Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Mechanistic-Informed Statistical Modelling for Pre-Harvest Apple Firmness Prediction Using Sensor-Based Data

A thesis submitted for the degree of

Master of Information Science

by

Ruizhe Yang



College of Sciences
Massey University
New Zealand

July 2026

Acknowledgements

First of all, I would like to express my profound gratitude to my main supervisor, Dr Mo Li, for her guidance, encouragement, and support. Her clear explanations, constant optimism, and immense patience were essential to my project. I also sincerely appreciate her careful revision of this thesis.

I am deeply thankful to my co-supervisor, Dr Matthieu Vignes, as well as Dr Nicolette Niemann, Dr Junqi Zhu, and Ms Yongqi Liang, for their scientific guidance, practical advice, encouragement, and support. I also deeply appreciate their revision of this thesis.

I would also like to thank my parents for their unconditional support during my graduate study. Finally I would also especially like to thank my husband, Dr. Xun Xiao, for his patience. Thank you for your spiritual support and company. When I wanted to give up, it was you who gave me encouragement. You have always said that no matter what decision I make, you will always be on my side. Thank you for everything. I love you always.

Statement of Originality

This is to certify that the intellectual content of this thesis is the product of my own work and that all assistance received in preparing this thesis, and all sources used, have been acknowledged.

Ruizhe Yang

Abstract

Accurately predicting pre-harvest apple fruit firmness is important for supporting harvest planning, fruit quality control, and post-harvest management. This thesis develops a mechanistic-informed statistical framework for predicting pre-harvest firmness in 'Royal Gala' apples. The proposed approach combines biological information about apple softening with fruit quality measurements, Near-Infrared (NIR) spectral data, and environmental data. More specifically, a two-phase one-knot spline model is first used to construct a biologically meaningful population-level baseline firmness trajectory over days after full bloom (`DAFB`). Two sequential statistical correction components are then considered: the $h(\cdot)$ model uses fruit-level variables and destructive NIR spectra, while the $g(\cdot)$ model uses date-level environmental information and longitudinal non-destructive NIR spectra. The model was developed using data from the 2021–2022 season. For the held-out cross-season evaluation using the 2024–2025 season, the two-phase decline structure was retained, the first assessment date was used to calibrate the season-specific initial firmness, and the two later assessment dates were used to evaluate prediction performance. The baseline model captured the broad decline in pre-harvest firmness. In the development season, the correction components produced only negligible improvements over the baseline, with RMSE remaining approximately 1.07 *kgf* and RPD remaining below 1.2. In the 2024–2025 evaluation, the environmental $g(\cdot)$ correction provided the best fruit-level performance, improving RMSE from 1.009 to 0.979 *kgf* and R^2 from 0.330 to 0.369. However, the corresponding RPD was only 1.285, indicating limited utility for precise individual-fruit prediction. Prediction errors were smaller when evaluated using assessment-date mean firmness, with an RMSE of 0.272 *kgf* for the environmental correction model, although this result was based on only two post-calibration dates. Overall, the results demonstrate a practical framework for integrating biological information with heterogeneous sensor-based data. The framework shows greater potential for describing population-level pre-harvest firmness development than for predicting individual-fruit firmness. More complete and consistently matched multi-season data are required before the model can be considered sufficiently robust for practical prediction.

Key Words: Apple firmness; Pre-harvest firmness prediction; mechanistic-informed statistical model; Near-Infrared spectral data; residual correction; environmental data; `Royal Gala`.

Contents

List of Figures **6**

List of Tables **7**

1 Introduction **8**

1.1 Background 8

1.2 Synopsis of New Contribution 12

2 Literature Review **13**

2.1 Apple Fruit and Economic Significance 13

2.2 Pre-harvest Apple Fruit Development 14

2.3 Factors Influencing Apple Firmness Development 15

2.4 Near-Infrared Spectroscopy for Apple Firmness Assessment 19

2.4.1 Introduction 19

2.4.2 Chemometric Methods for NIR Analysis 22

2.5 Existing Models for Apple Firmness Prediction 25

2.6 Research Gap and Motivation 28

2.7 Research Questions and Methodological Objective 29

3 Model Development Framework **30**

3.1 Data Scope and Alignment Challenge 30

3.2 Data Sources, Collection, and Availability 33

3.2.1 Data Collection Design and Measurement Procedures 33

3.2.2 Data Availability, Matching, and Seasonal Exclusions 36

3.2.3 Candidate Variable Availability and Selection 36

3.3 Notations for Model Output and Input . . . 37
3.3.1 Response of Interest . . . 37
3.3.2 Potential Model Input from Raw Data . . . 40
3.3.3 Data Alignment Challenges . . . 44
3.3.4 Model Formulation . . . 44
3.3.5 Fusion Specification . . . 46

4 Model Implementation **49**
4.1 Data Pre-processing and Validation . . . 50
4.1.1 Fruit Quality Measurement Data Cleaning . . . 50
4.1.2 NIR Spectral Data Validation and Pre-processing . . . 51
4.1.3 Environmental Data Pre-processing . . . 56
4.2 Model Development Based on the 2021-2022 Season . . . 57
4.2.1 Exploratory Modelling Using Generalised Additive Models . . . 57
4.2.2 Baseline Modelling . . . 59
4.2.3 Residual Correction Modelling . . . 61
4.3 Cross-Season Prediction Performance Evaluation Using Data Collected from the 2024–2025 Season . . . 65

5 Results and Comparisons **70**
5.1 Model Development Results and Comparisons . . . 70
5.2 Prediction Performance Evaluation Results for the 2024–2025 Season . . . 77

6 General Conclusions and Future Perspectives **86**
6.1 Conclusions . . . 86
6.2 Limitations and Future Work . . . 89

Bibliography **93**

List of Figures

2.1 Diagram of the factors influencing apple fruit firmness from a cellular perspective 17

2.2 Diagram of four factors influencing apple fruit firmness at harvest and during development. 20

3.1 Raw pre-harvest firmness measurements for the 2021–2022 season 31

3.2 Raw pre-harvest firmness measurements for the 2024–2025 season 32

3.3 Non-destructive NIR spectra from five example fruit records at each assessment date during the 2024–2025 season 35

3.4 Variable relationship diagram for pre-harvest apple fruit firmness prediction . . . 38

3.5 Comparison between simulations and observations of individual sugar components for the 2021–2022 pre-harvest period 39

3.6 Illustration of the response structure (not real data) 41

3.7 Illustration of the temporal alignment of high-frequency environmental variables with firmness assessment times 47

4.1 Predicted versus measured DMC for each pre-processing method, with the corresponding R^2 annotated for each panel in the $2021-2022$ season without the time variable 54

4.2 Predicted versus measured DMC for each pre-processing method, with the corresponding R^2 annotated for each panel during the later period (January–February 2025) in the 2024–2025 season without the time variable 56

5.1 Model development results for the 2021–2022 season 74

5.2 Cross-season prediction performance evaluation results for the 2024–2025 season 78

List of Tables

2.1 Examples of wavelength ranges used in published NIR calibration studies for quality attributes in the "Royal Gala" apple fruit 22

3.1 Summary of data sources, sample sizes, and analytical roles 34

4.1 The number of destructive NIR samples before and after matching (2021–2022 season) 51

4.2 Pre-processing modules used in the NIR analysis 52

4.3 Performance comparison of different pre-processing methods for NIR dry matter content prediction in the $2021 - 2022$ season without the time variable 53

4.4 The number of destructive NIR samples before and after matching (2024–2025 season) 53

4.5 Performance comparison of different pre-processing methods for NIR dry matter content prediction during the later period (January–February 2025) in the 2024–2025 season without the time variable 55

5.1 Measurement dates and sample sizes used in model development 71

5.2 Candidate and retained variables used in model development 72

5.3 Model performance for the 2021–2022 development season 75

5.4 Measurement dates and sample sizes used in the 2024–2025 cross-season evaluation 77

5.5 Predictor structures transferred to the 2024–2025 season 79

5.6 Model performance for the primary 2024–2025 cross-season evaluation 80

5.7 Sensitivity of baseline performance to the treatment of initial firmness 82

5.8 Date-mean prediction errors for the 2024–2025 evaluation 82

Chapter 1

Introduction

1.1 Background

Apple is one of the most important temperate fruit crops worldwide, and fruit quality is a key determinant of its commercial value (Kellerhals, 2009). In New Zealand, apple production is a major component of the export-oriented horticultural industry. For the year ending 30 June 2026, export revenue from apples and pears was forecast to reach NZ$1.3 billion, contributing to a record total horticultural export revenue of NZ$9.5 billion (Ministry for Primary Industries, 2026). This substantial export value highlights the economic importance of maintaining consistently high apple quality to meet international market requirements and sustain the competitiveness of the New Zealand apple industry.

In the apple industry, fruit quality is usually evaluated using a combination of external and internal traits (Harker et al., 2008). More specifically, external apple traits include fruit size, shape, skin colour, and appearance, whereas internal apple traits include firmness, texture, dry matter content (DMC), soluble solids content (SSC), acidity, and flavour (Harker et al., 2008). Among these quality attributes, apple firmness plays an important role in determining market quality and economic value (Musacchi and Serra, 2018) because it is closely related to fruit texture, consumer acceptance, storage ability, and market performance (Goncalves et al., 2017; Harker et al., 2003). For fresh apples, consumers generally prefer apples that are not too soft, since excessive softness is often associated with over-ripeness (Musacchi and Serra, 2018). Therefore, when apples become too soft, the eating experience may decline, and the fruit may be considered less desirable for the fresh market (Harker et al., 2002).

Furthermore, apple fruit firmness is also an important consideration for harvest timing and post-harvest management (Harker et al., 2002; Saei et al., 2011). More generally, apples must retain an acceptable level of firmness throughout storage, distribution, and retail to maintain their market quality (Małachowska and Tomala, 2023). This requirement is particularly important for the New Zealand apple industry because a substantial proportion of its production is exported to geographically distant markets. Consequently, New Zealand apples may undergo extended storage and long-distance transportation before reaching overseas retailers and consumers, making firmness retention an especially important quality objective. For commercial apple production, the ability to monitor and predict firmness before harvest can support harvest planning, quality control, and post-harvest management (Kumar and Singh, 2023).

More specifically, if apples are harvested too late, they may have reduced storage potential and lower market value (Goncalves et al., 2017). However, if apples are harvested too early, other quality traits, such as flavour and colour, may not have fully developed (Dar et al., 2019). Therefore, accurately predicting apple firmness during the pre-harvest period can help growers, researchers, and industry managers balance fruit maturity, quality, and storage requirements before the fruit enters the post-harvest supply chain (Goncalves et al., 2017; Małachowska and Tomala, 2023; Plotto et al., 1997). However, pre-harvest apple firmness prediction is challenging because apple fruit firmness is affected by multiple interacting biological, physical, and environmental processes throughout apple fruit development (Lachapelle et al., 2013; Volz et al., 2002; Chagné et al., 2014).

Moreover, Near-Infrared (NIR) spectral data provide a useful source of sensor-based information for pre-harvest apple firmness prediction. Traditional firmness measurement is destructive, meaning that the same apple cannot be repeatedly measured for firmness during development. In contrast, NIR spectroscopy can provide rapid and non-destructive information about internal fruit properties, such as dry matter content (DMC), soluble solids content (SSC), water status, and tissue-related spectral responses (Costa et al., 2009; Mendoza et al., 2011; Zhang et al., 2019). Therefore, NIR spectral data have the potential to support pre-harvest quality monitoring and provide additional predictors for apple firmness modelling. However, apple firmness is affected by multiple biological and physical processes, and the relationship between NIR spectra and firmness is often indirect and season-dependent (Møller et al., 2013; Xu et al., 2019). Therefore, using NIR spectral data alone may not be sufficient for robust pre-harvest firmness

prediction, and it is useful to combine NIR information with biological understanding of apple fruit softening and environmental factors that influence fruit development.

This study forms part of a broader digital twin programme at Plant & Food Research (PFR). The programme aims to integrate biological knowledge, fruit quality measurements, sensor observations, and environmental monitoring to develop a digital representation of apple fruit development on the tree. Within this broader programme, the present study focuses specifically on digitally modelling and describing changes in pre-harvest apple firmness. The firmness trajectory predicted in this thesis is intended to provide a firmness-related input to the wider digital fruit model developed by PFR, supporting future integration between sensor-based prediction and mechanistic fruit modelling.

As part of this programme, multiple complementary datasets were collected over four fruit-growing seasons, from 2021–2022 to 2024–2025, to monitor the pre-harvest development of 'Royal Gala' apples. These datasets include destructive fruit quality measurements; laboratory NIR spectra collected from fruit removed from the trees for destructive assessment; longitudinal non-destructive NIR spectra repeatedly collected in the field from labelled fruit while they remained attached to the tree; and hourly environmental observations from the local PFR weather station. Although the availability and completeness of these datasets varied among seasons, together they provide a multi-source and multi-season basis for developing and evaluating a digital representation of pre-harvest apple firmness.

Existing studies have generally used mechanistic and statistical approaches separately when modelling fruit firmness (Gwanpua et al., 2012; Walsh et al., 2020; Xiao and Li, 2022). Mechanistic models can represent known biophysical and biochemical processes and provide physiologically interpretable descriptions of fruit softening. However, they often require detailed prior knowledge, simplifying assumptions, and process variables that may not be fully observed in field datasets. In contrast, statistical models can use high-dimensional sensor data and flexibly capture empirical relationships without requiring all underlying mechanisms to be specified. However, they may provide limited biological interpretation and may not transfer reliably across developmental stages, measurement conditions, or growing seasons. Integrating these two approaches is therefore not trivial. Biological knowledge must be translated into a sufficiently parsimonious model structure, while heterogeneous datasets collected at different temporal and observational scales must be organised and aligned. At the same time, statistical flexibility

must be introduced without losing the biologically expected pattern of firmness decline. These considerations motivate the mechanistic-informed statistical approach used in this thesis.

To address these challenges, this study develops a mechanistic-informed statistical framework for predicting pre-harvest apple firmness in 'Royal Gala' apples. More specifically, biological information provides guidance on the expected pattern of apple firmness decline during fruit development. In particular, previous studies suggest that apple firmness generally declines over time, with a faster decline during the earlier development stage and a slower decline later in the pre-harvest period (Volz et al., 2003). Therefore, rather than treating apple firmness prediction as a purely black-box prediction problem, this thesis uses the known apple softening pattern to construct a biologically meaningful baseline curve over days after full bloom (DAFB).

However, a baseline curve alone may not fully explain the variation in apple firmness because fruit firmness is also affected by fruit-level properties, spectral information, and environmental conditions. Therefore, after constructing the baseline softening trajectory, this study further introduces statistical correction models to capture the remaining residual variation. More specifically, fruit-level variables and destructive NIR spectra are used to construct the first correction term. Environmental variables and longitudinal non-destructive NIR spectra are then used to construct the second correction term. In this way, the proposed fused model retains the interpretability of a mechanistic-informed firmness decline curve while allowing sensor-based data to adjust the prediction when useful information is available.

A further challenge arising from the datasets collected through the digital twin programme is that the available data sources are heterogeneous and are not observed at the same temporal or observational scale. In particular, destructive fruit quality measurements provide the response variable, apple firmness, but are available only at sparse assessment dates. Laboratory NIR spectra from destructively sampled fruit are paired with reference destructive measurements, whereas longitudinal non-destructive NIR spectra are repeatedly collected from intact labelled fruit and are therefore not directly matched to destructive firmness measurements at the individual-fruit level. In contrast, environmental variables are recorded hourly and need to be aligned with the much less frequent firmness assessment dates. Therefore, this thesis also develops a practical data-alignment and modelling framework that organises fruit quality measurements, destructive NIR spectra, non-destructive NIR spectra, estimated DMC, and environmental variables into a consistent structure for pre-harvest apple firmness prediction.

1.2 Synopsis of New Contribution

The main contribution of this thesis is the development of a mechanistic-informed statistical framework that combines the interpretability of a biologically guided firmness trajectory with the flexibility of data-driven correction models. Rather than applying a mechanistic or statistical model in isolation, the proposed framework first uses biological knowledge to define a low-complexity baseline representation of pre-harvest apple softening and then incorporates fruit-level measurements, destructive and non-destructive NIR spectral information, and environmental variables through sequential statistical correction components. A further contribution is the development of a practical strategy for organising and aligning destructive fruit measurements, longitudinal non-destructive NIR spectra, and high-frequency environmental data collected at different temporal and observational scales. The resulting framework is developed using data from the 2021–2022 season and evaluated using the holdout 2024–2025 season.

The rest of this thesis is organised as follows. Chapter 2 reviews the economic significance of apple fruit, pre-harvest apple fruit development, factors influencing apple firmness, Near-Infrared (NIR) spectroscopy for fruit quality assessment, chemometric methods for NIR analysis, and existing models for apple firmness prediction. It then identifies the research gap and presents the research questions and methodological objective addressed in this thesis. Chapter 3 introduces the model development framework, including the data scope, data sources, data exploration, model inputs, data-alignment challenges, model formulation and notation, and model-fusion specification. Chapter 4 describes the model implementation, including data cleaning, NIR validation, spectral pre-processing, environmental data alignment, baseline modelling, residual correction modelling, and the 2024–2025 prediction performance evaluation. Chapter 5 presents the results of the mechanistic-informed statistical model developed in Chapter 4, including the model-development results for the 2021–2022 season and holdout cross-season prediction performance evaluation results for the 2024–2025 season. Finally, Chapter 6 presents the general conclusions and future perspectives.

Chapter 2

Literature Review

2.1 Apple Fruit and Economic Significance

The apple fruit (*Malus domestica*) is the most widely cultivated temperate species in the genus *Malus*, grown mainly in temperate regions, some subtropical regions, and in some tropical highland regions (Kellerhals, 2009; Musacchi and Serra, 2018; Priyadarshan and Jain, 2009). It is one of the most popular fruits because of its economic value and notable nutritional and health benefits (Mierczak and Garus-Pakowska, 2024; Forsline et al., 2003; Musacchi and Serra, 2018; Johnston et al., 2002a). The major apple-producing countries include China, the United States, and Poland (Musacchi and Serra, 2018). Although smaller in total output in the world, New Zealand is a significant apple-producing country, where apples represent the greatest proportion of the fruit exports (Warrington, 2000; Aitken and Warrington, 2017). As a key component of New Zealand's horticultural export economy, a substantial proportion of the harvested apples is shipped to distant international markets (Warrington, 2000; McKenna et al., 2001). Apple firmness at harvest strongly influences quality retention during storage and long-distance distribution. It affects market acceptance and consumer satisfaction and is also important for storability and minimising post-harvest losses (Musacchi and Serra, 2018; Johnston et al., 2002a; Małachowska and Tomala, 2023). In addition, accurate pre-harvest firmness prediction can support harvest-timing and post-harvest management decisions. Overall, predicting pre-harvest apple fruit firmness and maintaining suitable apple fruit firmness throughout storage and transport are important for preserving apple fruit quality and commercial value (Li et al., 2018).

2.2 Pre-harvest Apple Fruit Development

Pre-harvest apple fruit development begins at bloom and proceeds through four important stages, including cell division, cell expansion, maturation, and ripening, which determine the structure and composition of the apple fruit at harvest (Volz et al., 2003, 2002; Gillaspy et al., 1993; Szymanska-Chargot et al., 2012; Denne, 1960). Therefore, understanding pre-harvest apple fruit development is important for apple firmness prediction modelling.

Cell division occurs during the first few weeks after bloom and contributes to formation of the cellular framework of the cortex (Musacchi and Serra, 2018; Denne, 1960; Blanpied and Wilde, 1968). During this early pre-harvest stage, cell division proceeds rapidly (Musacchi and Serra, 2018; Denne, 1960). As the apple fruit develops before harvest, cell division slows, and cell expansion becomes the dominant process, resulting in a substantial increase in apple fruit size and changes in tissue texture (Johnston et al., 2002a). This period of cell expansion is accompanied by the deposition and modification of primary cell wall materials, including cellulose, hemicellulose, and pectin (Johnston et al., 2002a; Redgwell and Fischer, 2002). Since these cell wall materials are closely related to cell adhesion and tissue strength, changes during cell division and cell expansion provide an important biological basis for understanding pre-harvest apple firmness development.

As apple fruit development progresses toward maturation, starch that accumulated during the earlier stages is gradually hydrolyzed into soluble sugars, such as fructose, glucose, and sucrose (Musacchi and Serra, 2018). Furthermore, this shift in carbohydrate metabolism increases the osmotic potential within the cells, improving flavour profiles and maintaining turgor, which is important for fruit firmness (Berüter, 1985; Chagné et al., 2014). Moreover, dry matter accumulation, including these sugars, is often used as an indicator of the quality potential of apple fruit at harvest (Toivonen and Lannard, 2020).

The final stage of the pre-harvest fruit development, the ripening stage, is characterized by significant biochemical changes (e.g., colour, aroma, and flavour), and a significant increase in the production of the ripening ethylene (Barry and Giovannoni, 2007; Fernández-Cancelo et al., 2022). Ethylene serves as a hormonal signal that initiates a cascade of physiological processes, most notably the activation of cell wall-modifying enzymes such as polygalacturonase (PG) and pectin methyl esterase (PME) (Johnston et al., 2002a; Ireland et al., 2014). Therefore, these

enzymes are responsible for the solubilisation and depolymerisation of pectin in the primary cell wall and middle lamella, which weakens cell-to-cell adhesion, leading to a decline in apple fruit firmness during apple fruit development (Johnston et al., 2002a). In addition to these physiological mechanisms, the temporal pattern of firmness decline during pre-harvest fruit development also provides useful information for modelling. According to Volz et al. (2003), firmness in 'Gala' apples declines rapidly during the earlier stage of fruit development, with the rate of decline slowing during the pre-harvest period. This suggests that the pre-harvest firmness trajectory can be represented by a monotone decline with different rates at different fruit developmental stages.

Overall, the interplay among cell division, cell expansion, carbohydrate metabolism, and ethylene-regulated cell wall modification during these stages determines the internal structure and textural properties of the apple at harvest (Su et al., 2022).

2.3 Factors Influencing Apple Firmness Development

For apple fruit, firmness is the most important quality attribute during the pre-harvest period, as it is not only one of the key indicators of eating quality, storability, and market value, but also serves as a maturity index used to determine the optimal picking time (Johnston et al., 2002a; Goncalves et al., 2017; Kviklienė et al., 2011). Importantly, apple firmness does not remain constant during pre-harvest apple fruit development (Denne, 1960; Johnston et al., 2002a; Volz et al., 2003). Instead, it gradually declines throughout the fruitlet growth period from full bloom until harvest (Volz et al., 2003). Especially, during this pre-harvest stage, apple fruit firmness is influenced by the combination of various factors, including environmental conditions, cellular structure, chemical composition, and orchard management (Ornelas-Paz et al., 2018; Su et al., 2022).

For environmental factors, atmospheric and soil conditions during the pre-harvest developing period influence apple fruit firmness through their impacts on xylem water potential (Ψ_x) (Landsberg et al., 1976; Lachapelle et al., 2013). Moreover, temperature, relative humidity, accumulated rainfall, and wind speed determine the vapour pressure deficit (VPD), which can regulate the rate of transpiration from leaves and apple fruit surfaces before harvest (Butler, 1976; Landsberg et al., 1976; Zanotelli et al., 2019). As a result, high temperature and low

humidity lead to a higher VPD and may accelerate water loss from leaves and fruit (Zanotelli et al., 2019; Landsberg et al., 1976). The increased water loss may lead to more negative xylem water potential (Ψ_x), reducing cellular turgor pressure and contributing to short-term variation in apple fruit firmness during the pre-harvest period (Lachapelle et al., 2013; Johnston et al., 2002a). Moreover, soil water potential reflects the availability of soil water for root uptake and influences the water status of apple trees during the pre-harvest apple fruit development (Zhang and Wang, 2023; Brighenti et al., 2024; Tao et al., 2023). More specifically, it is affected by different factors, including soil moisture content, evaporation rate, and irrigation frequency (Brighenti et al., 2024; George and Nissen, 1988; Butler, 1976). Furthermore, when soil water potential becomes more negative due to insufficient irrigation or rainfall, root water uptake is restricted, leading to a reduction in xylem water potential (Ψ_x) (Landsberg et al., 1976; Brighenti et al., 2024). Therefore, the cell turgor pressure is reduced, which may contribute to a loss of fruit firmness during the pre-harvest development of apple fruit (Lachapelle et al., 2013). Although the above studies provide empirical and physiological support for different parts of this pathway, the integrated quantification of the complete relationship from VPD through xylem water potential and cellular turgor to measured apple fruit firmness remains limited in apple studies (Carella et al., 2023; Steudle and Wieneke, 1985).

For cellular-level factors, apple fruit firmness is determined by the cell structural and functional properties of apple fruit tissues, which change dynamically during pre-harvest apple fruit development (Volz et al., 2002, 2003). In the early stage after full bloom, cell division occurs rapidly and contributes to the exponential phase of fruit growth (Denne, 1960). According to Johnston et al. (2002a), the primary cell wall consists of cellulose microfibrils embedded in a hemicellulose–pectin matrix. The pectic polymers act as the middle lamella to maintain cell adhesion and cell packing essential for apple fruit firmness (Johnston et al., 2002a; Redgwell and Fischer, 2002). Furthermore, as the apple fruitlet increases in size, cell expansion becomes predominant, and the associated loosening of the cell wall results in reduced apple fruit firmness (Su et al., 2022; Vicente et al., 2007; Qi et al., 2015). Therefore, changes in cell division, cell expansion, cell wall structure, and cell-to-cell adhesion provide an important cellular-level explanation for the apple firmness decline during fruit development. Figure 2.1 shows the diagram of the factors influencing apple fruit firmness from a cellular perspective.

