacity for all fields. Therefore, a buffer of water is retained in the soil below the crop's roots, which is tapped into as the root depth increases. After the first season, this extra water source can be depleted.

### 4.3 Simulation 2: MPC vs. local irrigation practice

In this second simulation, we compare the performance of the proposed MPC approach with that of the irrigation practice described in Section 3.1. In this local irrigation practice of Mozambique, the fields are irrigated every four days. Constraints are added on the number of fields that can be irrigated each day (Eq. (37)) to ensure the daily MPC controller does not irrigate more frequently than the local practice. Furthermore, both

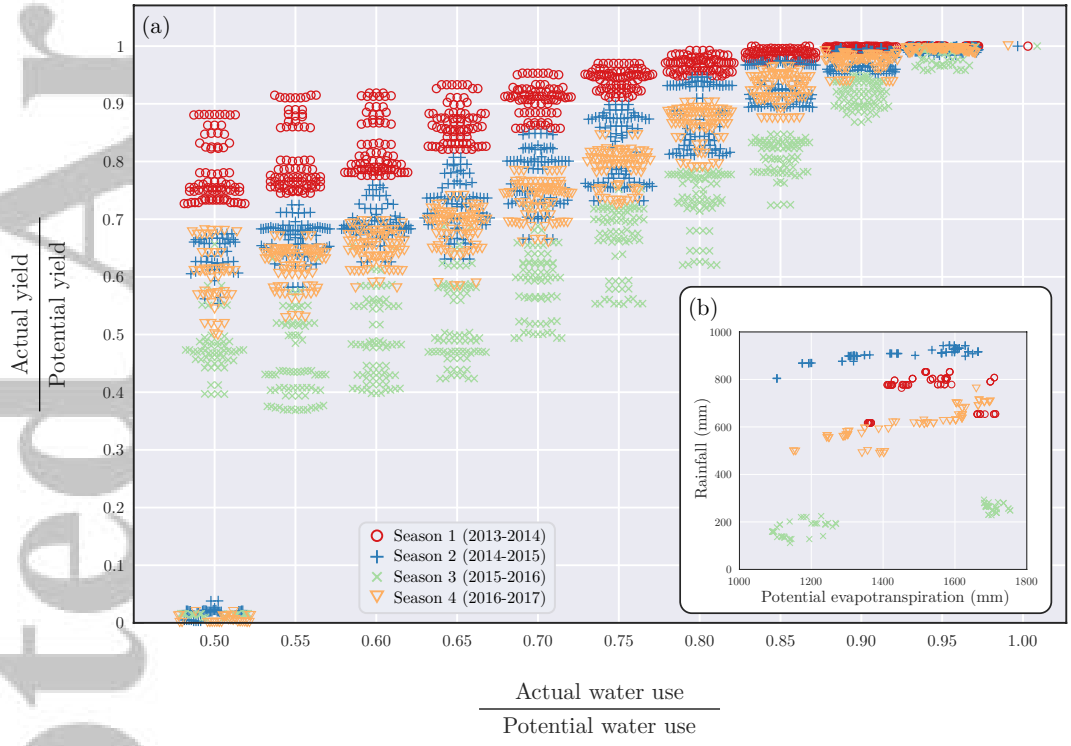


**Figure 6.** In (a) the effect of decreasing water use on the yield is depicted in beeswarm plots; for each restricted water use on the horizontal axis a collection of points is depicted that represent the relative yields of each field in the seasons. This figure shows the reduction in yield we can expect when water use is reduced under different atmospheric conditions. To clarify the spread of the results the potential evapotranspiration is plotted against the rainfall in (b).

the MPC and the local irrigation approach have no maximum water use constraint in this simulation.

