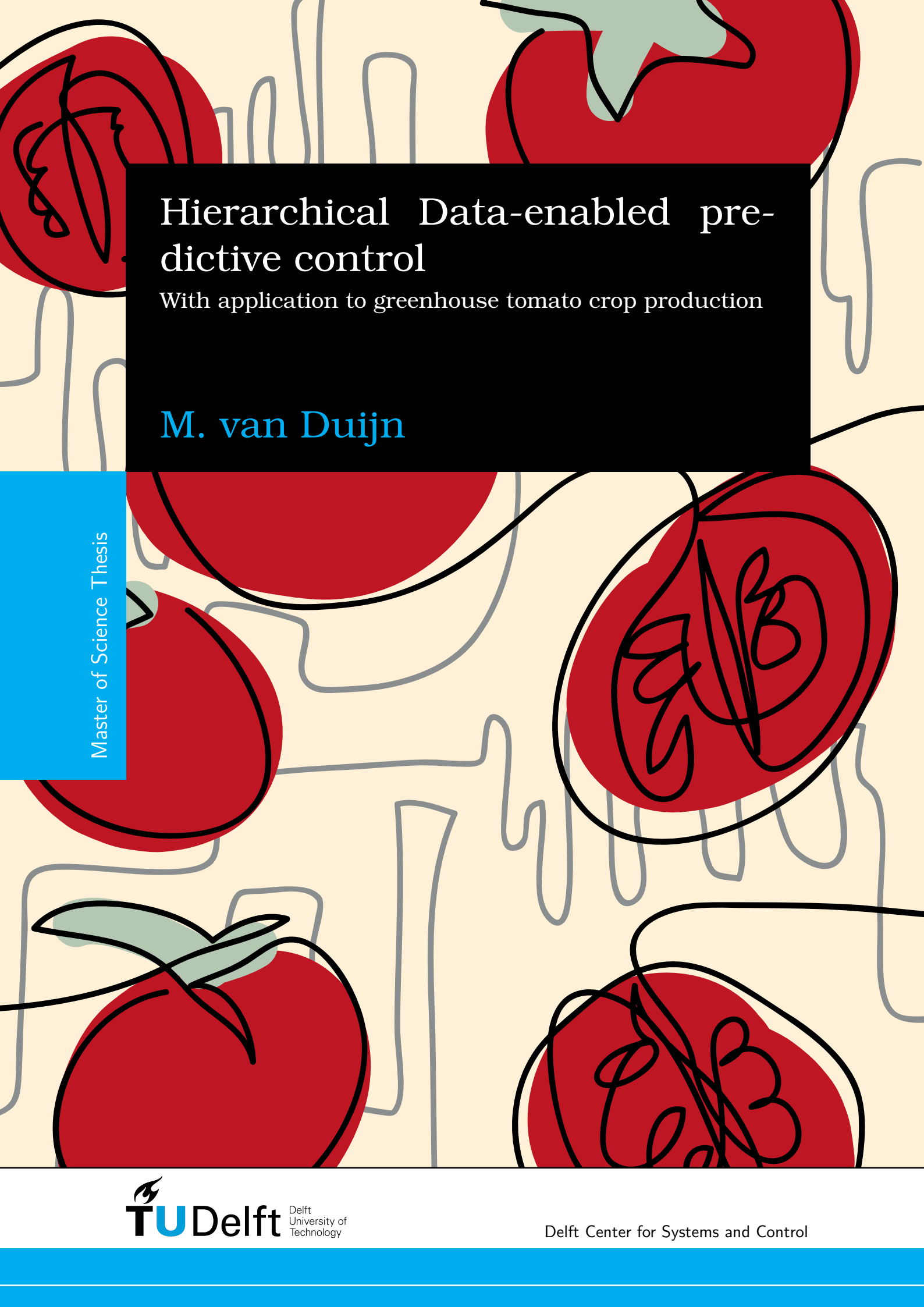


Hierarchical Data-enabled predictive control

With application to greenhouse tomato crop production

M. van Duijn

Master of Science Thesis

TU Delft Delft University of Technology

Delft Center for Systems and Control

Hierarchical Data-enabled predictive control

With application to greenhouse tomato crop production

Master of Science Thesis

For the degree of Master of Science in Systems and Control at Delft University of Technology

M. van Duijn

November 18, 2021

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of Technology

Abstract

With the rapidly growing world population and improving living standards of upcoming economies, the demand for fresh food is increasing vastly. Greenhouses have proven to be very effective in increasing crop yield since they offer a sheltered environment that can be controlled. Greenhouse growers need to take care of many processes, among which are maximising the crop yield while using as little resources as possible and controlling the indoor climate. An approach to alleviating demand for experienced growers is to automate processes within the greenhouse that aid in setpoint generation and setpoint tracking. Existing literature has presented multi-level architectures to cope with the two different timescales on which the two greenhouse subsystems act, i.e., the crop and greenhouse indoor climate subsystem. Current approaches in literature employ model-based techniques, which need tedious calibration per considered situation due to the influences of the local weather, environment and highly non-linear underlying physical processes. Therefore, these model-based approaches are scarcely used in the horticultural sector. Thus, this thesis proposes a data-driven control scheme that deals with long-term crop production control while taking into account resource usage.

The proposed control scheme uses temporal and functional decomposition of the overall problem to accommodate the different timescales on which the greenhouse indoor climate subsystem and crop subsystem act. With the focus on the entire growing season, this thesis emphasises the importance of the climate strategy for crop production control. The aim of this thesis is threefold. Firstly, a non-linear simulator model for the crop dynamics has been selected, adapted, and calibrated, using the Autonomous greenhouse challenge (AGC) dataset. Secondly, the overall crop production control problem has been decomposed into a hierarchical structure, including two subproblems with different objectives, acting on different timescales and using different system representations. The communication protocol between the two layers has been determined, and the setpoint tracking controller is established for the lower layer and tested on a benchmark temperature reference trajectory. Thirdly, the Data-enabled predictive control (DeePC) algorithm is leveraged for the crop production control problem and compared in simulation against a conventional Economic model predictive control (EMPC) setpoint generating controller. Both use the aforementioned established setpoint tracking controller for controlling the greenhouse towards the generated climate strategies.

In simulation, it is shown that the DeePC setpoint generating controller results in a climate strategy that results in a lower crop yield by 1.51% when compared to the EMPC climate strategy and 0.54% less when compared to the predefined benchmark reference trajectory. Furthermore, the resource usage needed for the climate strategy proposed by the DeePC setpoint generating controller uses 1.75% more heating resources when compared to the EMPC setpoint generating controller and 11.1% more when compared to a predefined temperature setpoint trajectory. The DeePC climate strategy results in 7.92% and 5.27% less electricity use when compared to the EMPC and predefined benchmark climate strategy, respectively. The DeePC climate strategy CO_2 usage is 11.4% less when compared to the EMPC climate strategy CO_2 usage and 20.8% less when compared to the predefined benchmark climate strategy.

The decrease in resource usage compensates for the lower crop yield of the DeePC climate strategy. According to the metric of net economic profit, the DeePC generated climate-strategy outperforms the benchmark climate strategy and EMPC climate strategy.

Contents

List of Figures

List of Tables

Acknowledgements

First of all, I would like to thank my supervisor, prof.dr.ir. Tamás Keviczky for his time, assistance and feedback during this thesis. It has been of tremendous help. Second, I would like to thank my other supervisor, Leonard Baart de la Faille, for his support, assistance, questions and inspiration from the real-life situation. It has made this thesis more tangible.

Besides my supervisors, I would like to thank my study buddies, roommates and numerous others for making this project more fun during the many hours I have spent on this Thesis, either at the TU Delft or at home during the pandemic. Lastly, I would like to thank my family and Sven for their unconditional support during this process.

Delft, University of Technology M. van Duijn
November 18, 2021

Chapter 1

Introduction

1-1 Motivation and problem formulation

The demand for food is ever increasing due to the world population growth and the increased living standard of many [2]. The world population is expected to increase to almost 10 billion people by 2050 [3]. One of the Sustainable development goals (SDG) is to end world hunger and ensure food security and nutritious food for humanity [4]. One approach to fulfilling this goal is to up-scale food production and make it more efficient using modern technology.

Glass greenhouse cultivation is the standard in the Netherlands since the greenhouse offers a favourable and protected environment for the crop in which the climate can be controlled [5]. Greenhouses have proven to yield high crop volumes per unit production area with efficient water use [6]. In the Netherlands alone, 1.5 billion tonnes of food are produced annually in glass greenhouses [7]. Energy, water, CO_2 and land are the four primary resources that are used for crop production [8].

Even though growing crops in greenhouses is beneficial, much energy is used. The Dutch horticultural sector amounts to 5% of the fossil fuel emissions in the Netherlands [7]. The Dutch government and the horticultural sector have signed an incentive to strive towards a CO_2 reduction of 3.5 megatonnes and climate-neutral glass greenhouse food production by 2040 [9] [10]. Optimal control policies can increase the efficiency of crop cultivation, i.e., water and energy use [7]. In this way, automatic control of the greenhouse-crop system helps tackle the growing food demand sustainably.

Another issue in the greenhouse sector is the need for experienced growers that manage all aspects of greenhouse crop production [11], especially in the trend that greenhouse production is implemented increasingly in developing countries [12]. Greenhouse growers need to make complex decisions that maximise crop yield while minimising resources. Automatic control can support the growers in this process [13]. Also, less experienced growers can then more easily operate greenhouse systems by employing more automated control strategies. In this way, the demand for and and shortage of the scarcely available experienced growers is alleviated.

Optimal control policies based on descriptive models of the system have proven their use in the control of complex industrial processes and systems [14]. Therefore, it has also been an often-used direction of research to control the indoor greenhouse climate [15]. However, both the indoor climate and crop are challenging to model systems which require extensive model calibration for each considered situation, depending on the greenhouse system, crop type, weather and location, amongst other aspects [16].

Since data has become more abundant, industrial process control has shifted towards data-driven control techniques. Also, Artificial intelligence (AI) algorithms have been widely studied in horticultural practices. However, these horticultural applications are primarily limited to plant observation activities [13]. The use of the data-driven methods for crop cultivation control is still limited [17]. Therefore, a gap in the research is the application of data-driven techniques for crop cultivation control. A novel data-driven control policy is the so-called Data-enabled predictive control (DeePC). It combines state estimation, output prediction, and optimal control calculation in one single optimisation based only on the system's measured input and output data. DeePC can accommodate a comprehensible system representation, cost function description and constraint implementation, and seems a promising data-based control scheme.

The above research directions and goals can be defined formally into the following research objective:

Design a data-driven control scheme that maximises crop production and uses the greenhouse actuators as efficiently as possible over the complete span of the growing season.

Therefore, this thesis aims to contribute to the implementation of a particular data-driven method for long-term crop control while minimising energy usage. This project thus provides a step towards automatically operated greenhouses for crop cultivation control to alleviate the demand for and shortage of experienced growers, helping increase crop production in glass greenhouses while using resources and inputs like energy, water, land and labour optimally.

1-2 Research question

The following research question is defined.

How can the novel DeePC algorithm be employed for the crop production control problem over the entire growing season to maximise the crop yield while minimising the control costs, using a long-term weather prediction?

The following sub-research questions are defined:

- **What model will be implemented as ground truth, and what will be the benchmark against which the proposed controllers are compared?**
  The performance of the data-driven control scheme will be evaluated, and as such, a simulator that acts as the ground truth for the crop subsystem is needed. Selecting a suitable model and adapting it to the system and data is the first step to answering the research question.

- **How will the different timescales on which the greenhouse subsystems act be dealt with?**
  The greenhouse system includes two subsystems, i.e., the slow crop subsystem and the fast greenhouse indoor climate subsystem. The overall problem spans over the entire growing season. Hence, a scheme to handle the two different timescales is needed to ensure the overall problem is tractable while including both timescales.
- **How will the novel DeePC algorithm be implemented to maximise the crop yield?**
  The novel DeePC algorithm is presented initially as a setpoint tracking control scheme without exogenous inputs to the system. Therefore, the algorithm needs to be leveraged to apply to the considered greenhouse-crop system and control problem.

1-3 Thesis contributions

From the research objectives, as stated in the previous section, the main goal of this project is to develop a data-driven control scheme that controls crop production over the entire growing season while minimising resource use and employing a long-term weather prediction. The following thesis contributions can be stated, following the research objectives and the results in simulation.

- **Extension and calibration of an available non-linear, mechanistic crop model to include all eight types of modern greenhouse actuators present in the Autonomous greenhouse challenge (AGC).**
  In this thesis, a non-linear model has been selected to use as a ground truth simulator for the crop subsystem to assess the performances of the proposed setpoint generating controllers. The original model is adapted to fit the considered system and calibrated to fit the data.
- **Design of a two-layer control architecture that deals with the greenhouse-crop system's different dynamical time scales.**
  A two-layer control scheme is proposed to deal with the two subsystems that act on different timescales. The upper layer operates at a slow sampling rate and uses a long prediction to generate optimal climate setpoints to account for foreseen weather circumstances. These climate setpoints are the reference greenhouse indoor air temperatures that influence crop growth. The lower layer tracks these climate setpoints, which operates at a faster sampling rate and ensures the greenhouse indoor climate is controlled.
- **An Economic model predictive control (EMPC) controller that generates temperature setpoints to control the tomato crop in the greenhouse.**
  The data-driven control scheme this thesis proposes is compared to an equivalent model-based scheme, which employs an EMPC setpoint generating controller in the upper layer. The EMPC controller uses the entire growing season as the prediction horizon in a shrinking horizon implementation and can control crop production.
- **Design of a leveraged DeePC control scheme that generates temperature setpoints to control the tomato crop in the greenhouse.**

This work presents a leveraged DeePC setpoint generating controller in the upper layer of the overall control scheme, capable of controlling crop production. This work is the first to employ the DeePC framework to output setpoints while considering an economic objective.

1-4 Thesis outline

The remainder of this thesis is structured as follows. In Chapter 2, the greenhouse-crop system at hand is introduced, and the scope of the thesis is given. Furthermore, the model used as ground truth and the proposed architecture into which the proposed controllers fit are shown. In Chapter 3, the controller used to obtain the actual control costs and greenhouse-crop states over the entire growing season is introduced, and the benchmark results obtained with this controller are shown. In Chapter 4, the first reference air temperature generating controller is presented. The preliminary results of this EMPC-type controller are shown. The second setpoint generating controller and accompanying preliminary results are shown in Chapter 5. This is a DeePC-type controller. The performances of the resulting synthesised controllers are then compared against the benchmark in simulation in Chapter 6. Finally, in Chapter 7, the research questions are answered, and recommendations for future work are given.

Chapter 2

Problem definition

In this chapter, the overall problem that this thesis considers will be introduced. The main focus is the crop yield of a greenhouse tomato crop. Therefore, first, the physical greenhouse system with tomato crop is described. Then, the mathematical model taken as the simulator ground truth is defined and how the model is adapted to, and calibrated with the available real-life data. The overall problem is decomposed into subproblems for tractability, and the control architecture on which the rest of this thesis is based is introduced.

2-1 The greenhouse-crop system

2-1-1 Assumptions and scope

The greenhouse-crop system is a non-linear, strongly coupled system, and many different types exist around the world, varying in either greenhouse or crop type. The thermodynamic variables within the greenhouse indoor climate subsystem are under the strong influence of the weather as exogenous inputs. In the literature, models ranging from very detailed process descriptions within the greenhouse to models describing just one thermodynamic variable are presented [15]. The crop subsystem is a complex biological system, with models ranging from describing individual cells to models describing just one crop biological variable. To arrive at a complete greenhouse-crop system description that is used for this thesis, and suitable for control purposes, the following most significant assumptions are made:

1. The greenhouse air is assumed a perfectly stirred tank, i.e., all thermodynamic variables are considered spatially constant.

2. The tomato crop is considered a big-fruit, big-leaf entity, i.e., it will be modelled in a so-called medium-grained fashion. Hence, there are only two measurable outputs, one variable describing the total mass of the tomato fruits and one variable representing the total mass of the leaves.

3. The fruit quality is considered constant.
4. Actuator setpoints are reached, e.g., the heating pipe temperature that the controller generates is the heating pipe temperature that is realised.
5. Perfect sensor data is assumed, implying that the sensor data that is used does not contain measurement noise or any other faults.
6. A single harvesting moment at the end of the growing season is considered.

The following aspects are focused on within this thesis, which does not encompass all aspects within crop control.

1. This thesis focuses on the greenhouse indoor climate, not the soil milieu, i.e., it is assumed the crop system is well fertilised and irrigated.
2. The influence of pests and diseases is ignored.
3. The greenhouse-crop system description and modelling are considered for the outdoor climate of the Netherlands and a Venlo-type greenhouse.
4. Air temperature is considered the primary thermodynamic variable, influencing crop growth the most [13].

2-1-2 System description

The greenhouse system consists of the greenhouse indoor climate subsystem and the crop subsystem, depicted schematically in Figure 2-1 below.

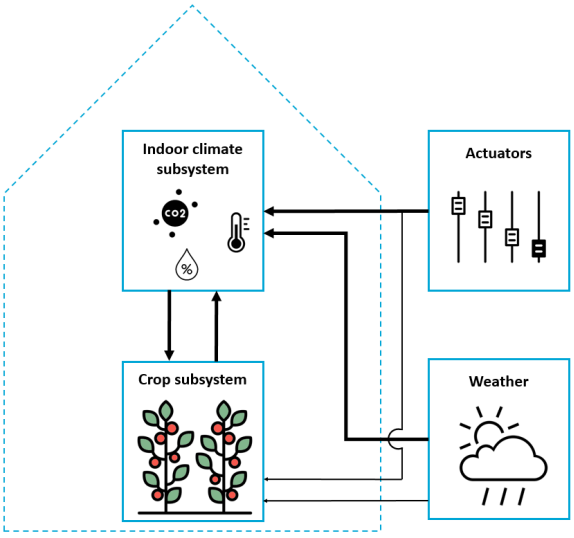


Figure 2-1: The greenhouse system consists of the indoor climate subsystem, the crop subsystem and the external input systems that are the actuators as control inputs and the weather as exogenous inputs.

The two subsystems in the greenhouse influence each other via the processes of photosynthesis, respiration and transpiration. As considered in this work, the subsystems include multiple types of variables, as depicted in Figure 2-2 below. This description is an adapted description of the description that is used later on as ground truth and matches the used dataset [18].

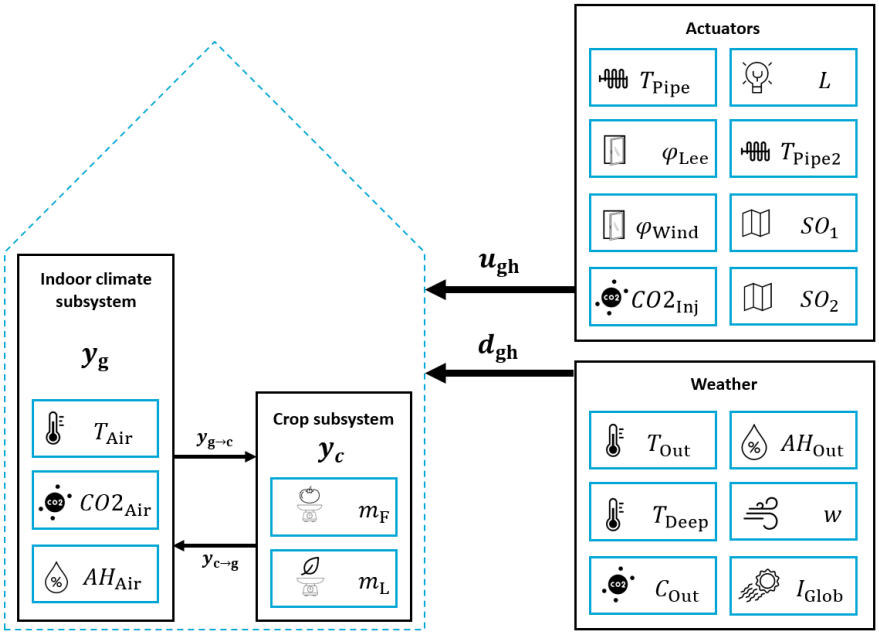


Figure 2-2: The variables of interest per subsystem.

The climatic variables of interest of the indoor climate subsystem are given by $\boldsymbol{y}_{\text{g}}$ in (2-1) below. These are the three climatic variables, present in the majority of the available indoor climate models [19]

$$\boldsymbol{y}_{\text{g}} = \begin{bmatrix} T_{\text{Air}} \\ C_{\text{Air}} \\ AH_{\text{Air}} \end{bmatrix}, \tag{2-1}$$

where T_{Air}, C_{Air} and AH_{Air} denote the air temperature in [C°], greenhouse air CO_2 levels in [g m^{-2}], and greenhouse air absolute humidity in [g m^{-3}], respectively.

The crop variables of interest are given by $\boldsymbol{y}_{\text{c}}$ in (2-2) below

$$\boldsymbol{y}_{\text{c}} = \begin{bmatrix} m_{\text{B}} \\ m_{\text{F}} \\ m_{\text{L}} \\ D \end{bmatrix}, \tag{2-2}$$

where m_{B}, m_{F}, m_{L} and D denote the assimilate buffer dry weight in [g m^{-2}], total crop fruit dry weight in [g m^{-2}], total crop leaf dry weight in [g m^{-2}], and the crop stage, respectively.

The control input vector to the greenhouse is given in (2-3) below

$$\boldsymbol{u}_{\text{gh}} = \begin{bmatrix} T_{\text{Pipe,1}} \\ T_{\text{Pipe,2}} \\ \phi_{\text{Lee}} \\ \phi_{\text{Wind}} \\ C_{\text{Inj}} \\ L \\ SO_1 \\ SO_2 \end{bmatrix}, \tag{2-3}$$

where $T_{\text{Pipe,1}}$, $T_{\text{Pipe,2}}$, are the heating pipe temperatures of the upper and lower pipe system in [°C]. Furthermore, ϕ_{Lee} and ϕ_{Wind} denote the leeward and windward side window opening in [%], C_{Inj} denotes the CO_2 injection rate in [g $s^{-1}m^{-3}$], L denotes the light intensity in [W m^{-2}], and SO_1 and SO_2 denote the screen opening of the temperature and climate screens in [%]. The eight control inputs that act on the greenhouse system act directly on the greenhouse indoor climate sub-system, i.e., $\boldsymbol{u}_{\text{gh}} = \boldsymbol{u}_{\text{g}}$. The control input vector $\boldsymbol{u}_{\text{gh}}$ is changed with respect to [18], which only consisted of four control inputs.

The control inputs that act directly on the crop subsystem are denoted by $\boldsymbol{u}_{\text{c}}$ in (2-4) below

$$\boldsymbol{u}_{\text{c}} = \begin{bmatrix} L \\ SO_1 \\ SO_2 \end{bmatrix}, \tag{2-4}$$

The exogenous inputs that act on the greenhouse are given by the vector $\boldsymbol{v}_{\text{gh}}$ in (2-5) below

$$\boldsymbol{v}_{\text{gh}} = \begin{bmatrix} C_{\text{Out}} \\ AH_{\text{Out}} \\ T_{\text{Out}} \\ I_{\text{Glob}} \\ w_{\text{Out}} \end{bmatrix}, \tag{2-5}$$

where C_{Out} is the outside CO_2 concentration in [g m^{-3}], AH_{Out} is the outside air absolute humidity in [g m^{-3}], T_{Out} is the outside air temperature in [°C], I_{Glob} is the solar radiation in [W m^{-2}], and w_{Out} is the wind speed in [m s^{-1}]. The five exogenous inputs that act on the greenhouse system act directly on the greenhouse indoor climate subsystem, i.e., $\boldsymbol{v}_{\text{gh}} = \boldsymbol{v}_{\text{g}}$.

The exogenous inputs that act directly on the crop subsystem are denoted by $\boldsymbol{v}_{\text{c}}$ and given in (2-6) below

$$\boldsymbol{v}_{\text{c}} = [I_{\text{Glob}}]. \tag{2-6}$$

2-1-3 Data preprocessing

The data that has been used in this thesis for modelling and control purposes is from the Autonomous greenhouse challenge (AGC) of 2019. The dataset contains data on the greenhouse indoor climate, weather variables, crop status and used controls. The data were collected during a 6-month tomato growing season in the Netherlands [13].

The data is not immediately applicable for the chosen system description type. Hence, the data preprocessing steps will be explained in this section. The dataset that has been used is the dataset of the Automatoes team, one of the five teams that competed in the AGC and won.

NaN data points

Some parts of the greenhouse indoor climate dataset and weather variables dataset contain many NaN data points. The missing entries that the NaN covers are fixed by linear interpolation between the two adjacent data points that have data. This is justified because only 71 NaN data points are present per variable, which is 0.15% of the total available measurements per variable.

Preprocessing of the control input data

To circumvent actuator dynamics in the control schemes, either the desired actuator setpoints or the realised actuator setpoints need to be used, and the difference between them should be neglected. In this thesis, the desired setpoints have been used as data, not the realised setpoints. It is assumed that the desired actuator setpoints are realised. This was done because in the dataset of the realised setpoints, some of the control inputs only contained NaN entries. Some other issues in the control input data are elaborated on per control input below. If a certain control input is not mentioned, the data is used without preprocessing or important note.

Heating pipe temperatures Heating pipe temperature measurements minima were 0 [°C] which is impossible since then the heating pipes would cool the greenhouse as the heating pipe temperature is then lower than the greenhouse temperature. Therefore, the heating pipe temperature data points were raised by 20 °C. The 20 °C was chosen because then the minimal and maximum values of the measurements coincide with the ground truth model [18].

Window openings The assumption is made that the control scheme has knowledge of the windward and leeward side window openings, since the window openings on different sides of the greenhouse with respect to the wind have different influence upon the ventilation [20].

CO_2 injection rate The CO_2 injection rate C_{Inj} in the dataset is defined in [kg ha^{-1} h^{-1}]. The following conversion formula has been used to change it to the unit of [g s^{-1} m^{-2}] that is used in the non-linear models that describe the system.

$$C_{\text{Inj}} = \frac{1}{3.6 \cdot 10^4} \tilde{C}_{\text{Inj}}, \tag{2-7}$$

where C_{Inj} is the CO_2 injection rate in [g^{-1}m^{-2}] and $\tilde{C}_{\text{Inj}}$ is the CO_2 injection rate in [kg ha^{-1}h^{-1}].

Artificial lighting In the AGC, two lighting systems, Light emitting diode (LED) and High pressure sodium (HPS), are used. Since the LED and HPS lighting systems work using different mechanisms, this needs to be taken into account. For the control schemes, the total intensity of both lighting systems has been used. The total intensity is converted from μmolm^{-2} to Wm^{-2} by the average conversion factor of 48.19 [Wm^{-2} μmol^{-1}]. For the conversion factors, the reader is referred to Section 6-1-2.

Preprocessing of the available crop data

The AGC dataset includes only the measured fruit dry weight, from day 65 onwards, with data samples about every five days after the 65^{th} day. Due to the nature of the non-linear model for which the AGC data will be used, the data is appended with $m_{\text{F}} = 5$ before the 65^{th} day. The non-linear model needs $m_{\text{F}} \neq 0$ to simulate, and a non-zero fruit dry weight is also required for the respiration. At the planting date, tomato plants already have a small albeit non-zero fruit and leaf dry weight [21]. If a particular crop variable is not mentioned, the data is used without preprocessing or important note.

Crop data interpolation The crop measurements have been interpolated for corresponding sampling rates for further use in model calibration, model identification and controller design. Crop measurements are costly to gather in reality since it is a manual process and labour intensive [22].

Fruit dry weight The dataset contains measurements of the total fruit fresh weight, and the Dry matter content (DMC) per measurement instance k. The total fruit dry weight is then calculated by

$$m_{\text{F}}(k) = C_{\text{DMC}}(k) \cdot \tilde{m}_{\text{F}}(k), \tag{2-8}$$

where $\tilde{m}_{\text{F}}$ is the total fruit fresh weight in [gm^{-2}] and m_{F} is the total fruit dry weight in [gm^{-2}] and C_{DMC} is the DMC at measurement instance k.

Leaf dry weight The AGC dataset does not contain measurements on the leaf dry weight. However, the leaf dry weight is the most important crop variable that couples the crop subsystem to the greenhouse indoor climate subsystem. How this is dealt with will be explained in Section 2-2-3.

Processing of the greenhouse indoor climate measurements

If a specific greenhouse indoor climate variable is not mentioned, then the data is used without preprocessing or important note.

Air CO_2 concentration The CO_2 measurements are in [ppm], while the mechanistic non-linear crop model uses CO_2 in [gm^{-3}]. The conversion factor below is used [24].

$$C_{\text{Air}} = 0.0409 \cdot \tilde{C}_{\text{Air}} MW_{\text{CO}_2}, \tag{2-9}$$

where $\tilde{C}_{\text{Air}}$ is the CO_2 concentration in [ppm], C_{Air} is the CO_2 concentration in [gm^{-3}] and MW_{CO_2} is the molecular weight of CO_2.

Air absolute humidity The available humidity measurements are the humidity deficit in the unit [gm^{-3}], which will be introduced in Section A-0-1, and the relative humidity in [%]. The system description as used in this thesis needs the humidity measurements in terms of the absolute humidity of the unit [gm^{-3}]. The conversion from the humidity deficit and relative humidity to absolute humidity is presented below [23].

$$AH_{\text{Air}} = \frac{RH_{\text{Air}} HD_{\text{Air}}}{100 - RH_{\text{Air}}}, \tag{2-10}$$

Implementation of the weather

Weather measurements of the following variables used as exogenous inputs to the greenhouse are available in the AGC dataset: T_{Air}, AH_{Out}, w_{Out}, I_{Glob}. These weather measurements are implemented directly.

Outside CO_2 concentration There are no measurements available of the outside CO_2 concentration. The average outside CO_2 concentrations are in the range of 410 ppm [24], which results in an average outside CO_2 concentration of 0.55 [gm^{-3}]. For the simulation and control purposes a standard normal distribution variable times 0.01 is added to this constant value at each sampling time.

2-1-4 Current practice and benchmark

To compare the proposed controllers that generate reference temperatures, a benchmark reference temperature is selected. This is based on the current practice in the horticultural sector.

Current practice

Reference temperatures that come from the seed companies are given to the growers that regulate towards these reference temperatures or deviate from it if the setpoints cannot be realised. The reference temperatures or so-called growing guidelines are predetermined and are subdivided into tomato growing stages that last a set number of days.

Since the guidelines differ per tomato type, the growing conditions presented in the section below are taken from the literature.

Optimal tomato growing conditions

Table (2-1) below shows the optimal growing temperatures that have emerged from the vast amount of simulation and growing results in the literature on tomatoes [25].

Period	Lower limit [°C]	Optimal [°C]	Upper limit [°C]
Day	15	24	32
Night	13.5	20	27

Table 2-1: Optimal growing temperatures for tomatoes

Benchmark set-up

This reference temperature trajectory is chosen as the benchmark set-up to compare the proposed temperature generating controllers with.

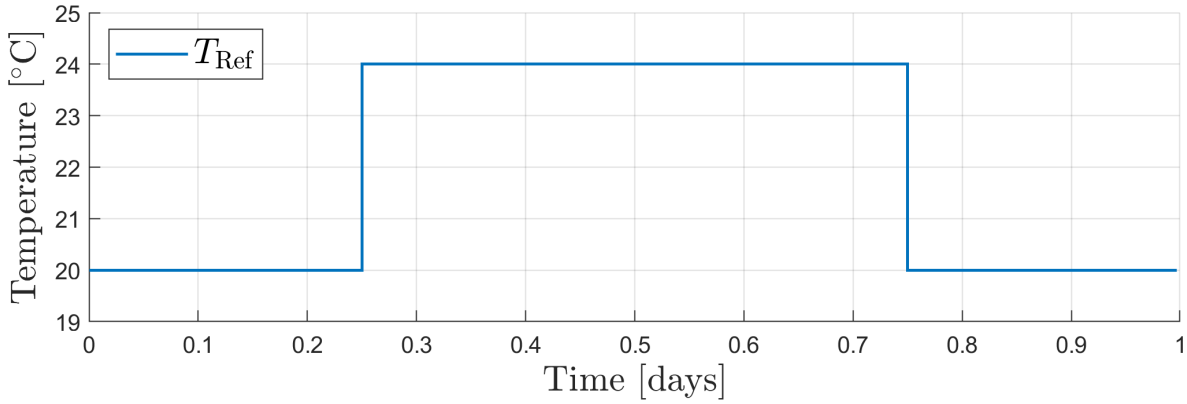


Figure 2-3: Daily benchmark reference temperature.

As can be seen, the diurnal temperature setpoint is higher than the nocturnal temperature setpoint.

2-2 Mechanistic non-linear crop model as ground truth

This section describes the crop model that has been used as ground truth to compare and validate the different control methods.

The greenhouse indoor climate model that is adapted and calibrated for a day of measurements is described in Section A-0-1. This model has not been used to control over the entire season, but it can be used for control purposes that consider only a day's worth of simulation.

Both models have been adapted and calibrated. The adapted versions will be given, and the changes, when compared to the original model, will be stated.

2-2-1 Crop mathematical model

The chosen crop model non-linearly describes the underlying physical processes that govern the growth of the tomato plant. The adapted version to include all actuator types present in the AGC dataset, is given.

The rate of change of the assimilate buffer weight $\dot{m}_{\text{B}}$ is given by

$$\dot{m}_{\text{B}} = P - b\left(f g_{\text{F}} m_{\text{F}} + v g_{\text{L}} \frac{m_{\text{L}}}{z}\right) - b\left(r_{\text{F}} m_{\text{F}} + r_{\text{L}} \frac{m_{\text{L}}}{z}\right), \tag{2-11}$$

where m_{B} denotes the assimilate buffer dry weight in [g m^2], P is the crop photosynthesis rate in [g s^{-1} m^{-2}], b is the buffer switching function [-], f is the fruit assimilate requirement

quotient [-], g_F is the relative fruit growth rate [s^{-1}], m_F is the total fruit dry weight [g m^{-2}], v is the vegetative requirement quotient [-], g_L is the relative fruit growth rate [s^{-1}], m_L is the total leaf dry weight [g m^{-2}], z is the leaf fraction of vegetative dry weight [-], r_F is the relative fruit respiration rate [s^{-1}], and r_L is the relative leaf respiration rate [s^{-1}], respectively.

The crop photosynthesis rate P is given by

$$P = P_m l \frac{I}{p_1 + I} \frac{C_{ppm}}{p_2 + C_{Air}}, \tag{2-12}$$

where P_m is the maximum photosynthesis rate [g $s^{-1}m^{-2}$], l is the Leaf area index (LAI) correction function [-], I is the Photosynthetically active radiation (PAR) [μmol $s^{-1}m^{-2}$], p_1 and p_2 are photosynthesis parameters, C_{ppm} is the CO_2 concentration [ppm], and C_{Air} is the CO_2 concentration [g m^{-3}], respectively.

The LAI correction function is given by

$$l = \frac{\left(\frac{m_L}{w_R}\right)^m}{1 + \left(\frac{m_L}{w_R}\right)^m}, \tag{2-13}$$

here, m_L is the leaf dry weight per ground area [g m^{-2}], and w_R [g m^{-2}] and m [-] are LAI correction function parameters.

The PAR I is given by

$$I = \eta m_p I_{Glob}, \tag{2-14}$$

where η is the radiation conversion factor [-], m_p [μmol J^{-1}] is the conversion factor from Watt to μmol, and I_{Glob} is the solar radiation [W m^{-2}], respectively.

The CO_2 [ppm] C_{ppm} is given by

$$C_{ppm} = \frac{10^6 R_G}{M_{CO_2} p_{atm}} (T_{Air} + T_0) C_{Air}, \tag{2-15}$$

where R_G is the universal gas constant [$Jmol^{-1}K^{-1}$], M_{CO_2} is the molar mass of CO_2 [kg], p_{atm} is the atmospheric air pressure [kPa], T_{Air} is the greenhouse indoor air temperature [°C], and T_0 is the absolute zero [°C], respectively.

The buffer switching function b is given by

$$b = 1 - e^{-b_1 m_B}, \tag{2-16}$$

where b_1 [m^2g^{-1}] is a buffer switching parameter and m_B is the assimilate buffer dry weight in [g m^{-2}]. When the assimilate buffer is almost empty, b tends to zero and when the assimilate buffer is almost full, b tends to one, governing the switching between the different biomasses within the crop.

The relative fruit growth rate g_F and relative leaf growth rate g_L are given by

$$g_F = (f_1 - f_2 D_P) Q_G^{\frac{T_{Air} - T_G}{10}}, \tag{2-17}$$

$$g_L = g_F v_1 e^{v_2 (T_{Air} - v_3)}, \tag{2-18}$$

where f_1 [s^{-1}] and f_2 [s^{-1}] are fruit growth rate coefficients, Q_{G} [-] is the Q_{10}-value for the temperature effect on the fruit growth rate, T_{G} [°C] is the growth rate reference temperature, and v_1 [-], v_2 [°C^{-1}] and v_3 [°C] are vegetative fruit growth ratio coefficients.

The relative fruit respiration rate r_{F} and relative leaf respiration rate r_{L} are given by

$$r_{\mathrm{F}} = M_{\mathrm{F}} Q_{\mathrm{R}}^{\frac{T_{\mathrm{Air}}-T_{\mathrm{G}}}{10}}, \tag{2-19}$$

$$r_{\mathrm{L}} = M_{\mathrm{L}} Q_{\mathrm{R}}^{\frac{T_{\mathrm{Air}}-T_{\mathrm{G}}}{10}}, \tag{2-20}$$

where Q_{R} [-] is the Q_{10}-value for the temperature effect on the maintenance respiration, and M_{F} [s^{-1}] and M_{L} [s^{-1}] are the fruit and leaf maintenance respiration coefficients, respectively.

The rate of change in fruit dry weight $\dot{m}_{\mathrm{F}}$ is given by

$$\dot{m}_{\mathrm{F}} = \Big(b g_{\mathrm{F}} - (1-b) r_{\mathrm{F}}\Big) m_{\mathrm{F}}, \tag{2-21}$$

here, the fruit growth is represented by the first term on the right-hand side and the fruit respiration is represented by the second term on the right-hand side.

The rate of change of leaf dry weight on the crop $\dot{m}_{\mathrm{L}}$ is given by

$$\dot{m}_{\mathrm{L}} = \Big(b g_{\mathrm{L}} - (1-b) r_{\mathrm{L}} - h_{\mathrm{L}}\Big) m_{\mathrm{L}}, \tag{2-22}$$

Here, the leaf growth is represented by the first term on the right-hand side and the leaf respiration is represented by the second term on the right-hand side.

The rate of change in the crop development stage $\dot{D}$ is given by

$$\dot{D} = d_1 + d_2 + \ln\left(\frac{T_{\mathrm{Air}}}{d_3}\right) + d_4 t, \tag{2-23}$$

where d_1 [s^{-1}], d_2 [s^{-1}], d_3 [°C], and d_4 [-] are plant development rate parameters, and t is the time [s].

The continuous-time dynamical model for the greenhouse crop is then given by

$$\dot{\boldsymbol{y}}_{\mathrm{c}} = f_{\mathrm{c}}(\boldsymbol{y}_{\mathrm{g}}, \boldsymbol{y}_{\mathrm{c}}, \boldsymbol{u}_{\mathrm{g}}, \boldsymbol{v}, t). \tag{2-24}$$

Crop model alterations

In (2-21) and (2-22), the terms that represent the harvest of the fruits and the pruning of the leaves are left out. Furthermore, compared to [18], the fruit and leaf harvest coefficients and the total harvest of fruits and leaves are left out of the model.