For chemical factors, changes in the biochemical composition of apple fruit tissues, such as

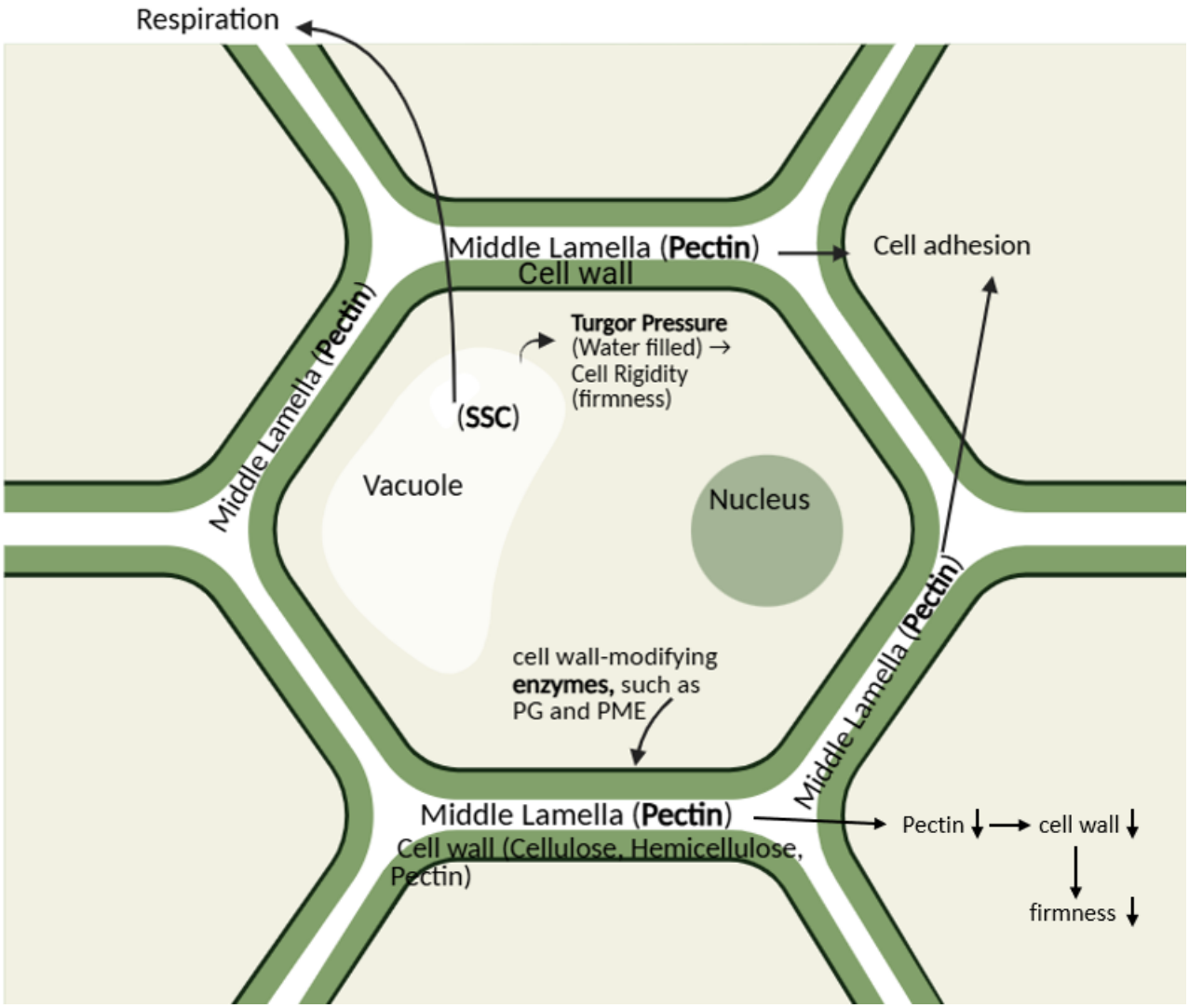


Figure 2.1: Diagram of the factors influencing apple fruit firmness from a cellular perspective

cell wall-modifying enzymes, soluble sugars, and dry matter content (DMC), play an important role in regulating apple firmness during the stages of the pre-harvest apple fruit development and maturation (Johnston et al., 2002a; Bonnin and Lahaye, 2013; Su et al., 2022; Saei et al., 2011). Among these, cell-wall-modifying enzymes, such as polygalacturonase (PG), pectin methyl esterase (PME), and glycosidases (e.g., β-galactosidase), cause pectin solubilisation and hemicellulose degradation in the middle lamella (Ben-Arie et al., 1979; Johnston et al., 2002a; Su et al., 2022; Wei et al., 2010). Changes in soluble sugar composition during pre-harvest apple fruit development may influence apple fruit firmness through their effects on osmotic regulation and cellular turgor pressure (Berüter, 1985; Johnston et al., 2002a). Changes in the concentrations of sucrose, fructose, glucose, and sorbitol influence cellular osmotic regulation and turgor pressure, which are essential for maintaining cell expansion, structural integrity, and fruit firm-

ness (Hansen, 1979; Berüter, 1985; Bieleski, 1969; Johnston et al., 2002a). In early apple fruit development, sucrose is absent, and sorbitol, glucose, and fructose levels decline until shortly before the June drop period (Berüter, 1985; Bieleski, 1969). As fruit growth progresses, the concentrations of sucrose and fructose accumulate steadily, while sorbitol concentration decreases markedly with the appearance of sorbitol dehydrogenase activity (Berüter, 1985). In addition, glucose shows a temporary decrease associated with starch accumulation before rising again at later developmental stages (Berüter, 1985). Moreover, during apple fruit development, starch reserves accumulated in early development are hydrolyzed into soluble sugars by amyloglucosidase, such as glucose, fructose, and sucrose, increasing soluble solids concentration (SSC), while sorbitol levels generally decrease, indicating a reduction in phloem import and a decline in fruit firmness (Berüter, 1985; Fishman and Génard, 1998; Berüter et al., 1997; Johnston et al., 2002a; Palmer et al., 2010).

With chemical changes, ethylene production is a major regulator in apple fruit development, especially during pre-harvest maturation and ripening (Blankenship and Unrath, 1988; Johnston et al., 2002a,b). Furthermore, ethylene production shows a significant increase during ripening, triggering the expression of several cell wall-modifying enzymes, such as polygalacturonase (PG) and pectin methyl esterase (PME) (Johnston et al., 2002a; Ireland et al., 2014). Increased ethylene concentration not only accelerates cell wall disassembly but also modulates osmotic balance through changes in soluble sugars, thereby contributing to the general decline in apple fruit firmness during maturation before harvest (Burg and Burg, 1967; Lau et al., 1986; Sun et al., 2021; Tacken et al., 2010). In addition, dry matter content (DMC) is associated with firmness at harvest (Chagné et al., 2014; Saei et al., 2011; Toivonen and Lannard, 2020; Palmer et al., 2010). More specifically, higher DMC reflects greater accumulation of structural carbohydrates, which strengthens cell-wall structure, supports turgor maintenance, and is often associated with increased SSC (Saei et al., 2011; Møller et al., 2013; Palmer et al., 2010). Therefore, these factors together contribute to the retention of apple texture and loss of firmness during pre-harvest apple fruit development and ripening.

For orchard management factors, according to Johnston et al. (2002a), various pre-harvest human practices, such as the thinning of crop load, irrigation, orchard floor management, harvest timing, and nutritional treatments, have a significant impact on apple firmness. For example, reducing crop load has been shown to enhance fruit firmness at the time of harvest (Saei et al.,

2011). Furthermore, there are various pre-harvest nutritional treatments that could influence apple firmness, such as calcium treatments that increase the cell wall integrity by cross-linking with pectins and the increase in membrane microviscosity that generally occurs during the ripening of apples (Johnston et al., 2002a; Sams, 1999; Bramlage, 1992). Therefore, orchard management practices before harvest are important external factors that may contribute to variation in pre-harvest apple firmness. However, although orchard management is an important factor, the orchard management information available in this study was limited. More specifically, the tree spacing variable was recorded in the 2021–2022 and 2022–2023 seasons, while this variable was not recorded in the 2023–2024 and 2024–2025 seasons. Moreover, other orchard management variables, such as thinning of crop load, irrigation, and nutritional treatments, were not available. Therefore, the orchard management factor variables were not included in our subsequent modelling study.

Overall, apple fruit firmness development during the pre-harvest period is influenced by the interaction of environmental, cellular-level, chemical, and orchard management factors. These factors influence apple fruit firmness through their effects on fruit water status, cell wall structure, biochemical composition, and orchard management practices. Therefore, apple firmness should be considered as the combined result of multiple interacting processes. Finally, Figure 2.2 summarises the four factors influencing apple fruit firmness during pre-harvest apple fruit development.

2.4 Near-Infrared Spectroscopy for Apple Firmness Assessment

2.4.1 Introduction

Near-Infrared (NIR) Spectroscopy is a spectroscopic non-destructive method that uses the light in the 700–2500 nm wavelength range to explore molecular composition and physical properties within biological samples (Grabska et al., 2023; Ciurczak et al., 2021). According to Ciurczak et al. (2021), NIR spectroscopy provides dense information through the detection of overtone and combination chemical bond motions, especially of water, sugar contents, and cell wall components. It is sensitive to oxygen-hydrogen (O-H), carbon-hydrogen (C-H), carbon-oxygen (C-O),

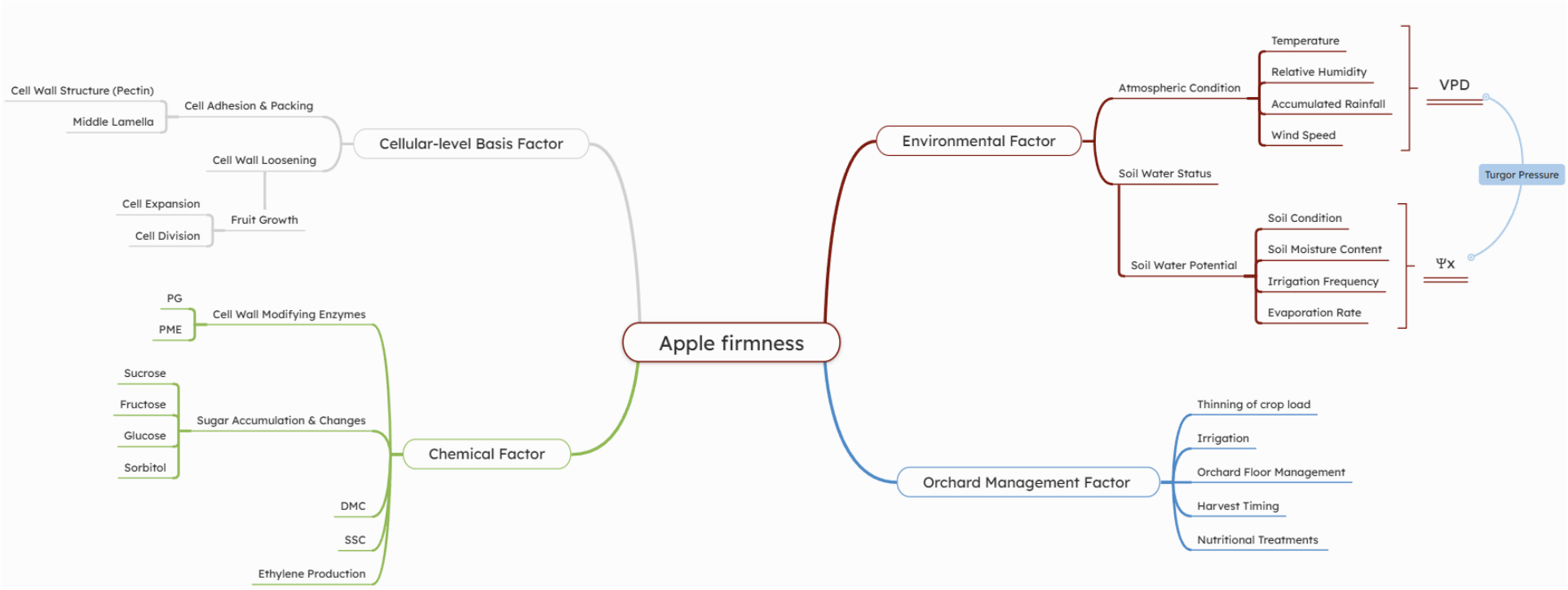


Figure 2.2: Diagram of four factors influencing apple fruit firmness at harvest and during development.

and nitrogen-hydrogen (N-H) bonds (Ciurczak et al., 2021). Compared with the visible spectroscopy region (400–700 nm), the changes in the NIR spectral region during apple fruit ripening are generally less dominated by pigment-related variation (Ciurczak et al., 2021). However, this does not mean that NIR spectra remain unchanged during apple fruit development. The NIR spectral response still varies with water content, soluble carbohydrates, dry matter accumulation, tissue structure, and cell-wall-related physical changes (Ciurczak et al., 2021; Nicolai et al., 2007). These ripening-related spectral changes make NIR spectroscopy potentially useful for apple fruit-quality assessment (Grabska et al., 2023). However, because firmness is determined by both chemical composition and mechanical tissue properties, its association with NIR spectra is generally more indirect than the association between NIR spectra and attributes such as DMC or SSC (Møller et al., 2013; Ma et al., 2021). In contrast, the visible region is more strongly affected by changes in pigment composition, including chlorophyll degradation and anthocyanin accumulation, and therefore generally shows more pronounced spectral changes during fruit ripening (McGlone et al., 2002).

Over the past few decades, NIR spectroscopy has been widely used in horticulture studies and commercial quality measurement of apple fruits (McGlone et al., 2002; Pandiselvam et al., 2022; Zhang et al., 2019; Nicolai et al., 2007). Due to fast measurement, minimum treatment of samples, and high-throughput screening potential, NIR spectroscopy has turned into an ideal tool for principal quality parameter measurement, such as dry matter content and soluble solids content (Pandiselvam et al., 2022; Zhang et al., 2019). In the context of pre-harvest apple firmness prediction, NIR spectroscopy is particularly useful because it can provide information about the internal properties of apple fruit without damaging the fruit (Costa et al., 2009). The reason is that traditional firmness measurement is destructive, and repeated destructive measurements on the same fruit are not possible during pre-harvest fruit development (Mendoza et al., 2011). Therefore, NIR spectroscopy provides a practical way to monitor apple fruit quality changes before harvest and to provide additional information for predicting apple firmness (Costa et al., 2009).

However, there are also some limitations; the transferability and stability of NIR calibrations can be threatened by growing season differences, cultivars, and measurement conditions (Xu et al., 2019). Moreover, the physiological basis of NIR responses to firmness is less evident compared to other attributes, such as soluble solids content (SSC), as the firmness of fruits is

influenced by various factors, such as cell turgor pressure and cell wall composition, which do not directly correspond with the chemical bonds targeted by NIR (Møller et al., 2013; Peirs et al., 2001; McGlone and Kawano, 1998). Therefore, NIR spectra can provide useful information for firmness prediction, but firmness may not be fully explained by spectral information only.

For apple fruit, the NIR spectral ranges that correspond to different quality attributes, such as firmness, dry matter content (DMC), and SSC, could be different between cultivars, reflecting underlying genetic diversity (Ciurczak et al., 2021). In particular, as discussed in the above Section 2.3, we know that the pre-harvest apple firmness is affected not only by chemical composition, but also by cell wall structure, water status, turgor pressure, and tissue mechanical properties. Therefore, the spectral response associated with firmness is often more complex and indirect than that of DMC or SSC. Furthermore, this complexity may present challenges when using NIR spectral data alone for pre-harvest apple firmness prediction modelling. Table 2.1 shows an example of wavelength ranges used in published NIR calibration studies for quality attributes in the "Royal Gala" apple fruit, which is the cultivar under test in our project (Ciurczak et al., 2021; Zhang et al., 2019; Wang et al., 2015; Ma et al., 2018). In particular, although the same wavelength ranges have been reported for SSC and apple firmness prediction, this does not imply that both attributes are predicted from the same spectral information (Grabska et al., 2023). The reason is that SSC is mainly associated with the absorption of soluble sugars and water, while firmness is influenced by additional factors such as tissue structure, cell wall properties, and water status (Cheng et al., 2022).

Table 2.1: Examples of wavelength ranges used in published NIR calibration studies for quality attributes in the "Royal Gala" apple fruit

Cultivar	Attribute	Published spectral range (nm)	Main spectral sensitivities	Reference
Royal Gala	SSC	500-1100	O–H, C-H	(Ciurczak et al., 2021)
	Firmness (N)	500-1100	O–H, C-H	(Ciurczak et al., 2021) (Zhang et al., 2019)
	Dry Matter	800-1000	O–H, C-H	(Ciurczak et al., 2021)

2.4.2 Chemometric Methods for NIR Analysis

After obtaining Near-Infrared (NIR) spectra, chemometric methods are used to extract the useful quality information of the fruit attributes that are hidden in the spectrum (Porep et al., 2015).

NIR data analysis includes two key steps: first, NIR data pre-processing to reduce noise and correct baseline drift; and second, developing chemometric models to correlate the processed spectra with specific fruit quality attributes (Stevens and Ramirez-Lopez, 2014; Çataltaş and Tütüncü, 2021; Xiao and Chen, 2022).

In the first step of NIR data pre-processing, commonly used techniques include Multiplicative Scatter Correction (MSC), Standard Normal Variate (SNV), Savitzky-Golay Filtering (with smoothing or derivatives), and mean centering and scaling (Çataltaş and Tütüncü, 2021; Stevens and Ramirez-Lopez, 2014). These pre-processing methods aim to reduce unwanted variability caused by light scattering, baseline shifts, and variations in effective path length, thereby improving model accuracy (Rinnan et al., 2009; Stevens and Ramirez-Lopez, 2014). Among these, SNV and MSC are particularly effective in correcting scatter effects and baseline, while Savitzky-Golay derivatives are widely used for enhancing spectral features (Savitzky and Golay, 1964; Naes et al., 1990; Geladi et al., 1985).

After pre-processing, chemometric models are built to translate spectral data into terms of reference measurements of fruit attributes, such as firmness, dry matter, and soluble solids content (Çataltaş and Tütüncü, 2021). According to Çataltaş and Tütüncü (2021); Yan (2025), the following chemometric models are mostly used for modelling these relationships:

- **Multiple Linear Regression (MLR)**: MLR is a statistical method used to establish a linear relationship between selected spectral variables and the target attribute (Mark and Workman Jr, 2018). However, MLR tends to perform poorly with high-dimensional NIR datasets, since its predictive performance is often poor when the number of input variables increases (Fritz and Berger, 2015).

- **Partial Least Squares Regression (PLSR)**: PLSR is one of the most widely used regression techniques in NIR spectroscopy applied to horticulture (Çataltaş and Tütüncü, 2021). Also, PLSR is a regression technique that combines the essential features of the principal component analysis (PCA) and MLR methods (Çataltaş and Tütüncü, 2021; Abdi and Williams, 2012). PLSR is particularly effective when the independent variables are highly correlated, making it well-suited for spectral data analysis (Næs and Martens, 1988).

- **Principal Component Regression (PCR)**: PCR is a statistical method that combines

PCA with linear regression to address the problem of collinearity (Mandel, 1982). Furthermore, the process begins by using PCA to extract the principal components of the spectral data. Among these, the components that account for the most variance in the data are selected, and a regression model is then built using these selected components (Mandel, 1982).

- **Support Vector Machine (SVM)**: SVM is a supervised learning method suitable for classification and regression problems (Cortes and Vapnik, 1995). Specifically, SVM is appropriate for defining non-linear associations between spectral measurements and quality characteristics of fruits (Sulistya et al., 2024).

- **Linear Discriminant Analysis (LDA)**: LDA is a statistical method used for both classification and dimensionality reduction problems (Çataltaş and Tütüncü, 2021; Izenman, 2013). Although LDA shares a conceptual similarity with PCA in reducing data dimensionality, its primary objective is different: LDA seeks to find a projection axis that maximizes the separation between different classes while minimising the variance within each class (Çataltaş and Tütüncü, 2021).

- **Partial Least Squares Discriminant Analysis (PLS-DA)**: PLS-DA is an extension of PLSR and is used to solve classification problems (Çataltaş and Tütüncü, 2021). Also, PLS-DA is particularly well-suited for analysing high-dimensional, collinear spectral datasets (Lee et al., 2018).

Overall, among the above methods, PLSR is the most relevant method to this thesis because it is widely used for analysing high-dimensional and highly correlated NIR spectral data in horticultural studies (Çataltaş and Tütüncü, 2021). PLSR is suitable for predicting continuous apple fruit quality attributes, e.g., DMC and apple fruit firmness (Çataltaş and Tütüncü, 2021). However, PLSR mainly models linear relationships, while the relationship between NIR spectra and pre-harvest apple firmness may be complex and indirect (Çataltaş and Tütüncü, 2021). Also, PCR can reduce the dimensionality of spectral data, but its components are not selected based on the response variable (Mandel, 1982). Moreover, non-linear methods, such as SVM, may capture more complex relationships, but they generally require larger datasets and careful parameter selection (Sulistya et al., 2024). Therefore, PLSR was selected in our study for NIR

spectral validation and DMC estimation.

Moreover, to evaluate the prediction performance of the built model from NIR spectral datasets, the cross-validation and independent test sets are commonly used (Walsh et al., 2020; Çataltaş and Tütüncü, 2021). There are various metrics to determine the accuracy of the built model, including the coefficient of determination (R^2), root mean square error of prediction (RMSEP), root mean square error of calibration (RMSEC), and residual predictive deviation (RPD) (Yan, 2025; Çataltaş and Tütüncü, 2021; Ncama et al., 2018). These measures provide an objective basis for comparing different modelling approaches and assessing the practical use of NIR-based predictions in fruit quality assessment (Çataltaş and Tütüncü, 2021).

The above methods describe how useful information can be extracted from NIR spectra. However, NIR spectroscopy is only one part of apple firmness assessment and is also one of the important aspects explored in this thesis. Apple fruit firmness changes systematically during apple fruit development and is influenced by biological and environmental processes that may not be fully represented by NIR spectral information alone. Therefore, the following section reviews the mechanistic and statistical modelling methods used for apple fruit firmness prediction.

2.5 Existing Models for Apple Firmness Prediction

In recent years, various models have been developed to predict apple fruit firmness, and these can be divided into two categories: mechanistic models and statistical models (Zude et al., 2006; Gwanpua et al., 2017; De Smedt et al., 2002).

The mechanistic models based on biophysical and biochemical mechanisms can describe the physiological processes of fruit softening, such as pectin degradation, enzyme activity, water transfer process, and ethylene production and regulation (Gwanpua et al., 2012). Mechanistic models also explain the effects of factors such as time, temperature, and initial maturity on firmness changes (Gwanpua et al., 2012; De Smedt et al., 2002). Moreover, mechanistic models have high interpretability because model parameters and variables have clear physiological meaning, which helps researchers to understand and simulate the softening mechanisms under different conditions. However, they require prior knowledge, simplified assumptions, and high computational ability (Xiao and Li, 2022). Gwanpua et al. (2012) developed a mechanistic model to simulate the pectin degradation processes and ethylene effects that lead to the changes

in firmness of "Braeburn" apple fruits during storage and shelf life. Róth et al. (2008) developed a mechanistic model to describe enzyme activities that lead to the changes in firmness of apple fruits during postharvest. To describe the development of apple fruit firmness during storage, De Smedt et al. (2002) developed a mechanistic model that simulates the water transfer process, which is associated with cellular turgor pressure and the hydrolysis of pectin in the middle lamella, weakening cell-to-cell adhesion. Gwanpua et al. (2017) developed a mechanistic model focused on ethylene production and regulation to describe how 1-methylcyclopropene (1-MCP) inhibits ethylene perception and delays ripening in apples. Similarly, Matabura (2022) developed a mechanistic model to simulate the ethylene production and regulation to describe the variability of the "Jonagold" apple fruit firmness during postharvest storage. Overall, these studies show that mechanistic models are useful for explaining the biological mechanisms underlying apple softening. However, most of these models have been developed for post-harvest storage or shelf-life conditions, rather than for predicting apple firmness during the pre-harvest period.

Compared to mechanistic models, statistical models rely on empirical relationships between observable variables and response variables, such as dry matter content (DMC), soluble solids content (SSC), and fruit firmness (McGlone et al., 2002; Zhang et al., 2019; Çataltaş and Tütüncü, 2021). In agricultural research, especially in fruit attribute prediction, the statistical models frequently use near-infrared (NIR) or Vis–NIR spectral data as explanatory variables for prediction (Fan et al., 2009; Shenk et al., 2007; Li et al., 2017; Xiao and Li, 2022; Peirs et al., 2001). Moreover, the statistical models do not require predefined biological mechanisms and can also capture complex nonlinear relationships (Burke, 1996; Walsh et al., 2020). However, the interpretability of statistical models is low, making it difficult to explain the physiological processes behind the predictions (Xiao and Li, 2022; Walsh et al., 2020). Also, statistical models struggle to handle time-related factors and usually require additional steps to incorporate temporal changes (Giovanelli et al., 2014; Lu and Bailey, 2005).

For the apple fruits, statistical models have been commonly used to predict apple fruit quality attributes, such as SSC, starch index, chlorophyll content, and firmness, using non-destructive NIR spectroscopy data or Vis–NIR spectral data (Marečková et al., 2022; Park et al., 2003; Giovanelli et al., 2014; Zude-Sasse et al., 2002; Menesatti et al., 2009). Among these statistical methods, models can be divided into linear latent variable models, such as partial least squares (PLS) regression, and non-linear machine learning methods, such as artificial neural networks

(ANNs) and convolutional neural networks (CNNs). Additionally, the linear latent variable model has the ability to manage high dimensionality and multicollinearity (Lu and Bailey, 2005). There are various studies that have demonstrated its effectiveness in predicting fruit firmness in apples at different periods (Mendoza et al., 2012; Lu and Bailey, 2005; Pirouz, 2006). For example, Mendoza et al. (2012) used partial least squares (PLS) regression to develop the calibration model to predict the "Jonagold", "Golden Delicious", and "Delicious" apple fruit firmness using Vis/NIR spectra collected at harvest. Similarly, Lu and Bailey (2005) used PLS regression to model "Golden Delicious" apple firmness changes under different storage conditions, including freshly harvested and cold-stored (one month and two months). Giovanelli et al. (2014) combined PLS regression and linear discriminant analysis (LDA) classification methods for the analysis of spectral data to predict 'Golden Delicious' apple firmness at postharvest. Overall, these studies demonstrate the usefulness of statistical and machine learning methods for apple firmness prediction, especially when spectral data are available. However, many of these models are developed for harvest-time, post-harvest, or storage, rather than predicting apple firmness during the pre-harvest period.

To capture the often non-linear relationships between spectral data and apple fruit attributes (e.g., firmness), researchers have also explored more advanced machine learning models. For example, the neural network (NN) regression approach has been applied to predict the firmness of 'Fuji' apples during storage based on NIR spectral data (Zeng et al., 2024) and to estimate the 'Fuji' apple firmness at different ripening stages based on Vis/NIR spectral data (Pourdarbani et al., 2020). Similarly, Malvandi et al. (2022) developed an NN model to develop a calibration model to predict apple firmness based on NIR spectral data during drying.

Although many researchers have studied mechanistic and statistical models for apple fruit firmness prediction separately, few published studies have focused on fusing these approaches for apple firmness prediction. Hybrid models that combine mechanistic and statistical methods aim to use the advantages of both: mechanistic models provide strong physiological interpretability and the ability to simulate underlying biological processes, while statistical models have the ability to capture complex data-driven relationships and handle high-dimensional input variables, such as NIR spectral data. For example, Van Der Sman and Sanders (2012) applied non-linear regression as a statistical method for parameter estimation within their mechanistic biological switch model to predict post-harvest apple firmness. Xiao and Li (2022) proposed

a fusion framework that combined a data-driven model with a mechanistic model for kiwifruit flesh firmness prediction. Although their published work was focused on Kiwifruit rather than apple fruit, it provides a useful example of how mechanistic and statistical models can be fused for fruit firmness prediction. Although these studies demonstrate the potential of integrating mechanistic modelling with statistical modelling, the incorporation of mechanistic information describing apple fruit softening into statistical models for pre-harvest firmness prediction remains insufficiently explored. In particular, limited attention has been given to combining mechanistic information with sensor data, such as NIR spectral data and environmental variables.

2.6 Research Gap and Motivation

Although a number of models have been developed for apple fruit firmness prediction, most studies have focused on harvest-time assessment, post-harvest storage, and fruit ripening processes. Comparatively fewer studies have investigated apple firmness prediction during the pre-harvest period, when fruit remains attached to the tree, and apple fruit firmness continuously changes throughout fruit development.

Furthermore, few existing mechanistic, statistical, and hybrid approaches have integrated biological information describing pre-harvest apple softening with fruit quality measurements, Near-Infrared (NIR) spectral data, and environmental information within a single prediction framework. These datasets are collected at different observational and temporal scales. For example, destructive firmness and fruit quality measurements are available only at selected assessment dates, non-destructive NIR spectra are repeatedly collected from the labelled intact apple fruit, and environmental variables are recorded at a higher frequency. Therefore, aligning these heterogeneous data sources and determining how each source can contribute to apple fruit firmness prediction is an additional challenge.