The optimization problem that is repeatedly solved by the daily MPC controller at each time step in simulations 2 and 3 belong to the class of mixed-integer linear programs (MILPs), which we solve using Gurobi (v8.1.1) (Gurobi Optimization, 2015) accessed via its MATLAB interface. Although Gurobi can find solutions with arbitrarily small optimality gaps to the globally optimal solution, the computational effort (measured in CPU time) can become too large based on problem size and the prescribed optimality gap. To complete our simulations to 'near' optimality within a reasonable time frame, we set the prediction horizon length to $N_{\mathrm{p}} = 60$ days and set the relative optimality gap tolerance to 1%. To account for pathological cases we limit the maximum computation time to 3600 seconds. This prediction horizon is still large enough to cover the first two growth stages completely (7 and 53 days, respectively).

In Figure 7, we show the amount of water used for irrigation in all four seasons against the final attained yields for all fields. For all fields, the MPC controller uses less water while producing an equal or higher yield than the local irrigation practice. The amount of water that the MPC controller can save with respect to the local practice seems to be proportional to the amount of rainfall in a season. This is especially noticeable in Figure 7 (c), where the difference in water use is small and the season was relatively dry (see Figure 6 (b)).

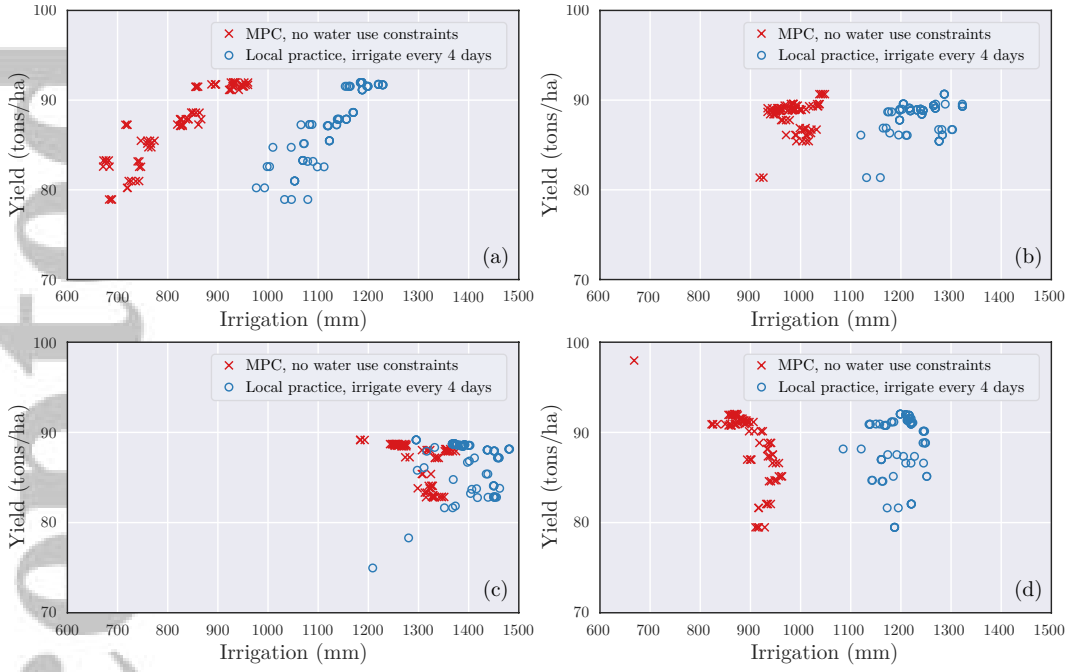


**Figure 7.** A comparison of the water productivity of the MPC controller and the local irrigation practice for each of the four growth seasons. In (a), (b), (c), and (d) the results of, respectively, season 1, 2, 3, and 4 are shown.

### 4.4 Simulation 3: Irrigation under water scarcity

To illustrate the water saving capabilities of our MPC approach over the local approach of simply irrigating every few days (described in Section 3.1), we now consider a simulation scenario in which available water volume is limited. As opposed to the first

simulation, we limit the total available water volume of all the fields *combined* to 80% of the potential total water use (established by the results from the first simulation). Therefore, in this simulation we solve the optimization problem from Eq. (33) for all fields at once (for one growth season at a time). The result is a maximum water volume per growth stage per field for a single growth season, such that the overall crop yield of that season is maximized. These maximum water volumes are then converted to irrigation depths using the surface areas of the fields. Finally, the irrigation depth per season per field is constrained using the inequality constraint from Eq. (38). Note that, similarly to the second simulation, the irrigation frequency is limited by imposing constraints on the number of fields that can be irrigated in a day.

