Field-Based Predictive Growth Modeling of Broccoli *(Brassica oleracea var. italica)* within Precision Agriculture

Understanding Field Growth Dynamics for Data-Driven Agriculture

N. Mateijsen

Master Thesis
Computer Science
December 11, 2025

TU Delft

Field-Based Predictive Growth Modeling of Broccoli (Brassica oleracea var. italica) within Precision Agriculture

Understanding Field Growth Dynamics for Data-Driven Agriculture

by

N. Mateijsen

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Thursday December 11, 2025 at 10:00 AM.

Student number: 4803760
Project duration: May 5, 2025 – December 11, 2025
Thesis committee:

Dr. C. C. S. Liem MMus,	TU Delft, supervisor
Dr. J. Sun,	TU Delft
Dr. ir. J. Wildenbeest,	Inholland University of Applied Sciences

TU Delft

Abstract

Broccoli for the fresh market in the Netherlands is still harvested manually, which is labor-intensive and increasingly difficult to sustain as seasonal labor declines. Existing mechanized harvesters cut entire fields at once and cannot account for plant-to-plant variation, leading to substantial losses when heads differ in maturity. These constraints motivate plant-level, data-driven growth modeling to support selective and more efficient harvesting.

This thesis investigates how field-based measurements can be used to model broccoli growth at the level of individual plants. The work addresses four problems: converting raw field video into plant-level growth curves, investigating which environmental parameters best describe the growth, determining whether cumulative temperature or thermal time better represents broccoli development, and evaluating how different growth models capture head diameter growth. A preliminary study using an external dataset creates the methodological foundation by analysing broccoli growth and benchmarking classical and neural models. A field study in a Dutch production environment expands this work through the development of a data-processing pipeline that includes head detection, plant identification, tracking, diameter estimation, and integration with local weather data. Across both studies, thermal time provides a more biologically meaningful predictor of development than cumulative temperature. Model comparison shows that classical parametric models capture general developmental trends, while a multi-layer perceptron achieves the highest predictive accuracy, with a mean absolute error of 0.583 cm, when multiple environmental variables are included. The results demonstrate that field-based predictive modeling can support precision-agriculture applications such as harvest planning and selective mechanized harvesting.

Preface

This thesis is written to conclude my Master's degree in Computer Science at the Delft University of Technology. The project is carried out in collaboration with the Hogeschool Inholland. Over the last months, I discovered that working with real field data and applying data analysis to practical agricultural problems can be both challenging and rewarding. Learning how computer science methods can support growers and help improve real production environments has been an interesting part of this project and has kept me motivated throughout the process.

Writing this thesis has been a lengthy process, and many people helped me move forward when the work became difficult or when I needed a different perspective. I want to thank Cynthia for her guidance and for asking direct questions that pushed me to think more clearly. Her feedback made me rethink many parts of the work and helped shape the final result. I also want to thank Jeroen for his support and for reminding me to look at the practical side of every choice I made. His input helped me keep the project grounded and realistic.

I am grateful to Bram and Surya for their help during the many stages of the project. They answered questions, helped solve problems in the field, and made sure I understood how the measurement setup worked. Their support made the work easier and more manageable. I also want to thank Verdonk for giving me access to their field and for sharing their insights on the practical side of growing broccoli. Their involvement helped connect the research to the real conditions in which broccoli is produced.

Finally, I want to thank my friends and family for their patience and support. They kept me focused and helped me keep perspective while finishing this thesis. Without their cooperation, this research would not have been possible.

Niels Mateijsen
Delft, December 2025

Contents

1 Introduction

1.1. Background and Motivation

Broccoli is widely regarded as a healthy, high-quality vegetable, being rich in fiber, calcium, and vitamin C. In The Netherlands, roughly 2,500 hectares are used for broccoli cultivation and are therefore an important part of the Dutch horticulture. Currently, the harvest of broccoli is still often done by hand. During the Dutch harvesting season (May to November), harvest workers visually assess and select each broccoli head individually, judging its size and maturity. This labor-intensive method involves a high physical burden on the pickers and is becoming increasingly challenging given the declining availability of seasonal laborers.

One promising solution to labor constraints and costs is mechanized harvesting. Single-pass machines for field-grown broccoli already exist, but they do not yet account for individual head-size variation, since all plants are harvested in bulk at the same time, regardless of their individual developmental status [3]. Therefore, a large portion of the harvested broccoli is not suited for commercial use. A study by Walton and Casada showed that in a single cutting, the maximum percentage of harvestable heads depends on the broccoli cultivar and can range from 30-90%, with most being under 50% [36]. This inefficiency highlights a growing interest in precision agriculture–based solutions that enable plant-specific monitoring and decision support, allowing harvesting operations to be guided by actual crop development rather than uniform field averages. However, recent developments have shown the potential for using broccoli harvest machines that cut plants based on individual broccoli head sizes [39].

Moreover, the Dutch Climate Agreement stipulates that by 2030 at least 50% of agricultural land should be managed under precision-agriculture methods [30, p. 143]. Similarly, in Greenports Nederland's Nationale Tuinbouwagenda, innovation is explicitly listed as one of six priority themes [14]. Therefore, these policy developments underline the importance of advancing precision-agriculture technologies to promote more sustainable and efficient agricultural practices in the Netherlands.

As such, there is a strong motivation to move beyond manual labor toward data-driven, plant-level solutions. For broccoli, this implies the need to monitor and model individual head growth, which can then be used for different applications. For example, previous work has shown that drones can monitor individual plant growth and detect issues such as uneven irrigation and necrosis [21]. This allows for individualized care of plants and opens the possibility of (automated) decision-making systems by using visual imagery collected from the drones. Another example is a recent study that demonstrated that recurrent neural network models integrating weather data can predict the optimal harvest date of broccoli with an average error below 2.5 days [25]. This enables growers to schedule harvest timing more precisely and reduce losses caused by harvesting too early or leaving heads in the field past their marketable size.

1.2. Research Questions

This thesis aims to investigate how predictive models based on field measurements can be used to understand the growth dynamics of broccoli. More specifically, the main research question is:

How can predictive models based on field measurements be used to understand the growth

dynamics of broccoli under Dutch growing conditions?

To address this question, it is broken down into four smaller sub-questions:

1. *How can field-based video recordings be converted into plant-level growth trajectories suitable for modeling?*
2. *Which environmental parameters best explain variation in broccoli growth dynamics?*
3. *Considering temperature representation, does cumulative temperature or thermal time better describe the underlying biological growth process?*
4. *How do various machine learning models differ in their ability to represent the growth dynamics of broccoli?*

1.3. Contributions

This thesis contributes to the field of data-driven precision agriculture through a combination of methodological development and applied growth modeling. This work involves the creation, processing, and validation of a new field dataset collected under commercial growing conditions. A core contribution, therefore, lies in the development of an end-to-end automated processing pipeline that transforms overhead RGB video and depth imagery into plant-level broccoli head diameter measurements. This pipeline consists of head detection, plant tracking, diameter estimation, and data filtering, and was designed to operate under realistic field conditions where variability in lighting and camera motion is present.

A contribution is the design and implementation of a novel plant identification and assignment strategy that enables tracking of individual plants across repeated measurement days. This approach estimates plant displacement based on camera movement and local velocity cues and integrates track-repair mechanisms to resolve temporary detection failures. By enabling assignment of measurements to plant ID, this method makes it possible to construct growth trajectories from raw field video.

In addition to the methodological contributions, this work provides quantitative findings on broccoli growth modeling. The thesis evaluates the suitability of cumulative temperature and thermal time as representations of developmental progress and validates whether thermal time is the more biologically meaningful predictor under the studied conditions. Furthermore, the research investigates which environmental parameters contribute most strongly to head diameter development, offering insight into the relative importance of temperature, humidity, soil moisture, solar radiation, and wind speed within a field setting.

Finally, this thesis conducts a comparative assessment of multiple predictive modeling approaches, ranging from parametric growth curves to neural architectures. By benchmarking these models, the study identifies which techniques provide the most accurate and robust predictions of broccoli growth under field conditions. Together, these methodological and analytical contributions demonstrate how automated field sensing can support scalable, plant-level growth monitoring and predictive modeling for future precision agriculture applications.

1.4. Structure

This thesis is embedded in a broader Dutch research project aimed at improving the prediction of broccoli growth and developing a more automated, plant-level measurement pipeline to support precision harvesting as described in Section 1.1. Accurate growth prediction is a central goal, but equally important is understanding what should be measured in the field and how these measurements can be reliably collected and processed to feed predictive models. Building such a pipeline requires both experimentation with machine learning techniques and hands-on experience with real-world data acquisition under commercial field conditions. This work builds on practical measurements and experiences collected directly in Dutch fields, while continuously comparing these findings with insights from existing literature, and in particular with the knowledge obtained from a preliminary study.

With this context, this thesis is organized into several chapters that build toward understanding and modeling broccoli growth. Chapter 2 provides background information on broccoli growth dynamics and relevant environmental factors through a review of related literature, concluding with the hypotheses that guide the research. Chapter 3 outlines the overall approach to growth modeling, describing the

different model types and feature-engineering strategies used to represent environmental influences on crop development. Chapter 4 presents a preliminary study conducted on an external dataset, which served to establish the methodological framework and benchmark model performance. Chapter 5 describes the field-based study carried out in a Dutch production environment, including data collection, preprocessing, and model application. Chapter 6 reports the results and comparative evaluation of the field-based study. Chapter 7 presents the main conclusions of the thesis, and finally Chapter 8 proposes suggestions for future research.

2

Background and Literature Review

This chapter provides the background for the growth modeling work in this thesis. It brings together two complementary perspectives. On one side, existing biological knowledge describes which environmental factors influence broccoli development and why these factors matter for head formation and growth. On the other side, data-driven modeling relies on measurable features that can support prediction. The chapter, therefore, reviews the main growth phases of broccoli, the environmental parameters that are known to affect development, and the modeling approaches that have been proposed in related work. These elements together form the basis for the hypotheses outlined at the end of the chapter and for the feature selection and modeling choices in the remainder of the thesis.

2.1. Growth Phases

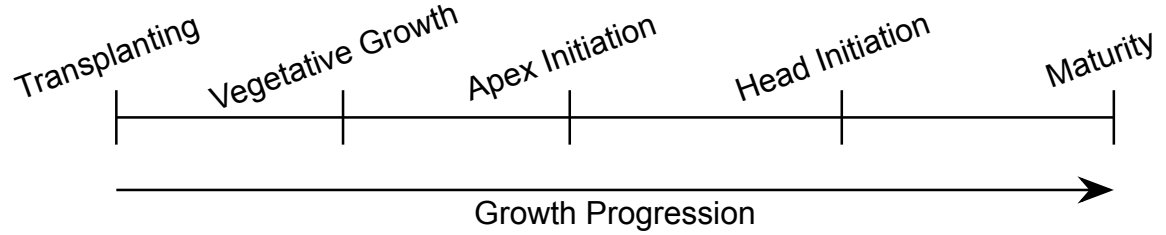


Figure 2.1: Growth stages of broccoli

Broccoli development proceeds through a series of growth stages, summarized in Figure 2.1. Five distinct growth phases are defined in the life cycle of broccoli: Transplanting, Vegetative Growth, Apex Initiation, Head Initiation, and Maturity.

Transplanting marks the beginning of field-based growth. At this point, young seedlings are moved from the nursery to the field. This marks the start of field-based growth.

Following transplanting, the plant enters the **vegetative growth** phase. During this stage, the plant focuses on producing leaves and establishing its canopy. Leaf production continues until the plant reaches a sufficient developmental stage to initiate reproductive growth.

Apex initiation occurs when the shoot apical meristem transitions from producing vegetative organs to reproductive structures. This stage represents the internal onset of head development, although no visible head is yet present.

Head initiation follows, during which the developing floral structures begin to differentiate. At this point, reproductive development becomes visible as the emerging head rises above the canopy.

The final stage is **maturity**, when the broccoli head has fully formed and reached a commercially acceptable size and structure for harvest. This stage concludes the crop cycle and determines the timing of the harvest.

This study focuses exclusively on the development from head initiation to maturity, since this is the period in which the head becomes visible, and its size can be measured reliably. In this study, growth

is defined as the diameter of the head, which cannot be assessed before head initiation, and therefore places the earlier phases outside the scope of the analysis. Although the focus lies on the period from head initiation to maturity, the earlier phases remain part of the same field growth cycle and define the developmental context from which the head emerges.

2.2. Environmental Parameters

Broccoli growth can be influenced by a number of environmental factors throughout the growth cycle. Air temperature influences developmental rate, while humidity affects transpiration and water balance. Soil moisture determines how effectively plants take up water, and solar radiation drives photosynthesis and energy availability. Wind exposure can alter plant structure through mechanical stress. In addition to these external factors, the timing of head initiation plays a critical role in determining when visible head development begins. These factors are relevant both from a biological perspective, where they help explain variation in growth, and from a modeling perspective, where they inform which features may be used in a predictive pipeline. The following subsections describe each of these factors in more detail.

2.2.1. Air Temperature

Air temperature is one of the most influential environmental factors determining the growth, development, yield, and final quality of broccoli heads [19, 34]. As a cool-weather crop, broccoli exhibits optimal growth under relatively mild climatic conditions and is particularly sensitive to temperature fluctuations outside this range. The species thrives between 5°C and 25°C, with an optimal growth window typically reported between 15°C and 23°C [19]. Within this range, physiological processes such as leaf expansion, photosynthesis, and reproductive development occur efficiently, resulting in uniform and compact head formation. Outside of this range, both low and high temperature stress can interfere with developmental processes, leading to growth retardation or morphological abnormalities [34].

To quantify the effect of temperature on plant development, many studies have introduced temperature-based growth models, often expressed as *Thermal Time* or *Growing Degree Days* (GDD). Thermal time models assume that development progresses linearly with accumulated effective temperature, but only within a biologically active temperature range. This range is defined by a lower base temperature (T_{base}) below which growth ceases, and an upper limit (T_{max}) beyond which physiological processes slow or stop. Temperatures outside this range contribute little to growth and are therefore excluded from the calculation. In contrast, a simple cumulative temperature model sums daily mean temperatures irrespective of their biological relevance, potentially overestimating development during periods of heat or cold stress. Thermal time models, therefore, could provide a more biologically grounded and realistic representation of developmental progress.

Reported values for T_{base} in broccoli vary between studies, ranging from 0°C [40, 15] to 3°C [10] and 7°C [12]. Likewise, the optimal temperature for growth has been estimated between 15°C and 17°C, while the maximum temperature threshold is typically found near 26°C [12]. These differences can be attributed to the use of different broccoli cultivars, which are chosen to match specific growing conditions, such as seasonal timing and local climate.

Elevated temperatures, especially those exceeding the optimal range, can deteriorate the quality of broccoli heads and reduce marketable yield [34]. Studies have shown that a rise in average air temperature by 4.4–5.4 °C can lead to a substantial reduction in marketable production, ranging from 42% to 92% due to degradation of head quality [19]. Effects of heat stress on the developing head include the formation of uneven flower bud sizes, the presence of leaves between flower buds, an uneven head surface, and reduced head weight and diameter. These temperature effects motivate the use of temperature-based features, such as cumulative temperature and thermal time, as core predictors in the modeling pipeline.

2.2.2. Humidity

Vapor pressure deficit (VPD) is a key indicator describing how humidity and temperature jointly affect plant physiology. It quantifies the difference between the saturation vapor pressure and the actual vapor pressure of the air and therefore reflects the drying power of the atmosphere. High VPD values indicate dry air with strong evaporative demand, while low values correspond to humid conditions with limited drying capacity. Because VPD is directly linked to transpiration, it strongly influences stomatal

conductance, photosynthesis, and overall plant water balance [16].

When VPD is moderate, plants can maintain open stomata, allowing efficient gas exchange and photosynthetic activity. At high VPD levels, however, excessive transpiration causes stomatal closure to prevent water loss, reducing carbon dioxide uptake and slowing growth. Conversely, very low VPD values result in air that is too humid, lowering transpiration to the point where nutrient transport and leaf cooling become insufficient [16]. General plant research indicates that extremes in VPD hinder growth [22]. Ding et al. confirmed this relationship in tomato cultivars and found that elevated VPD suppressed tomato fruit expansion by intensifying water loss relative to water import, leading to reduced fresh weight [9]. Given these physiological responses, VPD is included as a potential predictive feature to capture periods in which humidity-related stress may influence head development.

2.2.3. Soil Moisture

Previous research has shown that adequate soil water availability during and after head formation is critical for optimal head development [27, 18]. The results of a recent study demonstrated that water deficit conditions decreased the growth, yield, and quality of broccoli, while enhancing water use efficiency [18]. Akter et al. found a correlation between higher volumetric water contents and plant heights [1]. Furthermore, broccoli is a shallow-rooted crop and is particularly sensitive to water stress [18]. Studies also indicate that a dry-wet irrigation regime, where soil moisture is kept low until the start of head formation and increased thereafter, can yield head sizes similar to those achieved under continuous watering [27]. This suggests that the timing of water availability may be as important as total water input, with early deficit followed by adequate supply reducing water use without compromising marketable yield. Because these moisture conditions affect water uptake and growth, soil moisture is used as a predictive variable in the modeling framework.

2.2.4. Solar Radiation

Prior research has shown that intensified radiation, projected to rise by 32–75% under climate change scenarios, contributes to reductions in marketable yield by as much as 42–92% as a result of head quality deterioration under thermal stress [19]. Moreover, irrigation requirements increased by 14–61%, implying higher production costs. Additionally, Lindemann-Zutz et al. developed a process-based growth model that incorporates engineered features to simulate the influence of solar radiation on broccoli head development [23]. Their results demonstrated that radiation-driven variation in head induction timing could explain up to 79% of the observed variability in head size. Similarly, Marshall and Thompson found that although temperature accounted for most of the variation in broccoli development time, solar radiation contributed an additional 17.7% to the variation in the duration to maturity [26]. Since radiation influences energy availability and developmental timing, it is considered a predictive feature.

2.2.5. Wind Speed

Wind speed will be considered as a potential factor influencing broccoli morphology and growth. Increased wind exposure has been shown to reduce leaf area and promote the development of shorter, sturdier stems. These morphological changes are commonly observed in plants subjected to mechanical stress from wind [17, 38, 20]. However, studies applying mechanical stress factors, including simulated wind, reported no significant effect on total yield [20].

2.2.6. Head Initiation

Cultivar variation and temperature thresholds play a critical role in determining the timing and success of head initiation in broccoli. Studies consistently report that excessive temperatures can delay or prevent the transition from vegetative to reproductive development. For instance, Siomos et al. observed that the number of leaves required before head initiation increased with temperature, from 16.7 leaves at 13°C to 26.6 leaves at 30°C, indicating a delay in reproductive transition under warmer conditions [34]. Similarly, Lindemann-Zutz et al. reported a complete absence of head formation in the 'Ironman' cultivar at 25°C and above [23]. These findings align with broader evidence that head formation typically does not occur above 23–24°C, though some studies report failures only at temperatures exceeding 30°C [34]. Fellows et al. further demonstrated that in calabrese, apex initiation does not occur at extreme temperatures, failing both at 0°C and at 30–35°C [13].

Differences among cultivars further complicate temperature responses. Wurr et al. and Grevsen demonstrated that 'Shogun' exhibits delayed and abnormal head development above 20°C, whereas 'Caravel' and 'Emperor' show greater tolerance to elevated temperatures [41, 15]. This highlights the importance of cultivar-specific responses when predicting developmental outcomes under variable thermal regimes.

Different thermal time-based models have been used to model the required time until the head initiation phase. Temperature thresholds for these models are generally reported within the following ranges: base temperatures between 0 and 9.9°C, optimal temperatures between 10 and 21°C, and upper thresholds ranging from 18 to 35°C, depending on the cultivar [34]. Wurr et al. found that apex diameter growth peaks at around 15.6°C, with development largely ceasing below -2.8°C and above 23.6°C [41]. Furthermore, De Maria and Mourão estimated a base temperature of approximately 0.7°C and a thermal time requirement of roughly 680°C-days for head initiation in calabrese [8], a value close to those reported by Tan et al. for 'Fiesta' (670°Cd), 'Greenbelt' (612°Cd), and 'Marathon' (627°Cd) [35]. However, large within-cultivar variation for the 'Marathon' genotype limited the predictive accuracy of the thermal time model in some cases. Cammarano et al. further demonstrated that required thermal time increases for later transplanting dates, as crops develop more slowly under suboptimal temperatures later in the season [5].

Chilling treatments can accelerate head development. Miller et al. found that chilling 14-day-old seedlings under a diurnal cycle of 2°C to 14°C for 28 days advanced flowering by 16 days and reduced node number by 7–8 compared to non-chilled controls. In this context, nodes are the points on the stem where leaves or buds are attached. Prolonged chilling beyond 28 days provided no additional benefit, indicating a threshold beyond which vernalization is saturated [28]. This aligns with earlier work showing that low temperatures not only influence the rate of development but also the developmental threshold for transitioning into the head initiation phase.

2.3. Growth modeling

A wide range of modeling approaches has been used to represent plant development. Parametric growth models are commonly applied because they show biologically plausible structure while remaining computationally simple. Logistic functions have been used to model Brassica development rates and maturity timing under varying temperature conditions [40], capturing the characteristic slow–fast–slow progression of head expansion. Quadratic functions provide a more flexible empirical alternative and have been used to approximate growth responses across temperature regimes, although they lack biological constraints and can extrapolate unrealistically outside the observed range [32]. Grevsen introduced a broccoli-specific double-exponential model that better fits the expansion and deceleration of head diameter [15]. These parametric approaches are interpretable and require limited data, but this limits their ability to account for environmental variability or nonlinear interactions.

Neural models offer greater flexibility by learning nonlinear relationships directly from data. Multilayer perceptrons (MLPs) have shown competitive performance in agricultural prediction tasks where environmental variables interact in complex ways [6]. Long Short-Term Memory (LSTM) networks extend this capability by modeling temporal dependencies, allowing them to capture delayed effects of different conditions on growth. Alhnaity et al. demonstrated that LSTM models, which use sequential weather data, can predict tomato plant growth with high accuracy [2]. Neural models, therefore, provide strong predictive power but require more data and hyperparameter tuning, and their lack of interpretability can be a limitation when biological insight is important.

Mechanistic approaches aim to represent physiological processes directly rather than fitting observed growth trajectories. Lindemann-Zutz et al. developed a broccoli model that combines vernalization dynamics with radiation-driven dry-matter production and a model for biomass allocation to the head [24]. The model incorporates dynamic representations of leaf area development and light-use efficiency to simulate intercepted radiation and total dry matter.

Generative modeling represents a further direction in plant-growth prediction. Drees et al. proposed a temporal GAN that forecasts future plant appearances from drone images, using segmentation masks to estimate plant size and structural changes over time [11]. The predictions focus only on these images and do not include any environmental variables.

In this thesis, only a subset of the reviewed models is implemented, as the goal is to predict head diameter from field-based video measurements combined with environmental data. Parametric mod-

els such as the quadratic, logistic, and Grevsen formulations align well with this objective because they produce explicit diameter estimates, require limited inputs, and have been applied successfully in Brassica growth studies. Neural models, specifically MLPs and LSTMs, are included because they can integrate multiple environmental features and capture nonlinear or delayed effects that parametric curves cannot. In contrast, the mechanistic stochastic model by Lindemann-Zutz et al. predicts dry and fresh weight rather than diameter and requires detailed physiological parameters unavailable in a field-video context. Likewise, the temporal GAN approach focuses on synthesising future plant images and does not incorporate environmental variables or provide direct numeric diameter predictions. These methods, therefore, fall outside the scope of the modeling goals in this work.