Therefore, the purpose of our work is to develop a mechanistic-informed statistical model that combines mechanistic information with statistical modelling for predicting pre-harvest apple firmness. More specifically, the proposed model could incorporate mechanistic information describing apple fruit softening with data-driven relationships derived from fruit quality measurements, NIR spectral data, and environmental variables to evaluate whether the integration of this information improves pre-harvest apple firmness prediction.

2.7 Research Questions and Methodological Objective

The overall aim of this study is to develop and evaluate a mechanistic-informed statistical model to predict the pre-harvest apple fruit firmness using measured fruit quality data, Near-Infrared (NIR) spectral data, and high-frequency environmental data.

To achieve this aim, this thesis addresses two research questions and one methodological objective:

1. **Research Question One:** What is the most appropriate pre-harvest firmness response to model given the constraints of the available data in this study: firmness at the individual-fruit level, or a population-level firmness target (for example an average over time, tree, or orchard), given the available destructive firmness measurements and NIR data streams?

2. **Research Question Two:** Given the selected firmness response and the organised inputs, how can we construct a fused modelling framework that links the inputs to pre-harvest firmness by combining mechanistic information with data-driven techniques, and how does the resulting framework perform when evaluated on a holdout test season?

3. **Methodological Objective:** How can we define, clean, and organise the heterogeneous predictor variables from the raw datasets (physical measurements, outputs from other models, NIR on destructively tested fruits, longitudinal NIR on tracked fruits without destructive firmness, and environmental factors) into a consistent set of model inputs?

Finally, to address the above research questions and accomplish the methodological objective, the following chapters are used to achieve these aims: Chapter 3 addresses Research Question One and introduces the modelling framework relevant to Research Question Two; Chapter 4 addresses Research Question Two and the methodological objective by describing the data-processing procedures and implementation of the baseline and statistical correction models.

Chapter 3

Model Development Framework

In our study, firmness, fruit traits, NIR spectra, and environmental variables are all informative, but they are observed on different time scales and with incomplete overlap across years. Therefore, this creates a practical challenge: we need a model that is structured enough to reflect known biology, yet flexible enough to use the information carried by the NIR and environmental drivers. The data sources are heterogeneous not only in time scale but also in measurement noise, variable type (quantitative versus qualitative; discrete or binary versus continuous), dimensionality, and even distributional behaviour across variables of the same type. Therefore, the strategy of modelling in this thesis is to impose a physiologically reasonable baseline for the pre-harvest apple firmness decline first, and then introduce additional components to capture systematic deviations that the baseline does not explain.

3.1 Data Scope and Alignment Challenge

The available dataset includes repeated pre-harvest measurements of apple fruit firmness, diameter, weight, and soluble solids content (SSC), together with pre-harvest NIR spectra and hourly environmental variables. These data sources are misaligned at multiple levels:

- **Micro-level:** Environmental drivers are recorded hourly, whereas firmness is measured relatively sparsely.
- **Macro-level:** Multi-year data can show strong seasonal heterogeneity and incomplete overlap across variables.

- **System-level:** Apple fruit firmness is driven by a complex physiological system, but the key processes are only partially observed.

Therefore, simply merging multi-year data or purely black-box prediction models tends to be unreliable (Xiao and Li, 2022). Therefore, an important focus of this thesis is to align these data sources in a consistent way, so that information from high-frequency drivers and NIR can be used to support inference on apple firmness prediction at pre-harvest.

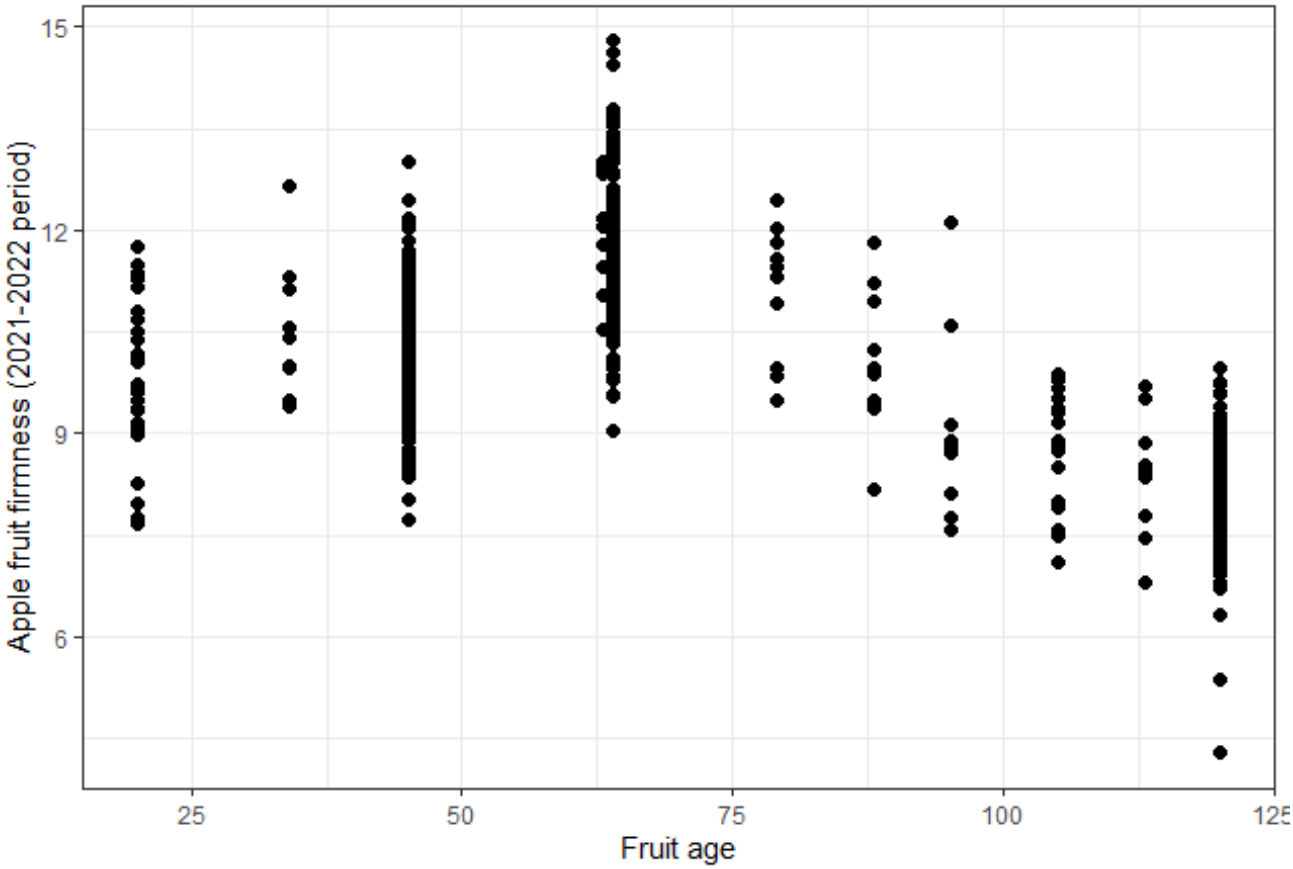


Figure 3.1: Raw pre-harvest apple fruit firmness (kgf) measurements against fruit age (DAFB) for the 2021–2022 season. Measurements before 60 DAFB are shown for completeness but were excluded from subsequent modelling because firmness could not be measured reliably when the fruit were small relative to the penetrometer probe.

For the pre-harvest window of interest (approximately 60–110 DAFB), apple fruit firmness is expected to decline over time, with a faster decline earlier in fruit development and a slower decline later in the pre-harvest period (Volz et al., 2003). However, the complete raw measurements shown in Figures 3.1 and 3.2 display an apparent increase in firmness at the earliest assessment dates, followed by a subsequent decline. This initial increase should not be interpreted as a physiological firmness pattern. Before approximately 60 DAFB, the fruit were small

relative to the penetrometer probe, and the probe could pass through the fruit or cause cracking rather than provide a reliable measurement of flesh firmness. The resulting early measurements were therefore considered unreliable. They are retained in the figures to show the complete raw datasets but were excluded from the subsequent modelling.

From approximately 60 DAFB onward, the retained measurements in both seasons show an overall decline in firmness with increasing fruit age. Therefore, the biologically expected firmness decline provides the basis for specifying a two-phase monotone baseline, rather than fitting the apparent increase–decrease pattern observed in the complete unfiltered raw data.

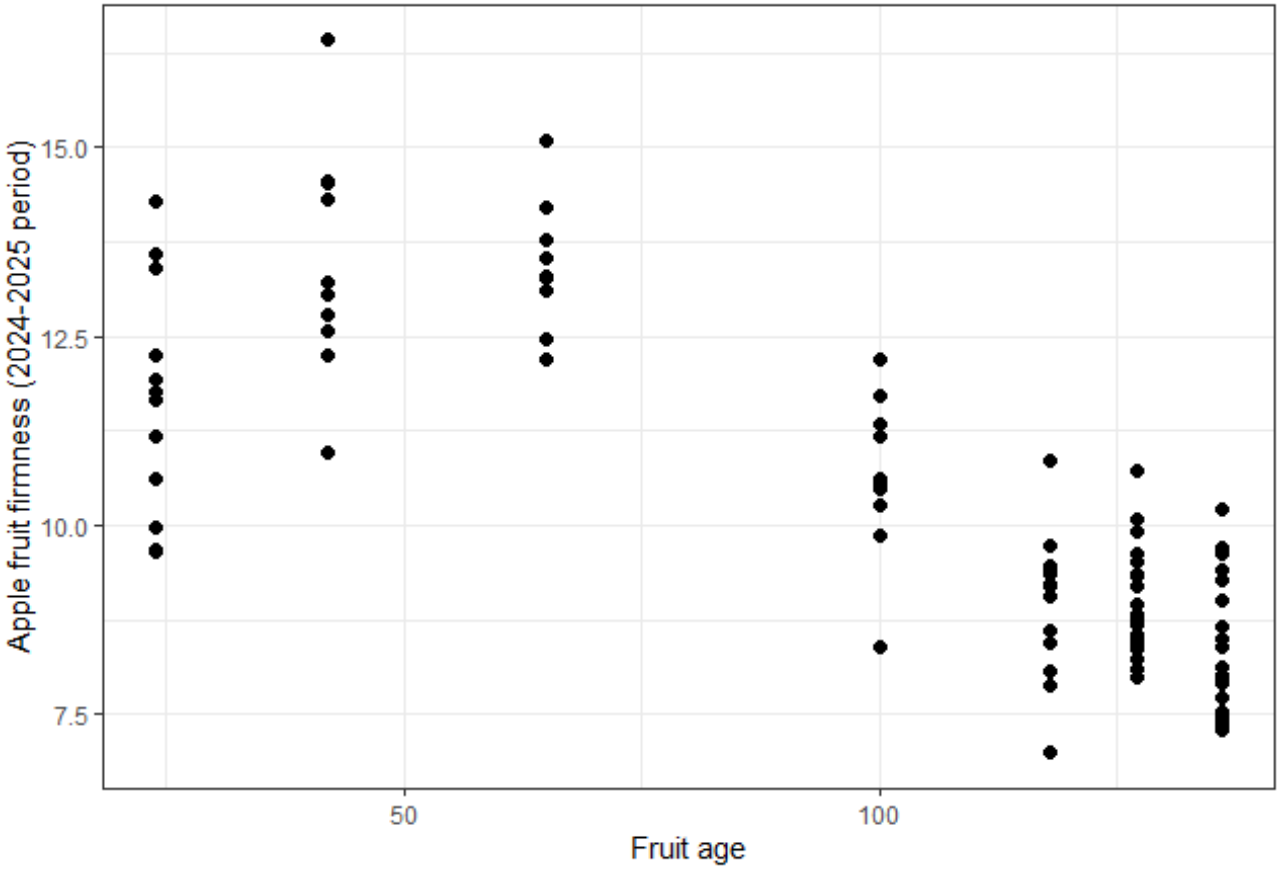


Figure 3.2: Raw pre-harvest apple fruit firmness (kgf) measurements against fruit age (DAFB) for the 2024–2025 season. Measurements before 60 DAFB are shown for completeness but were excluded from subsequent modelling because firmness could not be measured reliably when the fruit were small relative to the penetrometer probe.

3.2 Data Sources, Collection, and Availability

The datasets analysed in this thesis were collected as part of a broader digital twin programme at Plant & Food Research (PFR). The present thesis uses a subset of the fruit-quality, NIR spectral, and environmental data collected under this wider programme, with a specific focus on modelling pre-harvest apple firmness. Because the field and laboratory data collection was undertaken within the broader programme rather than designed specifically for this thesis, the following description reports the collection details that could be verified from the records made available for the present analysis. Complete metadata for the orchard block, planting architecture, number of sampled trees, and some destructive measurement devices were not available in the records supplied for this thesis; therefore, no unverified details are reported.

The data used in this study came from four complementary sources: destructive fruit quality assessments, destructive NIR spectra, longitudinal non-destructive NIR spectra, and environmental records. These data sources differed in their seasonal availability, sample size, measurement type, temporal resolution, and role in the modelling framework. Table 3.1 provides an overview of the available data before the detailed data-collection procedures and exclusion decisions are described in the following subsections.

3.2.1 Data Collection Design and Measurement Procedures

Apple fruit traits were assessed approximately fortnightly, weather permitting, from 20 days after full bloom until harvest. Data collection covered four fruit-growing seasons, from November 2021 to February 2025. Pre-harvest fruit quality data were collected from a total of 1,270 'Royal Gala' fruit. The reference measurements, including dry matter content (DMC), fruit firmness, and soluble solids content (SSC), were destructive and were therefore conducted after the fruit had been removed from the trees.

Two types of NIR spectral data were collected using the same portable F-750 Produce Quality Meter (Felix Instruments, Camas, WA, USA). The instrument had a measurement range of 285–1200 nm and a spectral resolution of 3 nm (Zhang et al., 2019). The two NIR datasets differed in their sampling design and measurement setting rather than in the instrument used.

For the NIR data associated with destructive fruit assessments, ten fruit were harvested from each sampled tree at each assessment, comprising five fruit from the upper canopy and five from

Table 3.1: Summary of data sources, sample sizes, and analytical roles for pre-harvest apple firmness prediction.

Data source	**Seasons**	**Sample size and coverage**	**Variables included**	**Measurement type**	**Role in this thesis**
Fruit quality assessment data	2021–2022 to 2024–2025	1,270 fruit-level destructive assessments collected across four seasons before exclusions	DAFB, fruit firmness, DMC, SSC, fruit diameter, and fruit weight	Destructive fruit-quality measurements collected at discrete assessment dates	Provides the main response variable, apple fruit firmness, and candidate fruit-level predictors for the $h(\cdot)$ correction model.
Destructive NIR spectral data	2021–2022 and 2024–2025 used; 2023–2024 excluded	519 raw spectra: 412 in 2021–2022 and 107 in 2024–2025. After matching with reference DMC for DMC validation, 427 spectra were retained: 335 and 92, respectively	NIR spectra measured on fruit with destructive reference measurements	One-shot laboratory spectra collected from detached, intact fruit before destructive reference testing	Used to assess the quality and general usability of the NIR data through DMC validation and to provide high-dimensional spectral predictors for the $h(\cdot)$ correction model.
Non-destructive NIR spectral data	2021–2022, 2023–2024, and 2024–2025 available; 2023–2024 excluded from the main analysis	Ten labelled fruits from a single tree were repeatedly measured; the number of spectral records varied by season	Repeated NIR spectra from labelled fruit that remained attached to the tree	Longitudinal non-destructive field measurements collected across the pre-harvest period	Used to describe temporal changes in spectral profiles and to provide date-level spectral predictors for the complete $g(\cdot)$ correction model.
Environmental data	2021–2022 to 2024–2025 available; 2021–2022 and 2024–2025 used in the main analysis	Hourly records. Each complete 14-day look-back window contained 336 observations. Observation periods varied by season.	Air temperature, relative humidity, and derived vapour pressure deficit (VPD)	Hourly weather-station data from the local PFR site at Havelock North	Used to construct date-level environmental predictors for the $g(\cdot)$ correction model after temporal alignment and dimension reduction.

Note: Sample-size units differ among data sources. Fruit quality data are counted as destructively assessed fruit, NIR data are counted as spectral records, and environmental data are counted as hourly observations or complete look-back windows. The destructive NIR counts distinguish the raw spectra available before matching from the spectra successfully matched with fruit-quality reference records.

below eye level to account for within-tree variability. The harvested fruit were transported to the laboratory, where their NIR spectra were collected before the destructive reference measurements were performed. In this thesis, the term "destructive NIR spectral data" does not mean that the NIR measurement itself damaged the fruit. Instead, it refers to NIR spectra collected from fruit that had been removed from the tree and were subsequently subjected to destructive reference measurements.

For the longitudinal non-destructive NIR measurements, ten fruit were labelled on a single tree and repeatedly measured in the field while they remained attached to the tree. This allowed the spectral development of individual fruit to be tracked throughout the pre-harvest period without removing them from the tree.

Figure 3.3 shows examples of the longitudinal non-destructive NIR spectra collected during the 2024–2025 season.

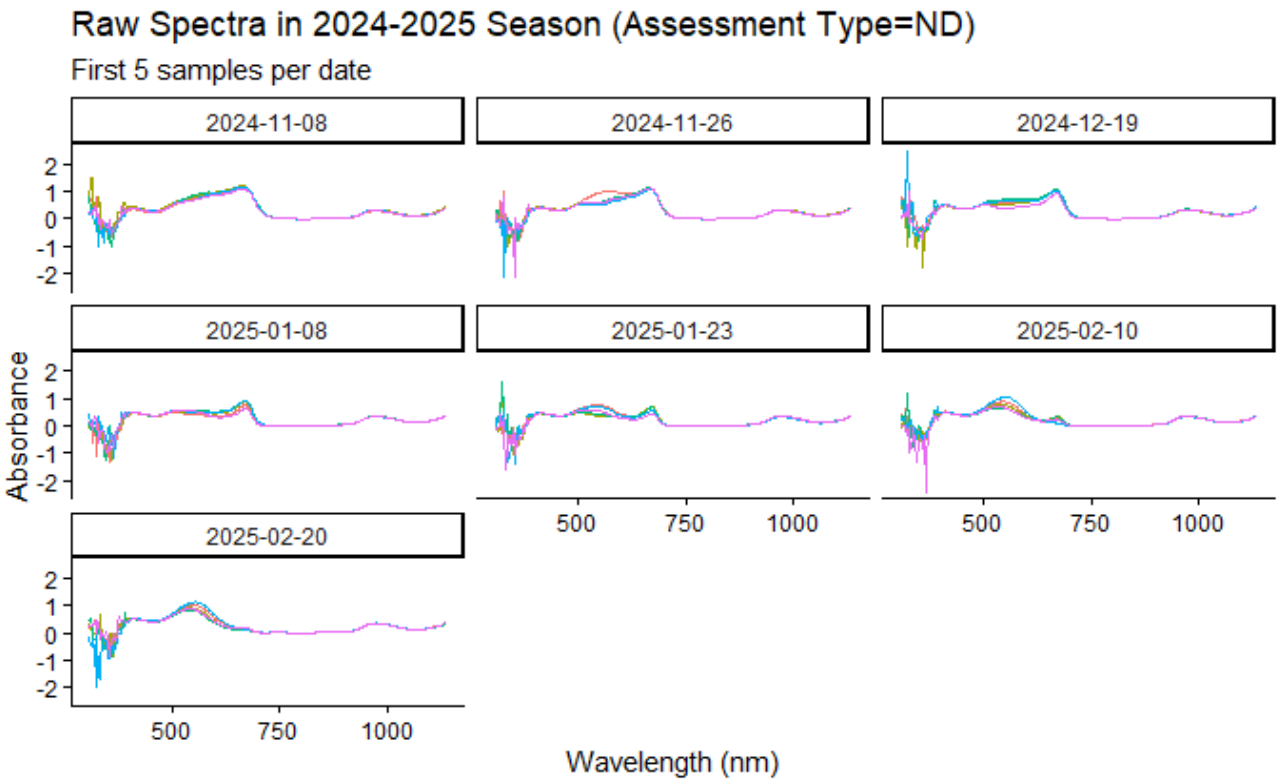


Figure 3.3: Non-destructive NIR spectra from five example fruit records at each assessment date during the 2024–2025 pre-harvest period. The spectra were collected in the field while the labelled fruit remained attached to the tree.

Environmental data were obtained from the nearest local weather station at Havelock North, located on the PFR site. Hourly weather records corresponding to the fruit-assessment periods

were downloaded from the MetWatch website in CSV format.

3.2.2 Data Availability, Matching, and Seasonal Exclusions

As summarised in Table 3.1, fruit quality assessment data were available across four growing seasons, but the availability and completeness of the NIR data varied among seasons. Longitudinal non-destructive NIR spectra were recorded during the 2021–2022, 2023–2024, and 2024–2025 seasons. However, for the 2023–2024 season, the corresponding reference DMC measurements were available only from 24 February 2024 onward. Therefore, pre-harvest reference DMC measurements required for the NIR validation procedure were not available for this season.

As a result, the 2023–2024 NIR data could not be processed and evaluated consistently with the other seasons and were excluded from the main analysis. The 2021–2022 and 2024–2025 seasons were consequently retained for model development and holdout cross-season prediction performance evaluation, respectively.

Across these two retained seasons, a total of 519 raw destructive NIR spectra were initially available: 412 spectra from the 2021–2022 season and 107 spectra from the 2024–2025 season. After the spectra were matched with the corresponding fruit quality and reference DMC records, 335 spectra were retained for 2021–2022 and 92 spectra were retained for 2024–2025, giving a total matched sample size of 427.

The raw destructive NIR measurements covered seven assessment dates in 2021–2022 and eight assessment dates in 2024–2025. After matching with reference DMC for DMC validation, spectra remained available for four and seven assessment dates, respectively. The total of 519 refers only to the one-shot laboratory spectra collected from destructively sampled fruit and does not include the repeated longitudinal spectra collected from labelled fruit on the tree.

Overall, the data availability assessment identified the 2021–2022 and 2024–2025 seasons as the two seasons containing sufficiently complete combinations of fruit quality, NIR spectral, and environmental data for the main modelling analysis.

3.2.3 Candidate Variable Availability and Selection

Based on the literature review, a conceptual variable-relationship diagram was developed to summarise the main biological and environmental pathways associated with pre-harvest apple

firmness. As shown in Figure 3.4, the diagram includes photosynthesis in the leaves, sugar metabolism in the fruit, dry matter accumulation, starch synthesis and accumulation, turgor pressure, cell-wall integrity, and fruit firmness (Lakso, 2018; Negi et al., 2025; Tijero et al., 2021; Jahed and Hirst, 2023; Musacchi and Serra, 2018). Environmental variables, including air temperature, relative humidity, and vapour pressure deficit (VPD), may influence fruit firmness indirectly through their effects on photosynthesis, fruit water status, and sugar accumulation. The diagram was used to organise the potential variables considered for the pre-harvest firmness prediction model.

The availability of the candidate variables was then checked against the data collected through the larger digital twin programme. Several variables were confirmed to be available for this study, including soluble solids content (SSC), dry matter content (DMC), apple fruit firmness, NIR spectral data, air temperature, and relative humidity. VPD could also be derived from the measured air temperature and relative humidity.

However, some potential variables were not included because their reliability in the current dataset required further validation. In particular, turgor pressure and individual sugar components were not used in the subsequent modelling. Figure 3.5 compares the available simulations and observations of individual sugar components for the 2021–2022 pre-harvest period. The simulated sucrose concentrations appear reasonably consistent with the observations, whereas the simulated glucose and fructose concentrations show substantial deviations. Therefore, the individual sugar components were excluded from this study.

3.3 Notations for Model Output and Input

In this section, the notation is introduced to define the available data and to formalize the alignment problem on the days after full bloom (`DAFB`) time scale.

3.3.1 Response of Interest

The response variable of interest is pre-harvest apple firmness. For fruit quality measurements, apple fruit firmness is measured at a sparse set of assessment times after full bloom. Therefore, we denote the sample firmness for fruit i by $y_i(t_i)$, where i indexes apple fruit and t_i denotes the assessment time point for fruit i in days after full bloom (`DAFB`). Measurements began at

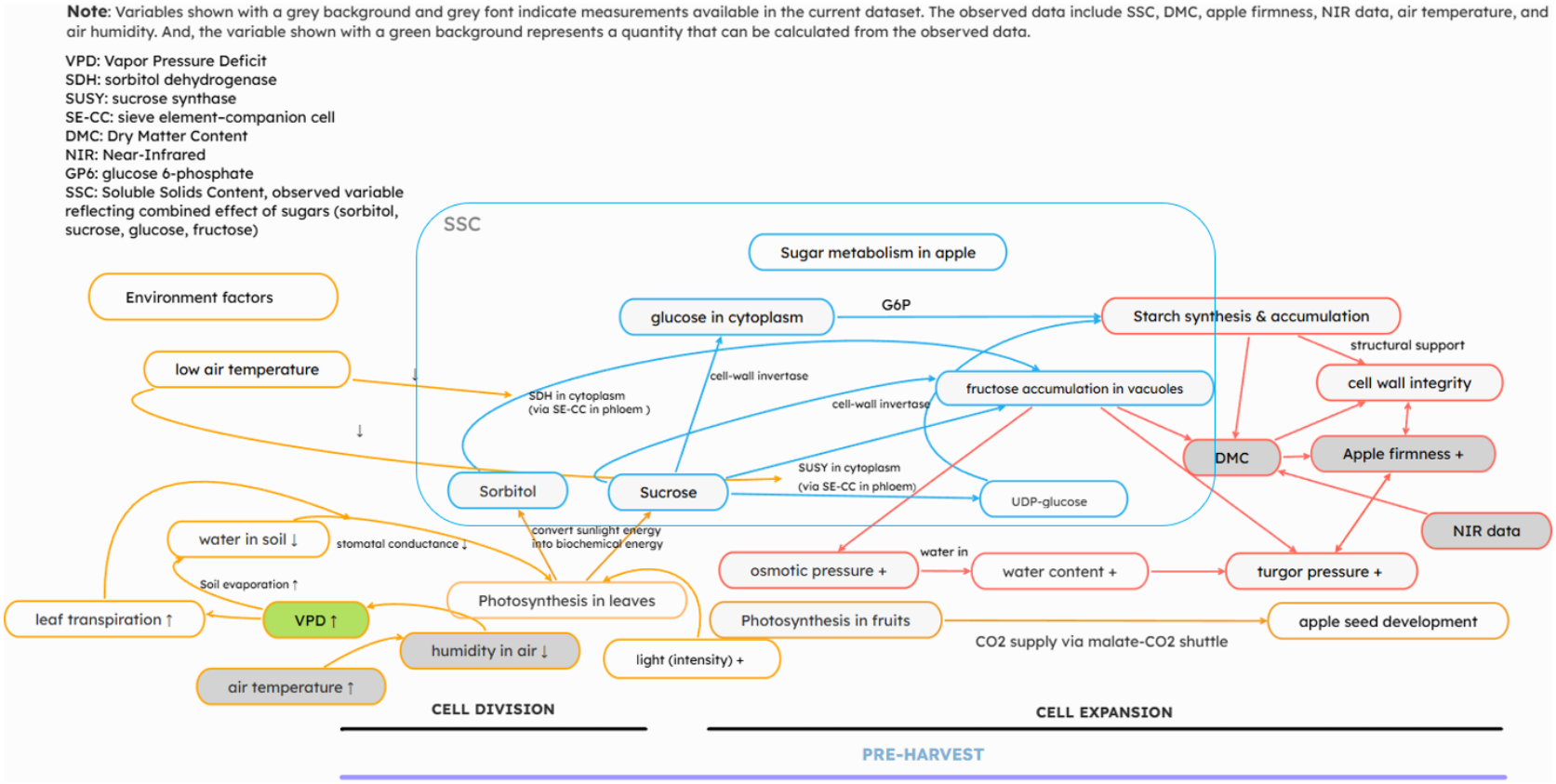


Figure 3.4: Variable relationship diagram for pre-harvest apple fruit firmness prediction. Variables shown with a grey background and grey text represent measurements available in the current dataset; variables shown with a green background represent quantities that can be derived from the observed data; orange represents environmental factors; blue represents sugar components; and red represents internal fruit states and firmness-related physiological traits.