When the real plantation in Xinavane faces drought, the local staff reduces the irrigation frequency in an attempt to reduce the water use. The workers always irrigate the fields to field capacity, in order to minimize the number of revisits to each field. We compare the water use of the local irrigation practice with an irrigation frequency of 4, 5, 6, and 7 days with that of the MPC controller. The results are listed in Table 2. Only the MPC controller is able to meet the water volume constraints; the local approach uses too much water in all of the seasons for all the irrigation frequencies. Note that even though the MPC controller irrigates in the preseason the total amount of irrigation for the season is not exceeded.

**Table 2.** Actual water use ($\times 10^6$ m$^3$) of all fields combined for each season. Note that in the MPC results the preseason amounts of irrigation are added to the total water use of the subsequent season.

| | Season 1 | Season 2 | Season 3 | Season 4 |
|---|---|---|---|---|
| Constrained water volume | 2.98 | 3.65 | 4.88 | 3.46 |
| Local practice, 4 day irrigation frequency | 5.16 | 5.64 | 6.38 | 5.47 |
| Local practice, 5 day irrigation frequency | 4.97 | 5.39 | 6.15 | 5.27 |
| Local practice, 6 day irrigation frequency | 4.74 | 5.18 | 5.83 | 4.85 |
| Local practice, 7 day irrigation frequency | 4.56 | 4.94 | 5.43 | 4.40 |
| MPC, 4 day irrigation frequency | 2.80* | 3.58* | 4.75* | 3.39* |

* Feasible seasons.

## 5 Conclusions and Future Research

The optimization problem of allocating water over multiple fields of a plantation in order to maximize the seasonal yield or profit is highly complex, as we have to consider many factors, such as different crop stages, restrictions on seasonal water volume, available machinery (e.g. sprinklers), and human operators. In this work, we propose a methodology to reduce this complex problem into two separate optimal control problems, which are solved using a two-level structure consisting of a seasonal irrigation planner and a daily MPC irrigation controller. The seasonal irrigation planner considers the allocation of the available water of the upcoming growth season over the crop stages of each field, such that the overall yield is maximized. For the optimization of this water allocation, we assume that within a crop stage the allocated water is distributed perfectly temporally and spatially such that the water stress is minimized. This minimization of the daily water stress is realized by an MPC controller that decides on the daily irrigation amounts of each field, such that the water stress is minimized for all fields, while not exceeding the total water amounts that were allocated for each field by the seasonal irrigation planner. Although the realized precipitation and temperature in the season may be different from those of the optimal plan proposed by the seasonal irrigation sched-

uler, the MPC controller receives daily measurements and consequently updates it's policy accordingly to correct for the deviations from the optimal plan.

For the seasonal irrigation planner we describe the total expected yield over multiple fields by the summation of crop yields computed by the multiplicative compounding yield function from (Raes et al., 2006). As this function is highly nonlinear and nonconvex, the difficulty to find optimal solutions scales exponentially with the number of fields. Therefore, we approximate it using an alternative problem formulation. The result is an efficient formulation of the seasonal irrigation planner that determines the optimal water distribution over the fields to maximize the overall yield, subject to local operational constraints such as restrictions on seasonal water volume or fertilizer amounts.