2.4. Hypotheses

The hypotheses in this study start from the expectation that broccoli development during the head growth phase can be described by the parameters outlined in Section 2.2. These hypotheses serve two complementary roles. From a biological perspective, they summarise established knowledge about how temperature, humidity, soil moisture, solar radiation, wind, and developmental timing are understood to influence growth. From a predictive perspective, they motivate which factors may be relevant as features in the modeling pipeline, even when formal hypothesis testing is not feasible. The field data used in this thesis covers only a limited range of environmental conditions, which constrains the extent to which individual effects can be evaluated directly. The following sections, therefore, present biologically informed expectations that guide feature selection and interpretation, rather than claims that each parameter can be tested with the available dataset.

2.4.1. Air Temperature

It is hypothesized that broccoli growth and head development will be most favorable when air temperatures remain within the optimal range of 15°C to 23°C during the period from head initiation to harvest. Temperatures outside this range, particularly those exceeding the upper threshold of 26°C, are expected to negatively impact both the quality and marketable yield of the crop. Based on literature, it is also hypothesized that broccoli development can be accurately predicted using thermal time models with base temperatures between 0°C and 7°C, though the specific base temperature that best fits observed growth patterns may vary depending on cultivar and local conditions. Furthermore, it is expected that elevated temperatures will correlate with visual and structural defects in head formation, such as uneven bud development, the presence of leaves within the head, and reduced head size.

These expectations reflect established biological knowledge. In the Verdonk dataset, however, temperatures remained within a relatively narrow and mostly favorable range, which limits the ability to evaluate heat stress or to identify optimal and suboptimal temperature thresholds directly. In this thesis, temperature-related hypotheses therefore act as biologically motivated expectations rather than strictly testable claims.

2.4.2. Humidity

It is hypothesized that high VPD values will correspond to reduced stomatal conductance and lower photosynthetic activity, which leads to decreased broccoli growth. Under such conditions, increased transpiration results in greater water loss, while stomatal closure limits the carbon dioxide uptake, which inhibits biomass accumulation. On the other hand, it is also hypothesized that very low VPD values indicate high relative humidity and low evaporative demand that may limit growth by reducing transpiration rates to levels that reduce nutrient transport and leaf cooling. It is therefore expected that the collected data will show optimal broccoli development under intermediate VPD conditions.

While these relationships are supported by plant physiological research, the humidity levels in the Verdonk dataset show limited extremes, which constrain the extent to which reduced growth under very high or very low VPD can be observed. The hypothesis is therefore included as a biological expectation, with the understanding that only modest variation can be explored in the available data.

2.4.3. Soil Moisture

It is hypothesized that higher soil moisture levels will be associated with increased head size. It is further hypothesized that water deficit conditions in the field data will correlate with reduced growth rates, head quality, and lower yields. Given the sensitivity of shallow-rooted broccoli plants to water

stress, even short-term moisture deficits are expected to impact development. As per the water stress levels described by Akter et al., deficit conditions for loam soils are defined by values under 15% volumetric water content [1].

These expectations follow findings from controlled irrigation studies. The field dataset used here, however, does not contain clear water deficit periods or a wide range of moisture levels, meaning that strong correlations between moisture and head size cannot be tested directly. This hypothesis, therefore, provides a biological context, while the data-driven analysis can only explore limited variation in soil moisture.

2.4.4. Solar Radiation

Based on previous work, it is hypothesized that solar radiation contributes to the growth of broccoli heads. While temperature remains the primary driver of head growth, solar radiation appears to play a complementary role in influencing growth dynamics. It is therefore expected that, within non-stressful temperature ranges, increased solar radiation during the head development phase will be associated with a larger head size and potentially a shorter time to maturity.

Since only one growth cycle was measured, the variability in radiation levels remains limited and reflects only the specific timing within that particular season, which restricts the extent to which radiation-related effects can be examined. As a result, this hypothesis serves mainly as a biologically informed expectation rather than a factor that can be examined in depth with the available observations.

2.4.5. Wind Speed

It is hypothesized that increased wind speed will lead to small observable changes in plant architecture, such as reduced leaf area and more compact, structurally reinforced stems. Despite these adaptations, it is further hypothesized that wind speed variation will not significantly affect broccoli yield. The expectation is that the collected data will not show reductions in the growth of the head.

Although wind-related morphological responses are documented, this study focuses on predicting head diameter rather than structural adaptations. Wind speed is therefore included in the dataset as a potential feature, but it is not expected to contribute meaningfully to the final predictive performance.

2.4.6. Head Initiation

It is hypothesized that the timing of head initiation in broccoli, defined as the first visible emergence of the head following vegetative growth, mainly depends on thermal time from the point of transplanting, but this pattern changes depending on the interaction between average temperature and cultivar-specific thermal sensitivity. While previous studies suggest a relatively narrow range of optimal temperatures for head formation, it is expected that, within our dataset, the cultivar used will differ in both the total thermal time required and the upper temperature thresholds beyond which head initiation is delayed or suppressed. Specifically, under moderate conditions (approximately 8–22°C), thermal time models are anticipated to accurately predict head initiation, but under warmer regimes, the rate of development will decline or stall, leading to a divergence from linear thermal accumulation. This is expected to be cultivar dependent, with some showing delayed or failed head initiation at mean daily temperatures exceeding 23°C, while others may sustain development at slightly higher temperatures.

The biological mechanisms behind head initiation are well established, but this study covers only a single growth cycle in one field. Since no variation across fields or seasons is available and head initiation cannot be observed directly, it is included here only as a relevant stage in the full growth cycle and in the timing of harvest, without further analysis in this work.

3

Growth Modeling

This chapter presents the models developed to predict the growth of broccoli. The primary objective of these models is to estimate the diameter of broccoli heads over time, enabling precision agriculture applications such as optimal harvest planning and mechanized harvesting. Several modeling approaches are introduced, ranging from simple linear models to more complex deep learning architectures. Each model is evaluated on its ability to accurately predict growth trajectories under varying environmental conditions. Following the description of the models, the chapter also discusses the approach used to tune hyperparameters for the neural-based models.

3.1. Overview of Growth Models

This section gives a more detailed explanation of the different models used in this study. The subsections describe the individual models, ranging from simple parametric formulations to neural models.

3.1.1. Baseline

The baseline model is defined as a simple linear regression model fitted on the entire training set. It captures the relationship between broccoli head diameter and thermal time using a strictly positive linear function. This model provides a benchmark against which more complex growth models can be evaluated. To construct the baseline model, a standard ordinary least squares regression is applied using thermal time as the independent variable and broccoli diameter as the dependent variable. The resulting linear function takes the following form:

$$\hat{d} = \max(a \cdot T + b, 0) \tag{3.1}$$

where T represents thermal time, a is the slope of the linear regression, and b is the intercept. The maximum operator ensures that the predicted diameter remains non-negative, which is a biologically necessary constraint since the broccoli head diameter cannot fall below zero.

The baseline model is intentionally kept very simple. It makes a strong assumption that broccoli growth progresses linearly with thermal time and does not include any other variables. This simplicity has both advantages and disadvantages. On the one hand, the model is computationally efficient, highly interpretable, and easy to implement. These properties make it useful as an initial approximation and as a reference model against which more sophisticated models can be evaluated. On the other hand, the linear form limits its ability to capture biologically realistic growth patterns.

3.1.2. Quadratic

The quadratic model is a natural extension of the baseline linear model, introducing a non-linear component to capture curvature in the growth trajectory. This model assumes that the relationship between broccoli head diameter and thermal time can be described by a second-degree polynomial. Such a form allows the model to express a broader range of growth dynamics, like an acceleration in growth. The general formulation of the quadratic model is given by:

$$\hat{d} = a + b \cdot T + c \cdot T^2 \tag{3.2}$$

Where T denotes thermal time and a, b, and c the parameters of the model. These parameters are estimated from the training data using non-linear least squares fitting.

The inclusion of a quadratic term enables the model to accommodate both concave and convex growth trends. However, this flexibility also introduces the risk of producing unrealistic extrapolations outside the observed data range. To mitigate such risks, the model is fitted exclusively on the training data, and its predictions are only interpreted within a meaningful thermal time interval. The quadratic model provides a simple yet more expressive alternative to the linear baseline and serves as a useful intermediary between purely linear and more complex non-linear growth formulations.

3.1.3. Logistic

The logistic growth model is a classical representation of bounded biological growth, in which the rate of increase accelerates initially and then slows as the system approaches a limiting maximum. This behavior aligns well with observed patterns in broccoli development, where head diameter increases rapidly during intermediate stages and gradually levels off as maturation is reached.

In this study, two versions of the logistic model are implemented: a one-dimensional model based solely on thermal time, and a two-dimensional model that additionally incorporates soil moisture. The 1D model assumes that growth is exclusively a function of thermal time, while the 2D variant captures the joint influence of temperature and soil moisture.

The 1D logistic model is defined as:

$$\hat{d} = \frac{L}{1 + \exp(-k \cdot (T - T_0))} \tag{3.3}$$

Where T denotes thermal time, L is the asymptotic maximum diameter, k controls the steepness of the curve, and T_0 is the inflection point where the growth rate is maximal. This formulation was fitted to the training data using non-linear regression, and predictions were subsequently made on both training and test sets.

To extend the model's capacity to account for environmental variation, a 2D logistic model was constructed by integrating the soil moisture, which represents volumetric water content measured at a depth of 30cm. This feature was selected as this is believed to be the second most important variable in model growth based on previous work.

The 2D logistic model is given by:

$$\hat{d} = \frac{L}{1 + \exp(-(a \cdot T + b \cdot S + c))} \tag{3.4}$$

Where T is thermal time, S is the soil moisture, L represents the upper asymptote of the logistic function, as in the 1D case, and a, b, and c are model parameters that govern the influence of each input. The model thus incorporates both inputs using a linear combination within the exponent, enabling it to capture interaction effects between temperature and moisture on the growth rate.

3.1.4. Grevsen

The Grevsen model is a biologically inspired non-linear growth model originally derived from the work of Grevsen et al. [15], and adapted in the formulation proposed by Wang et al. [37]. It is specifically designed to represent the asymptotic and decelerating nature of crop development. The model was initially presented in logarithmic form as:

$$\ln(\hat{d}) = a - b \cdot \exp(-c \cdot T) \tag{3.5}$$

Where $\hat{d}$ denotes the estimated head diameter, T thermal time, and a, b, and c the parameters of the model. However, for this research, a simplified variant is used, consistent with the implementation found in [37]. The resulting expression avoids logarithmic transformation and is directly defined as:

$$\hat{d} = a \cdot \exp(-\exp(-b \cdot (T - c))) \tag{3.6}$$

A variant of the model, referred to as the biased Grevsen model, is also used, which introduces an additional bias term, allowing the entire growth curve to shift vertically. This biased form is given by:

$$\hat{d} = a \cdot \exp(-\exp(-b \cdot (T - c))) + d \tag{3.7}$$

These Grevsen models are particularly suitable for modeling the growth of broccoli head diameter due to their capacity to capture the nature of biological growth processes. The double-exponential structure reflects an initial phase of rapid expansion followed by a gradual deceleration as the plant approaches physiological maturity.

Moreover, the model has been shown to generalize reasonably well across seasons. In the study by Wang et al., the model was validated using cross-year predictions, where a model trained on data from one year was used to predict head diameter in another [37]. While early-stage predictions yielded an acceptable correlation with drone-measured values ($r^2 > 0.57$), performance declined at later stages. However, this occurred under conditions where the broccoli heads had already grown beyond market standards and would not be left unharvested in real production scenarios. As such, the reduced accuracy in the late phase does not undermine the model's practical utility. Additionally, Wang et al. also observed that although one-to-one prediction accuracy declined over time, the overall distribution of predicted versus actual head sizes remained similar. This indicates that the model is still useful for approximating the general trends in broccoli development.

3.1.5. Multi-Layer Perceptron

The multi-layer perceptron (MLP) is a feedforward neural network model designed to approximate complex, non-linear relationships between environmental features and broccoli head diameter. In contrast to the previously described parametric models, the MLP is a purely data-driven approach that learns directly from the structure and scale of the input data. This flexibility allows the model to capture subtle interactions and dependencies that may not be easily expressed in closed-form equations. An additional advantage of the MLP is its ability to incorporate a wide range of input features, making it highly adaptable to diverse sources of environmental data.

The MLP architecture consists of an input layer, one or more hidden layers, and a single output neuron corresponding to the predicted diameter $\hat{d}$. Each hidden layer is followed by a rectified linear unit (ReLU) activation function, and dropout regularization is optionally applied to prevent overfitting. The final output layer is a linear transformation that maps the learned feature representation to a continuous scalar output.

The model is trained using mini-batch gradient descent with the Adam optimizer and MAE as the loss function. The input and output features are scaled using the `MinMaxScaler` to ensure numerical stability and to accelerate convergence. The entire training process is performed over 20 epochs, and model performance is evaluated after each epoch on both the training and test sets. The architecture and training parameters used in this model are based on the optimal configuration identified during the hyperparameter tuning procedure described in Section 3.3.

The MLP model represents a flexible and powerful approach to growth modeling. Unlike predefined functional models, it does not impose specific assumptions on the form of the growth curve, making it well-suited for learning from complex datasets. Furthermore, its architecture can easily accommodate multiple input variables, enabling the integration of additional explanatory features such as soil properties or sensor data. However, this comes at the cost of interpretability and requires careful validation to avoid overfitting.

3.1.6. Long Short-Term Memory

The long short-term memory (LSTM) model is a recurrent neural network architecture specifically designed to capture temporal dependencies in sequential data. Unlike feedforward models such as the multi-layer perceptron, which operate on static input vectors, LSTMs process sequences of inputs and are therefore well-suited for time series tasks. In the context of broccoli growth modeling, this capability could be valuable, as plant development is a temporal process influenced by evolving environmental conditions over time, and the LSTM also takes weather data from previous days into account.

The LSTM operates by maintaining an internal memory state that is updated at each time step based on the current input and the previous state. This structure enables the model to retain relevant information from earlier time steps while discarding noise, effectively learning which parts of the sequence are most important for the prediction. The model is composed of a sequence-processing layer followed by a linear transformation that maps the hidden state at the final time step to a single scalar output, representing the estimated head diameter $\hat{d}$.

In this study, the LSTM is trained on fixed-length input sequences of environmental features and past measurements. For each plant, the data is segmented into overlapping sequences. Each input

sequence consists of a temporally ordered set of feature vectors, and the corresponding target is the diameter at the next time step. This structure allows the model to learn how historical environmental conditions relate to subsequent growth outcomes.

The LSTM model is trained using mini-batch gradient descent with the Adam optimizer, and MAE is used as the loss function. Inputs and targets are normalised using the `MinMaxScaler` to ensure numerical stability and faster convergence during training. The model configuration used corresponds to the best-performing parameter combination identified through the hyperparameter tuning procedure described in Section 3.3.

By explicitly modeling sequences, the LSTM can use temporal structure in the data that other models ignore. This includes trends, lags, and accumulations in the feature values that may influence growth in complex and delayed ways. Such dynamics are difficult to represent using static models and highlight the potential of recurrent architectures in agricultural time series modeling. However, this flexibility also introduces increased model complexity, making LSTMs more computationally demanding and harder to interpret compared to simpler alternatives.

3.2. Feature Engineering

In this section, we describe the derived features that are used to model broccoli head growth. These features are constructed from raw weather measurements and diameter data, and are designed to capture biologically relevant factors such as effective thermal exposure and atmospheric moisture demand. In particular, we focus on vapor pressure deficit and thermal time, two well-established indicators of crop development potential. Additionally, we further refine the thermal time feature by introducing an offset that estimates the unobserved early-stage development. All of these features are added to the dataset of both the Tanashi and the Verdonk use cases.

3.2.1. Vapor Pressure Deficit

VPD is an environmental variable that reflects the drying power of the air and is closely linked to plant transpiration and stomatal regulation, as described in Section 2.2.2. High VPD values typically increase transpiration rates, which can accelerate water loss and stress the plant, while low VPD values reduce transpiration but may also limit nutrient transport and cooling.

VPD is calculated using the equation:

$$VPD = e_s \cdot \left(1 - \frac{RH}{100}\right), \tag{3.8}$$

where RH is the relative humidity and e_s is the saturation vapor pressure. For temperatures above 0°C, e_s can be estimated using Tetens' formula [29]:

$$e_s = 0.61078 \cdot \exp\left(\frac{17.27T}{T + 237}\right), \tag{3.9}$$

Where T is the air temperature in degrees Celsius and e_s is expressed in kilopascals.

In this study, VPD is computed daily using the daily mean temperature and relative humidity. The resulting values are then used as environmental features in the modeling process to capture the effect of atmospheric moisture demand on broccoli growth.

3.2.2. Thermal Time

While cumulative temperature is a direct summation of daily mean temperatures, thermal time is a biologically informed measure that has been widely adopted in agronomic literature to model growth, as mentioned in Section 2.2.1. Thermal time accounts for the fact that plant development only proceeds within a biologically active temperature range. Temperatures below a base threshold or above a maximum threshold are not assumed to contribute to physiological development. This allows thermal time to focus on periods when meaningful growth can occur.

Mathematically, the growing degree days (GDD) for day t is defined as:

$$GGD_t = \begin{cases} 0 & \text{if } \bar{T} < T_{base} \\ T_{max} - T_{base} & \text{if } \bar{T} > T_{max} \\ \bar{T} - T_{base} & \text{otherwise} \end{cases}$$

where $\bar{T}$ is the mean temperature, and T_{base} and T_{max} are the base and maximum temperatures, which are parameters chosen for the specific application.

The daily growing degree value depends on how the observed temperature compares to the biological threshold range. Temperatures below the base value contribute nothing, and temperatures above the upper limit are capped. If $\bar{T}$ falls below T_{base}, the daily contribution is zero. If $\bar{T}$ exceeds T_{max}, the contribution is set to $T_{max} - T_{base}$. Otherwise, the contribution is $\bar{T} - T_{base}$. For the preliminary study, the values are set as $T_{base} = 2℃$ and $T_{max} = 26℃$ as described in the literature.

Thermal time TT is then defined as the running sum of daily GDD values:

$$TT_t = \sum_{i=0}^{t} GDD_i \tag{3.10}$$

3.2.3. Thermal Time Offset

One limitation of the datasets is that the broccoli measurements do not start from the moment of head initiation, that is, the point where the head diameter is effectively zero. As a result, a portion of the early growth phase is missing from the data. This means that for all broccoli plants, regardless of whether they initiated head formation early or late, the thermal time values begin accumulating from an arbitrary point in the growth cycle. As such, different individuals can exhibit the same thermal time despite having developed at different rates. This introduces bias in the temporal alignment of growth data and limits the biological interpretability of thermal time as a predictor.

Given that thermal time is the most important feature in our modeling pipeline, increasing its accuracy is especially important. To address this, we introduce a thermal time offset, which estimates the amount of thermal exposure that is unaccounted for before the first measurement. This offset aims to reconstruct the missing portion of each broccoli's development trajectory.

The approach is based on the assumption that average broccoli growth is approximately linear over time when plotted against thermal time. Using this relationship, we can estimate how much thermal time would have been required to reach the observed initial diameter, had measurements begun at head initiation.

To compute the offset, we begin by fitting a simple linear model relating head diameter to thermal time across the full dataset. The model takes the form:

$$\hat{d} = \alpha \cdot T + \beta \tag{3.11}$$

where $\hat{d}$ is the estimated head diameter, T is thermal time, α the global slope and β the bias. Using this global slope, we estimate the thermal time offset for each broccoli i as the amount of thermal exposure required to produce the observed initial diameter:

$$\Delta TT_i = \frac{d_{i,0}}{\alpha}, \tag{3.12}$$

Where $d_{i,0}$ is the initial observed diameter of broccoli i, and ΔTT is the estimated missing thermal time before the first measurement. The adjusted thermal time for broccoli i at time t is then computed as:

$$\tilde{TT}_{i,t} = TT_{i,t} + \Delta TT_i. \tag{3.13}$$

This correction effectively shifts each growth curve backward in thermal time to better reflect its true developmental timeline.

3.3. Hyperparameter Tuning

To ensure optimal predictive performance, hyperparameter tuning was conducted for both the MLP and LSTM models. Given the sensitivity of neural networks to architectural and training choices, systematic exploration of hyperparameter configurations is essential. The goal of the tuning process is to identify the parameter combinations that minimize the MAE on validation data.

For the MLP, the hyperparameters tuned included the number and size of hidden layers, the learning rate, and the dropout rate. A grid search was conducted over a predefined parameter space using 10-fold cross-validation. Each fold involved standardising the input and target features using a `MinMaxScaler`, training the model over 20 epochs, and evaluating performance on the validation fold. The configurations yielding the lowest average validation MAE were selected for final model training. This procedure ensured a robust comparison across different model capacities and regularisation settings.

For the LSTM model, the tuning process also followed a grid search strategy, but with additional attention to the temporal structure of the data. Specifically, a grouped 10-fold cross-validation was applied, where folds were split based on the broccoli plant ID. This approach ensured that all measurements of a given plant were restricted to either the training or validation set within each fold, preventing data leakage. The parameters explored included the sequence length, the number and length of hidden units, the learning rate, and the batch size. For each configuration, the model was trained for 20 epochs, and performance was evaluated using the MAE on the validation set.

In both cases, the parameter combinations were ranked based on the mean and standard deviation of the validation MAE across folds. The configuration with the lowest average validation MAE was used for the final experiments of both studies. This procedure provided a principled and data-driven approach to model selection, ensuring that the reported results were not biased by arbitrary or suboptimal design choices.

3.4. Feature Importance

To gain insight into the contribution of individual input features to model predictions, permutation feature importance analysis was performed for the MLP and LSTM models. These models, while flexible and powerful, are inherently non-transparent: they do not offer a direct or interpretable mapping between inputs and outputs. In contrast to mechanistic or parametric models, neural networks operate as black boxes, making it difficult to understand how specific input variables influence the predicted broccoli head diameter. Feature importance analysis helps overcome this limitation by measuring how much each input contributes to the model's predictive accuracy.

Permutation feature importance evaluates the effect of randomly permuting each input feature on the model's prediction error. For every feature, its values are shuffled across the dataset, thereby breaking any relationship between that feature and the target variable. The model is then used to make predictions on this altered data, and the performance is re-evaluated. If a feature is important for accurate prediction, then disrupting its values should degrade the model's performance. On the other hand, if permuting a feature has little effect, likely, the feature is not strongly used by the model. This procedure is repeated multiple times for each feature to reduce the effect of randomness in the permutations. The average increase in prediction error is reported as a measure of importance, and confidence intervals can be computed to quantify uncertainty.

Applying this method to the MLP and LSTM models allows identification of which features most affect the predicted broccoli diameter. This analysis enhances the interpretability of neural network models and offers valuable guidance for feature selection and understanding the underlying drivers of growth in data-driven modeling.

4

Preliminary Study: Tanashi

For the development and validation of our broccoli growth prediction models, we made use of an external dataset described in the study by Wang et al., which is called 'Tanashi' [37]. This dataset was obtained from multi-year field experiments conducted in Japan and comprises detailed measurements of individual broccoli heads using a combination of destructive and non-destructive techniques. This preliminary study is used to establish a methodological foundation for the growth-modeling work in this thesis. By analyzing the noise characteristics, testing smoothing strategies, and benchmarking several modeling approaches, we can identify which models are most suitable for representing broccoli head-diameter development and which smoothing method is the most suitable. We will use the same models as described in Section 3.1. Based on prior literature, we hypothesize that parametric models will capture broad developmental trends, while neural models will achieve higher predictive accuracy when more environmental variables are included.