The original model by [18] implements a continuous harvesting relationship, whereas in reality, the harvesting takes place at discrete time instances [26]. Crop management decisions such as leaf pruning, fruit harvesting and truss and stem density management actions are only valuable for describing the crop development when the considered crop is described in a fine-grained fashion, meaning that the individual tomatoes and leaves are modelled [13]. This is

not the case for this thesis, and hence any intermediate harvesting decisions and actions are omitted, and a single harvest at the end of the season is assumed [27].

In the AGC dataset, the desired stem density and number of trusses are available [13], but not the number of tomatoes per truss. Therefore, to make the data compatible with the chosen model, again, a single harvest at the end of the season is considered.

2-2-2 Discretisation of the Continuous-time (CT) model

The crop and greenhouse indoor climate models as described in Section 2-2-1 and Section A-0-1 are CT models discretised using an Euler discretisation, which is given in (2-25) below

$$\boldsymbol{y}(k+1) = \boldsymbol{y}(k) + \tau \dot{\boldsymbol{y}}(k), \tag{2-25}$$

where τ is the sampling time. Here, $\boldsymbol{y}$ can be either crop or greenhouse indoor climate outputs.

2-2-3 Model calibration

The adapted crop model is calibrated using the measured climate outputs, exogenous inputs and control input data. The general calibration cost function J_{Gen} that is minimised for the parameter calibration is given below

$$\begin{aligned} J_{\text{Gen}}(w_h, y_{hj}, p) &= \sum_{h=1}^{L} \sum_{i=1}^{M} \sum_{j=1}^{N} w_h \Big(\hat{y}_h \left(t_i, p\right) - y_{hj} \left(t_i\right) \Big)^2, \\ p^* &= \arg\min J_{\text{Gen}}(w_h, y_{hj}, p), \end{aligned} \tag{2-26}$$

where L, M and N denote the number of outputs, time instances and replicates per time instance, respectively [28]. w_h denotes the relative weight for each output, $\hat{y}_h(t_i, p)$ denotes the simulated output y_h at time t_i and $y_{hj}(t_i)$ denotes, at time t_i, the j^{th} measurement replicate of y_h.

Crop model calibration

The crop model has been calibrated using the AGC data. Since the AGC measurement data includes only the fruit dry weight m_{F}, the crop model is calibrated only using the fruit dry weight as output.

Since the crop exhibits slow dynamics, the calibration of the crop model needs to be performed over the entire growing season [18].

The calibration cost function for the crop model calibration, which is a simplified version of (2-26) is given by

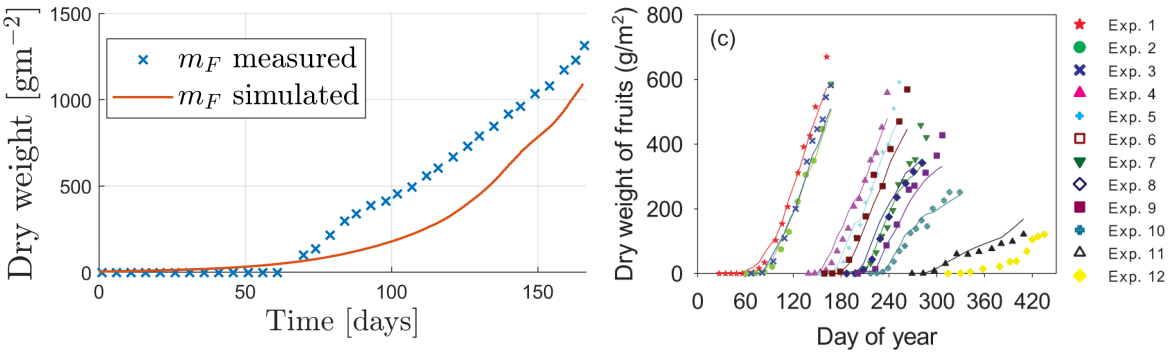
$$\begin{aligned} J(m_{\text{F}}, p) &= \sum_{h=1}^{L} \left(\hat{m}_{\text{F}} \left(t, p\right) - m_{\text{F}} \left(t\right) \right)^2, \\ p^* &= \arg\min J(m_{\text{F}}, p), \end{aligned} \tag{2-27}$$

where L is the time instance of the last m_{F} measurement of the growing season, $\hat{m}_{\text{F}} \left(t, p\right)$ is the m_{F} prediction at time instance t for chosen parameter vector p, and $m_{\text{F}} \left(t\right)$ is the measured

m_{F} at time instance t. The resulting optimisation problem is a non-linear program and solved using the MATLAB fmincon solver. The non-linear program is due to the non-linear model that is used to obtain $\hat{m}_{\mathrm{F}}$.

Since only the fruit weight was available, not all crop model parameters have been tuned. The parameters that were used in the calibration procedure are given in Section B-1.

Calibrated fruit weight The resulting simulated fruit weight which emerges from the calibrated non-linear model can be seen in Figure 2-4 below.



(a) Calibrated m_{F}.

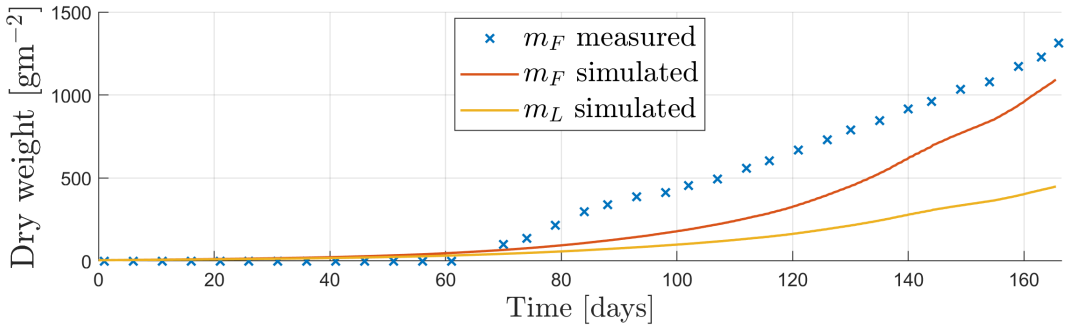
(b) Other crop measurements and model, from [1].

Figure 2-4: Calibrated non-linear crop model fruit weight.

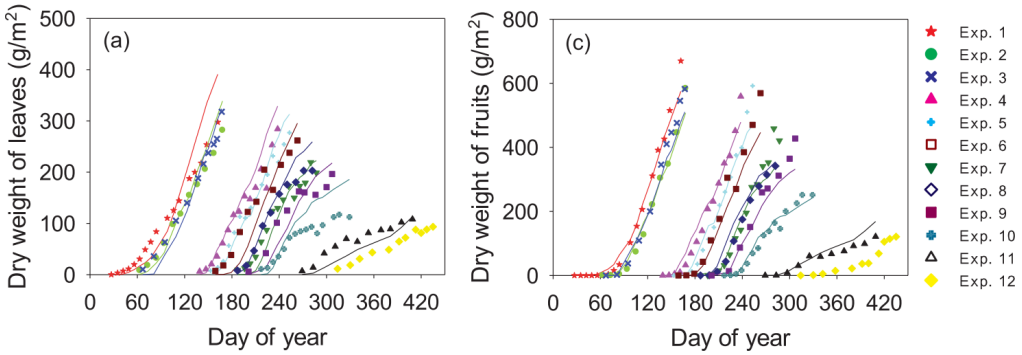
There is no exact match between the simulated m_{F} and the measured m_{F}. This can be explained by the fact that the model uses a big-fruit, big-leaf description whereas in reality the tomato crop is an intricate and complicated biological system encompassing more than just two output variables. Therefore, the model might not describe all processes or be able to capture the true dynamics of the underlying processes of the tomato crop. However, the order of magnitude is the same.

Relationship fruit weight and leaf weight from calibrated crop model Since no measurements were available of the leaf dry weight, the simulation results of m_{L} using the calibrated model can be used to draw some conclusions on the modelled relationship between m_{F} and m_{L}.

The emerging relationship between m_{F} and m_{L} from this model is that m_{L} and m_{F} are almost proportional, which agrees with real-life experiments and the literature on this topic [1] [21]. This is presented in Figure 2-5 below.



(a) Model relationship m_{F} and m_{L}.



(b) Other crop measurements and model from [1].

Figure 2-5: Other crop measurements and model, from [1]. From these two plots, the proportional relationship between m_{F} and m_{L} is confirmed. This proportional relationship also emerges from the calibrated non-linear crop model.

From Figure 2-5 it can be seen that in real-life, a proportional relationship between m_{F} and m_{L} is present as well. The contours of the leaf dry weight per experiment number are shaped almost identically as the contours of the fruit dry weight per experiment number, only with different measurement values.

2-3 Problem decomposition

The dynamics of the greenhouse climate variables operate on a timescale of minutes to hours, whereas the crop dynamics operate on a timescale of days. To ensure all dynamics are captured when the problem is approached in a one-layer control structure, the problem needs to operate at the fast dynamical timescale, meaning that the optimisation problem might become intractable over long optimisation periods. The long optimisation period is, on the other hand, needed to capture the crop dynamics. Therefore, shifting towards a decomposed structure can overcome this tractability issue [29].

The decomposition of the overall system under control occurs via different approaches: functional decomposition, temporal decomposition or spatial decomposition. The basis of functional decomposition is assigning functionally different control objectives to the subsequent different layers in a hierarchical dependence [30]. Temporal decomposition occurs when either the systems' dynamics are characterised on different timescales or when the disturbances

evolve on different timescales. Spatial decomposition occurs when the system includes two spatially individual components that are linked.

Hierarchical optimal control can be useful in at least two types of control problems. The first is when the overall system under control exhibits dynamic behaviour that evolves in two different timescales. The second is whenever the optimisation and control algorithms compute the optimal targets and the effective control actions at different rates [29]. The greenhouse-crop cultivation control problem is characterised by dynamics on different timescales: the fast climatic dynamics and crop physiological processes, and the slow crop growth processes. The two-layer structure that emerges from the dynamics' two timescales can be extended to multiple levels.

The subproblems can all be assigned their respective control objective and submodel. A typical information structure is for the upper layer to determine the setpoints for the lower layer to track [31].

In this section, the two decomposition types are introduced. Hereafter, the control architecture into which the remainder of this thesis fits is presented, along with the linear models used later on in this thesis.

2-3-1 Temporal system decomposition

Two-timescale systems are commonly referred to as singularly perturbed systems [29]. The considered system (2-25) can be decomposed into two separate systems, that are linked via several, arbitrary interconnecting variables, into

$$\begin{cases} \boldsymbol{x}^s(k^{\mathrm{s}}+1) &= \boldsymbol{f}^s(\boldsymbol{x}^s(k^{\mathrm{s}}), \boldsymbol{x}^f(k^{\mathrm{s}}), \boldsymbol{u}(k^{\mathrm{s}}), \boldsymbol{d}^s(k^{\mathrm{s}}), k^{\mathrm{s}}), \\ \boldsymbol{y}^s(k^{\mathrm{s}}) &= \boldsymbol{g}^s(\boldsymbol{x}^{\mathrm{s}}(k^{\mathrm{s}}), \boldsymbol{x}^f(k^{\mathrm{s}}), \boldsymbol{u}(k^{\mathrm{s}}), \boldsymbol{n}^s(k^{\mathrm{s}}), k^{\mathrm{s}}), \end{cases} \tag{2-28}$$

$$\begin{cases} \boldsymbol{x}^f(k^{\mathrm{f}}+1) &= \boldsymbol{f}^f(\boldsymbol{x}^s(k^{\mathrm{f}}), \boldsymbol{x}^f(k^{\mathrm{f}}), \boldsymbol{u}(k^{\mathrm{f}}), \boldsymbol{d}^f(k^{\mathrm{f}}), k^{\mathrm{f}}), \\ \boldsymbol{y}^f(k^{\mathrm{f}}) &= \boldsymbol{g}^f(\boldsymbol{x}^s(k^{\mathrm{f}}), \boldsymbol{x}^f(k^{\mathrm{f}}), \boldsymbol{u}(k^{\mathrm{f}}), \boldsymbol{n}^f(k^{\mathrm{f}}), k^{\mathrm{f}}), \end{cases} \tag{2-29}$$

where $\boldsymbol{f}^s : \mathbb{R}^{n^{\mathrm{s}}_{\boldsymbol{x}}} \times \mathbb{R}^{n^{\mathrm{s}}_{\boldsymbol{x}}} \to \mathbb{R}^{n^{\mathrm{s}}_{\boldsymbol{x}}}$ and for the slow subsystem $\boldsymbol{x}(k^{\mathrm{s}}) \in \mathbb{R}^{n^{\mathrm{s}}_{\boldsymbol{x}}}$, $\boldsymbol{u}(k^{\mathrm{s}}) \in \mathbb{R}^{n^{\mathrm{s}}_{\boldsymbol{u}}}$, $\boldsymbol{y}(k^{\mathrm{s}}) \in \mathbb{R}^{n^{\mathrm{s}}_{\boldsymbol{y}}}$, $\boldsymbol{d}(k) \in \mathbb{R}^{n^{\mathrm{s}}_{\boldsymbol{d}}}$ and $\boldsymbol{n}(k^{\mathrm{s}}) \in \mathbb{R}^{n^{\mathrm{s}}_{\boldsymbol{n}}}$ and for the fast subsystem $\boldsymbol{f}^f : \mathbb{R}^{n^{\mathrm{f}}_{\boldsymbol{x}}} \times \mathbb{R}^{n^{\mathrm{f}}_{\boldsymbol{x}}} \to \mathbb{R}^{n^{\mathrm{f}}_{\boldsymbol{x}}}$ and $\boldsymbol{x}(k^{\mathrm{f}}) \in \mathbb{R}^{n^{\mathrm{f}}_{\boldsymbol{x}}}$, $\boldsymbol{u}(k^{\mathrm{f}}) \in \mathbb{R}^{n^{\mathrm{f}}_{\boldsymbol{u}}}$, $\boldsymbol{y}(k^{\mathrm{f}}) \in \mathbb{R}^{n^{\mathrm{f}}_{\boldsymbol{y}}}$, $\boldsymbol{d}(k^{\mathrm{f}}) \in \mathbb{R}^{n^{\mathrm{f}}_{\boldsymbol{d}}}$ and $\boldsymbol{n}(k^{\mathrm{f}}) \in \mathbb{R}^{n^{\mathrm{f}}_{\boldsymbol{n}}}$, where all the symbols denote the standard variables and $\boldsymbol{n}$ is measurement noise. $n^{\mathrm{s}}_{\boldsymbol{x}}$, $n^{\mathrm{s}}_{\boldsymbol{u}}$, $n^{\mathrm{s}}_{\boldsymbol{y}}$ and $n^{\mathrm{f}}_{\boldsymbol{x}}$, $n^{\mathrm{f}}_{\boldsymbol{u}}$, $n^{\mathrm{f}}_{\boldsymbol{y}}$ denote the dimensions of the state vector, input vector and output vector of the fast and slow subsystem, respectively. The overall system under control can be decomposed in more than two layers, giving rise to the following description for subsystem i

$$\begin{cases} \boldsymbol{x}^i(t+1) &= \boldsymbol{f}^i(\boldsymbol{x}^i(t), \boldsymbol{x}^j(t), \ldots, \boldsymbol{x}^{N_s}, \boldsymbol{u}(t), \boldsymbol{d}^i(t), t), \\ \boldsymbol{y}^i(t) &= \boldsymbol{g}^i(\boldsymbol{x}^i(t), \boldsymbol{x}^j(t), \ldots, \boldsymbol{x}^{N_s}, \boldsymbol{u}(t), \boldsymbol{n}^i(t), t), \end{cases} \tag{2-30}$$

with N_s denoting the number of subsystems.

2-3-2 Functional system decomposition

When two sub-problems with different control objectives are treated in separate control layers, functional decomposition takes place [30]. The layers should include various types of partial objectives. The objectives can consist of economic criteria and tracking criteria, for example.

A two-layer structure is often adopted with the upper layer calculating the reference setpoints that the lower level must track. The upper layer is then the so-called supervisory constraint control layer, and the lower layer is the direct control layer. Then, the initial conditions for the lower level Dynamic real-time optimisation (DRTO) are estimated by measurements of the process [31].

The general formulation of the two layered DRTO-problem is as follows.

$$\begin{aligned}
\min_{\boldsymbol{u}^{\mathrm{s}}_{\mathrm{gh}}(k^{\mathrm{s}})} \quad & J^{\mathrm{s}}\left(\boldsymbol{x}^{\mathrm{s}}(k^{\mathrm{s}}), \boldsymbol{u}^{\mathrm{s}}(k^{\mathrm{s}}), \boldsymbol{d}^{\mathrm{s}}(k^{\mathrm{s}}), \boldsymbol{x}^{\mathrm{f}}(k^{\mathrm{s}}), N^{\mathrm{s}}_{\mathrm{h}}\right), \\
\text{s.t.} \quad & \boldsymbol{x}^{\mathrm{s}}(j^{\mathrm{s}}+1) = \boldsymbol{f}^{\mathrm{s}}(\boldsymbol{x}^{\mathrm{s}}(j^{\mathrm{s}}), \boldsymbol{u}^{\mathrm{s}}(j^{\mathrm{s}}), \boldsymbol{d}^{\mathrm{s}}(j^{\mathrm{s}})), \forall j^{\mathrm{s}} = k^{\mathrm{s}}, \ldots, k^{\mathrm{s}} + N^{\mathrm{s}}_{\mathrm{h}} - 1, \\
& \boldsymbol{x}^{\mathrm{s}}(k^{\mathrm{s}}) = \boldsymbol{x}^{\mathrm{s}}_0, \\
& \boldsymbol{y}^{\mathrm{s}}(j^{\mathrm{s}}) = \boldsymbol{g}^{\mathrm{s}}(\boldsymbol{x}^{\mathrm{s}}(j^{\mathrm{s}}), \boldsymbol{u}^{\mathrm{s}}(j^{\mathrm{s}}), \boldsymbol{d}^{\mathrm{s}}(j^{\mathrm{s}})), \forall j^{\mathrm{s}} = k^{\mathrm{s}}, \ldots, k^{\mathrm{s}} + N^{\mathrm{s}}_{\mathrm{h}} - 1, \\
& \boldsymbol{u}^{\mathrm{s}}(j^{\mathrm{s}}) \in \mathcal{U}^{\mathrm{s}}, \forall j^{\mathrm{s}} = k^{\mathrm{s}}, \ldots, k^{\mathrm{s}} + N^{\mathrm{s}}_{\mathrm{h}} - 1, \\
& \boldsymbol{x}^{\mathrm{s}}(j^{\mathrm{s}}) \in \mathcal{X}^{\mathrm{s}}, \forall j^{\mathrm{s}} = k^{\mathrm{s}}, \ldots, k^{\mathrm{s}} + N^{\mathrm{s}}_{\mathrm{h}} - 1, \\
& \boldsymbol{y}^{\mathrm{s}}(j^{\mathrm{s}}) \in \mathcal{Y}^{\mathrm{s}}, \forall j^{\mathrm{s}} = k^{\mathrm{s}}, \ldots, k^{\mathrm{s}} + N^{\mathrm{s}}_{\mathrm{h}} - 1,
\end{aligned} \tag{2-31}$$

where $J^{\mathrm{s}} = J^{\mathrm{eco}}$ in DRTO. The general lower layer optimisation problem is then given by

$$\begin{aligned}
\min_{\boldsymbol{u}^{f}_{\mathrm{gh}}(k^{\mathrm{f}})} \quad & J^{\mathrm{f}}\left(\boldsymbol{x}^{\mathrm{f}}(k^{\mathrm{f}}), \boldsymbol{u}^{\mathrm{f}}(k^{\mathrm{f}}), \boldsymbol{d}^{\mathrm{f}}(k^{\mathrm{f}}), \boldsymbol{x}^{\mathrm{s}}(k^{\mathrm{s}}), N^{\mathrm{f}}_{\mathrm{h}}\right), \\
\text{s.t.} \quad & \boldsymbol{x}^{\mathrm{f}}(j^{\mathrm{f}}+1) = \boldsymbol{f}^{\mathrm{f}}(\boldsymbol{x}^{\mathrm{f}}(j^{\mathrm{f}}), \boldsymbol{u}^{\mathrm{f}}(j^{\mathrm{f}}), \boldsymbol{d}^{\mathrm{f}}(j^{\mathrm{f}})), \forall j^{\mathrm{f}} = k^{\mathrm{f}}, \ldots, k^{\mathrm{f}} + N^{\mathrm{f}}_{\mathrm{h}} - 1, \\
& \boldsymbol{x}^{\mathrm{f}}(k^{\mathrm{f}}) = \boldsymbol{x}^{\mathrm{f}}_0, \\
& \boldsymbol{y}^{\mathrm{f}}(j^{\mathrm{f}}) = \boldsymbol{g}^{\mathrm{f}}(\boldsymbol{x}^{\mathrm{f}}(j^{\mathrm{f}}), \boldsymbol{u}^{\mathrm{f}}(j^{\mathrm{f}}), \boldsymbol{d}^{\mathrm{f}}(j^{\mathrm{f}})), \forall j^{\mathrm{f}} = k^{\mathrm{f}}, \ldots, k^{\mathrm{f}} + N^{\mathrm{f}}_{\mathrm{h}} - 1, \\
& \boldsymbol{u}^{\mathrm{f}}(j^{\mathrm{f}}) \in \mathcal{U}^{\mathrm{f}}, \forall j^{\mathrm{f}} = k^{\mathrm{f}}, \ldots, k^{\mathrm{f}} + N^{\mathrm{f}}_{\mathrm{h}} - 1, \\
& \boldsymbol{x}^{\mathrm{f}}(j^{\mathrm{f}}) \in \mathcal{X}^{\mathrm{f}}, \forall j^{\mathrm{f}} = k^{\mathrm{f}}, \ldots, k^{\mathrm{f}} + N^{\mathrm{f}}_{\mathrm{h}} - 1, \\
& \boldsymbol{y}^{\mathrm{f}}(j^{\mathrm{f}}) \in \mathcal{Y}^{\mathrm{f}}, \forall j^{\mathrm{f}} = k^{\mathrm{f}}, \ldots, k^{\mathrm{f}} + N^{\mathrm{f}}_{\mathrm{h}} - 1,
\end{aligned} \tag{2-32}$$

where $J^{\mathrm{f}} = J^{\mathrm{track}}$ in DRTO.

2-3-3 Control architecture

The proposed control architecture is shown in Figure 2-6 below. This control architecture decomposes the overall problem of controlling the crop towards maximum fruit yield while minimising the resource usage with a temporal system decomposition and a functional system decomposition as discussed above.

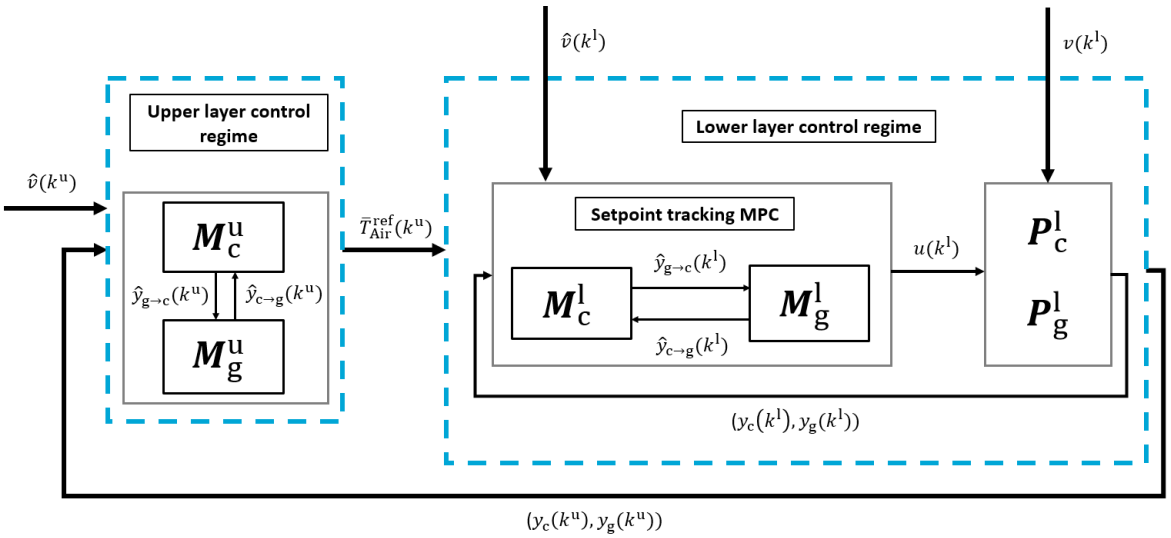


Figure 2-6: Proposed control architecture.

here, $M_{\mathrm{c}}^{\mathrm{s}}$ and $M_{\mathrm{g}}^{\mathrm{s}}$ are the crop control model and indoor climate control model used for the upper layer setpoint generating control scheme, $M_{\mathrm{c}}^{\mathrm{f}}$ and $M_{\mathrm{g}}^{\mathrm{f}}$ are the crop control model and indoor climate control model used for the lower layer setpoint tracking control scheme, and $P_{\mathrm{c}}^{\mathrm{l}}$ and $P_{\mathrm{g}}^{\mathrm{l}}$ are the crop system plant prediction model and indoor climate plant prediction model at lower layer sampling times, respectively.

The other upper layer entities seen in the figure represent the following, all at upper layer sampling time k^{s}. $\bar{T}_{\mathrm{Air}}(k^{\mathrm{s}})$ is the reference greenhouse indoor air temperature generated by the upper layer control regime at time instance k^{s}. $\hat{\boldsymbol{v}}^{\mathrm{s}}(k^{\mathrm{s}})$ is the upper layer weather prediction with sampling time τ^{s}. $\hat{\boldsymbol{y}}_{\mathrm{g}\to\mathrm{c}}(k^{\mathrm{s}})$ is the upper layer greenhouse indoor climate output vector passed to the upper layer crop model, $\hat{\boldsymbol{y}}_{\mathrm{c}\to\mathrm{g}}(k^{\mathrm{s}})$ is the upper layer crop output vector passed to the upper layer greenhouse indoor climate model.

The other lower layer entities seen in the figure represent the following, all at lower layer sampling time k^{f}. $\hat{\boldsymbol{v}}(k^{\mathrm{f}})$ is the lower layer weather prediction with sampling time τ^{l}. $\hat{\boldsymbol{y}}_{\mathrm{g}\to\mathrm{c}}(k^{\mathrm{f}})$ is the lower layer greenhouse indoor climate output vector passed to the lower layer crop model, $\hat{\boldsymbol{y}}_{\mathrm{c}\to\mathrm{g}}(k^{\mathrm{f}})$ is the lower layer crop output vector passed to the upper layer greenhouse indoor climate model. $\boldsymbol{v}(k^{\mathrm{f}})$ is the lower layer true weather. $\boldsymbol{u}(k^{\mathrm{f}})$ is the control input vector as calculated by the lower layer control regime that is implemented on the simulation models that act as plant. $\boldsymbol{y}_{\mathrm{c}}(k^{\mathrm{s}})$ and $\boldsymbol{y}_{\mathrm{g}}(k^{\mathrm{s}})$ are the measured simulated crop and greenhouse indoor climate outputs that are measured on the lower layer and send each upper layer sampling time k^{s} to the upper layer control regime.

The upper layer does not contain a plant description; the plant descriptions are used in the lower layer. Hence an upper layer plant description is ominous.

In this setting, the upper or lower layer models can be swapped when better system descriptions become available.

To summarise, the upper layer control regime uses system descriptions with a larger sampling time and long-term weather predictions to generate greenhouse indoor air temperature setpoints that are sent to the lower layer control regime. Two control strategies are proposed for

this upper layer control regime. The lower layer control regime is a linear Model predictive control (MPC) setpoint tracking controller that uses the same linear model for the greenhouse indoor climate subsystem for control and simulation but uses a linear crop model for control and a non-linear crop simulation model.

Sampling time choice

To separate the overall problem, which is to maximise yield while keeping control inputs small, suitable models need to be constructed to use per layer.

The upper layer problem sampling time needs to be as large as possible to optimise over a large time horizon, preferably the entire growing season, while keeping computation time as small as possible. A large sampling time is needed for the crop model simulation. Furthermore, the sampling rate needs to capture the dynamical trends of the fast dynamical indoor climate subsystem. Therefore, an upper layer sampling time $\tau^{\mathrm{s}} = 6$ hours has been chosen for the upper layer models. The different diurnal and nocturnal dynamics of the greenhouse indoor climate are correctly captured using this sampling time.

The lower layer models need to capture the fast dynamics of the indoor climate subsystem. Since the AGC data has a sampling time of 5 minutes and also the non-linear model employs a sampling rate of 5 minutes, a lower layer sampling time $\tau^{\mathrm{f}} = 5$ minutes has been chosen.

The effect of the different sampling rates can be seen in Figures 2-14 to 2-17.

Communication

The communication protocol between the two layers needs to be established to implement the multilayer control schemes. The basic communication rules are given as follows, where the time instants are defined in terms of the fast timescale, i.e., the time instants with super- or subscript f, from the lower, faster-moving layer [32].

Basic rules:

1. At every instant k^{f}, each level is supposed to know the current value of its state and control;

2. At time $v^{\mathrm{s}}k^{\mathrm{s}}$ the high level communicates to the low level its current control value $u^{\mathrm{s}}(k^{\mathrm{s}})$ and the references $\left(\bar{x}^{\mathrm{s}}(k^{\mathrm{f}}), \bar{u}^{\mathrm{u}}(k^{\mathrm{f}})\right)$

Switching

In Figure 2-7 below, a schematic representation can be seen of the switching moment between the upper layer and lower layer.

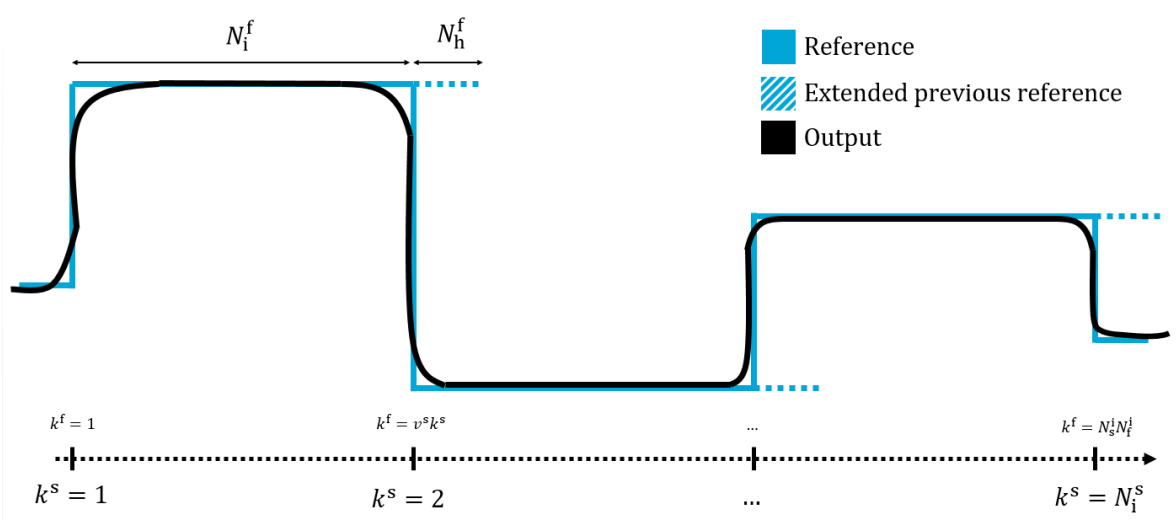


Figure 2-7: Switching protocol with extended previous reference trajectory. With k^s the current upper layer time instance, k^f the current lower layer time instance, N_i^f the number of implemented lower layer iterations per upper layer iteration k^s, N_h^f the lower layer prediction horizon, $v^s = N_i^f$ the number of lower layer iterations per upper layer iteration and N_i^s the number of implemented upper layer iterations.

where k^s the current upper layer time instance, k^f the current lower layer time instance, N_i^f the number of implemented lower layer iterations per upper layer iteration k^s, N_h^f the lower layer prediction horizon, $v^s = N_i^f$ the number of lower layer iterations per upper layer iteration and N_i^s the number of implemented upper layer iterations.

Note that in this communication scheme, the assumption is made that after the last lower layer iteration, the system is 'paused' while the upper layer recalculates the greenhouse indoor air reference temperature trajectory over the season. This is acceptable if the calculation time per upper layer iteration is significantly less than the time an upper layer iteration needs, which is the case. On average, the calculation time of one upper layer iteration amounts to approximately 1 minute, while the time one upper layer iteration encompasses is 6 hours, and the time one lower layer iteration encompasses is 5 minutes.

Proposed control schemes and models in the architecture

Several aspects of the control architecture overview above need to be filled with either a system representation or a control scheme. The proposed control schemes will be explained in this section.

- For the upper layer setpoint generating scheme and system representation, two controllers are proposed. The first proposed controller is a linear Economic model predictive control (EMPC), with a corresponding linear greenhouse model, elaborated in Chapter 4. The second proposed controller is a leveraged Data-enabled predictive control (DeePC) controller with corresponding data-based system representation, elaborated in Chapter 5.

- For the lower layer setpoint control regime, linear MPC is implemented, elaborated in Chapter 3.

- For the lower layer plant description of the crop subsystem, the non-linear calibrated mechanistic model is used. For the lower layer plant description of the greenhouse indoor climate subsystem, the same linear model for control and simulation is used.

2-4 Model use in the control architecture

In this section, the linear models used for the MPC control schemes as mentioned above are derived and characterised. The linear models are compared to the measurement data and against each other. Last, the model mismatch between the linear crop models and the non-linear crop model is characterised.

2-4-1 Prediction-error method (PEM) identified linear models

Besides implementing data-driven control methods and calibrating a mechanistic model to the available data to investigate whether it is possible to control the greenhouse system during the entire crop cycle, simple Linear time-invariant (LTI) models have been identified using the PEM [33].

To obtain the state-space system matrices, the following objection function is minimised.

$$J_{\mathrm{N}}(\boldsymbol{y}, A, B, C, D) = \sum_{t=1}^{N} e^2(t), \tag{2-33}$$

where the difference between the predicted output and the measured output at time t is denoted by $e(t)$. N denotes the number of samples.

Model characteristics

In Table 2-2, the model characteristics of the different identified models can be found. The following number of states per model resulted in the models with the best fit.

Sampling time	5min			6h		
	$\boldsymbol{y}_{\mathrm{g}}$	$\boldsymbol{y}_{\mathrm{c}}$	$\boldsymbol{y}_{\mathrm{gh}}$	$\boldsymbol{y}_{\mathrm{g}}$	$\boldsymbol{y}_{\mathrm{c}}$	$\boldsymbol{y}_{\mathrm{gh}}$
Number of states	3	4	7	3	4	5

Table 2-2: Number of states of the PEM identified linear models.

Step response analysis

The step response analysis in this section will cover the greenhouse indoor climate subsystem and is performed for the linear model with sampling time $\tau^{\mathrm{f}} = 5\mathrm{min}$.

For the crop subsystem, the following variable types can be considered as 'inputs' to the system for a step response analysis: $\boldsymbol{u}_{\mathrm{c}}$, $\boldsymbol{v}_{\mathrm{c}}$ and $\boldsymbol{y}_{\mathrm{g}\to\mathrm{c}}$. The crop subsystem is considered separately, since a step response analysis cannot be spoken of but rather a temperature dependency that lasts over the entire season.

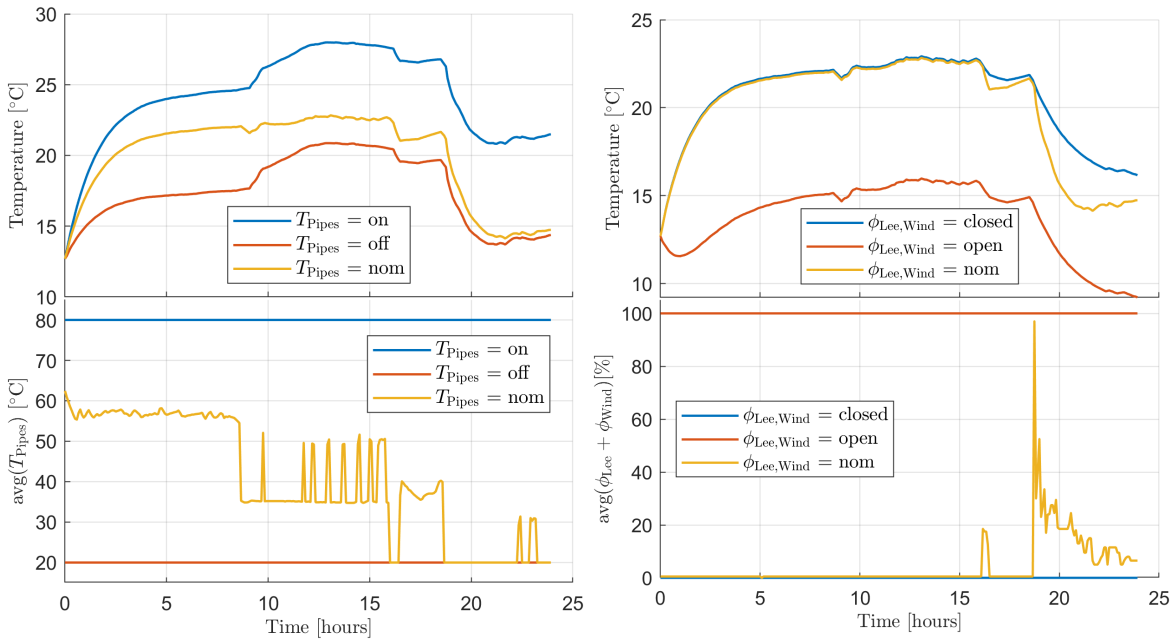
For the greenhouse indoor climate subsystem, the following variable types that can be considered as 'inputs' to the system will be covered in this specific type of step response analysis: $\boldsymbol{u}_{\mathrm{c}}$, $\boldsymbol{v}_{\mathrm{c}}$ and $\boldsymbol{y}_{\mathrm{c}\to\mathrm{g}}$.

For this step response analysis, all variables, except the variable considered for the step response, will be kept not at their minimal values but rather at the measured values for the specific day in the season the step response is considered for. This is because it is not a regular dynamical system, e.g., closing the windows is not giving no input but also giving an input.

First, the variables will be considered individually, but also the control inputs will be considered together. The weather type of variables will be considered individually and together as well.

The step responses for the first day of the growing season are depicted in Figures 2-8 to 2-13 below, to illustrate the influence of the different variables. Thereafter, the step responses are quantified.

In Figure 2-8 below, the "step response" of T_{Air} to $T_{\text{Pipe1,2}}$ and $\phi_{\text{Lee,Wind}}$ can be seen.