The MPC controller minimizes the daily water stress by scheduling irrigation such that the soil moisture of the fields is regulated within a water-stress-free zone. This requires a model of the interaction between the soil, the atmosphere, and the crop. A simple water balance model is created for which the saturation dynamics are modeled explicitly using conditionally switched depletion dynamics. Furthermore, evapotranspiration is predicted using a linear model based on remote-sensing measurements of the canopy cover, temperature data, and evaporation and transpiration data obtained from closed-loop simulations in AquaCrop-OS. The effect of water stress on the crop yield is included in the cost function of the daily irrigation controller using an additive crop production function. This provides the daily controller with information on which fields have priority in receiving water. The two-level approach presented can handle resource and hydraulic infrastructure constraints. Therefore, the approach is generic as it is not restricted to a specific irrigation method, crop, soil type, or local environment.

The performance of our two-level approach is evaluated on three closed-loop simulations in AquaCrop-OS of a sugarcane plantation (from Tongaat Hulett) located in Xinavane, Mozambique. The parameters in AquaCrop-OS are set to match the parameters of the real plantation. Then, using the local standard irrigation practice (irrigate each field every few days) from the Xinavane sugarcane plantation the growth seasons from 2013 till 2017 are recreated. This data set is then used to identify the water balance model parameters and evapotranspiration model parameters. The goal of the simulations is to investigate the maximum (theoretical) benefit that can be achieved using our optimal control approach.

In the first simulation, we study the effect of water scarcity on the season yield under our proposed two-level approach. The results give an indication of the decline of yield as a function of water availability that can be expected. This can be useful information for decision makers to decide on the water distribution of upcoming seasons. In the second simulation, the water productivity of the proposed two-level approach is compared to a simple heuristic of irrigating the fields to field capacity every four days. Our two-level approach achieves the same seasonal yield as the local irrigation practice, while using up to 30% less water. In the third simulation, the total water use is restricted to 80% of the potential water use and the capabilities of the local irrigation practice and our optimal control approach in producing feasible irrigation schedules are tested. Our control approach is able to satisfy the water use constraint while maximizing the water productivity. The local approach of irrigation every few days to field capacity was not able to constrain seasonal water use to a feasible amount.

A next step towards implementation of closed-loop algorithms such as the one proposed in our paper, is creating reliable observers for soil moisture. Such an observer could be designed by combining in situ and satellite remote-sensing data approaches to create soil moisture measurements, which can then be incorporated into the control loop. Furthermore, our future work will explore stochastic MPC methods for hedging land, water (and labor) allocation against uncertainty in water availability within the proposed framework. Another step to take for implementation is to evaluate the effect of the un-

certainty of rainfall. However, under sparse rainfall events the effect of the rainfall uncertainty on the performance will likely be minimal as the controller can change the irrigation schedules following a measurement after a rainfall event has occurred. Nonetheless, the effect that uncertainty of rainfall has on irrigation control will need to be researched more in depth to render our approach feasible for a broader range of climates.

## Acronyms

**CGDD** Cumulative Growing Degree Days
**FAO** Food and Agriculture Organization of the United Nations
**GDD** Growing Degree Days
**MLD** Mixed-Logic Dynamical
**MPC** Model Predictive Control
**MILP** Mixed-Integer Linear Programming
**NDVI** Normalized Difference Vegetation Index
**PCHIP** Piecewise Cubic Hermite Interpolating Polynomial
**RMSE** Root Mean Square Error
**RZD** Root Zone Depletion
**TAW** Total Available Water
**VAF** Variance Accounted For

## Acknowledgments

The authors want to thank Tongaat Hulett for providing insight into the daily operation of their sugarcane plantations during multiple visits in 2018 and 2019, specifically the agricultural department of Tongaat Hulett Xinavane. Furthermore, we would like to thank the Netherlands Enterprise Agency for subsidizing the IWACA-Tech project, which this work is part of. The settings and parameters used for (initializing) AquaCrop-OS are available from a publicly accessible repository, see: `http://doi.org/10.4121/uuid:4fb3a35f-1786-45ee-a2f8-65a391fa86d0`.

## References

Allen, R., Pereira, L., Raes, D., & Smith, M. (1998). Crop evapotranspiration-guidelines for computing crop water requirements. *Journal of Irrigation and Drainage*, *300*(9), D05109.