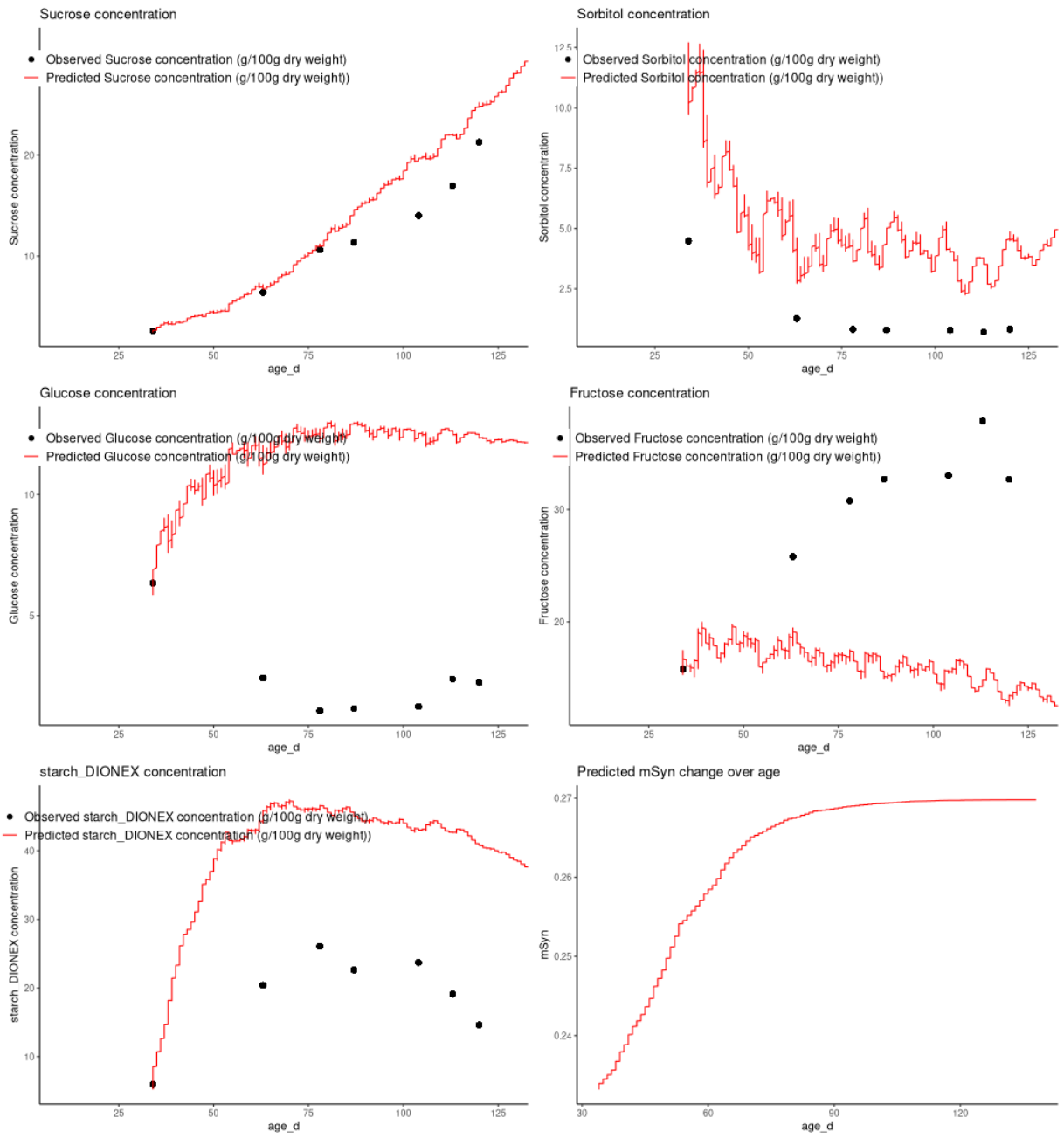


Figure 3.5: Comparison between simulations and observations of individual sugar components for the 2021–2022 pre-harvest period. *Source: Yongqi Liang, provided via personal communication.*

approximately $20 - 24$ DAFB. However, due to an inconsistency in probe type before 60 DAFB, only observations with $t_i \geq 60$ DAFB were retained. More specifically, in the 2021–2022 season, the probe used for firmness measurement was relatively large compared with the small apple fruit size at this early stage, causing the apple fruit to crack during testing.

Ideally, our target is to predict firmness for each apple fruit over the pre-harvest period, namely, to achieve an individual apple fruit firmness curve $y_i(t)$. However, in practice, this goal is difficult to achieve with the current datasets. The destructive test yields a limited set of measurements per fruit. Various key input variables that could explain differences between fruits are unavailable or very limited. For example, ethylene is not measured, and observations on individual sugar components are quite limited. Therefore, in this thesis, we choose to focus on predicting the overall mean firmness curve $\bar{y}(t)$ shown in Figure 3.6, which is more stable and better supported by the available data. Furthermore, in future work, we can target predictions of average firmness at the tree level, or for a defined block within the orchard, which may be more practical to achieve and more useful from a management perspective. Notably, the population-level mean curve $\bar{y}(t)$ is the target of the baseline component, while the complete fused model produces fruit-level predictions by modelling systematic deviations from this mean using $h(\cdot)$ and $g(\cdot)$.

3.3.2 Potential Model Input from Raw Data

The input variables can be roughly divided into the following five categories:

1. **Fruit quality measurements (destructive):** These fruit quality variables are measured destructively at a limited number of assessment times after full bloom, including fruit diameter, weight, dry matter content (DMC), and soluble solids content (SSC). Let t denote a continuous time index in days after full bloom (DAFB). In the data, the apple fruit i is assessed only at its test time t_i, so these destructively measured variables are denoted by

$$D_i(t_i),\ W_i(t_i),\ \mathrm{DMC}_i(t_i),\ \mathrm{SSC}_i(t_i),$$

   respectively, where i indexes fruit. However, when we discuss the underlying evolution over time, we write $D_i(t)$, $W_i(t)$, $\mathrm{DMC}_i(t)$, and $\mathrm{SSC}_i(t)$ as functions of continuous time t, indicating that they are only observed at the discrete times t_i.

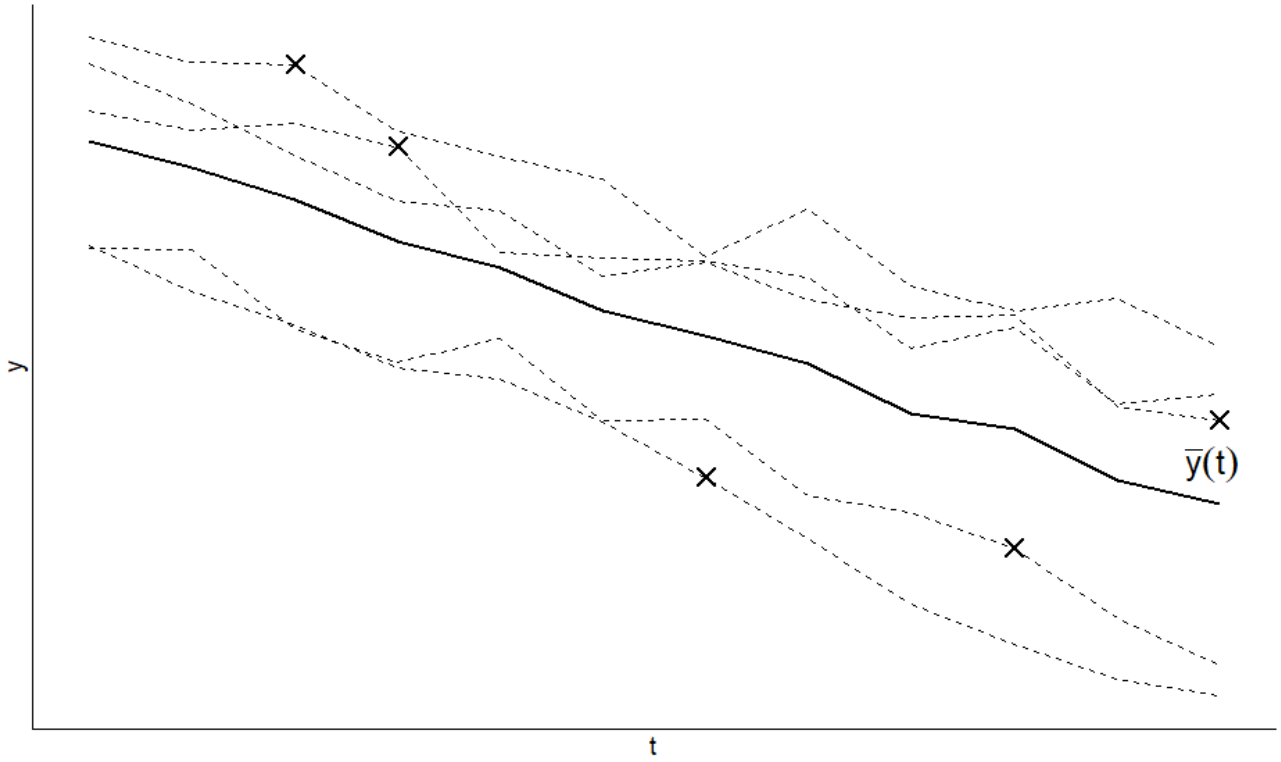


Figure 3.6: Illustration of the response structure (not real data), where cross points represent destructively measured firmness measurements $y_i(t_i)$ at tested times t_i, the dashed curves indicate the underlying pre-harvest firmness curves for individual fruits that we aim to predict, and the solid curve represents the average apple firmness curve that we try to predict, denoted by $\bar{y}(t)$.

2. **One-shot NIR spectral (destructive):** The one-shot Near-Infrared (NIR) spectra variable is measured destructively at a limited number of apple fruits at the selected assessment times after full bloom, and some spectra are missing. In our thesis, let i index fruit and let j index wavelength. Therefore, we could write the NIR spectrum for fruit i measured at time t_i as a function of wavelength:

$$X_i(t_i, \lambda_j), \qquad j = 1, \ldots, J,$$

where $\{\lambda_j\}_{j=1}^{J}$ are the recorded wavelengths. For convenience, we collect the spectrum into a vector

$$\mathbf{X}_i = \big(X_i(t_i, \lambda_1), \ldots, X_i(t_i, \lambda_J)\big)^\top.$$

3. **Longitudinal NIR spectral (non-destructive):** The longitudinal Near-Infrared (NIR) spectra are measured non-destructively using a portable NIR device (F-750 Produce Qual-

ity Meter, Felix Instruments, Camas, WA, USA) for a small set of labelled apples on a single tree at selected measurement times after full bloom. For simplicity, we do not distinguish individual fruits and treat the spectra as repeated measurements over time. In our thesis, let t denote time in days after full bloom (`DAFB`) on a continuous scale. In the data, spectra are recorded at discrete measurement times t_k (`DAFB`), so the observed longitudinal non-destructive NIR spectrum is written as

$$Z(t_k, \lambda_j), \qquad k = 1, \ldots, K; j = 1, \ldots, J,$$

where $\{\lambda_j\}_{j=1}^{J}$ are the recorded wavelength. For convenience, we collect each measurement into a vector

$$\mathbf{Z}(t_k) = \big(Z(t_k, \lambda_1), \ldots, Z(t_k, \lambda_J)\big)^\top.$$

When we discuss how the spectrum may change over time, we write $Z(t, \lambda_j)$ as a function of the continuous time t, indicating that it is only observed at the measurement times t_k. Notably, the labelled fruit used for longitudinal non-destructive NIR measurements were not directly matched with fruit-level destructive firmness measurements. Therefore, the relationship between these longitudinal spectra and apple firmness could not be directly validated in our study. Instead, the one-shot destructive NIR spectra were matched with the measured DMC measurements and used for NIR spectral dataset validation, supporting the use of longitudinal non-destructive NIR spectra in the apple firmness prediction model.

4. **Environmental factors:** For the environmental variables, sensors record measurements at a high frequency (hourly). Let $\tau \in \{\tau_1, \tau_2, \ldots, \tau_n\}$ denote the recorded time stamps in calendar time. Since our modelling time scale is days after full bloom (`DAFB`), we let $t(\tau)$ be the corresponding time in days after full bloom (`DAFB`) and map each time stamp τ to its corresponding `DAFB` time $t(\tau)$ in our analysis via the following mapping function:

$$t(\tau) = \mathrm{DAFB}(\tau),$$

We denote relative humidity, temperature, and the derived vapour pressure deficit (VPD) at time stamp τ by

$$\mathrm{RH}(t(\tau)),\ T(t(\tau)),\ V(t(\tau)).$$

respectively, indicating that the environmental values observed at the time stamp τ are expressed on the `DAFB` scale. To align the high-frequency environmental series with firmness observations, for each firmness assessment time t_i we choose a "look-back" interval $(t_i - \Delta,\ t_i]$, where Δ is a chosen window length, e.g., 7 or 14 days. Then, environmental variables are summarised over this interval to form aligned covariates at t_i, for example, using window-based averages or other rolling summaries. Moreover, the process of calculating vapour pressure deficit (VPD) from air temperature and relative humidity used in our analysis is detailed as follows:

Let T denote air temperature (in °C) and RH denote relative humidity (in %). Therefore, the saturation vapour pressure (in kPa) is computed as

$$e_s(T) = 0.611 \exp\left(\frac{17.4\,T}{239 + T}\right),$$

Then, the vapour pressure of the air (kPa) is

$$e_a(T, \mathrm{RH}) = \frac{\mathrm{RH}}{100}\, e_s(T),$$

Finally, VPD of the external air (kPa) is

$$\mathrm{VPD}(T, \mathrm{RH}) = e_s(T) - e_a(T, \mathrm{RH}).$$

5. **Estimated DMC:** The estimated DMC is not a directly measured physical fruit quality variable. Instead, it is estimated from the one-shot destructive NIR spectra variable using a partial least squares regression (PLSR) model, where the measured DMC variable is used as the reference. Therefore, we could write the estimated DMC value as $\widehat{\mathrm{DMC}}_i(t_i)$ for fruit i at assessment time t_i. Notably, because the estimated DMC is estimated rather than measured directly, it contains prediction error with uncertainty from the PLSR modelling step. In particular, this uncertainty is not propagated into the subsequent apple fruit firmness prediction model in our current study. Instead, the estimated DMC is treated as an available predictor in the apple fruit firmness prediction model.

These five categories distinguish fruit-level destructive measurements, one-shot spectra, lon-

gitudinal spectra, environmental data and model-derived predictors according to their observational and temporal scales. If more data of different types are collected, they can be classified into one of the five categories. If needed, new categories can be created to accommodate more complex data. To ensure the interpretability of the model and quantify the contribution of each category, the data of a specific category shall be incorporated into our modelling framework in a customised form, which will be explored in the following sections.

3.3.3 Data Alignment Challenges

The alignment challenge is that the environmental series $\{\mathrm{RH}(t(\tau)),\, T(t(\tau)),\, V(t(\tau))\}$ are recorded hourly, while the destructive measurements (including apple firmness) are only available at the assessment times $\{t_i\}$. The one-shot NIR spectra $X_i(t_i, \lambda_j)$ are only available at selected assessment times, and some spectra are missing. The modelling goal is to align the heterogeneous inputs,

$$\big\{D_i(t_i),\, W_i(t_i),\, \mathrm{DMC}_i(t_i),\, \mathrm{SSC}_i(t_i),\, X_i(t_i, \lambda_j),\, Z(t_k, \lambda_j),\, \mathrm{RH}(t(\tau)),\, T(t(\tau)),\, V(t(\tau)), \widehat{\mathrm{DMC}}_i(t_i)\big\},$$

to the assessment times, and use them to support the prediction of the pre-harvest firmness outcome $y_i(t_i)$.

3.3.4 Model Formulation

Finally, in this section, we formulate our modelling strategy for predicting apple fruit firmness during the pre-harvest period. The available datasets are heterogeneous and incomplete, and they are collected at different time scales with different levels of missingness. Therefore, rather than fitting a single black-box model directly to all variables, we adopt a structured formulation informed by the mechanistic information reviewed in Chapter 2. More specifically, the model first separates the common pre-harvest softening pattern from fruit-specific departures. The common pattern, including the initial firmness value, the possible presence of a knot, the expected decline over time, is motivated by the biological understanding of firmness decline during pre-harvest apple fruit development. Then, we apply a data-driven statistical correction model to explain these departures using the available data.

We start by decomposing the firmness of apple fruit i at time t_i into an overall mean compo-

nent and a fruit-specific correction term that combines mechanistic information with data-driven:

$$y_i(t) = \bar{y}(t) + \Delta y_i(t) + \varepsilon_i, \tag{3.1}$$

where $y_i(t)$ denotes the apple fruit firmness for the ith fruit at time t (`DAFB`). The $\bar{y}(t)$ represents the overall mean firmness curve over time, using the collected mechanistic information to capture the common softening process described by apple fruits. $\Delta y_i(t)$ represents the deviation of fruit i from the mean curve and will serve as the systematic correction term that captures the variation associated with the available fruit-level, NIR spectral, and environmental covariates. Moreover, ε_i represents the measurement error and the remaining unexplained fruit-level variation.

Specifically, we model the correction term $\Delta y_i(t)$ using the available input variables. This correction includes two generic data-driven components:

1. $h(\cdot)$ component summarizes the observed apple fruit information collected at the assessment time t_i.

2. $g(\cdot)$ component summarizes information from time-varying data sources.

Therefore, Equation (3.1) can be rewritten as follows:

$$\begin{aligned} y_i(t_i) = \bar{y}(t_i) &+ h\big(D_i(t_i),\, W_i(t_i),\, \mathrm{DMC}_i(t_i),\, \mathrm{SSC}_i(t_i),\, \mathbf{X}_i(t_i),\, \widehat{\mathrm{DMC}}_i(t_i)\big) \\ &+ g\big(\mathbf{Z}(\cdot),\, \mathrm{RH}(t(\tau)),\, T(t(\tau)),\, V(t(\tau))\big) + \varepsilon_i. \end{aligned} \tag{3.2}$$

where $h(\cdot)$ and $g(\cdot)$ are generic data-driven functions, which will be constructed during modelling and validation. The inputs in $h(\cdot)$ are measured at the discrete time t_i. In contrast, the longitudinal NIR records $\mathbf{Z}(\cdot)$ and the environmental series $\{\mathrm{RH}(t(\tau)),\, T(t(\tau)),\, V(t(\tau))\}$ are collected on their own time grids at hourly-level time stamps τ. $\mathbf{Z}(\cdot)$ in function $g(\cdot)$ is a surrogate for internal fruit condition and $\mathrm{RH}(t(\tau))$, $T(t(\tau))$, $V(t(\tau))$ in function $g(\cdot)$ are surrogates for external environment.

In addition, for $\bar{y}(t)$ in equation (3.1), we do not embed the full mechanistic process into the predictive model during the pre-harvest period. Instead, we use the available mechanistic information from the raw data and literature review to construct a baseline mean firmness curve $\bar{y}(t)$ that captures the overall softening pattern over time. In particular, the mechanistic information provides guidance on the key features of this baseline curve, such as the initial apple

firmness value, the possible presence of a knot, and a two-phase structure before harvest.

Specifically, to incorporate high-frequency environmental variables into the firmness model, we considered summarising the environmental variables over time intervals, e.g., considering a two-week time interval or a one-month time interval. Therefore, the environmental influence at each assessment time was represented by recent weather conditions rather than the full cumulative history of weather.

However, modelling all possible cumulative effects becomes increasingly complex when more data are collected over time. There are two main reasons: first, weather summaries from neighbouring intervals are often highly correlated, leading to collinearity and unstable estimation if all historical intervals are included; second, firmness is measured only at a limited number of assessment times, so including many interval-specific predictors would create a high-dimensional input relative to the effective sample size. Therefore, as a practical first step, we adopt a simplified alignment strategy and use only the most recent interval window to represent the environmental influence at each assessment time; for example, the second interval is used for $t^{(2)}$, the third interval for $t^{(3)}$, and so on. Notably, this look-back window is used as a pragmatic alignment choice rather than as an empirically established optimal period, since environmental effects may also operate over shorter, longer, or cumulative time scales. Figure 3.7 illustrates the alignment of high-frequency environmental variables with the firmness-assessment times.

Overall, this subsection describes the general formulation of the mechanistic-informed statistical model for pre-harvest apple fruit firmness prediction. To apply the above equations in practice, the baseline curve, the $h(\cdot)$ and $g(\cdot)$ correction terms, and their corresponding input variables need to be constructed and estimated. Therefore, the following subsection describes the fusion procedure used to make the equations operational, including baseline estimation, covariate construction and alignment, residual correction modelling, and final fused prediction.

3.3.5 Fusion Specification

Following the above subsections, we express the real apple firmness as Equation 3.2 and the model fusion is implemented using the following four-step procedure, which is described below:

- **Step 1:** Estimation of the mechanistic baseline depends on known mechanistic information. We first estimate $\bar{y}(t_i)$ using the two-window structure, representing the early and

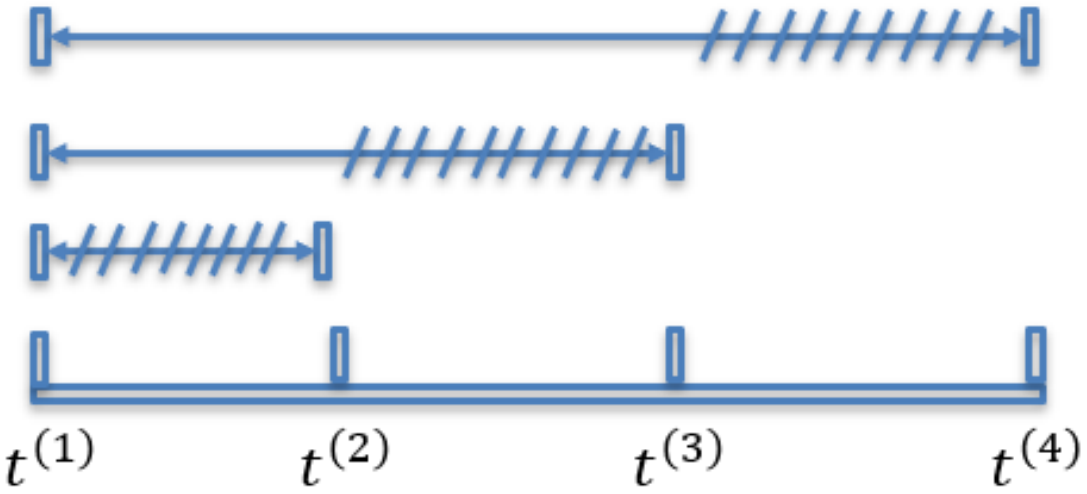


Figure 3.7: Illustration of the temporal alignment of high-frequency environmental variables with firmness assessment times. For each assessment time $t^{(k)}$, the hatched segment shows the selected time interval over which weather measurements are summarised and used as environmental inputs for predicting firmness at $t^{(k)}$. In our project, we only use the most recent interval rather than the full cumulative enviromental history.

later stages of firmness decline during the pre-harvest period. In practice, $\bar{y}(\cdot)$ can be specified as a change-point model with different decline rates before and after the change point.

- **Step 2:** Construction and alignment of covariates. For each firmness observation time t, we construct matched predictor inputs. In particular, the destructive NIR spectra are aligned with destructively tested fruit at the assessment time and used as fruit-level spectral predictors. The non-destructive NIR spectra are summarised at the measurement-date level, and then reduced to a small number of functional principal component scores. For the environmental variables, observations from the preceding interval are extracted for each assessment time and represented as environmental curves. Then, functional principal component analysis is used to reduce these high-dimensional environmental curves into low-dimensional scores. In particular, since the environmental variables are shared by apple fruits measured on the same date, the environmental predictors are aggregated to the measurement-date level before fitting the $g(\cdot)$ correction model.

- **Step 3:** Residual modelling to construct the statistical correction terms. After fitting the

baseline model, the $h(\cdot)$ correction model is fitted to the fruit-level baseline residuals using elastic net regression. Furthermore, after the $h(\cdot)$ correction, the remaining residuals are aggregated to the measurement-date level. More specifically, the environmental-only $g(\cdot)$ correction is fitted using weighted linear regression. Then, the complete $g(\cdot)$ correction, incorporating environmental and non-destructive NIR functional principal component scores, is fitted using the weighted ridge regression.

- **Step 4:** Fused prediction. For prediction at a new time point $t^\star$, we combine the fitted baseline and correction components:

$$\begin{aligned}\widehat{y}(t^\star) = \widehat{\bar{y}}(t^\star) + \widehat{h}\big(D(t^\star),\, W(t^\star),\, \mathrm{DMC}(t^\star),\, \mathrm{SSC}(t^\star),\, \mathbf{X}(t^\star),\, \widehat{\mathrm{DMC}}_i(t^\star)\big) \\ + \widehat{g}\big(\mathbf{Z}(\cdot),\, \mathrm{RH}(t^\star),\, T(t^\star),\, V(t^\star)\big).\end{aligned} \tag{3.3}$$

  where $\widehat{g}(t^\star)$ is evaluated using the aligned NIR and environmental features at $t^\star$. This yields a final prediction that retains the mechanistic decline pattern while incorporating additional information from Vis/NIR spectra and environmental conditions in a controlled way.

In particular, the proposed model-fusion approach uses existing datasets from previous seasons as inputs, including destructive fruit measurements, destructive and non-destructive NIR spectra, and environmental factors. In addition, the framework can incorporate estimated input variables from PFR's fruit model, which are generated via mechanistic modelling. Overall, the above sections define the overall modelling framework and the roles of different input categories. The detailed implementation steps, including data cleaning, data pre-processing, NIR spectral data pre-processing, environmental alignment, dimension reduction, and model fitting, are presented in Chapter 4.

Chapter 4

Model Implementation

This chapter describes the implementation of the proposed mechanistic-informed statistical modelling framework for pre-harvest apple fruit firmness prediction. Based on the model formulation in Chapter 3, the developed model applies a multi-step fusion strategy. Namely, the sample apple fruit firmness is decomposed into a biologically motivated baseline softening trajectory and two additional statistical correction terms. More specifically, the baseline model first builds a mean firmness curve over days after full bloom (`DAFB`) to capture the common pre-harvest softening trajectory, using the known mechanistic information. Then, residual modelling is used to construct two correction terms sequentially. The first correction term $h(\cdot)$ models the residual between the sample firmness and the fitted baseline firmness using fruit-level measurements and destructive Near-Infrared (NIR) spectral variables. After the first correction term modelling, the second correction term ($g(\cdot)$) further models the remaining residual using environmental information and the non-destructive NIR spectral variables.