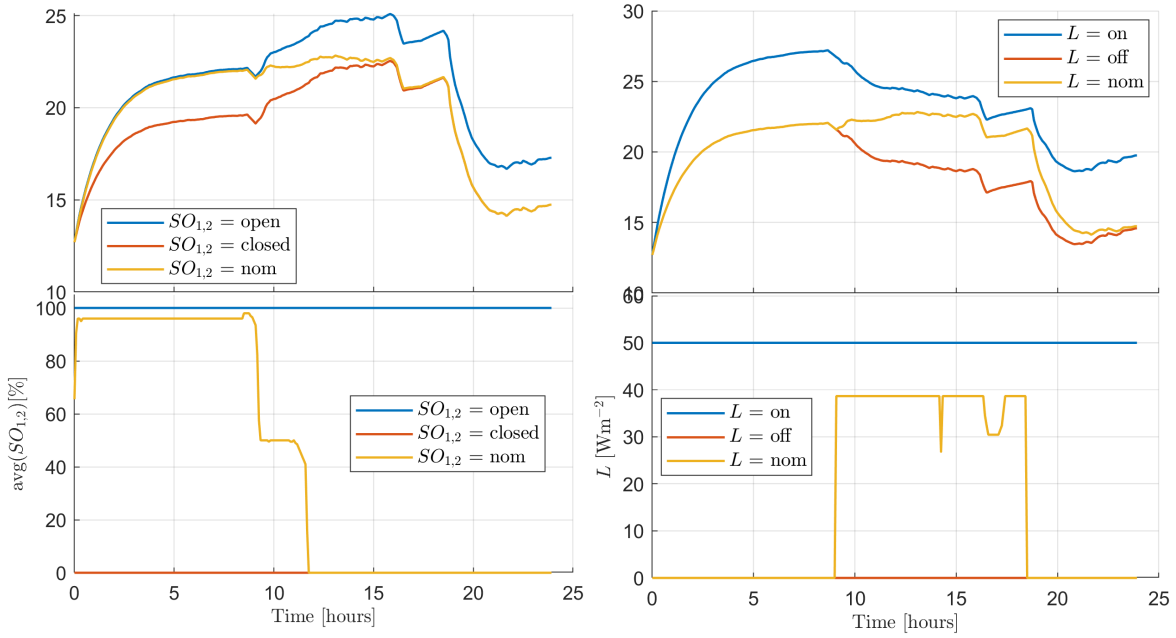


(a) "Step response" of T_{Air} to $T_{\text{Pipe1,2}}$.

(b) "Step response" of T_{Air} to $\phi_{\text{Lee,Wind}}$.

Figure 2-8: "Step response" of T_{Air} to $T_{\text{Pipe1,2}}$ and $\phi_{\text{Lee,Wind}}$ 1st day of the season.

In Figure 2-9 below, the "step response" of T_{Air} to $SO_{1,2}$ and L can be seen.

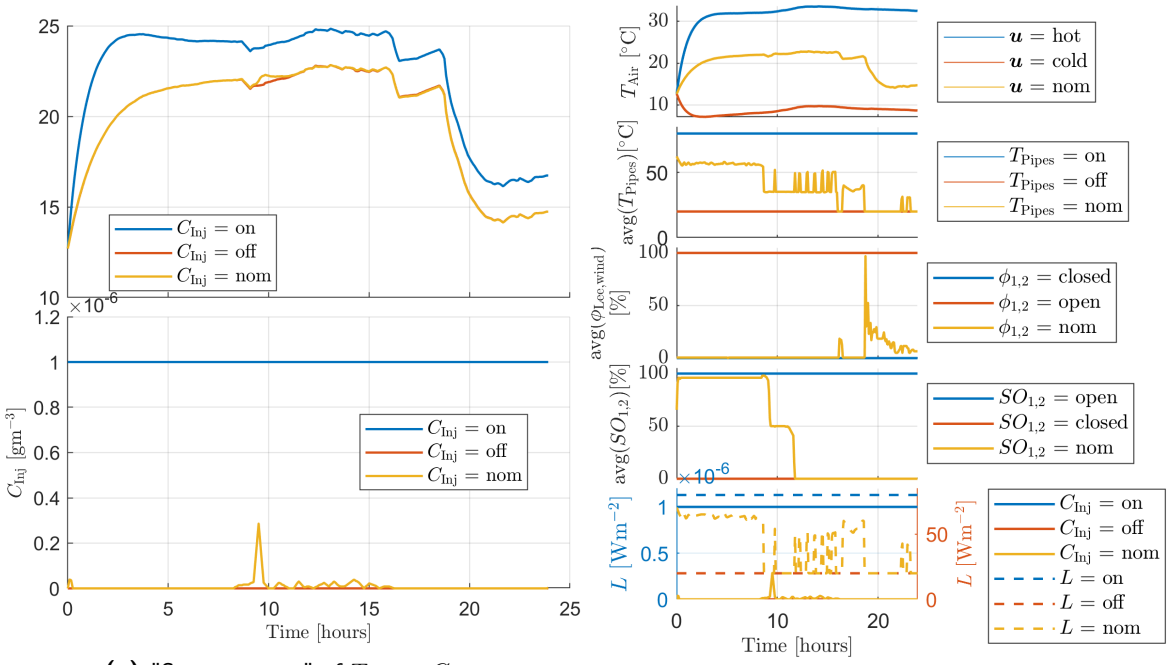


(a) "Step response" of T_{Air} to $SO_{1,2}$.

(b) "Step response" of T_{Air} to L.

Figure 2-9: "Step response" of T_{Air} to $SO_{1,2}$ and L 1st day of the season.

In Figure 2-10 below, the "step response" of T_{Air} to C_{Inj} and combined $\boldsymbol{u}$ can be seen.

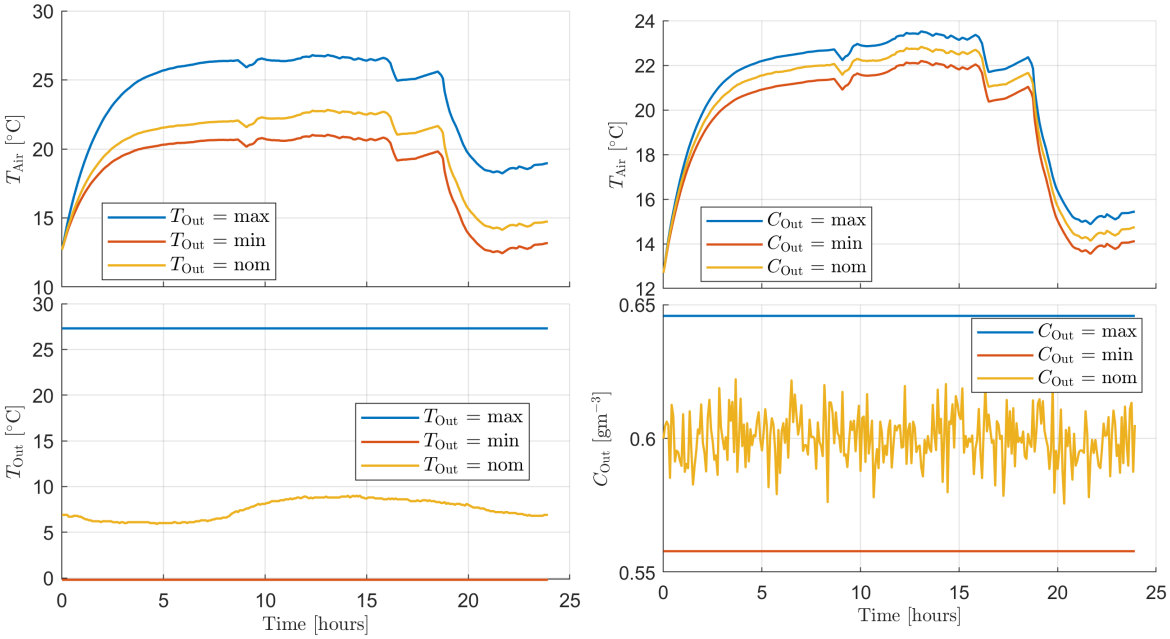


(a) "Step response" of T_{Air} to C_{Inj}.

(b) "Step response" of T_{Air} to $\boldsymbol{u}$.

Figure 2-10: "Step response" of T_{Air} to C_{Inj} and combined $\boldsymbol{u}$ 1st day of the season.

In Figure 2-11 below, the "step response" of T_{Air} to T_{Out} and C_{Out} can be seen.

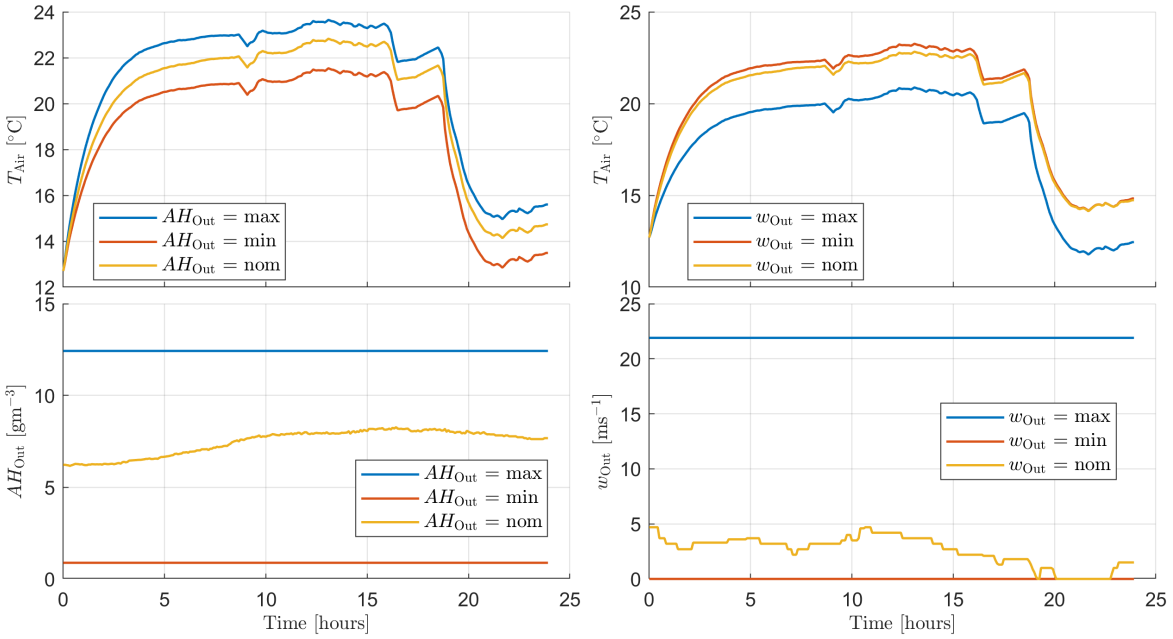


(a) "Step response" of T_{Air} to T_{Out}.

(b) "Step response" of T_{Air} to C_{Out}.

Figure 2-11: "Step response" of T_{Air} to T_{Out} and C_{Out} 1st day of the season.

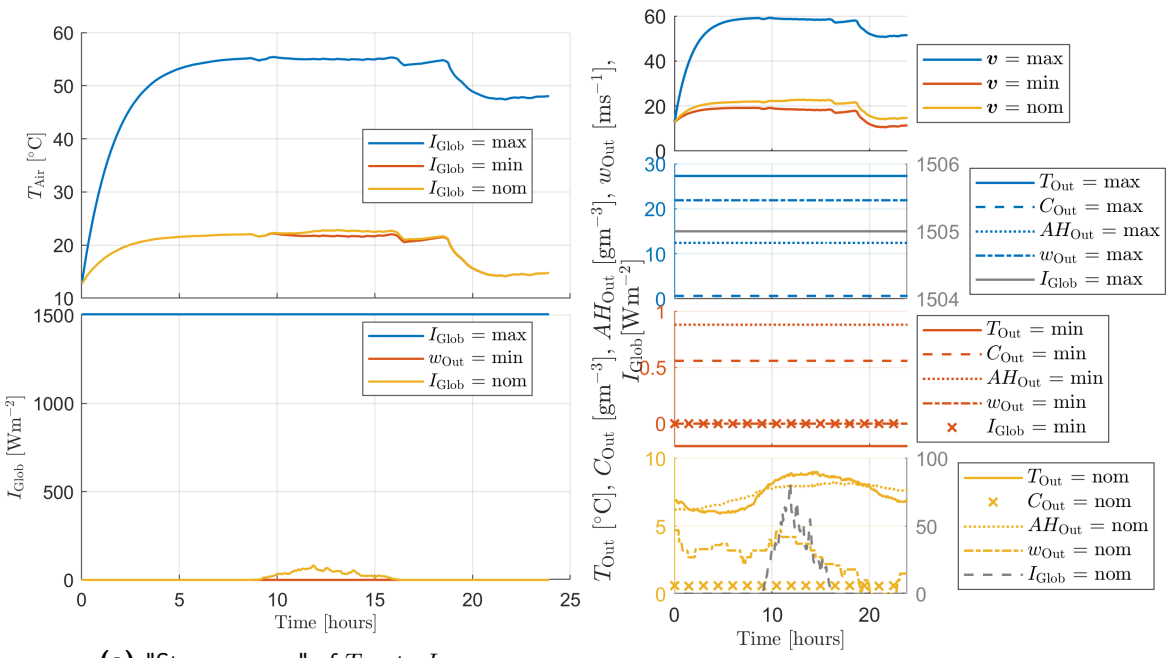
In Figure 2-12 below, the "step response" of T_{Air} to AH_{Air} and w_{Out} individually can be seen.



(a) "Step response" of T_{Air} to AH_{Out}. **(b)** "Step response" of T_{Air} to w_{Out}.

Figure 2-12: "Step response" of T_{Air} to AH_{Out} and w_{Out} 1st day of the season.

In Figure 2-13 below, the "step response" of T_{Air} to I_{Glob} individually and $\boldsymbol{v}$ together can be seen.



(a) "Step response" of T_{Air} to I_{Glob}. **(b)** "Step response" of T_{Air} to $\boldsymbol{v}$.

Figure 2-13: "Step response" of T_{Air} to I_{Glob} and combined $\boldsymbol{v}$ 1st day of the season.

The step responses are quantified in Table 2-3 below.

Table 2-3: An overview of the heating capacities per control input type in the greenhouse.

Actuator type	**Time to reach daily max(T_{Air})**					**+1 °C rise**
	Min [min]	Day	Max [min]	Day	Avg [min]	Avg [min]
T_{Pipes}	20	161	945	151	336	94
$\phi_{\text{Lee,Wind}}$	220	30	1030	54	627	240
$SO_{1,2}$	215	56	980	151	668	253
L	20	161	1020	84	704	195
C_{Inj}	10	161	955	151	475	33
All	10	161	605	151	98	12

The first five columns with values are calculated per day. Per actuator type, the actuators are set to the mode that heats the greenhouse the most and it is measured daily how many minutes it takes to reach the daily maximum temperature. The day for which it takes the least and most amount of time and the respective times are given as well. Furthermore, the average daily time it takes per actuator type to reach the maximum temperature also given.

The second five columns with values show the times it takes per actuator type to heat the greenhouse air with one degree. This is calculated for every sampling instance in the greenhouse.

2-4-2 Model mismatch

In this section, first, the linear PEM-identified models are compared to the data they have been identified with, in Figure 2-15 to Figure 2-17 below. Then, the PEM-identified crop models are compared to the calibrated non-linear crop model that will be used as a ground truth simulator in terms of the temperature dependency.

PEM models and data comparison

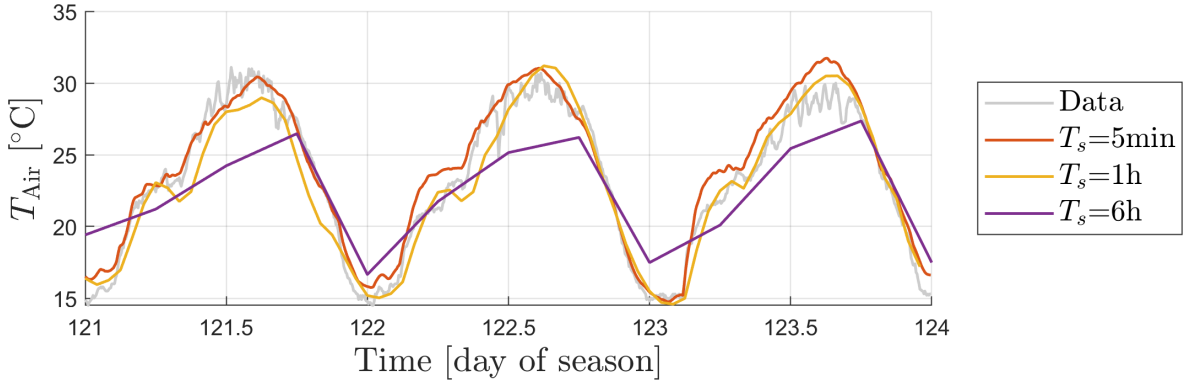


Figure 2-14: Linear PEM identified model T_{Air} comparison.

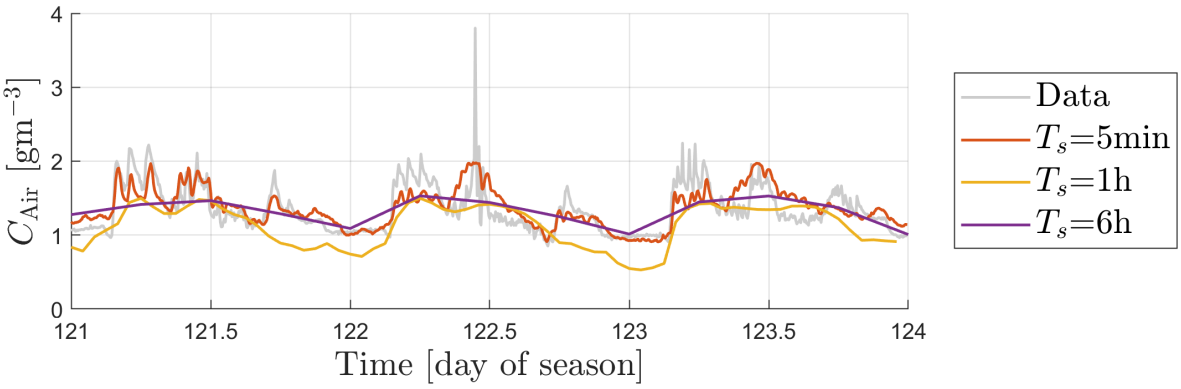


Figure 2-15: Linear PEM identified model C_{Air} comparison.

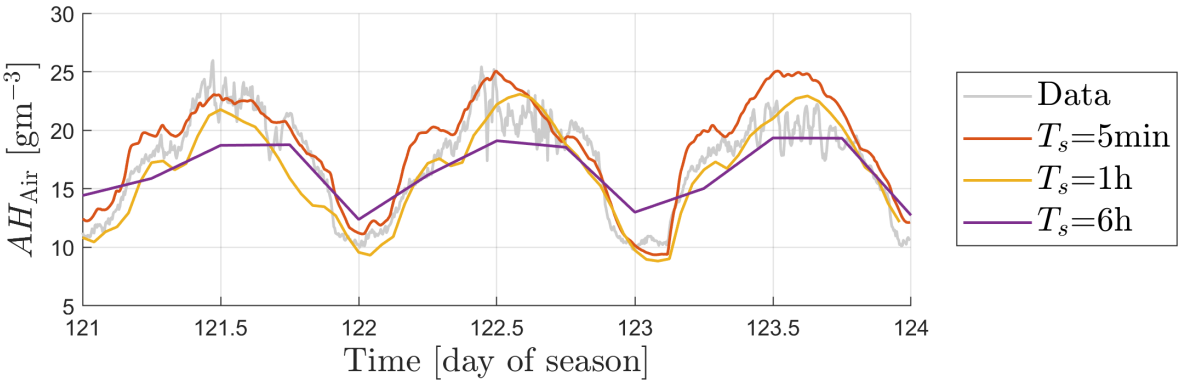


Figure 2-16: Linear PEM identified model AH_{Air} comparison.

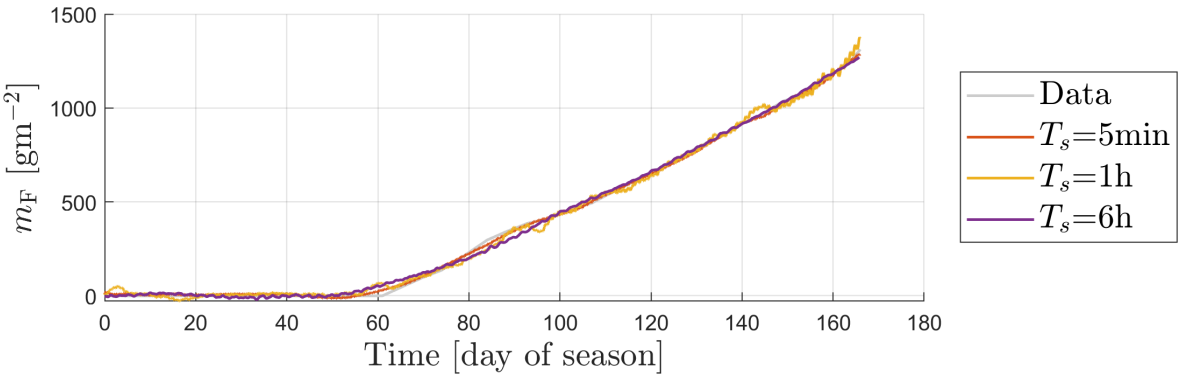


Figure 2-17: Linear PEM identified model m_{F} comparison.

It can be concluded that the model with the smallest sampling time matches the measurement data the best, while the model with the longest sampling time still captures the overall trend of the dynamics but shows quantitatively different behaviour.

Temperature dependency different crop models

The crop model is dependent on different variables; however since this thesis aims to generate greenhouse indoor air temperature setpoints for the overall problem, the temperature dependency is of importance.

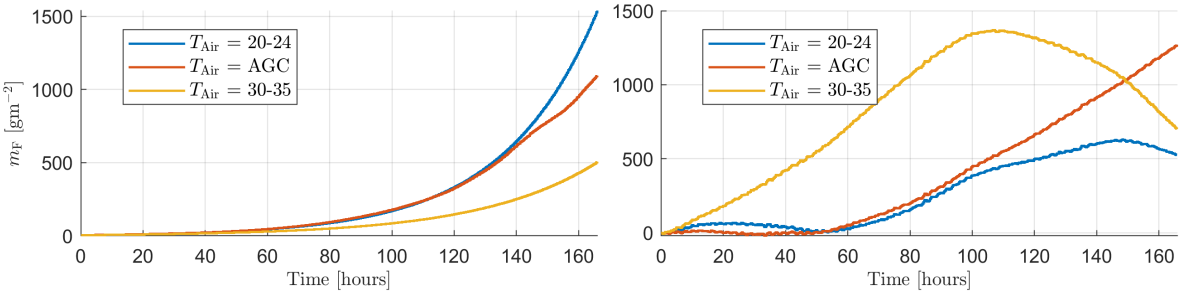


Figure 2-18: Temperature dependency of the non-linear crop model (left) and the linear crop model (right) with sampling time $\tau^s = 6$h.

The dynamics of the linear crop model do not resemble the dynamics of the non-linear model. Therefore, using the linear model in the upper layer EMPC setpoint generating controller can lead to a different temperature strategy than is optimal for the non-linear model. This is acceptable because, in reality, there is also a model mismatch.

2-5 Conclusive remarks on the problem definition

In this chapter, the problem has been introduced. Some conclusive remarks concerning to the interpretation of implementing these tools in the remainder of this work can be made.

The greenhouse-crop system has been introduced, the available data has been manipulated to use, and the benchmark to which the synthesised controllers will be compared is defined. The benchmark has been selected based on the literature. However, it can be noted that it does not change during the season, which would not apply in real life when considering the changing weather conditions. Hence, this influences the performance of this benchmark strategy.

A non-linear crop model has been introduced that is compatible with the set-up of the system description mentioned above. The adapted version is given, and it is calibrated. Note that the calibration was performed only over the fruit dry weight, using one dataset. Furthermore, note that for true crop control, this crop system representation is not sufficient in the broad sense since all crop control mechanisms that result from crop management decisions such as pruning and harvesting are omitted. However, for crop control via the greenhouse indoor climate strategy, it is sufficient.

The overall crop production control has been decomposed to make the overall problem tractable and to be able to employ the faster changing dynamical greenhouse indoor climate to control the slow varying dynamical crop system. The established communication protocol is based on the different timescales and sends the upper layer generated setpoints to track to the lower layer. Note that the lower layer main goal is to track the upper layer generated

reference, i.e., it is assumed that the generated setpoint can achieve the highest performance in crop yield and resource usage. This neglects the possible model mismatch and might hence result in performance degradation. This model degradation originates from the different models used in this thesis, of which the linear models are introduced and characterised after the control architecture has been set. The linear models show correct correspondence to the data that has been used to identify the models.

This chapter has proposed the tools, scope, and framework used for the controller synthesis in the remainder of this thesis. The aforementioned issues are considered, but the efficiency of different controllers for crop production control via the greenhouse indoor climate strategy can be investigated with said tools.

Chapter 3

Tracking Model predictive control (MPC)

Thus far, the different model-based system representation's of the greenhouse-crop cultivation system have been presented. Also, the overall problem has been decomposed via temporal system decomposition and spatial decomposition into a two-layered structure shown in Figure 2-6. This chapter describes the synthesis and evaluation of the setpoint tracking controller in the lower layer subproblem, which will be a Quadratic Programming (QP) MPC controller. The chapter starts with the introduction of the MPC framework and used linear model. Hereafter the MPC problem will be recast into the used QP framework. Subsequent sections will then present the obtained results of the lower layer controller tracking the benchmark greenhouse indoor air reference temperature trajectory and analyse them before moving on to the other setpoint generating generating controllers.

3-1 MPC framework

MPC uses a model description of the system under control. The algorithm calculates the optimal control input over the prediction horizon to reach an objective, which is often a reference tracking objective, which is depicted in Figure 3-1. Without loss of generality, a system is often represented with the following non-linear state-space representation

$$\begin{aligned} \boldsymbol{x}(k+1) &= \boldsymbol{f}(\boldsymbol{x}(k), \boldsymbol{u}(k)), \\ \boldsymbol{y}(k) &= \boldsymbol{g}(\boldsymbol{x}(k), \boldsymbol{u}(k)), \end{aligned} \tag{3-1}$$

where $\boldsymbol{x}(k) \in \mathbb{R}^{n_{\boldsymbol{x}}}$ is the system state vector, $\boldsymbol{u}(k) \in \mathbb{R}^{n_{\boldsymbol{u}}}$ is the control input vector and $\boldsymbol{y}(k) \in \mathbb{R}^{n_{\boldsymbol{y}}}$ is the system output vector at time k. $\boldsymbol{f} : \mathbb{R}^{n_{\boldsymbol{x}}} \times \mathbb{R}^{n_{\boldsymbol{x}}} \to \mathbb{R}^{n_{\boldsymbol{x}}}$ and $\boldsymbol{g} : \mathbb{R}^{n_{\boldsymbol{x}}} \times \mathbb{R}^{n_{\boldsymbol{x}}} \to \mathbb{R}^{n_{\boldsymbol{y}}}$ are the system state evolution and system output functions.

Thus, when an explicit system model is available, an optimisation problem can be solved at timestep k that is subject to the state and control input constraints and the system dynamics

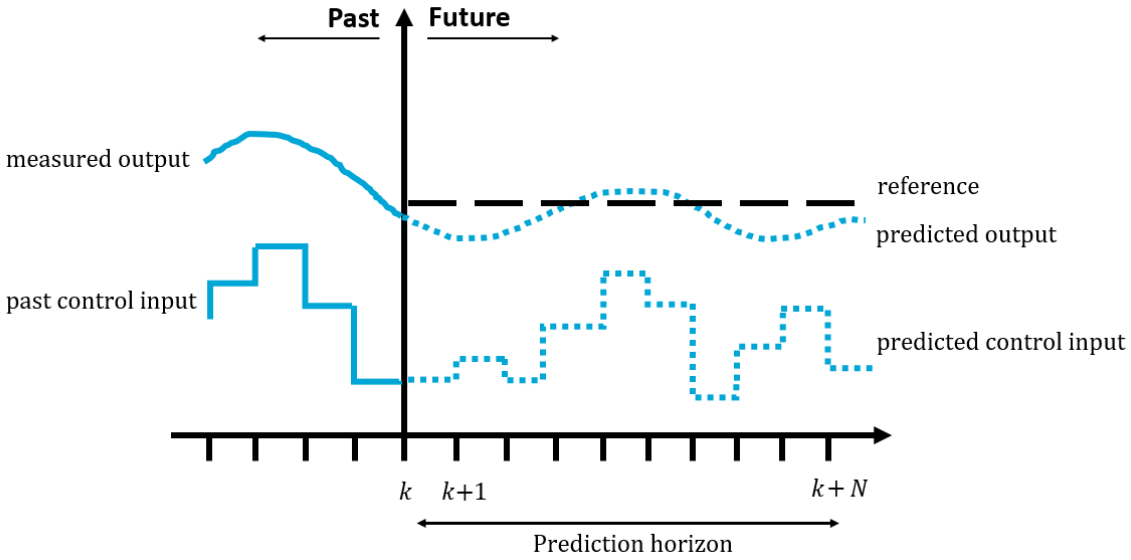


Figure 3-1: MPC scheme.

[34]. The general optimisation problem for the non-linear MPC case is given in (3-2) below.

$$\begin{aligned}
\min_{\boldsymbol{u}} \quad & J_{\mathrm{N}}(\boldsymbol{x},\boldsymbol{u},\boldsymbol{r}) = \sum_{k=0}^{N-1} \ell\left(\boldsymbol{x}(j),\boldsymbol{u}(j),\boldsymbol{r}(j)\right) + V_{\mathrm{f}}\left(\boldsymbol{x}_N,\boldsymbol{r}_N\right), \\
\text{s.t.} \quad & \boldsymbol{x}(j+1) = \boldsymbol{f}(\boldsymbol{x}(j),\boldsymbol{u}(j)), \quad \forall (j) \in \{k,\ldots,k+N-1\}, \\
& \boldsymbol{y}(j) = \boldsymbol{g}(\boldsymbol{x}(j),\boldsymbol{u}(j)), \quad \forall (j) \in \{k,\ldots,k+N-1\}, \\
& \boldsymbol{x}_0 = \hat{\boldsymbol{x}}(j), \\
& \boldsymbol{u}(j) \in \mathcal{U}, \quad \forall (j) \in \{k,\ldots,k+N-1\}, \\
& \boldsymbol{x}(j) \in \mathcal{X}, \quad \forall (j) \in \{k,\ldots,k+N-1\},
\end{aligned} \tag{3-2}$$

here $J_N(\boldsymbol{x},\boldsymbol{u},\boldsymbol{r})$ is the to-be minimised cost function. The first term, $\ell(\boldsymbol{x}_j,\boldsymbol{u}_j,\boldsymbol{r}_j)$ is the stage cost which is computed at each prediction time index j and $V_{\mathrm{f}}(\boldsymbol{x}_N,\boldsymbol{r}_N)$ is the terminal cost at the end of the prediction horizon, $j = N$.

$N \in \mathbb{Z}_{>0}$ is the prediction horizon, $\boldsymbol{u} = (\boldsymbol{u}_k,\cdots,\boldsymbol{u}_{k+N-1})$ are the control input vectors and arguments of the optimisation problem, $\boldsymbol{x} = (\boldsymbol{x}_k \ldots \boldsymbol{x}_{k+N-1})$ are the predicted state vectors, $\boldsymbol{y} = (\boldsymbol{y}_k \ldots \boldsymbol{y}_{k+N-1})$ are the predicted system output vectors and $\boldsymbol{r} = (\boldsymbol{r}_k,\ldots,\boldsymbol{r}_{k+N-1})$ are the output reference trajectory vectors. Furthermore, the initial state $\boldsymbol{x}_0$ for the optimisation problem is the state estimation $\hat{\boldsymbol{x}}(k)$ at time instance k at which the optimisation problem is solved. If full state information is available, $\hat{\boldsymbol{x}}(k) = \boldsymbol{x}(k)$. If no full state information is available, the state is often estimated via an observer. $\mathcal{U}$ and $\mathcal{X}$ are the control input and the state constraint sets, respectively.

The non-linear MPC algorithm is then given in Algorithm (1) below [35].

Algorithm 1 MPC algorithm

Input: System functions $\boldsymbol{f}(\boldsymbol{x}(k), \boldsymbol{u}(k))$, prediction horizon N, constraint sets $\mathcal{U}$ and $\mathcal{X}$;

1: Obtain initial state estimate $\hat{\boldsymbol{x}}(k)$.
2: Solve (3-2) for the optimal input sequence $\boldsymbol{u}^* = (\boldsymbol{u}_0^*, \cdots, \boldsymbol{u}_{N-1}^*)$.
3: Apply only the first input $\boldsymbol{u}(k) = \boldsymbol{u}_0^*$.
4: Set k to $k+1$.
5: Return to 1.

3-2 Linear Discrete-time (DT) state-space prediction model

For the lower layer setpoint tracking controller, a linear state-space description is used as introduced in Chapter 2. The prediction model that has been used is given by

$$\begin{cases} \boldsymbol{x}(k+1) = A\boldsymbol{x}(k) + B_u\boldsymbol{u}(k) + B_v\boldsymbol{v}(k), \\ \boldsymbol{y}(k) = C\boldsymbol{x}(k), \end{cases} \tag{3-3}$$

where $\boldsymbol{y}(k)$ is obtained by stacking $\boldsymbol{y} = [\boldsymbol{y}_\text{g} \quad \boldsymbol{y}_\text{c}]^\top$ as given by (2-1) and (2-2), $\boldsymbol{x}(k)$ are internal states, $\boldsymbol{u}(k)$ are all the eight control inputs to the greenhouse as given by $\boldsymbol{u}_\text{gh}$ in (2-3) and $\boldsymbol{v}(k)$ are the five weather exogenous inputs to the greenhouse as given by $\boldsymbol{v}_\text{gh}$ in (2-5).

As mentioned in Chapter 2, the dimensions of the implemented state-space system description are as follows. $A \in \mathbb{R}^{7\times7}$, $B_{\boldsymbol{u}} \in \mathbb{R}^{7\times8}$, $B_{\boldsymbol{v}} \in \mathbb{R}^{7\times5}$ and $C \in \mathbb{R}^{5\times7}$, and the sampling time for the implemented system description is, as mentioned in Chapter 2, $\tau^l = 300$ [s]. The pole locations of the linear system are plotted in Figure 3-2 below.

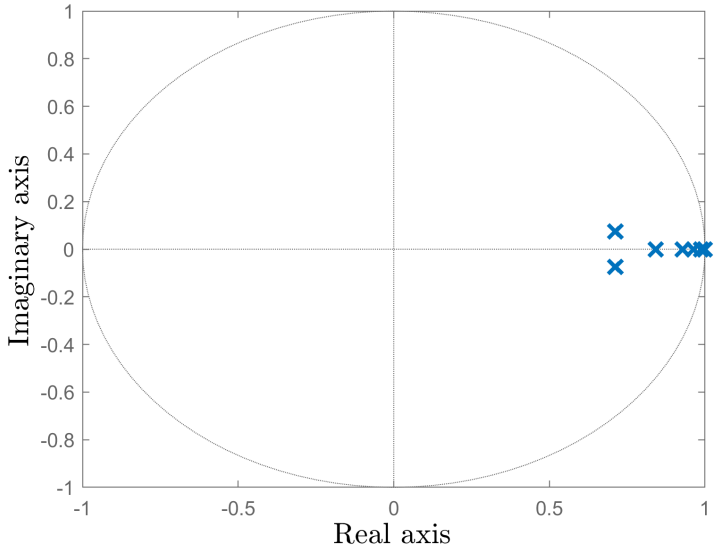


Figure 3-2: Pole locations of the used greenhouse linear model for the lower layer setpoint tracking control.

All but one pole are located in the interior of the unit circle. Except for the pole that numerically resides on the unit circle, the system would be stable. The corresponding natural

frequencies and damping ratios per pole are

$$\omega_n = \left|\frac{\ln(z)}{\tau^l}\right| = \begin{bmatrix} 9.63 \cdot 10^{-8} \\ 4.29 \cdot 10^{-5} \\ 1.27 \cdot 10^{-4} \\ 2.52 \cdot 10^{-4} \\ 5.76 \cdot 10^{-4} \\ 1.17 \cdot 10^{-3} \\ 1.17 \cdot 10^{-3} \end{bmatrix}, \quad \zeta_n = -\cos((\angle z)) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0.96 \\ 0.96 \end{bmatrix}, \tag{3-4}$$

As can be seen, five of the seven of the modes are critically damped with damping ratio of 1.

3-2-1 Augmented system

Extra tracking error states are added to obtain an augmented system that takes into account the deviation of the reference. The extra states $\boldsymbol{q}(k)$ are given by

$$\boldsymbol{q}(k+1) = (\boldsymbol{y}(k) - \boldsymbol{r}(k)), \tag{3-5}$$

The augmented system's matrices are given by

$$\begin{cases} \begin{bmatrix} \boldsymbol{x}(k+1) \\ \boldsymbol{q}(k+1) \end{bmatrix} = \begin{bmatrix} A & \mathbf{0}_{n_{\boldsymbol{x}} \times n_{\boldsymbol{y}}} \\ C & \mathbf{0}_{n_{\boldsymbol{y}} \times n_{\boldsymbol{y}}} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}(k) \\ \boldsymbol{q}(k) \end{bmatrix} + \begin{bmatrix} B_{\boldsymbol{u}} \\ \mathbf{0}_{n_{\boldsymbol{y}} \times n_{\boldsymbol{u}}} \end{bmatrix} \boldsymbol{u}(k) + \begin{bmatrix} B_{\boldsymbol{v}} \\ \mathbf{0}_{n_{\boldsymbol{y}} \times n_{\boldsymbol{v}}} \end{bmatrix} \boldsymbol{v}(k) + \begin{bmatrix} \mathbf{0}_{n_{\boldsymbol{x}} \times n_{\boldsymbol{y}}} \\ -I \end{bmatrix} \boldsymbol{r}(k), \\ \boldsymbol{y}(k) = \begin{bmatrix} C & \mathbf{0}_{n_{\boldsymbol{x}} \times n_{\boldsymbol{y}}} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}(k) \\ \boldsymbol{q}(k) \end{bmatrix}, \end{cases} \tag{3-6}$$

The state-space formulation of the augmented system is given by

$$\begin{cases} \tilde{\boldsymbol{x}}(k+1) = \tilde{A}\tilde{\boldsymbol{x}}(k) + \tilde{B}_{\boldsymbol{u}}\boldsymbol{u}(k) + \tilde{B}_{\mathrm{v}}\boldsymbol{v}(k) + \tilde{B}_{\boldsymbol{r}}\boldsymbol{r}(k), \\ \boldsymbol{y}(k) = \tilde{C}\tilde{\boldsymbol{x}}(k), \end{cases} \tag{3-7}$$

where variables with a tilde denote augmented system variables. It can also be rewritten as

$$\begin{cases} \tilde{\boldsymbol{x}}(k+1) = \tilde{A}\tilde{\boldsymbol{x}}(k) + \tilde{B}\tilde{\boldsymbol{u}}(k), \\ \boldsymbol{y}(k) = \tilde{C}\tilde{\boldsymbol{x}}(k), \end{cases} \tag{3-8}$$

where $\tilde{B}$ is given by

$$\tilde{B} = \begin{bmatrix} B_{\boldsymbol{u}} & B_{\boldsymbol{v}} & \mathbf{0}_{n_{\boldsymbol{x}} \times n_{\boldsymbol{y}}} \\ \mathbf{0}_{n_{\boldsymbol{y}} \times n_{\boldsymbol{u}}} & \mathbf{0}_{n_{\boldsymbol{y}} \times n_{\boldsymbol{v}}} & -I_{n_{\boldsymbol{y}} \times n_{\boldsymbol{y}}} \end{bmatrix}, \tag{3-9}$$

and the augmented input vector $\tilde{\boldsymbol{u}}$ is given by

$$\tilde{\boldsymbol{u}} = \begin{bmatrix} \boldsymbol{u}(k) \\ \boldsymbol{v}(k) \\ \boldsymbol{r}(k) \end{bmatrix}, \tag{3-10}$$

The dimensions of the state-space matrices of the augmented system are as follows. $\tilde{A} \in \mathbb{R}^{(n_{\boldsymbol{x}}+n_{\boldsymbol{y}})\times(n_{\boldsymbol{x}}+n_{\boldsymbol{y}})}$, $\tilde{B}_{\boldsymbol{u}} \in \mathbb{R}^{(n_{\boldsymbol{x}}+n_{\boldsymbol{y}})\times(n_{\boldsymbol{u}})}$, $\tilde{B}_{\boldsymbol{v}} \in \mathbb{R}^{(n_{\boldsymbol{x}}+n_{\boldsymbol{y}})\times(n_{\boldsymbol{v}})}$, $\tilde{B}_{\boldsymbol{r}} \in \mathbb{R}^{(n_{\boldsymbol{x}}+n_{\boldsymbol{y}})\times(n_{\boldsymbol{y}})}$, $\tilde{C} \in \mathbb{R}^{(n_{\boldsymbol{y}})\times(n_{\boldsymbol{x}}+n_{\boldsymbol{y}})}$ and $\tilde{D} \in \mathbb{R}^{(n_{\boldsymbol{y}})\times(n_{\boldsymbol{x}}+n_{\boldsymbol{y}})}$. With abuse of notation, the remainder of this chapter uses variables without tilde to denote the augmented system, which will be used further on.