Bemporad, A., & Morari, M. (1999). Control of systems integrating logic, dynamics, and constraints. *Automatica*, *35*(3), 407–427.

Bras, R., & Cordova, J. (1981). Intraseasonal water allocation in deficit irrigation. *Water Resources Research*, *17*(4), 866–874.

Burek, P., Satoh, Y., Fischer, G., Kahil, M., Scherzer, A., Tramberend, S., ... Wiberg, D. (2016). *Water Futures and Solution - Fast Track Initiative* (IIASA Working Paper). Laxenburg, Austria.

Camacho, E., & Bordons, C. (1995). *Model Predictive Control in the Process Industry*. Springer.

Champagne, C., Berg, A., Belanger, J., McNairn, H., & De Jeu, R. (2010). Evaluation of soil moisture derived from passive microwave remote sensing over agricultural sites in canada using ground-based soil moisture monitoring networks. *International Journal of Remote Sensing*, *31*(14), 3669–3690.

Delgoda, D., Saleem, S., Malano, H., & Halgamuge, M. (2016). Root zone soil moisture prediction models based on system identification: Formulation of the theory and validation using field and AQUACROP data. *Agricultural Water Management*, *163*, 344–353.

Doorenbos, J., & Kassam, A. (1979). Yield response to water. *Irrigation and Drainage Paper*, *33*.

Dudley, N., Howell, D., & Musgrave, W. (1971). Optimal intraseasonal irrigation water allocation. *Water Resources Research*, *7*(4), 770–788.

Estes, J., Jensen, J., & Tinney, L. (1978). Remote sensing of agricultural water demand information: A California study. *Water Resources Research*, *14*(2), 170–176.

Foster, T., Brozović, N., Butler, A., Neale, C., Raes, D., Steduto, P., ... Hsiao, T. (2017). AquaCrop-OS: An open source version of FAO's crop water productivity model. *Agricultural Water Management*, *181*, 18–22.

Fritsch, F., & Butland, J. (1984). A method for constructing local monotone piecewise cubic interpolants. *SIAM Journal on Scientific and Statistical Computing*, *5*(2), 300–304.

Gelcer, E., Fraisse, C., Zotarelli, L., Perondi, D., Malia, A., Ecole, C., & Migliaccio, K. (2018). A smart irrigation tool to determine the effects of ENSO on water requirements for tomato production in Mozambique. *Water*, *10*.

Georgiou, P., & Papamichail, D. (2008). Optimization model of an irrigation reservoir for water allocation and crop planning under various weather conditions. *Irrigation Science*, *26*(6), 487–504.

Goswami, S., Gamon, J., Vargas, S., & Tweedie, C. (2015). *Relationships of NDVI, Biomass, and Leaf Area Index (LAI) for six key plant species in Barrow, Alaska* (Tech. Rep.). PeerJ PrePrints.

Gurobi Optimization, I. (2015). Gurobi optimizer reference manual. *URL http://www. gurobi. com*.

Hanks, R. (1974). Model for predicting plant yield as influenced by water use. *Agronomy Journal*, *66*(5), 660–665.

Jensen, J. (1968). Water consumption by agricultural plants. *Water Deficit and Plant Growth*, *2*, 1–22.

Jensen, M., Wright, J., & Pratt, B. (1971). Estimating soil moisture depletion from climate, crop and soil data. *Transaction of the ASAE*, *14*(5), 954–959.

Kamble, B., Kilic, A., & Hubbard, K. (2013). Estimating crop coefficients using remote sensing-based vegetation index. *Remote Sensing*, *5*(4), 1588–1602.

Mao, Y., Liu, S., Nahar, J., Liu, J., & Ding, F. (2018). Soil moisture regulation of agro-hydrological systems using zone model predictive control. *Computers and Electronics in Agriculture*, *154*, 239–247.