4.1. Method

This section outlines the methodological framework used in the preliminary study. First, the datasets, field measurements, and the drone-based imaging pipeline used to construct plant-level growth trajectories are described. Then we define the evaluation criteria applied to quantify predictive performance. Next, we describe the cross-validation and statistical testing procedures used to assess differences between modeling approaches. Finally, we explain the experimental setup used to compare alternative temperature representations under identical modeling conditions. Together, these sections define the complete experimental protocol used to ensure reproducible and consistent evaluation of all models presented in this study.

4.1.1. Data Acquisition

The dataset includes data collected over three consecutive years (2020-2022), capturing both spatial and temporal variations in broccoli development. The data contains repeated measurements of individual plants from the 'Jet Dome' cultivar throughout their growing seasons, with non-destructive measurements, enabling time-series growth tracking, and destructive measurements. The exact number of available sample sizes per year was manually inspected to be:

- **2020:** 480 broccoli heads were measured non-destructively on 4 different dates, and 563 individuals were destructively sampled.
- **2021:** 400 broccoli heads were measured non-destructively on 5 different dates, and 557 individuals were destructively sampled.
- **2022:** 180 broccoli heads were measured non-destructively on 21 different dates. There is no data available on destructive samples.

Destructive measurements were conducted indoors after harvesting the broccoli heads, allowing for more precise measurement of traits such as head diameter and weight. Non-destructive measurements

Figure 4.1: An example image of the Tanashi dataset on which the broccoli diameters are extracted.

were performed directly in the field. These involved image-based assessment of the head diameter, enabling repeated observations of the same plant across multiple dates.

The imaging-based data acquisition pipeline used for non-destructive measurements involved low-altitude drone flights. Aerial surveys were conducted using RGB cameras mounted on commercial drones (e.g., DJI Phantom 4 and Mavic 2 Pro), with image resolutions of 5472 × 3648 pixels. Drone flights were executed at altitudes between 10 and 15 meters. Ground control points (GCPs) and RTK-GNSS were used to achieve georeferencing accuracy.

To extract head diameter information, a multi-stage computer vision pipeline was employed. This consisted of:

1. **Broccoli position detection** using YOLOv5 during the seedling stage, which was feasible due to the high contrast between plants and soil.

2. **Head segmentation** using a pretrained BiSeNet v2 network. The segmentation model was fine-tuned using a limited set of manually annotated training images.

3. **Geometric conversion** from pixel-based segmentations to real-world units (centimeters), using backward-projected spatial transformations derived from photogrammetry.

The drone-based head diameter estimates were validated against field measurements and achieved an RMSE of approximately 7-12 mm between measurement dates, which is comparable to state-of-the-art close-range image-based methods.

4.1.2. Metrics

The predictive performance of all models in this preliminary study is evaluated using the Mean Absolute Error (MAE). MAE measures the average magnitude of the difference between predicted and observed head diameters, expressed in centimeters. Formally, for a set of predictions $\hat{y_0} \dots \hat{y_n}$ and corresponding ground truths $y_0 \dots y_n$:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{4.1}$$

MAE is chosen because it provides a direct and interpretable link between model performance and the practical decisions that depend on accurate diameter estimates. Since broccoli head diameter is measured in centimeters and directly determines harvest timing and marketability, an error metric

expressed in the same units allows the results to be read without further transformation or statistical interpretation. A grower can immediately assess what an average error of, for example, 0.6 cm means in relation to commercial size thresholds and cutting decisions. As a result, MAE aligns the evaluation criterion with both the biological meaning of the target variable and the operational interpretability of prediction accuracy, making it the most suitable metric for analyzing broccoli head-growth models in this preliminary study.

To ensure that the MAE values reported in this study are reproducible and reliable, the metric is always computed under controlled and consistent evaluation conditions. All models are evaluated using the same splits, preprocessing steps, and fixed random seeds, which prevents variation caused by stochastic effects or inconsistent data partitioning. Reliability is further supported by reporting the mean and standard deviation of MAE across folds rather than relying on a single measurement. This approach ensures that the reported performance reflects typical model behavior and is not driven by favorable or unfavorable data splits. As a result, the MAE values presented throughout this study represent stable, repeatable estimates of performance rather than incidental outcomes of individual training runs.

4.1.3. Model Comparison

All models in this study are compared using a 10-fold cross-validation procedure. The dataset is partitioned into ten equally sized folds. In each iteration, nine folds are used for training, and one fold is used for testing. This process rotates until every fold has served once as the test set. This approach ensures that every data point contributes to both training and evaluation, which provides a more robust estimate of model performance than a single train-test split. Because the folds differ in composition, the resulting variation in performance across them reflects how sensitive a model is to sampling fluctuations in the data, offering an indication of the model's reliability. Model performance within each fold is quantified using the MAE. After completing all ten iterations, the average MAE across folds is taken as the model's overall performance indicator, and the standard deviation across folds reflects variability due to differences in data partitions.

To determine whether differences in performance between models are statistically meaningful, pairwise statistical tests are performed for every combination of models. Specifically, we apply paired t-tests to the fold-wise MAE values. The tests are paired because each model is evaluated on the same folds. Every fold yields one MAE value per model, and these values naturally form linked observations. A paired test, therefore, controls for fold-specific effects, such as difficulty or noise in particular subsets of the data. Using an unpaired test would treat the MAE values from different models as independent samples, which would ignore this natural correspondence and reduce statistical power. The paired t-test instead isolates performance differences attributable to the models themselves, providing a more sensitive and appropriate comparison in this setting. The statistical significance of the pairwise comparisons will be annotated using the standard star notation (* for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$).

Together, the cross-validation procedure and paired statistical testing offer a consistent and reproducible framework for evaluating and comparing the predictive accuracy of the growth models in this study.

4.1.4. Thermal Time vs Cumulative Temperature

To evaluate whether thermal time or cumulative temperature provides a more effective representation of broccoli development, we conduct a model comparison in which the same models are trained twice: once using thermal time as the main temperature-derived feature, and once using cumulative temperature. All other inputs, preprocessing steps, and cross-validation folds remain identical between the two runs. This ensures that any performance differences can be attributed specifically to the temperature representation rather than to differences in data handling or model configuration.

The evaluation focuses on two aspects. First, we compare the average MAE across folds between the thermal time and cumulative temperature versions of each model. This indicates which representation yields more accurate predictions. Second, we assess whether these differences are statistically significant by applying paired tests on the fold-wise MAE values.

This setup allows us to determine not only which representation performs better, but also whether the advantage is large enough to be considered meaningful rather than coincidental.

4.2. Utility of the Dataset

The dataset offers several aspects that make it a valuable resource for modeling broccoli growth. First, the dataset contains a good number of broccoli across all years, which ensures enough data points for evaluation. Furthermore, it includes measurements of head diameter, which is the primary feature of interest for growth prediction and harvest optimization. Next to this, the same individual broccoli plants were measured repeatedly over the course of the entire growth cycle. This temporal consistency enables us to create time-series models that capture the plant growth dynamics, which is essential for predicting harvest-ready size based on weather. Moreover, the precise location of the experimental field is described in the original publication. This allows for the retrieval of weather data from external sources, such as weather APIs, which we can use in our growth models. In this study, the Open-Meteo weather API was used to retrieve the weather variables like hourly temperature, humidity, and radiation data from the exact coordinates of the experimental site.

However, the data collected in 2020 and 2021 suffer from a limited number of measurement dates (four and five, respectively), which are deemed insufficient to model continuous growth curves. As such, these two subsets do not meet the temporal requirements necessary for our modeling objectives. For this reason, we will exclusively use the 2022 portion of the dataset, which comprises 180 broccoli plants with 21 diameter measurements each. This data contains enough measurement dates to model, validate, and evaluate the different models.

4.3. Diameter Smoothing

In order to model broccoli growth over time, accurate measurements of head diameter are essential. However, initial inspection of the dataset revealed the presence of noise in the diameter measurements. Given that broccoli head growth is a monotonic biological process during the main development phase, decreases in diameter across measurement dates are not physiologically plausible. This requires a detailed review of the raw data and the use of smoothing methods to reduce measurement artifacts. Each broccoli head in the dataset is measured from four angles, and the maximum of these four values is taken as the final diameter, following the methodology described in the original dataset publication. To illustrate the nature of the noise, Figure 4.2 shows the diameter measurements of a single broccoli, plotted separately for each angle. The values sometimes decrease, as seen, for example, on the 24th.

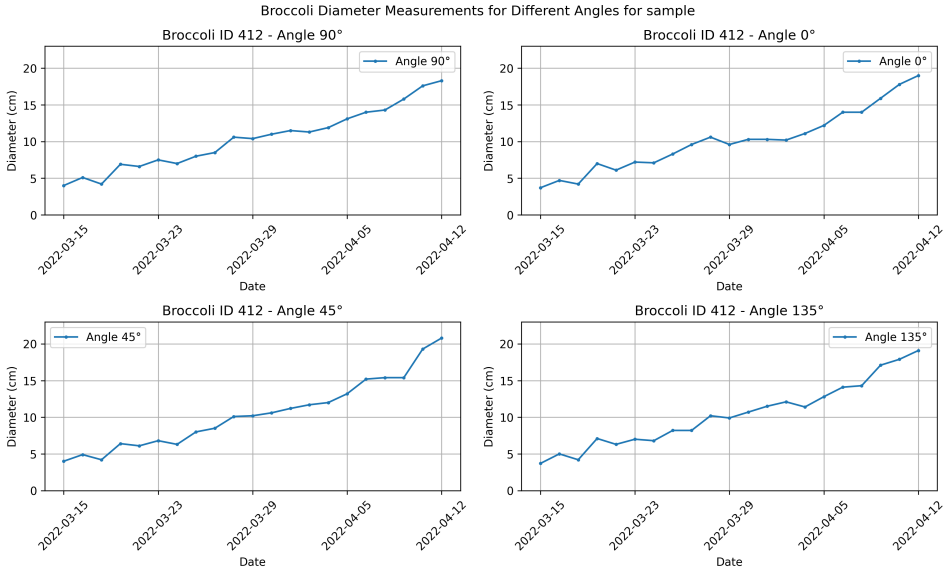


Figure 4.2: Diameter over time for a single broccoli measured from four angles.

Even after aggregating the values across angles using common statistics like minimum, mean, and maximum, such fluctuations persist. This indicates that the noise is not specific to individual angles but affects the measurement process more broadly.

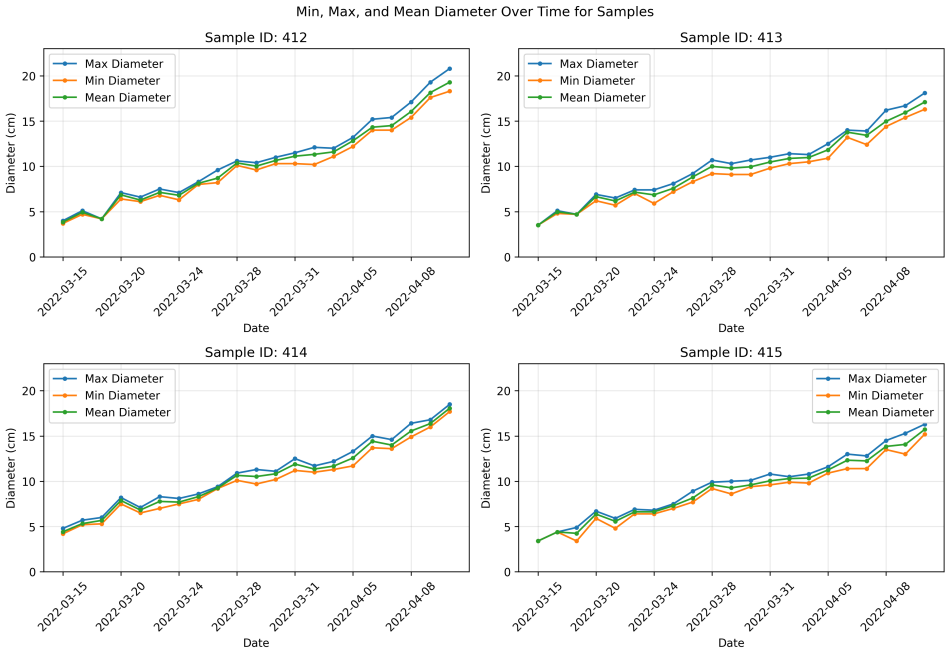


Figure 4.3: Diameter values for four broccoli over time, showing the minimum, mean, and maximum of the four angles.

To quantify the extent of these implausible fluctuations, we define a *negative dip* as any instance where a diameter measurement decreases compared to the previous time point. This metric was computed per angle across all broccoli. The goal is to understand whether noise is angle-dependent or systematic across all viewpoints.

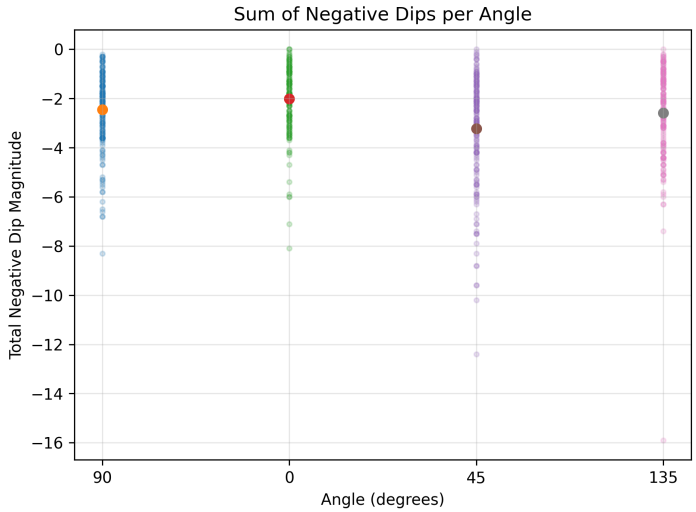


Figure 4.4: Sum of the magnitude of all negative dips per angle, across all broccoli. Each dot represents one broccoli; the mean per angle is shown.

Figure 4.4 shows that substantial dips are present in all four angles. This results in a cumulative negative dip value with a mean between -2cm and -4cm for a single broccoli growth cycle. This value can reach up to a magnitude of 16cm as seen in the 135° data. The growth period for this broccoli can be seen in Figure 4.5. The distributions of the negative dips are similar across all angles, which indicates that the number of dips is not dependent on the angle from which the diameter was measured.

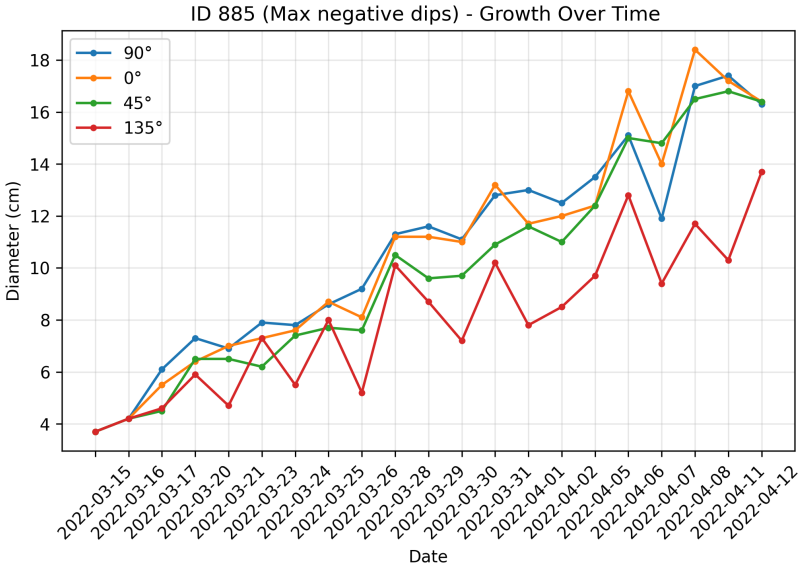


Figure 4.5: Diameter values for the broccoli with the highest cumulative negative dips, shown per angle.

To investigate whether the noise is tied to particular stages in the growth cycle, we aggregated the count of negative dips per measurement date. Figure 4.6 shows the distribution of the number of negative dips per day. This analysis was performed both per angle and on the maximum diameter. For both configurations the distributions are nearly identical the negative dip counts, which confirms that the noise is not angle-specific and persists even by taking the maximum diameter value over all angles.

Furthermore, the counts are not uniformly distributed over all the dates. Instead, there are certain dates with substantially more dips than others. This suggests that the measurement noise may not be random but could be the result of external conditions or procedural inconsistencies on specific days. There are several potential explanations for this. One possibility is that varying environmental conditions, such as changes in lighting, fog, or rain, affected the quality of the drone images or the segmentation accuracy. Another explanation could be inconsistencies in flight altitude, angle, or camera calibration during data acquisition.

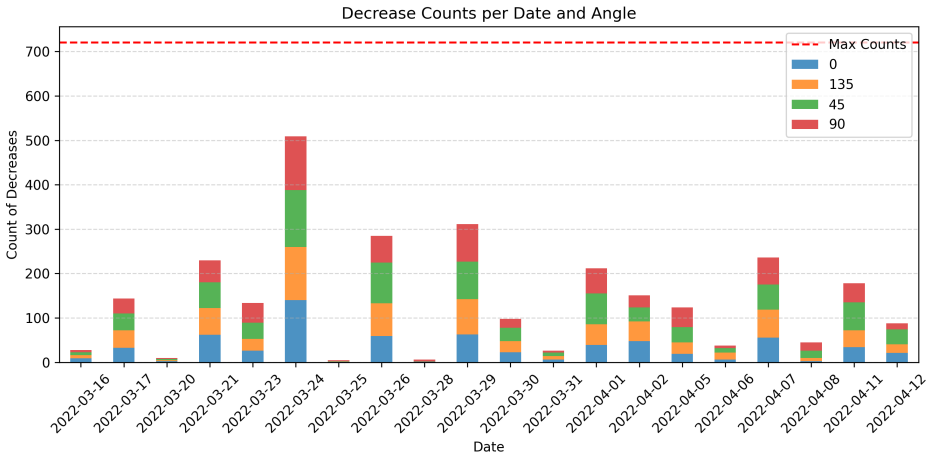


Figure 4.6: Count of negative dips per measurement date, grouped by angle.

Having established that the maximum diameter contains the same noise problem, we focus the rest of our analysis on this metric, as it is the one ultimately used in modeling. We analyzed the distribution of the magnitude of negative dips in the maximum diameter across all broccoli, shown in Figure 4.7. The distribution begins at zero and exhibits a power-law-like shape, with the frequency of dips decreasing

rapidly as the magnitude increases. The largest observed dip reaches a value of 2.4cm, which is considerable given the typical growth increments in broccoli head development.

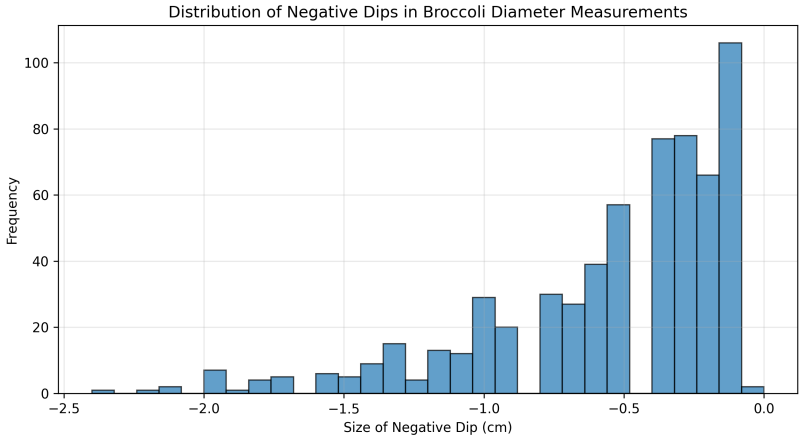
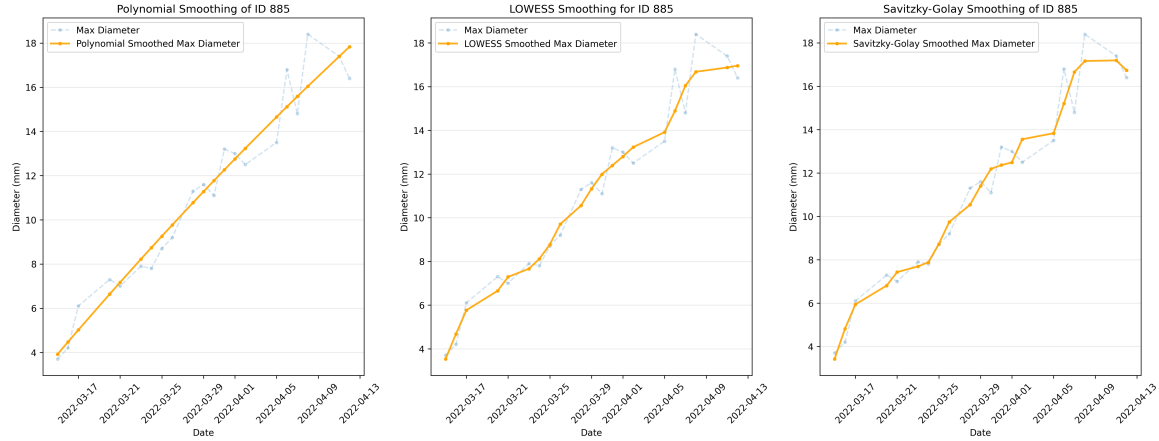


Figure 4.7: Distribution of the magnitude of negative dips for the maximum diameter across all broccoli. Most dips are small, but large deviations occur.

To address this issue, we apply three different smoothing techniques to the max diameter signal: polynomial regression, Locally Weighted Scatterplot Smoothing (LOWESS), and the Savitzky-Golay filter. Each method attempts to fit a plausible growth curve to the noisy data while preserving the overall trend and shape of the true signal.



(a) (b) (c)

Figure 4.8: Comparison of smoothing methods applied to the maximum diameter of the noisiest broccoli: (a) Polynomial regression, (b) LOWESS, and (c) Savitzky-Golay.

The first approach to smoothing the diameter measurements is polynomial regression. The core idea behind this method is to approximate the noisy diameter trajectory of each broccoli with a continuous, smooth function. Specifically, a second-degree polynomial is fitted to the data. This type of function can capture general upward or downward trends with some degree of curvature, while avoiding abrupt local fluctuations and ensuring monotonicity. Figure 4.8a illustrates the result of polynomial smoothing on the broccoli sample with the largest cumulative negative dips. The raw data, shown as a dotted line, shows large fluctuations and multiple decreases in diameter. In contrast, the fitted polynomial curve is nearly linear, suggesting that the quadratic term has very limited influence on the overall fit. This indicates that while the polynomial has effectively removed the implausible dips, it has also eliminated the finer-grained structure of the original growth pattern. As such, while this method provides a smooth approximation, it may oversimplify the underlying growth dynamics.

The second smoothing method we apply is locally weighted scatterplot smoothing (LOWESS). Unlike global approaches such as polynomial regression, LOWESS operates by fitting simple models,

typically low-degree polynomials, locally around each point in the data. Each fitted value is obtained by weighting nearby observations based on their distance to the target point, giving more influence to points closer to the target and less to those farther away. In our implementation, we use a fraction of 0.3 of the data when estimating each diameter, which provides a balance between smoothness and local adaptability. This results in a smooth curve that adapts to local variations while still reducing high-frequency noise. LOWESS is particularly well-suited for data where the underlying trend is expected to be smooth but not necessarily captured by a single global function. In the context of broccoli growth, this allows us to retain moderate local variations that may reflect meaningful differences in growth rate, while filtering out unlikely short-term decreases in head diameter. Figure 4.8b shows the result of applying LOWESS smoothing to the broccoli with the largest cumulative negative dips. Compared to the raw measurements, the smoothed curve follows the overall increasing trend while still capturing some subtle curvature present in the original data.