Notably, the model implementation was conducted in `R`. In our study, the 2021–2022 season is mainly used to construct and fit the proposed mechanistic-informed statistical modelling. Then, the 2024–2025 season is used as a holdout test season to evaluate the predictive performance of the fitted mechanistic-informed model in a new season. Section 4.1 of this Chapter describes the data pre-processing and validation steps used to prepare the fruit quality measurement data, NIR spectral data, estimated DMC, and environmental data for modelling. In Section 4.2, the baseline model is implemented to capture the common pre-harvest apple firmness decline over `DAFB`. Also, the residual correction model is implemented in this Section, including the fruit-level ($h(\cdot)$) correction term and the date-level ($g(\cdot)$) correction term. Finally, Section 4.3 presents the

prediction performance on the 2024–2025 season.

4.1 Data Pre-processing and Validation

To implement the proposed mechanistic-informed statistical modelling framework, we performed various data pre-processing and validation steps for different raw datasets that could be used for the modelling, including fruit quality measurement data cleaning, Near-Infrared (NIR) spectral data validation, NIR spectral data pre-processing, dry matter content (DMC) estimated from spectral NIR data, and environmental data pre-processing.

4.1.1 Fruit Quality Measurement Data Cleaning

The fruit quality measurement datasets from the 2021–2022 and 2024–2025 seasons were checked and cleaned separately. Based on the discussion in Chapter 3, the variables used in the modelling include measurement date, days after full bloom (`DAFB`), fruit number, assessment type, fruit weight, fruit diameter, soluble solids content (SSC), dry matter content (DMC), and apple fruit firmness. However, in the raw datasets, these variables were not always recorded using the same names across seasons. Therefore, the above variables need to be renamed and standardised before the two seasons are combined.

For the 2021–2022 season, the raw fruit quality data were filtered to retain destructive observations only. The variable `DestructiveND` was used to identify destructive measurements. Only observations collected during the pre-harvest period were retained. The selected measurement window started from the first available assessment date after 60 `DAFB` (`DAFB` = 63) and ended at the final pre-harvest assessment date before harvest, which was 07 February 2022 (`DAFB` = 113). Furthermore, the observations with missing apple firmness values were removed because apple fruit firmness is the response variable in our study.

For the 2024–2025 season, the raw fruit quality data were cleaned using a similar process. Firstly, only observations with cultivar recorded as `Gala` were retained. Only observations collected during the pre-harvest period were retained. The selected measurement window started from the first available assessment date after 60 `DAFB` (`DAFB` = 65) and ended at the final pre-harvest assessment date before harvest, which was 10 February 2025 (`DAFB` = 118). Furthermore, the observations with missing apple firmness values were removed because apple fruit firmness

is the response variable in our study.

After the above cleaning steps, the two datasets were converted into a consistent format. The standardised variables include `season`, `cultivar`, `measure_date`, `full_bloom_date`, `dafb`, `assessment_type`, `fruit_number`, `weight`, `Diameter`, `ssc`, `DMC_ref`, and `firmness`. Finally, the cleaned 2021–2022 and 2024–2025 datasets were then combined into one modelling dataset.

4.1.2 NIR Spectral Data Validation and Pre-processing

Before model construction, the destructive NIR spectra were validated using measured dry matter content (DMC) as the reference variable. This analysis assessed the quality of the spectral data and whether the spectra contained useful information for predicting an important internal fruit quality variable. The longitudinal non-destructive NIR spectra could not be directly validated against apple firmness because matched fruit-level firmness measurements were unavailable. Therefore, the results provide indirect support for their general usability rather than direct validation for firmness prediction.

For the 2021–2022 season, the files corresponding to destructive measurements were extracted from the absorbance spectral files and combined into one dataset. The spectra were matched with the fruit quality measurement data using measurement date and within-date sample order, and the sample counts were checked for each date. Table 4.1 summarises the numbers of destructive NIR samples before and after matching. In this season, the number of observations varied substantially across measurement dates. The original sample-size distribution was retained, and no weighting or resampling was applied to balance the dates. Consequently, measurement dates with larger sample sizes may have had greater influence on the fitted PLSR models, which should be considered when interpreting the validation results.

Table 4.1: The number of destructive NIR samples before and after matching with the fruit quality measurement data (2021–2022 season)

Measurement Date	Sample Counts	Sample Retained
2021-12-02	210	173
2021-12-20	10	9
2021-12-21	152	148
2022-01-05	10	–
2022-01-14	10	5
2022-01-21	10	–
2022-01-31	10	–

For the validation analysis, wavelengths from 390 to 1101 nm were retained. All-zero and near-constant wavelength columns, incomplete observations, and duplicated spectral rows were removed. Individual zero values within otherwise retained spectra were treated as missing values and linearly interpolated along the wavelength direction. Potential outliers were identified using principal component analysis (PCA) scores and PLSR residuals, with samples showing large Mahalanobis distances or large residuals removed from the validation dataset.

Different spectral pre-processing methods were compared, including multiplicative scatter correction (MSC), standard normal variate (SNV), Savitzky–Golay smoothing and derivatives, standardisation, and combinations of these modules. Table 4.2 summarises the modules included in each pre-processing method.

Table 4.2: Pre-processing modules and their combinations used in the NIR analysis

Method Name	**Scatter correction**		**Savitzky–Golay**			**Standardisation**
	MSC	**SNV**	$m = 0$	$m = 1$	$m = 2$	
`raw`	×	×	×	×	×	×
`msc`	✓	×	×	×	×	×
`snv`	×	✓	×	×	×	×
`sg(deriv1)`	×	×	×	✓	×	×
`sg(deriv2)`	×	×	×	×	✓	×
`msc+sg(deriv1)`	✓	×	×	✓	×	×
`snv+sg(deriv1)`	×	✓	×	✓	×	×
`msc+smooth`	✓	×	✓	×	×	×
`msc+smooth+scale`	✓	×	✓	×	×	✓
`msc+sg(deriv1)+scale`	✓	×	×	✓	×	✓
`sg(deriv1)+scale`	×	×	×	✓	×	✓
`snv+sg(deriv1)+scale`	×	✓	×	✓	×	✓

Note: The Savitzky–Golay filter used a polynomial order of $p = 2$ and a window length of $w = 11$ wavelength points. The derivative orders $m = 0$, $m = 1$, and $m = 2$ denote smoothing, first derivative, and second derivative, respectively. Standardisation was performed wavelength-wise by centring each wavelength variable and scaling it to unit standard deviation. When multiple modules were included, they were applied in the following order: MSC or SNV, Savitzky–Golay processing, and standardisation.

For the 2021–2022 season, two situations were compared: (1) the full period without the time variable and (2) the full period with the time variable. According to Xiao and Li (2022), root mean squared error (RMSE), mean absolute error (MAE), bias, coefficient of determination (R^2), and the optimal number of PLS components were used for model comparison. The full period without the time variable gave the best overall results. As shown in Table 4.3, SNV achieved the highest R^2 of 0.735 and the lowest RMSE of 0.604. Because the 2021–2022 season was used for model development, SNV was selected as the common spectral pre-processing method for the

subsequent modelling framework. Figure 4.1 shows the corresponding predicted and measured DMC values for the different pre-processing methods.

Table 4.3: Performance comparison of different pre-processing methods for NIR dry matter content prediction in the $2021 - 2022$ season without the time variable

Method	**RMSE**	**MAE**	**Bias**	R^2	**RPD**	**Components**
`raw`	0.657	0.511	-1.26×10^{-15}	0.686	1.785	15
`msc`	0.665	0.514	-1.31×10^{-15}	0.678	1.762	14
`snv`	0.604	0.465	6.55×10^{-15}	0.735	1.943	22
`sg(deriv1)`	0.636	0.492	5.89×10^{-16}	0.706	1.844	18
`sg(deriv2)`	0.651	0.517	2.49×10^{-16}	0.692	1.802	12
`msc+sg(deriv1)`	0.640	0.490	1.73×10^{-15}	0.702	1.832	18
`snv+sg(deriv1)`	0.638	0.487	-3.29×10^{-16}	0.704	1.838	18
`msc+smooth`	0.654	0.503	7.90×10^{-16}	0.689	1.793	17
`msc+smooth+scale`	0.675	0.513	1.06×10^{-16}	0.669	1.738	15
`msc+sg(deriv1)+scale`	0.667	0.510	1.01×10^{-16}	0.676	1.757	14
`sg(deriv1)+scale`	0.693	0.526	1.59×10^{-17}	0.651	1.693	8
`snv+sg(deriv1)+scale`	0.672	0.513	8.48×10^{-17}	0.672	1.746	13

For the 2024–2025 season, the same validation framework was applied. Only 'Gala' apples with destructive measurements and available reference DMC values were retained. Table 4.4 summarises the numbers of destructive NIR samples before and after matching with the fruit quality measurement data.

Table 4.4: The number of destructive NIR samples before and after matching with the fruit quality measurement data (2024–2025 season)

Measurement Date	**Sample Counts**	**Sample Retained**
2024-11-08	12	12
2024-11-26	10	10
2024-12-19	10	10
2025-01-08	10	10
2025-01-23	10	10
2025-02-10	15	–
2025-02-19	20	20
2025-02-28	20	20

The retained counts in Tables 4.1 and 4.4 refer specifically to spectra successfully matched with reference DMC for the DMC validation analysis. Spectra without reference DMC could still be retained for the firmness model when they were successfully matched with the firmness outcome and the common firmness-model predictors.

Four situations were compared for this season: (1) the full period without the time variable, (2) the full period with the time variable, (3) the later period from January to February without

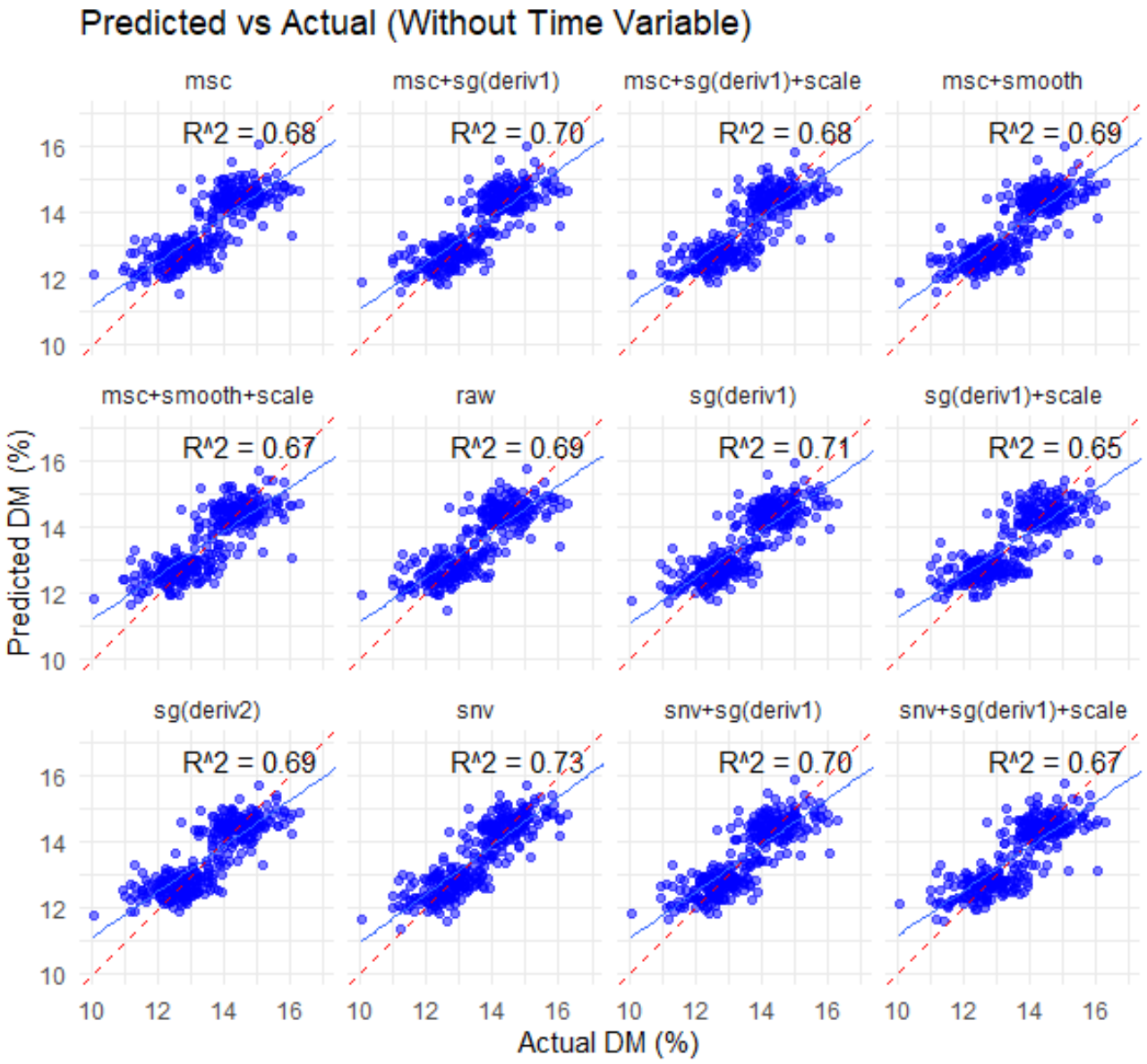


Figure 4.1: Predicted versus measured DMC for each pre-processing method, with the corresponding R^2 annotated for each panel in the $2021-2022$ season without the time variable. The red dashed line is the ideal $y = x$ line and the blue solid line is the fitted regression line.

the time variable, and (4) the later period from January to February with the time variable. The later period without the time variable gave the best overall results. As shown in Table 4.5, the Savitzky-Golay first derivative achieved the highest R^2 of 0.649 and the lowest RMSE of 1.04 in this season-specific comparison. However, the 2024–2025 season was treated as a holdout test season, and these results were not used to reselect or optimise the pre-processing method

for the final modelling framework. Instead, SNV, which had been selected using the 2021–2022 model-development data, was retained as the common pre-processing method. Figure 4.2 shows the corresponding predicted and measured DMC values.

Table 4.5: Performance comparison of different pre-processing methods for NIR dry matter content prediction during the later period (January–February 2025) in the 2024–2025 season without the time variable

Method	**RMSE**	**MAE**	**Bias**	R^2	**RPD**	**Components**
`raw`	1.080	0.866	-5.92×10^{-17}	0.617	1.616	9
`msc`	1.190	0.941	7.11×10^{-16}	0.541	1.476	8
`snv`	1.130	0.850	9.47×10^{-16}	0.587	1.556	9
`sg(deriv1)`	1.040	0.843	-5.92×10^{-16}	0.649	1.688	10
`sg(deriv2)`	1.410	1.11	-6.51×10^{-16}	0.357	1.247	4
`msc+sg(deriv1)`	1.170	0.946	-1.18×10^{-15}	0.553	1.496	8
`snv+sg(deriv1)`	1.350	1.06	-3.85×10^{-16}	0.405	1.296	6
`msc+smooth`	1.340	1.05	-2.96×10^{-16}	0.413	1.305	6
`msc+smooth+scale`	1.280	0.994	-1.18×10^{-17}	0.466	1.368	7
`msc+sg(deriv1)+scale`	1.110	0.886	-8.59×10^{-16}	0.601	1.583	8
`sg(deriv1)+scale`	1.140	0.880	-4.44×10^{-16}	0.575	1.534	4
`snv+sg(deriv1)+scale`	1.090	0.855	-1.78×10^{-16}	0.617	1.616	8

Overall, the validation results indicate that the destructive NIR spectra contained useful information for predicting DMC and support the use of NIR-derived variables in the subsequent modelling framework. Although the best-performing pre-processing method differed between the two season-specific analyses, SNV was applied consistently to the destructive and non-destructive NIR spectra from both seasons. This common treatment was selected using the 2021–2022 model-development data and retained to ensure consistency and comparability across datasets, while avoiding adaptation of the modelling procedure to the 2024–2025 holdout test season.

For the subsequent firmness modelling, SNV was applied consistently to the destructive NIR spectra from both seasons. The SNV-transformed spectra were used both as candidate predictors in the fruit-level $h(\cdot)$ correction model and as inputs to the PLSR model for estimated DMC. For the longitudinal non-destructive NIR spectra, all-zero and near-constant wavelength columns, incomplete rows, and duplicated spectra were removed before applying SNV. The transformed spectra were then summarised at the measurement-date level for the $g(\cdot)$ correction model.

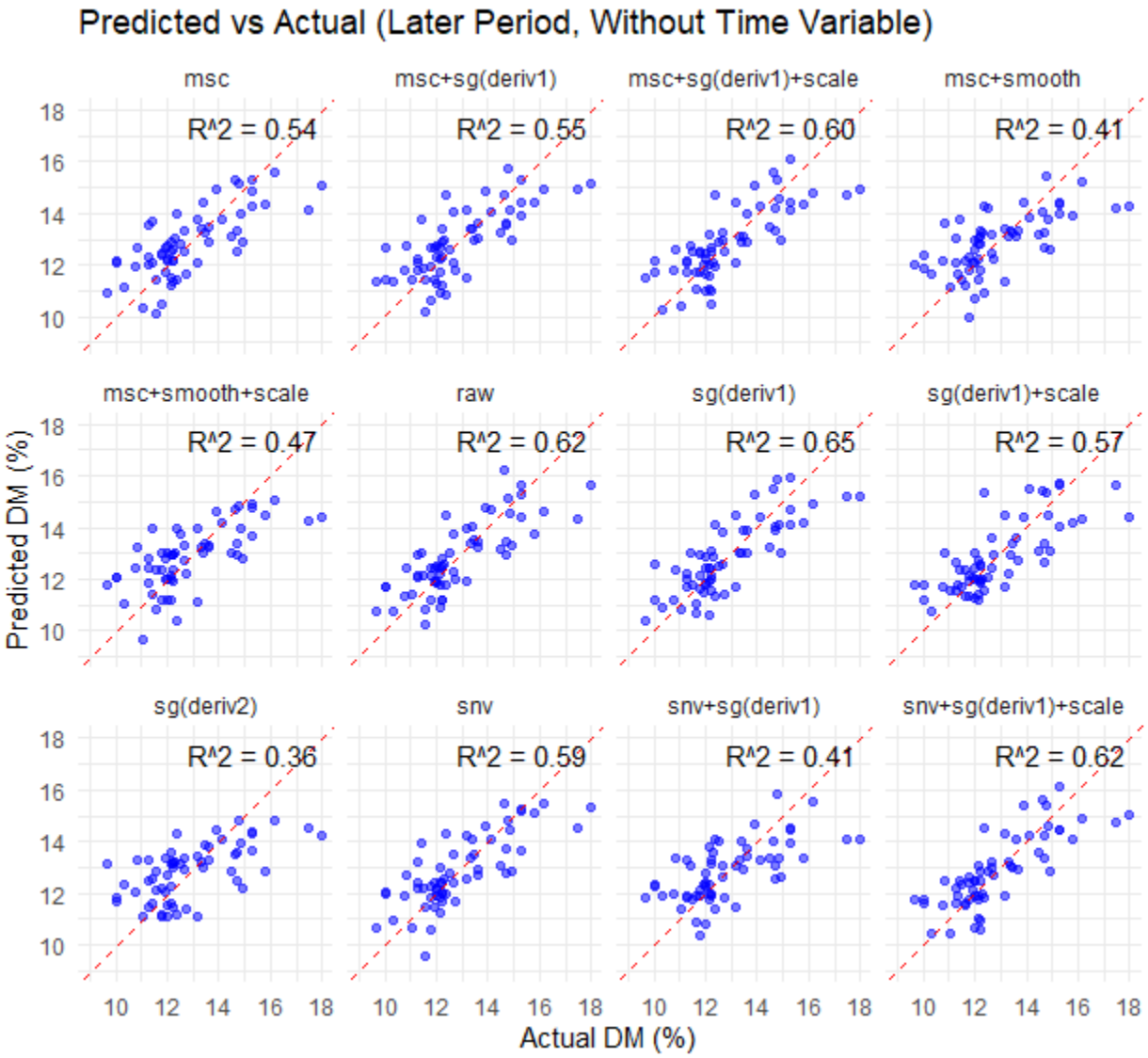


Figure 4.2: Predicted versus measured DMC for each pre-processing method, with the corresponding R^2 annotated for each panel during the later period (January–February 2025) in the 2024–2025 season without the time variable

4.1.3 Environmental Data Pre-processing

The environmental data were pre-processed before being included in the $g(\cdot)$ correction model. Air temperature and relative humidity were extracted from the hourly weather records, and vapour pressure deficit (VPD) was calculated in kPa using the equations given in Section 3.3.2. Because the VPD distribution was right-skewed, a logarithmic transformation was applied before standardisation.

To preserve the separation between model development and cross-season evaluation, the means and standard deviations used to standardise air temperature, relative humidity, and log-

transformed VPD were estimated using the 2021–2022 development-period records only. The same transformation parameters were then applied unchanged to the 2024–2025 environmental records.

For each unique firmness assessment date, a 14-day look-back window was constructed from the hourly environmental data. Each complete window contained 336 hourly observations. Since all fruit assessed on the same date shared the same environmental history, one environmental curve was constructed per assessment date rather than repeating the same curve for every fruit. Functional principal component analysis was fitted using the development-season date-level curves, and the 2024–2025 curves were subsequently projected onto the fitted development-season basis.

Notably, these environmental data measurements from the weather station were treated as proxies for local orchard-level environmental conditions, and the direct measurements of the microclimate within the tree canopy were not available.

4.2 Model Development Based on the 2021-2022 Season

4.2.1 Exploratory Modelling Using Generalised Additive Models

To investigate the relationship between days after full bloom (DAFB) and apple fruit firmness, generalised additive models (GAMs) were first used as an exploratory tool. The purpose of this analysis was to assess whether the firmness trajectory showed evidence of nonlinearity and to determine an appropriate level of complexity for the subsequent baseline model. The GAM analysis was used only to explore the shape of the firmness trajectory and was not used as the final prediction model.

A GAM extends a conventional linear regression model by replacing some linear predictor terms with unknown smooth functions. A general GAM can be written as (Wood, 2017)

$$g(\mu_i) = \beta_0 + \sum_{j=1}^{p} f_j(x_{ij}), \qquad \mu_i = \mathrm{E}(Y_i \mid x_{i1}, \ldots, x_{ip}), \tag{4.1}$$

where Y_i is the response for observation i, $g(\cdot)$ is a link function, β_0 is the intercept, x_{ij} is the jth predictor for observation i, and $f_j(\cdot)$ is an unknown smooth function estimated from the data.

In the present analysis, the response variable was apple fruit firmness and the main continuous

predictor was fruit age in DAFB. For the untransformed firmness response, the basic Gaussian GAM with an identity link can be expressed as

$$Y_i = \beta_0 + f(t_i) + \varepsilon_i, \qquad \varepsilon_i \overset{\text{iid}}{\sim} N(0, \sigma^2), \tag{4.2}$$

where Y_i denotes the firmness of fruit observation i, t_i denotes its fruit age in DAFB, and $f(t_i)$ represents the smooth firmness trajectory over fruit development.

The complexity of an estimated smooth term can be summarised by its effective degrees of freedom (edf). An edf close to 1 indicates an approximately linear relationship, whereas an edf greater than 1 indicates increasing evidence of curvature. Thus, the edf values were used here to assess whether the relationship between DAFB and firmness could be represented by a low-complexity trajectory or required a highly flexible nonlinear function.

Five GAM specifications were compared. Let c_i denote the growing season associated with observation i, α_{c_i} denote the parametric effect of season, $f(t)$ denote a common smooth function shared by both seasons, and $f_{c_i}(t)$ denote a season-specific smooth function. The five model structures can be written as

$$\begin{aligned}
\mathcal{M}_1 : &\quad Y_i = \beta_0 + f(t_i) + \varepsilon_i, \\
\mathcal{M}_2 : &\quad Y_i = \beta_0 + \alpha_{c_i} + f(t_i) + \varepsilon_i, \\
\mathcal{M}_3 : &\quad Y_i = \beta_0 + \alpha_{c_i} + f_{c_i}(t_i) + \varepsilon_i, \\
\mathcal{M}_4 : &\quad \log(Y_i) = \beta_0 + \alpha_{c_i} + f(t_i) + \varepsilon_i, \\
\mathcal{M}_5 : &\quad \log(Y_i) = \beta_0 + \alpha_{c_i} + f_{c_i}(t_i) + \varepsilon_i.
\end{aligned} \tag{4.3}$$

The models therefore differed in whether seasonal differences were ignored, represented by a parametric season effect, or represented by season-specific smooth functions, and whether firmness was modelled on its original or logarithmic scale. The following results were obtained:

- `model1`: When the two seasons were pooled without accounting for seasonal differences, the smooth term had a high edf of 8.161. This indicates that a highly flexible function was required to describe the pooled data.

- `model2`: After including season as a parametric effect, the edf decreased substantially to 1.894. This suggests that much of the apparent complexity in the pooled model was

attributable to between-season differences rather than to strong within-season nonlinearity.

- `model3`: When season-specific smooth functions were fitted, the 2021–2022 season had an edf of 2.725, whereas the 2024–2025 season had an edf of 1.623. Thus, the firmness trajectory in 2021–2022 showed a low-complexity nonlinear pattern, while the 2024–2025 trajectory was closer to linear.

- `model4`: After applying a logarithmic transformation to firmness and including season as a parametric effect, the common smooth term had an edf of 1.906. This indicates that the overall relationship remained mildly nonlinear on the logarithmic scale.

- `model5`: When season-specific smooth functions were fitted on the logarithmic scale, the 2021–2022 season had an edf of 2.917, whereas the 2024–2025 season had an edf of approximately 1. This indicates that the 2024–2025 season followed an approximately linear trajectory, while the 2021–2022 season retained a low-complexity but meaningful nonlinear pattern.

Overall, apple fruit softening is biologically expected to follow a mildly nonlinear trajectory during the pre-harvest period (Volz et al., 2003). The GAM results indicated that the 2021–2022 season contained clearer evidence of such a low-complexity nonlinear pattern and was therefore considered more suitable for baseline estimation. This result supports the subsequent use of a low-complexity one-knot spline model to represent the baseline firmness trajectory. In contrast, the 2024–2025 season was approximately linear over the available modelling period and was therefore retained for holdout cross-season prediction performance evaluation with season-specific recalibration. Consequently, the 2021–2022 season was used for model development and refinement, whereas the 2024–2025 season was used to evaluate the fitted modelling framework.

4.2.2 Baseline Modelling

To build the baseline model, the 2021–2022 season data were used to fit a two-phase one-knot linear regression spline model with a literature-based fixed initial firmness. Furthermore, this baseline model was implemented using the development function `fit_baseline_tp()` in `R`. This is consistent with the exploratory GAM results, which indicate that the apple firmness softening trend is mildly nonlinear rather than highly complex. Therefore, a low-complexity spline model

is appropriate for baseline construction, because it can capture the main apple firmness decline trend while still keeping the model simple and interpretable. Notably, the initial firmness is used as a season-specific calibration value. Therefore, although it is fixed during the fitting of a given baseline model, its value can be adjusted when the fitted baseline structure is transferred to a different season, especially for further testing. The following eight steps for constructing the two-phase regression spline baseline model using the development function `fit_baseline_tp()` in `R` are described below:

- **Step 1:** For the 2021–2022 season, the data were split into a training set and a test set based on the `DAFB` range. Specifically, the training set contained observations with non-missing `DAFB` and firmness values and with `DAFB` ≤ 100, whereas the test set contained observations at the target prediction fruit age, namely `DAFB` $= 105$.