Martens, B., Gonzalez Miralles, D., Lievens, H., Van Der Schalie, R., De Jeu, R., Fernández-Prieto, D., ... Verhoest, N. (2017). GLEAM v3: satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, *10*(5), 1903–1925.

Mayne, D., Rawlings, J., Rao, C., & Scokaert, P. (2000). Constrained model predictive control: Stability and optimality. *Automatica*, *36*(6), 789 - 814. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0005109899002149` doi: https://doi.org/10.1016/S0005-1098(99)00214-9

McCarthy, A., Hancock, N., & Raine, S. (2014). Simulation of irrigation control strategies for cotton using model predictive control within the VARIwise simulation framework. *Computers and Electronics in Agriculture*, *101*, 135–147.

McMaster, G., & Wilhelm, W. (1997). Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology*, *87*(4), 291–300.

Moutonnet, P. (2002). Yield response factors of field crops to deficit irrigation. In *Deficit Irrigation Practices.*

Mulla, D. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*, *114*(4), 358–371.

Nahar, J., Liu, S., Mao, Y., Liu, J., & Shah, S. (2019). Closed-loop scheduling and control for precision irrigation. *Industrial & Engineering Chemistry Research*.

Park, Y., Shamma, J., & Harmon, T. (2009). A receding horizon control algorithm for adaptive management of soil moisture and chemical levels during irrigation. *Environmental Modelling & Software*, *24*(9), 1112–1121.

Priestley, C., & Taylor, R. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, *100*(2), 81–92.

Protopapas, A., & Georgakakos, A. (1990). An optimal control method for real-time irrigation scheduling. *Water Resources Research*, *26*(4), 647–669.

Raes, D., Geerts, S., Kipkorir, E., Wellens, J., & Sahli, A. (2006). Simulation of yield decline as a result of water stress with a robust soil water balance model. *Agricultural Water Management*, *81*(3), 335–357.

Raes, D., Steduto, P., Hsiao, T., & Fereres, E. (2009). AquaCropâĂŤThe FAO crop model to simulate yield response to water: II. Main algorithms and software description. *Agronomy Journal*, *101*(3), 438–447.

Raes, D., Steduto, P., Hsiao, T., & Fereres, E. (2018). AquaCrop Version 6.0-6.1 Reference Manual [Computer software manual]. Rome, Italy.

Robertson, M., & Donaldson, R. (1998). Changes in the components of cane and sucrose yield in response to drying-off of sugarcane before harvest. *Field Crops Research*, *55*(3), 201–208.

Saleem, S., Delgoda, D., Ooi, S., Dassanayake, K., Liu, L., Halgamuge, M., & Malano, H. (2013). Model predictive control for real-time irrigation scheduling. *IFAC Proceedings Volumes*, *46*(18), 299–304.

Smilovic, M., Gleeson, T., & Adamowski, J. (2016). Crop kites: Determining crop-water production functions using crop coefficients and sensitivity indices. *Advances in Water Resources*, *97*, 193–204.

Steduto, P., Hsiao, T. C., Fereres, E., Raes, D., et al. (2012). *Crop yield response to water* (Vol. 1028). fao Rome.

Stewart, J., Cuenca, R., Pruitt, W., Hagan, R., & Tosso, J. (1977). Determination and utilization of water production functions for principal California crops. *W-67 CA Contributing Project Report, University of California, Davis, USA*.

Trout, T., Johnson, L., & Gartung, J. (2008). Remote sensing of canopy cover in horticultural crops. *HortScience*, *43*(2), 333–337.

Wardlaw, R., & Barnes, J. (1999). Optimal allocation of irrigation water supplies in real time. *Journal of Irrigation and Drainage Engineering*, *125*(6), 345–354.

*Water Energy Nexus: Excerpt from the World Energy Outlook 2016* (Tech. Rep.). (2016). Paris, France.

Zhang, H., Anderson, R., & Wang, D. (2015). Satellite-based crop coefficient and regional water use estimates for Hawaiian sugarcane. *Field Crops Research*, *180*, 143–154.