The third smoothing method we consider is the Savitzky-Golay filter. This technique smooths a signal by applying a moving window across the data and, within each window, fitting a low-degree polynomial using least-squares regression. The fitted polynomial is then evaluated at the central point of the window, and this value is used as the smoothed output. This process is repeated for every point in the series, producing a smoothed signal that locally approximates the data while enforcing a consistent polynomial structure. Figure 4.8c shows the result of applying Savitzky-Golay smoothing to the broccoli with the largest cumulative negative dips using a window length of 7 and a second-degree polynomial. Again, the smoothed curve preserves the general upward trend while still reflecting some of the subtle patterns in the original data.

Figure 4.9 presents a comparison of all three smoothing methods applied to the broccoli with the highest cumulative negative dips. The polynomial fit results in a nearly linear curve and does not capture any of the localized structure present in the raw data. Both LOWESS and Savitzky-Golay, on the other hand, produce curves that closely follow the local trends in the data, preserving subtle changes in the growth trajectory. One notable observation is that the Savitzky-Golay method shows a slight decrease in diameter at the final time step, which is not consistent with expected growth behavior. This behavior is likely a consequence of edge artifacts, which are known to occur for the Savitzky-Golay method due to reduced support at the boundaries [33]. In contrast, the LOWESS method provides a similarly smooth and locally responsive curve but does not exhibit such boundary effects.

Given this observation, we will not use the Savitzky-Golay smoothing in the remainder of this study. Since its behavior is highly similar to that of LOWESS, we select it as the preferred method for smoothing diameter measurements before model fitting. However, we will also include the polynomial smoothing in our model evaluation to compare its impact on predictive performance relative to LOWESS.

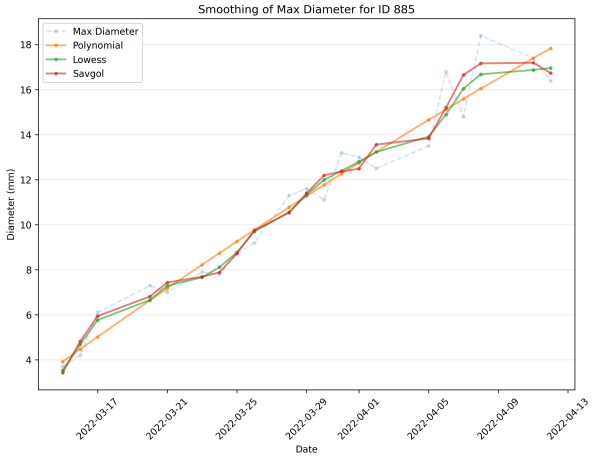


Figure 4.9: Comparison of all three smoothing methods applied to the maximum diameter of the same broccoli.

4.4. Results

This section presents the results of the preliminary study. It begins with the outcomes of the addition of an offset to the cumulative temperature representation. Afterwards, the results of the hyperparam-

eter tuning procedure for the neural models are examined, followed by an overall comparison of all model classes evaluated on the Tanashi dataset. The results then examine how the choice of temperature representation influences predictive accuracy by comparing models trained on thermal time and cumulative temperature. Finally, the section contains results for the feature importance analysis for the neural models to provide insight into which environmental variables contribute most strongly to the predictions. These findings are revisited under real field conditions in Section 6.3.

4.4.1. Thermal Time Offset

Figure 4.10 compares broccoli head diameter plotted against the original thermal time and the adjusted thermal time after applying the offset for the Tanashi dataset. In the uncorrected representation, many data points align in dense vertical bands. This indicates that large numbers of plants share identical thermal time values while being at clearly different developmental stages. This effect results from all plants accumulating thermal time from the same reference point despite having initiated head growth at different moments.

After applying the thermal time offset, the scatter becomes more evenly distributed along the thermal time axis. Individual growth trajectories show improved separation, producing a pattern that better reflects biological variability in developmental timing between plants. The vertical dispersion at given thermal time values is reduced, indicating a clearer mapping between accumulated thermal exposure and observed head size.

This result shows that the offset improves the temporal alignment of individual growth curves by correcting for unobserved early development. As a consequence, the adjusted thermal time representation produces a more realistic depiction of broccoli growth dynamics and provides a better foundation for model training and evaluation.

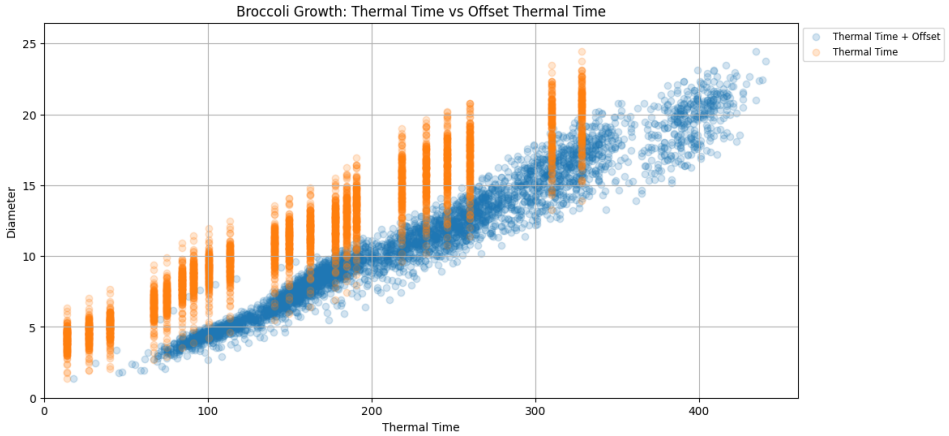


Figure 4.10: Diameter plotted against original thermal time and adjusted thermal time with offset of the Tanashi dataset.

4.4.2. Hyperparameter Tuning Results

For the MLP model, the tuning results indicate that higher dropout rates consistently led to worse predictive performance. The learning rate and batch size had a minimal effect on performance. The values explored in the grid search all ensured that the training loss plateaued within the fixed number of epochs used, meaning that these parameters were not a limiting factor for convergence. The final selected configuration for the MLP used a learning rate of 0.0005, a batch size of 64, a dropout rate of 0.0, and 2 hidden layers with 128 and 64 neurons, respectively.

For the LSTM model, the tuning results reveal a clear negative correlation between sequence length and performance. Shorter sequence lengths consistently resulted in lower validation MAE values, with the best results obtained when the sequence length was at its minimum value. This outcome suggests that broccoli growth is not strongly dependent on previous weather and that incorporating multiple previous days' data introduces additional noise rather than providing useful predictive information. As such, it is hypothesized that an MLP is better suited for this task than a recurrent architecture, as it relies only on the current set of input features without attempting to model temporal dependencies, which appear to be irrelevant for accurate growth prediction. As with the MLP, the learning rate and

batch size had minimal influence on performance, with all tested values ensuring convergence within the set number of epochs. The final selected configuration for the LSTM used a learning rate of 0.001, a batch size of 32, a hidden state size of 64, and a sequence length of 2.

In conclusion, the hyperparameter tuning process confirmed that architectural choices have a large impact on performance. For the MLP, limiting or excluding dropout improved accuracy, while for the LSTM, minimizing sequence length was critical for optimal results. These findings reinforce the idea that broccoli growth modeling benefits more from static feature-based prediction than from sequential temporal modeling.

4.4.3. Overall Comparison Across Models

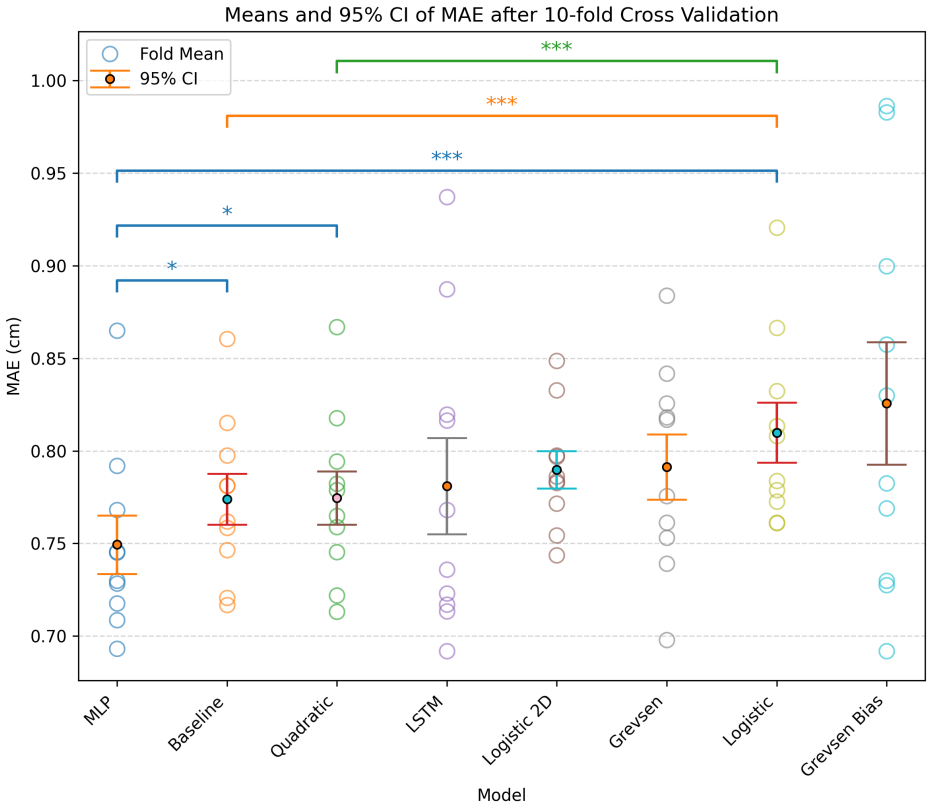


Figure 4.11: Fold-level MAE results for each model on the Tanashi dataset using LOWESS smoothing, including mean values with 95% confidence intervals and significance annotations from paired t-tests. Significance levels are denoted by conventional asterisks: * for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$. The MLP model shows the lowest error across all models and is significantly better than some models.

Model	MAE (Mean)	MAE (Std)
MLP	**0.749**	0.050
Baseline	0.774	0.044
Quadratic	0.774	0.045
LSTM	0.781	0.082
Logistic 2D	0.790	**0.032**
Grevsen	0.791	0.055
Logistic 1D	0.810	0.051
Grevsen Bias	0.826	0.105

Table 4.1: MAE performance comparison across models. The MLP model has the lowest MAE and the Logistic 2D model has the lowest variance.

The results of the 10-fold cross-validation on the external dataset reveal clear differences in predictive performance among the evaluated models, as measured by the MAE. Figure 4.11 presents the fold-level MAE distributions for all models, along with the mean and associated 95% confidence intervals. Table 4.1 shows for each model its mean MAE and the standard deviation. Finally, Figure 4.12 shows the predictions of all models on an example broccoli growth trajectory.

The MLP achieves the lowest overall MAE of $0.749(\pm 0.05)$ with low variability across folds, indicating both high predictive accuracy and stability. The performance advantage of the MLP over many other models is supported by the paired t-tests, with statistically significant improvements over the *baseline*, *quadratic*, and *logistic* models. Its strength lies in the ability to incorporate multiple weather-related features and base its predictions solely on the current weather state, without relying on temporal sequences. This allows the MLP to capture complex non-linear growth patterns that extend beyond the influence of temperature alone, resulting in a robust and accurate predictive model.

The *baseline* and *quadratic* thermal-time models achieve very similar performance, with MAE values of approximately $0.774(\pm 0.044)$, and both significantly outperform the *logistic* model. The results of the *baseline* and *quadratic* models are almost identical. This is explainable due to the fact that, although the quadratic model allows for a quadratic term, in practice this fitted quadratic coefficient has little influence, resulting in a shape that is very close to linear. Consequently, both models behave similarly, with the *baseline* representing a straight line between the average starting and ending points, and the quadratic model approximating this same linear trend.

The *Grevsen* and *Grevsen Bias* models perform less competitively in this dataset, with MAE values of $0.791(\pm 0.051)$ and $0.826(\pm 0.105)$ respectively. While the *Grevsen* model remains reasonably close to the best-performing temperature-only models, the *Grevsen Bias* variant suffers from higher error and higher variability across folds, indicating that the additional bias term does not yield a predictive advantage in this scenario.

The *logistic 2D* model, which incorporates soil moisture alongside thermal time, achieves an MAE of $0.790(\pm 0.032)$, performing similarly but more stably to the *Grevsen* model and better than the *logistic* model in its single-feature form. The slight improvement over the 1D logistic confirms the added value of including a second environmental variable. However, its lack of statistical superiority over the MLP and LSTM suggests that while additional features can be beneficial, model flexibility and non-linear learning capacity are still important determinants of performance.

The LSTM achieves an MAE of $0.781(\pm 0.082)$, but shows relatively high variance across folds. This is likely due to its sequence-based nature, which introduces additional noise rather than providing useful temporal information. The absence of statistically significant differences between the LSTM and other models indicates that its recurrent structure does not provide a clear advantage in this application. The observed negative correlation between sequence length and performance suggests that temporal dependencies beyond the current weather contribute little to improving predictive accuracy for broccoli growth.

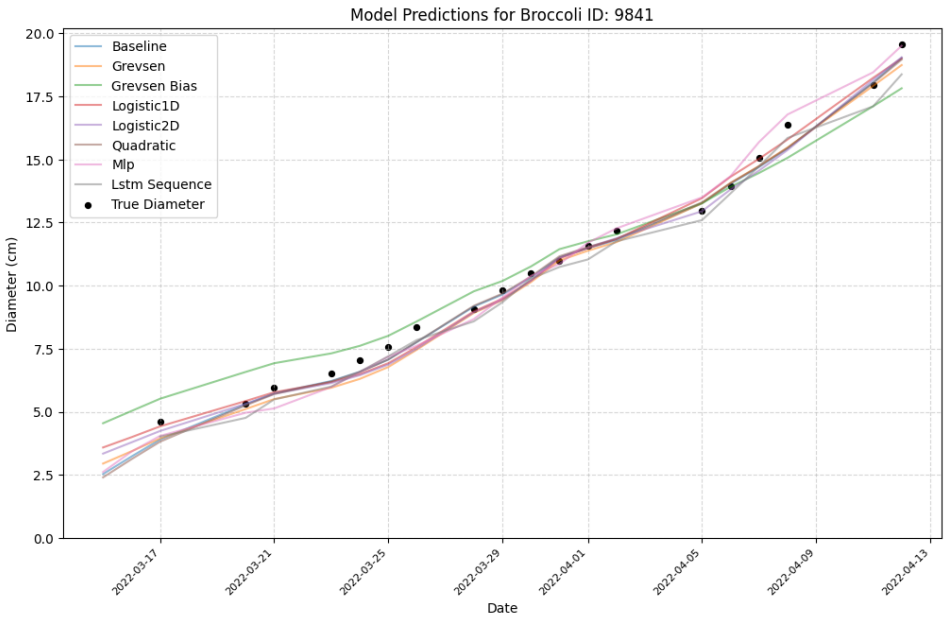


Figure 4.12: An example broccoli sample with predicted values from all models. Most models follow the trend of the example broccoli growth well.

4.4.4. Thermal Time vs. Cumulative Temperature

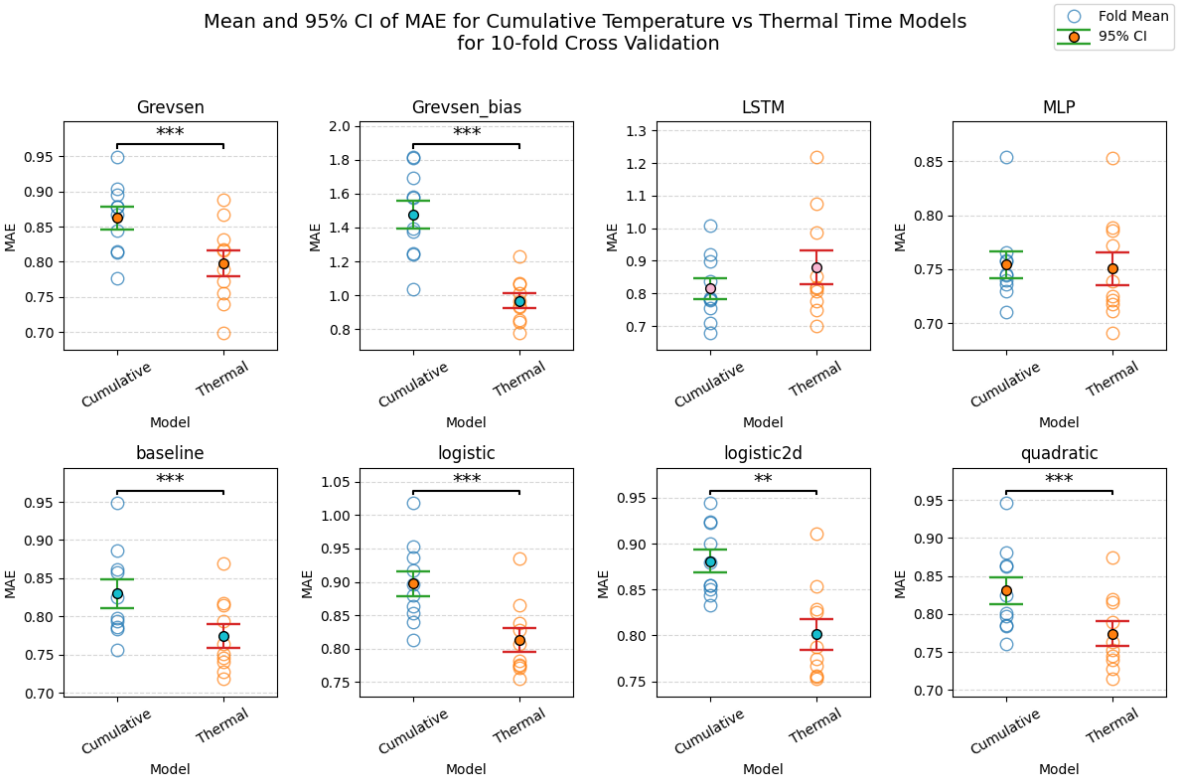


Figure 4.13: Cross-validated MAE for models using cumulative temperature vs. thermal time. Each subplot shows the fold-level MAE, mean with 95% confidence interval, and significance annotation. Significance levels are denoted by conventional asterisks: * for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$. All models, except the neural models, show significantly better results when using Thermal Time.

The fold-level MAE distributions are visualized in Figure 4.13. For each model, a pair of vertical scatter plots is shown, one for the cumulative temperature configuration and one for the thermal time configuration. The plots indicate individual fold MAE scores, the mean MAE with 95% confidence intervals, and the statistical significance of the difference where applicable.

These findings demonstrate that thermal time is a more biologically grounded and effective temporal feature for predictive modeling of broccoli growth. Across the eight evaluated models, six show statistically significant improvements when thermal time is used instead of cumulative temperature, with most achieving highly significant results ($p < 0.001$). This indicates that accounting for biologically active temperature ranges, by excluding ineffective extremes, leads to more accurate growth predictions.

The only models for which the difference is not statistically significant are the MLP and LSTM. A plausible explanation for this result is that both models incorporate multiple additional environmental features beyond temperature alone. As a consequence, the relative influence of the temporal predictor (thermal time versus cumulative temperature) is reduced. A similar pattern is observed in the 2D logistic model, which includes soil moisture as an additional input. Although this model still benefits from the use of thermal time, the performance difference is less pronounced than in the corresponding 1D logistic model, which depends solely on temperature-based input.

Overall, the results indicate that thermal time is a more effective temperature-based predictor, particularly when time is the main explanatory variable. When additional features are included in the model, thermal time still offers an advantage, though its relative influence is reduced. A direct replication of this analysis on the Verdonk dataset is presented in Section 6.4.

4.4.5. Feature Importance

To better understand how the MLP model makes its predictions, permutation feature importance analysis was conducted to determine which environmental variables most strongly influence the predicted broccoli head diameter. The results of this analysis are summarized in Figure 4.14. Each bar in the figure represents the mean increase in prediction error after permuting a specific feature, with error bars showing the standard deviation across 20 repeated permutations.

The results clearly show that *Thermal Time* is by far the most influential variable, with an importance score of 3.69 ± 0.06. This confirms that accumulated temperature, constrained by biological limits, is the primary driver of broccoli development and head growth. The remaining features contribute much less to the model's predictive power, indicating that while environmental conditions influence growth, their effect is secondary to temperature-based development.

In comparison, *maximum temperature*, *soil moisture at shallow depth*, and *mean humidity* show only a very slight influence, with importance scores an order of magnitude lower (between 0.28 and 0.35). These factors provide only minimal improvement to predictive accuracy and are likely associated with short-term effects on plant stress or water balance rather than direct growth progression.

All remaining features, including those related to soil temperature, solar radiation, and precipitation, contribute negligibly to the model. Their low importance indicates that they do not play a meaningful role in predicting head diameter within the observed environmental range. Overall, the analysis confirms that broccoli growth is primarily driven by Thermal Time, with other environmental variables having only minimal influence. Feature importance is re-evaluated on the field dataset in Section 6.5.

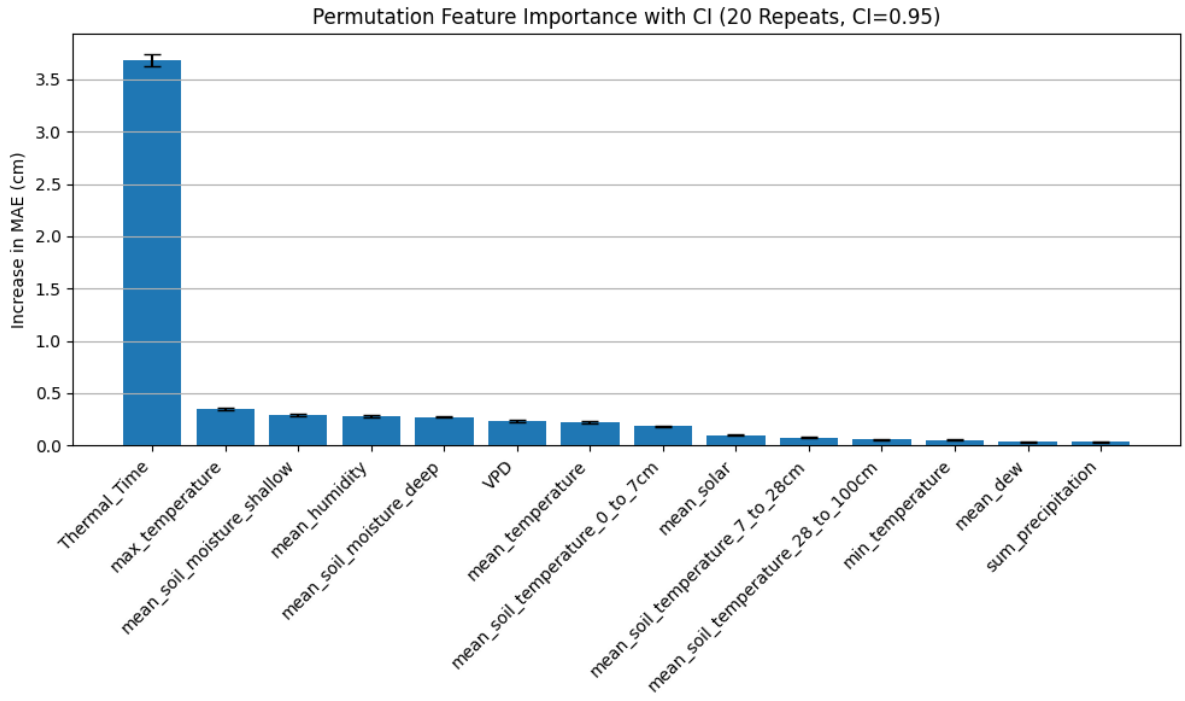


Figure 4.14: Permutation feature importance results for the MLP model trained on the Tanashi dataset. Bars show the mean increase in MAE after feature permutation, with standard deviation across repetitions as error bars.