3-3 QP controller design

Since a linear state-space system description is employed, the optimisation problem can be recast into a QP problem to increase the problem's solving speed; a QP program can be solved in polynomial time [36].

The cost function to be minimised in order to track the temperature in the greenhouse is defined in (3-11)

$$J_N\left(\boldsymbol{x},\boldsymbol{u},\boldsymbol{r}\right)=\sum_{j^{\mathrm{f}}=k^{\mathrm{f}}}^{k^{\mathrm{f}}+N-1}\boldsymbol{x}(j^{\mathrm{f}})^{\top}Q_c\boldsymbol{x}(j^{\mathrm{f}}), \tag{3-11}$$

where Q_c is the state cost matrix that contains tuning variables in this control setting. Q_c is set as $\mathrm{diag}([0,0,0,0,0,1,0,0,0,0])$, to penalise the deviation from the reference temperature trajectory. Note that this results in a positive semi-definite cost and also no terminal cost is implemented. Here, superscript f denotes that the problem deals with the lower layer, fast-dynamical system. N is the control horizon, which is equal to the prediction horizon in this case. For this application, $N=12$, a control horizon of 1 hour, seemed a good compromise between speed and performance.

The optimisation problem for the MPC controller is given by

$$\begin{aligned}
\boldsymbol{u}_N^{*}=\ & \underset{\boldsymbol{u}_N}{\operatorname{argmin}} \quad J_N\left(\boldsymbol{x}_N,\boldsymbol{u}_N\right), \\
\text{s.t.}\quad & \boldsymbol{x}(j^{\mathrm{f}}+1)=A\boldsymbol{x}(j^{\mathrm{f}})+B_{\boldsymbol{u}}\boldsymbol{u}(j^{\mathrm{f}})+B_{\boldsymbol{v}}\boldsymbol{v}(j^{\mathrm{f}}) && \forall j^{\mathrm{f}}=k^{\mathrm{f}},\ldots,k^{\mathrm{f}}+N-1,\\
& \boldsymbol{y}(j^{\mathrm{f}})=C\boldsymbol{x}(j^{\mathrm{f}}) && \forall j^{\mathrm{f}}=k^{\mathrm{f}},\ldots,k^{\mathrm{f}}+N-1,\\
& A_{\boldsymbol{u},\mathrm{u}}\boldsymbol{u}(j^{\mathrm{f}})\leq b_{\boldsymbol{u},\mathrm{u}} && \forall j^{\mathrm{f}}=k^{\mathrm{f}},\ldots,k^{\mathrm{f}}+N-1,\\
& A_{\boldsymbol{u},\mathrm{l}}\boldsymbol{u}(j^{\mathrm{f}})\geq b_{\boldsymbol{u},\mathrm{l}} && \forall j^{\mathrm{f}}=k^{\mathrm{f}},\ldots,k^{\mathrm{f}}+N-1,\\
& A_{\boldsymbol{y},\mathrm{u}}\boldsymbol{y}(j^{\mathrm{f}})\leq b_{\boldsymbol{y},\mathrm{u}} && \forall j^{\mathrm{f}}=k^{\mathrm{f}},\ldots,k^{\mathrm{f}}+N-1,\\
& A_{\boldsymbol{y},\mathrm{l}}\boldsymbol{y}(j^{\mathrm{f}})\geq b_{\boldsymbol{y},\mathrm{l}} && \forall j^{\mathrm{f}}=k^{\mathrm{f}},\ldots,k^{\mathrm{f}}+N-1,\\
& \boldsymbol{x}(k^{\mathrm{f}})=\boldsymbol{x}_0.
\end{aligned} \tag{3-12}$$

Note that in this setting, full state information is assumed and also the short-term weather prediction will be identical to the real implemented weather, i.e., $\hat{\boldsymbol{v}}_{\mathrm{gh}}=\boldsymbol{v}_{\mathrm{gh}}$. The problem can be rewritten as a QP problem by rewriting the sequence of outputs $\boldsymbol{x}(1),\cdots,\boldsymbol{x}(N)$ as a linear combination of the vector of control inputs $\boldsymbol{u}_N$, exogeneous inputs $\boldsymbol{v}_N$, references $\boldsymbol{r}_N$, and the current state $\boldsymbol{x}(k)$ so the equality constraint can be eliminated. This is done as

follows:

$$\begin{aligned}\boldsymbol{x}_N &= \underbrace{\begin{bmatrix} B_u & 0 & \dots & 0 \\ AB_u & B_u & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}B_u & A^{N-2}B_u & \dots & B_u \end{bmatrix}}_{\Gamma_N^{\boldsymbol{u}}} \boldsymbol{u}_N + \underbrace{\begin{bmatrix} A \\ A^2 \\ \vdots \\ A^N \end{bmatrix}}_{\Gamma_N^{\boldsymbol{x}}} \boldsymbol{x}(k) \\ &+ \underbrace{\begin{bmatrix} B_v & 0 & \dots & 0 \\ AB_v & B_v & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}B_v & A^{N-2}B_v & \dots & B_v \end{bmatrix}}_{\Gamma_N^{\boldsymbol{v}}} \boldsymbol{v}_N + \underbrace{\begin{bmatrix} B_r & 0 & \dots & 0 \\ AB_r & B_r & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}B_r & A^{N-2}B_r & \dots & B_r \end{bmatrix}}_{\Gamma_N^{\boldsymbol{r}}} \boldsymbol{r}_N,\end{aligned} \tag{3-13}$$

The vector of control inputs is defined as

$$\boldsymbol{u}_N = \begin{bmatrix} \boldsymbol{u}(k) \\ \vdots \\ \boldsymbol{u}(k+N-1) \end{bmatrix}, \tag{3-14}$$

the vector of exogenous inputs is defined as

$$\boldsymbol{v}_N = \begin{bmatrix} \boldsymbol{v}(k) \\ \vdots \\ \boldsymbol{v}(k+N-1) \end{bmatrix}, \tag{3-15}$$

$\boldsymbol{y}_N$ is defined as

$$\boldsymbol{y}_N = \begin{bmatrix} \boldsymbol{y}(k) \\ \vdots \\ \boldsymbol{y}(k+N-1) \end{bmatrix}, \tag{3-16}$$

and $\boldsymbol{r}_N$ is defined as

$$\boldsymbol{r}_N = \begin{bmatrix} \boldsymbol{r}(k) \\ \vdots \\ \boldsymbol{r}(k+N-1) \end{bmatrix}, \tag{3-17}$$

in the same manner as $\boldsymbol{u}_N$. The cost function J_N can now be rewritten as follows.

$$J_N = \boldsymbol{x}_N^\top \underbrace{\begin{bmatrix} Q_c & & \\ & \ddots & \\ & & Q_c \end{bmatrix}}_{\mathcal{Q}} \boldsymbol{x}_N, \tag{3-18}$$

Now, using (3-13), the cost function for the QP of the augmented system is given by

$$\begin{aligned} V_N &= \left(\Gamma_{\boldsymbol{u}}\boldsymbol{u} + \Gamma_{\boldsymbol{v}}\boldsymbol{v}_N + \Gamma_{\boldsymbol{x}}\boldsymbol{x}(k) + \Gamma_{\boldsymbol{r}}\boldsymbol{r}_N\right)^\top \mathcal{Q} \left(\Gamma_{\boldsymbol{u}}\boldsymbol{u} + \Gamma_{\boldsymbol{v}}\boldsymbol{v}_N + \Gamma_{\boldsymbol{x}}\boldsymbol{x}(k) + \Gamma_{\boldsymbol{r}}\boldsymbol{r}_N\right) \\ &= \boldsymbol{u}_N^\top H \boldsymbol{u}_N + \boldsymbol{c}(\boldsymbol{x}(k))^\top \boldsymbol{u}_N + V, \end{aligned} \tag{3-19}$$

where the quadratic term matrix for the augmented system H, and the linear term matrix for the augmented system $\boldsymbol{c}(\boldsymbol{x}(k))$ are given by

$$H = \Big(\Gamma_{\boldsymbol{u}}\boldsymbol{u}_N\Big)^\top \mathcal{Q}\Big(\Gamma_{\boldsymbol{u}}\boldsymbol{u}_N\Big), \tag{3-20}$$

$$\boldsymbol{c}(\boldsymbol{x}(k)) = 2\Big(\boldsymbol{v}_N^\top\Gamma_{\boldsymbol{v}} + \boldsymbol{r}_N^\top\Gamma_{\boldsymbol{r}} + \boldsymbol{x}(k)^\top\Gamma_{\boldsymbol{x}}\Big). \tag{3-21}$$

Note that the quadratic cost matrix H used in this problem formulation is not positive definite but rather positive semi-definite due to the positive semi-definite $\mathcal{Q}$ and Q_c.

The constraints on the control inputs and system outputs can also be rewritten as linear inequalities in terms of the free variable $\boldsymbol{u}_N$. The constraints are combined for every timestep:

$$\begin{aligned}
\begin{bmatrix} A_{\boldsymbol{y},\mathrm{u}} & & \\ & \ddots & \\ & & A_{\boldsymbol{y},\mathrm{u}} \end{bmatrix} \boldsymbol{x}_N = \mathcal{A}_{\boldsymbol{y},\mathrm{u}}\boldsymbol{y}_N \quad &\leq \begin{bmatrix} b_{\boldsymbol{y},\mathrm{u}} \\ \vdots \\ b_{\boldsymbol{y},\mathrm{u}} \end{bmatrix} = \beta_{\boldsymbol{y},\mathrm{u}}, \\
\begin{bmatrix} A_{\boldsymbol{u},\mathrm{u}} & & \\ & \ddots & \\ & & A_{\boldsymbol{u},\mathrm{u}} \end{bmatrix} \boldsymbol{u}_N = \mathcal{A}_{\boldsymbol{u},\mathrm{u}}\boldsymbol{u}_N \quad &\leq \begin{bmatrix} b_{\boldsymbol{u},\mathrm{u}} \\ \vdots \\ b_{\boldsymbol{u},\mathrm{u}} \end{bmatrix} = \beta_{\boldsymbol{u},\mathrm{u}}, \\
\begin{bmatrix} A_{\boldsymbol{y},\mathrm{l}} & & \\ & \ddots & \\ & & A_{\boldsymbol{y},\mathrm{l}} \end{bmatrix} \boldsymbol{x}_N = \mathcal{A}_{\boldsymbol{y},\mathrm{l}}\boldsymbol{y}_N \quad &\leq \begin{bmatrix} b_{\boldsymbol{y},\mathrm{l}} \\ \vdots \\ b_{\boldsymbol{y},\mathrm{l}} \end{bmatrix} = \beta_{\boldsymbol{y},\mathrm{l}}, \\
\begin{bmatrix} A_{\boldsymbol{u},\mathrm{l}} & & \\ & \ddots & \\ & & A_{\boldsymbol{u},\mathrm{l}} \end{bmatrix} \boldsymbol{u}_N = \mathcal{A}_{\boldsymbol{u},\mathrm{l}}\boldsymbol{u}_N \quad &\leq \begin{bmatrix} b_{\boldsymbol{u},\mathrm{l}} \\ \vdots \\ b_{\boldsymbol{u},\mathrm{l}} \end{bmatrix} = \beta_{\boldsymbol{u},\mathrm{l}}.
\end{aligned} \tag{3-22}$$

Then, again using (3-13), the constraints are written as

$$\begin{bmatrix} \mathcal{A}_{\boldsymbol{y},\mathrm{u}}\tilde{\Gamma}_{\boldsymbol{u}} \\ \mathcal{A}_{\boldsymbol{y},\mathrm{l}}\tilde{\Gamma}_{\boldsymbol{l}} \\ \mathcal{A}_{\boldsymbol{u},\mathrm{u}} \\ \mathcal{A}_{\boldsymbol{u},\mathrm{l}} \end{bmatrix} \boldsymbol{u}_N \leq \begin{bmatrix} \beta_{\boldsymbol{y},\mathrm{u}} - \mathcal{A}_{\boldsymbol{y},\mathrm{u}}\tilde{\Gamma}_{\boldsymbol{x}}\boldsymbol{x}(k) - \mathcal{A}_{\boldsymbol{y},\mathrm{u}}\tilde{\Gamma}_{\boldsymbol{v}}\boldsymbol{v}_N - \mathcal{A}_{\boldsymbol{y},\mathrm{u}}\tilde{\Gamma}_{\boldsymbol{r}}\boldsymbol{r}_N \\ \beta_{\boldsymbol{y},\mathrm{l}} - \mathcal{A}_{\boldsymbol{y},\mathrm{l}}\tilde{\Gamma}_{\boldsymbol{x}}\boldsymbol{x}(k) - \mathcal{A}_{\boldsymbol{y},\mathrm{l}}\tilde{\Gamma}_{\boldsymbol{v}}\boldsymbol{v}_N - \mathcal{A}_{\boldsymbol{y},\mathrm{l}}\tilde{\Gamma}_{\boldsymbol{r}}\boldsymbol{r}_N \\ \beta_{\boldsymbol{u},\mathrm{u}} \\ \beta_{\boldsymbol{u},\mathrm{l}} \end{bmatrix}, \tag{3-23}$$

where $\mathcal{A}_{\boldsymbol{y},\mathrm{l}}$ and $\mathcal{A}_{\boldsymbol{u},\mathrm{l}}$ accommodate the rewriting of the greater than inequality constraint to a smaller than inequality constraint. $A_{\boldsymbol{y},\mathrm{u}}$, $A_{\boldsymbol{y},\mathrm{l}}$ are the constraint matrices for the upper and lower constraints on the outputs, respectively. $b_{\boldsymbol{y},\mathrm{u}}$ and $b_{\boldsymbol{y},\mathrm{l}}$ are the upper and lower bounds on the outputs. These can be employed to induce bounds on the climate conditions such that the climate cannot be harmful for the tomato crop. The Γ's in (3-23) with tildes denote the same Γ's as in (3-13) only with system matrix C times every entry to obtain the output vector over the entire prediction horizon in QP form.

The constraints on the control inputs are set as follows

$$\underbrace{\begin{bmatrix}20\\20\\0\\0\\0\\0\\0\\0\end{bmatrix}}_{b_{\boldsymbol{u},\mathrm{l}}} \leq \underbrace{\begin{bmatrix}T_{\mathrm{Pipe,1}}\\T_{\mathrm{Pipe,1}}\\\phi_{\mathrm{Lee}}\\\phi_{\mathrm{Wind}}\\C_{\mathrm{Inj}}\\L\\SO_1\\SO_2\end{bmatrix}}_{\boldsymbol{u}} \leq \underbrace{\begin{bmatrix}80\\80\\100\\100\\1\cdot 10^6\\50\\100\\100\end{bmatrix}}_{b_{\boldsymbol{u},\mathrm{u}}}, \tag{3-24}$$

to match the minimal and maximal actuator values. The windward side window opening is kept to maximally half the leeward side window opening, to diminish the different effect of the ventilation rate on the windward side. This is implemented as

$$2\phi_{\mathrm{Lee}}(j^{\mathrm{f}}) \geq \phi_{\mathrm{Wind}}(j^{\mathrm{f}}), \quad \forall j^{\mathrm{f}} = k^{\mathrm{f}}, \ldots, k^{\mathrm{f}} + N_{\mathrm{h}}^{\mathrm{f}} - 1, \tag{3-25}$$

and put into QP form in $\mathcal{A}_\phi$ and b_ϕ.

Summarising, the output-feedback control law in QP form that tracks the reference air temperature is given by

$$\begin{aligned} \boldsymbol{u}_N^* = \quad & \underset{\boldsymbol{u}_N}{\operatorname{argmin}} \quad \boldsymbol{u}_N^T H \boldsymbol{u}_N + \mathbf{f}(\boldsymbol{x}(k))^T \boldsymbol{u}_N, \\ \text{s.t.} \quad & \begin{bmatrix}\mathcal{A}_{\boldsymbol{u},\mathrm{u}}\\\mathcal{A}_{\boldsymbol{u},\mathrm{l}}\\\mathcal{A}_\phi\end{bmatrix} \boldsymbol{u}_N \leq \begin{bmatrix}\beta_{\boldsymbol{u},\mathrm{u}}\\\beta_{\boldsymbol{u},\mathrm{l}}\\\beta_\phi\end{bmatrix}, \end{aligned} \tag{3-26}$$

where the use of the output constraints is omitted, since the tracking achieves the desired reference temperature very well.

In the receding horizon approach, just the first element of $\boldsymbol{u}_N^*$ is implemented and the QP problem is solved again at every timestep $k \in \{1, \ldots, N\}$.

Algorithm 2 Temperature setpoint tracking linear MPC algorithm

Input: System matrices $(A, B_{\boldsymbol{u}}, B_{\boldsymbol{v}}, C, D)$, prediction horizon N, constraint sets $\mathcal{U}$ and $\mathcal{X}$;

1: Obtain initial state estimate $\hat{\boldsymbol{x}}(k)$.
2: Solve (3-26) for the optimal input sequence $\boldsymbol{u}^* = (\boldsymbol{u}_0^*, \cdots, \boldsymbol{u}_{N-1}^*)$.
3: Apply only the first input $\boldsymbol{u}(k) = \boldsymbol{u}_0^*$.
4: Set k to $k+1$.
5: Return to 1.

3-4 QP benchmark results

In this section, the results from the synthesised tracking controller for the benchmark reference temperature trajectory will be shown. The results can be divided into the following parts: the tracking capabilities, the control costs associated with this particular tracking, the

resulting fruit yield and net economic profit. In this chapter, only the tracking capabilities, associated control costs and resulting fruit yield are presented. The net economic profit is only introduced in Chapter 6 in the comparison section that compares the three reference temperature trajectories. First, the tracking capabilities of the designed linear QP MPC controller are presented. Note that the model used for the controller design is the same as the plant prediction model, and full state information is available; there is no model mismatch.

3-4-1 Tracking capabilities

The QP MPC controller can track the benchmark greenhouse indoor air reference temperature over the entire season. To illustrate this, three days are shown in Figure 3-3, one at the beginning, middle, and end of the growing season. Also, the corresponding weather exogenous inputs, calculated optimal control inputs and the tracked greenhouse indoor air temperature are shown, for the first and last day of the growing season.

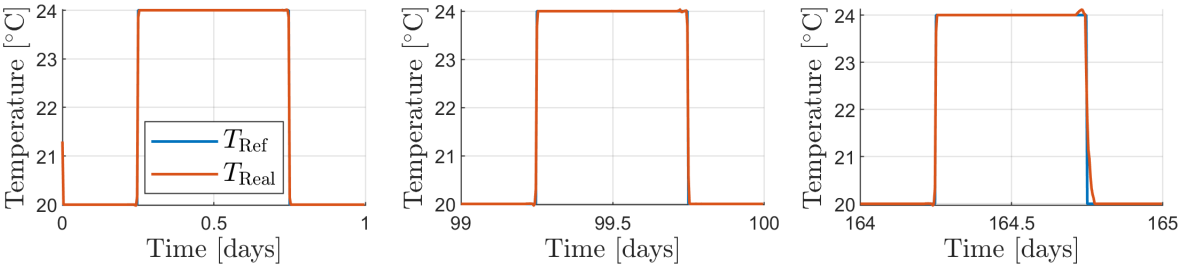


Figure 3-3: Three days in the season with tracked T_{Air}.

The tracking accuracy of the controller is expressed in terms of the Root mean square error (RMSE):

$$\text{RMSE} = \sqrt{\frac{\sum_{k=1}^{N_{\text{i}}^{\text{l}} N_{\text{i}}^{\text{u}}} (\boldsymbol{r}_{1,k} - \boldsymbol{y}_{1,k})^2}{N_{\text{i}}^{\text{l}} N_{\text{i}}^{\text{u}}}}, \tag{3-27}$$

which is RMSE = 0.0871 °C.

The tracking error for three selected days is shown in Figure 3-4 below.

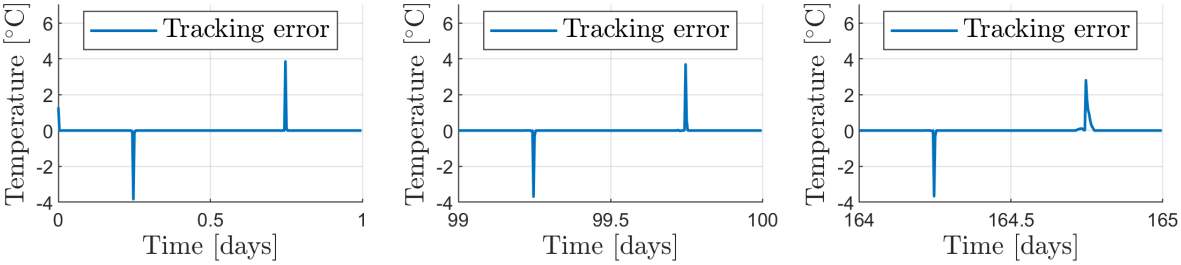
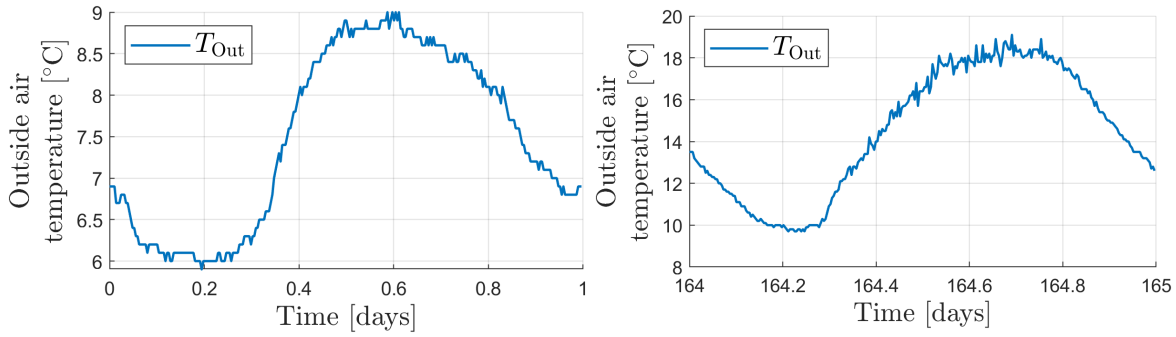


Figure 3-4: Three days in the season with tracking error.

3-4-2 Control costs

The associated calculated optimal controls for the two days are presented in this section. To illustrate the effect of the weather on the generated control inputs, the weather variables for the first and last day of the growing season are shown as well.

The outside air temperature at the beginning and end of the season are presented in Figure 3-5 below.

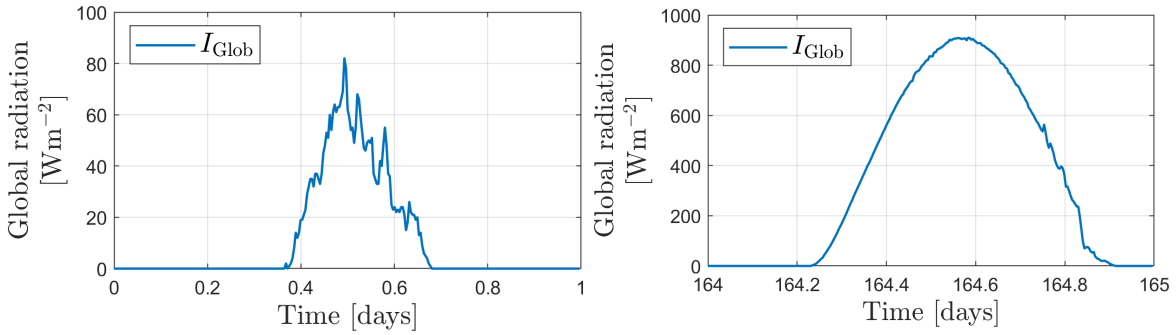


(a) T_{Out} for the 1st day of the season.

(b) T_{Out} for the 165th day of the season.

Figure 3-5: T_{Out} for the first and last day of the season.

The global radiation at the beginning and end of the season are presented in Figure 3-6 below.

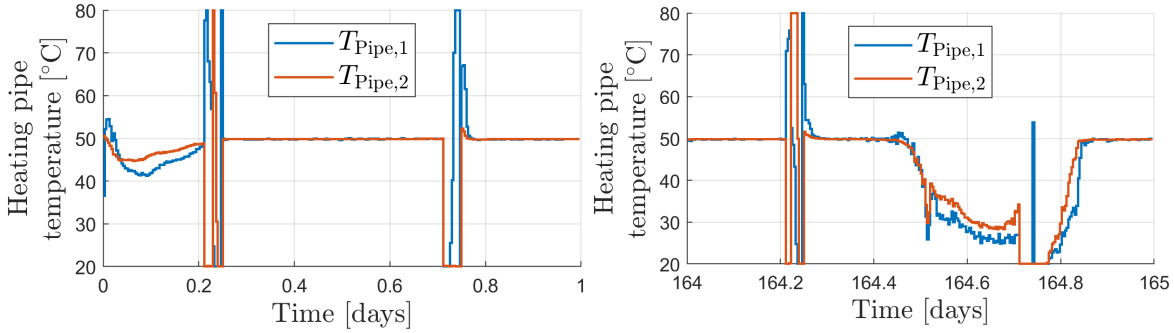


(a) I_{Glob} for the 1st day of the season.

(b) I_{Glob} for the 165th day of the season.

Figure 3-6: I_{Glob} for the first and last day of the season.

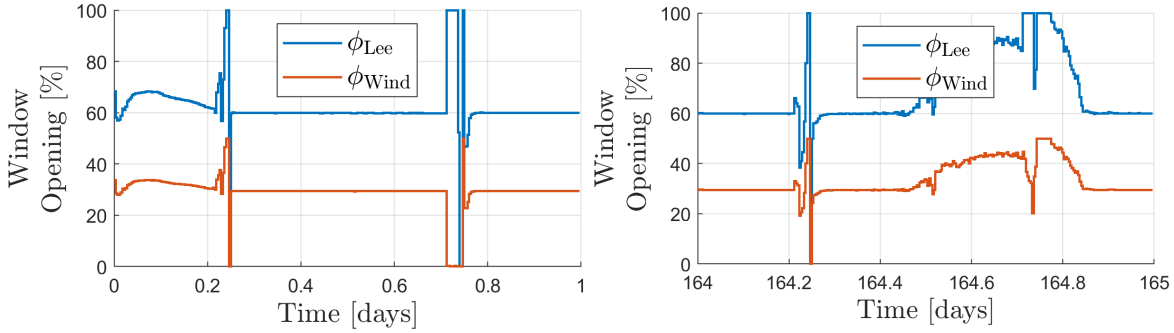
The calculated optimal heating pipe inputs are presented in Figure 3-7 below.



(a) T_{Pipe1} and T_{Pipe2} for the 1st day of the season. **(b)** T_{Pipe1} and T_{Pipe2} for the 165th day of the season.

Figure 3-7: T_{Pipe1} and T_{Pipe2} for the first and last day of the season.

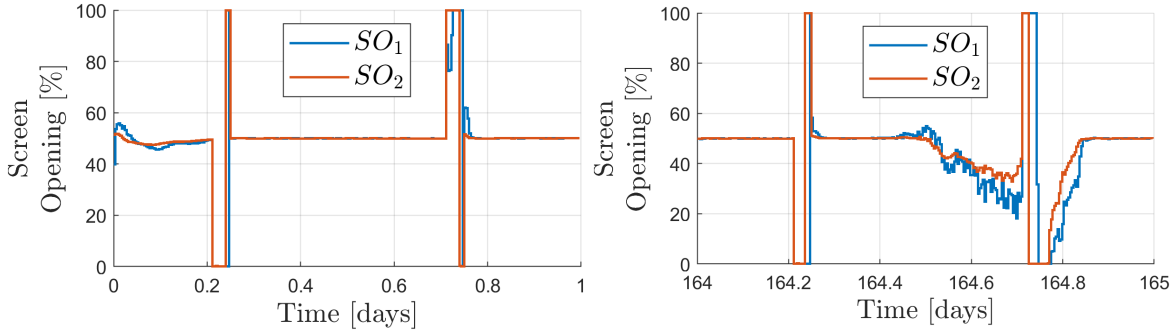
The calculated optimal window openings for the beginning and end of the season are presented in Figure 3-8 below.



(a) ϕ_{Lee} and ϕ_{Wind} for the 1st day of the season. **(b)** ϕ_{Lee} and ϕ_{Wind} for the 165th day of the season.

Figure 3-8: ϕ_{Lee} and ϕ_{Wind}, the leeward and windward side window openings, for the first and last day of the season.

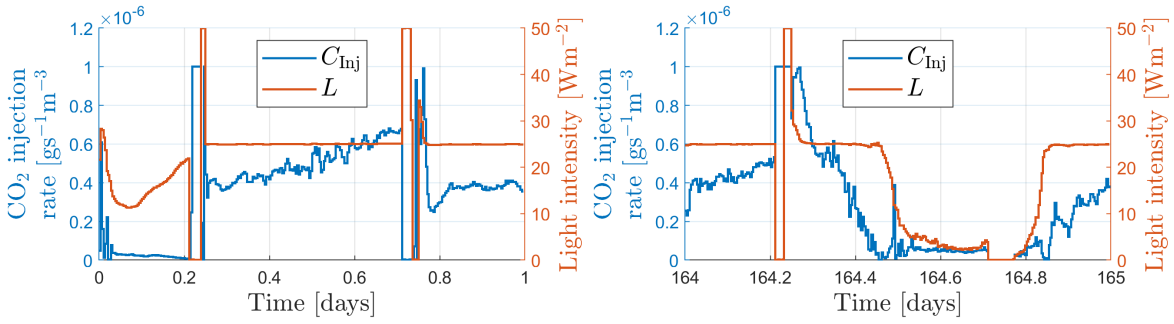
The calculated optimal screen openings are presented in Figure 3-9 below.



(a) SO_1 and SO_2 for the 1st day of the season. **(b)** SO_1 and SO_2 for the 165th day of the season.

Figure 3-9: SO_1 and SO_2 for the first and last day of the season.

The calculated optimal light intensity and CO_2 injection rate are presented in Figure 3-10 below.

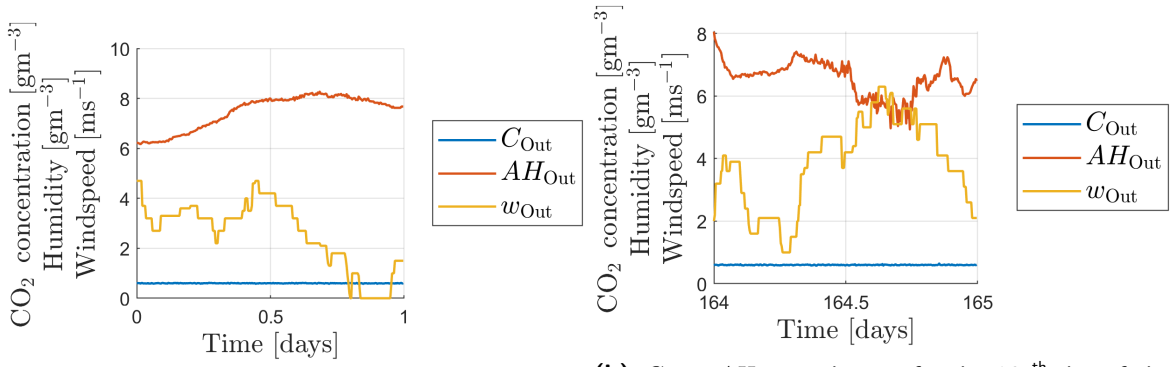


(a) C_{Inj} and L for the 1st day of the season.

(b) C_{Inj} and L for the 165th day of the season.

Figure 3-10: C_{Inj} and L for the first and last day of the season. It can be seen that the lights are used as heating source.

The outside CO_2 concentration, humidity and wind speed at the beginning and end of the season are presented in Figure 3-11 below.



(a) C_{Out}, AH_{Out} and w_{Out} for the 1st day of the season.

(b) C_{Out}, AH_{Out} and w_{Out} for the 165th day of the season.

Figure 3-11: C_{Out}, AH_{Out} and w_{Out} for the first and last day of the season.

Fruit yield

The resulting fruit yield as simulated by the non-linear model as plant is presented in Figure 3-12 below.

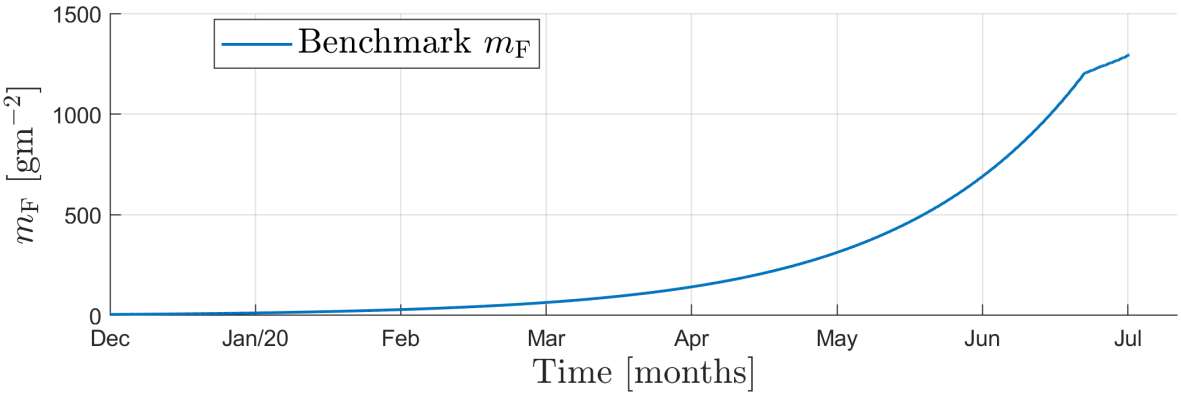


Figure 3-12: Fruit yield resulting from the benchmark reference trajectory.

The amount of kg fresh tomato weight resulting from tracking the predefined reference temperature trajectory is 807 kg when considering the same greenhouse production compartment as the Autonomous greenhouse challenge (AGC).

3-5 Conclusive remarks on tracking QP MPC

In this chapter, the lower layer setpoint tracking controller has been synthesised. Some remarks on the further use are made.

The designed QP MPC temperature setpoint tracking controller will be identical for both upper layer setpoint generating controllers. This is because this controller is merely used to follow the proposed climate strategy and to calculate which controls are needed to achieve the proposed climate strategy.

The controller can achieve the benchmark reference temperature trajectory of 20 degrees at night and 24 degrees by day during the entire growing season under changing weather circumstances. For further use, the upper layer control strategies can infer reference temperatures that are not easily compatible with the weather conditions at that time, e.g., a very high temperature setpoint at the beginning of the season. Thus from these preliminary benchmark results, it cannot be concluded that every climate strategy is achievable by this lower layer setpoint tracking controller.

The controls are now merely calculated to track the proposed temperature setpoints. However, to achieve optimal performance in terms of resource usage, a term could be included that penalises the control input. However, since the focus is to test which upper layer control strategy proposes the optimal climate strategy, this trade-off in the lower layer is omitted.

It can be concluded that the synthesised controller is usable for comparing the climate strategy generating controllers.

Chapter 4

Upper layer setpoint generating Model predictive control (MPC)

Up until now, the greenhouse indoor climate subsystem and crop subsystem were described using different models. The overall crop production control problem was decomposed into a two-layered structure. A linear MPC controller was described and implemented for a benchmark greenhouse indoor air reference temperature. In this chapter, an Economic model predictive control (EMPC) controller will be synthesised that generates reference greenhouse indoor air temperatures that will be tracked by the lower layer controller of the overall decomposed problem. First, the applicable EMPC theory will be described and leveraged for the particular crop production problem in Section 4-1 and Section 4-2. Then, the results in combination with the lower layer setpoint tracking controller will be presented in Section 4-3.

4-1 EMPC theory

For non-linear systems, the traditional quadratic cost that is encountered in tracking MPC can logically be replaced by some economic criterion. In this way, output constraints can be directly integrated in the optimisation objective [37].

EMPC unifies the economic optimisation of a system with a receding horizon policy [38]. The main difference with nominal tracking MPC is the more general cost function of the form

$$J_{\text{eco}} = \|q(\boldsymbol{u})\| - \|\boldsymbol{y}\|, \tag{4-1}$$

where $q(\boldsymbol{u})$ represents the control input costs and $\boldsymbol{y}$ is some expression of the economic revenue related to the process outputs. The cost function can be any cost function with some economic parameters.

The economic goal functions encountered in the optimisation problem for the crop cultivation control problem vary in the exact implementation. Generally, they include a type of expression of the dry matter production, the control input costs, control input rates and sometimes penalty functions on the different control inputs [39].

4-2 Fruit yield maximisation EMPC

In this section, the theory will be used for the crop production problem at hand. First, the used prediction model, which has been introduced briefly in Chapter 2, will be covered again. Then, the used cost metric and constraints, and the implications of these aspects will be presented.

4-2-1 Prediction model

For the upper layer setpoint generating MPC controller, a linear state-space description of the greenhouse system, as introduced in Chapter 2 has been adopted

$$\begin{cases} \boldsymbol{x}^{\mathrm{s}}_{\mathrm{gh}}(k^{\mathrm{s}}+1) = A^{\mathrm{s}}_{\mathrm{gh}}\boldsymbol{x}^{\mathrm{s}}_{\mathrm{gh}}(k^{\mathrm{s}}) + B^{\mathrm{s}}_{\boldsymbol{u}_{\mathrm{gh}}}\boldsymbol{u}^{\mathrm{s}}_{\mathrm{gh}}(k^{\mathrm{s}}) + B^{\mathrm{s}}_{\boldsymbol{v}_{\mathrm{gh}}}\boldsymbol{v}^{\mathrm{s}}_{\mathrm{gh}}(k^{\mathrm{s}}), \\ \boldsymbol{y}^{\mathrm{s}}_{\mathrm{gh}}(k^{\mathrm{s}}) = C^{\mathrm{s}}_{\mathrm{gh}}\boldsymbol{x}^{\mathrm{s}}_{\mathrm{gh}}(k^{\mathrm{s}}), \end{cases} \tag{4-2}$$

here superscript s denotes the upper layer, slow-varying subproblem variables, the subscript gh denotes that the entire greenhouse system is considered, $\boldsymbol{x}^{\mathrm{s}}_{\mathrm{gh}}(k)$ are internal states, $\boldsymbol{u}^{\mathrm{s}}_{\mathrm{gh}}(k)$ are all the eight control inputs to the greenhouse and $\boldsymbol{v}^{\mathrm{s}}_{\mathrm{gh}}(k)$ are the five weather exogenous inputs to the greenhouse and $\boldsymbol{y}^{\mathrm{s}}_{\mathrm{gh}}(k)$ are the five outputs of the greenhouse. Since the states are internal states, they do not have a physical meaning. They are not full-state measurable. Since the state needs to be estimated only once every six hours, once per upper layer iteration, this does not have big implications if the state estimation is not perfect.