- **Step 2:** Using the training data, the empirical mean firmness curve was constructed by grouping observations at each `DAFB`. For each `DAFB`, the mean firmness, standard deviation (sd), and sample size were calculated.

- **Step 3:** The reference starting `DAFB`, denoted by t_0, was set as the minimum `DAFB` in the training set. Fruit age was then centred at t_0. The literature-based fixed initial apple firmness was set as $F_0 = 12.1$. Therefore, the sample mean firmness measured in the lab was treated as a deviation from F_0.

- **Step 4:** A search function was defined to identify the optimal knot over the `DAFB` interval from 70 to 100. For each candidate knot, a two-phase linear spline model with the fixed initial firmness F_0 was fitted. Also, the model performance was evaluated using `AIC`, `BIC`, and `RMSE`. As a result, only those candidate knots that satisfy the literature-based condition of a steeper decline before the knot and a slower decline after the knot are retained.

- **Step 5:** The best knot was selected based on the candidate knot results. Furthermore, the final two-phase baseline model was fitted to the empirical mean firmness data using weighted least squares, where the weights were given by the number of observations at each `DAFB`.

- **Step 6:** Using the final fitted model, the fitted baseline firmness values were obtained over

a prediction grid from the minimum `DAFB` to the target prediction `DAFB`. The corresponding slopes before and after the estimated knot were also recorded.

- **Step 7:** The baseline residual for each fruit was calculated as the difference between the sample apple firmness and the predicted baseline firmness. Moreover, the predicted baseline firmness values were also attached back to the fruit-level observations.

- **Step 8:** The sample mean firmness curve and the fitted baseline curve were constructed for visualisation. Therefore, the sample apple firmness trajectory and the fitted-plus-predicted baseline apple firmness trajectory can be compared directly on the same figure.

4.2.3 Residual Correction Modelling

After the baseline model was fitted, residual correction modelling was used to determine whether the available fruit-level, spectral, and environmental information could explain variation that was not captured by the baseline firmness trajectory. The correction procedure was implemented sequentially. First, the fruit-level correction model $h(\cdot)$ was fitted to the residuals from the baseline model. Second, the remaining residuals were aggregated to the measurement-date level and modelled using the date-level correction model $g(\cdot)$.

For fruit i measured at time t_i, the residual from the baseline model was defined as

$$r_i^{(0)} = y_i(t_i) - \widehat{\bar{y}}(t_i), \tag{4.4}$$

where $y_i(t_i)$ is the observed fruit firmness and $\widehat{\bar{y}}(t_i)$ is the fitted two-phase baseline firmness at the corresponding DAFB.

Fruit-level correction model $h(\cdot)$

To allow the same fruit-level correction structure to be developed using the 2021–2022 season and subsequently applied to the 2024–2025 season, the candidate predictor set was restricted to variables available in both seasons. These predictors included fruit weight, soluble solids content (SSC), and the SNV-transformed destructive NIR wavelengths common to the two seasonal datasets. Fruit diameter, measured DMC, and estimated DMC were not included in this common correction model because they were not consistently available in the 2024–2025 season. The

estimated DMC analysis described in Section 4.1.2 was therefore used primarily to assess the general quality and usability of the destructive NIR spectra rather than as an input to the final cross-season correction model.

Because the destructive NIR spectra contained a large number of highly correlated wavelength variables, elastic net regression was used to model the baseline residuals. The model was fitted using the `glmnet` function in `R` with $\alpha = 0.5$. The regularisation parameter λ was selected by leave-one-assessment-date-out cross-validation. More specifically, all fruit observations collected on the same assessment date were assigned to the same cross-validation fold. This prevented observations sharing the same date-specific growing and measurement conditions from being divided between the fitting and validation subsets.

The $h(\cdot)$ model was fitted using the 2021–2022 observations with `DAFB` < 100, while observations at `DAFB` $= 105$ were retained as a within-season prediction set. Let $\widehat{h}(\mathbf{x}_i)$ denote the predicted baseline residual from the fitted elastic net model, where $\mathbf{x}_i$ contains the available fruit-level and destructive NIR predictors. The firmness prediction after the first correction was calculated as

$$\widehat{y}_i^{(h)}(t_i) = \widehat{\bar{y}}(t_i) + \widehat{h}(\mathbf{x}_i). \tag{4.5}$$

The residual remaining after the $h(\cdot)$ correction was then

$$r_i^{(1)} = y_i(t_i) - \widehat{y}_i^{(h)}(t_i). \tag{4.6}$$

These remaining residuals were used as the response for the subsequent $g(\cdot)$ correction models.

Date-level correction model $g(\cdot)$

The environmental variables and longitudinal non-destructive NIR summaries were shared by fruit assessed on the same date. Therefore, fruit-level observations did not provide independent information for estimating their effects. To avoid treating repeated date-level information as independent, the residuals after the $h(\cdot)$ correction were aggregated to the measurement-date level.

For assessment date d, the mean remaining residual was defined as

$$\bar{r}_d^{(1)} = \frac{1}{n_d} \sum_{i:\, d(i)=d} r_i^{(1)}, \tag{4.7}$$

where n_d is the number of fruit observations available on date d and $d(i)$ denotes the assessment date of fruit i. The value n_d was subsequently used as a regression weight, so that dates based on more fruit observations received greater weight while the date remained the independent unit of analysis.

For the environmental predictors, one 14-day look-back curve was constructed for each unique assessment date. Each complete curve contained 336 hourly observations. Functional principal component analysis was fitted to the development-season date-level curves, and the first functional principal component score from the VPD curves was used as the environmental predictor. The environmental-only correction model was fitted using weighted linear regression:

$$\bar{r}_d^{(1)} = \gamma_0 + \gamma_1 z_{d1}^{\mathrm{VPD}} + \eta_d, \tag{4.8}$$

where z_{d1}^{VPD} is the first VPD functional principal component score for assessment date d. The regression was weighted by n_d.

For the complete $g(\cdot)$ correction model, longitudinal non-destructive NIR spectra were also summarised over the 14-day look-back period ending at each destructive assessment date. The SNV-transformed spectra within each window were averaged to obtain one representative date-level spectrum. Functional principal component analysis was then fitted to the date-level spectra from the 2021–2022 development season, and the first non-destructive NIR functional principal component score was retained.

The complete correction model used the first VPD score and the first non-destructive NIR score as predictors. Because the number of development-season assessment dates was small and the two predictors could be correlated, weighted ridge regression was used. The fitted model can be represented as

$$\bar{r}_d^{(1)} = \delta_0 + \delta_1 z_{d1}^{\mathrm{VPD}} + \delta_2 z_{d1}^{\mathrm{NIR}} + \eta_d, \tag{4.9}$$

where z_{d1}^{NIR} is the first functional principal component score from the longitudinal non-destructive

NIR spectra. The ridge model was fitted using `glmnet` with $\alpha = 0$, using n_d as the observation weight. The regularisation parameter λ was selected by leave-one-assessment-date-out cross-validation.

The environmental and non-destructive NIR FPCA mean functions and eigenfunctions were estimated using the 2021–2022 development-season data only. The fitted bases were subsequently retained when projecting the within-season prediction date and the 2024–2025 data. Thus, the dimension-reduction structure was not re-estimated using the evaluation observations.

The predicted date-level correction was assigned to all fruit measured on the corresponding assessment date. The environmental-only and complete corrected predictions were therefore calculated as

$$\widehat{y}_i^{(hg,\mathrm{env})}(t_i) = \widehat{y}_i^{(h)}(t_i) + \widehat{g}_{\mathrm{env}}\{d(i)\}, \tag{4.10}$$

$$\widehat{y}_i^{(hg,\mathrm{complete})}(t_i) = \widehat{y}_i^{(h)}(t_i) + \widehat{g}_{\mathrm{complete}}\{d(i)\}. \tag{4.11}$$

Consequently, four model forms were compared:

1. the two-phase baseline model;
2. the baseline plus the common fruit-level $h(\cdot)$ correction model;
3. the baseline plus $h(\cdot)$ and the environmental $g(\cdot)$ correction model; and
4. the complete model incorporating both environmental and longitudinal non-destructive NIR information in $g(\cdot)$.

For each comparison, all four models were evaluated using the same set of observations with complete inputs. Fruit-level performance was assessed using R^2, residual predictive deviation (RPD), root mean square error (RMSE), mean absolute error (MAE), and mean square error (MSE). Because the number of fruit observations differed substantially among assessment dates, date-balanced RMSE and MAE were also calculated by first obtaining the error within each date and then giving each date equal weight. In addition, errors were calculated after averaging the observed and predicted firmness values within each assessment date to assess how well the

models represented the population-level mean firmness trajectory. The observations at DAFB = 105 were evaluated separately as the within-season prediction set.

4.3 Cross-Season Prediction Performance Evaluation Using Data Collected from the 2024–2025 Season

After the mechanistic-informed statistical model was developed using the 2021–2022 season, the fitted modelling framework was transferred to the 2024–2025 season for a holdout cross-season prediction performance evaluation. The purpose of this evaluation was to examine whether the two-phase firmness-decline structure and the statistical correction components developed from one growing season could provide useful predictions in a different seasonal population.

The 2024–2025 evaluation used a season-specific initial-firmness calibration. Therefore, it should be interpreted as a cross-season evaluation following limited first-date calibration rather than as a completely uncalibrated external prediction. The first assessment date in the 2024–2025 season was used only to estimate the season-specific initial firmness, while the subsequent assessment dates were retained for prediction performance evaluation.

After applying the data-cleaning and pre-processing procedures described in Chapter 4, the 2024–2025 destructive fruit quality measurements, destructive NIR spectra, longitudinal non-destructive NIR spectra, and environmental data were prepared using the same variable definitions and processing procedures as the 2021–2022 development data. In particular, the destructive and longitudinal non-destructive NIR spectra were transformed using standard normal variate (SNV), and only wavelengths available in both seasons were used in the transferred model.

Season-Specific Calibration of the Baseline

The two-phase baseline structure, including the selected knot and slopes, was retained from the 2021–2022 development model. However, the overall initial firmness level differed between the two seasons. The initial firmness for the 2024–2025 season was therefore calibrated using the first available destructive firmness assessment on 19 December 2024 at DAFB = 65.

The transferred baseline can be written as

$$\widehat{\bar{y}}_{\text{new}}(t; F_0) = F_0 + \widehat{q}(t), \tag{4.12}$$

where $\widehat{q}(t)$ represents the fitted deviation from the initial firmness implied by the two-phase baseline developed using the 2021–2022 season. Let t_c denote the calibration DAFB and let $\mathcal{C}$ denote the set of fruit measured on the calibration date. The season-specific initial firmness was calculated as

$$\widehat{F}_{0,2024\text{–}2025} = \frac{1}{n_c} \sum_{i \in \mathcal{C}} \left[y_i(t_c) - \widehat{q}(t_c) \right], \tag{4.13}$$

where n_c is the number of destructive firmness observations on the calibration date. Based on the ten available firmness observations at $\texttt{DAFB} = 65$, this procedure gave

$$\widehat{F}_{0,2024\text{–}2025} = 13.682 \ \textit{kgf}. \tag{4.14}$$

This calibrated value is the baseline intercept required for the mean prediction at the first assessment date to agree with the corresponding observed mean firmness. It is therefore not simply equal to the observed mean firmness at that date.

To prevent the same observations from being used for both calibration and evaluation, all observations from 19 December 2024 were excluded from the reported performance metrics. The primary cross-season evaluation was consequently conducted using the two later assessment dates, $\texttt{DAFB} = 100$ and $\texttt{DAFB} = 118$, containing 10 and 15 fruit observations, respectively.

A strict-transfer sensitivity analysis was also conducted by retaining the original development-season value $F_0 = 12.1$ without any 2024–2025 calibration. In addition, a fixed-intercept scenario with $F_0 = 13.5$ was calculated to assess the sensitivity of the results to the treatment of the season-specific initial firmness. These alternative scenarios were used for comparison, whereas the first-date-calibrated evaluation was treated as the primary cross-season analysis.

Transfer of the Fruit-Level $h(\cdot)$ Correction

The fruit-level $h(\cdot)$ model used the common predictor structure described in Section 4.2. Its candidate predictors were fruit weight, SSC, and the SNV-transformed destructive NIR wave-

lengths available in both seasons. Fruit diameter, measured DMC, and estimated DMC were not included because these variables were not consistently available in the 2024–2025 season.

The elastic net model was fitted using only the 2021–2022 development data. Its coefficients and selected regularisation parameter were then retained and applied directly to the 2024–2025 predictor matrix. The $h(\cdot)$ model was not refitted using the 2024–2025 firmness outcomes. For fruit i in the new season, the prediction after the fruit-level correction was

$$\widehat{y}_{i,\text{new}}^{(h)}(t_i) = \widehat{\bar{y}}_{\text{new}}(t_i) + \widehat{h}(\mathbf{x}_{i,\text{new}}), \tag{4.15}$$

where $\mathbf{x}_{i,\text{new}}$ contains the common fruit-level and destructive NIR predictors.

Transfer of the Date-Level $g(\cdot)$ Corrections

The environmental variables and longitudinal non-destructive NIR spectra were incorporated at the measurement-date level. All pre-processing quantities required for their transfer were estimated from the 2021–2022 development data only.

For the environmental correction, one complete 14-day VPD look-back curve was constructed for each 2024–2025 assessment date. Each curve contained 336 hourly observations. The VPD records were transformed using the centring and scaling parameters estimated from the 2021–2022 environmental data. The resulting 2024–2025 VPD curves were then projected onto the environmental FPCA mean function and eigenfunction fitted using the development-season dates. The first VPD functional principal component score was entered into the previously fitted weighted linear $g(\cdot)$ correction model.

For the complete $g(\cdot)$ correction, longitudinal non-destructive NIR spectra collected during the 14-day period ending at each destructive assessment date were used. The spectra were processed using SNV and averaged within each assessment-date window to obtain one representative spectrum per date. These date-level spectra were then projected onto the non-destructive NIR FPCA basis estimated from the 2021–2022 development-season data. The first non-destructive NIR functional principal component score was combined with the first VPD score and entered into the fitted weighted ridge correction model.

Neither the environmental nor the non-destructive NIR FPCA basis was re-estimated using the 2024–2025 observations. Similarly, the coefficients of the environmental and complete $g(\cdot)$

models were retained from the 2021–2022 development analysis. The predicted date-level corrections were assigned to all fruit measured on the corresponding assessment date.

The resulting cross-season predictions were therefore

$$\widehat{y}_{i,\text{new}}^{(hg,\text{env})}(t_i) = \widehat{y}_{i,\text{new}}^{(h)}(t_i) + \widehat{g}_{\text{env}}\{d(i)\}, \tag{4.16}$$

$$\widehat{y}_{i,\text{new}}^{(hg,\text{complete})}(t_i) = \widehat{y}_{i,\text{new}}^{(h)}(t_i) + \widehat{g}_{\text{complete}}\{d(i)\}, \tag{4.17}$$

where $d(i)$ denotes the assessment date associated with fruit i.

Performance Evaluation

The primary cross-season performance evaluation used the 25 fruit observations from `DAFB` $= 100$ and `DAFB` $= 118$. The calibration observations at `DAFB` $= 65$ were not included in the performance calculations. The same complete set of fruit observations was used to compare the following four model forms:

1. the calibrated two-phase baseline model;
2. the baseline plus the common fruit-level $h(\cdot)$ correction model;
3. the baseline plus $h(\cdot)$ and the environmental $g(\cdot)$ correction model; and
4. the complete model incorporating both environmental and longitudinal non-destructive NIR information in $g(\cdot)$.

Fruit-level prediction performance was assessed using R^2, residual predictive deviation (RPD), root mean square error (RMSE), mean absolute error (MAE), mean square error (MSE), and prediction bias. Bias was defined as predicted minus observed firmness.

Because the two evaluation dates contained different numbers of fruit, date-balanced RMSE and MAE were also calculated by first evaluating the prediction errors separately within each date and then giving the two dates equal weight. In addition, the observed and predicted firmness values were averaged within each assessment date, and date-mean RMSE and MAE were calculated to examine how well the model represented the population-level firmness trajectory.

Since only two dates remained after calibration, the date-mean results were treated as descriptive measures rather than as a robust independent validation of trajectory-level performance.

Chapter 5

Results and Comparisons

In this Chapter, we present the results and comparisons of the developed mechanistic-informed statistical modelling framework for pre-harvest apple fruit firmness prediction. As described in Chapter 4, the 2021–2022 season dataset was used to develop and fit the fused model, including the baseline model, the fruit-level correction model $h(\cdot)$, and the date-level correction model $g(\cdot)$. Then, the fitted modelling framework was applied to the 2024–2025 season as holdout cross-season prediction performance evaluation with season-specific recalibration to evaluate whether the proposed fused model could be used for a new season and to assess its predictive performance on new data.

The results are presented by comparing different model forms, including the baseline model, the baseline plus $h(\cdot)$ correction model, the baseline plus $h(\cdot)$ and environmental $g(\cdot)$ correction model, and the complete model including both environmental and non-destructive Near-Infrared (NIR) information in $g(\cdot)$. Finally, the comparisons use several performance metrics, including R^2, `RMSE`, `MAE`, `MSE`, and RPD to evaluate and compare the predictive performance of each model component and to examine whether the fused model improved apple firmness prediction while retaining the biologically meaningful declining trend over days after full bloom (DAFB).

5.1 Model Development Results and Comparisons

Following the model implementation steps described in Chapter 4, the 2021–2022 season dataset was used to develop and evaluate the proposed mechanistic-informed statistical model. In the baseline model, the initial apple firmness value was set to $F_0 = 12.1$, and the selected knot was

located at DAFB = 70. After merging the fruit quality measurements with the destructive NIR spectral data and applying the required cleaning procedures, observations with DAFB < 100 were used for model development, while observations at DAFB = 105 were retained as a within-season prediction set.

Based on the cleaned fruit-level modelling dataset, the baseline model and the common fruit-level correction model $h(\cdot)$ were first fitted. The environmental variables were then aligned with the fruit observations using 14-day look-back windows and included in the date-level $g(\cdot)$ correction model. Finally, the longitudinal non-destructive NIR spectra were summarised over the corresponding 14-day windows at the measurement-date level and incorporated into the complete $g(\cdot)$ correction model. To permit a direct comparison, all four models were evaluated using the same observations for which the complete set of required inputs was available. Table 5.1 summarises the measurement dates and sample sizes retained at each stage of model development.

Table 5.1: Measurement dates and sample sizes retained at each stage of model development for the 2021–2022 season.

Measurement date	DAFB	Before merging	Common $h(\cdot)$ dataset	Environmental $g(\cdot)$ dataset	Complete $g(\cdot)$ dataset
2021-12-20	63	10	10	10	10
2021-12-21	64	150	145	145	145
2022-01-05	79	10	0	0	0
2022-01-14	88	10	9	0	0
2022-01-21	95	10	10	10	10
2022-01-31	105	10	9	9	9
2022-02-07	113	10	0	0	0

Table 5.2 presents the candidate variables, fitting methods, and retained inputs for each model component. To ensure that the same fruit-level correction could subsequently be applied to the 2024–2025 season, the common $h(\cdot)$ model considered only fruit weight, SSC, and the SNV-transformed destructive NIR wavelengths available in both seasons. The elastic net model retained an intercept of approximately -0.043 *kgf*, while the fitted SSC coefficient was approximately 4.6×10^{-17} and no fruit-weight or NIR wavelength variable was retained. Therefore, the common $h(\cdot)$ component was effectively an intercept adjustment rather than a meaningful fruit-level predictor. The environmental $g(\cdot)$ model used the first FPCA score of the 14-day VPD curve, while the complete $g(\cdot)$ model additionally used the first FPCA score from the longitudinal non-destructive NIR spectra.

Table 5.2: Candidate variables and retained inputs used in model development.

Model component	**Candidate variables**	**Selection or fitting method**	**Retained inputs**
Baseline model	`DAFB`	Two-phase one-knot spline	`DAFB`
Common $h(\cdot)$ correction model	Fruit weight, SSC, and common SNV-transformed destructive NIR wavelengths	Elastic net regression with leave-one-assessment-date-out cross-validation	No meaningful fruit-level predictor was retained; the fitted correction was effectively intercept-only
Environmental $g(\cdot)$ correction model	First FPCA score of the 14-day VPD curve	Date-level weighted linear regression	First FPCA score of VPD
Complete $g(\cdot)$ correction model	First FPCA score of VPD and the first FPCA score from the longitudinal non-destructive NIR spectra	Date-level weighted ridge regression with leave-one-assessment-date-out cross-validation	First FPCA score of VPD and the first FPCA score from the longitudinal non-destructive NIR spectra

Figure 5.1 compares the observed sample mean firmness curve with the trajectories from the baseline model, the baseline plus the common $h(\cdot)$ correction model, the baseline plus $h(\cdot)$ and environmental $g(\cdot)$ correction model, and the complete model incorporating both environmental and longitudinal non-destructive NIR information in $g(\cdot)$. All four model trajectories show a general decline in mean apple firmness as `DAFB` increases. The corrected curves remain close to the baseline curve, which is consistent with the small numerical contribution of the correction terms reported below. The observation at `DAFB` $= 88$ is shown in the figure to describe the available destructive measurements, but it was not included in the common four-model numerical comparison because complete date-level inputs were unavailable at this date.

The observed destructive measurements show a local "valley", in which the sample mean firmness decreases at `DAFB` $= 95$ and then increases slightly at `DAFB` $= 105$. This isolated increase is not interpreted as evidence of a general physiological re-firming pattern. Because firmness was measured destructively, different fruit were sampled at each assessment date. The connected sample means therefore do not represent repeated measurements of the same individual fruit. In addition, only ten fruit observations were available at `DAFB` $= 95$, and nine were available at `DAFB` $= 105$. The date-specific means are consequently sensitive to between-fruit heterogeneity, sampling variation, and measurement error. The approximate confidence intervals for these dates also overlap substantially, suggesting that the apparent increase may be within the sampling uncertainty. Furthermore, `DAFB` $= 105$ was treated as a held-out prediction point and was not used to fit the baseline model.

The fitted curves are intended to estimate the underlying population-level firmness trajectory rather than reproduce every local fluctuation in the destructive sample means. The models combine a biologically motivated monotone baseline with statistical corrections based on the available fruit-level, spectral, and environmental information. The correction terms can adjust the baseline when stable information is present in the predictors, but they do not reproduce the isolated valley because it was not supported by a stable predictor relationship. Therefore, the generally decreasing model trajectories are considered more consistent with the expected pre-harvest softening process than interpreting the point-to-point increase as a true reversal of apple fruit softening (Volz et al., 2003).

The vertical error bars in Figure 5.1 represent approximate 95% confidence intervals around the observed mean firmness or mean predicted firmness at each assessment date. They illustrate

the uncertainty associated with the estimated mean curves and should be considered when interpreting differences between neighbouring assessment dates.

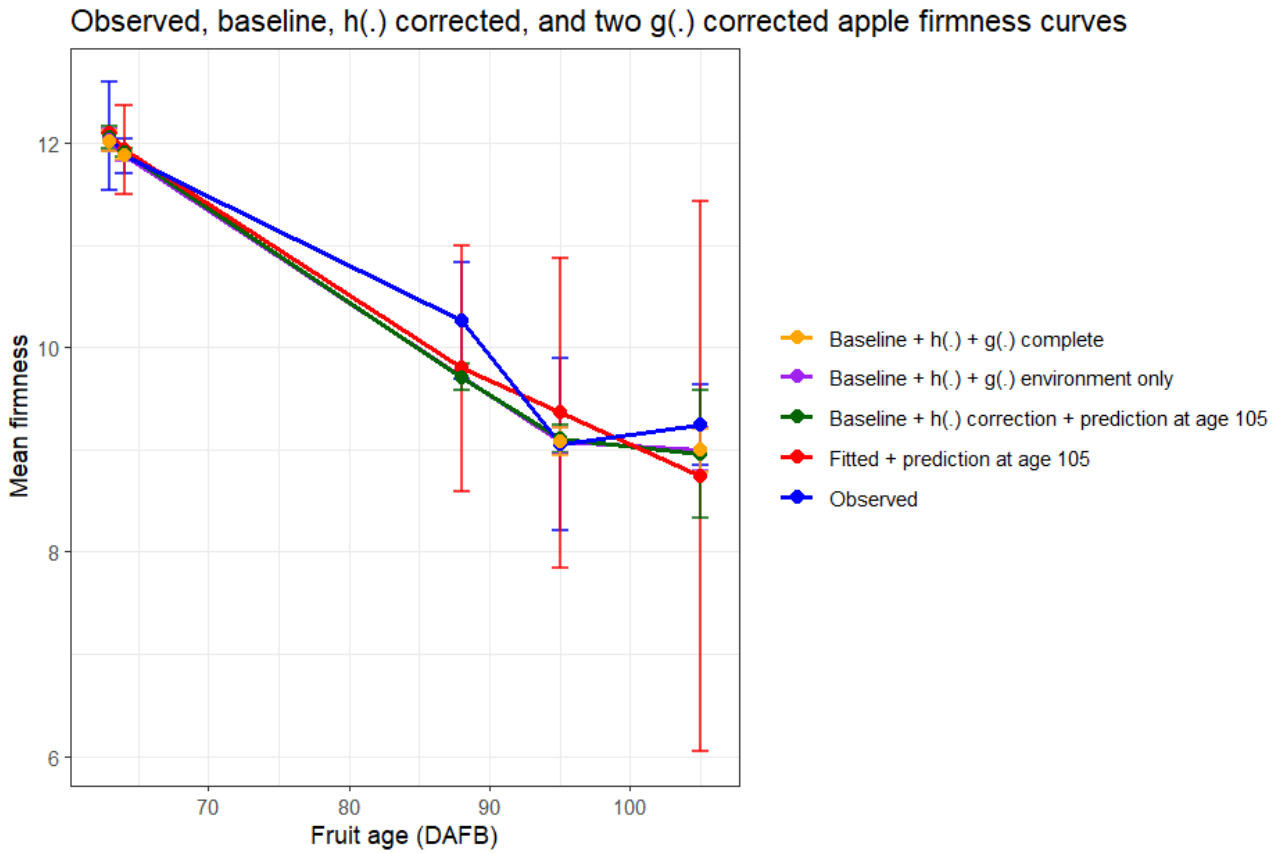


Figure 5.1: Model development results for the 2021–2022 season, including the observed sample mean firmness curve and the trajectories from the baseline model, the baseline plus the common $h(\cdot)$ correction model, the baseline plus $h(\cdot)$ and environmental $g(\cdot)$ correction model, and the complete $g(\cdot)$ correction model. The observations at `DAFB` = 105 were retained as the within-season prediction set. The vertical error bars represent approximate 95% confidence intervals around the observed mean firmness or mean predicted firmness at each assessment date. Apple firmness is expressed in *kgf*.

To compare the performance of the different model components, several metrics were calculated, including R^2, residual predictive deviation (RPD), root mean square error (RMSE), mean absolute error (MAE), and mean square error (MSE). The development metrics were calculated from the common set of 165 fruit observations at three assessment dates. The within-season holdout metrics were calculated from nine fruit observations at `DAFB` = 105. Table 5.3 summarises the results.