4.5. Conclusions

The preliminary study on the Tanashi dataset provided a first validation of the modeling framework developed in this research. The analysis offered valuable insights into both the quality of the data and the performance of the different modeling approaches. Although the dataset contained a sufficient number of measurements to study broccoli growth, it also showed clear measurement noise. The analysis of negative dips in head diameter revealed systematic fluctuations across measurement angles and dates, suggesting that part of the noise originated from varying imaging conditions rather than random variation. Smoothing was therefore essential to obtain biologically plausible, monotonic growth curves. Among the evaluated techniques, LOWESS proved to be the most reliable method, as it removed implausible fluctuations while still preserving local changes in growth and avoiding artifacts near the boundaries. For this reason, LOWESS was selected as the preferred preprocessing approach for all subsequent modeling.

Across the eight evaluated models, the MLP achieved the lowest mean absolute error in the preliminary study, with an average MAE of 0.75 cm across the 10 cross-validation folds. In the Tanashi dataset, broccoli heads grew approximately 0.68 cm per day, meaning that the MLP's residual error corresponds to 1.1 days of growth. This level of error is acceptable for most growth-monitoring and model-comparison purposes, since harvest decisions are typically made on a multi-day basis rather than at sub-daily precision. The LSTM, despite its ability to model temporal dependencies, did not outperform the MLP. Its MAE of 0.78 cm and the highest overall standard deviation of 0.08 cm indicate that incorporating historical environmental sequences introduces more variability than useful predictive structure for this dataset. This suggests that broccoli head diameter in the preliminary study is explained primarily by the current environmental state rather than by accumulated temporal patterns, making static models such as the MLP more suitable. Overall, the MLP not only achieved the best statistical performance but did so with an error magnitude that is small relative to both daily biological growth and the operational precision required for practical applications, making it the most effective model in this preliminary study.

The comparison between cumulative temperature and thermal time demonstrated that thermal time provides a more biologically meaningful and statistically better predictor of broccoli development. For six out of eight models, using thermal time significantly reduced prediction errors, confirming that filtering out temperature extremes improves accuracy. The effect was most pronounced for models that rely

mainly on temperature-based inputs, while the difference was smaller and non-significant for the MLP and LSTM, which include additional environmental variables. Overall, thermal time can be considered the most reliable temperature-related feature for modeling broccoli growth.

Finally, the feature importance analysis provided clear evidence that thermal time is by far the most influential variable, confirming its dominant role in explaining head diameter variation. Other features, such as maximum temperature, soil moisture, and humidity, contributed only slightly, and the remaining environmental variables had a negligible impact. This confirms that, within the observed environmental conditions, broccoli growth is mainly determined by temperature rather than by short-term fluctuations in other weather factors.

In summary, the preliminary study validated the modeling pipeline, established LOWESS as the preferred smoothing technique, identified thermal time as the most effective predictor, and showed that the MLP achieved the most accurate predictions. These findings form a strong foundation for applying the same modeling framework to the newly collected field data in the Verdonk study.

5

Field-Based Data Acquisition and Processing

To complement the external dataset and evaluate the performance of the developed models under controlled conditions, we collected a new dataset in a commercial broccoli field, and it is therefore named after the grower, Verdonk Broccoli. This dataset captures plant-level observations over time and includes environmental measurements from the field. The following sections describe the measurement setup used in the field, the structure of the data collection process, and the preprocessing pipeline used to convert raw video data into usable inputs for growth modeling. The field measurements used in this study were collected by a researcher from Inholland at the field of Verdonk. This included preparing the measurement setup, installing and maintaining the camera harness, and performing the video recordings during the season.

5.1. Measurement Setup

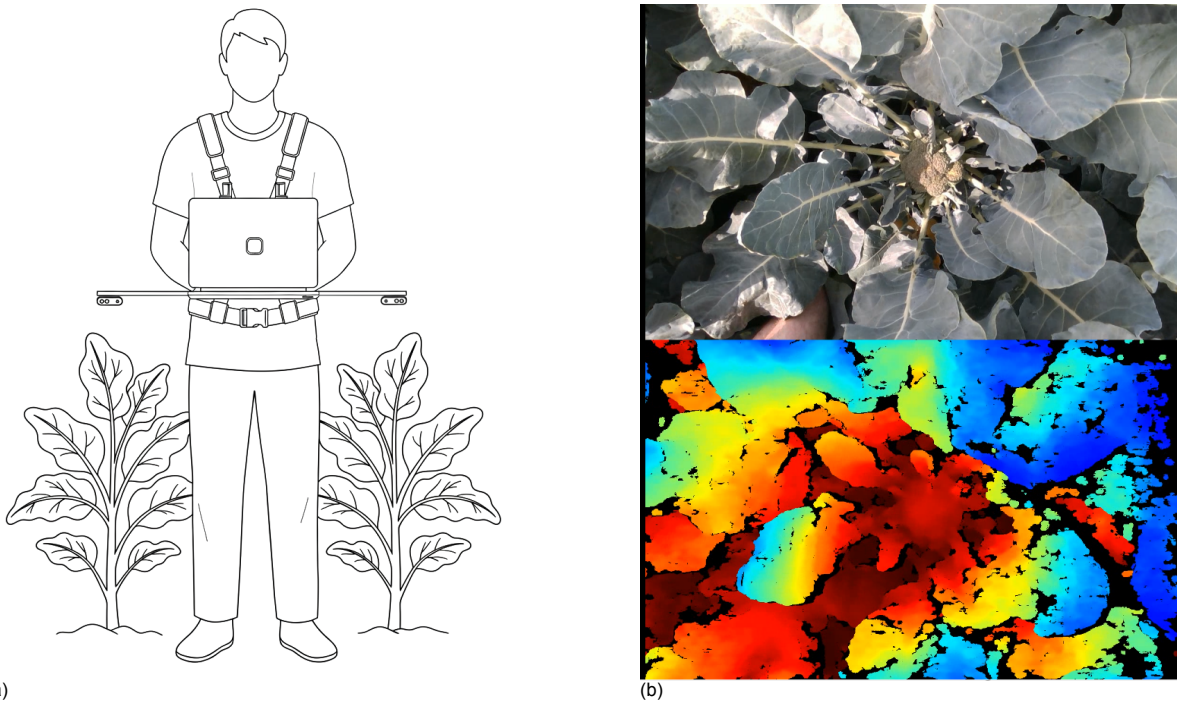
To collect broccoli field data, a custom measurement setup was developed to capture high-resolution, repeatable observations of individual broccoli plants over time. The setup was designed to be mobile, nondestructive to the plants, and capable of operating in outdoor conditions without the need for fixed installations.

At the core of the setup is a laptop mounted on the operator using two adjustable straps to secure the laptop around the shoulders, holding it stably at hip height. This height ensures that the recording devices are well-positioned above the broccoli heads. Two Intel RealSense D415 RGB-D cameras are attached to either side of the laptop. These cameras extend outward such that each records a separate crop line as the operator walks through the field. A schematic representation of this setup can be seen in Figure 5.1a. The setup allows for two simultaneous measurements of two parallel broccoli rows, maximizing spatial coverage and ensuring that each plant is visible from above. These cameras capture both video images, used for broccoli head detection and segmentation, and depth information. An example of such a frame is shown in Figure 5.1b. Together, these are used to estimate head diameter, which serves as a time series input for the different growth models.

The operator walks through a corridor between two parallel rows of broccoli plants, with each camera capturing one of the rows from above. This configuration allows for the simultaneous observation of two crop lines during each pass through the field. The walking path ensures that all plants along both rows are recorded in a single traversal.

5.2. Data Collection

The data collection took place in a broccoli field where a fixed experimental layout was established to ensure consistent measurements over time. This layout was designed to reflect a similar scenario for future automated crop monitoring, as motivated by the precision-agriculture objectives outlined in Chapter 1. Although manual measurements would have been possible to produce a dataset for model evaluation, this approach would have severely limited the dataset size due to the manpower required



(a) (b)

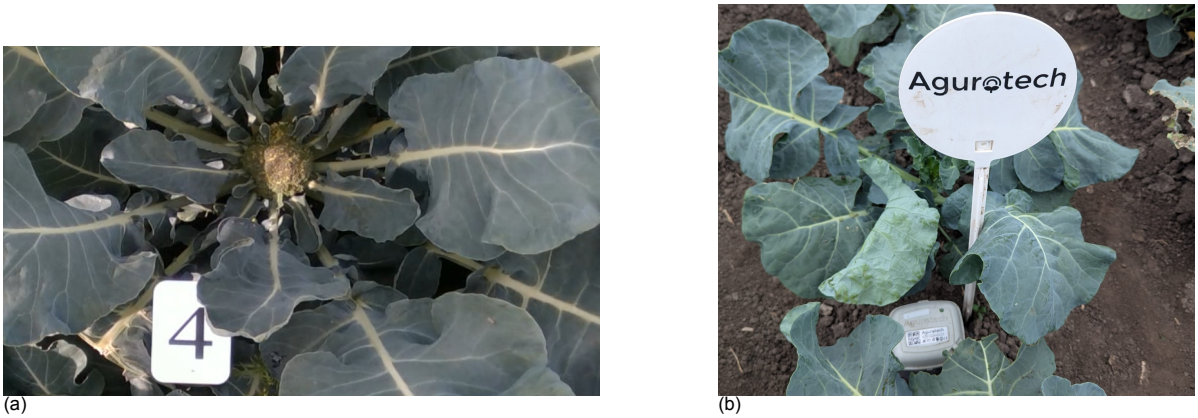
Figure 5.1: a) Schematic representation of the setup used to extract RGB-D data from Verdonk's broccoli fields. b) Example camera output where the top row shows the RGB frame and the bottom row shows the frame's depth channel, which is color-coded to show the depth.

to repeatedly measure hundreds of individual plants. Moreover, a central objective of this study is to investigate how field video can be transformed into reliable plant-level growth data, making the development and validation of an automated acquisition pipeline an integral part of the research rather than a purely instrumental choice.

In total, 12 measurement days were performed from the 30th of June to the 25th of July, covering a full broccoli growth cycle from early head initiation to harvest maturity. Every broccoli plant in the experiment belonged to the *Ironman* cultivar, which is the most used crop by Verdonk. The field was structured around two parallel crop lines, each consisting of 300 individual broccoli plants, and located in the North-Holland province of The Netherlands. These two lines were measured simultaneously using the dual-camera setup described in the previous section, with each camera capturing one crop line from above.

To help with spatial referencing and ensure consistency across measurement days, the crop lines were segmented into smaller units. Visual markers were placed at regular intervals along the rows, dividing each line into 10 segments of 30 broccoli plants. Each marker was positioned between two adjacent plants, at the boundary between segments. These markers provided smaller and more manageable videos, helped with finding a specific plant for validation of methods, and enabled more reliable assignment of plant IDs across different measurement days. An example of such a marker is visible in Figure 5.2a.

For measuring soil-specific data, Agurotechs' Saturnia Parva climate pole, provided by Vertify, was used. At the center of the two crop lines, the sensor was installed to continuously record environmental conditions. The pole measured soil temperature, moisture, and electrical conductivity at depths of 15 and 30 cm. Its central position ensured that the recorded climate data were representative of the conditions experienced by all plants within the experimental area. Figure 5.2b shows how this climate pole is installed in the field.



(a) (b)

Figure 5.2: a) Example frame where a marker is clearly visible, indicating the start of a new segment. b) Photo of the installed climate pole.

5.3. Head Detection

This section describes the approach used to detect broccoli heads in the field video data, which represents the first step of the automated processing pipeline. Reliable head detection is essential for all subsequent stages, including plant identification, tracking, and diameter estimation, since detection errors propagate directly through the pipeline and affect measurement quality.

5.3.1. Preliminary Head Detection Model

A preliminary head detection model was previously developed at Inholland to automate the detection of broccoli heads from RGB imagery under field conditions. This model, based on the YOLOv8n architecture, demonstrated reliable detection performance across diverse lighting and occlusion scenarios, achieving an F1-score of 96% and further cross-validated on an external dataset by Blok et al. with an F1-score of 98% [4]. The associated training and evaluation code from this work served as the basis for the development of the head detection model used in this study.

5.3.2. Model Training for This Study

Building upon the earlier Inholland research, the same YOLOv8n architecture was adopted as the basis for the head detection model used in this study. The model was trained from scratch using newly collected data from the Verdonk broccoli fields, ensuring that the detection performance was tailored to the specific visual characteristics of this dataset.

Dataset and Annotation

The training dataset consisted of manually annotated RGB images extracted from the recorded measurement videos. From each of the seven measurement days, 50 random frames were selected in which at least one broccoli plant was visible. Each annotation contained a bounding box around a visible broccoli head. Not all selected frames contained plants with a head. Some of the selected frames did not contain visible heads, either because the plant had not yet reached its head initiation phase on earlier dates or because the head had already been harvested on later dates. This resulted in a total of 250 annotated samples rather than the extracted 350.

The original dataset, therefore, showed an imbalance in the representation of measurement days, with some days contributing substantially more images than others. This imbalance led to reduced model sensitivity to small broccoli heads, which were primarily present during the earlier growth stages. In contrast, the performance on large broccoli heads did not suffer from this imbalance. This can be explained by the higher visual distinctiveness of mature heads, which show a more defined contour and larger bounding box areas, making them easier to detect even when fewer examples are present in the training data. To address the reduced performance on smaller heads, additional annotated frames from the first two measurement days were added, improving both temporal balance and the representation of earlier growth stages. As shown in Figure 5.3, the resulting final dataset contains a slightly higher proportion of images from the early measurement days, ensuring better representation of smaller broccoli heads and improving detection performance across developmental stages. Figure 5.4 shows one example annotated video frame for each measurement date.

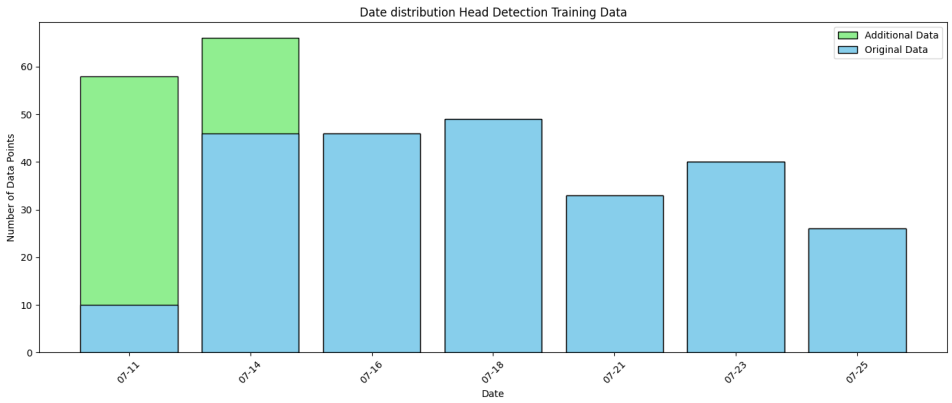


Figure 5.3: Distribution of annotated images per measurement day for the original and the final dataset. The stacked bar chart shows the addition of early-stage samples to improve model sensitivity to small broccoli heads.

Figure 5.4: Examples of 7 annotated video frames, each from a different measurement date.

Model Training and Evaluation

The YOLOv8n model was trained using the complete dataset described above, with an 80/20 split for the training and test sets. The IoU threshold was set to 0.6, consistent with the previous implementation.

Model performance was evaluated on the independent test set using standard object detection metrics. The model achieved a precision of 0.9305, a recall of 0.9872, and an F1-score of 0.958. The mAP@50 was 0.988, while the stricter mAP@50-95 metric reached 0.808. These results demonstrate that the retrained YOLOv8n model is capable of accurately detecting broccoli heads across different developmental stages, including early growth phases that were underrepresented in the original annotation set.

5.4. Plant ID Assignment

A challenge in constructing a consistent time series of broccoli growth is the assignment of consistent plant IDs across all measurement days. Since the data are collected from video recordings while walking through the field, there is no inherent spatial reference system available that marks each plant individually. To address this, a slot-based identification method was developed, which is based on the estimation of camera motion throughout the video as seen in Figure 5.5.

For each measurement session, the video for each segment was manually trimmed such that the first and last frames both show a broccoli plant positioned at the center of the image. These two boundary frames serve as ground truth for the spatial placement of the slot centers and are therefore used to correctly infer the boundaries of all intermediate slots.

To identify the relative position of each broccoli plant within the segment, we first estimated the total vertical motion of the camera throughout the video. This was achieved by computing frame-to-frame pixel displacements using OpenCV's Enhanced Correlation Coefficient (ECC) algorithm with affine motion estimation. To improve the robustness of the ECC registration under variable lighting conditions, each frame was preprocessed using Contrast Limited Adaptive Histogram Equalization (CLAHE). This enhances local contrast, allowing for more accurate alignment, particularly in outdoor environments with uneven illumination.

For each pair of consecutive frames, the vertical translation component was extracted and accumulated to compute the total vertical camera displacement over the segment. Given that each segment always contains exactly 30 plants, we divided the total vertical movement into 30 equally spaced slots. Each slot corresponds to a single broccoli plant. As the operator walks through the segment, each

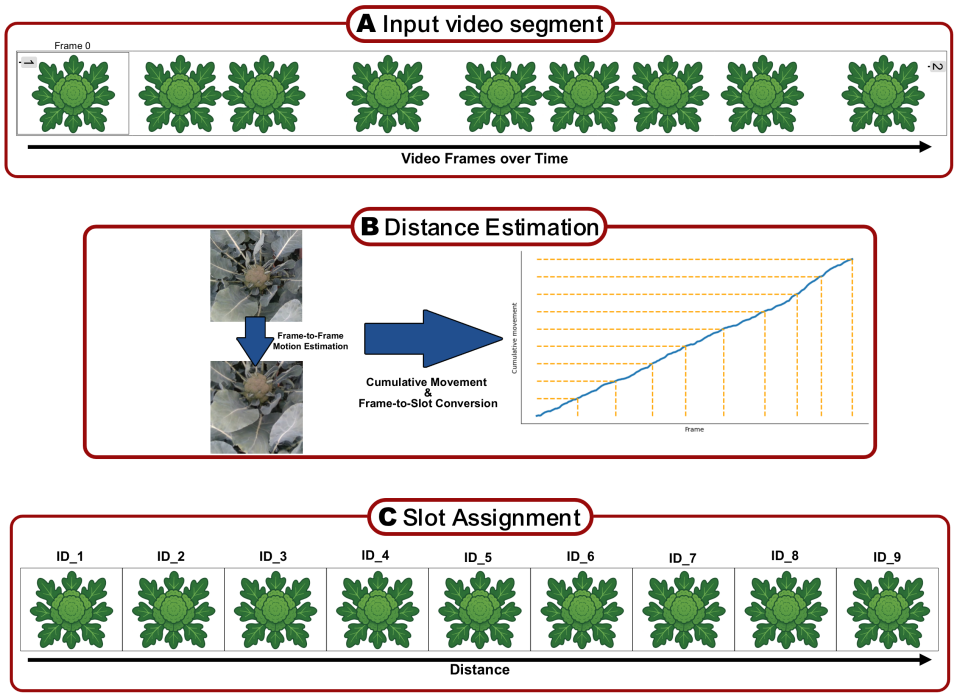


Figure 5.5: Overview of the slot-based broccoli identification pipeline: (A) trimmed input video segment with visible broccoli plants and segment markers, (B) frame-to-frame camera motion estimation, cumulative displacement, and frame-to-slot conversion, and (C) uniform slot assignment of plant IDs across the segment.

detection is assigned to a slot based on the cumulative vertical displacement at the time of detection. This effectively links detections to plant positions based on their relative location in the video. To ensure proper boundary handling, the first and last slots are assigned only half the width of the others, reflecting the fact that the first and last frames center a broccoli plant in the middle of the frame. This avoids assigning detections outside the visible range of the segment.

5.5. Tracking

After the individual broccoli detections were obtained for each frame, a tracking pipeline was implemented to transform these isolated detections into temporally consistent trajectories, or *tracks*, each corresponding to a single broccoli plant. The purpose of this step was to ensure that every detected broccoli head received a unique ID that persisted across consecutive frames, thereby enabling the combination of different measurements of the same broccoli for a more consistent data point. The process integrated both spatial and temporal reasoning and was designed to be robust to detection noise, camera motion, and false positive detections. A visual representation of 3 intermediate steps can be seen in Figure 5.6

5.5.1. Detection Linking

The tracking algorithm operates sequentially across video frames. For each frame, all detections that pass a spatial filter are extracted. Detections whose horizontal position lies outside the central crop area are removed to prevent the inclusion of detections that correspond to other broccoli columns. For every detection, the center coordinates of its bounding box are calculated, forming a set of two-dimensional points that represent the spatial positions of the broccoli heads in that frame.

To associate detections across frames, a cost matrix is constructed that quantifies the dissimilarity between the positions of currently active tracks and the detections in the new frame. The spatial distance between the bounding box centers serves as the main matching criterion, while a temporal penalty is added to account for skipped frames. The total association cost between a track and a detection is defined as a linear combination of spatial and temporal components:

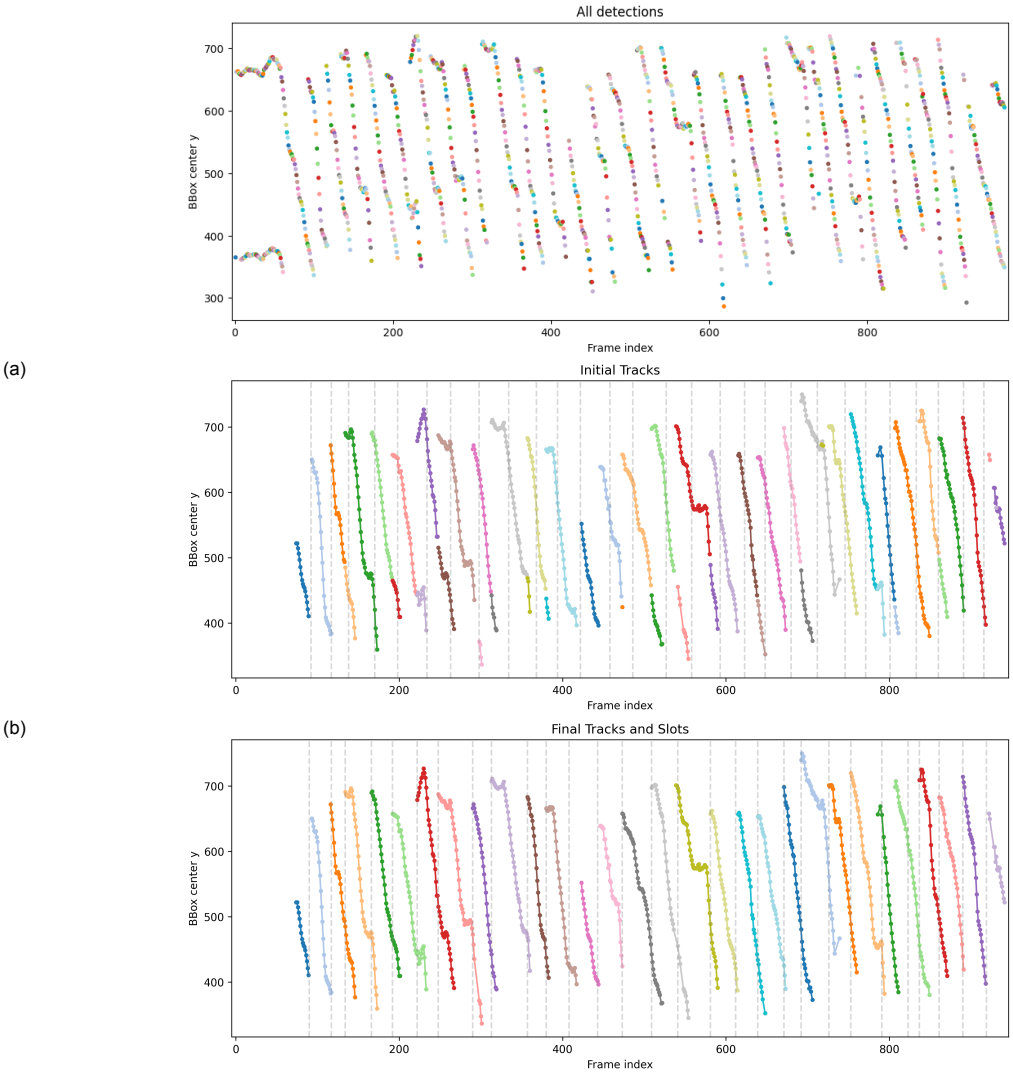


Figure 5.6: Overview of the tracking and slot refinement process. (a) Broccoli detections before trimming of video and track creation, showing all detections of a single video. (b) Initial tracking results after detection linking, including the preliminary slot boundaries derived from cumulative camera displacement. (c) Final refined tracks after velocity-based repair and slot boundary refinement, showing continuous trajectories and corrected plant slot alignment.

$$C_{ij} = \alpha|y_i - y_j| + \beta|x_i - x_j| + \gamma\Delta t \tag{5.1}$$

where (x_i, y_i) and (x_j, y_j) denote the centers of two detections which are part of two distinct tracks, Δt is the frame gap since the track was last updated, and α, β, and γ are weighting parameters controlling the contribution of each term. In this study, the parameters were empirically set to $\alpha = 1$, $\beta = 0$, and $\gamma = 10$. The horizontal weight β was set to zero because the horizontal position of the broccoli heads remains relatively constant across frames, as the camera movement primarily occurs along the vertical (y) axis. Consequently, variations in x carry little information for matching detections between frames. A maximum matching distance threshold of 100 ensures that detections that are too far from any active track initiate a new trajectory rather than being incorrectly assigned.