As mentioned in Chapter 2, the dimensions of the implemented state-space system description are as follows. $A^{\mathrm{s}}_{\mathrm{gh}} \in \mathbb{R}^{7\times 7}$, $B^{\mathrm{s}}_{\boldsymbol{u}_{\mathrm{gh}}} \in \mathbb{R}^{7\times 8}$, $B^{\mathrm{s}}_{\boldsymbol{v}_{\mathrm{gh}}} \in \mathbb{R}^{7\times 5}$ and $C^{\mathrm{s}}_{\mathrm{gh}} \in \mathbb{R}^{5\times 7}$, and the sampling time for the implemented system description is, as mentioned in Chapter 2, $\tau^s = 6$ [h]. The pole locations of the linear system are plotted in Figure 4-1 below.

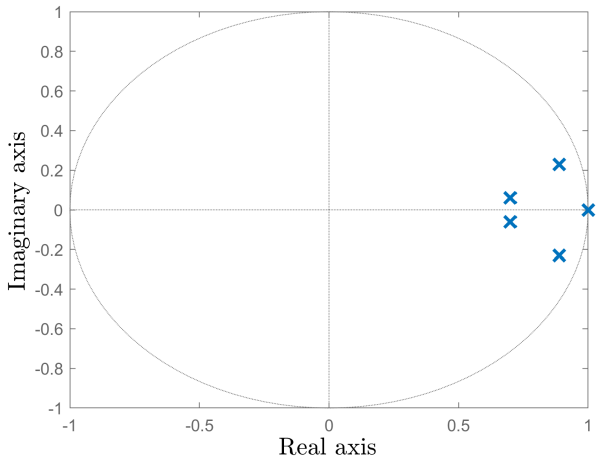


Figure 4-1: Pole locations of the used greenhouse linear model for the upper layer setpoint tracking control.

All poles are within the unit circle, except one that numerically lies on the unit circle. The system would be stable if not for the pole located on the unit circle. The corresponding

natural frequencies and damping ratios per pole are

$$\omega_n = \left|\frac{\ln(z)}{\tau^l}\right| = \begin{bmatrix} 6.9 \cdot 10^{-13} \\ 1.2 \cdot 10^{-5} \\ 1.2 \cdot 10^{-5} \\ 1.7 \cdot 10^{-5} \\ 1.7 \cdot 10^{-5} \end{bmatrix}, \quad \zeta_n = -\cos((\angle z)) = \begin{bmatrix} -1 \\ 0.32 \\ 0.32 \\ 0.97 \\ 0.97 \end{bmatrix}, \tag{4-3}$$

here, the negative damping would imply that the system is unstable. However, this negative damping corresponds to the fruit weight and leaf weight dynamics; it originates from the pole furthest from the origin. The notion of stability does not apply directly to these kind of dynamics. Thus, it will not have any implications on the climate control. The natural frequencies of the linear system used for the upper layer crop production control are smaller when compared to the natural frequencies of the linear system used in the lower layer climate control, which were stated in Chapter 3.

4-2-2 Cost metric for EMPC

The upper layer economic optimisation considers the fruit weight m_{F} and the control costs over the entire growing season. A single harvest is considered at the end of the growing season, giving rise to a shrinking horizon. The general aforementioned economic cost function is then as follows for the particular crop production control problem as described in this thesis

$$J(\boldsymbol{u}^{\mathrm{s}}_{\mathrm{gh}}) = \sum_{j^{\mathrm{s}}=1}^{N^{\mathrm{s}}_{\mathrm{i}}-k^{\mathrm{s}}+1} \Big(\ell_{\mathrm{cost}}(j^{\mathrm{s}})\Big) - m_{\mathrm{F}}(N^{\mathrm{s}}_{\mathrm{i}}), \tag{4-4}$$

In the above equation, $m_{\mathrm{F}}(N^{\mathrm{s}}_{\mathrm{i}})$ is the total fruit dry weight at the end of the season, $N^{\mathrm{s}}_{\mathrm{i}} = 664$ is the number of upper layer iterations to cover the whole growing season from the beginning, and j^{s} is the j^{th} time instance in the upper layer prediction. Furthermore, k^{s} is the current time instance and $\ell_{\mathrm{cost}}(j^{\mathrm{s}})$ is the stage cost associated with the calculated control inputs per upper layer prediction time instance j^{s} and given below.

$$\ell_{\mathrm{cost}}(j^{\mathrm{s}}) = (\boldsymbol{u}^{\mathrm{s}}_{\mathrm{gh}}(j^{\mathrm{s}}))^{\top} R^{\mathrm{s}}_{\boldsymbol{u}_{\mathrm{gh}}} \boldsymbol{u}^{\mathrm{s}}_{\mathrm{gh}}(j^{\mathrm{s}}), \tag{4-5}$$

where $R^{\mathrm{s}}_{\boldsymbol{u}_{\mathrm{gh}}}$ is the control cost matrix for the upper layer subproblem and given by $R_{\boldsymbol{u}_{\mathrm{gh,s}}} = \mathrm{diag}([1, 0, 0, 10^6, 1, 1, 0, 0])$. The economic cost of the actuators that were implemented includes the costs for the electricity of the lighting system and the heating energy consumption of the two heating pipe systems, in accordance with the economic performance metric in [22]. Hence, the economic control costs induced by the usage of the temperature and black-out screens and the opening of the windows is neglected.

This cost function is a non-convex cost function due to the second term representing the economic revenue from the fruit yield.

4-2-3 EMPC constraints for crop production control

The variables occurring in the EMPC optimisation need to be constrained. The controls are constrained in a similar fashion as in Chapter 3, and restated for clarity in (4-6). In the

EMPC scheme, the outputs of the greenhouse indoor climate are constrained as well. This is due to the fact that the optimisation problem is predicting over the entire growing season and might find an optimum with corresponding to-be tracked greenhouse indoor climate variables that are harmful to the crop at the current time. The crop outputs are not constrained, as these are maximised.

The constraints on the greenhouse indoor climate variables are set to

$$\underbrace{\begin{bmatrix}15\\0\\0\\0\\0\end{bmatrix}}_{b^{\text{low}}_{\boldsymbol{y}_{\text{gh,s}}}} \leq \underbrace{\begin{bmatrix}T_{\text{Air}}\\C_{\text{Air}}\\AH_{\text{Air}}\\m_{\text{F}}\\m_{\text{L}}\end{bmatrix}}_{\boldsymbol{y}_{\text{gh,s}}} \leq \underbrace{\begin{bmatrix}30\\50\\10\\1\cdot10^{6}\\1\cdot10^{6}\end{bmatrix}}_{b^{\text{upp}}_{\boldsymbol{y}_{\text{gh,s}}}}, \qquad \underbrace{\begin{bmatrix}20\\20\\0\\0\\0\\0\\0\\0\end{bmatrix}}_{b^{\text{low}}_{\boldsymbol{u}_{\text{gh,s}}}} \leq \underbrace{\begin{bmatrix}T_{\text{Pipe,1}}\\T_{\text{Pipe,1}}\\\phi_{\text{Lee}}\\\phi_{\text{Wind}}\\C_{\text{Inj}}\\L\\SO_1\\SO_2\end{bmatrix}}_{\boldsymbol{u}_{\text{gh,s}}} \leq \underbrace{\begin{bmatrix}80\\80\\100\\100\\1\cdot10^{6}\\50\\100\\100\end{bmatrix}}_{b^{\text{upp}}_{\boldsymbol{u}_{\text{gh,s}}}}, \tag{4-6}$$

here 15 °C and 30 °C for T_{Air} are chosen according to the tomato growing temperatures as described in Chapter 2. The bounds for C_{Air} and AH_{Air} are set to prevent unfavourable conditions for the tomato crop. The upper bounds on the fruit and leaf dry weight are set very high to approximate no bound, and the lower bound for the crop outputs are set to zero since negative fruit or leaf dry weight is not possible in reality. The bounds on the control inputs are determined by the minimum and maximum values the actuators can achieve. The upper and lower bound matrices for the entire prediction horizon are then constructed as follows.

$$\beta^{\text{upp}}_{\boldsymbol{y}_{\text{gh,s}}} = \begin{bmatrix} b^{\text{upp}}_{\boldsymbol{y}_{\text{gh,s}}} \\ \vdots \\ b^{\text{upp}}_{\boldsymbol{y}_{\text{gh,s}}} \end{bmatrix}, \qquad \beta^{\text{low}}_{\boldsymbol{y}_{\text{gh,s}}} = \begin{bmatrix} b^{\text{low}}_{\boldsymbol{y}_{\text{gh,s}}} \\ \vdots \\ b^{\text{low}}_{\boldsymbol{y}_{\text{gh,s}}} \end{bmatrix}, \tag{4-7}$$

$$\beta^{\text{upp}}_{\boldsymbol{u}_{\text{gh,s}}} = \begin{bmatrix} b^{\text{upp}}_{\boldsymbol{u}_{\text{gh,s}}} \\ \vdots \\ b^{\text{upp}}_{\boldsymbol{u}_{\text{gh,s}}} \end{bmatrix}, \qquad \beta^{\text{low}}_{\boldsymbol{u}_{\text{gh,s}}} = \begin{bmatrix} b^{\text{low}}_{\boldsymbol{u}_{\text{gh,s}}} \\ \vdots \\ b^{\text{low}}_{\boldsymbol{u}_{\text{gh,s}}} \end{bmatrix}, \tag{4-8}$$

where $b^{\text{upp}}_{\boldsymbol{u}_{\text{gh,s}}}$ is repeated $N^{\text{s}}_{\text{i}} - k^{\text{s}} + 1$ times. The constraint matrices over the entire growing season are given by

$$\begin{bmatrix} \mathcal{A}^{\text{upp}}_{\boldsymbol{y}_{\text{gh,s}}}\Gamma_{\boldsymbol{u}_{\text{gh,s}}} \\ \mathcal{A}^{\text{low}}_{\boldsymbol{y}_{\text{gh,s}}}\Gamma_{\boldsymbol{u}_{\text{gh,s}}} \\ \mathcal{A}^{\text{upp}}_{\boldsymbol{u}_{\text{gh,s}}} \\ \mathcal{A}^{\text{low}}_{\boldsymbol{u}_{\text{gh,s}}} \end{bmatrix} \boldsymbol{u}_N \leq \begin{bmatrix} \beta^{\text{upp}}_{\boldsymbol{y}_{\text{gh,s}}} - \mathcal{A}^{\text{upp}}_{\boldsymbol{y}_{\text{gh,s}}}\Gamma_{\boldsymbol{x}_{\text{gh,s}}(k^{\text{s}})}\boldsymbol{x}(k) - \mathcal{A}^{\text{upp}}_{\boldsymbol{y}_{\text{gh,s}}}\Gamma_{\boldsymbol{v}_{\text{gh,s}}} \\ \beta^{\text{low}}_{\boldsymbol{y}_{\text{gh,s}}} - \mathcal{A}^{\text{low}}_{\boldsymbol{y}_{\text{gh,s}}}\Gamma_{\boldsymbol{x}_{\text{gh,s}}(k^{\text{s}})}\boldsymbol{x}(k) - \mathcal{A}^{\text{low}}_{\boldsymbol{y}_{\text{gh,s}}}\Gamma_{\boldsymbol{v}_{\text{gh,s}}} \\ \beta^{\text{upp}}_{\boldsymbol{u}_{\text{gh,s}}} \\ \beta^{\text{low}}_{\boldsymbol{u}_{\text{gh,s}}} \end{bmatrix}, \tag{4-9}$$

where $\mathcal{A}^{\text{upp}}_{\boldsymbol{y}_{\text{gh,s}}}$ and $\mathcal{A}^{\text{low}}_{\boldsymbol{y}_{\text{gh,s}}}$ accommodate the rewriting of the greater than inequality constraint to a smaller than inequality constraint. The constraint matrices are structured in the same manner as described in Chapter 3, only with $\mathcal{A}^{\text{upp}}_{\boldsymbol{y}_{\text{gh,s}}}$, $\mathcal{A}^{\text{low}}_{\boldsymbol{y}_{\text{gh,s}}}$, $\mathcal{A}^{\text{upp}}_{\boldsymbol{u}_{\text{gh,s}}}$, $\mathcal{A}^{\text{low}}_{\boldsymbol{u}_{\text{gh,s}}}$, $\Gamma_{\boldsymbol{u}_{\text{gh,s}}}$, $\Gamma_{\boldsymbol{v}_{\text{gh,s}}}$, $\Gamma_{\boldsymbol{x}_{\text{gh,s}}(k^{\text{s}})}$ and $\Gamma_{\boldsymbol{r}_{\text{gh,s}}}$ with the system matrices of the used upper layer greenhouse system prediction model. The difference is only that the respective matrices shrink along with the shrinking horizon implementation.

4-2-4 EMPC optimisation problem and algorithm for crop production control

The optimisation problem that is solved in one upper layer iteration k^{s} is given by

$$
\begin{aligned}
\min_{\boldsymbol{u}^{\mathrm{s}}} \quad & J_{\mathrm{N}}(\boldsymbol{x}_{N_{\mathrm{h}}^{\mathrm{s}}}, \boldsymbol{u}_{\mathrm{gh}}, \boldsymbol{v}_{\mathrm{gh}}^{\mathrm{s}}) = \sum_{j^{\mathrm{s}}=1}^{N_{\mathrm{i}}^{\mathrm{s}}-k^{\mathrm{s}}+1} \Big(\ell_{\mathrm{cost}}(j^{\mathrm{s}}) \Big) - m_{\mathrm{F}}(N_{\mathrm{i}}^{\mathrm{s}}), \\
\text{s.t.} \quad & \boldsymbol{x}_{\mathrm{gh}}(j^{\mathrm{s}}+1) = A_{\mathrm{gh}}\boldsymbol{x}_{\mathrm{gh}}(j^{\mathrm{s}}) + B_{\boldsymbol{u}_{\mathrm{gh}}}\boldsymbol{u}_{\mathrm{gh}}(j^{s}) + B_{\boldsymbol{v}_{\mathrm{gh}}}\boldsymbol{v}_{\mathrm{gh}}(j^{\mathrm{s}}), \quad \forall j^{\mathrm{s}} \in \{k^{\mathrm{s}}, \ldots, k^{\mathrm{s}} + N_{\mathrm{h}}^{\mathrm{s}} - 1\}, \\
& \boldsymbol{y}_{\mathrm{gh}}(j^{\mathrm{s}}) = C_{\mathrm{gh}}\boldsymbol{x}_{\mathrm{gh}}(j^{s}), \quad \forall j^{\mathrm{s}} \in \{k^{\mathrm{s}}, \ldots, k^{\mathrm{s}} + N_{\mathrm{h}}^{\mathrm{s}} - 1\}, \\
& \boldsymbol{x}_{\mathrm{gh}}^{\mathrm{s}}(0) = \hat{\boldsymbol{x}}_{\mathrm{gh}}(k^{\mathrm{s}}), \\
& \boldsymbol{u}_{\mathrm{gh}}(j^{\mathrm{s}}) \in \mathcal{U}_{\mathrm{gh}}^{\mathrm{s}}, \quad \forall j^{\mathrm{s}} \in \{k^{\mathrm{s}}, \ldots, k^{\mathrm{s}} + N_{\mathrm{h}}^{\mathrm{s}} - 1\}, \\
& \boldsymbol{x}_{\mathrm{gh}}(j^{\mathrm{s}}) \in \mathcal{X}_{\mathrm{gh}}^{\mathrm{s}}, \quad \forall j^{\mathrm{s}} \in \{k^{\mathrm{s}}, \ldots, k^{\mathrm{s}} + N_{\mathrm{h}}^{\mathrm{s}} - 1\},
\end{aligned}
\tag{4-10}
$$

here $J_{\mathrm{N}}(\boldsymbol{x}_{N_{\mathrm{h}}^{\mathrm{s}}}, \boldsymbol{u}_{\mathrm{gh}}, \boldsymbol{v}_{\mathrm{gh}}^{\mathrm{s}})$ is the to-be minimised cost function. The first term, $\ell(\boldsymbol{x}_j, \boldsymbol{u}_j, \boldsymbol{r}_j)$ is the stage cost which is computed at each prediction time index j and $V_{\mathrm{f}}(\boldsymbol{x}_N, \boldsymbol{r}_N)$ is the terminal cost at the end of the prediction horizon, $j = N_{\mathrm{h}}^{\mathrm{s}}$.

The EMPC algorithm for the crop production problem is then given in Algorithm (3) below [35].

Algorithm 3 EMPC for crop production control algorithm

Input: System matrices $(A_{\mathrm{gh}}^{\mathrm{s}}, B_{\boldsymbol{u}_{\mathrm{gh}}}^{\mathrm{s}}, B_{\boldsymbol{v}_{\mathrm{gh}}}^{\mathrm{s}}, C_{\mathrm{gh}}^{\mathrm{s}}, D_{\mathrm{gh}}^{\mathrm{s}})$, prediction horizon $N_{\mathrm{h}}^{\mathrm{s}}$, constraint sets $\mathcal{U}_{\mathrm{gh}}^{\mathrm{s}}$ and $\mathcal{X}_{\mathrm{gh}}^{\mathrm{s}}$, weather prediction $\boldsymbol{v}_{\mathrm{gh}}^{\mathrm{s}}$;

1: Obtain initial state estimate $\hat{\boldsymbol{x}}_{\mathrm{gh}}(k^{\mathrm{s}})$.
2: Solve (4-10) for the optimal input sequence $(\boldsymbol{u}_{\mathrm{gh}}^{\mathrm{s}})^* = ((\boldsymbol{u}_0^{\mathrm{s}})^*, \cdots, (\boldsymbol{u}_{N_{\mathrm{h}}^{\mathrm{s}}}^{\mathrm{s}})^*)$ and corresponding optimal greenhouse indoor climate outputs.
3: Apply only the first input $\boldsymbol{u}_{\mathrm{gh}}(k^{\mathrm{s}}) = (\boldsymbol{u}_{\mathrm{gh},0}^*)$.
4: Set k^{s} to $k^{\mathrm{s}} + 1$.
5: Return to 1.

When the crop production is controlled in this one-layered approach, the fast dynamics of the greenhouse indoor climate are not taken into consideration. To ensure the indoor climate variables are taken care of, the climate control extra layer is needed that operates at sampling time τ^{f}.

4-2-5 EMPC in the two-layer crop production control problem

Following the one-layer optimisation, the two-layer algorithm for the crop production control problem with setpoint generating upper layer EMPC controller is given in Algorithm (4). In contrast to the one-layer approach, the two-layer optimisation can deal with the slow and fast dynamics of the greenhouse system and employ a prediction horizon that spans the whole growing season.

Algorithm 4 EMPC two-layer crop production control problem algorithm

Input: System matrices $(A^{\mathrm{s}}_{\mathrm{gh}}, B^{\mathrm{s}}_{\boldsymbol{u}_{\mathrm{gh}}}, B^{\mathrm{s}}_{\boldsymbol{v}_{\mathrm{gh}}}, C^{\mathrm{s}}_{\mathrm{gh}}, D^{\mathrm{s}}_{\mathrm{gh}})$, upper layer prediction horizon $N^{\mathrm{s}}_{\mathrm{h}}$, constraint sets $\mathcal{U}^{\mathrm{s}}_{\mathrm{gh}}$, $\mathcal{Y}^{\mathrm{s}}_{\mathrm{gh}}$ and $\mathcal{X}^{\mathrm{s}}_{\mathrm{gh}}$, weather prediction $\boldsymbol{v}^{\mathrm{s}}_{\mathrm{gh}}$, lower layer prediction horizon $N^{\mathrm{f}}_{\mathrm{h}}$ constraint sets $\mathcal{U}^{\mathrm{f}}_{\mathrm{gh}}$ and $\mathcal{X}^{\mathrm{f}}_{\mathrm{gh}}$;

1: Obtain initial state estimate $\hat{\boldsymbol{x}}_{\mathrm{gh}}(k^{\mathrm{s}})$.
2: Solve (4-10) for the optimal input sequence $(\boldsymbol{u}^{\mathrm{s}}_{\mathrm{gh}})^* = \left((\boldsymbol{u}^{\mathrm{s}}_0)^*, \cdots, (\boldsymbol{u}^{\mathrm{s}}_{N^{\mathrm{s}}_{\mathrm{h}}})^*\right)$ and corresponding optimal greenhouse indoor air reference temperature trajectory $T^*_{\mathrm{Air}} = (T^*_{\mathrm{Air},0}, \cdots, T^*_{\mathrm{Air},N-1})$
3: Sample first reference $T^*_{\mathrm{Air},0}$ to lower layer sampling rate.
4: **for** $k^{\mathrm{f}} \in \{1, \ldots, N^{\mathrm{f}}_{\mathrm{i}}\}$
   a: Obtain initial state estimate $\hat{\boldsymbol{x}}_{\mathrm{gh}}(k^{\mathrm{f}}_{\mathrm{gh}})$.
   b: Solve (3-2) for the optimal input sequence $(\boldsymbol{u}^{\mathrm{f}}_{\mathrm{gh}})^* = \left((\boldsymbol{u}^{\mathrm{f}}_0)^*, \cdots, (\boldsymbol{u}^{\mathrm{f}}_{N^{\mathrm{f}}_{\mathrm{h}}})^*\right)$.
   c: Apply only the first input $\boldsymbol{u}(k^{\mathrm{f}}) = (\boldsymbol{u}^{\mathrm{f}}_{\mathrm{gh},0})^*$.
   d: Set k^{f} to $k^{\mathrm{f}} + 1$.
   e: Return to 4.
5: **end for**
6: Set k^{s} to $k^{\mathrm{s}} + 1$ and update initial state estimate $\hat{\boldsymbol{x}}_{\mathrm{gh}}(k^{\mathrm{s}})$.
7: Return to 1.

4-3 EMPC fruit maximisation results

The two-layer approach for crop production control is implemented, of which the results are presented in this section. The setpoint generating EMPC algorithm results are divided into the climate strategy, resource usage, fruit yield and net economic profit.

4-3-1 EMPC climate strategy results

In Figure 5-4 below, the generated reference temperatures over the entire season are shown.

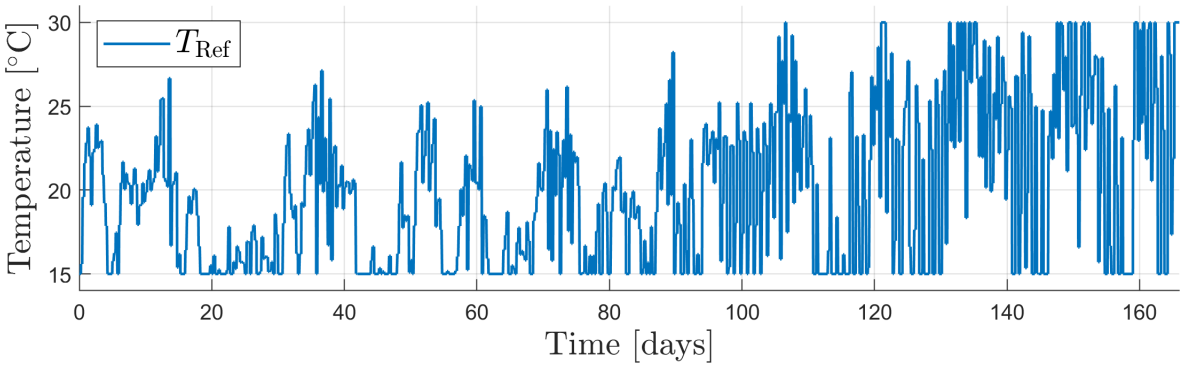


Figure 4-2: Reference temperatures generated by EMPC controller over the entire growing season.

Here, the output constraint set $\mathcal{Y}^{\mathrm{s}}_{\mathrm{gh}}$ implements the upper and lower bound on the generated reference temperature by $T_{\mathrm{Air}} \geq 15°\mathrm{C}$ and $T_{\mathrm{Air}} \leq 30°\mathrm{C}$, which the generated temperature

setpoints satisfy. The weekly averages for the generated EMPC greenhouse indoor air temperature setpoints and corresponding average other greenhouse indoor variables are shown in Figures 4-3 to 4-5 below.

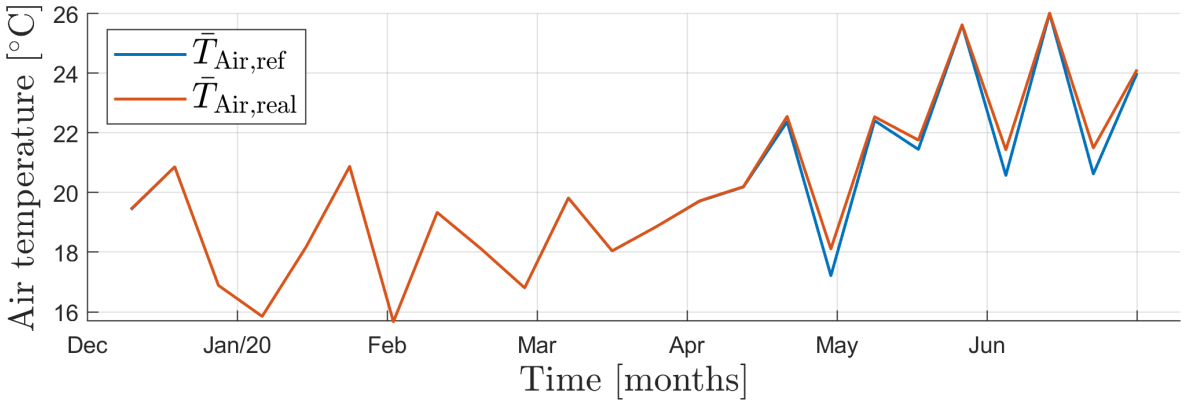


Figure 4-3: Weekly average of $T_{\text{Air,ref}}$ and realised T_{Air}, resulting from the EMPC climate strategy.

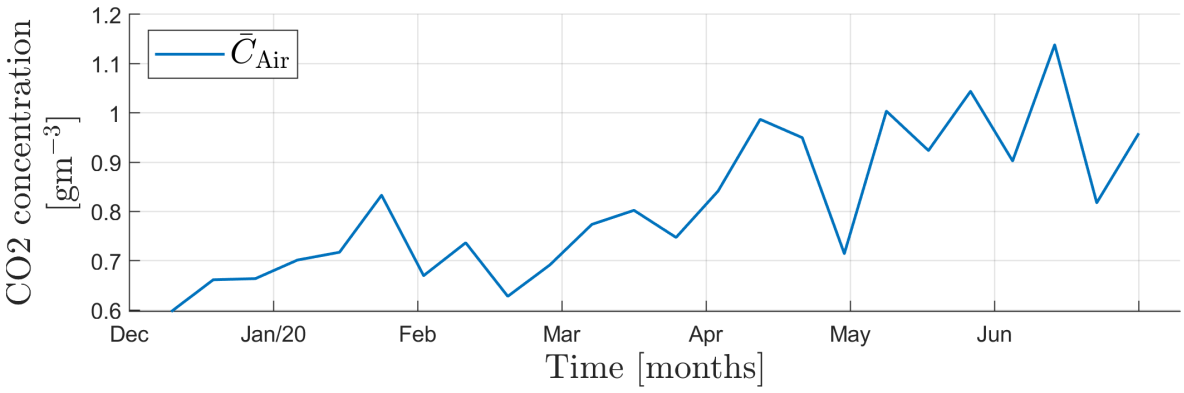


Figure 4-4: Weekly average of C_{Air} that accompanies the tracked $T_{\text{Air,ref}}$, resulting from the EMPC climate strategy.

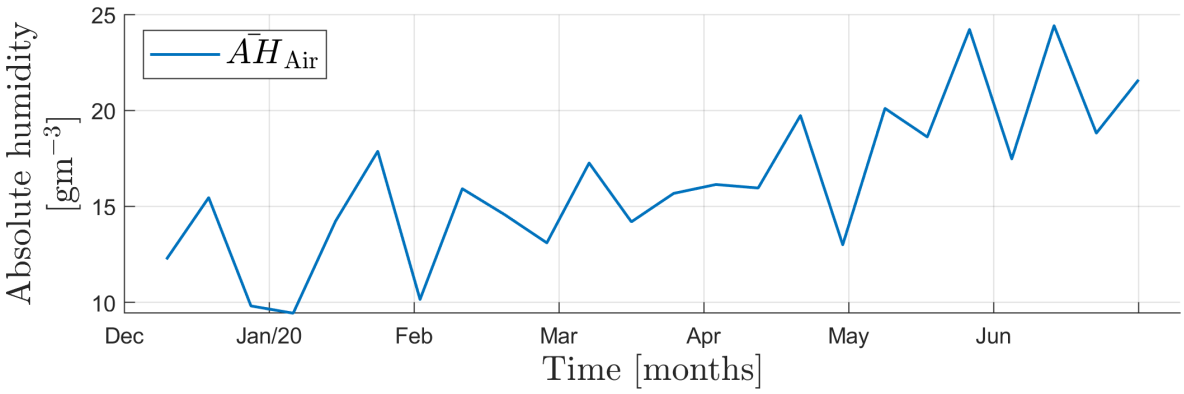


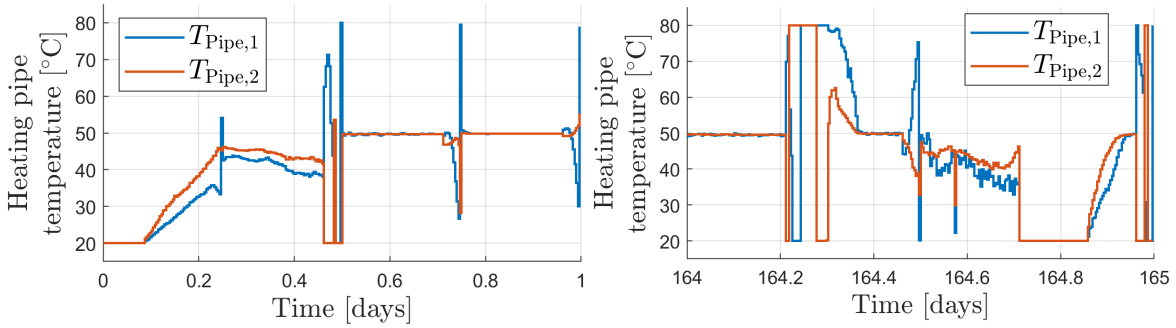
Figure 4-5: Weekly average of AH_{Air} that accompanies the tracked $T_{\text{Air,ref}}$, resulting from the EMPC climate strategy.

It can be noted that the tracking performs worse at the end of the growing season. This can be explained due to the larger difference between the nocturnal and diurnal setpoints. The cooling cannot be achieved in the available time with these fluctuations.

4-3-2 EMPC resource usage

The weather conditions, crop conditions and climate strategy vary throughout the growing season and the calculated control inputs to follow the EMPC climate strategy vary accordingly. For illustrative purposes, the optimal controls calculated by the lower layer setpoint tracking controller are depicted for the first and last day of the growing season in Figures 4-6 to 4-9 below.

The calculated optimal heating pipe inputs are presented in Figure 4-6 below.

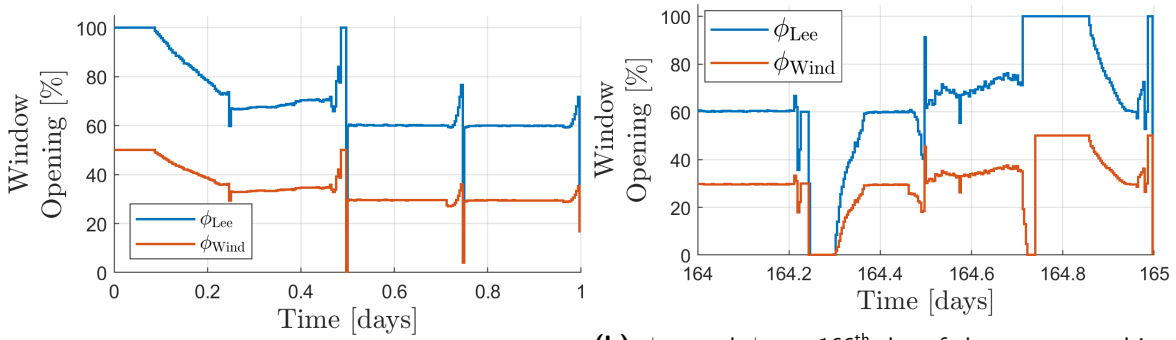


(a) T_{Pipe1} and T_{Pipe2} 1st day of the season.

(b) T_{Pipe1} and T_{Pipe2} 166th day of the season.

Figure 4-6: T_{Pipe1} and T_{Pipe2} first and last day of the season, resulting from the EMPC climate strategy.

The calculated optimal leeward side and windward side window openings, ϕ_{Lee} and ϕ_{Wind}, for the beginning and end of the season are presented in Figure 4-7 below.

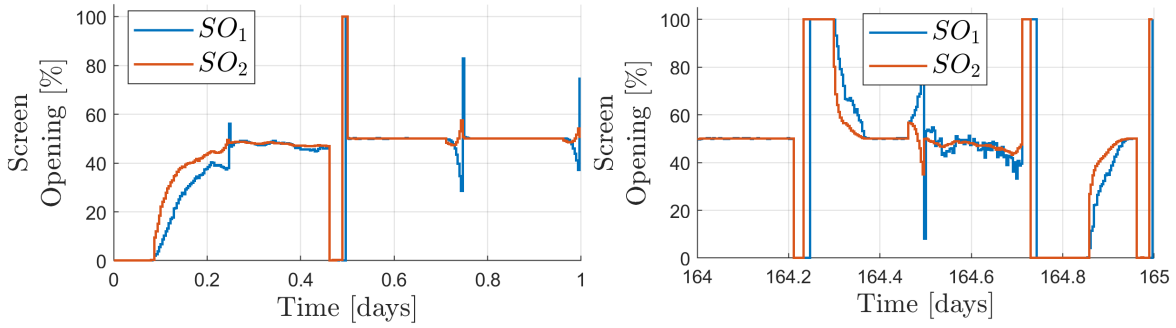


(a) ϕ_{Lee} and ϕ_{Wind} 1st day of the season.

(b) ϕ_{Lee} and ϕ_{Wind} 166th day of the season, resulting from the EMPC climate strategy.

Figure 4-7: ϕ_{Lee} and ϕ_{Wind}, the leeward and windward side window openings, for the first and last day of the season, resulting from the EMPC climate strategy.

The calculated optimal screen openings are presented in Figure 5-10 below.

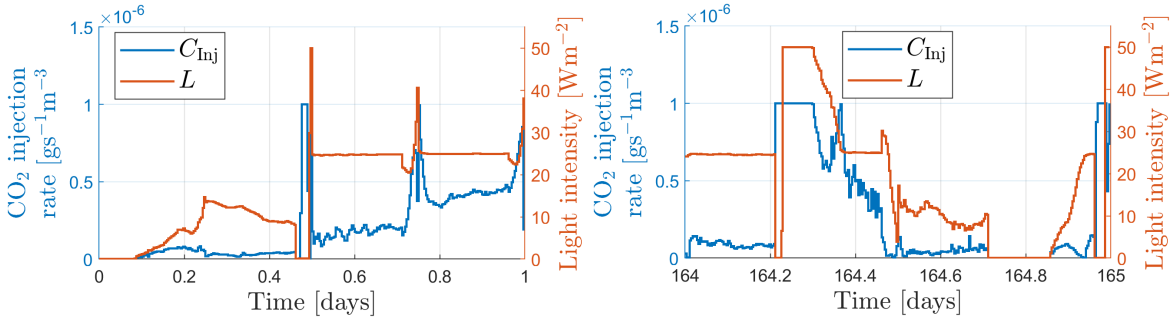


(a) SO_1 and SO_2 1st day of the season.

(b) SO_1 and SO_2 166th day of the season.

Figure 4-8: SO_1 and SO_2 first and last day of the season, resulting from the EMPC climate strategy.

The calculated optimal light intensity and CO_2 injection rate are presented in Figure 4-9 below.



(a) C_{Inj} and L 1st day of the season.

(b) C_{Inj} and L 166th day of the season.

Figure 4-9: C_{Inj} and L first and last day of the season, resulting from the EMPC climate strategy.

4-3-3 EMPC fruit yield results

The fruit yield resulting from the EMPC climate strategy on the ground truth crop simulator is presented in Figure 4-10 below.

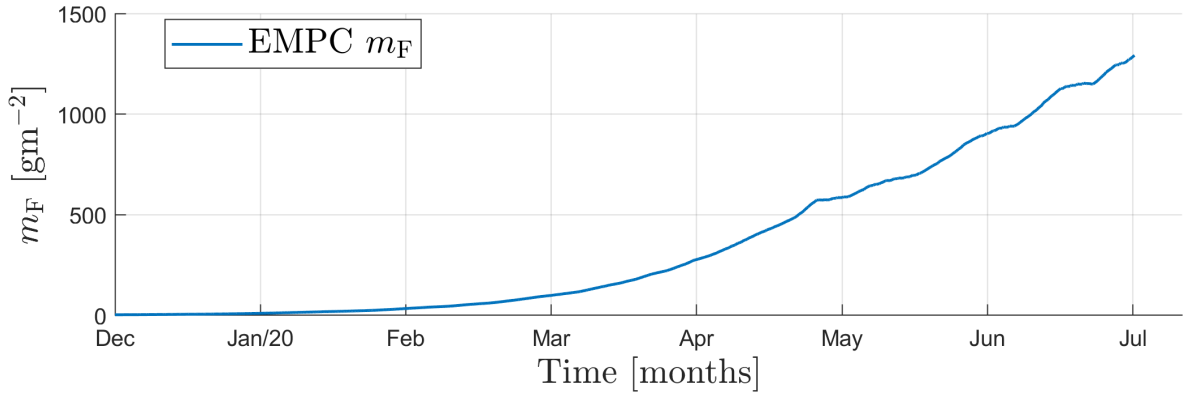


Figure 4-10: Fruit yield, resulting from the EMPC climate strategy.

The resulting amount of fresh tomatoes from the EMPC strategy is 815 kg for the particular greenhouse department considered in this thesis and the Autonomous greenhouse challenge (AGC). This is more than the crop yield resulting from the benchmark reference temperature strategy. The benchmark reference temperature strategy yielded 806 kg of fresh tomatoes.

4-4 Conclusive remarks on fruit maximisation EMPC

In this chapter, the upper layer EMPC controller is presented that generates reference temperatures that are sent to the climate tracking controller. Some conclusive remarks on the interpretation of this controller will be given before this controller will be compared to the novel Data-enabled predictive control (DeePC) setpoint generating controller in the next chapter.

The model-based two-layered control scheme that uses an EMPC setpoint generating controller and MPC setpoint tracking controller is capable of controlling the crop production. The effectiveness depends on the accuracy of the model description, which was covered in Chapter 2.

The temperature strategy resulting from the EMPC controller tends to a higher reference temperature nearing the end of the growing season. This is logical when considering the weather conditions, which imply hotter outside air temperature and higher global radiation at the end of the growing season.

The controller generated reference temperatures fluctuate more at the end of the season when compared to the reference temperatures at the beginning of the season. This can indicate harsher control strategies needed to achieve these setpoints.

Chapter 5

Upper layer setpoint generating Data-enabled predictive control (DeePC)

So far, the overall system has been introduced according to a non-linear crop model and linear climate models. The overall problem has been decomposed into two layers, of which the lower layer employs a setpoint tracking Model predictive control (MPC) controller which has been described. The first controller that generates the reference air temperatures has been discussed, which is an Economic model predictive control (EMPC) controller. In this chapter, the second setpoint generating controller will be synthesised, which is a leveraged DeePC controller. First, the DeePC preliminary theory will be introduced in Section 5-1. Hereafter, it will be leveraged for a new type of DeePC controller in Section 5-2. Last, the preliminary results will be discussed in Section 5-3.