For the development data, the baseline model achieved $R^2 = 0.282$, RPD = 1.184, and RMSE = 1.071 *kgf*. Adding the common $h(\cdot)$ correction increased R^2 by only 0.003 and reduced RMSE by approximately 0.002 *kgf*. The environmental $g(\cdot)$ model gave the numerically highest

Table 5.3: Fruit-level model performance for the 2021–2022 development season and within-season holdout.

Model	**Data**	n	R^2	**RPD**	**RMSE (kgf)**	**MAE (kgf)**	**MSE (kgf^2)**
Baseline	Development	165	0.282	1.184	1.071	0.852	1.146
Baseline + common $h(\cdot)$	Development	165	0.285	1.186	1.069	0.848	1.142
Baseline + common $h(\cdot)$ + $g(\cdot)$ environment	Development	165	0.286	1.187	1.068	0.844	1.140
Baseline + common $h(\cdot)$ + $g(\cdot)$ complete	Development	165	0.285	1.186	1.068	0.847	1.141
Baseline	Holdout	9	−0.773	0.797	0.762	0.668	0.581
Baseline + common $h(\cdot)$	Holdout	9	−0.910	0.768	0.791	0.701	0.626
Baseline + common $h(\cdot)$ + $g(\cdot)$ environment	Holdout	9	−1.278	0.703	0.864	0.780	0.746
Baseline + common $h(\cdot)$ + $g(\cdot)$ complete	Holdout	9	−1.028	0.745	0.815	0.728	0.664

Note: Apple firmness was measured in kilogram-force (*kgf*). Therefore, RMSE and MAE are expressed in *kgf*, whereas MSE is expressed in *kgf*^{2}. The R^2 and RPD values are dimensionless.

development R^2 of 0.286 and the lowest RMSE of 1.068 *kgf*, but the RMSE reduction relative to the baseline was only approximately 0.3%. The complete $g(\cdot)$ model did not improve on the environmental model. These small differences do not provide evidence of a substantive improvement from the correction components in the development data.

The date-balanced analysis led to the same conclusion. When each of the three development dates was given equal influence, the RMSE values were 1.090, 1.086, 1.080, and 1.083 *kgf* for the baseline, common $h(\cdot)$, environmental $g(\cdot)$, and complete $g(\cdot)$ models, respectively. Therefore, the highly unequal sample sizes across dates are an important limitation, but they do not materially change the model ranking or the conclusion that the correction terms provided little additional fruit-level information.

Because apple firmness was measured in kilogram-force, the RMSE of approximately 1.07 *kgf* represents an error of approximately 9.1% relative to the mean observed development firmness of 11.71 *kgf*. This remains a substantial error for predicting the firmness of an individual fruit. The development RPD values were also approximately 1.18 for all four models, well below 2, which indicates limited utility for quantitative individual-fruit prediction.

The practical interpretation is more favourable when the objective is to describe the average firmness trajectory. After averaging observed and predicted firmness within each assessment date, the development RMSE values were 0.180, 0.152, 0.107, and 0.134 *kgf* for the baseline, common $h(\cdot)$, environmental $g(\cdot)$, and complete $g(\cdot)$ models, respectively. However, these are in-sample descriptive results based on only three development dates, and the baseline was itself estimated from date-level mean firmness. They therefore show that the fitted models reproduce the development-season mean trajectory, but should not be interpreted as independent evidence

of predictive accuracy.

For the within-season holdout at DAFB = 105, the baseline model had the lowest RMSE and MAE. Each correction component worsened the holdout performance. The negative holdout R^2 values should be interpreted cautiously because the holdout contained only nine fruit at a single assessment date and therefore contained no variation in DAFB. In this setting, R^2 assesses only whether the models explain differences among individual fruit relative to the holdout mean; it does not assess the firmness trajectory over time. The baseline prediction of the assessment-date mean had an absolute error of 0.503 *kgf*, approximately 5.4% of the observed holdout mean of 9.24 *kgf*. The common $h(\cdot)$, environmental $g(\cdot)$, and complete $g(\cdot)$ models produced larger date-mean errors of 0.546, 0.647, and 0.580 *kgf*, respectively.

The remaining unexplained fruit-level variation should not be attributed solely to poor data quality. Rather, it reflects a combination of genuine biological heterogeneity, unrecorded local growing conditions, sampling variability, and measurement error. Fruit collected from the same tree and on the same assessment date may still differ in canopy position, light exposure, water status, crop load, and maturity. These factors can produce substantial differences in firmness, but only some were represented by the predictors available in this study (Musacchi and Serra, 2018; Chagné et al., 2014; Lachapelle et al., 2013; Saei et al., 2011).

The limited contribution of the $g(\cdot)$ components is also related to the observational scale of their predictors. The environmental and longitudinal NIR summaries were constructed at the measurement-date level and were shared by all fruit assessed on the same date. These variables could potentially explain differences in mean firmness among dates, but they could not explain variation among individual fruit within a date. In addition, only three distinct development dates were available for fitting the date-level corrections. The negligible fruit-level improvement therefore does not imply that environmental conditions or longitudinal NIR information are biologically unimportant; rather, the current data structure provided limited information for estimating stable date-level effects.

Overall, the 2021–2022 results show that the biologically informed baseline captured the dominant population-level decline in firmness. The common $h(\cdot)$ correction was effectively an intercept adjustment, while the environmental and complete $g(\cdot)$ corrections provided only negligible improvement in the development data and did not improve the within-season holdout prediction. The framework therefore has greater potential for describing a population-level pre-

harvest firmness trajectory than for precise individual-fruit prediction under the current data conditions.

5.2 Prediction Performance Evaluation Results for the 2024–2025 Season

After the mechanistic-informed statistical model was developed using the 2021–2022 season, the fitted modelling framework was transferred to the held-out 2024–2025 season. The purpose of this cross-season evaluation was to examine whether the two-phase firmness-decline structure and the statistical correction components developed from one growing season could provide useful predictions in a different seasonal population.

As described in Section 4.3, the first 2024–2025 assessment date was used to calibrate the season-specific initial firmness. The two-phase baseline structure, including the knot and the two decline rates, was retained from the 2021–2022 development model. Calibration using the ten destructive firmness observations collected on 19 December 2024 at `DAFB` = 65 gave $F_0 = 13.682$ *kgf*. To separate calibration from performance evaluation, the observations from this first date were excluded from all primary performance metrics. The primary evaluation therefore used 25 fruit observations from the two later assessment dates, `DAFB` = 100 and `DAFB` = 118.

Table 5.4 summarises the sample sizes and analytical roles of the three 2024–2025 assessment dates. The calibration used all ten available firmness observations at `DAFB` = 65. Eight of these observations also had the complete common predictor set required for the four-model comparison shown in Figure 5.2, but none of the first-date observations were included in Table 5.6.

Table 5.4: Measurement dates, analytical roles, and sample sizes used in the 2024–2025 cross-season prediction performance evaluation.

Measurement date	DAFB	Role	Available firmness observations	Common-model complete cases
2024-12-19	65	Calibration	10	8
2025-01-23	100	Evaluation	10	10
2025-02-10	118	Evaluation	15	15

Table 5.5 presents the predictor structures transferred to the 2024–2025 season. The common $h(\cdot)$ model considered fruit weight, SSC, and the SNV-transformed destructive NIR wavelengths available in both seasons. However, no meaningful non-intercept predictor was retained. The

SSC coefficient was effectively zero, and the fitted $h(\cdot)$ component acted mainly as a small intercept adjustment of approximately -0.043 *kgf*. The environmental $g(\cdot)$ model used the first VPD FPCA score, while the complete $g(\cdot)$ model combined the first VPD score with the first longitudinal non-destructive NIR score. All coefficients and FPCA bases were estimated from the 2021–2022 development data and then retained for the 2024–2025 evaluation.

Figure 5.2 compares the observed sample mean firmness curve with the predictions from the calibrated baseline model, the baseline plus common $h(\cdot)$ correction model, the baseline plus $h(\cdot)$ and environmental $g(\cdot)$ correction model, and the complete model incorporating both environmental and longitudinal non-destructive NIR information in $g(\cdot)$. The calibration date at `DAFB` $= 65$ is retained in the figure to show the full seasonal trajectory, but it is not an independent prediction point and does not contribute to the performance metrics.

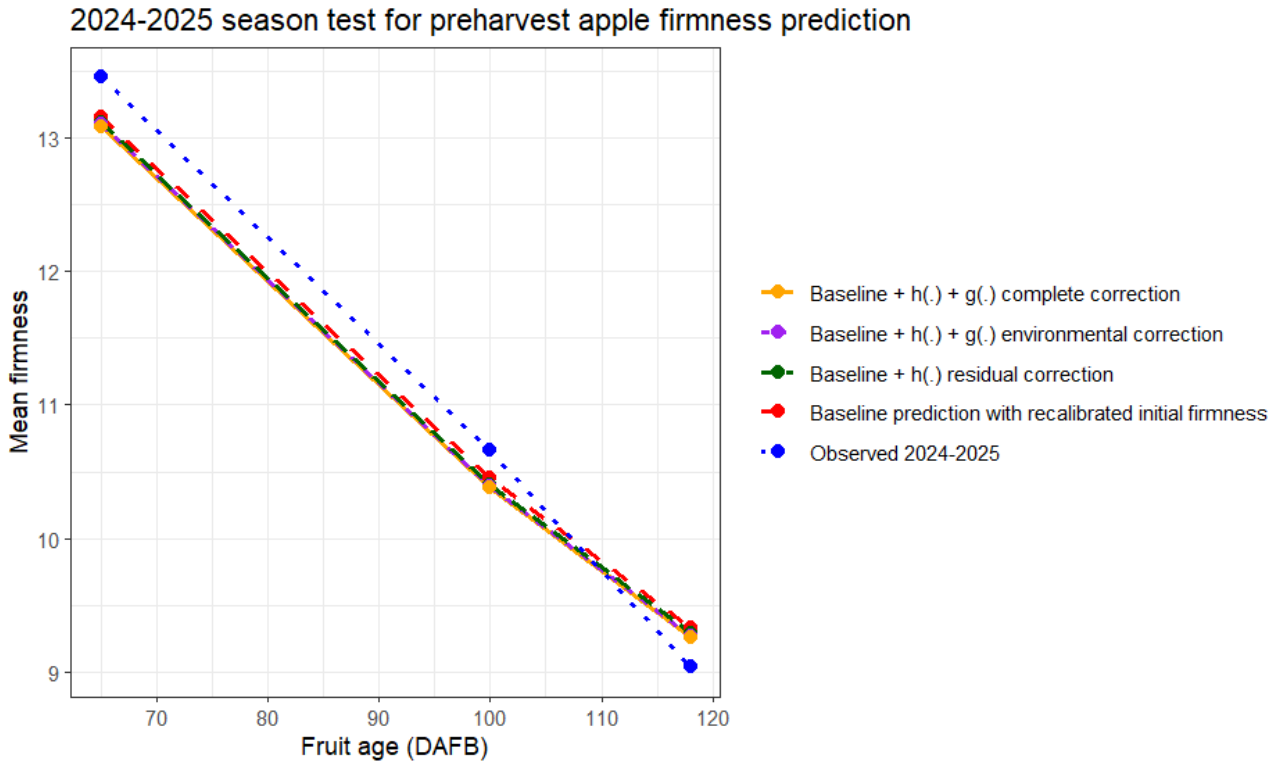


Figure 5.2: Cross-season prediction performance evaluation results for the 2024–2025 season, including the observed sample mean firmness curve and the predicted curves from the calibrated baseline model, the baseline plus common $h(\cdot)$ correction model, the baseline plus $h(\cdot)$ and environmental $g(\cdot)$ correction model, and the complete $g(\cdot)$ correction model. The observations at `DAFB` $= 65$ were used for initial-firmness calibration and are shown only to provide context; the reported performance metrics were calculated using `DAFB` $= 100$ and `DAFB` $= 118$. Apple firmness is expressed in *kgf*.

All four predicted curves followed the general decline in observed mean firmness across the

Table 5.5: Predictor structures and fitted components transferred to the 2024–2025 cross-season prediction performance evaluation.

Model component	**Candidate or transferred predictors**	**Development and transfer method**	**Fitted component used in evaluation**
Baseline model	`DAFB`	Two-phase one-knot spline fitted using the 2021–2022 season, with first-date calibration of F_0 for 2024–2025	Transferred knot and decline rates with $F_0 = 13.682\ kgf$
Common $h(\cdot)$ correction model	Fruit weight, SSC, and common SNV-transformed destructive NIR wavelengths	Elastic net regression fitted using 2021–2022 development data and transferred without refitting	No meaningful non-intercept predictor retained; effectively an intercept-only adjustment
Environmental $g(\cdot)$ correction model	First FPCA score of the 14-day VPD look-back curve	Development-season FPCA followed by weighted linear regression; basis and coefficients retained for transfer	First FPCA score of VPD
Complete $g(\cdot)$ correction model	First FPCA score of VPD and first FPCA score from the longitudinal non-destructive NIR spectra	Development-season FPCA followed by weighted ridge regression; bases and coefficients retained for transfer	First FPCA score of VPD and first FPCA score from the longitudinal non-destructive NIR spectra

pre-harvest period. This indicates that the two-phase structure transferred from the 2021–2022 season, after calibration of its initial level, captured the principal direction of the 2024–2025 population-level softening trajectory. The corrected curves remained close to the baseline curve, indicating that the statistical correction components made relatively small adjustments.

Table 5.6 presents the fruit-level performance results for the 25 observations at the two post-calibration evaluation dates. Bias was defined as predicted minus observed firmness. The environmental $g(\cdot)$ model achieved the highest R^2 and RPD and the lowest RMSE and absolute bias. The complete $g(\cdot)$ model had a marginally lower MAE, but its R^2, RPD, and RMSE were slightly worse than those of the environmental model.

Table 5.6: Fruit-level model performance for the primary 2024–2025 cross-season evaluation after first-date calibration. The calibration date was excluded, leaving 25 fruit observations across two assessment dates.

Model	R^2	**RPD**	**RMSE (kgf)**	**MAE (kgf)**	**MSE (kgf^2)**	**Bias (kgf)**
Baseline	0.330	1.247	1.009	0.728	1.018	0.272
Baseline + common $h(\cdot)$	0.344	1.260	0.998	0.713	0.996	0.229
Baseline + common $h(\cdot)$ + $g(\cdot)$ environment	0.369	1.285	0.979	0.707	0.959	0.092
Baseline + common $h(\cdot)$ + $g(\cdot)$ complete	0.354	1.270	0.990	0.706	0.980	0.192

Note: Bias is defined as predicted minus observed firmness. Positive values indicate average overprediction. The R^2 and RPD values are dimensionless.

Relative to the baseline model, the common $h(\cdot)$ correction reduced RMSE from 1.009 to 0.998 *kgf*, a reduction of approximately 1.1%. Adding the environmental $g(\cdot)$ correction reduced RMSE to 0.979 *kgf*, an improvement of approximately 2.9% relative to the baseline, and reduced average bias from 0.272 to 0.092 *kgf*. The complete $g(\cdot)$ model reduced RMSE by approximately 1.9% relative to the baseline. Therefore, the environmental correction provided the largest numerical improvement, although the differences among the four models remained small.

The same conclusion was obtained when the two assessment dates were given equal weight. Date-balanced RMSE decreased from 1.009 *kgf* for the baseline model to 0.988 *kgf* for the environmental correction model, a reduction of approximately 2.0%. Thus, the small improvement was not caused solely by the difference between the sample sizes of 10 and 15 fruit at the two evaluation dates.

Despite these numerical improvements, the fruit-level predictive performance remained limited. The observed mean firmness in the evaluation population was 9.688 *kgf*, so RMSE values of approximately 0.98–1.01 *kgf* corresponded to errors of about 10% of the average firmness. In

addition, all RPD values were between 1.247 and 1.285, remaining well below 2. These results indicate that none of the models provided sufficiently accurate individual-fruit predictions for operational use. The environmental correction should therefore be interpreted as a small improvement within an overall model whose individual-fruit prediction ability remains modest.

Sensitivity to Initial-Firmness Calibration

The treatment of the season-specific initial firmness had a substantial effect on the cross-season results. Table 5.7 compares the baseline model under three initial-firmness scenarios. In the strict transfer scenario, the development-season value $F_0 = 12.1$ was retained without calibration. This produced an RMSE of 1.702 *kgf* and a bias of -1.404 *kgf*, indicating systematic underprediction in the 2024–2025 season.

A fixed-value sensitivity scenario using $F_0 = 13.5$ and all three assessment dates produced a much higher R^2 of 0.765, an RPD of 2.096, and an RMSE of 0.963 *kgf*. However, this value of F_0 increased the baseline level by 1.4 *kgf*, almost exactly offsetting the -1.404 *kgf* bias observed under strict transfer. It therefore removed nearly all of the average seasonal level difference. In addition, this scenario included all three widely separated assessment dates, which produced a larger observed firmness variance and consequently a higher R^2 and RPD.

The primary first-date-calibrated evaluation avoids using the calibration observations in the performance calculation. After the first date was excluded, the observed firmness standard deviation was 1.258 *kgf*, which was almost identical to the development population standard deviation of 1.267 *kgf*. Under these more comparable response distributions, the development baseline ($R^2 = 0.282$, RMSE = 1.071 *kgf*) and the calibrated cross-season baseline ($R^2 = 0.330$, RMSE = 1.009 *kgf*) showed broadly similar performance. Thus, the cross-season results no longer suggest that prediction in the held-out season was substantially better than model development.

The sensitivity analysis shows that the absolute firmness level was not fully transferable between seasons. The two-phase decline structure retained useful information, but a season-specific estimate of the initial level was required. This finding is biologically plausible because the initial firmness of a seasonal population may vary with growing conditions, maturity status, crop load, water status, and other orchard-level factors. The calibrated framework therefore separates two components: a transferable developmental decline shape and a season-specific

Table 5.7: Sensitivity of the 2024–2025 baseline performance to the treatment of the season-specific initial firmness.

Scenario	F_0	Dates evaluated	R^2	RPD	RMSE (kgf)	MAE (kgf)	Bias (kgf)
Strict transfer without calibration	12.100	3	0.266	1.185	1.702	1.503	−1.404
Fixed-value sensitivity scenario	13.500	3	0.765	2.096	0.963	0.709	−0.004
First-date calibration; calibration date excluded	13.682	2	0.330	1.247	1.009	0.728	0.272

Note: The scenarios use different evaluation sets. The fixed-value scenario includes all three dates, whereas the primary first-date-calibrated scenario excludes the calibration date. Therefore, the numerical values should be interpreted as a sensitivity analysis rather than as a direct ranking.

starting level.

Population-Level Mean Firmness

Because the baseline component was designed to represent a population-level firmness trajectory, performance was also examined after averaging observed and predicted firmness within each assessment date. Table 5.8 presents the date-mean errors for the two post-calibration evaluation dates. Date-mean R^2 and RPD are not reported because only two dates were available and these statistics would be highly unstable.

Table 5.8: Prediction errors after averaging observed and predicted firmness within each of the two post-calibration assessment dates.

Model	Date-mean RMSE (kgf)	Date-mean MAE (kgf)	Date-mean bias (kgf)
Baseline	0.338	0.256	0.221
Baseline + common $h(\cdot)$	0.312	0.256	0.178
Baseline + common $h(\cdot)$ + $g(\cdot)$ environment	0.272	0.269	0.038
Baseline + common $h(\cdot)$ + $g(\cdot)$ complete	0.292	0.255	0.141

The date-mean errors were much smaller than the individual-fruit errors. The environmental correction model achieved a date-mean RMSE of 0.272 *kgf*, corresponding to approximately 2.8% of the mean observed firmness across the two dates. This supports the interpretation that the framework has greater potential for monitoring average firmness development than for predicting the firmness of a specific fruit. However, these results are based on only two post-calibration dates and should be regarded as descriptive evidence rather than a robust validation of date-level prediction.

Biological and Modelling Interpretation

Biologically, the broad decline captured by the baseline model is consistent with the expected softening of apple fruit during pre-harvest development. Fruit softening is associated with cell expansion, modification and solubilisation of cell-wall components, weakening of cell-to-cell adhesion, changes in cellular turgor, and ethylene-regulated ripening (Volz et al., 2003; Johnston et al., 2002a; Ireland et al., 2014). The cross-season results therefore suggest that developmental age, represented by DAFB, captured the dominant population-level signal, while the season-specific initial level required recalibration.

The common $h(\cdot)$ correction provided little additional information because no meaningful fruit-level predictor was retained. This does not imply that SSC, fruit weight, or destructive NIR spectra are biologically unrelated to firmness. Rather, their relationships with residual firmness were not sufficiently stable under the common predictor structure available in both seasons. Important fruit-level variables, including DMC and fruit diameter, were not consistently available in the evaluation season, and other sources of fruit-to-fruit variation, such as canopy position, light exposure, crop load, and water status, were not fully represented.

The environmental $g(\cdot)$ correction produced the best fruit-level RMSE and substantially reduced average bias, suggesting that the recent VPD trajectory contained some information about differences in mean firmness between the two evaluation dates. However, the improvement was small and was estimated from only a few development and evaluation dates. The result therefore provides preliminary support for including environmental information but does not establish a stable cross-season environmental effect.

Adding longitudinal non-destructive NIR information did not improve performance beyond the environmental correction. One possible reason is that the two later 2024–2025 non-destructive NIR FPCA scores fell outside the range represented by the development-season dates, particularly at the final assessment. The complete model was therefore partly extrapolating along the NIR component. In addition, the longitudinal spectra were summarised at the date level and were not directly paired with destructive firmness measurements from the same fruit. These features limited the model's ability to learn a stable individual-level relationship between longitudinal spectral change and firmness.

Comparison with Previous Studies

Direct numerical comparison with previous firmness-prediction studies requires caution because published studies differ in cultivar, developmental stage, reference measurement, response range, sample size, sensor technology, and validation design. Nevertheless, previous results offer useful context for the comparatively modest individual-fruit performance obtained in this study.

For example, Ma et al. (2021) developed a specialised Vis–NIR spatially resolved spectroscopy system for post-harvest apples and reported prediction R^2 values of 0.92 for SSC and 0.87 for firmness, with corresponding RMSEP values of 0.35% and 0.71 N. Similarly, Marečková et al. (2022) developed cultivar-specific NIR firmness calibrations using repeated large-scale sampling of stored commercial apples. Their calibration models achieved R^2 values above 0.91 and RPD values above 2.3, with the best model reaching $R^2 = 0.94$ and RPD = 2.6. These values are stronger than the fruit-level results obtained here.

However, these studies used closely matched spectroscopy-reference datasets, relatively large calibration populations, and controlled post-harvest measurement procedures. The present study addressed a different and more heterogeneous problem by aligning destructive measurements, laboratory spectra, longitudinal on-tree spectra, and hourly environmental data across pre-harvest development and between growing seasons. The numerical differences should therefore not be interpreted as a direct ranking of the models.

The current findings are also consistent with the widely reported difficulty of predicting firmness from NIR spectra. Firmness is a mechanical property influenced by cell-wall structure, cell adhesion, tissue architecture, and cellular turgor, whereas DMC and SSC are more directly associated with chemical constituents producing NIR absorption features. For example, McGlone and Kawano (1998) reported prediction R^2 values of 0.90 for both DMC and SSC in kiwifruit but only 0.66 for firmness. A firmness model fitted to a narrower fruit population reached $R^2 = 0.76$ but transferred poorly to independent fruit. In apples, Møller et al. (2013) also found stronger NIR performance for DMC and SSC than for firmness and related firmness more closely to the physical state and distribution of water within the fruit. The modest and unstable contribution of NIR predictors in the present study is therefore scientifically plausible, although it remains insufficient for practical individual-fruit prediction.

Previous studies also show that fusion can improve fruit-quality prediction when the informa-

tion sources are complementary and measured on well-matched samples. Mendoza et al. (2012) combined acoustic firmness, bioyield firmness, Vis–NIR spectroscopy, and spectral-scattering measurements from 6,535 apples. Their fused sensor models reduced firmness prediction errors by approximately 6.2%–24.9%, depending on cultivar, year, and sensor combination. Acoustic and bioyield sensors contributed information more directly related to mechanical fruit properties, while spectral measurements provided complementary optical and chemical information.

A mechanistic–data-driven fusion approach has also been developed for kiwifruit firmness. Xiao and Li (2022) combined a mechanistic kinetic softening model with functional regression using at-harvest NIR spectra from the same individual fruit. Their fused model outperformed conventional PLS and support vector machine benchmarks, although the improvement was not sufficient to support immediate industrial application. These studies show that fusion can improve prediction, but it does not automatically overcome indirect sensor–firmness relationships, limited sample size, incomplete matching, or population differences. These limitations and their implications for future data collection and model development are discussed further in Section 6.2.

Overall, the 2024–2025 results show that the two-phase developmental decline structure retained useful cross-season information, but the absolute firmness level required first-date calibration. After the calibration date was excluded, development and cross-season fruit-level errors were of a similar magnitude. The environmental correction produced a small reduction in RMSE and bias, whereas the common fruit-level and complete longitudinal NIR corrections provided little additional benefit. The framework was substantially more accurate for assessment-date mean firmness than for individual-fruit firmness, supporting its potential as a population-level monitoring approach rather than an operational individual-fruit prediction tool.

Chapter 6

General Conclusions and Future Perspectives

6.1 Conclusions

Pre-harvest apple fruit firmness prediction is important for harvest planning, fruit quality management, and post-harvest decision-making, but it is challenging because firmness reflects interacting biological, physical, environmental, and measurement-related processes. This thesis developed a mechanistic-informed statistical framework for modelling pre-harvest firmness in 'Royal Gala' apples. The framework combined a biologically guided baseline softening trajectory with statistical correction components based on fruit measurements, NIR spectral data, and environmental information. The principal conclusions are organised below according to the two research questions and the methodological objective introduced in Chapter 2.

Research Question One: What firmness response should be modelled?

The population-level mean firmness trajectory was selected as the principal baseline response. Destructive firmness measurements did not provide repeated firmness observations from the same fruit, and the longitudinal non-destructive NIR measurements were not directly matched with individual-fruit destructive firmness measurements. In addition, several variables that could explain differences among individual fruit were unavailable or incomplete. Under these conditions, a complete longitudinal firmness curve for each fruit was not sufficiently supported by the available data.

The baseline component therefore targeted the mean pre-harvest firmness curve, $\bar{y}(t)$, over days after full bloom (DAFB). The statistical $h(\cdot)$ and $g(\cdot)$ components were then used to examine whether the available fruit-level and date-level information could explain systematic departures from this mean. The results confirmed that the current data support estimation of the population-level trajectory more strongly than precise prediction of individual-fruit firmness.

Methodological Objective: How can the heterogeneous data be organised?

The methodological objective was addressed by developing a consistent processing and alignment framework for the different data sources. Destructive fruit-quality measurements and one-shot laboratory NIR spectra were cleaned and matched at the fruit and assessment-date levels. Standard normal variate transformation was applied consistently to the spectral data, and PLSR modelling of DMC was used to assess the general quality and usability of the destructive NIR spectra.