5.5.2. Velocity-Based Track Repair

Although the frame-to-frame association captures most trajectories correctly, it is intentionally designed to be conservative in merging detections to minimize the risk of incorrect associations between distinct broccoli heads. To reconstruct continuous plant trajectories, a post-processing procedure based on estimated vertical motion was applied. This method exploits the regular camera movement along the

crop row, where each broccoli appears to move vertically through the frame at a near-constant rate.

For each track, the average vertical velocity is estimated using the global displacement between detections. Using this velocity, the algorithm predicts the expected position of a track in the following or preceding frames. When another track starts or ends close to this predicted position and within a defined spatial gap, both tracks are merged. This merging is constrained by several checks to prevent erroneous associations: (1) tracks are not merged if an intermediate object lies between them in the frame sequence, (2) excessive positional jumps are disallowed, and (3) tracks are not merged if the predicted positional difference exceeds the expected vertical movement range of a broccoli head between consecutive frames, ensuring that only spatially plausible displacements are considered. This iterative repair process is repeated until no further merges are possible, producing smooth and complete broccoli trajectories.

5.5.3. Slot Boundary Refinement and ID Assignment

The initial slot boundaries derived from camera motion estimation provide a first approximation of plant positions, but this approach is not perfect. Small inaccuracies in the estimation of frame-to-frame displacement accumulate progressively over time, resulting in increasing spatial drift along the vertical axis. This accumulation of error is minimal near the start and end of the segment, where the first and last frames are manually aligned, but it typically reaches its maximum in the center of the segment. As a result, the initially uniform slot division can deviate from the true spatial arrangement of the broccoli plants.

To correct these accumulated errors, the slot boundaries are refined using the actual broccoli detections obtained from the tracking results. The refinement algorithm analyzes pairs of adjacent broccoli trajectories to identify the frame at which their bounding boxes are most symmetrically positioned with respect to the center of the frame. This symmetry represents the location of the true boundary between two neighboring plants. When this point deviates from the initially estimated slot boundary, a correction is applied by locally shifting the boundary position to align with the observed symmetry in the detections.

This process is repeated across all adjacent track pairs, progressively refining each slot boundary based on empirical detection data rather than on accumulated camera displacement alone. Through this refinement, the spatial alignment of the slots becomes more consistent with the visual evidence from the video, effectively correcting for drift and improving the accuracy of plant ID assignment across the segment.

5.5.4. Output and Validation

The final output consists of a list of trackings indexed by frame and track identifier, where each entry contains the bounding box, confidence score, median depth, and estimated head diameter. The resulting tracks were visualized by plotting the vertical position of each broccoli over time, which revealed clear, continuous trajectories corresponding to the plants' passage through the camera's field of view. To further validate and manually refine the results, a small GUI was developed. This interface allowed the visual inspection of each frame with the corresponding track and slot identifiers overlaid on the video. Using this tool, erroneous detections could be interactively removed, and incorrect slot assignments could be corrected. This manual validation step ensured that false positives were excluded and that each broccoli track was correctly associated with its physical plant slot, resulting in a clean dataset for subsequent growth analysis.

5.6. Diameter Estimation

After the head detection model has generated bounding boxes for all broccoli heads, these detections are then used to estimate the physical diameter of each head. The initial formula used to compute the diameter is defined as:

$$D = \frac{D_{px} \cdot H_1}{f_x} \tag{5.2}$$

Where D_{px} denotes the diameter in pixels, H_1 the depth estimated from the camera, and f_x the focal length. However, this approximation consistently underestimated the true diameter of the broccoli head. This error is explained by the depth used in the calculation being measured at the center of the crown,

whereas the diameter is determined by the edges. Since the depth at the edges is greater than at the center, this leads to the difference.

To reduce this, a correction factor based on empirical measurements was introduced. The updated formula accounted for an empirically derived ratio C_b. This enhancement yields the new formula:

$$D = \frac{D_{px} \cdot H_1 \cdot (f_x + D_{px} \cdot C_b)}{f_x^2} \tag{5.3}$$

This revised approach provided a more accurate estimation of crown diameters. This ratio is empirically determined at 0.1.

To estimate the depth H_1 used in these calculations, a cropped region of interest was defined at the center of each detection's bounding box. This region covered 50% of the box width and height to exclude background pixels and noisy border regions. The median depth value within this region was then taken as the representative depth of the broccoli head. The diameter in pixels, D_{px}, was approximated as the largest side length of the bounding box, providing a consistent estimate of the visible crown size independent of orientation. By combining the pixel-based diameter, the median depth, and the intrinsic camera parameters, the physical diameter was computed for each detection using Equation 5.3.

After the creation of the broccoli tracks, the diameter estimation was applied to each detection belonging to a given track. This means that for every frame in which a broccoli was detected, a diameter value was calculated using Equation 5.3. As a result, each track contains multiple estimates of the same broccoli head diameter.

The diameter estimations occasionally produced outlier values caused by partial visibility or temporary occlusions of the broccoli heads. In particular, detections where the broccoli was located near the borders of the frame and detections that suffered from occlusions by leaves tended to underestimate the true diameter as the head was only partially visible. To address this, the lower 50th percentile of all estimated diameters within each track was removed before calculating the final value. The median of the remaining estimates was then taken to represent the final diameter of that broccoli at the time of measurement. This approach effectively reduces the influence of extreme or underestimated values, resulting in more stable and reliable measurements for subsequent growth modeling. Specifically, it reduces the mean standard deviation of measurements from a single broccoli from 0.61 cm to 0.14 cm. Figure 5.7 shows an example of all measurements of a single broccoli and the final extracted diameter value.

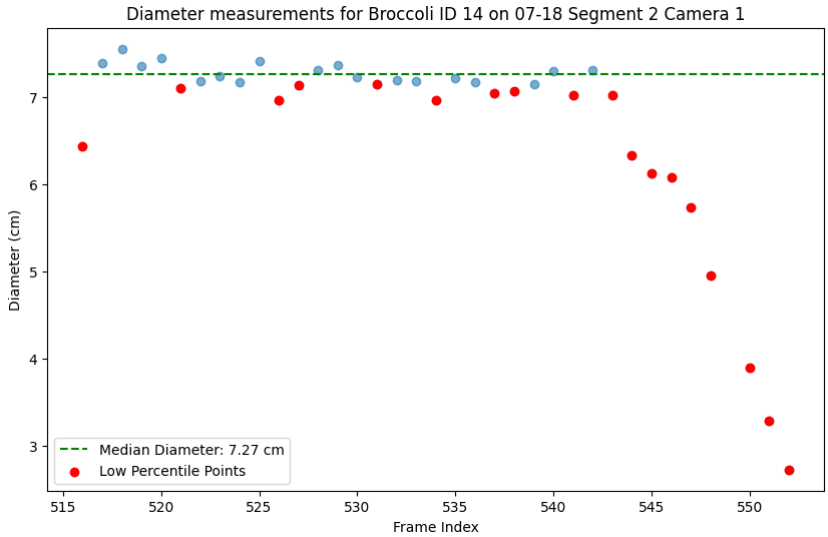


Figure 5.7: An example of all measurements of one broccoli from one video. The lower 50th percentile points are not taken into account when taking the median diameter of all measurements.

5.7. Data Filtering

Before using the extracted diameter data for model development, several filtering steps were applied to remove unreliable measurements and biologically implausible growth patterns. These steps ensured that only consistent and realistic growth trajectories were retained for the subsequent analysis. The filtering procedure was applied sequentially and focused on removing plants or individual measurements that showed inconsistent temporal behavior or measurement artifacts.

5.7.1. Decreasing Diameter

As broccoli head growth is a monotonic process during the main development phase, any decrease in measured diameter over time is biologically implausible and therefore likely the result of detection errors or partial occlusions. To identify such cases, the time series of each broccoli was inspected for negative differences in diameter between consecutive measurement dates. If a decrease was detected at any point, the entire broccoli track was removed from the dataset. This procedure eliminated all plants that exhibited at least one dip in diameter across the recorded period, retaining only those with continuously increasing growth profiles.

5.7.2. Initial Growth Slope

In the early stages of the measurement period, the detection model occasionally produced false positive detections in frames where no visible broccoli head was yet present. These false detections typically resulted in abnormally low or even negative initial growth rates when compared to the subsequent measurements. To identify and correct these cases, the first two diameter measurements of each broccoli were used to calculate an initial growth slope, expressed in centimeters per day.

The distribution of these initial slopes was then analyzed, and the lowest 2% were considered abnormal, as they represented unrealistic or biologically implausible growth rates. For all plants in this range, only the first measurement was removed, since the following measurements generally corresponded to valid detections once the broccoli head became visible. This selective removal step effectively eliminated false early detections while retaining the reliable portions of each broccoli's growth trajectory.

5.7.3. Local Growth Slopes

Next, the local slope between every pair of consecutive measurements was computed for each broccoli. This local slope quantifies the short-term growth rate and allows for detecting inconsistent or negative growth intervals within otherwise valid trajectories. Measurements whose local growth slope fell below the 2nd percentile of the overall slope distribution were flagged as abnormal and removed. These instances typically correspond to measurements with partial occlusion of the broccoli crown.

5.7.4. Final Dataset

After applying all filtering steps, the dataset was reduced from 599 broccoli growth curves, with a total of 2,736 individual measurements, to 546 curves containing 2,639 measurements. The remaining dataset contained smooth, monotonic growth trajectories that accurately reflected the biological development of the broccoli heads.

To ensure the correctness of the filtering procedure, the removed data points were manually verified by inspecting the corresponding frame images. This visual validation confirmed that the excluded measurements indeed originated from incorrect detections or incomplete visibility of the broccoli heads. Some examples of these removals are shown in Figure 5.8. Overall, the final dataset consisted of clean, realistic, and continuous growth profiles suitable for further modeling and analysis.

Figure 5.8: Examples of removed detections. The first row shows three cases of initial growth slope removals. The second row presents three cases of local growth slope removals and decreasing diameters.

5.8. Weather Data

To complement the visual measurements of broccoli growth, environmental data were collected for the Verdonk field experiments. These data provide the necessary climatic and soil context required to model and interpret the observed growth patterns. Two primary sources of weather data were used: direct field measurements from the climate pole installed on-site, and meteorological data obtained from the Open-Meteo API.

The climate pole, positioned centrally between the two measured crop rows, continuously recorded soil and environmental variables throughout the growing period. Measurements were taken approximately every thirty minutes at two soil depths, 15 cm and 30 cm, corresponding to the upper and lower root zones of the broccoli plants. For each depth, the pole measured the soil temperature, electrical conductivity, and volumetric water content. These values indicate soil thermal conditions, salinity, and moisture availability, which are key factors that influence water uptake and nutrient transport in the plant.

In addition to the in-field measurements, broader atmospheric variables were retrieved using the Open-Meteo API, which sources its information directly from the Royal Netherlands Meteorological Institute (KNMI). The retrieved variables included near-surface air temperature, relative humidity, precipitation, wind speed, and incoming shortwave radiation. Together, these features capture the main environmental drivers of plant development.

Because the meteorological and soil variables were recorded hourly, they were aggregated to daily values to create a uniform dataset suitable for integration with the growth measurements. Specifically, for each day, the mean, minimum, and maximum were computed for air temperature, while relative humidity, wind speed, shortwave radiation, electrical conductivity, soil temperature, and volumetric water content were averaged. Precipitation was summed over the day to capture total daily rainfall.

Finally, additional derived features were engineered using the methods described in Section 3.2. Specifically, the VPD was calculated from temperature and relative humidity to quantify the atmospheric moisture demand, while thermal time was calculated to describe how plant development accumulates over time as a result of temperature. An additional offset as described in Section 3.2.3 was applied to this thermal time to account for the early growth period that was not captured before the first measurement, ensuring that the results remain consistent with the modeling approach used for the external dataset. Together with the climate pole sensor and the API weather data, these features finalize the environmental factors that will be used in the growth modeling and are analysed in Chapter 6.

6

Field Study Results

This chapter presents the predictive modeling results obtained from the field dataset collected in the Verdonk study, as described in Chapter 5. This chapter is parallel to the preliminary analysis on the Tanashi dataset reported in Section 4.4, enabling a direct comparison between the external dataset and the field measurements. Using the same modeling approaches and evaluation metrics introduced in Chapter 3, the results in this chapter quantify model performance on the Verdonk dataset and assess the influence of environmental features under commercial growing conditions.

6.1. Thermal Time Offset

Figure 6.1 shows the effect of the thermal time offset on the field-based Verdonk dataset. Similar to the preliminary study, the original thermal time representation produces clustering of diameter measurements at common thermal time values, despite visible differences in plant development. This confirms that variations in head initiation timing also occur in the realistic production setting.

After applying the offset correction, the growth trajectories become more dispersed and better aligned along the developmental timeline. Individual plants exhibit clearer progression patterns, and diameter variation at shared thermal time values decreases. This indicates that the offset effectively compensates for missing early-stage development and harmonizes growth curves across different plants.

These results confirm that the thermal time offset improves the biological consistency of the field dataset, not only in a controlled benchmark study but also under real commercial growing conditions. The adjusted representation supports more reliable learning of growth patterns and enhances the interpretability of downstream predictive modeling results.

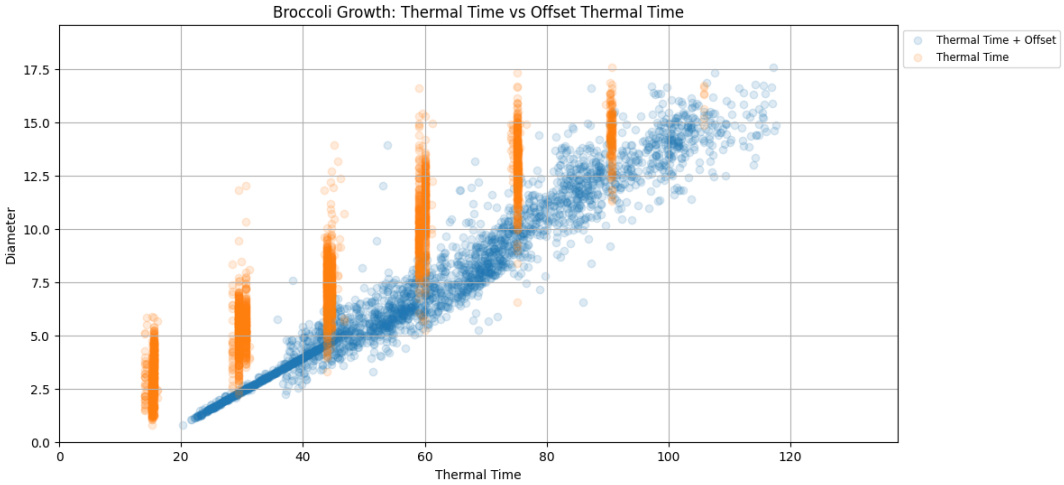


Figure 6.1: Diameter plotted against original thermal time and adjusted thermal time with offset of the Verdonk dataset.

6.2. Hyperparameter Tuning Results

For the MLP model, the tuning results for the field-based dataset largely confirm the trends observed in the preliminary study. Higher dropout rates again resulted in reduced predictive accuracy, while the learning rate and batch size showed only minor influence on model performance. All tested configurations converged within the fixed number of training epochs, indicating that the learning rate range was sufficiently broad to ensure stable optimization. The final selected configuration for the MLP used a learning rate of 0.001, a batch size of 256, a dropout rate of 0, and three hidden layers with 128, 256, and 128 neurons, respectively. This configuration achieved the lowest mean validation MAE across folds and was therefore used in all experiments.

For the LSTM model, similar patterns to those seen in the preliminary study emerged. Performance decreased with increasing sequence length, with shorter sequences consistently achieving lower validation errors. This again suggests that broccoli growth is primarily determined by the current environmental conditions rather than by longer historical dependencies. Incorporating extended temporal information appears to introduce noise rather than informative structure. The influence of the learning rate and batch size remained limited, with all configurations reaching convergence within the fixed training schedule. The optimal LSTM configuration was found with a learning rate of 0.01, a batch size of 32, a hidden state size of 64, and a sequence length of 2.

The hyperparameter tuning outcomes confirm the observations made in the preliminary study. The MLP again performed best when no dropout was applied. Similarly, the LSTM achieved optimal results when the temporal dependency was kept minimal, with the shortest tested sequence length yielding the lowest validation error. Together, these results highlight the importance of systematic hyperparameter tuning, as small architectural and training adjustments, such as dropout rate or sequence length, can substantially influence the model's ability to learn meaningful relationships from the available data.

6.3. Overall Comparison Across Models

Model	MAE (Mean)	MAE (Std)
MLP	**0.583**	0.038
Logistic 2D	0.632	0.056
Quadratic	0.642	0.035
Baseline	0.658	0.034
Logistic	0.664	**0.029**
Grevsen	0.665	0.053
LSTM	0.803	0.093
Grevsen Bias	0.854	0.075

Table 6.1: MAE performance comparison across models. The MLP model has the lowest MAE, and the Logistic 1D model has the lowest variance.

This section directly addresses Research Question 4 by comparing the ability of different models to represent broccoli growth dynamics. The results of the 10-fold cross-validation on the field-based dataset demonstrate clear differences in predictive performance among the evaluated models. Figure 6.2 displays the fold-level MAE distributions for all models, along with the mean values and 95% confidence intervals. Table 6.1 shows for each model its mean MAE and the standard deviation. Finally, Figure 6.3 shows the predictions of all models on an example broccoli growth trajectory.

The MLP model achieves the lowest overall MAE of 0.583 (± 0.038) with relatively low variability across folds, confirming its robustness and predictive accuracy under field conditions. The improvement of the MLP compared to the other models is statistically significant for most pairwise comparisons, except for the *logistic 2D* model. Similar to the findings from the Tanashi dataset, the superior performance of the MLP can be attributed to its ability to model non-linear relationships between multiple environmental features while relying on only the current weather state instead of including previous days.

The *logistic 2D* model, which includes both thermal time and soil moisture, achieves the second-best performance with an MAE of 0.632 (± 0.056). This confirms that incorporating soil moisture information improves predictions relative to the one-dimensional *logistic* model with an MAE of 0.664 (± 0.029). The improvement is statistically significant compared to the *Grevsen Bias* and *LSTM* model.

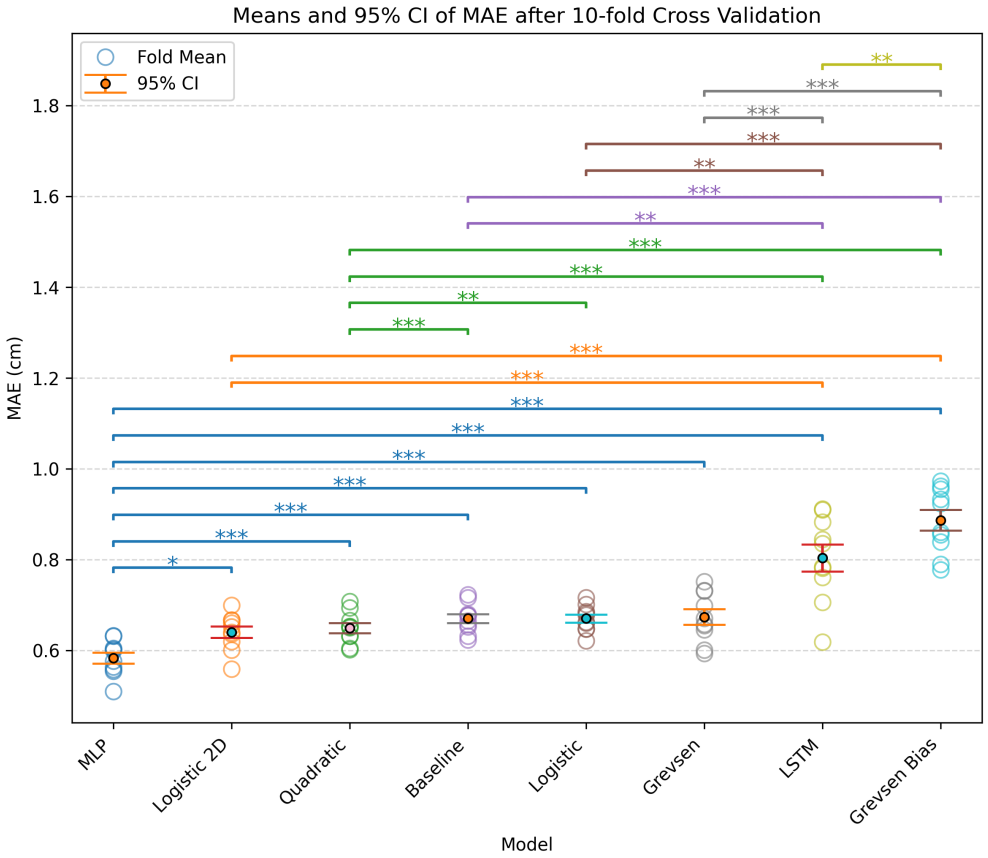


Figure 6.2: Fold-level MAE results for each model on the Verdonk dataset, including mean values with 95% confidence intervals and significance annotations from paired t-tests.

The *quadratic* model performs slightly better than the *baseline* model, with MAE values of 0.642 (± 0.035) and 0.658 (± 0.034), respectively, and this difference is statistically significant. Similar to the findings from the preliminary study, the fitted quadratic coefficient is minimal, indicating that the inclusion of a second-order term provides only limited curvature to the growth trajectory. As a result, the predicted growth curves of both models are nearly identical. This suggests that broccoli growth progresses in a nearly linear manner with respect to thermal time, and that higher-order terms do not meaningfully enhance predictive accuracy. However, the small difference in performance is still significant.