5-1 DeePC theory

The novel DeePC algorithm, as presented in [40], uses a parametric-free system description to compute optimal controls in a receding horizon fashion, while using real-time output feedback. It replaces the system identification, state estimation and future trajectory prediction with one single optimisation problem. The algorithm allows the implementation of input and output constraints, useful for safety criteria. The formal equivalence with MPC for deterministic Linear time-invariant (LTI) systems has been proven. This section first introduces the systems' behavioural theory. Then, the DeePC algorithm is presented.

5-1-1 Systems behavioural theory

Behavioural system theory is employed in DeePC for the used non-parametric system representation. Behavioural system theory considers the subspace of the signal space in which the

trajectories of the dynamical system exist. In contrast with classical system theory, where particular system representations are employed, this theory describes dynamical systems in the more general terms of their behaviour. A dynamical system is captured by the following definitions [41], taken directly from [40]:

Definition 5.1 (Dynamical system)**.** A *dynamical system* is a 3-tuple $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathscr{B})$ where $\mathbb{Z}_{\geq 0}$ is the discrete-time axis, $\mathbb{W}$ is a signal space, and $\mathscr{B} \subseteq \mathbb{W}^{\mathbb{Z}_{\geq 0}}$ is the behaviour.

Definition 5.2 (Linear dynamical system)**.** $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathscr{B})$ is *linear* if $\mathbb{W}$ is a vector space and $\mathscr{B}$ is a linear subspace of $\mathbb{W}^{\mathbb{Z}_{\geq 0}}$.

Definition 5.3 (Time invariant dynamical system)**.** $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathscr{B})$ is *time invariant* if $\mathscr{B} \subseteq \delta\mathscr{B}$ where $\delta : \mathbb{W}^{\mathbb{Z}_{\geq 0}}$ is the forward time shift defined by $(\delta w)(t) = w(t+1)$ and $\delta\mathscr{B} = \{\delta w | w \in \mathscr{B}\}$.

Definition 5.4 (Complete dynamical system)**.** $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathscr{B})$ is *complete* if $\mathscr{B}$ is closed in the topology of pointwise convergence.

$\mathcal{L}^{m+p}$ are the class of systems $(\mathbb{Z}_{\geq 0}, \mathbb{R}^{m+p}, \mathscr{B})$ that satisfy the definitions (5.2)-(5.4), where $m, p \in \mathbb{Z}_{\geq 0}$. Furthermore, dynamical system $\mathbb{W}$ of $(\mathbb{Z}_{\geq 0}, \mathbb{R}^{m+p}, \mathscr{B})$ is finite-dimensional if it satisfies (5.2)-(5.3) [42]. Proceeding, only the behaviour $\mathscr{B}$ is used to denote a dynamical system $\mathcal{L}^{m+p}$.

The two so-called identifiability conditions are formulated that are needed and sufficient for LTI systems to describe the dynamical system from data [42]:

Definition 5.5 (Controllable)**.** A system $\mathscr{B} \in \mathcal{L}^{m+p}$ is *controllable* if for every $T \in \mathbb{Z}_{\geq 0}, w^1 \in \mathscr{B}_T, w^2 \in \mathscr{B}$ there exists signal $w \in \mathscr{B}$ and $T' \in \mathbb{Z}_{\geq 0}$ such that $w_t = w_t^1$ for $1 \leq t \leq T$ and $w_t = w_{t-T-T'}^2$ for $t > T + T'$.

Definition 5.6 (Persistently exciting of order L)**.** Let $L, T \in \mathbb{Z}_{\geq 0}$ such that $T \geq L$. The signal $u = \text{col}(u_1, \ldots, u_T) \in \mathbb{R}^{Tm}$ is *persistently exciting of order L* if the Hankel matrix

$$\mathscr{H}_L(u) := \begin{bmatrix} u_1 & \cdots & u_{T-L+1} \\ \vdots & \ddots & \vdots \\ u_L & \cdots & u_T \end{bmatrix} \tag{5-1}$$

is of full row rank.

Definition (5.5) implies that any two trajectories of a controllable behavioural system can be patched together in finite time. Definition (5.6) implies that the full dynamics of the systems' behaviour is captured by an output sequence that results from exciting the system with a persistently exciting input signal.

Equivalence of representations

The dynamical system' behavioural representation $\mathscr{B}$ is equivalent to, amongst other parametric system representations, the Discrete-time (DT) state-space representation denoted by $\mathscr{B}(A, B, C, D) = \{\text{col}(u, y) \in (\mathbb{R}^{m+p})^{\mathbb{Z}_{\geq 0}} \mid \exists x \in (\mathbb{R}^n)^{\mathbb{Z}_{\geq 0}} \text{s.t.} \quad \sigma x = Ax + Bu, y = Cx + Du\}$.

The following notions are needed to collect the data for the parametric-free system representation of dynamical system $\mathscr{B}$. The state-space *minimal representation* is the input/output/state (I/S/O) representation of smallest order $\boldsymbol{n}(\mathscr{B})$. The *lag* $\ell(\mathscr{B})$ of a dynamical system is defined as the smallest integer number $\ell \in \mathbb{Z}_{\geq 0}$ such that $\mathscr{O}_\ell(A,C) := \mathrm{col}(C, CA, \ldots, CA^{\ell-1})$, the observability matrix, has rank $\boldsymbol{n}(\mathscr{B})$. The number of data points T must be at least $(n_{\boldsymbol{u}}+1)(t+\boldsymbol{n}(\mathscr{B})-1)$ to satisfy the condition of persistence of excitation. Last, the lower triangular Toeplitz matrix $\mathscr{T}_N(A,B,C,D)$ consisting of (A,B,C,D) is defined by

$$\mathscr{T}_N(A,B,C,D) := \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ CA^{N-2}B & \cdots & CB & D \end{bmatrix}. \tag{5-2}$$

Then, the equivalence of representations relies on the following two lemmas [41]:

Lemma 5.1 (Unique system state). Let $\mathscr{B} \in \mathcal{L}^{m+p}$ and $\mathscr{B}(A,B,C,D)$ a minimal I/S/O representation. Let $T_{\mathrm{p}}, T_{\mathrm{f}} \in \mathbb{Z}_{>0}$ with $T_{\mathrm{p}} \geq \boldsymbol{\ell}(\mathscr{B})$ and $\mathrm{col}(u_{\mathrm{ini}}, u_{\mathrm{fut}}, y_{\mathrm{ini}}, y_{\mathrm{fut}}) \in \mathscr{B}_{T_{\mathrm{p}}+T_{\mathrm{f}}}$. Then there exists a unique $x_{\mathrm{ini}} \in \mathbb{R}^{\boldsymbol{n}(\mathscr{B})}$ such that [43]

$$y = \mathscr{O}_N(A,C)x_{\mathrm{ini}} + \mathscr{T}_N(A,B,C,D)u_{\mathrm{fut}} \tag{5-3}$$

Lemma 5.2 (Fundamental Lemma). Consider a controllable system $\mathscr{B} \in \mathcal{L}^{m+p}$. Let $T, t \in \mathbb{Z}_{>0}$, and $w = \mathrm{col}(u,y) \in \mathscr{B}_T$. Assume u to be persistently exciting of order $t + \boldsymbol{n}(\mathscr{B})$. Then $\mathrm{colspan}(\mathscr{H}_t(w)) = \mathscr{B}_t$.

Lemma (5.1) implies that, given a sufficiently long window of initial system data $\mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}})$, the state to which the input sequence u_{ini} drives the system is unique. Also, if the system matrices (A,B,C,D) and the input trajectory u and corresponding output trajectory y are known, the initial system state trajectory x_{ini} can be computed. The *Fundamental Lemma* (5.2) implies that a persistently exciting input sequence generates a finite length output sequence that can construct any output trajectory of a controllable LTI system.

5-1-2 DeePC algorithm

Data collection

DeePC uses a parametric-free system representation which consists of offline measured sequence of input and output data $\boldsymbol{u}^{\mathrm{d}} = \mathrm{col}(\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{\mathrm{T}}) \in \mathbb{R}^{Tn_{\boldsymbol{u}}}$ and $\boldsymbol{y}^{\mathrm{d}} = \mathrm{col}(\boldsymbol{y}_1, \cdots, \boldsymbol{y}_{\mathrm{T}}) \in \mathbb{R}^{Tn_{\boldsymbol{y}}}$ with the assumption that $\boldsymbol{u}^{\mathrm{d}}$ is persistently exciting. The data is put into Hankel matrix form and split into two parts, the so-called *past data* and *future data*, as follows

$$\begin{bmatrix} U_{\mathrm{p}} \\ U_{\mathrm{f}} \end{bmatrix} = \left[\begin{array}{cccc} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \ldots & \boldsymbol{u}_{T-T_{\mathrm{p}}-T_{\mathrm{f}}+1} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{u}_{T_{\mathrm{p}}} & \boldsymbol{u}_{T_{\mathrm{p}}+1} & \ldots & \boldsymbol{u}_{T-T_{\mathrm{f}}} \\ \hline \boldsymbol{u}_{T_{\mathrm{p}}+1} & \boldsymbol{u}_{T_{\mathrm{p}}+2} & \ldots & \boldsymbol{u}_{T-T_{\mathrm{f}}+1} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{u}_{T_{\mathrm{p}}+T_{\mathrm{f}}} & \boldsymbol{u}_{T_{\mathrm{p}}+T_{\mathrm{f}}+1} & \ldots & \boldsymbol{u}_T \end{array}\right], \tag{5-4}$$

For the input data and output data, this results in

$$\begin{bmatrix} U_\mathrm{p} \\ U_\mathrm{f} \end{bmatrix} := \mathscr{H}_{T_\mathrm{p}+T_\mathrm{f}}\left(\boldsymbol{u}^\mathrm{d}\right), \quad \begin{bmatrix} Y_\mathrm{p} \\ Y_\mathrm{f} \end{bmatrix} := \mathscr{H}_{T_\mathrm{p}+T_\mathrm{f}}\left(\boldsymbol{y}^\mathrm{d}\right), \tag{5-5}$$

where $U_\mathrm{p} \in \mathbb{R}^{(T_\mathrm{p} n_{\boldsymbol{u}})\times(T-T_\mathrm{p}-T_\mathrm{f}+1)}$, the *past input data* consists of the first T_p block rows of $\mathscr{H}_{T_\mathrm{p}+T_\mathrm{f}}(\boldsymbol{u}^\mathrm{d})$ and $U_\mathrm{f} \in \mathbb{R}^{(T_\mathrm{f} n_{\boldsymbol{u}})\times(T-T_\mathrm{p}-T_\mathrm{f}+1)}$, the *future input data*, consists of the last T_f block rows of $\mathscr{H}_{T_\mathrm{p}+T_\mathrm{f}}(\boldsymbol{u}^\mathrm{d})$. Similar for the *past output data* $Y_\mathrm{p} \in \mathbb{R}^{(T_\mathrm{p} n_{\boldsymbol{y}})\times(T-T_\mathrm{p}-T_\mathrm{f}+1)}$ and the *future output data* $Y_\mathrm{f} \in \mathbb{R}^{(T_\mathrm{f} n_{\boldsymbol{y}})\times(T-T_\mathrm{p}-T_\mathrm{f}+1)}$. The subscript p denotes the past data that is used for initial condition estimation of the underlying system state, and the subscript f denotes the future data that is used to predict the future system trajectories. The data needs to satisfy the following condition: $T \geq (n_{\boldsymbol{u}} + n_{\boldsymbol{y}} + 1)(T_\mathrm{p} + T_\mathrm{f} + n_{\boldsymbol{y}}) - 1$.

Using the Fundamental Lemma presented in [40], the dynamical system $\mathscr{B}$ can be described parametric-free if there exists $g \in \mathbb{R}^{T-T_\mathrm{p}-T_\mathrm{f}+1}$ such that

$$\begin{bmatrix} U_\mathrm{p} \\ Y_\mathrm{p} \\ U_\mathrm{f} \\ Y_\mathrm{f} \end{bmatrix} g = \begin{bmatrix} \boldsymbol{u}_\mathrm{ini} \\ \boldsymbol{y}_\mathrm{ini} \\ \boldsymbol{u}_\mathrm{fut} \\ \boldsymbol{y}_\mathrm{fut} \end{bmatrix}. \tag{5-6}$$

Furthermore, if $T_\mathrm{p} \geq \ell(\mathscr{B})$, the uniqueness result which is presented in [40] implies that a unique $\boldsymbol{x}_\mathrm{ini} \in \mathbb{R}^{n(\mathscr{B})}$ exists for which the output $\boldsymbol{y}$ is uniquely determined. Future system trajectories can be predicted given an initial trajectory $\mathrm{col}(\boldsymbol{u}_\mathrm{ini}, \boldsymbol{y}_\mathrm{ini}) \in \mathscr{B}_\mathrm{ini}$ and the offline collected data U_p, U_f, Y_p and Y_f. Also, given a desired reference output trajectory $\boldsymbol{y}$, the corresponding feed-forward control input sequence $\boldsymbol{u}$ can be computed.

DeePC algorithm

The optimisation problem solved in the DeePC algorithm is then given by

$$\begin{aligned} \min_{g,\boldsymbol{u},\boldsymbol{y}} \quad & \sum_{k=0}^{T_\mathrm{f}-1} \left(\|\boldsymbol{y}_k - \boldsymbol{r}_{t+k}\|_Q^2 + \|\boldsymbol{u}_k\|_R^2 \right), \\ \text{s.t.} \quad & \begin{bmatrix} U_\mathrm{p} \\ Y_\mathrm{p} \\ U_\mathrm{f} \\ Y_\mathrm{f} \end{bmatrix} g = \begin{bmatrix} \boldsymbol{u}_\mathrm{ini} \\ \boldsymbol{y}_\mathrm{ini} \\ \boldsymbol{u}_\mathrm{fut} \\ \boldsymbol{y}_\mathrm{fut} \end{bmatrix}, \\ & \boldsymbol{u}_k \in \mathcal{U}, \quad \forall k \in \{0, \ldots, T_\mathrm{f}-1\}, \\ & \boldsymbol{y}_k \in \mathcal{Y}, \quad \forall k \in \{0, \ldots, T_\mathrm{f}-1\}, \end{aligned} \tag{5-7}$$

where $N_p \in \mathbb{Z}_{>0}$ is the prediction horizon, $\boldsymbol{r} = (\boldsymbol{r}_0, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_{N_p}) \in (\mathbb{R}^{n_{\boldsymbol{y}}})^{\mathbb{Z}_{\geq 0}})$ the reference trajectory, $\mathrm{col}(\boldsymbol{u}_\mathrm{ini}, \boldsymbol{y}_\mathrm{ini}) \in \mathscr{B}_{T_\mathrm{p}}$ the past input/output data, $\mathcal{U} \subseteq \mathbb{R}^{n_{\boldsymbol{u}}}$ the input constraint set, $\mathcal{Y} \subseteq \mathbb{R}^{n_{\boldsymbol{y}}}$ the output constraint set, $Q \in \mathbb{R}^{n_{\boldsymbol{y}} \times n_{\boldsymbol{y}}}$ the output cost matrix, $R \in \mathbb{R}^{n_{\boldsymbol{u}} \times n_{\boldsymbol{u}}}$ the input cost matrix.

The key difference with the equivalent MPC optimisation problem is that the Input-output (I/O) samples completely replace the system model and state estimate in the optimisation problem. The DeePC algorithm is now given by

Algorithm 5 DeePC algorithm

Input: $\text{col}(\boldsymbol{u}^{\text{d}}, \boldsymbol{y}^{\text{d}}) \in \mathscr{B}_T$, reference trajectory $\boldsymbol{r} \in \mathbb{R}^{N_p}$, past input/output data $\text{col}\,(\boldsymbol{u}_{\text{ini}}, \boldsymbol{y}_{\text{f}}) \in \mathscr{B}_{T_{\text{f}}}$, constraint sets $\mathcal{U}$ and $\mathcal{Y}$, and performance matrices Q and R

1: Solve (5-7) for $g^\star$.
2: Compute the optimal input sequence $\boldsymbol{u}^\star = U_f g^\star$.
3: Apply input $(\boldsymbol{u}(t), \ldots, \boldsymbol{u}(t+s)) = (\boldsymbol{u}_0^\star, \ldots, \boldsymbol{u}_s^\star)$ for some $s \leq N_{\text{p}} - 1$.
4: Set t to $t+s$ and update $\boldsymbol{u}_{\text{ini}}$ and $\boldsymbol{y}_{\text{ini}}$ to the T_{f} most recent input/output measurements.
5: Return to 1.

A schematic representation of the DeePC algorithm is presented in Figure 5-1

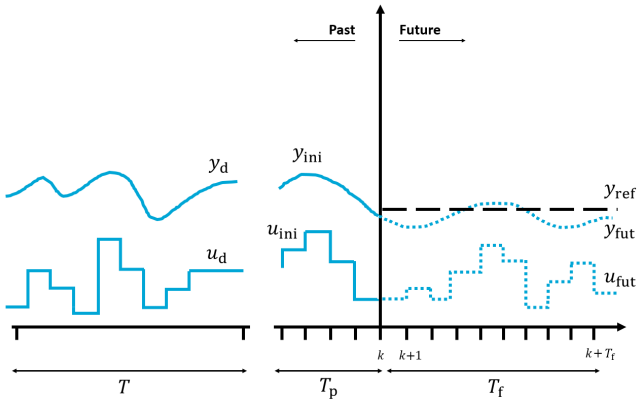


Figure 5-1: DeePC scheme.

5-2 Fruit yield maximisation DeePC

In this section, the DeePC formulation is leveraged to perform the fruit yield maximisation in a parametric-free form. First, the future crop output prediction using the systems' behavioural theory is implemented. Then, future $T_{\text{Air,ref}}$ trajectories using a leveraged DeePC algorithm are shown. For the fruit yield maximisation subproblem, a sampling time of $T_{\text{s}} = 6$[h] is used to match the sampling time of the upper layer EMPC implementation.

5-2-1 DeePC greenhouse-crop system description

The crop production control problem includes, besides the inputs and outputs, also the exogenous weather inputs. When considering the problem in the setting that the considered system is the crop subsystem, the greenhouse indoor climate outputs are also implemented as exogenous inputs. Here, data Hankel matrices are constructed of the recorded exogenous input signals $\boldsymbol{v}_{\text{g}}^{\text{d}} = \text{col}(\boldsymbol{v}_{\text{g}}(1), \ldots, \boldsymbol{v}_{\text{g}}(T))$ and greenhouse indoor climate signals $\boldsymbol{y}_{\text{g}}^{\text{d}} = \text{col}(\boldsymbol{y}_{\text{g}}(1), \ldots, \boldsymbol{y}_{\text{g}}(T))$ as follows

$$\mathscr{H}_{T_{\text{p}}+T_f}(\boldsymbol{v}_{\text{g}}) := \begin{bmatrix} \boldsymbol{v}_{\text{g}}(1) & \cdots & \boldsymbol{v}_{\text{g}}(T - T_{\text{p}} - T_{\text{f}} + 1) \\ \vdots & \ddots & \vdots \\ \boldsymbol{v}_{\text{g}}(T_{\text{p}} + T_{\text{f}}) & \cdots & \boldsymbol{v}_{\text{g}}(T) \end{bmatrix}, \tag{5-8}$$

$$\mathscr{H}_{T_\mathrm{p}+T_f}(\boldsymbol{y}_\mathrm{g}) := \begin{bmatrix} \boldsymbol{y}_\mathrm{g}(1) & \cdots & \boldsymbol{y}_\mathrm{g}(T - T_\mathrm{p} - T_\mathrm{f} + 1) \\ \vdots & \ddots & \vdots \\ \boldsymbol{y}_\mathrm{g}(T_\mathrm{p} + T_\mathrm{f}) & \cdots & \boldsymbol{y}_\mathrm{g}(T) \end{bmatrix}. \tag{5-9}$$

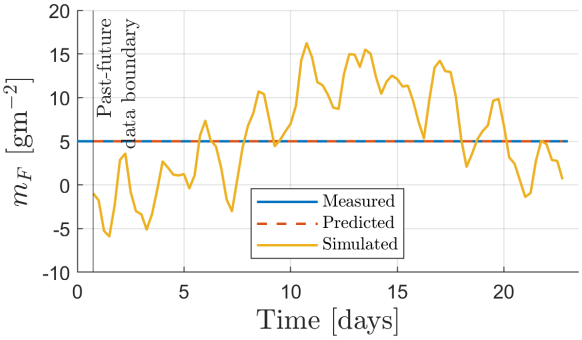
As such, the data are partitioned in the same manner into past and future parts as

$$\begin{bmatrix} V_\mathrm{p} \\ V_\mathrm{f} \end{bmatrix} := \mathscr{H}_{T_\mathrm{p}+N}\left(\boldsymbol{v}^\mathrm{d}\right), \quad \begin{bmatrix} Y_\mathrm{g,p} \\ Y_\mathrm{g,f} \end{bmatrix} := \mathscr{H}_{T_\mathrm{p}+N}\left(\boldsymbol{y}_\mathrm{g}^\mathrm{d}\right), \tag{5-10}$$

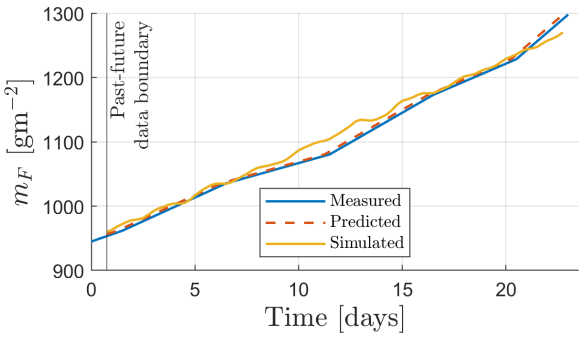
and then the equality constraints in (5-7), that are the parametric-free system representations, are changed to

$$\begin{bmatrix} U_\mathrm{g,p} \\ Y_\mathrm{c,p} \\ V_\mathrm{g,p} \\ Y_{c,\mathrm{p}} \\ U_\mathrm{g,f} \\ V_\mathrm{g,f} \\ Y_\mathrm{g,f} \end{bmatrix} g = \begin{bmatrix} \boldsymbol{u}_\mathrm{g,ini} \\ \boldsymbol{y}_\mathrm{c,ini} \\ \boldsymbol{v}_\mathrm{g,ini} \\ \boldsymbol{y}_\mathrm{g,ini} \\ \boldsymbol{u}_\mathrm{g,fut} \\ \boldsymbol{v}_\mathrm{g,fut} \\ \boldsymbol{y}_\mathrm{g,fut} \end{bmatrix}, \tag{5-11}$$

One would expect $\boldsymbol{y}_\mathrm{c,ini}$, $\boldsymbol{v}_\mathrm{c,fut}$, $\boldsymbol{u}_\mathrm{c,ini}$ and $\boldsymbol{u}_\mathrm{c,fut}$ since the crop subsystem is the considered system in the fruit maximisation subproblem. However, the greenhouse indoor climate variables $\boldsymbol{y}_g$ are influenced by all weather variables and all greenhouse actuators, which in turn influence the crop dynamics. Therefore, not $\boldsymbol{y}_\mathrm{c,ini}$, $\boldsymbol{v}_\mathrm{g,fut}$, $\boldsymbol{u}_\mathrm{c,ini}$ and $\boldsymbol{u}_\mathrm{c,fut}$ but rather $\boldsymbol{y}_\mathrm{g,ini}$, $\boldsymbol{v}_\mathrm{g,fut}$, $\boldsymbol{u}_\mathrm{g,ini}$ and $\boldsymbol{u}_\mathrm{g,fut}$ are included in the parametric-free system representation. This is also to include the control cost forecast depending on the long-term weather prediction into the upper layer control problem. When all other trajectories are known, this system representation can be used to predict future outputs $\boldsymbol{y}_\mathrm{c,fut}$, by solving for g via a least-squares minimisation. Vice-versa, if desired future outputs $\boldsymbol{y}_\mathrm{c,fut}$ are known, corresponding future inputs $\boldsymbol{u}_\mathrm{g,fut}$ can be calculated. In Figures 5-2a to 5-2b, the future crop output prediction via the parametric-free system representation at the beginning and end of the growing season is shown.



(a) Predicted crop outputs by DeePC and sys ID linear model at the beginning of the growing season.



(b) Predicted crop outputs by DeePC and sys ID linear model at the end of the growing season.

Figure 5-2: DeePC future crop output prediction capabilities compared to the linear identified model and measurement data at the beginning and end of the growing season.

The prediction is accurate because measurements of the crop over the entire season are used for the offline system representation construction. Because of the long sampling time, which leads

to only 664 data points for the whole season, this can be implemented without tractability issues. It can be concluded that the prediction is accurate enough to use for DeePC fruit yield maximisation.

5-2-2 DeePC crop production control algorithm

After the parametric-free system representation has been adapted to fit the greenhouse-crop system, the leveraged optimisation problem and corresponding algorithm for the crop production control DeePC will be presented in this section. First, the optimisation problem will be given, then the parameter choices are elaborated on. Finally, the issues corresponding to the implementation over the entire season are mentioned.

Crop production control DeePC

First, the optimisation problem for the linear deterministic case will be given. Hereafter, the optimisation problem for the non-linear or stochastic case will be shown, which is ultimately implemented due to the stochastic nature of the exogenous inputs and the non-linear behaviour of the system at hand.

Linear crop production control DeePC To employ the DeePC setting for fruit yield optimisation, the following optimisation problem set-up needs to be solved:

$$\begin{aligned}
\min_{g} \quad & \sum_{k=0}^{T_{\mathrm{c,f}}-1} \left(\|\boldsymbol{u}_k\|_R^2 \right) - C_{\mathrm{F}} Y_{\mathrm{c,f}} g(T_{\mathrm{c,f}}), \\
\text{s.t.} \quad & \begin{bmatrix} U_{\mathrm{g,p}} \\ Y_{\mathrm{c,p}} \\ V_{\mathrm{g,p}} \\ Y_{\mathrm{g,p}} \\ V_{\mathrm{g,f}} \end{bmatrix} g = \begin{bmatrix} \boldsymbol{u}_{\mathrm{g,ini}} \\ \boldsymbol{y}_{\mathrm{c,ini}} \\ \boldsymbol{v}_{\mathrm{g,ini}} \\ \boldsymbol{y}_{\mathrm{g,ini}} \\ \boldsymbol{v}_{\mathrm{g,fut}} \end{bmatrix}, \\
& \boldsymbol{u}_k \in \mathcal{U}, \quad \forall k \in \{0, \dots, T_{\mathrm{c,f}} - 1\}, \\
& \boldsymbol{y}_k \in \mathcal{Y}, \quad \forall k \in \{0, \dots, T_{\mathrm{c,f}} - 1\}.
\end{aligned} \tag{5-12}$$

Here, the first term represents the costs associated with the controls and the second term represents the revenue obtained from from the fruit weight. C_{F} is a weighing factor for the fruit dry weight revenue, and $Y_{\mathrm{c,f}} g(T_{\mathrm{c,f}})$ denotes the last entry of the vector $Y_{\mathrm{c,f}} g$, which is the predicted m_{F} at the end of the prediction horizon.

Non-linear or stochastic crop production control DeePC When the DeePC scheme is used for non-linear or stochastic systems, the optimisation with strict equality constraints might not be able to find a solution. The greenhouse-crop system is under influence of the stochastic weather exogenous inputs. Therefore, the equality constraints are softened, resulting in the

following optimisation procedure:

$$\begin{aligned}
\min_{g} \quad & \sum_{k=0}^{T_{\mathrm{c,f}}-1} \left(\|\boldsymbol{u}_k\|_R^2 \right) - C_{\mathrm{F}} Y_{\mathrm{c,f}} g(T_{\mathrm{c,f}}) \\
& + \lambda_{\mathrm{V_{g,p}}} \|V_{\mathrm{c,p}} g - v_{\mathrm{c,ini}}\|_2^2 + \lambda_{\mathrm{V_{g,f}}} \|V_{\mathrm{c,f}} g - v_{\mathrm{c,fut}}\|_2^2 \\
\text{s.t.} \quad & \begin{bmatrix} U_{\mathrm{g,p}} \\ Y_{\mathrm{c,p}} \\ Y_{\mathrm{g,f}} \end{bmatrix} g = \begin{bmatrix} \boldsymbol{u}_{\mathrm{g,ini}} \\ \boldsymbol{y}_{\mathrm{c,ini}} \\ \boldsymbol{y}_{\mathrm{g,ini}} \end{bmatrix} \\
& \boldsymbol{u}_k \in \mathcal{U}, \forall k \in \{0, \ldots, T_{\mathrm{c,f}} - 1\}, \\
& \boldsymbol{y}_k \in \mathcal{Y}, \forall k \in \{0, \ldots, T_{\mathrm{c,f}} - 1\},
\end{aligned} \tag{5-13}$$

here, $\lambda_{\mathrm{V_{g,p}}}$ and $\lambda_{\mathrm{V_{g,f}}}$ are regularisation parameters.

Parameter choices

In the crop production control DeePC setting, multiple parameters need to be chosen and tuned. The parameters are the time parameters needed to split the measured data into past and future parts and the costs on the different variables in the leveraged DeePC algorithm.

Choice of time parameters To predict the future crop outputs, the following values were chosen: $T_{\mathrm{c,p}} = 4$, $T_{\mathrm{c,f}} = 87$ and $T_{\mathrm{c}} = 664$. These values were chosen such that measurements of the entire growing season were used and to satisfy the condition $T_{\mathrm{c}} \geq (n_{\boldsymbol{u}_{\mathrm{c}}} + n_{\boldsymbol{y}_{\mathrm{g}\to\mathrm{c}}} + n_{\boldsymbol{v}_{\mathrm{c}}} + 1)(T_{\mathrm{c,p}} + T_{\mathrm{c,f}} + n_{\boldsymbol{y}_{\mathrm{c}}}) - 1$.

This results in DeePC data matrices with the following dimensions. $U_{\mathrm{g,p}} \in \mathbb{R}^{(n_{\boldsymbol{u}_{\mathrm{g}}} T_{\mathrm{c,p}}) \times (T_{\mathrm{c}} - T_{\mathrm{c,p}} - T_{\mathrm{c,f}} + 1)}$, $Y_{\mathrm{c,p}} \in \mathbb{R}^{(n_{\boldsymbol{y}_{\mathrm{c}}} T_{\mathrm{c,p}}) \times (T_{\mathrm{c}} - T_{\mathrm{c,p}} - T_{\mathrm{c,f}} + 1)}$, $V_{\mathrm{g,p}} \in \mathbb{R}^{(n_{\boldsymbol{v}_{\mathrm{g}}} T_{\mathrm{c,p}}) \times (T_{\mathrm{c}} - T_{\mathrm{c,p}} - T_{\mathrm{c,f}} + 1)}$, $Y_{\mathrm{g,p}} \in \mathbb{R}^{(n_{\boldsymbol{y}_{\mathrm{g}}} T_{\mathrm{c,p}}) \times (T_{\mathrm{c}} - T_{\mathrm{c,p}} - T_{\mathrm{c,f}} + 1)}$, $U_{\mathrm{g,f}} \in \mathbb{R}^{(n_{\boldsymbol{u}_{\mathrm{g}}} T_{\mathrm{c,f}}) \times (T_{\mathrm{c}} - T_{\mathrm{c,p}} - T_{\mathrm{c,f}} + 1)}$, $Y_{\mathrm{c,f}} \in \mathbb{R}^{(n_{\boldsymbol{y}_{\mathrm{c}}} T_{\mathrm{c,f}}) \times (T_{\mathrm{c}} - T_{\mathrm{c,p}} - T_{\mathrm{c,f}} + 1)}$, $V_{\mathrm{g,f}} \in \mathbb{R}^{(n_{\boldsymbol{v}_{\mathrm{g}}} T_{\mathrm{c,f}}) \times (T_{\mathrm{c}} - T_{\mathrm{c,p}} - T_{\mathrm{c,f}} + 1)}$, $Y_{\mathrm{g,f}} \in \mathbb{R}^{(n_{\boldsymbol{y}_{\mathrm{g}}} T_{\mathrm{c,f}}) \times (T_{\mathrm{c}} - T_{\mathrm{c,p}} - T_{\mathrm{c,f}} + 1)}$

Tuning of weighing parameters The parameter that sets the relation between the control costs and the output costs is C_{F}. A value of $C_{\mathrm{F}} = 3 \cdot 10^4$ resulted in temperature setpoints that are reasonable.

The control cost matrix R has been set to $R = \mathrm{diag}([1, 0, 0, 10^6, 1, 1, 0, 0])$, to penalise the actuators that are taken into account in the final cost metric. These actuators are the heating pipe systems, the lights and the CO_2 injection.

$\lambda_{V_{\mathrm{g,p}}} = \lambda_{V_{\mathrm{g,f}}}$ were set to 10^4.

Seasonly DeePC issues

When implementing the crop production control DeePC for the entire season, some issues are encountered that correspond to the assumptions and conditions that are placed on the data due to the particular DeePC data-based framework. These issues are the initialisation at the beginning of the season, the shrinking horizon implementation and the effect of not optimising over the entire growing season for the first 574 DeePC iterations.

DeePC initialisation Data is available for 166 days of the growing season. For the upper layer subproblem with corresponding sampling time, this means there are data trajectories with 664 time instances. At time instance $k^{\mathrm{s}} = 1$, an additional $T_{\mathrm{c,p}}$ measurement time instances are needed for the first $T_{\mathrm{c,p}}$ iterations to be able to pass initial trajectories to the DeePC algorithm. This is accommodated for by taking the first $T_{\mathrm{c,p}}$ entries of $\boldsymbol{y}_{\mathrm{c}}$, $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{y}_{\mathrm{g}}$, flipping them and appending them at the beginning of the dataset, as schematically depicted in Figure 5-3 below.

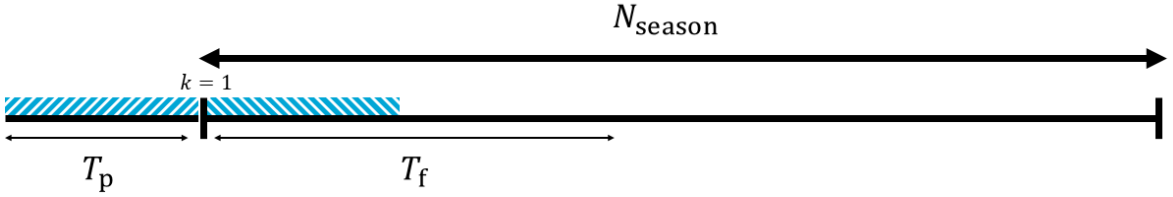


Figure 5-3: DeePC data manipulation so that the first $T_{\mathrm{c,p}}$ iterations also possess initial trajectories. N_{season} is the total number of iterations for the whole growing season. At the first time instance the problem is solved, $k^{\mathrm{s}} = 1$, T_{p} datapoints are needed, as $\boldsymbol{y}_{\mathrm{c,ini}}$, $\boldsymbol{u}_{\mathrm{ini}}$, $\boldsymbol{v}_{\mathrm{c,ini}}$ and $\boldsymbol{y}_{\mathrm{g,ini}}$ are needed in 5-13. In the picture, it is shown that this is circumvented by taking the first available T_{p} datapoints and flipping them.

Note that the ratios in the figure do not correspond to the true implemented ratios.

Accommodation scheme for final fruit weight maximisation When maximising the fruit yield, the time the final iteration of the prediction horizon represents is of importance. This is because, in reality, the fruit weight does not always increase but at times also decreases due to maintenance respiration. The iteration to maximise for is the 3$^{\mathrm{rd}}$ time instance of the day, which is the end of the midday.

Due to the nature of the DeePC set-up, not the entire growing season can be used for the upper layer fruit maximisation procedure. This is because of the conditions that are imposed on the T parameters in the DeePC set-up. Hence, here an accommodation scheme is needed. Therefore, not a fixed prediction horizon $T_{\mathrm{c,f}}$ is used, but rather an alternating one, given by $T_{\mathrm{c,f}} = [\underbrace{\mathcal{T}_{\mathrm{k}} \dots \mathcal{T}_{\mathrm{k}}}_{144\mathrm{x}} \quad \mathcal{T}_{\mathrm{end}}]$. Here, $\mathcal{T}_{\mathrm{k}} = [87 \quad 86 \quad 85 \quad 84]$ and $\mathcal{T}_{\mathrm{end}} = [83 \dots 1]$.

Shrinking horizon DeePC When the algorithm proceeds, at some point, the DeePC implementation needs to have a shrinking horizon implementation, as can be seen in the paragraph above. The shrinking horizon DeePC implementation and its implications will be presented in this section.

When a shrinking horizon is implemented, the data matrices $U_{\mathrm{g,p}}$, $Y_{\mathrm{c,p}}$, $V_{\mathrm{g,p}}$ and $Y_{\mathrm{g,p}}$ that are used to estimate the current state do not change in dimension. The data matrices that are used to predict the future outputs $U_{\mathrm{g,f}}$, $Y_{\mathrm{c,f}}$, $V_{\mathrm{g,f}}$ and $Y_{\mathrm{g,f}}$ do change in dimension with changing horizon. This happens as follows:

$$\mathscr{H}_{T_{\mathrm{p}}+T_f}(\boldsymbol{d}) := \left[\begin{array}{c|cc} \boldsymbol{d}_1 & \cdots & \boldsymbol{d}_{T-T_{\mathrm{p}}-T_f+1} \\ \hline \vdots & \ddots & \vdots \\ \boldsymbol{d}_{T_{\mathrm{p}}+T_f} & \cdots & \boldsymbol{d}_T \end{array}\right], \tag{5-14}$$

here, $\boldsymbol{d}$ is any offline collected data trajectory $\boldsymbol{u}_{\mathrm{g}}$, $\boldsymbol{y}_{\mathrm{c}}$, $\boldsymbol{v}_{\mathrm{g}}$ or $\boldsymbol{y}_{\mathrm{g}}$. The Hankel matrix used at iteration k^{s} is the entire matrix depicted in (5-14), and the Hankel matrix used at iteration $k^{\mathrm{s}}+1$ is the down-right part of the previous data matrix.

The resulting two-layer crop production control including the upper layer setpoint generating DeePC controller is then given in Algorithm 6 below.

Algorithm 6 DeePC two-layer crop production control problem algorithm

Input: $\mathrm{col}(\boldsymbol{u}_{\mathrm{g}}^{\mathrm{d}}, \boldsymbol{y}_{\mathrm{c}}^{\mathrm{d}}, \boldsymbol{y}_{\mathrm{g}}^{\mathrm{d}}, \boldsymbol{v}_{\mathrm{g}}^{\mathrm{d}}) \in \mathscr{B}_T$, constraint sets $\mathcal{U}^{\mathrm{s}}$ and $\mathcal{Y}^{\mathrm{s}}$, and performance matrices Q and R

1: Solve (5-13) for $g^{\star}$.
2: Compute optimal greenhouse climate outputs $\boldsymbol{y}_{\mathrm{g}}^{\star} = Y_{\mathrm{g,f}} g^{\star}$.
3: Send air temperature reference $T_{\mathrm{Air,ref}}$ to lower layer.
4: **for** $k^{\mathrm{f}} \in \{1, \ldots, N_{\mathrm{i}}^{\mathrm{f}}\}$
   a: Obtain initial state estimate $\hat{\boldsymbol{x}}_{\mathrm{gh}}(k_{\mathrm{gh}}^{\mathrm{f}})$.
   b: Solve (3-2) for the optimal input sequence $(\boldsymbol{u}_{\mathrm{gh}}^{\mathrm{f}})^* = \big((\boldsymbol{u}_0^{\mathrm{f}})^*, \cdots, (\boldsymbol{u}_{N_{\mathrm{h}}^{\mathrm{f}}}^{\mathrm{f}})^*\big)$.
   c: Apply only the first input $\boldsymbol{u}(k^{\mathrm{f}}) = (\boldsymbol{u}_{\mathrm{gh},0}^{\mathrm{f}})^*$.
   d: Set k^{f} to $k^{\mathrm{f}}+1$.
   e: Return to 4.
5: **end for**
6: Set k^{s} to $k^{\mathrm{s}}+1$ and update $\boldsymbol{u}_{\mathrm{g,ini}}$, $\boldsymbol{y}_{\mathrm{c,ini}}$, $\boldsymbol{v}_{\mathrm{g,ini}}$ and $\boldsymbol{y}_{\mathrm{g,ini}}$ to the T_{ini} most recent input/output measurements.
7: Return to 1.