The *logistic* and *Grevsen* models yield nearly identical MAE values with an MAE of 0.664 (± 0.029) and 0.665 (± 0.053), respectively, confirming that the differences between these biologically inspired formulations are minor in this dataset. The *Grevsen Bias* variant performs notably worse with an MAE of 0.854 (± 0.075), and this underperformance is statistically significant compared to all other models except the *LSTM*, indicating that the added bias term introduces unnecessary flexibility that does not correspond to meaningful biological variation.

The *LSTM* model, designed to capture temporal dependencies, shows the highest variability and one of the poorest mean performances with an MAE 0.803 (± 0.093). This result is consistent with the observations from the Tanashi dataset, where sequence-based learning did not improve accuracy. The relatively weak performance and large fold-level variance indicate that sequential dependencies across measurement days do not enhance predictive power and that the added complexity of recurrent modeling introduces noise rather than useful information.

Overall, the ranking of the models reveals both consistencies and differences compared to the preliminary study. The MLP again emerges as the best-performing model in both datasets, confirming its robustness and ability to generalize across experimental conditions. Likewise, the *Grevsen Bias* model consistently ranks as the worst-performing approach, showing that the addition of a free bias term systematically reduces predictive reliability.

However, some differences in the intermediate rankings are observed between the two studies. These deviations are small and likely result from measurement uncertainty in the field-based data rather than from true model performance differences, although cultivar variation may also play a minor role. As such, these shifts are of limited practical importance, and the focus lies primarily on the consistent outliers: the MLP, which performs best in both datasets, and the *Grevsen Bias* model, which remains the least effective across studies.

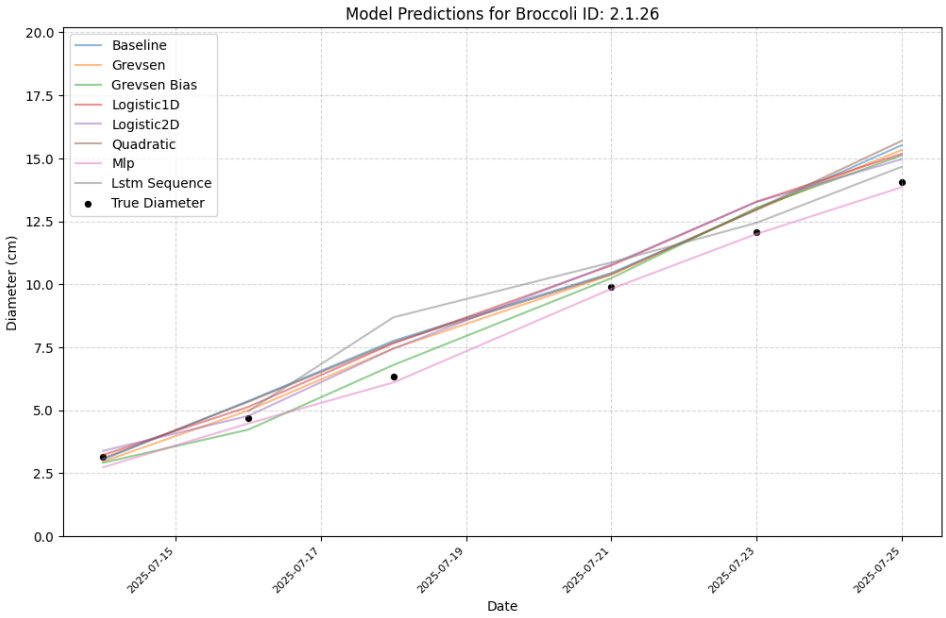


Figure 6.3: An example broccoli sample with predicted values from all models. Most models follow the trend of the example broccoli growth well.

6.4. Thermal Time vs. Cumulative Temperature

This section addresses Research Question 3 by empirically comparing the two temperature representations. As in the preliminary study, this experiment compared the predictive performance of models trained using cumulative temperature against those using thermal time as the primary air temperature feature. Both configurations were applied to the same eight growth models, and performance was evaluated through ten-fold cross-validation. The MAE per fold was calculated for each model and each configuration. The fold-level results were then compared using paired two-sided t-tests to determine whether the differences in prediction error between the cumulative temperature and thermal time representations were statistically significant. The resulting distributions of the cross-validated MAE scores are presented in Figure 6.4.

Across the tested models, the results align with the findings from the Tanashi dataset. The models based only on temperature inputs, such as the baseline, quadratic, logistic, and Grevsen models, all perform significantly better when thermal time is used instead of cumulative temperature. This confirms that integrating biologically informed temperature thresholds (T_{base} and T_{max}) enhances predictive accuracy by excluding temperature extremes that do not contribute to physiological development. In contrast, models that incorporate additional explanatory variables show a less pronounced difference between the two configurations.

In contrast, models that incorporate additional explanatory variables show a less pronounced difference between the two configurations. Specifically, all thermal time models achieved statistically significant improvements over their cumulative temperature counterparts, except for the MLP, LSTM, and logistic2D models. The absence of a significant difference for the MLP and LSTM models is consistent with the preliminary study and can be attributed to their multi-variable nature. Both models utilize additional environmental features beyond temperature, such as humidity, radiation, and soil moisture, which collectively reduce the relative influence of the temporal feature.

In this study, the logistic2D model also showed no statistically significant difference between the two configurations. A plausible explanation for this deviation from the Tanashi results is the improved

quality of the soil moisture data in the Verdonk experiment. Whereas the Tanashi dataset relied on API-derived moisture estimates, the Verdonk dataset includes direct sensor measurements from the field, providing more accurate and representative soil moisture information. As a result, the inclusion of soil moisture sensor data contains enough information for the relative improvement of thermal time over cumulative temperature to be reduced such that there is no statistical significance anymore. This is supported by the comparison with the 1D logistic model, which does show a statistically significant difference between the two temperature representations.

In summary, the results demonstrate that thermal time is a more biologically grounded and effective temporal feature for modeling broccoli head growth. Models that rely only on temperature-based inputs, such as the baseline, quadratic, logistic, and Grevsen models, show statistically significant improvements when thermal time is used instead of cumulative temperature. This confirms that incorporating biologically meaningful thresholds filter out temperature extremes, leading to more accurate predictions of developmental progress.

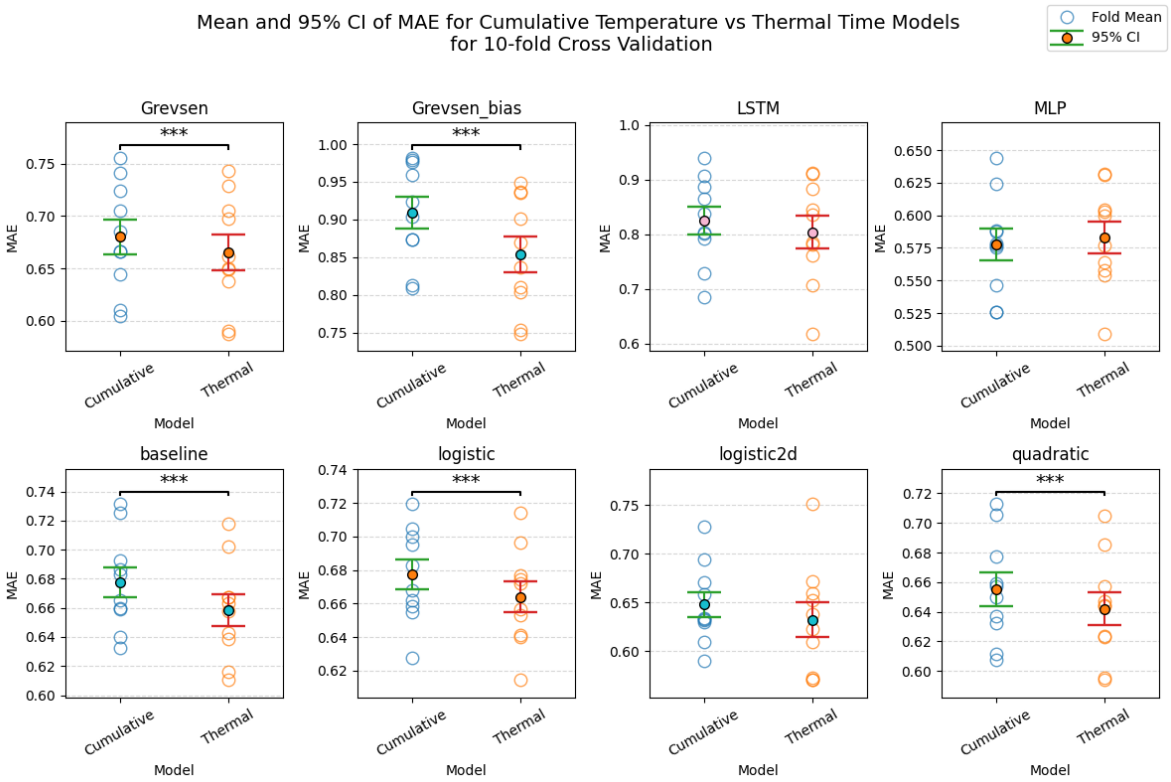


Figure 6.4: Cross-validated MAE for models using cumulative temperature vs. thermal time. Each subplot shows the fold-level MAE, mean with 95% confidence interval, and significance annotation. Significance levels are denoted by conventional asterisks: * for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$.

6.5. Feature Importance

This section contributes to Research Question 2 by evaluating which environmental parameters most strongly influence predictive performance. To investigate which environmental factors most strongly influence the predictions of the MLP model, permutation feature importance analysis was performed. This approach quantifies the contribution of each input variable by measuring the increase in model prediction error after randomly permuting that feature. Features that cause a larger rise in prediction error when permuted are considered more influential in determining broccoli head diameter. Each feature was permuted 20 times to account for random variation, and the mean increase in MAE with its 95% confidence intervals was computed.

The results of the analysis are shown in Figure 6.5. As expected, *Thermal Time* remains the most influential variable, with an importance score of 3.27 ± 0.05. This finding aligns with the outcome from the preliminary study and confirms that temperature-based accumulation within biologically active

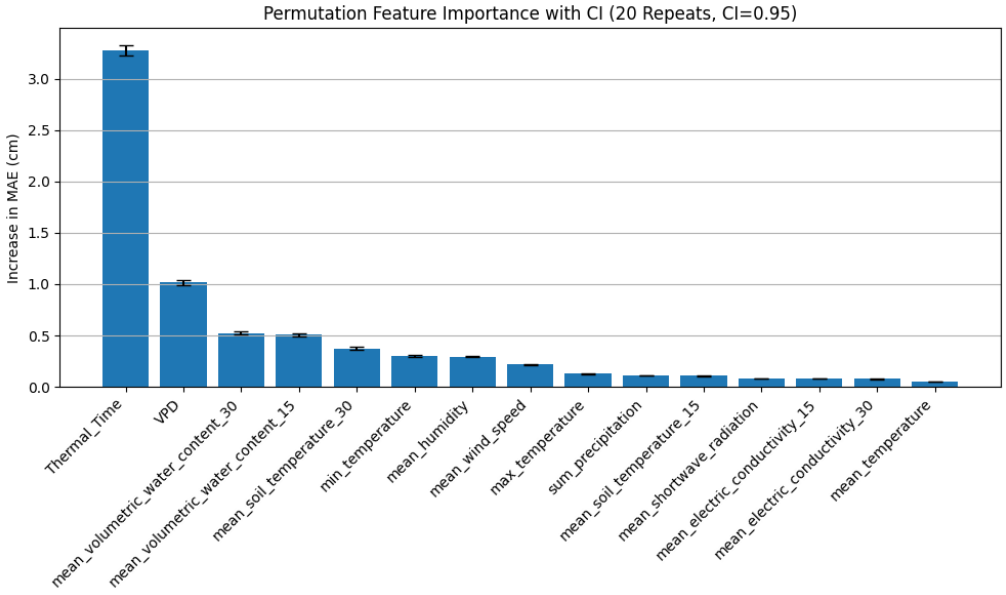


Figure 6.5: Permutation feature importance results for the MLP model trained on the Verdonk dataset. Bars show the mean increase in MAE after feature permutation, with 95% confidence intervals across 20 repetitions as error bars.

thresholds is the dominant driver of broccoli head development.

The second most important variable is the *VPD*, with a considerably lower importance score of 1.02 ± 0.02. This suggests that humidity-related stress, expressed through VPD, plays a notable but secondary role in growth. It reflects the influence of atmospheric moisture demand on transpiration and, consequently, on the plant's ability to sustain photosynthetic activity under varying field humidity conditions.

Among soil-related variables, the *volumetric water content* at both 15 cm (0.51 ± 0.01) and 30 cm (0.52 ± 0.02) depths ranks next in importance. These features indicate that soil moisture availability contributes to predictive accuracy, likely by capturing variations in water supply that influence the growth rate and stress tolerance of the plants. The similar magnitude of the two depths suggests that both shallow and deeper root-zone moisture are relevant to the Ironman cultivar's water uptake.

All remaining environmental variables, including air temperature, soil temperature, humidity, wind speed, radiation, and precipitation, show only a negligible influence on the model's predictions. This indicates that within the relatively uniform and moderate weather conditions of the Verdonk field, these factors contribute little additional information to the prediction of head diameter.

6.6. Validation of T_{base} and T_{max}

To determine the optimal values of T_{base} and T_{max} in the thermal time calculation, a grid search was performed across a predefined range of possible thresholds. The base temperature T_{base} was varied from $0℃$ to $10℃$ in increments of $1℃$. Similarly, the maximum temperature T_{max} was varied from $20℃$ to $32℃$ in increments of $1℃$. For each combination of these parameters, daily minimum and maximum temperatures were clipped to the corresponding thresholds, and the resulting growing degree days were calculated. Thermal time was then computed for each plant over the growing period. To correct for the unobserved early growth phase before the first measurement, the offset method described in Section 3.2.3 was applied, ensuring consistent temporal alignment of all growth trajectories.

For each pair of T_{base} and T_{max}, the four growth models that use only thermal time as an explanatory variable (baseline, quadratic, logistic, and Grevsen) were fitted. A 10-fold cross-validation was applied, and the MAE was calculated for each model across folds. The overall performance for each parameter combination was determined as the mean MAE averaged across all four models. The pair of T_{base} and T_{max} values with the lowest mean MAE was selected as the optimal parameter combination.

The results of the grid search identified an optimal T_{base} of $3℃$ and a T_{max} of $23℃$. These values yielded the lowest average MAE across the four evaluated models, indicating that this temperature

range best represents the biologically active window for broccoli growth within the observed field conditions. These findings are consistent with previous studies on temperature-dependent growth of broccoli and related Brassica species, where reported T_{base} values for broccoli typically range between $0℃$ and $7℃$ [40, 15, 10, 12]. Likewise, the found T_{max} of $23℃$ lies close to the reported value of $26℃$ in the study by Dufault [12]. The identified thresholds, therefore, fall well within the range observed in existing literature, supporting the plausibility of the obtained values.

6.7. Interpretations and Limitations

This section interprets the results of the field-based study in the context of the research goals and examines the limitations that affect their validity and generalisability. The discussion connects model performance to practical applicability, assesses how dataset size and environmental variability influence the findings, and reflects on methodological constraints in both data acquisition and modeling.

6.7.1. Predictive Accuracy

Before focusing on the MLP specifically, it is important to briefly position its performance relative to the other evaluated models. Parametric approaches, including the baseline linear, quadratic, logistic, and Grevsen model, captured the global growth trend but showed larger errors and reduced flexibility in modeling variability and environmental interactions. The LSTM model did not outperform the static MLP, likely due to the limited importance of previous days with respect to the growth rate of broccoli. Across both studies, the MLP achieved the lowest mean error, justifying the focus on this model in the following discussion.

The MLP model achieved an MAE of 0.58 cm on the Verdonk dataset, with a standard deviation of only 0.04 cm across 10 cross-validation folds. This low variability demonstrates that the model generalizes consistently across plants and measurement subsets, indicating strong robustness under field conditions.

To evaluate the practical meaning of this error, it can be related to the average daily head diameter increase. In the Verdonk field data, broccoli heads expanded by approximately 1.06 cm per day, while in the external Tanashi dataset, this growth rate averaged 0.68 cm per day. The residual error of 0.58 cm, therefore, corresponds to around half a day of growth under Verdonk conditions and around one day of growth under the slower Tanashi conditions. This prediction error is small enough for harvest planning applications, where harvest decisions are typically made on a multi-day basis rather than at the level of individual hours. In the Verdonk study, for example, three harvest events occurred within a seven-day window, indicating that the model's average error of less than one day of growth is sufficiently precise to support reliable harvest timing estimation.

The MLP model achieved an MAE of 0.58 cm on the field-based dataset, which is very accurate when considered in the context of commercial broccoli production. Harvest-ready broccoli typically exceeds a head diameter of 12 cm, implying that the residual model error represents less than five percent of the final size. From a practical perspective, such an error margin is acceptable for field-level decision-making tasks, including harvest scheduling and yield estimation, where predictions are aggregated across large plant populations. The model closely follows the overall growth trajectory, which is particularly relevant for population-scale forecasting. Nevertheless, small systematic offsets between predicted and measured diameters were observed for individual plants. These deviations likely result from variations in individual plant characteristics, such as genetic variation or slight measurement noise.

Additionally, prediction accuracy depends not only on the trained model itself but also on the accuracy of its input variables. Environmental parameters such as air temperature, humidity, soil moisture, and solar radiation are subject to measurement noise and forecast uncertainty, which directly propagate into the predictions. Short-term forecasts, specifically temperature, which is the most important input, typically maintain relatively high input accuracy and therefore support reliable growth predictions. However, as predictions are extended to more days ahead, forecast errors in weather variables accumulate and increase total prediction uncertainty, gradually reducing the reliability of diameter estimates. In this work, such uncertainty propagation was not explicitly modeled, but is possible using stochastic modeling.

The measurement uncertainty in the diameter estimation process remains a contributing factor to the observed prediction error. The standard deviation of the median of the upper 50th percentile mea-

surements, from which the final diameter values were derived, was 0.14 cm, corresponding to about 24% of the MLP's overall MAE. This indicates that while measurement noise accounts for only part of the residual error, it still contributes meaningfully to the total uncertainty and helps explain the remaining variability that the model cannot capture.

Moreover, the estimation pipeline assumes that broccoli heads can be approximated as perfectly circular objects. In reality, head shapes deviate from this idealized geometry, introducing an approximation error that is not reducible through improved modeling alone. This further highlights that the reported MAE reflects a combined effect of modeling limitations, sensor noise, and geometric assumptions.

An important practical consideration is that harvesting decisions are based on head weight rather than diameter, and the economic value of broccoli does not increase linearly with size. Heads below approximately 500 g are not suitable for the fresh market and therefore represent no commercial value, while heads that grow beyond roughly 700 g imply a loss of usable product, as the excess mass is economically wasted. In practice, the optimal harvest target lies just above the 500 g threshold, maximizing marketable yield while minimizing overgrowth. As such, an underestimate may delay harvest and push heads toward the overgrowth region, while an overestimate may trigger harvest too early, resulting in sub-threshold heads that cannot be sold. As a result, the risk of prediction errors is most pronounced for plants close to the 500 g threshold, where the margin for error is minimal.

Overall, these findings demonstrate that the MLP provides accurate and consistent growth predictions under field conditions. Its low cross-validation variability, small relative error compared to daily growth, and limited influence of measurement noise collectively indicate that the model performs close to the practical limit of attainable accuracy for this dataset. Therefore, the MLP can be considered sufficiently reliable for operational use in harvest scheduling and yield forecasting, offering a robust foundation for data-driven decision support in precision broccoli production.

6.7.2. Model Selection and Practical Applicability

Although the MLP model demonstrates the highest predictive accuracy among all tested approaches, its implementation requires the integration of multiple environmental variables, which are obtained through an API connection to external meteorological databases and a climate pole. This dependency on multiple data sources introduces potential challenges in operational settings. In particular, API-based access to high-frequency weather data may involve subscription costs. Secondly, a climate pole or similar sensor should be bought to provide accurate measurements of soil moisture. Therefore, while the MLP model is the most accurate, its practical deployment must consider both data accessibility and economic feasibility.

The primary variable that benefits from direct local sensing is soil moisture, which is not reliably captured by regional meteorological services and shows field-specific variability. However, even for soil moisture, measurements from nearby sensor installations or representative field probes may provide sufficiently accurate estimates for predictive purposes, particularly when deploying models over larger production areas. As such, the need for an on-site climate pole must be weighed against cost considerations and the required level of prediction accuracy.

For theoretical cases where data connections are constrained, alternative models provide suitable compromises between complexity and accuracy. The two-dimensional logistic model represents a strong secondary option. It requires only two environmental inputs, namely thermal time and soil moisture, and still achieves reasonable predictive accuracy. Because of the reduction of input variables and the known formula, the model is biologically more interpretable, and its parameters make it more transparent.

In minimal sensor setups where only temperature information is available, the quadratic model provides a practical fallback solution. Although it lacks the complex modeling capacity of the MLP and the environmental sensitivity of the logistic2D model, its simplicity enables robust predictions with minimal data requirements.

In this context, the choice between models depends on the trade-off between available sensor infrastructure and the desired prediction accuracy. For precision agriculture applications where real-time weather data can be reliably accessed, the MLP remains the preferred model. In contrast, for low-cost or small-scale operations, the logistic2D or quadratic model may be more appropriate.

6.7.3. Generalization and Data Limitations

The conclusions drawn from this study are based on a dataset representing a single cultivar, grown within one specific field under relatively uniform weather conditions. While this controlled setup enabled the systematic evaluation of model behavior and reduced confounding factors, it also limits the generalizability of the findings. The environmental variability during the measurement period was modest, with only small fluctuations in temperature, humidity, and soil moisture. As a result, the models were primarily exposed to growth conditions that remained close to the optimal range for broccoli cultivation. Under such stable circumstances, differences between models may appear less pronounced, since extreme environmental effects, such as heat stress or droughts, were largely absent.

This limited variability implies that the models were validated within a narrow climatic window, and their robustness under more diverse or challenging conditions remains uncertain. Diameter sizes may behave differently when extended to broader temperature ranges or different soil moisture levels. Therefore, while the MLP model achieved the lowest prediction error within this dataset, its better performance should be understood in the context of the relatively stable conditions of this study, rather than as evidence of universal applicability across all growing environments. Some confidence in this finding can nevertheless be derived from the fact that the same conclusion was reached in the preliminary study, which was conducted under different environmental and geographical conditions and also identified the MLP as the best-performing model.

7

Conclusions

This study set out to investigate how predictive models based on field measurements can be used to describe broccoli growth under Dutch growing conditions. This was addressed through four research questions focusing on (1) the construction of plant-level growth trajectories from field video, (2) the influence of environmental factors on head development, (3) the representation of temperature through cumulative temperature versus thermal time, and (4) the comparative performance of different modeling approaches. The conclusions synthesise the findings from both the preliminary study and the Verdonk field study.

7.1. Field-Based Data Processing Pipeline

The first research question describes how field-based video recordings can be converted into plant-level growth curves that are suitable for modeling. This process involves several consecutive steps, starting with automated head detection and ending with the integration of environmental data. A YOLOv8n-based model, retrained on field images from the Verdonk dataset, accurately detected broccoli heads across developmental stages, achieving an F1-score of 0.958. To ensure consistent plant identification across frames and measurement days, a novel slot-based ID system was developed using camera motion estimation, effectively linking detections to fixed plant positions.

Detected heads were then tracked over time to form continuous trajectories, corrected for missed links and drift using velocity-based repair and slot boundary refinement. Each detection was used to estimate physical head diameter based on RGB-D data, corrected for depth curvature, and filtered to remove outliers. This reduced measurement variability and produced stable diameter estimates. A series of filtering steps ensured data quality by removing implausible decreases in diameter and abnormal growth slopes.

Finally, in-field weather data from a climate pole were combined with external meteorological data retrieved from the Open-Meteo API. These were aggregated into daily values and used to calculate derived variables such as vapor pressure deficit and thermal time. The resulting dataset provides accurate, temporally aligned growth trajectories that form the foundation for the predictive modeling of broccoli development.

7.2. Environmental Parameters

The second research question examined which environmental parameters most strongly influence broccoli growth dynamics. This was addressed through permutation feature importance analysis, conducted for both the preliminary study and the Verdonk field dataset.

Across both datasets, Thermal Time emerged as by far the most influential predictor. In the preliminary study, it achieved an importance score of 3.69, while a similarly dominant contribution was observed in the Verdonk study at 3.27. This consistent result confirms that accumulated temperature within biologically active thresholds is the primary driver of broccoli head development, and that temperature-based progression governs growth more strongly than any other measured variable.

In the preliminary study, all remaining environmental features showed only marginal importance. Maximum temperature, shallow soil moisture, and mean humidity contributed weakly, with importance

scores between 0.28 and 0.35, and all other factors, including soil temperature, radiation, and precipitation, were negligible. This indicated that while short-term variations in water balance or atmospheric stress may influence growth locally, these effects were not as pronounced as temperature-driven variables.

The Verdonk field study revealed a more differentiated pattern. While Thermal Time remained dominant, VPD emerged as the second most influential feature with an importance of 1.02, indicating a measurable role of atmospheric moisture demand in regulating transpiration and stomatal behavior. In addition, soil moisture at both 15 cm and 30 cm depths contributed moderately to predictive performance (0.51 and 0.52, respectively), suggesting that water availability across the root zone affects growth rate and stress tolerance of the Ironman cultivar under field conditions. The similar importance values at both depths imply the relevance of both surface and deeper soil water reserves for plant uptake.

As in the preliminary study, all remaining environmental variables showed negligible importance in the Verdonk analysis. Their limited contribution reflects the relatively stable and moderate environmental conditions during the single observed growing season, which restricted the range of stress responses that could be captured.

Taken together, the feature importance results from both datasets demonstrate that broccoli growth is mainly controlled by temperature accumulation, as represented by Thermal Time. Secondary influences from humidity-related stress and soil moisture contribute only slightly to the growth variation. Other environmental parameters play only a minor role in explaining growth dynamics within the conditions examined in this thesis.

7.3. Thermal Time vs. Cumulative Temperature

The third research question compared cumulative temperature and thermal time to determine which measure better describes the biological growth process. Both approaches were evaluated using identical model configurations, and their prediction errors were compared through cross-validation. The results show that models based solely on temperature, such as the baseline, quadratic, logistic, and Grevsen models, significantly performed better when using thermal time. This improvement confirms that applying biologically meaningful temperature thresholds enhances predictive accuracy by excluding non-contributory temperature extremes.

For models that also incorporated additional environmental variables, such as the MLP, LSTM, and logistic2D models, the difference between the two temperature representations was not statistically significant. This reduced effect can be explained by the inclusion of other weather variables like humidity, radiation, and soil moisture, which already capture part of the environmental variability related to growth. In particular, the logistic2D model's improvement is attributed to the higher accuracy of the soil moisture sensor data in the Verdonk field when compared to the preliminary study, which uses API data for this variable and is therefore less accurate.

In addition, the optimal temperature thresholds for the thermal time calculation were validated through a grid search across a range of possible values for T_{base} and T_{max}. The best-performing combination was found at $T_{\text{base}} = 3°\text{C}$ and $T_{\text{max}} = 23°\text{C}$, representing the biologically active window for broccoli growth under the measured field conditions. These values align with previously reported temperature ranges for broccoli cultivars, confirming that the chosen thresholds are physiologically realistic and well-suited for modeling growth within temperate climates.

In summary, thermal time provides a more biologically grounded and effective temporal variable for modeling broccoli growth. It improves accuracy for models driven mainly by temperature inputs, while its advantage becomes less pronounced in multi-variable models where additional environmental factors already capture much of the relevant physiological variation.

7.4. Model Comparison

The fourth research question focused on the ability of different models to represent the growth dynamics of broccoli. To evaluate this, eight models were compared using 10-fold cross-validation on the field-based dataset. The results show clear differences in predictive accuracy, with the MLP achieving the lowest mean absolute error of 0.583 and demonstrating high robustness across folds. Its performance can be attributed to its ability to capture complex non-linear relationships between multiple environmental variables, relying only on the current weather state without requiring temporal sequences. This

outcome closely mirrors the findings from the preliminary study, where the MLP likewise outperformed parametric growth curves, indicating that its flexibility generalizes beyond the external dataset to real field conditions.

In contrast, the LSTM model, which uses the same environmental variables as the MLP but additionally incorporates historical weather information, performed worse. This behavior was also observed in the preliminary study, where the inclusion of sequential information did not lead to improved predictive accuracy. Together, these results indicate that temporal dependencies between consecutive measurement days do not enhance growth prediction for broccoli under the studied conditions. The current environmental state, combined with a cumulative temperature representation, appears sufficient to describe plant development, and adding historical observations introduces additional noise rather than valuable temporal structure.

Overall, the consistent ranking of model performance across both the preliminary and the Verdonk field studies demonstrates that the MLP model provides the most accurate and robust representation of broccoli growth dynamics. This agreement between studies strengthens confidence in the validity of the modeling approach and confirms that temporal modeling offers no advantage in this application, while growth can be effectively predicted using instantaneous environmental inputs alongside a cumulative-based temperature feature.

7.5. Implications

The outcomes of this thesis have several direct implications for the practical deployment of predictive monitoring systems within precision agriculture. First, the performance of thermal time as the primary growth driver, confirmed in both the preliminary study and the Verdonk field study, implies that reliable growth prediction can be achieved with relatively limited sensor inputs. Accurate air temperature data alone already enables baseline performance, while the inclusion of a small set of additional variables, particularly vapor pressure deficit and soil moisture, further improves predictive accuracy. This indicates that fully instrumented, high-cost infrastructures are not a strict requirement for operational application, and that many farms could make use of predictive modeling through publicly available meteorological data complemented by a small number of low-cost field sensors.

Second, the demonstrated performance of the MLP model over both parametric and temporal neural models implies that practical systems do not need to rely on complex sequence-based architectures or high-frequency historical data streams. Growth predictions can be generated using daily environmental data combined with cumulative temperature metrics. This substantially lowers computational requirements and eases system integration, making it feasible to run prediction pipelines on standard agricultural data platforms without specialized hardware or advanced machine-learning infrastructure.

Third, the ability to extract stable, plant-level diameter trajectories from field video demonstrates that non-destructive crop monitoring can realistically support selective harvesting strategies. Accurate head-size predictions enable growers to move beyond field-wide cutting approaches and toward more targeted harvest scheduling, either by identifying sections of fields that are ready for harvest or by enabling size-based selection in mechanized harvesting systems. This has the potential to reduce premature harvesting losses and decrease the volume of produce rejected for being undersized or overmature.

Finally, the low dependence of the best-performing models on dense sensor networks and temporal modeling increases the scalability of such systems. Predictive services could be extended to multiple fields or farms using shared regional weather data sources while maintaining acceptable accuracy. This makes the approach accessible not only to large commercial producers but also to smaller operations, supporting broader adoption of data-driven decision support tools across the sector.

Together, these implications show that field-based predictive growth modeling is not only methodologically feasible but also practically deployable. With limited additional investment in sensing and data integration, growers can use predictive models to support harvest planning, mechanized selection, and general crop monitoring, thereby improving both efficiency and marketable yield.

8

Future Work

While this research demonstrates that field-based predictive modeling can provide accurate estimates of broccoli head growth, several directions remain to further enhance both the precision and applicability of such models in practical agricultural settings. The following subsections discuss different suggestions for future work concerning data acquisition, computer vision, and optimization.

8.1. Improving Diameter Estimation through Segmentation

The current method for diameter estimation relies on the depth measured at the center of the broccoli crown. This approach introduces a systematic underestimation due to the curvature of the broccoli head, where the true diameter corresponds to a depth at the contour rather than at the center. A segmentation-based method would mitigate this error by explicitly extracting the head boundary and determining the diameter along the contour points using the depth map. Such an approach would also allow partial head reconstruction in cases of occlusion, where the full crown is not visible, by estimating the radius from a visible subset of the head.

A similar approach by Blok et al. is to use a learned model that directly extracts object size from depth data by first segmenting the crop in the RGB image with a segmentation network and then fitting a circle to the resulting mask, converting the pixel radius to a diameter through depth values from the depth image [4]. This method, demonstrated for broccoli heads, shows that combining RGB segmentation with depth-based geometric fitting can provide accurate diameter estimates even under substantial occlusion. Using these methods, we can get a more reliable estimation of the broccoli head diameter when occlusion by, for example, leaves takes place.

8.2. Time Interval of Field Measurements

The measurement period of the Verdonk field experiment covered 14 days from the beginning of head initiation to harvest, with only seven distinct measurement days. While this was considered sufficient for constructing growth trajectories, a higher measurement interval would enable a more detailed characterization of growth dynamics and improve model calibration. Increasing the number of measurement days would better capture variations and allow for finer validation of model predictions. However, this would also increase the operational workload, as each measurement day requires the presence of a human operator in the field. Future research could therefore explore the integration of automated or semi-autonomous data collection systems to increase frequency without additional manual labor.

8.3. Autonomous Drones

An ideal precision agriculture workflow would replace human camera operators with automated image-collection platforms capable of operating independently across production fields, identifying points of interest, and enabling continuous large-scale monitoring while reducing labor demands. Such data acquisition does not necessarily need to rely exclusively on drones. Alternative approaches include cameras mounted on standard agricultural machinery performing routine field operations, as well as satellite-based Earth observation systems that provide regular, large-area coverage. Each of these

platforms can, in principle, support the generation of visual crop monitoring data required for growth modeling. However, it would require high-definition imagery, which in satellite observations might not be operational and economically feasible.

Drones represent one particularly flexible option within this broader monitoring landscape, as they can offer high spatial resolution and targeted data collection. Stakeholders in the agricultural sector already expect increasing levels of autonomous aerial sensing to become necessary for efficient and sustainable production in the near future [31]. In addition, an existing drone-based pipeline has been developed by Wang et al. to measure the head diameters of broccoli for entire fields [37], demonstrating the technical feasibility of this approach.

However, the analysis in [31] shows that the Dutch regulations do not yet permit this level of autonomy. Fully autonomous missions fall within the Specific category, requiring operator training, registration, an operational manual, and a Specific Operations Risk Assessment (SORA). These requirements are manageable, but the critical barrier is the prohibition of Beyond Visual Line of Sight (BLOS) operations without direct pilot oversight. As a result, current agricultural drones remain automated rather than autonomous. They follow pre-programmed paths while a pilot keeps them within visual range, which makes the replacement of human operators still economically and operationally unfeasible.

There is, however, some encouraging progress. Recent developments show that at least one operator in the Netherlands has received authorization to fly multiple agricultural drones simultaneously within visual range as part of a broader precision-agriculture innovation initiative [7]. This suggests that regulation and industry are beginning to adapt, which could open a pathway toward more scalable image-based crop monitoring systems in the near future. If regulatory evolution continues and technical safeguards improve, the economic feasibility of automated or semi-automated aerial monitoring may become viable for a broader range of farms.

8.4. Expanding to Multiple Fields and Weather Conditions

The current study was limited to a single field and growth cycle, resulting in homogeneous environmental conditions. Within the Verdonk dataset, the observed weather ranges were narrow: daily mean air temperatures varied approximately between 11.7 - 24.4 °C, vapor pressure deficit ranged from 0.36–0.54 kPa, and volumetric soil moisture fluctuated between 23.3 - 39.7% at 15 cm and 26.4 - 38.0% at 30 cm depth. Mean solar radiation remained within 226.4 - 305.4 W/m^2, and no prolonged periods of extreme heat, drought stress, or excessive humidity were observed. These conditions are representative of a typical Dutch summer growing season but do not capture the full spectrum of weather extremes reported in earlier literature. As a result, the predictive models developed in this thesis are effectively validated only within this limited environmental domain.

To improve generalization and capture a broader range of weather effects, future studies should include measurements from multiple fields, ideally representing different geographic locations and seasonal conditions. Measuring the same cultivar in different climates would help improve how the models capture environmental effects and could show how it copes with non-ideal conditions.

With multiple fields, additional analyses would become possible. Cross-field validation could be used to test how well a model trained under one set of conditions performs in another, providing insight into the generalizability of both mechanistic and data-driven models. Moreover, the broader variability in weather would allow a more detailed sensitivity analysis of environmental factors, revealing how variables such as vapor pressure deficit and soil moisture interact under stress conditions. Finally, fitting models across several environments would make it possible to evaluate whether parameters such as T_{base} and T_{max} vary between regions or seasons, leading to a more robust and transferable formulation of thermal time.

8.5. Measurement Setup

Future improvements to the field measurement setup should focus on enhancing both the reliability of plant detection and the spatial accuracy of the recorded data. Two key aspects that warrant attention are illumination control and spatial referencing.

The detection model used in this study occasionally struggled to identify broccoli heads under challenging lighting conditions, particularly during the early growth stages when the developing head remains partially hidden since it sits low in the plant and surrounding leaves block sunlight. In such cases, the center of the plant is often shaded, resulting in insufficient contrast between the head and the sur-

rounding foliage. These illumination issues led to missed or uncertain detections in several recordings. Adding an external light source near the cameras would mitigate this limitation by providing a more stable and uniform lighting condition across all measurement sessions. Such illumination would allow the detection model to more reliably identify small or occluded heads during head initiation, thereby improving the consistency and completeness of the diameter dataset.

In addition to improving lighting conditions, the spatial accuracy of plant tracking can be enhanced through the integration of a GPS module. In the current setup, the distance travelled by the operator is estimated using the ECC algorithm, as described in Section 5.4. Although this method provides an approximate measure of displacement, it is susceptible to cumulative error and drift. The inclusion of a GPS unit would provide a reduction in the uncertainty of the estimated camera motion and enabling more reliable assignment of plant identities across measurement days.

Bibliography

[1] Ayasha Akter, Shapla Akhter, and Md Mokter Hossain. Water stresses alter the growth, yield and quality of broccoli (Brassica Oleracea var. italica). *Scientific Reports*, 15(1), 12 2025.

[2] Bashar Alhnaity, Simon Pearson, Georgios Leontidis, and Stefanos Kollias. Using Deep Learning to Predict Plant Growth and Yield in Greenhouse Environments. *CoRR*, abs/1907.00624, 2019.

[3] Pieter M. Blok, Ruud Barth, and Wim van den Berg. Machine vision for a selective broccoli harvesting robot. *IFAC-PapersOnLine*, 49(16):66–71, 10 2016.

[4] Pieter M. Blok, Eldert J. van Henten, Frits K. van Evert, and Gert Kootstra. Image-based size estimation of broccoli heads under varying degrees of occlusion. *Biosystems Engineering*, 208:213–233, 8 2021.

[5] Davide Cammarano, Mark A. Taylor, Jacqueline A. Thompson, Gladys Wright, Andrew Faichney, Richard Haacker, Andrew Orr, and Philip J. White. Predicting dates of head initiation and yields of broccoli crops grown throughout Scotland. *European Journal of Agronomy*, 116:126055, 5 2020.

[6] Jessica N. Castillo and Jose R. Muñoz. Mathematical Model for Broccoli Growth Prediction Based on Artificial Networks. *Lecture Notes in Electrical Engineering*, 984 LNEE:845–859, 2023.

[7] Wiebe de Jager. Dutch drone operator first to receive approval for simultaneous use of three agricultural drones, 11 2025. Accessed: 2025-11-30.

[8] Isabel De Maria Mourão and Luís Miguel Brito. Empirical models for harvest date prediction in Broccoli (Brassica oleracea L. var. italica Plenck). *Acta Horticulturae*, 539:47–53, 2000.

[9] Juping Ding, Xiaocong Jiao, Ping Bai, Yixin Hu, Jiayu Zhang, and Jianming Li. Effect of vapor pressure deficit on the photosynthesis, growth, and nutrient absorption of tomato seedlings. *Scientia Horticulturae*, 293:110736, 2 2022.

[10] M.T. Diputado Jr and M.A. Nichols. The effect of sowing date and cultivar on the maturity characteristics of broccoli (Brassica oleraceae var. Italica)). *Acta Horticulturae*, 247(247):59–66, 9 1989.

[11] Lukas Drees, Laura Verena Junker-Frohn, Jana Kierdorf, and Ribana Roscher. Temporal Prediction and Evaluation of Brassica Growth in the Field using Conditional Generative Adversarial Networks. *CoRR*, abs/2105.07789, 5 2021.

[12] Robert J. Dufault. Determining Heat Unit Requirements for Broccoli Harvest in Coastal South Carolina. *Journal of the American Society for Horticultural Science*, 122(2):169–174, 3 1997.

[13] Jane R. Fellows, R. J. Reader, and D. C.E. Wurr. A model for leaf production and apex development in calabrese. *Journal of Horticultural Science*, 72(2):327–337, 1997.

[14] Greenports Nederland. Nationale tuinbouwagenda 2019–2030: Circular horticulture in practice – the priorities of the horticulture cluster united in greenports nederland. Technical report, Greenports Nederland, 2019. First edition, June 2019.

[15] K. Grevsen. Effects of temperature on head growth of broccoli (Brassica oleracea L. var. italica): Parameter estimates for a predictive model. *Journal of Horticultural Science and Biotechnology*, 73(2):235–244, 1998.

[16] Charlotte Grossiord, Thomas N. Buckley, Lucas A. Cernusak, Kimberly A. Novick, Benjamin Poulter, Rolf T.W. Siegwolf, John S. Sperry, and Nate G. McDowell. Plant responses to rising vapor pressure deficit. *New Phytologist*, 226:1550–1566, 6 2020.

[17] Laurie Hodges and James R. Brandle. Windbreaks for fruit and vegetable crops. Extension Circular EC1779, University of Nebraska–Lincoln Extension, Lincoln, NE, September 2006. Revised September 11, 2006.

[18] Ehab A. Ibrahim, Noura E.S. Ebrahim, and Gehan Z. Mohamed. Mitigation of water stress in broccoli by soil application of humic acid. *Scientific Reports*, 14:1–10, 12 2024.

[19] Konstantinos Koularmanis, Pavlos Tsouvaltzis, and Anastasios Siomos. Impact of elevated temperature and solar radiation on broccoli (brassica oleraceae var. italica plenck) cultivation. *Horticulturae 2025, Vol. 11, Page 187*, 11:187, 2 2025.

[20] Joyce G. Latimer. Drought or Mechanical Stress Affects Broccoli Transplant Growth and Establishment but Not Yield. *HortScience*, 25(10):1233–1235, 10 1990.

[21] Cheng Ju Lee, Ming Der Yang, Hsin Hung Tseng, Yu Chun Hsu, Yu Sung, and Wei Ling Chen. Single-plant broccoli growth monitoring using deep learning with UAV imagery. *Computers and Electronics in Agriculture*, 207:107739, 4 2023.

[22] Cherubino Leonardi, Soraya Guichard, and Nadia Bertin. High vapour pressure deficit influences growth, transpiration and quality of tomato fruits. *Scientia Horticulturae*, 84:285–296, 6 2000.

[23] Karsten Lindemann-Zutz. *Head size variation within broccoli (Brassica oleracea var. italica) plantings, causes and prediction for decision support*. Doctor of horticultural sciences (dr. rer. hort.), Gottfried Wilhelm Leibniz Universität Hannover, Hannover, Germany, 2015. Dissertation approved by the Faculty of Natural Sciences.

[24] Karsten Lindemann-Zutz, Andreas Fricke, and Hartmut Stützel. Prediction of time to harvest and its variability of broccoli (Brassica oleracea var. italica) part II. Growth model description, parameterisation and field evaluation. *Scientia Horticulturae*, 200:151–160, 3 2016.

[25] Mykhailo Lohachov, Ryoji Korei, Kazuo Oki, Koshi Yoshida, Issaku Azechi, Salem Ibrahim Salem, and Nobuyuki Utsumi. RNN-Based Approach for Broccoli Harvest Time Forecast. *Agronomy 2024, Vol. 14, Page 361*, 14(2):361, 2 2024.

[26] B. Marshall and R. Thompson. A Model of the Influence of Air Temperature and Solar Radiation on the Time to Maturity of Calabrese Brassica oleracea var italica. *Annals of Botany*, 60(5):513–519, 11 1987.

[27] A. R. Maurer. Response of broccoli to five soil water regimes. *Canadian Journal of Plant Science*, 56:953–959, 1976. Contribution no. 228, received 14 May 1976, accepted 12 July 1976.

[28] Conrad H. Miller. Diurnal Temperature Cycling Influences Flowering and Node Numbers of Broccoli. *HortScience*, 23(5):873–875, 9 2022.

[29] John L. Monteith and Mike H. Unsworth. Steady-State Heat Balance: (i) Water Surfaces, Soil, and Vegetation. *Principles of Environmental Physics*, pages 217–247, 1 2013.

[30] Rijksoverheid Nederland. Klimaatakkoord. Technical report, Rijksoverheid, June 2019.

[31] Chelsey Oldenburger. Vliegend in het grijze gebied: De juridische knelpunten van autonoom dronevliegen. Master's thesis, Hogeschool Inholland Rotterdam, Rotterdam, Netherlands, 2025.

[32] Jørgen E. Olesen and Kai Grevsen. Effects of temperature and irradiance on vegetative growth of cauliflower (brassica oleracea l. botrytis) and broccoli (brassica oleracea l. italica). *Journal of Experimental Botany*, 48:1591–1598, 8 1997.

[33] Michael Schmid, David Rath, and Ulrike Diebold. Why and How Savitzky-Golay Filters Should Be Replaced. *ACS Measurement Science Au*, 2(2):185–196, 4 2022.

[34] Anastasios S. Siomos, Konstantinos Koularmanis, and Pavlos Tsouvaltzis. The impacts of the emerging climate change on broccoli (brassica oleracea l. var. italica plenck.) crop. *Horticulturae 2022, Vol. 8, Page 1032*, 8:1032, 11 2022.

[35] D. K.Y. Tan, C. J. Birch, A. H. Wearing, and K. G. Rickert. Predicting broccoli development I. Development is predominantly determined by temperature rather than photoperiod. *Scientia Horticulturae*, 84(3-4):227–243, 6 2000.

[36] L. R. Walton and J. H. Casada. Evaluation of Broccoli Varieties for Mechanical Harvesting. *Applied Engineering in Agriculture*, 4(1):5–7, 1988.

[37] Haozhou Wang, Tang Li, Erika Nishida, Yoichiro Kato, Yuya Fukano, and Wei Guo. Drone-Based Harvest Data Prediction Can Reduce On-Farm Food Loss and Improve Farmer Income. *Plant Phenomics*, 5:0086, 1 2023.

[38] F. H. Whitehead. Experimental studies of the effect of wind on plant growth and anatomy: Iii. soil moisture relations. *New Phytologist*, 62:80–85, 1963.

[39] World Agritech. Roboveg unveils autonomous broccoli harvester, 10 2021.

[40] D. C. E. Wurr, Jane R. Fellows, and Angela J. Hambidge. The influence of field environmental conditions on calabrese growth and development. *Journal of Horticultural Science*, 66(4):495–504, 1 1991.

[41] D. C.E. Wurr, J. R. Fellows, K. Phelps, and R. J. Reader. Vernalization in calabrese (Brassica oleracea var. italica) - a model for apex development. *Journal of Experimental Botany*, 46(10):1487–1496, 10 1995.