5-3 Fruit maximisation DeePC results

The results for the setpoint generating DeePC algorithm are divided into the climate strategy, resource usage, fruit yield and net economic profit. These results will be presented in this section.

5-3-1 DeePC climate strategy results

In Figure 5-4 below, the generated reference temperatures over the entire season are shown.

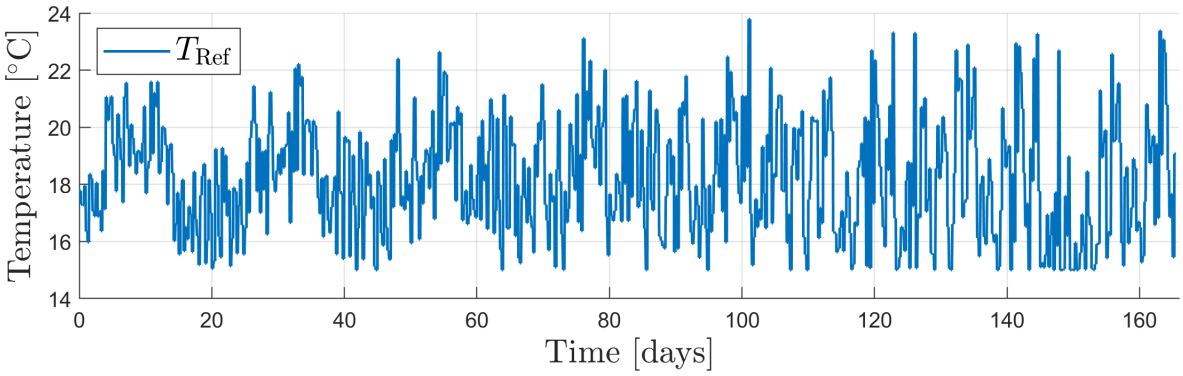


Figure 5-4: Reference temperatures generated by DeePC over the entire growing season.

Here, the output constraint set $\mathcal{Y}^{\text{s}}$ implements the upper and lower bound on the generated reference temperature by $T_{\text{Air}} \geq 15°\text{C}$ and $T_{\text{Air}} \leq 30°\text{C}$, which the generated temperature setpoints satisfy. The weekly averages for the generated DeePC greenhouse indoor air temperature setpoints and corresponding average other greenhouse indoor variables are shown in Figures 5-5 to 5-7 below.

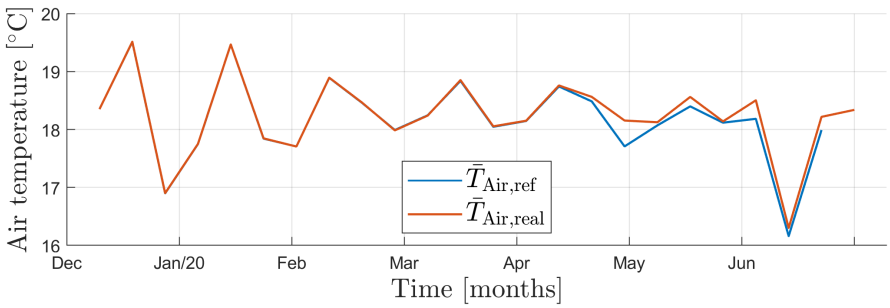


Figure 5-5: Weekly average of $T_{\text{Air,ref}}$, resulting from the DeePC climate strategy.

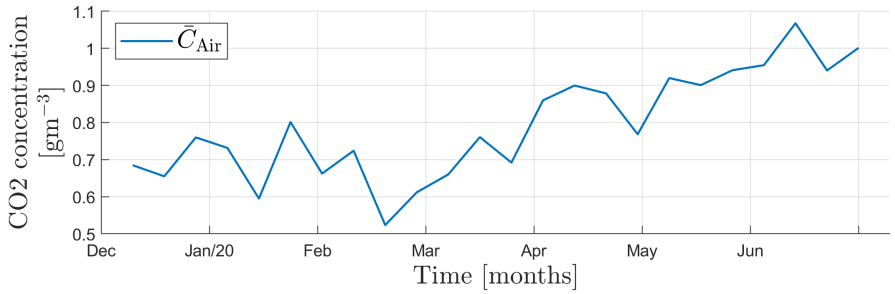


Figure 5-6: Weekly average of C_{Air} that accompanies the tracked $T_{\text{Air,ref}}$, resulting from the DeePC climate strategy.

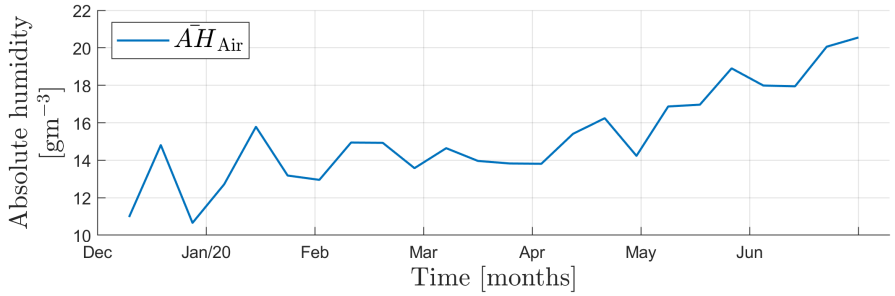
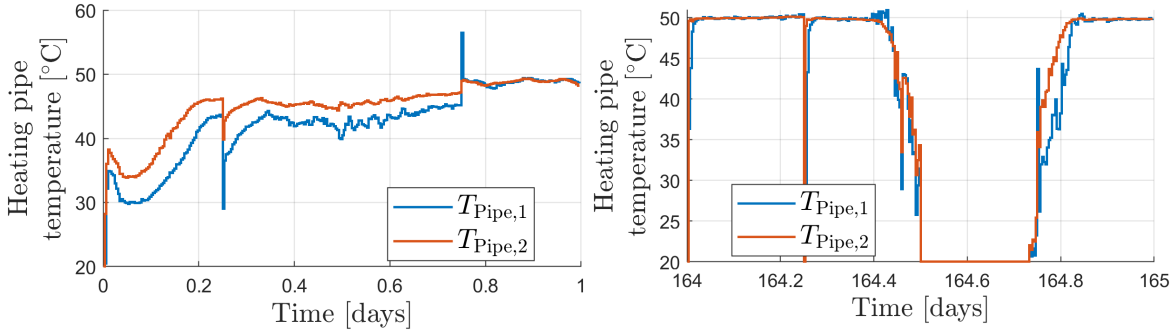


Figure 5-7: Weekly average of AH_{Air} that accompanies the tracked $T_{\text{Air,ref}}$, resulting from the DeePC climate strategy.

5-3-2 DeePC resource usage

Over the season, the weather conditions but also the climate strategy varies. Thus, the calculated control inputs vary too. Not every day of the season will be shown, but rather the control inputs calculated by the lower layer setpoint tracking controller to obtain the DeePC climate strategies for the first and last day of the growing season are shown, as these days vary the most in terms of weather conditions and crop conditions. The calculated optimal heating pipe inputs are presented in Figure 5-8 below.

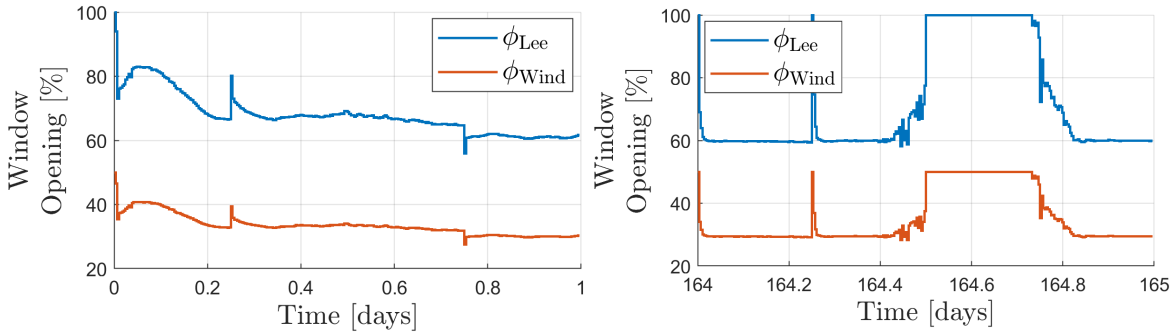


(a) T_{Pipe1} and T_{Pipe2} 1st day of the season.

(b) T_{Pipe1} and T_{Pipe2} 165th day of the season.

Figure 5-8: T_{Pipe1} and T_{Pipe2} for the first and last day of the season, resulting from the DeePC climate strategy.

The calculated optimal window openings are presented in Figure 5-9 below.

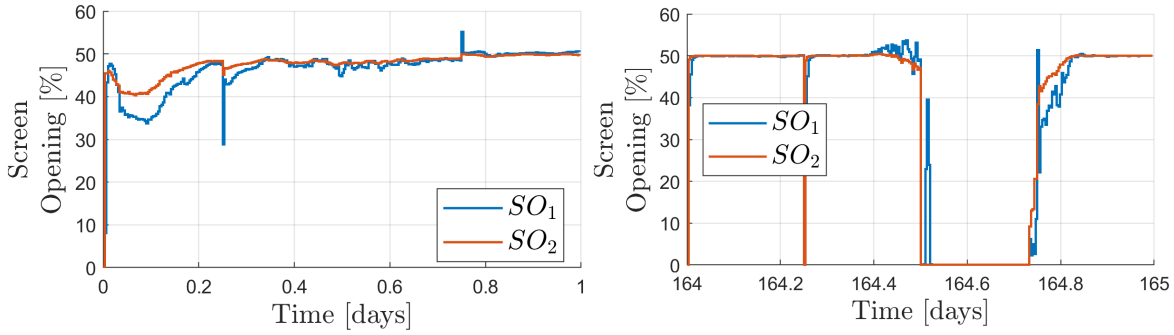


(a) ϕ_{Lee} and ϕ_{Wind} 1st day of the season.

(b) ϕ_{Lee} and ϕ_{Wind} 165th day of the season.

Figure 5-9: ϕ_{Lee} and ϕ_{Wind} for the first and last day of the season, resulting from the DeePC climate strategy.

The calculated optimal screen openings are presented in Figure 5-10 below.

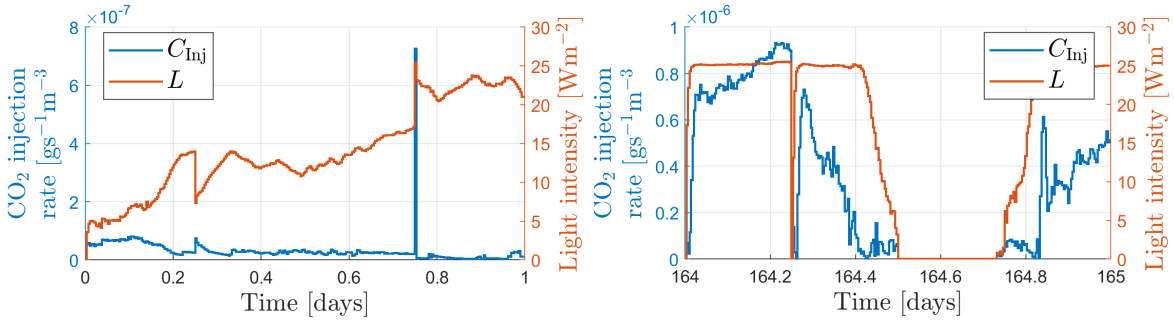


(a) SO_1 and SO_2 1st day of the season.

(b) SO_1 and SO_2 166th day of the season.

Figure 5-10: SO_1 and SO_2 for the first and last day of the season, resulting from the DeePC climate strategy.

The calculated optimal light intensity and CO_2 injection rate are presented in Figure 5-11 below.



(a) C_{Inj} and L 1st day of the season.

(b) C_{Inj} and L 165th day of the season.

Figure 5-11: C_{Inj} and L for the first and last day of the season, resulting from the DeePC climate strategy.

The heating pipes and lights are less used at the end of the season. The windows are used to a greater extent to cope with the warmer weather conditions imposed by the hotter outside temperature and higher global radiation. Also, the screens and lights are employed to control the greenhouse indoor climate towards the setpoints.

5-3-3 DeePC fruit yield results

The fruit yield resulting from the DeePC climate strategy on the ground truth crop simulator is presented in Figure 5-12 below.

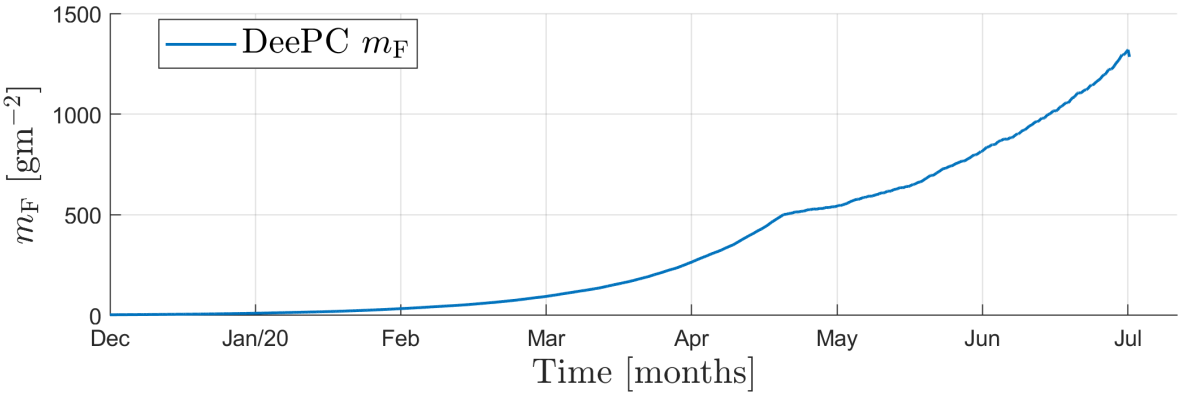


Figure 5-12: Fruit yield, resulting from the DeePC climate strategy.

The amount of kg fresh tomato weight resulting from the DeePC strategy is 802 kg for the considered greenhouse compartment.

5-4 Conclusive remarks on fruit yield maximisation DeePC

It has been shown that the computational burden incurred by using DeePC instead of model-based equivalents is higher [44]. However, when using the DeePC setting in the upper layer of the crop production control problem, this is not a big issue since the time elapsed by one upper layer iteration is six hours. The average computation time of the DeePC algorithm over the season is 41 seconds, thus not incurring too much delay in the control of the greenhouse. During the calculation time of the new setpoint, the old setpoint can be used as the current setpoint until the setpoint is updated in the next lower layer iteration.

The DeePC cannot predict the future outputs over the entire growing season due to the implications imposed by the DeePC set-up, which is contradicting with the original control design objective. This states that the overall problem spans the whole growing season. However, the proposed control architecture with the DeePC is able to handle the two timescales that are present. Moreover, the prediction horizon of the upper layer spans multiple weeks. It can be argued that this is long enough for the slow timescale dynamical system and long enough to take into account the long-term weather prediction.

When the entire growing season of crop data is used for the parametric-free system representation of the crop subsystem, the data must be collected somewhere in the past of a greenhouse

compartment with the crop. This implies that the data used to describe the greenhouse system at hand does not originate from the true system under control, meaning that there is a description mismatch.

Chapter 6

Performance comparison

The performances of the two setpoint generating controllers are evaluated and compared against the predefined benchmark reference temperature trajectory in this chapter. The performance metrics used to compare the results are in terms of the generated climate strategy by the controllers, resource usage, fruit yield and net economic profit.

6-1 Comparison metrics

6-1-1 Climate strategies

The most important aspects of the climate strategy are the temperature strategy, and whether the setpoint tracking controller achieves the devised temperature strategy outputted by the upper layer setpoint generating controller. For the two setpoint generating controllers' average CO_2 and humidity levels, the reader is referred to Chapter 4 and Chapter 5.

The weekly average temperatures were presented already in the chapters that introduce the setpoint generating controllers. In Figure 6-1 below, they are plotted together for comparison.

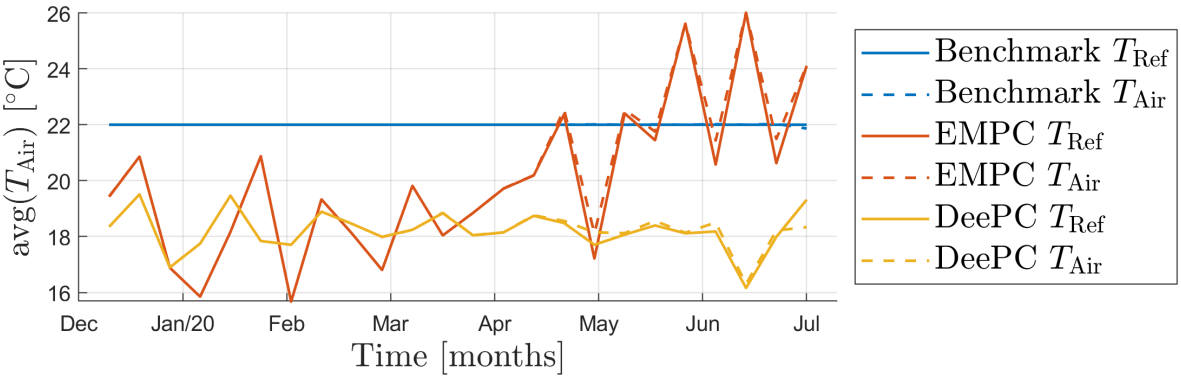


Figure 6-1: Generated reference greenhouse indoor air temperatures by the different setpoint generating controllers and realised greenhouse indoor air temperatures.

The Root mean square error (RMSE) of the greenhouse indoor air temperature is 0.942 °C for the Economic model predictive control (EMPC) setpoint generating controller and 0.535 °C for the Data-enabled predictive control (DeePC) setpoint generating controller. This implies that the climate strategy that the EMPC controller generates is less suitable for the time of the season. The RMSE for the predefined benchmark reference temperature is 0.087 °C.

6-1-2 Resource use and costs

The resource usages are calculated from the implemented control cost. The control costs used in the cost function do not yet represent true resource usages or incurred economic costs by the actuators, which will be the metric used to compare the methods. Therefore, the control costs need this conversion to true economic costs. From the Autonomous greenhouse challenge (AGC) evaluation, the metrics of interest are:

- Heating energy usage [MJ]
- Electricity energy usage [kWh]
- Daily CO2 dosage [g m^{-2}].

From the above measures, the following three efficiencies are deduced:

- Energy use efficiency for heat [MJ kg^{-1} tomato]
- Energy use efficiency for electricity [kWh kg^{-1} tomato]
- CO_2 use efficiency [kg CO_2 dosage kg^{-1} tomato].

The calculations from the applied control input vectors to resource usage are explained in the sections below. The used greenhouse data considered a greenhouse with production area of $A_{\text{prod}} = 62.5$ [m^2] and total area of $A_{\text{gh}} = 96$ [m^2] [13]. The height taken for the calculations is $H_{\text{gh}} = 6$ [m] [45].

Heating energy consumption

The heating energy consumptions by the two heating pipe systems are derived as follows. First, the heat release [W m^{-2}] from the two heating pipes are calculated as [13]:

$$Q_{\text{Pipes}} = 0.62(T_{\text{Pipe,1}} - T_{\text{Air}}) + 2.1(T_{\text{Pipe,2}} - T_{\text{Air}}). \tag{6-1}$$

Hereafter, the heat release Q_{Pipes} is converted to [MJ day^{-1}].

Electricity consumption

The electricity consumption by the artificial lighting system is calculated by converting the measured electricity consumption to kWh. The data on which the control strategies are based includes two lighting systems, High pressure sodium (HPS) and Light emitting diode (LED) lighting. However, the calculated optimal control inputs only give optimal lighting settings for one lighting system. The two lighting systems of the data both have different electricity consumption rates, LED even per light colour. Since the control schemes calculate only one lighting setting, an average electricity consumption rate is chosen, which is the average of the HPS electricity consumption rate and the average LED electricity consumption rate.

LED electricity consumption rates [Wm^{-2}]:

- Blue: 7.27
- Red: 25.3
- Farred: 6.23
- White: 22.72.

This results in an average LED electricity consumption rate when assuming all colours occur evenly in the LED light, of: 15.38 [W m^{-2}].

The average electricity consumption of the lighting system is then $EC = 48.19$ [Wm^{-2}]. Hereafter, the electricity consumption is converted to kWh.

Economic costs by implemented actuator controls

From the resource usages of heating energy, electricity and CO_2, economic costs are calculated by multiplying the usages with given prices. The following prices are used. Electricity has an on-peak price of €0.08 per kWh (07:00-23:00 h) and an off-peak price of €0.04 per kWh (23:00-7:00 h). Heating energy has a fixed price of € 0.03 per kWh. Last, CO_2 costs € 0.08 per kg up to 12 [kg m^{-2}] and € 0.20 per kg above 12 [kg m^{-2}] [13].

In this thesis, the controllers did not take into account the different electricity costs at different times. Hence, the performance metric might be skewed. However, it is insightful for comparison purposes between the controllers since none of the controllers takes the cost into account. Thus both controllers have the performance degradation due to this fact. In Table 6-1 below, the calculated resource usages and accompanying efficiencies for the EMPC and DeePC setpoint generating controllers and the predefined benchmark reference setpoint trajectory are presented.

Table 6-1: An overview of the resource use efficiency per setpoint generating controller.

T_{Ref} generating controller	Heat		Electricity		CO_2	
	Usage [MJ]	**Efficiency [MJ kg^{-1}]**	**Usage [kWh]**	**Efficiency [kWh kg^{-1}]**	**Usage [kg]**	**Efficiency [kg kg^{-1}]**
Benchmark	$1.258{\cdot}10^4$	15.6	$1.670{\cdot}10^4$	21.1	640	0.79
EMPC	$1.374{\cdot}10^4$	16.9	$1.718{\cdot}10^4$	21.1	572	0.70
DeePC	$1.398{\cdot}10^4$	17.4	$1.582{\cdot}10^4$	19.7	507	0.63

It is noted that the climate strategy of the EMPC setpoint generating controller results in medium efficient heating energy usage, less efficient electricity usage and medium efficient CO_2 usage. The DeePC setpoint generating controller results in the least efficient heating energy usage according to the chosen metrics. However it also results in the most efficient electricity and CO_2 usages. The predefined benchmark reference temperature results in the most efficient heating energy usage, medium electricity usage efficiency and the least efficient CO_2 usage.

6-1-3 Fruit yield revenue

Since in this problem setting, a single harvest at the end of the growing season is considered, the price for the tomatoes is based on the final price according to the used dataset. The price is 2€ per kg tomato [13].

In Figure 6-2 below, the resulting fruit yields from the three climate strategies are presented.

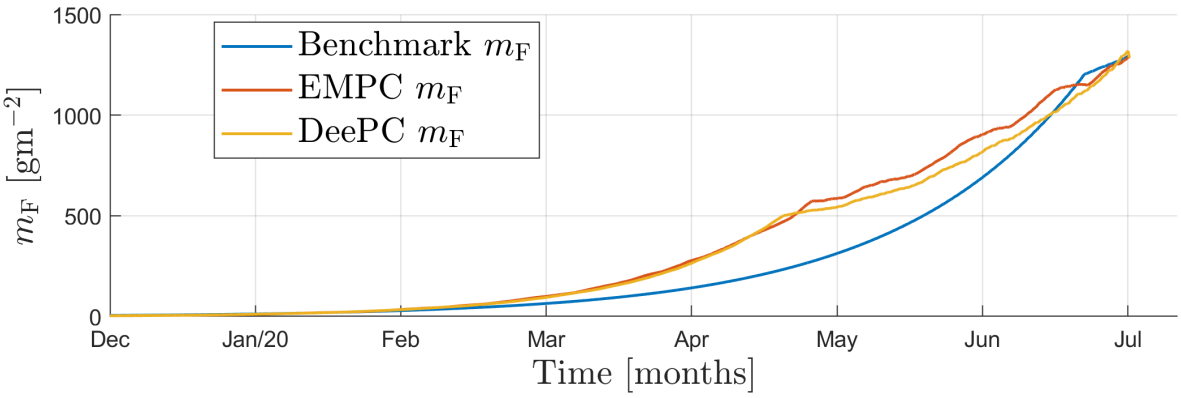


Figure 6-2: Resulting fruit yield from the climate strategies obtained from the setpoint generating EMPC and DeePC controllers and the predefined benchmark reference trajectory.

The two setpoint generating controllers, EMPC and DeePC, result in higher fruit yield rate during the season and lesser fruit yield rate near the end of the season. This could be explained by the fruit respiration. Also, it can be seen that the benchmark temperature strategy results in a smoother fruit yield rate but in a lesser fruit yield rate at the very end of the season. This could be explained by some modelling error in the ground truth model.

It can be concluded that the different climate strategies result in comparable fruit yields at the end of the growing season. The DeePC climate strategy results in 802 kg, the EMPC

climate strategy results in 815 kg and the predefined benchmark climate strategy results in 806 kg.

6-1-4 Net economic profit

The net economic profit is the revenue minus the costs associated with the implemented controls. In Table 6-2 below, the economic measures for the two setpoint generating controllers and predefined benchmark reference trajectory are presented. These are calculated as described in Section 6-1-2.

Table 6-2: Total costs, total income, and net profit (€/m2) for temperature generating control strategies. Calculations were performed according to 6-1-2.

T_{Ref} **generating controller**	**Total cost** [€/m²]	**Total income** [€/m²]	**Net profit** [€/m²]
Benchmark	20.27	25.82	5.55
EMPC	21.00	26.08	5.08
DeePC	19.19	25.68	6.49

It can be concluded that according to these metrics, the DeePC setpoint generating controllers results in the most profit, mainly because the economic costs incurred by the actuators are the least of the three control strategies.

6-2 Discussion

As stated in Section 6-1-1, the climate strategy generated by the EMPC controller might be the least suitable since the RMSE of the EMPC controller is the highest. Since the average temperature generated by the EMPC controller is higher at the end of the season, which would be expected when taking into account the higher outside temperature and higher global radiation, this less suitable climate strategy can be explained by the strong fluctuations in the nocturnal and diurnal setpoints.

The resource usage accompanying the DeePC generated climate strategy is 1.75% higher for the heating when compared to the EMPC controller and 11.1% higher compared to the predefined benchmark climate strategy. Even though the average temperature of the climate strategy generated by the DeePC is, overall, lower than the average temperature of the climate strategy generated by the EMPC controller, the heating usage is somewhat higher. This can be explained by the fact that the nocturnal setpoints of the DeePC climate strategy at the beginning of the season are higher than those of the EMPC climate strategy, resulting in more required heating power. The average temperature of the DeePC climate strategy is only significantly lower than the average temperature of the EMPC climate strategy after half of the season. At this point in time, not as much heating power is required due to the weather conditions that result in warmer greenhouse indoor climate conditions. When compared to the benchmark temperature strategy, the required heating power is even higher because the benchmark temperatures are not very harsh and lie more towards the seasonal average.

The DeePC resource usage for electricity is 7.92% less when compared to the EMPC controller electricity use and 5.27% less when compared to the predefined benchmark climate strategy. The lighting is less for the DeePC climate strategy when compared to the EMPC and predefined benchmark climate strategy also due to the lower average temperature, since the lighting systems were also used as heating sources.

The used CO_2 by the DeePC generated climate strategy is 11.4% less than the used CO_2 for the EMPC climate strategy and 20.8% less than the used CO_2 for the predefined benchmark climate strategy. This can be due to multiple factors. The lower crop yield of the DeePC strategy might partially be explained by this.

The fruit yield analysis has been performed using a non-linear mechanistic calibrated crop model that describes the underlying physical processes of the crop. The fruit yield obtained by the setpoint generating DeePC controller obtains 1.51% fewer tomatoes than the setpoint generating EMPC controller and 0.54% fewer tomatoes than the predefined reference climate strategy. Both the DeePC and the EMPC controllers use different system descriptions than the non-linear crop model that is used as a plant prediction model. This can result in a different temperature strategy than the optimal one for the respective prediction model.

Overall it can be stated that, even though the DeePC strategy results in a slightly lesser fruit yield, the decrease in resource usage compensates for this in the overall economic criterion of net profit.

Chapter 7

Conclusion and discussion

7-1 Conclusions

This thesis aimed to assess whether the crop production can be controlled by data-driven control methods, maximising the fruit yield while minimising the resource usage and employing a long-term weather prediction. Accordingly, the applicability of the novel Data-enabled predictive control (DeePC) algorithm has been investigated in a proposed hierarchical structure. This work has focused on implementing control strategies on the long-term control of the crop instead of merely the daily greenhouse indoor climate control. As such, the proposed control scheme includes the effect of the long-term crop dynamics and weather dynamics throughout the season while the short-term greenhouse indoor climate dynamics are also considered. The three subquestions, accompanying the main research question as introduced in Chapter 1, are recalled to analyse whether the proposed control scheme accomplishes the set requirements.

- **What model will be implemented as a ground truth and what will be the benchmark against which the proposed controllers are compared?**
  In Section 2-2, the non-linear, mechanistic model that describes the crop as a separate subsystem in the greenhouse is introduced [18]. The model needed to describe the underlying physical processes. The crop model has been chosen to describe the tomato crop in a so-called medium-grained fashion, i.e., the tomatoes and leaves are modelled as one weight. This is to keep the dimensionality of the overall crop production control problem workable. To this end, the non-linear crop model has been adapted to include all modern greenhouse controls that were present in the Autonomous greenhouse challenge (AGC) dataset. The main selection criterion for the non-linear crop model was that the greenhouse indoor climate variables and crop variables are described in the same manner as the dataset does. Thereafter, the non-linear crop model has been calibrated to use as a ground truth model for the proposed setpoint generating controller comparison.
- **How will the different timescales on which the greenhouse subsystems act be dealt with?**

The proposed controllers employ a two-layer structure that is introduced in Section 2-3-3. The architecture arises from a temporal and functional decomposition of the overall problem. Here, the upper layer operates at a slower sampling rate and generates climate setpoints which the faster-operating lower layer will track. As such, both timescales are incorporated in the overall problem while addressing multiple objectives. In this work, the focus is on the upper layer control. Hence, the lower layer controller will be a Model predictive control (MPC) controller which is an often-used technique in the greenhouse climate setpoint tracking control as found in the literature [15] [39] [46]. This lower layer setpoint tracking controller is introduced in Chapter 3. Subsequently, the model-based Economic model predictive control (EMPC) control algorithm for the setpoint generation has been proposed in Chapter 4 that will be compared against the DeePC setpoint generating controller, introduced in Chapter 5.

- **How will the novel DeePC algorithm be implemented to maximise the fruit yield?**
  The DeePC algorithm has been described for setpoint tracking in a fashion similar to setpoint tracking MPC. To this aim, the DeePC algorithm has been extended to include two types of exogenous inputs: the weather variables and the greenhouse indoor climate variables that act on the crop subsystem. With the available data and the conditions imposed by the DeePC set-up, the prediction horizon of the DeePC algorithm does not span the entire growing season. However, it does span a longer time period than is possible with a one-layer approach that implements the timescale of the lower layer. After the two types of exogenous inputs are included, the cost function has been adapted to the economic type of cost function and the stochastic nature of the weather variables.

The hierarchical decomposition of the overall problem can be employed to include different control techniques for the upper layer or lower layer when desired. One drawback of the DeePC method for crop production control over the entire growing season is that crop measurement data is needed at each time instance that the problem is solved again to update the initial trajectories. This is costly and perhaps impossible to acquire at the chosen upper layer sampling time since it is currently measured manually [13]. Moreover, it can be concluded that even though DeePC has shown to be computationally heavier than the equivalent MPC [44], for the long-term crop production control problem this is not a problem since the upper layer timescale operates at such a slow sampling rate that the computation time of the setpoint is significantly less than the time expired during one upper layer instance. Concluding, this thesis has identified that the DeePC algorithm can control the crop production, maximising the fruit yield while minimising the control resource usage when it is implemented in a hierarchical fashion.

7-2 Recommendations

The conventional EMPC and newly DeePC setpoint generating controllers can achieve sufficient fruit growth via the outputted reference greenhouse indoor air temperatures. However, multiple aspects could be further explored or improved. To improve the applicability of the data-driven control scheme for the crop production control problem, a further study could assess the following topics.

- **Control cost prediction implementation:** To decrease the resource usage further in terms of the economic costs implied by them, a distinction could be made between the peak hours and off-peak hours of the and electricity costs. To extend further, a prediction could be implemented that uses a forecast of the electricity and fuel cost to include in the optimisation to take cost fluctuations into account over the long term.
- **Crop management:** One of the main challenges in crop production control is how to manipulate the crop characteristics, e.g., plant load, stem density, such that the long-term crop objective of the best tomatoes is achieved. A medium-grained model has been used that does not describe the individual fruits, stems and leaves in this work. The proposed structure and control scheme can be implemented for a more detailed crop simulator which can then be employed to also output crop management decisions [1]. Also, the crop can be employed in the greenhouse indoor climate control since the crop influences the climate via the processes of photosynthesis, respiration and transpiration.
- **Weather implementation:** In this thesis, the weather prediction for the upper layer is different from the weather prediction used for the lower layer. However, within the lower layer itself, the weather prediction is identical to the true implemented weather. To make the problem more realistic, different weather conditions than those predicted should be implemented.
- **Chance-constrained MPC:** Similarities can be seen between building climate management and greenhouse indoor climate management. In this thesis, the stochastic nature of the weather variables was not considered. Certain control schemes implemented in the building climate management employ chance-constrained MPC and could be used as an inspiration for the greenhouse indoor climate control [47] [48]. In this way, the stochastic weather variables are accounted for which can improve the results of the controllers.
- **Lower layer setpoint tracking control:** In this work, the lower layer MPC controller uses positive semi-definite quadratic cost on the control inputs. This results in a non-unique set of solutions in the lower layer setpoint tracking problem since then the overall problem is ill-posed. To overcome this, a scheme can be implemented that penalises the deviation of the target control inputs in a positive definite manner to ensure offset-free control whenever disturbances are present [49] [50]. Implementing a quadratic cost on the control inputs instead of the difference between the control inputs and the target control inputs results in a trade-off between tracking accuracy and resource usage.
- **System representation update:** In this work, the system representation that is used for the setpoint generating and setpoint tracking controllers is the same over the entire growing season. Multiple extensions regarding the system representation update could improve the results, which include:
  - Update of the crop system representation as soon as more recent crop measurements become available. The crop exhibits different dynamics throughout the season due to the increasing biomass. Furthermore, the crop data that is initially used for the system representation is measured offline, meaning the data does not originate from the same greenhouse compartment. Hence updating will make the representation of the considered crop more accurate. The update could regard

either the data-driven DeePC description or the identified linear models that are used in the EMPC setpoint generating controller. This will result in an online type of controller, like Online data-enabled predictive control (ODeePC) or adaptive or online MPC schemes [51].

– Use of a variable upper layer sampling time to account for the phenomenon that the daytime in the winter is not the same as the daytime in summer. Hence, a scheme that accommodates this in the nocturnal and diurnal setpoint generation can be implemented.

Appendix A

Non-linear greenhouse indoor climate model

A-0-1 Greenhouse indoor climate mathematical model

A non-linear mechanistic climate model has been selected and adapted to include the eight types of modern greenhouse actuators for the greenhouse indoor climate. However, the calibration procedure for the Autonomous greenhouse challenge (AGC) dataset did not render a model with the same parameters over the entire season that was capable of describing the greenhouse indoor climate over the entire season.

The model has been selected as it included T_{Air}, C_{Air} and AH_{Air} as main variables, which is needed to match the AGC dataset used for this thesis. Furthermore, the mathematical formulations that describe the rate of change in these thermodynamic variables are based on the physical relations between the different variables within the greenhouse. Among which are ventilation and crop evaporation. It included a range of modern actuators but needed adaptation to include all eight types of modern greenhouse actuators present in the AGC dataset.

In this section, the adapted equations are given, after which the adaptations are stated and explained.

Greenhouse air temperature rate of change $\dot{T}_{\text{Air}}$

$$\dot{T}_{\text{Air}} = \frac{1}{C_{\text{Air}}}\left(Q_{\text{Vent}} + Q_{\text{Covscr}} + Q_{\text{Pipe1}} + Q_{\text{Pipe2}} + Q_{\text{Sun}} + Q_{\text{Trans}} + Q_{\text{Cond}} + Q_{\text{Lamp}}\right), \tag{A-1}$$

where Q_{Vent} [J m^{-2}s^{-1}] is the ventilative heat loss, Q_{Covscr} [J m^{-2}s^{-1}] is the conductive heat loss through the cover and screens, Q_{Pipe1} [J m^{-2}s^{-1}] is the heat gain due to heating pipe system 1, Q_{Pipe2} [J m^{-2}s^{-1}] is the heat gain due to heating pipe system 2, Q_{Sun} [J m^{-2}s^{-1}] is the radiative heat gain, Q_{Trans} [J m^{-2}s^{-1}] heat loss due to crop transpiration, Q_{Cond} [J m^{-2}s^{-1}]

is the heat gain due to condensation at the cover, and Q_{Lamp} [J m^{-2}s^{-1}] is the heat gain due to the artificial lighting.

The heat loss due to ventilation Q_{Vent} is given by

$$Q_{\text{Vent}} = k_{\text{v}}(T_{\text{Out}} - T_{\text{Air}}), \tag{A-2}$$

here, k_{v} is the ventilation heat transfer coefficient [W°C^{-1}m^{-2}], T_{Out} is the outside air temperature [°C], and T_{Air} is the greenhouse indoor air temperature [°C].

The ventilation heat transfer coefficient k_{v} is given by

$$k_{\text{v}} = \rho_{\text{a}} c_{\text{p}} \Phi_{\text{v}}, \tag{A-3}$$

where ρ_{a} is the specific mass of air [kg m^{-3}], c_{p} is the air specific heat at constant pressure [J°C^{-1}kg^{-1}], and Φ_{v} is the ventilation flux [m s^{-1}]. The adapted ventilation flux, to include the screens, is given by

$$\Phi_{\text{v}} = \left((1-(1-P_{\text{Scr1}})(100-SO_1))(1-(1-P_{\text{Scr2}})(100-SO_2))\right)\left(\frac{\sigma\phi_{\text{Lee}}}{1+\chi\phi_{\text{Lee}}}+\zeta+\xi\phi_{\text{Wind}}\right)w+\psi, \tag{A-4}$$

where σ [%$^{-1}$], χ [%$^{-1}$], ζ [-], ξ [%$^{-1}$] and ψ [m s^{-1}] are ventilation rate parameters, and ϕ_{Lee} [%] and ϕ_{Wind} [%] are the window openings on the leeward and windward side of the greenhouse.

The adapted conductive heat loss through the cover and screens Q_{CovScr} is given by

$$Q_{\text{CovScr}} = k_{\text{r}} \frac{\left((1-(1-P_{\text{QScr1}})(1-\frac{SO_1}{100}))(1-(1-P_{\text{QScr2}})(1-\frac{SO_2}{100}))\right)}{1+\left((1-(1-P_{\text{QScr1}})(1-\frac{SO_1}{100}))+(1-(1-P_{\text{QScr2}})(1-\frac{SO_2}{100}))\right)}, \tag{A-5}$$

where the greenhouse cover and the screens are modelled as three heat conductors in a series composition [52]. P_{QScr1} [-] and P_{QScr2} [-] are the heat conductances of the temperature (SO_1) and climate (SO_2) screens, respectively.

The heat gain due to heating pipe system 1 Q_{Pipe1} [Wm^{-2}] and the heat gain due to heating pipe system 2 Q_{Pipe2} [Wm^{-2}] are given by

$$Q_{\text{Pipe1}} = \alpha_1(T_{\text{Pipe1}} - T_{\text{Air}}), \quad Q_{\text{Pipe2}} = \alpha_2(T_{\text{Pipe2}} - T_{\text{Air}}), \tag{A-6}$$

where α_i [Wm^{-2} °C^{-1}] is the pipe air heat transfer coefficient for heating pipe system i, T_{Pipe1} [°C] is the temperature of the upper heating pipe system, and T_{Pipe2} [°C] is the temperature of the lower heating pipe system.

The adapted heat transfer coefficient of heating pipe system i to T_{Air} α is given by

$$\alpha_i = \upsilon_i \sqrt{\tau_{\text{Pipe,i}} + \sqrt{|T_{\text{Air}} - T_{\text{Pipe,i}}|}}, \tag{A-7}$$

where υ_i [W m^{-2} °C$^{-3/4}$] and $\tau_{\text{Pipe,i}}$ [°C$^{-1/2}$] are heat transfer coefficient parameters of the two heating pipe systems.

The heat gain due to the solar radiation Q_{Sun} is given by

$$Q_{\text{Sun}} = \eta\left(1-(1-\tau_{\text{Scr1}})\left(1-\frac{SO_1}{100}\right)\right)\left(1-(1-\tau_{\text{Scr2}})\left(1-\frac{SO_2}{100}\right)\right)I_{\text{Glob}}, \tag{A-8}$$

where τ_{Scr1} [-] and τ_{Scr2} [-] are the transmissivities of the blackout and energy screens, and I_{Glob} is the solar global radiation [W m^{-2}].

The heat loss due to the crop transpiration Q_{Trans} is given by

$$Q_{\text{Trans}} = -\lambda E, \tag{A-9}$$

here λ [J g^{-1}] is the water vaporisation energy, and E [g s^{-1} m^{-2}] is the crop transpiration rate.

The water vaporisation energy λ is given by

$$\lambda = l_1 - l_2 T_{\text{Air}}, \tag{A-10}$$

where l_1 [J g^{-1}] and l_2 [J g^{-1}°C^{-1}] are water vaporisation energy coefficients.

The crop transpiration rate E is given by

$$E = \frac{s\eta I_{\text{Glob}} + \rho_{\text{a}} c_{\text{p}} D_{\text{Air}} g_{\text{b}}}{\lambda\left(s + \gamma(1 + \frac{g_{\text{b}}}{g}\right)}, \tag{A-11}$$

where s [kPa°C^{-1}] denotes the slope of the saturated water vapour pressure curve, D_{g} [kPa] denotes the greenhouse air vapour pressure deficit, g_{b} [ms^{-1}] denotes the leaf boundary layer conductance, γ [kPa°C^{-1}] denotes the apparent psychometric constant, and g [ms^{-1}] denotes the leaf conductance.

The slope of saturated water vapour pressure curve s is given by

$$s = s_1 T_{\text{Air}}^2 + s_2 T_{\text{Air}} + s_3, \tag{A-12}$$

where s_1 [kPa °C^{-3}], s_2 [kPa °C^{-2}] and s_3 [kPa °C^{-1}] are coefficients of the saturated water vapour pressure curve.

The greenhouse air vapour pressure deficit D_{g} [kPa] is given by

$$D_{\text{Air}} = p^*_{\text{Air}} - p_{\text{Air}}, \tag{A-13}$$

where p^*_{Air} is the air saturated vapour pressure [kPa] and p_{Air} is the the greenhouse air vapour pressure [kPa]. These are given by

$$p^*_{\text{Air}} = a_1 e^{\frac{a_2 T_{\text{Air}}}{a_3 + T_{\text{Air}}}}, \tag{A-14}$$

$$p_{\text{g}} = \Lambda T_{\text{Air}} A H_{\text{Air}}, \tag{A-15}$$

where a_1 [kPa], a_2 [-] and a_3 [°C] are saturation vapour pressure parameters.

The leaf conductance g [mm s^{-1}] is given by

$$g = g_1\left(1 - g_2 e^{-g_3 I_{\text{Glob}}}\right) e^{-g_4 C_{\text{Air}}}, \tag{A-16}$$

where g_1 [mm s^{-1}], g_2 [-], g_3 [s m^2 μmol^{-1}] and g_4 [m^{3}g^{-1}] are leaf conductance parameters.

The heat gain due to condensation at the greenhouse cover Q_{Cond} is given by

$$Q_{\text{Cond}} = \frac{\lambda}{1 + \epsilon} M_{\text{c}}, \tag{A-17}$$

The condensation mass flow at the greenhouse cover M_c is given by

$$M_C = \begin{cases} m_1 \left|T_g - T_c\right|^{m_2} \left(W_{Air} - W_c^*\right), & \text{if } W_g > W_c^*, \\ 0, & \text{if } W_g \leq W_c^*, \end{cases} \tag{A-18}$$

where m_1 [g s^{-1} m^{-2}] and m_2 [-] are mass transfer parameters, W_{Air} [-] is the greenhouse air humidity ratio at vapour pressure p_{Air}, and W_c^* is the cover humidity ratio at saturated vapour pressure p_{Air}^*. These are given by

$$W_{Air}(p) = \frac{\omega p_{Air}}{p_{atm} - p_{Air}}, \tag{A-19}$$

$$W(p) = \frac{\omega p_{Air}^*}{p_{atm} - p_{Air}^*}, \tag{A-20}$$

where ω [-] is a humidity ratio parameter and p_{atm} is the atmospheric air pressure [kPa].

The heat gain due to the lamps Q_{Lamp} [W m^{-2}] is given by

$$Q_{Lamp} = \eta P_{Eff} \frac{L}{100}, \tag{A-21}$$

where P_{Eff} [Wm^{-2}] is the electric power efficiency of the artificial lighting, and L [%] is the percentage of switched on lights.

Greenhouse air CO_2 concentration rate of change $\dot{C}_{Air}$

The greenhouse air CO_2 concentration change $\dot{C}_{Air}$ is given by

$$\dot{C}_{Air} = \left(\frac{V_{Air}}{A_{Air}}\right)^{-1} \left(\Phi_v (C_{Out} - C_{Air}) + C_{Inj} + R - \mu P\right), \tag{A-22}$$

where $\left(\frac{V_{Air}}{A_{Air}}\right)^{-1}$ is the average greenhouse height.

Greenhouse air absolute humidity rate of change $\dot{AH}_{Air}$

The greenhouse air absolute humidity change $\dot{AH}_{Air}$ is given by

$$\dot{AH}_{Air} = \left(\frac{V_{Air}}{A_{Air}}\right)^{-1} \left(E - \Phi_v (AH_{Air} - AH_{Out}) - M_c\right), \tag{A-23}$$

The continuous-time dynamical model for the greenhouse climate is then given by

$$\dot{\boldsymbol{y}}_g = \boldsymbol{g}_g(\boldsymbol{y}_g, \boldsymbol{y}_c, \boldsymbol{u}_g, \boldsymbol{v}). \tag{A-24}$$

The rationale behind the alterations

Modern greenhouses encompass a wide variety of actuators to control the indoor climate, besides the four control inputs originally included in the model by [18]. The AGC included eight actuators. Thus, to be able to calibrate the model to the AGC data and deploy this data for system identification techniques, the non-linear model needed adaptation.

The rate of change of the greenhouse soil temperature T_{Soil} is omitted since there are no measurements available of this variable, and the mathematical description merely depends on the deep soil temperature T_{Deep} of which also no data is available. The outward heat flux through the soil that is missed by ommitting T_{Soil} as a greenhouse climate model state can be accounted for in the calibration of k_{r}, the conductive heat loss through the cover. The heating pipe temperatures T_{Pipe1} and T_{Pipe2} are directly incorporated in the rate of change of the greenhouse air temperature $\dot{T}_{\mathrm{Air}}$, omitting the heating system model.

Furthermore, Q_{Sun} and Q_{CovScr} are adapted to include the screens, Q_{Pipe2} is added in a similar fashion as Q_{Pipe1} to include the second heating pipe system and Q_{Lamp} is added to include the lamps [52].

Appendix B

Model parameters

B-1 Crop non-linear model calibrated parameters

Symbol	Description	Unit	Nominal value	Calibrated value
η	Radiation conversion factor	-	0.7	0.3345
τ_{Scr1}	Transmittivity of the energy curtain	-	-	0.1053
τ_{Scr2}	Transmittivity of the black-out curtain	-	-	0.3042
b_1	Buffer switching function coefficient	m^2g^{-1}	2.7	0.9851
b_F	Fruit respiration requirement coefficient	-	-	0.1041
d_1	Plant development rate parameter	s^{-1}	$2.1332{\cdot}10^{-7}$	$2{\cdot}10^{-7}$
d_2	Plant development rate parameter	s^{-1}	$2.4664{\cdot}10^{-7}$	$1.9731{\cdot}10^{-8}$
d_3	Plant development rate parameter	°C	20.0	-
d_4	Plant development rate parameter	-	$7.4966{\cdot}10^{-11}$	$1.852{\cdot}10^{-7}$
f	Fruit assimilate requirement quotient	-	1.20	$1.7469{\cdot}10^{-12}$
f_1	Fruit growth rate coefficient	s^{-1}	$8.1019{\cdot}10^{-7}$	$1.8382\ {\cdot}10^{-6}$
f_2	Fruit growth rate coefficient	s^{-1}	$4.6296{\cdot}10^{-6}$	$3.5201\ {\cdot}10^{-7}$
m	Leaf area index (LAI) correction function parameter	-	2.511	0.9587
M_{CO_2}	CO_2 molar mass	kg	0.0044	-
M_{F}	Fruit maintenance respiration coefficient	s^{-1}	$1.157{\cdot}10^{-7}$	$6.7656{\cdot}10^{-6}$
M_{L}	Vegetative maintenance respiration coefficient	s^{-1}	$2.894{\cdot}10^{-7}$	$1.9023{\cdot}10^{-6}$
m_{p}	Watt to μmol conversion factor	μmol J^{-1}	4.57	1.6674
p_1	Net-photosynthesis parameter	g m^{-3}	577	249.2543
p_2	Net-photosynthesis parameter	g $s^{-1}m^{-2}$	221	21.3008
p_{atm}	Atmospheric air pressure	kPa	101.0	-
P_{L}	Light parameter	μmol $J^{-1}m^{-2}$	-	$144.2043{\cdot}10^{4}$
P_{m}	Maximum photosynthesis rate	g $s^{-1}m^{-2}$	$2.2538\ {\cdot}10^{-3}$	277.0579
Q_{G}	Fruit growth rate temperature Q_{10}-value	-	1.0	0.4852
Q_{R}	Maintenance respiration Q_{10}-value	-	2.0	0.1775
R_{G}	Gas constant	J $mol^{-1}K^{-1}$	8.3144	-
T_0	Absolute zero	°C	273.15	-
T_{G}	Growth rate temperature effect reference temperature	°C	20.0	4.6010
v	Vegetative assimilate requirement quotient	-	1.23	0.3334
v_1	Vegetative fruit growth ratio parameter	-	1.3774	2.46
v_2	Vegetative fruit growth ratio parameter	°C	-0.168	$-1{\cdot}10^{-8}$
v_3	Vegetative fruit growth ratio parameter	°C	19.0	-
$\frac{V_{\text{gh}}}{A_{\text{gh}}}$	Average greenhouse height	m	3	-
w_{R}	LAI correction function parameter	g m^{-2}	32.23	7.9711
z	Leaf fraction of vegetative dry weight	-	0.6081	657.5746

B-2 Greenhouse indoor climate non-linear model

Symbol	Description	Unit	Nominal value	Calibrated value
γ	Apparent psychometric constant	kPa $^{\circ}C^{-1}$	0.067	-
ϵ	Inside outside cover heat resistance ratio	-	3.0	
ζ	Ventilation rate parameter	-	$2.7060 \cdot 10^{-5}$	-
Λ	Pressure constant	N m $^{\circ}C^{-1}g^{-1}$	0.46152	-
μ	CO_2CH_2O molar weight fraction	-	1.4467	-
ν	Ventilation rate parameter	$\%^{-1}$	$3.68 \cdot 10^{-5}$	-
η	Radiation conversion factor	-	0.7	0.3345
ξ	Ventilation rate parameter	$\%^{-1}$	$6.3233 \cdot 10^{-5}$	-
ρ_a	Specific mass of air	kg m^{-3}	1.29	-
σ	Ventilation rate parameter	$\%^{-1}$	$7.1708 \cdot 10^{-5}$	-
τ	Pipe air heat transfer coefficient parameter	$^{\circ}C^{-1/2}$	3.0	-
τ_{Scr1}	Transmittivity of the energy curtain	-	-	0.1053
τ_{Scr2}	Transmittivity of the black-out curtain	-	-	0.0.3042
υ	Pipe air heat transfer coefficient parameter	W m^{-2}	0.74783	-
χ	Ventilation rate parameter	$\%^{-1}$	0.0156	-
ψ	Ventilation rate parameter	ms^{-1}	$7.4 \cdot 10^{-5}$	-
ω	Humidity ratio parameter	-	0.622	-
a_1	Saturation vapour pressure parameter	kPa	0.611	-
a_2	Saturation vapour pressure parameter	-	17.27	-
a_3	Saturation vapour pressure parameter	$^{\circ}C$	239	-
C_{Air}	Air specific heat at constant pressure	J $kg^{-1}K^{-1}$	$1 \cdot 10^3$	-
C_p	Specific heat of water at constant pressure	J $^{\circ}C^{-1}m^{-2}$	4180	-
g_1	Leaf conductance parameter	mm s^{-1}	20.3	
g_2	Leaf conductance parameter	-	0.44	
g_3	Leaf conductance parameter	s m^2 μmol^{-1}	$2.5 \cdot 10^{-3}$	
g_4	Leaf conductance parameter	m^3g^{-1}	$3.1 \cdot 10^{-4}$	
g_b	Boundary layer conductance	mm s^{-1}	10	
k_r	Roof heat transfer coefficient	W $^{\circ}C^{-1}$ m^{-2}	7.9	-
l_1	Vaporisation energy coefficient	J g^{-1}	$2.501 \cdot 10^6$	-
l_2	Vaporisation energy coefficient	J g^{-1} $^{\circ}C^{-1}$	$2.381 \cdot 10^3$	-
m_1	Mass transfer parameter	g $s^{-1}m^{-2}$	$1.0183 \cdot 10^{-3}$	-
m_2	Mass transfer parameter	-	0.33	-
M_P	Watt to μmol conversion factor	μmol J^{-1}	4.57	-
P_{Eff}	Light parameter to heat		-	-
P_G	Photosynthetically active radiation (PAR) to global radiation ratio	-	0.475	-
P_{QScr1}	Cover screen heat parameter		-	-
P_{QScr2}	Cover screen heat parameter		-	-
s_1	Saturated water vapour pressure curve slope parameter	kPa $^{\circ}C^{-1}$	$1.8407 \cdot 10^{-4}$	-
s_2	Saturated water vapour pressure curve slope parameter	kPa $^{\circ}C^{-2}$	$9.7838 \cdot 10^{-4}$	-
s_3	Saturated water vapour pressure curve slope parameter	kPa $^{\circ}C^{-3}$	0.051492	-

Bibliography

[1] X.-R. Fan, M.-Z. Kang, E. Heuvelink, P. D. Reffye, and B.-G. Hu, "A knowledge-and-data-driven modeling approach for simulating plant growth: A case study on tomato growth," *Ecological Modelling*, vol. 312, pp. 363–373, 2015.

[2] J. Cheeseman, *Food Security in the Face of Salinity, Drought, Climate Change, and Population Growth*. Elsevier Inc., 2016.

[3] Population Division United Nations Department of Economic and Social Affairs, *World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)*, vol. 12. New York: United Nations, 2019.

[4] R. Pérez-Escamilla, "Food security and the 2015-2030 sustainable development goals: From human to planetary health," *Current Developments in Nutrition*, vol. 1, no. 7, pp. 1–8, 2017.

[5] S. H. Wittwer and N. Castilla, "Protected cultivation of horticultural crops worldwide," *HortTechnology*, vol. 5, no. 1, pp. 6–23, 1995.

[6] C. Stanghellini, "Horticultural Production in Greenhouses: Efficient Use of Water," *International Symposium on Growing Media and Soilless Cultivation*, vol. 1034, pp. 25–32, 2014.

[7] H. F. de Zwart, *Analyzing energy-saving options in greenhouse cultivation using a simulation model*. PhD thesis, Wageningen University of Research, 1996.

[8] L. Graamans, E. Baeza, A. van den Dobbelsteen, I. Tsafaras, and C. Stanghellini, "Plant factories versus greenhouses: Comparison of resource use efficiency," *Agricultural Systems*, vol. 160, no. November 2017, pp. 31–43, 2018.

[9] C. Schouten, "Wijziging Convenant CO2 emissieruimte binnen het CO2-sectorsysteem glastuinbouw voor de periode 2013-2020," *Staatscourant*, vol. 66263, no. 149, pp. 1–5, 2018.

[10] Stuurgroep Glastuinbouw, "Notitie Klimaatakkoord Glastuinbouw," 2018.

[11] B. D. Sparks, "What Is the Current State of Labor in the Greenhouse Industry?," 2018.

[12] C. van Rijswick, "World Vegetable Map 2018: More than Just a Local Affair," *RaboResearch Food & Agribusiness*, no. January, pp. 1–4, 2018.

[13] S. Hemming, F. de Zwart, A. Elings, A. Petropoulou, and I. Righini, "Cherry tomato production in intelligent greenhouses-sensors and ai for control of climate, irrigation, crop yield, and quality," *Sensors (Switzerland)*, vol. 20, no. 22, pp. 1–30, 2020.

[14] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[15] G. van Straten and E. J. van Henten, "Optimal Greenhouse Cultivation Control: Survey and Perspectives," *IFAC Proceedings Volumes*, vol. 43, no. 26, pp. 18–33, 2010.

[16] A. Ramírez, F. Rodríguez, and M. Berenguel, "Calibration and Validation of Complex and Simplified Tomato Growth Models for Control Purposes in the Southeast of Spain," *International Workshop on Models for Plant Growth and Control of Product Quality in Horticultural Production*, vol. 654, pp. 147–154, 2004.

[17] T. Morimoto and Y. Hashimoto, "AI approaches to identification and control of total plant production systems," *Control Engineering Practice*, vol. 8, no. 5, pp. 555–567, 2000.

[18] F. Tap, *Economics-based optimal control of greenhouse tomato crop production*. PhD thesis, Wageningen University of Research, 2000.

[19] I. L. López-Cruz, E. Fitz-Rodríguez, R. Salazar-Moreno, A. Rojano-Aguilar, and M. Kacira, "Development and analysis of dynamical mathematical models of greenhouse climate: A review," *European Journal of Horticultural Science*, vol. 83, no. 5, pp. 269–279, 2018.

[20] T. de Jong, *Natural ventilation of large multi-span greenhouses*. PhD thesis, Wageningen University of Research, 1990.

[21] E. Heuvelink, *Tomato growth and yield: quantitative analysis and synthesis*. PhD thesis, Wageningen University of Research, 1996.

[22] S. Hemming, F. De Zwart, A. Elings, I. Righini, and A. Petropoulou, "Remote control of greenhouse vegetable production with artificial intelligence—greenhouse climate, irrigation, and crop production," *Sensors (Switzerland)*, vol. 19, no. 8, 2019.

[23] C. M. Lewis and D. J. Chambers, "Humidity," *Anaesthesia and Intensive Care Medicine*, vol. 22, no. 1, pp. 54–57, 2020.

[24] J. Penuelas, et. al., "Increasing atmospheric CO2 concentrations correlate with declining nutritional status of European forests," *Communications Biology*, vol. 3, no. 1, pp. 1–11, 2020.

[25] S. R. Adams, K. E. Cockshull, and C. R. Cave, "Effect of temperature on the growth and development of tomato fruits," *Annals of Botany*, vol. 88, no. 5, pp. 869–877, 2001.

[26] E. Heuvelink, F. Kempkes, and L. Marcelis, "Van klimaatregeling naar plantsturing," 2009.

[27] E. M. Nederhoff, *Effects of CO2 concentration on photosynthesis, transpiration and production of greenhouse fruit vegetable crops.* PhD thesis, Wageningen University of Research, 1993.

[28] Y. Bard, *Nonlinear Parameter Estimation.* Academic Press New York and London, 1974.

[29] R. Scattolini, "Architectures for distributed and hierarchical Model Predictive Control - A review," *Journal of Process Control*, vol. 19, no. 5, pp. 723–731, 2009.

[30] P. Tatjewski, "Advanced control and on-line process optimization in multilayer structures," *Annual Reviews in Control*, vol. 32, no. 1, pp. 71–85, 2008.

[31] L. Würth, R. Hannemann, and W. Marquardt, "A two-layer architecture for economically optimal process control and operation," *Journal of Process Control*, vol. 21, no. 3, pp. 311–321, 2011.

[32] B. Picasso, C. Romani, and R. Scattolini, "On the design of hierarchical control systems with MPC," *Proceedings of the European Control Conference*, 2009.

[33] L. Ljung, *System Identification - Theory for the User.* Prentice Hall, 1999.

[34] R. Amrit, J. B. Rawlings, and D. Angeli, "Economic optimization using model predictive control with a terminal cost," *Annual Reviews in Control*, vol. 35, no. 2, pp. 178–186, 2011.

[35] A. Bemporad and M. Morari, "Robust Model Predictive Control: A survey," *Robustness in identification and control*, pp. 207–226, 1999.

[36] M. K. Kozlov, S. P. Tarasov, and L. G. Khachiyan, "The polynomial solvability of convex quadratic programming," *USSR Computational Mathematics and Mathematical Physics*, vol. 20, no. 5, pp. 223–228, 1980.

[37] S. Engell, "Feedback control for optimal process operation," *Journal of Process Control*, vol. 17, no. 3, pp. 203–219, 2007.

[38] M. Ellis, H. Durand, and P. D. Christofides, "A tutorial review of economic model predictive control methods," *Journal of Process Control*, vol. 24, no. 8, pp. 1156–1178, 2014.

[39] X. Blasco, M. Martínez, J. M. Herrero, C. Ramos, and J. Sanchis, "Model-based predictive control of greenhouse climate for reducing energy and water consumption," *Computers and Electronics in Agriculture*, vol. 55, no. 1, pp. 49–70, 2007.

[40] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the DeePC," *18th European Control Conference*, pp. 307–312, 2019.

[41] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, "A note on persistency of excitation," *Systems and Control Letters*, vol. 54, no. 4, pp. 325–329, 2004.

[42] I. Markovsky, J. C. Willems, S. Van Huffel, and B. De Moor, "Exact and Approximate Modeling of Linear Systems: A Behavioral Approach," *Society for Industrial and Applied Mathematics*, 2006.

[43] I. Markovsky and P. Rapisarda, "Data-driven simulation and control," *International Journal of Control*, vol. 81, no. 12, pp. 1946–1959, 2008.

[44] L. G. J. Kerkhof, *Optimal Control of Autonomous Greenhouses - A data-driven approach.* MSc thesis, Delft University of Technology, 2020.

[45] M. M. Peet and G. Welles, "Greenhouse tomato production importance of the industry," *Crop production science in horticulture*, vol. 13, pp. 257–304, 2005.

[46] M. Y. El Ghoumari, D. Megías, J. I. Montero, and J. Serrano, "Model predictive control of greenhouse climatic processes using on-line linearisation," *European Control Conference*, pp. 3452–3457, 2001.

[47] F. Oldewurtel, et. al., "Energy Efficient Building Climate Control using Stochastic Model Predictive Control and Weather Predictions," *Proceedings of the 2010 American control conference*, pp. 5100–5105, 2010.

[48] F. Oldewurtel, C. N. Jones, and M. Morari, "A Tractable Approximation of Chance Constrained Stochastic MPC based on Affine Disturbance Feedback," *47th IEEE conference on decision and control*, pp. 4731–4736, 2008.

[49] F. Borrelli and M. Morari, "Offset free model predictive control," *Proceedings of the IEEE Conference on Decision and Control*, pp. 1245–1250, 2007.

[50] D. Limon, I. Alvarado, T. Alamo, and E. F. Camacho, "MPC for tracking piecewise constant references for constrained linear systems," *Automatica*, vol. 44, no. 9, pp. 2382–2387, 2008.

[51] S. Baros, C.-Y. Chang, G. E. Colón-Reyes, and A. Bernstein, "Online Data-Enabled Predictive Control," *arXiv preprint arXiv:2003.03866*, 2020.

[52] W. J. P. Kuijpers, D. Katzin, S. V. Mourik, D. J. Antunes, S. Hemming, and M. J. G. V. D. Molengraft, "Lighting systems and strategies compared in an optimally controlled greenhouse," *Biosystems Engineering*, vol. 202, pp. 195–216, 2020.

Glossary

List of Acronyms

AGC	Autonomous greenhouse challenge
AI	Artificial intelligence
CT	Continuous-time
DeePC	Data-enabled predictive control
DMC	Dry matter content
DRTO	Dynamic real-time optimisation
DT	Discrete-time
EMPC	Economic model predictive control
HPS	High pressure sodium
I/O	Input-output
I/S/O	input/output/state
LAI	Leaf area index
LED	Light emitting diode
LTI	Linear time-invariant
MPC	Model predictive control
ODeePC	Online data-enabled predictive control
PAR	Photosynthetically active radiation
PEM	Prediction-error method
QP	Quadratic Programming
RMSE	Root mean square error
SDG	Sustainable development goals

Nomenclature

List of Symbols

Variables

$\mathcal{A}_{\boldsymbol{u},\mathrm{l}}$	Lower constraint matrix on inputs Quadratic Programming (QP) form
$\mathcal{A}_{\boldsymbol{u},\mathrm{u}}$	Upper constraint matrix on inputs QP form
$\mathcal{A}_{\boldsymbol{y},\mathrm{l}}$	Lower constraint matrix on outputs QP form
$\mathcal{A}_{\boldsymbol{y},\mathrm{u}}$	Upper constraint matrix on outputs QP form
$\tilde{A}$	Augmented state-space matrix A
A	State-space matrix A
$A_{\boldsymbol{u},\mathrm{l}}$	Lower constraint matrix on inputs
$A_{\boldsymbol{u},\mathrm{u}}$	Upper constraint matrix on inputs
$A_{\boldsymbol{y},\mathrm{l}}$	Lower constraint matrix on outputs
$A_{\boldsymbol{y},\mathrm{u}}$	Upper constraint matrix on outputs
AH_{Air}	Greenhouse indoor absolute humidity
AH_{Out}	Outside air absolute humidity
$\beta^{\mathrm{low}}_{\boldsymbol{u}_{\mathrm{gh,s}}}$	Lower bounds on greenhouse system inputs over entire prediction horizon
$\beta^{\mathrm{low}}_{\boldsymbol{y}_{\mathrm{gh,s}}}$	Lower bounds on greenhouse system outputs over entire prediction horizon
$\beta^{\mathrm{upp}}_{\boldsymbol{u}_{\mathrm{gh,s}}}$	Upper bounds on greenhouse system inputs over entire prediction horizon
$\beta^{\mathrm{upp}}_{\boldsymbol{y}_{\mathrm{gh,s}}}$	Upper bounds on greenhouse system outputs over entire prediction horizon
$\beta_{\boldsymbol{u},\mathrm{l}}$	Lower on inputs in QP form
$\beta_{\boldsymbol{u},\mathrm{u}}$	Upper on inputs in QP form
$\beta_{\boldsymbol{y},\mathrm{l}}$	Lower on outputs in QP form
$\beta_{\boldsymbol{y},\mathrm{u}}$	Upper bounds on outputs in QP form
$\mathcal{A}^{\mathrm{low}}_{\boldsymbol{u}_{\mathrm{gh,s}}}$	Lower constraint matrix for greenhouse system inputs over entire prediction horizon
$\mathcal{A}^{\mathrm{low}}_{\boldsymbol{y}_{\mathrm{gh,s}}}$	Lower constraint matrix for greenhouse system outputs over entire prediction horizon
$\mathcal{A}^{\mathrm{upp}}_{\boldsymbol{u}_{\mathrm{gh,s}}}$	Upper constraint matrix for greenhouse system inputs over entire prediction horizon
$\mathcal{A}^{\mathrm{upp}}_{\boldsymbol{y}_{\mathrm{gh,s}}}$	Upper constraint matrix for greenhouse system outputs over entire prediction horizon
$\mathscr{B}$	System behaviour

$\tilde{B}$	Augmented state-space matrix B
$\tilde{B}_{\boldsymbol{u}}$	Augmented state-space matrix $B_{\boldsymbol{u}}$
$\tilde{B}_{\boldsymbol{v}}$	Augmented state-space matrix $B_{\boldsymbol{v}}$
B	State-space matrix B
$b^{\text{low}}_{\boldsymbol{u}_{\text{gh,s}}}$	Lower bounds on greenhouse system inputs
$b^{\text{low}}_{\boldsymbol{y}_{\text{gh,s}}}$	Lower bounds on greenhouse system outputs
$b^{\text{upp}}_{\boldsymbol{u}_{\text{gh,s}}}$	Upper bounds on greenhouse system inputs
$b^{\text{upp}}_{\boldsymbol{y}_{\text{gh,s}}}$	Upper bounds on greenhouse system outputs
$b_{\boldsymbol{u},\text{l}}$	Lower on inputs
$b_{\boldsymbol{u},\text{u}}$	Upper on inputs
$B_{\boldsymbol{u}}$	State-space matrix $B_{\boldsymbol{u}}$
$B_{\boldsymbol{v}}$	State-space matrix $B_{\boldsymbol{v}}$
$b_{\boldsymbol{y},\text{l}}$	Lower on outputs
$b_{\boldsymbol{y},\text{u}}$	Upper bounds on outputs
$\boldsymbol{c}(\boldsymbol{x}(k))^\top$	Linear terms cost function QP controller
$\tilde{\boldsymbol{c}}(\boldsymbol{x}(k))^\top$	Linear terms cost function QP controller
$\tilde{C}$	Augmented state-space matrix C
$\tilde{C}_{\text{Air}}$	CO_2 concentration in original unit
$\tilde{C}_{\text{Inj}}$	CO_2 injection rate original unit
C	State-space matrix C
C_{Air}	Greenhouse indoor CO_2 concentration
C_{DMC}	Dry matter content (DMC)
C_{F}	Fruit dry weight importance factor
C_{Inj}	CO_2 injection rate
C_{Out}	Outside CO_2 concentration
δ	Forward time shift operator
$\tilde{D}$	Augmented state-space matrix D
D	Crop stage
D	State-space matrix D
$\boldsymbol{f}^f$	System state rate of change function fast subproblem
$\boldsymbol{f}^s$	System state rate of change function slow subproblem
$\boldsymbol{g}^f$	System output rate of change function fast subproblem
$\boldsymbol{g}^s$	System output rate of change function slow subproblem
$\Gamma^{\boldsymbol{u}}_N$	Linear relation matrix from input to output
$\Gamma^{\boldsymbol{v}}_N$	Linear relation matrix from exogenous input to output
$\Gamma^{\boldsymbol{x}}_N$	Linear relation matrix from state to output
$\tilde{\Gamma}_{\boldsymbol{u}}$	Linear relation matrix from input to output augmented system
$\tilde{\Gamma}_{\boldsymbol{v}}$	Linear relation matrix from exogenous input to output augmented system
$\tilde{\Gamma}_{\boldsymbol{x}}$	Linear relation matrix from state to output augmented system
$\boldsymbol{g}_{\text{c}}$	Crop non-linear model function

$\boldsymbol{g}_{\mathrm{g}}$ Greenhouse indoor climate non-linear model function
$\mathcal{H}_L(u)$ Hankel matrix of signal u of order L
$\tilde{H}$ Quadratic terms cost function QP controller
H Quadratic terms cost function QP controller
HD_{Air} Greenhouse indoor air humidity deficit
I_{Glob} Global radiation
$I_{n_a \times n_b}$ Identity matrix with n_a rows and n_b columns
J Cost function
ℓ Stage cost per time instance
ℓ System lag
$\lambda_{\mathrm{V_{c,f}}}$ Regularisation parameter for future weather data
$\lambda_{\mathrm{V_{c,p}}}$ Regularisation parameter for past weather data
$\mathcal{L}$ Class of systems
λ Eigenvectors
$M_{\mathrm{c}}^{\mathrm{f}}$ Crop control model for the fast subproblem
$M_{\mathrm{g}}^{\mathrm{f}}$ Greenhouse indoor climate control model for the fast subproblem
$M_{\mathrm{c}}^{\mathrm{s}}$ Crop control model for the slow subproblem
$M_{\mathrm{g}}^{\mathrm{s}}$ Greenhouse indoor climate control model for the slow subproblem
$MW_{\mathrm{CO_2}}$ Molecular weight CO_2
$P_{\mathrm{c}}^{\mathrm{f}}$ Crop prediction model for the fast subproblem
$P_{\mathrm{g}}^{\mathrm{f}}$ Greenhouse indoor climate prediction model for the fast subproblem
$P_{\mathrm{c}}^{\mathrm{s}}$ Crop prediction model for the slow subproblem
$P_{\mathrm{g}}^{\mathrm{s}}$ Greenhouse indoor climate prediction model for the slow subproblem
m_{B} Assimilate buffer dry weight
$\tilde{m}_{\mathrm{F}}$ Total fruit fresh weight
m_{F} Total fruit dry weight
m_{L} Total leaf dry weight
$\boldsymbol{n}$ Measurement noise
$\boldsymbol{n}$ System order
$N_{\mathrm{h}}^{\mathrm{f}}$ Lower layer prediction horizon
$N_{\mathrm{i}}^{\mathrm{f}}$ Number of implemented lower layer iterations per upper layer iteration
$N_{\mathrm{h}}^{\mathrm{s}}$ Upper layer prediction horizon
$N_{\mathrm{i}}^{\mathrm{s}}$ Number of implemented upper layer iterations
$n_{\boldsymbol{u}}$ Dimension of input vector
$n_{\boldsymbol{v}}$ Dimension of exogenous input vector
$n_{\boldsymbol{x}}$ Dimension of state vector
$n_{\boldsymbol{y}}$ Dimension of output vector
N Prediction horizon
$\mathbf{0}_{n_a \times n_b}$ Zeros matrix with n_a rows and n_b columns
p Parameter vector for the non-linear crop or greenhouse indoor climate model

P_{L}	Lighting system intensity
ϕ_{Lee}	Leeward side window opening
ϕ_{Wind}	Windward side window opening
$\boldsymbol{q}$	Extra error states
$\mathcal{Q}$	Rewritten output cost matrix
Q_{c}	Output cost matrix
$\boldsymbol{r}$	Reference vector
$\mathbb{R}$	Set of real numbers
R	Input cost matrix
$R_{\boldsymbol{u}_{\mathrm{gh,s}}}$	Upper layer control cost matrix
RH_{Air}	Greenhouse indoor air relative humidity
SO_1	Temperature screens opening
SO_2	Climate screens opening
$\bar{T}_{\mathrm{Air}}$	Greenhouse indoor air reference temperature
$\mathcal{T}_{\mathrm{k}}$	Shrinking horizon lengths
$\mathscr{T}$	Toeplitz matrix
τ	Sampling time
τ^{f}	Sampling time fast subproblem
τ^{s}	Sampling time slow subproblem
T_{Air}	Greenhouse indoor air temperature
T_{Out}	Outside air temperature
T_{Upp}	Lower heating pipe system temperature
T_{Upp}	Upper heating pipe system temperature
T_{Ref}	Reference temperature
$\boldsymbol{u}$	Input vector
$\boldsymbol{u}^*$	Optimal control input vector
$\boldsymbol{u}^f$	Input vector fast subproblem
$\boldsymbol{u}^s$	Input vector slow subproblem
$\boldsymbol{u}_{\mathrm{c}}$	Crop subsystem inputs
$\boldsymbol{u}_{\mathrm{gh}}$	Greenhouse system inputs
$\boldsymbol{u}_{\mathrm{g}}$	Crop subsystem inputs
$\boldsymbol{y}_{\mathrm{fut}}$	Future output trajectory
$\mathcal{U}$	Input constraint set
$\tilde{\boldsymbol{u}}$	Augmented input vector
u_{fut}	Future input trajectory
U_{f}	Input data Hankel matrix to predict future output
u_{ini}	Initial input trajectory
U_{p}	Input data Hankel matrix to estimate current state
Y_{f}	Output data Hankel matrix to predict future output
y_{ini}	Initial output trajectory

Y_{p}	Output data Hankel matrix to estimate current state
$\boldsymbol{v}$	Exogenous input vector
$\boldsymbol{v}^f$	Exogenous vector fast subproblem
$\boldsymbol{v}^s$	Exogenous vector slow subproblem
$\boldsymbol{v}_{\mathrm{c}}$	Crop subsystem exogenous outputs
$\boldsymbol{v}_{\mathrm{fut}}$	Future exogenous trajectory
$\boldsymbol{v}_{\mathrm{gh}}$	Greenhouse system exogenous outputs
$\boldsymbol{v}_{\mathrm{g}}$	Crop subsystem exogenous outputs
$\boldsymbol{v}_{\mathrm{ini}}$	Initial exogenous input trajectory
$\hat{\boldsymbol{v}}^{\mathrm{f}}$	Weather prediction fast subproblem
$\hat{\boldsymbol{v}}^{\mathrm{s}}$	Weather prediction slow subproblem
$\in$	In
$\tilde{V}_{\boldsymbol{x}}$	Constant terms cost function QP controller
$V_{\boldsymbol{x}}$	Constant terms cost function QP controller
V_{f}	Exogenous input data Hankel matrix to predict future output
V_{f}	Terminal cost
V_{p}	Exogenous input data Hankel matrix to estimate current state
$\mathbb{W}$	Signal space
w_{Out}	Wind speed
w_h	Relative weight for each output
$\boldsymbol{x}$	State vector
$\boldsymbol{x}^f$	State vector fast subproblem
$\boldsymbol{x}^s$	State vector slow subproblem
$\boldsymbol{x}_{\mathrm{c}}$	Crop subsystem states
$\boldsymbol{x}_{\mathrm{gh}}$	Greenhouse system states
$\boldsymbol{x}_{\mathrm{g}}$	Crop subsystem states
$\hat{\boldsymbol{x}}$	State vector estimate
$\mathcal{X}$	State constraint set
$\tilde{\boldsymbol{x}}$	Augmented state vector
$\boldsymbol{x}_0$	Initial state vector
$\boldsymbol{y}$	Output vector
$\boldsymbol{y}^f$	Output vector fast subproblem
$\boldsymbol{y}^s$	Output vector slow subproblem
$\boldsymbol{y}_{\mathrm{c}}$	Crop subsystem outputs
$\boldsymbol{y}_{\mathrm{gh}}$	Greenhouse system outputs
$\boldsymbol{y}_{\mathrm{g}}$	Crop subsystem outputs
$\hat{y}_{hj}$	Simulated output y_h at time t_i
$\mathcal{Y}$	Output constraint set
y_{hj}	j^{th} measurement replicate of y_h at time t_i
$\mathbb{Z}$	Set of complex numbers