For the final cross-season firmness model, the fruit-level $h(\cdot)$ candidate set was restricted to variables available in both seasons: fruit weight, SSC, and common SNV-transformed destructive NIR wavelengths. Hourly environmental data were aligned with each firmness assessment using 14-day look-back windows. Because environmental conditions and longitudinal non-destructive NIR summaries were shared by fruit assessed on the same date, these inputs were represented at the measurement-date level. Functional principal component analysis was used to reduce the environmental curves and longitudinal NIR spectra to low-dimensional scores. The transformations and FPCA bases were estimated from the 2021–2022 development data and then retained for the 2024–2025 cross-season evaluation. This framework provided a practical way to combine fruit-level, spectral, and high-frequency environmental information without treating repeated date-level inputs as independent observations.

Research Question Two: How did the fused framework perform?

The biologically informed baseline was represented by a two-phase one-knot spline. For the 2021–2022 development season, the initial firmness was set to $F_0 = 12.1$ *kgf* and the selected knot was located at DAFB 70. The fitted baseline reproduced the broad monotone decline in

pre-harvest firmness, with a faster decline before the knot and a slower decline afterwards.

At the individual-fruit level, the statistical corrections provided little additional improvement during model development. The baseline model achieved $R^2 = 0.282$, RPD = 1.184, and RMSE = 1.071 *kgf*. The environmental correction gave the numerically best development result, but only increased R^2 to 0.286 and reduced RMSE to 1.068 *kgf*. The common $h(\cdot)$ model retained no meaningful non-intercept predictor and was effectively a small intercept adjustment. At the within-season holdout date, DAFB 105, the baseline model had the lowest RMSE of 0.762 *kgf*, while all correction models produced larger errors. Thus, the correction components did not demonstrate stable within-season predictive gains.

The 2024–2025 analysis also showed that the absolute firmness level was not directly transferable between seasons. Strict transfer of the 2021–2022 baseline produced an RMSE of 1.702 *kgf* and a bias of -1.404 *kgf*. The first 2024–2025 assessment date was therefore used to calibrate the season-specific initial firmness, giving $F_0 = 13.682$ *kgf*, and that calibration date was excluded from the primary performance evaluation. The remaining evaluation used 25 fruit observations at DAFB 100 and 118.

After first-date calibration, the baseline model achieved $R^2 = 0.330$, RPD = 1.247, and RMSE = 1.009 *kgf*. The environmental $g(\cdot)$ correction produced the best overall fruit-level result, with $R^2 = 0.369$, RPD = 1.285, and RMSE = 0.979 *kgf*. This represented an RMSE reduction of approximately 2.9% relative to the calibrated baseline. The complete model including longitudinal non-destructive NIR information did not improve on the environmental model. Because all fruit-level RPD values remained below 1.3 and RMSE remained close to 1 *kgf*, the resulting accuracy was insufficient for precise individual-fruit prediction.

The interpretation was more favourable at the population level. After averaging firmness within each of the two post-calibration assessment dates, the environmental model achieved a date-mean RMSE of 0.272 *kgf*, corresponding to approximately 2.8% of the mean observed firmness. This result suggests that the framework has greater potential for monitoring average firmness development than for predicting the firmness of a particular fruit. However, only two post-calibration dates were available, so this date-mean result should be treated as descriptive rather than as a robust validation of population-level prediction.

Overall, the study demonstrates a workable framework for combining a biologically interpretable softening trajectory with heterogeneous sensor-based data. The two-phase baseline

captured the dominant pre-harvest decline and its shape retained useful information across seasons, but the seasonal firmness level required calibration. The fruit-level and longitudinal NIR corrections did not provide stable additional value, while the environmental correction produced only a small cross-season improvement. The proposed framework should therefore be regarded as a proof of concept for population-level firmness monitoring rather than an operational tool for individual-fruit prediction.

6.2 Limitations and Future Work

The conclusions above should be interpreted in light of several limitations concerning seasonal coverage, sampling design, data matching, predictor availability, and practical validation.

Seasonal Coverage and Initial-Firmness Calibration

Only two seasons contained sufficiently complete combinations of fruit quality, NIR, and environmental data for the main analysis. The 2021–2022 season was used for model development, while the 2024–2025 season provided a single held-out cross-season evaluation. Moreover, the first 2024–2025 assessment date was required to calibrate the season-specific initial firmness, leaving only two dates for primary performance evaluation. The current results therefore cannot establish how consistently the baseline shape and correction components would transfer across a broader range of seasonal conditions.

Future studies should collect complete data over additional growing seasons and use repeated leave-one-season-out evaluation. The season-specific initial level could then be incorporated explicitly, for example through a hierarchical season effect or a predefined early-season calibration procedure. Such a design would distinguish more clearly between the transferable developmental decline and the season-specific starting firmness.

Sampling Balance and Predictor Consistency

The usable observations were unevenly distributed across assessment dates. In the 2021–2022 development data, 145 complete observations came from DAFB 64, whereas only approximately ten observations were available at the other retained dates. Although the date-balanced analysis

led to the same model ranking, more even sampling would provide stronger information about the complete firmness trajectory and reduce dependence on a single date.

Predictor availability also differed between seasons. DMC and fruit diameter, which may be associated with firmness, were not sufficiently complete in 2024–2025 and therefore could not be included in the common cross-season $h(\cdot)$ model. Under the remaining common predictor set, no meaningful fruit-level predictor was retained. Future data collection should use the same variables, instruments, and assessment schedule in every season and should aim for more complete matching among fruit-quality measurements, destructive spectra, longitudinal spectra, and environmental records.

Individual-Fruit Linkage and Biological Heterogeneity

The destructive design meant that different fruit were used at different firmness assessment dates. Consequently, the observed date means did not represent repeated firmness measurements from the same fruit. In addition, the longitudinal non-destructive NIR spectra were not directly linked with later destructive firmness measurements from the same labelled fruit. This limited the ability to determine whether within-fruit spectral changes were associated with subsequent firmness.

Where feasible, future experiments should repeatedly scan labelled fruit on the tree and then obtain a destructive firmness reference from the same fruit at a scheduled endpoint. More detailed fruit- and orchard-level information, including tree identity, canopy position, light exposure, crop load, irrigation, and indicators of fruit or tree water status, would also help explain fruit-to-fruit heterogeneity. With such data, hierarchical or mixed-effects models could separate variation among seasons, trees, assessment dates, and individual fruit.

Sensor Information and Biological Representation

Firmness is a mechanical property influenced by cell-wall structure, cell adhesion, tissue architecture, and cellular turgor. Its relationship with NIR spectra is therefore more indirect than the relationships of DMC or SSC with NIR absorption. In the present study, the later 2024–2025 longitudinal NIR scores also extended beyond the range represented by the development-season dates, so the complete model partly relied on extrapolation. Additional seasons and better

matched fruit-level reference data are needed before a stable longitudinal NIR correction can be established. Environmental conditions around individual fruit may differ from those recorded by a local weather station. In the present study, the environmental predictors were derived from a local weather station rather than sensors located within the apple-tree canopy. Therefore, the recorded air temperature, relative humidity, and derived VPD may not fully represent the conditions experienced by individual fruit because of local differences in shading, transpiration, and wind exposure. This difference may have limited the ability of the environmental $g(\cdot)$ correction to capture the effects of environmental conditions on fruit firmness. Future studies using environmental measurements collected within the canopy may help improve the environmental correction.

In addition, variation in optical path length was not explicitly considered in this study. Diffuse-reflectance NIR measurements can be influenced by light scattering and changes in effective optical path length. Although SNV pre-processing was applied to reduce the effects of spectral scattering, changes in optical path length during fruit development may not have been fully removed. More specifically, changes in fruit tissue structure during fruit development may affect the scattering properties of the fruit and therefore influence the measured NIR spectra. As a result, some of the spectral variation may reflect changes in optical path length rather than changes in chemical composition. Therefore, this may have limited the contribution of the NIR-based correction components to apple firmness prediction.

Future studies could combine NIR spectroscopy with complementary measurements more directly related to mechanical firmness, such as acoustic or bio-yield measurements. Direct observations of relevant physiological states, including fruit or tree water status, xylem water potential, turgor-related measures, or ripening indicators, could also strengthen the biological basis of the corrections. With more assessment dates, environmental effects could be represented using multiple or distributed look-back periods rather than a single fixed 14-day window. The mechanistic component could also be extended beyond a biologically constrained curve to include measured or modelled physiological states when sufficiently reliable data become available.

Practical Validation and Uncertainty

The fruit-level RMSE remained approximately 10% of average firmness, and RPD remained below 1.3, indicating that the current framework is not sufficiently accurate for operational individual-fruit prediction. The smaller date-mean errors are encouraging for population-level monitoring, but they were calculated from only two post-calibration dates and were not compared with a formally established industry decision tolerance.

Future evaluation should define the intended decision context before judging model usefulness. For example, the required accuracy may differ between monitoring an orchard-level softening trend, scheduling a follow-up assessment, and classifying individual fruit. Prediction intervals, date-specific bias, calibration uncertainty, and explicitly defined industry tolerances should be reported alongside RMSE, MAE, R^2, and RPD. If model-derived inputs such as estimated DMC are included in future versions, their uncertainty should also be propagated into the final firmness predictions.

In summary, the present work establishes an interpretable foundation for combining biological knowledge with heterogeneous sensor data, but further progress depends more on improved study design and matched multi-season data than on increasing model complexity alone. A larger, more balanced, and longitudinally linked dataset would allow the framework to be evaluated more rigorously and would determine whether its promising population-level performance can be translated into a reliable decision-support tool.

Bibliography

Abdi, H. and Williams, L. J. (2012). Partial least squares methods: Partial least squares correlation and partial least square regression. In *Computational Toxicology: Volume II*, pages 549–579. Springer.

Aitken, A. G. and Warrington, I. J. (2017). Fresh facts: New Zealand horticulture 2017. Industry report, The New Zealand Institute for Plant & Food Research Limited, Auckland, New Zealand.

Barry, C. S. and Giovannoni, J. J. (2007). Ethylene and fruit ripening. *Journal of Plant Growth Regulation*, 26(2):143–159.

Ben-Arie, R., Kislev, N., and Frenkel, C. (1979). Ultrastructural changes in the cell walls of ripening apple and pear fruit. *Plant Physiology*, 64(2):197–202.

Berüter, J. (1985). Sugar accumulation and changes in the activities of related enzymes during development of the apple fruit. *Journal of Plant Physiology*, 121(4):331–341.

Berüter, J., Feusi, M. E. S., and Rüedi, P. (1997). Sorbitol and sucrose partitioning in the growing apple fruit. *Journal of Plant Physiology*, 151(3):269–276.

Bieleski, R. (1969). Accumulation and translocation of sorbitol in apple phloem. *Australian Journal of Biological Sciences*, 22(3):611–620.

Blankenship, S. M. and Unrath, C. (1988). Internal ethylene levels and maturity of 'Delicious' and 'Golden Delicious' apples destined for prompt consumption. *Journal of the American Society for Horticultural Science*, 113(1):88–91.

Blanpied, G. and Wilde, M. H. (1968). A study of the cells in the outer flesh of developing McIntosh apple fruits. *Botanical Gazette*, 129(3):173–183.

Bonnin, E. and Lahaye, M. (2013). Contribution of cell wall modifying enzymes on the texture of fleshy fruits: The example of apple. *Journal of the Serbian Chemical Society*, 78(3):417–427.

Bramlage, W. J. (1992). Interactions of orchard factors and mineral nutrition on quality of pome fruit. In *International Symposium on Pre-and Postharvest Physiology of Pome-fruit 326*, pages 15–28.

Brighenti, S., Tagliavini, M., Comiti, F., Aguzzoni, A., Giuliani, N., Abdelkader, A. B., Penna, D., and Zanotelli, D. (2024). Drip irrigation frequency leads to plasticity in root water uptake by apple trees. *Agricultural Water Management*, 298:108870.

Burg, S. P. and Burg, E. A. (1967). Molecular requirements for the biological activity of ethylene. *Plant Physiology*, 42(1):144–152.

Burke, H. B. (1996). Statistical analysis of complex systems in biomedicine. In *Learning from data: Artificial intelligence and statistics V*, pages 251–258. Springer.

Butler, D. (1976). Estimation of the transpiration rate in an apple orchard from net radiation and vapour pressure deficit measurements. *Agricultural Meteorology*, 16(2):277–289.

Carella, A., Massenti, R., and Lo Bianco, R. (2023). Testing effects of vapor pressure deficit on fruit growth: A comparative approach using peach, mango, olive, orange, and loquat. *Frontiers in Plant Science*, 14:1294195.

Çataltaş, Ö. and Tütüncü, K. (2021). A review of data analysis techniques used in near-infrared spectroscopy. *Avrupa Bilim ve Teknoloji Dergisi*, (25):475–484.

Chagné, D., Dayatilake, D., Diack, R., Oliver, M., Ireland, H., Watson, A., Gardiner, S. E., Johnston, J. W., Schaffer, R. J., and Tustin, S. (2014). Genetic and environmental control of fruit maturation, dry matter and firmness in apple (Malus× domestica Borkh.). *Horticulture Research*, 1:14046.

Cheng, J., Guo, W., Zhang, Z., Zeng, S., and Wang, Z. (2022). Optical properties of 'Gala'(Malus pumila) apple pulp and their relationship with internal quality. *Infrared Physics & Technology*, 123:104210.

Ciurczak, E. W., Igne, B., Workman Jr, J., and Burns, D. A. (2021). *Handbook of near-infrared analysis*. CRC press.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

Costa, G., Noferini, M., Fiori, G., Torrigiani, P., et al. (2009). Use of Vis/NIR spectroscopy to assess fruit ripening stage and improve management in post-harvest chain. *Fresh Prod*, 1:35–41.

Dar, J. A., Wani, A. A., Ahmed, M., Nazir, R., Zargar, S. M., and Javaid, K. (2019). Peel colour in apple (Malus× domestica Borkh.): An economic quality parameter in fruit market. *Scientia Horticulturae*, 244:50–60.

De Smedt, V., Barreiro, P., Verlinden, B., Veraverbeke, E., De Baerdemaeker, J., and Nicolaı, B. (2002). A mathematical model for the development of mealiness in apples. *Postharvest Biology and Technology*, 25(3):273–291.

Denne, M. P. (1960). The growth of apple fruitlets, and the effect of early thinning on fruit development. *Annals of Botany*, 24(95):397–406.

Fan, G., Zha, J., Du, R., and Gao, L. (2009). Determination of soluble solids and firmness of apples by Vis/NIR transmittance. *Journal of Food Engineering*, 93(4):416–420.

Fernández-Cancelo, P., Muñoz, P., Echeverría, G., Larrigaudière, C., Teixidó, N., Munné-Bosch, S., and Giné-Bordonaba, J. (2022). Ethylene and abscisic acid play a key role in modulating apple ripening after harvest and after cold-storage. *Postharvest Biology and Technology*, 188:111902.

Fishman, S. and Génard, M. (1998). A biophysical model of fruit growth: simulation of seasonal and diurnal dynamics of mass. *Plant, Cell & Environment*, 21(8):739–752.

Forsline, P. L., Aldwinckle, H. S., Dickson, E. E., Luby, J. J., and Hokanson, S. C. (2003). Collection, maintenance, characterization, and utilization of wild apples of Central Asia. In Janick, J., editor, *Horticultural Reviews: Wild Apple and Fruit Trees of Central Asia*, volume 29, pages 1–61. John Wiley & Sons, Hoboken, NJ.

Fritz, M. and Berger, P. D. (2015). Can you relate in multiple ways? multiple linear regression and stepwise regression. In *Improving the user experience through practical data analytics*, pages 239–269. Morgan Kaufmann.

Geladi, P., MacDougall, D., and Martens, H. (1985). Linearization and scatter-correction for near-infrared reflectance spectra of meat. *Applied Spectroscopy*, 39(3):491–500.

George, A. and Nissen, R. (1988). The effects of temperature, vapour pressure deficit and soil moisture stress on growth, flowering and fruit set of custard apple (Annona cherimola× Annona squamosa) 'African Pride'. *Scientia Horticulturae*, 34(3-4):183–191.

Gillaspy, G., Ben-David, H., and Gruissem, W. (1993). Fruits: A developmental perspective. *The Plant Cell*, 5(10):1439.

Giovanelli, G., Sinelli, N., Beghi, R., Guidetti, R., and Casiraghi, E. (2014). NIR spectroscopy for the optimization of postharvest apple management. *Postharvest Biology and Technology*, 87:13–20.

Goncalves, M. W., Argenta, L. C., and MARTIN, M. S. D. (2017). Maturity and quality of apple fruit durinig the harvest period at apple industry. *Revista Brasileira de Fruticultura*, 39(5):e–825.

Grabska, J., Beć, K. B., Ueno, N., and Huck, C. W. (2023). Analyzing the quality parameters of apples by spectroscopy from Vis/NIR to NIR region: A comprehensive review. *Foods*, 12(10):1946.

Gwanpua, S. G., Verlinden, B., Hertog, M., Bulens, I., Van de Poel, B., Van Impe, J., Nicolaï, B., and Geeraerd, A. (2012). Kinetic modeling of firmness breakdown in 'Braeburn' apples stored under different controlled atmosphere conditions. *Postharvest Biology and Technology*, 67:68–74.

Gwanpua, S. G., Verlinden, B. E., Hertog, M. L., Nicolai, B. M., and Geeraerd, A. H. (2017). A mechanistic modelling approach to understand 1-mcp inhibition of ethylene action and quality changes during ripening of apples. *Journal of the Science of Food and Agriculture*, 97(11):3802–3813.

Hansen, P. (1979). ^{14}C-studies on apple trees. IX. seasonal changes in the formation of fruit constituents and their subsequent conversions. *Physiologia Plantarum*, 47(3):190–194.

Harker, F., Maindonald, J., Murray, S., Gunson, F., Hallett, I., and Walker, S. (2002). Sensory interpretation of instrumental measurements 1: Texture of apple fruit. *Postharvest Biology and Technology*, 24(3):225–239.

Harker, F. R., Gunson, F. A., and Jaeger, S. R. (2003). The case for fruit quality: An interpretive review of consumer attitudes, and preferences for apples. *Postharvest Biology and Technology*, 28(3):333–347.

Harker, F. R., Kupferman, E. M., Marin, A. B., Gunson, F. A., and Triggs, C. M. (2008). Eating quality standards for apples based on consumer preferences. *Postharvest Biology and Technology*, 50(1):70–78.

Ireland, H. S., Gunaseelan, K., Muddumage, R., Tacken, E. J., Putterill, J., Johnston, J. W., and Schaffer, R. J. (2014). Ethylene regulates apple (Malus× domestica) fruit softening through a dose× time-dependent mechanism and through differential sensitivities and dependencies of cell wall-modifying genes. *Plant and Cell Physiology*, 55(5):1005–1016.

Izenman, A. J. (2013). Linear discriminant analysis. In *Modern multivariate statistical techniques: regression, classification, and manifold learning*, pages 237–280. Springer.

Jahed, K. R. and Hirst, P. M. (2023). Fruit growth and development in apple: A molecular, genomics and epigenetics perspective. *Frontiers in Plant Science*, 14:1122397.

Johnston, J. W., Hewett, E. W., and Hertog, M. L. (2002a). Postharvest softening of apple (Malus domestica) fruit: A review. *New Zealand Journal of Crop and Horticultural Science*, 30(3):145–160.

Johnston, J. W., Hewett, E. W., Hertog, M. L., and Harker, F. R. (2002b). Temperature and ethylene affect induction of rapid softening in 'Granny Smith' and 'Pacific Rose™' apple cultivars. *Postharvest Biology and Technology*, 25(3):257–264.

Kellerhals, M. (2009). Introduction to apple (Malus× domestica). In *Genetics and genomics of Rosaceae*, pages 73–84. Springer.

Kumar, A. and Singh, B. K. (2023). Pre-harvest treatments affecting the post-harvest quality of fresh produce. In *Postharvest Management of Fresh Produce*, pages 1–17. Elsevier.

Kviklienė, N., Kviklys, D., Valiuškaitė, A., Viskelis, P., Uselis, N., Lanauskas, J., and Buskiene, L. (2011). Effect of harvest date on fruit maturity, quality and storability of 'Lodel' apples. *Journal of Food, Agriculture & Environment*, 9(3-4):132–135.

Lachapelle, M., Bourgeois, G., and DeEll, J. R. (2013). Effects of preharvest weather conditions on firmness of 'McIntosh' apples at harvest time. *HortScience*, 48(4):474–480.

Lakso, A. N. (2018). Apple. *Handbook of environmental physiology of fruit crops*, pages 3–42.

Landsberg, J., Blanchard, T., and Warrit, B. (1976). Studies on the movement of water through apple trees. *Journal of Experimental Botany*, 27(4):579–596.

Lau, O., Liu, Y., and Yang, S. (1986). Effects of fruit detachment on ethylene biosynthesis and loss of flesh firmness, skin color, and starch in ripening 'Golden Delicious' apples. *Journal of the American Society for Horticultural Science*, 111(5):731–734.

Lee, L. C., Liong, C.-Y., and Jemain, A. A. (2018). Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: A review of contemporary practice strategies and knowledge gaps. *Analyst*, 143(15):3526–3539.

Li, B., Lecourt, J., and Bishop, G. (2018). Advances in non-destructive early assessment of fruit ripeness towards defining optimal time of harvest and yield prediction – A review. *Plants*, 7(1):3.

Li, M., Pullanagari, R. R., Pranamornkith, T., Yule, I. J., and East, A. R. (2017). Quantitative prediction of post storage 'Hayward' kiwifruit attributes using at harvest Vis-NIR spectroscopy. *Journal of Food Engineering*, 202:46–55.

Lu, R. and Bailey, B. B. (2005). NIR measurement of apple fruit soluble solids content and firmness as affected by postharvest storage. In *2005 ASAE Annual Meeting*, page 1. American Society of Agricultural and Biological Engineers.

Ma, T., Li, X., Inagaki, T., Yang, H., and Tsuchikawa, S. (2018). Noncontact evaluation

of soluble solids content in apples by near-infrared hyperspectral imaging. *Journal of Food Engineering*, 224:53–61.

Ma, T., Xia, Y., Inagaki, T., and Tsuchikawa, S. (2021). Rapid and nondestructive evaluation of soluble solids content (SSC) and firmness in apple using Vis–NIR spatially resolved spectroscopy. *Postharvest Biology and Technology*, 173:111417.

Małachowska, M. and Tomala, K. (2023). Apple quality during shelf-life after long-term storage and simulated transport. *Agriculture*, 13(11):2045.

Malvandi, A., Kapoor, R., Feng, H., and Kamruzzaman, M. (2022). Non-destructive measurement and real-time monitoring of apple hardness during ultrasonic contact drying via portable NIR spectroscopy and machine learning. *Infrared Physics & Technology*, 122:104077.

Mandel, J. (1982). Use of the singular value decomposition in regression analysis. *The American Statistician*, 36(1):15–24.

Marečková, M., Danková, V., Zelenỳ, L., and Suran, P. (2022). Non-destructive near infrared spectroscopy externally validated using large number sets for creation of robust calibration models enabling prediction of apple firmness. *Journal of Near Infrared Spectroscopy*, 30(2):97–104.

Mark, H. and Workman Jr, J. (2018). *Chemometrics in spectroscopy*. Elsevier.

Matabura, V. V. (2022). Modelling of firmness variability of Jonagold apple during postharvest storage. *Journal of Food Science and Technology*, 59(4):1487–1498.

McGlone, V. A., Jordan, R. B., and Martinsen, P. J. (2002). Vis/NIR estimation at harvest of pre-and post-storage quality indices for 'Royal Gala' apple. *Postharvest Biology and Technology*, 25(2):135–144.

McGlone, V. A. and Kawano, S. (1998). Firmness, dry-matter and soluble-solids assessment of postharvest kiwifruit by NIR spectroscopy. *Postharvest Biology and Technology*, 13(2):131–141.

McKenna, M., Le Heron, R., and Roche, M. (2001). Living local, growing global: Renegotiating the export production regime in New Zealand's pipfruit sector. *Geoforum*, 32(2):157–166.

Mendoza, F., Lu, R., Ariana, D., Cen, H., and Bailey, B. (2011). Integrated spectral and image analysis of hyperspectral scattering data for prediction of apple fruit firmness and soluble solids content. *Postharvest Biology and Technology*, 62(2):149–160.

Mendoza, F., Lu, R., and Cen, H. (2012). Comparison and fusion of four nondestructive sensors for predicting apple fruit firmness and soluble solids content. *Postharvest Biology and Technology*, 73:89–98.

Menesatti, P., Zanella, A., D'Andrea, S., Costa, C., Paglia, G., and Pallottino, F. (2009). Supervised multivariate analysis of hyper-spectral NIR images to evaluate the starch index of apples. *Food and Bioprocess Technology*, 2(3):308–314.

Mierczak, K. and Garus-Pakowska, A. (2024). An overview of apple varieties and the importance of apple consumption in the prevention of non-communicable diseases – A narrative review. *Nutrients*, 16(19):3307.

Ministry for Primary Industries (2026). Situation and outlook for primary industries: June 2026. Forecast report, Ministry for Primary Industries, Wellington, New Zealand.

Møller, S. M., Travers, S., Bertram, H. C., and Bertelsen, M. G. (2013). Prediction of postharvest dry matter, soluble solids content, firmness and acidity in apples (cv. Elshof) using NMR and NIR spectroscopy: A comparative study. *European Food Research and Technology*, 237(6):1021–1024.

Musacchi, S. and Serra, S. (2018). Apple fruit quality: Overview on pre-harvest factors. *Scientia Horticulturae*, 234:409–430.

Naes, T., Isaksson, T., and Kowalski, B. (1990). Locally weighted regression and scatter correction for near-infrared reflectance data. *Analytical Chemistry*, 62(7):664–673.

Næs, T. and Martens, H. (1988). Principal component regression in NIR analysis: viewpoints, background details and selection of components. *Journal of Chemometrics*, 2(2):155–167.

Ncama, K., Magwaza, L. S., Mditshwa, A., and Tesfay, S. Z. (2018). Application of visible to near-infrared spectroscopy for non-destructive assessment of quality parameters of fruit. In *Infrared Spectroscopy-Principles, Advances, and Applications*. IntechOpen.

Negi, S., Kumar, P., Kumar, A., Kumar, V., and Irfan, M. (2025). Environmental variables affecting apple fruit development and bioactive compounds: Biochemical insights and biotechnological advances. *Journal of Plant Growth Regulation*, pages 1–20.

Nicolai, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., and Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology*, 46(2):99–118.

Ornelas-Paz, J. d. J., Quintana-Gallegos, B. M., Escalante-Minakata, P., Reyes-Hernández, J., Pérez-Martínez, J. D., Rios-Velasco, C., and Ruiz-Cruz, S. (2018). Relationship between the firmness of Golden Delicious apples and the physicochemical characteristics of the fruits and their pectin during development and ripening. *Journal of food science and technology*, 55(1):33–41.

Palmer, J. W., Harker, F. R., Tustin, D. S., and Johnston, J. (2010). Fruit dry matter concentration: A new quality metric for apples. *Journal of the Science of Food and Agriculture*, 90(15):2586–2594.

Pandiselvam, R., Prithviraj, V., Manikantan, M., Kothakota, A., Rusu, A. V., Trif, M., and Mousavi Khaneghah, A. (2022). Recent advancements in NIR spectroscopy for assessing the quality and safety of horticultural products: A comprehensive review. *Frontiers in Nutrition*, 9:973457.

Park, B., Abbott, J. A., Lee, K.-J., Choi, C., and Choi, K. (2003). Near-infrared diffuse reflectance for quantitative and qualitative measurement of soluble solids and firmness of Delicious and Gala apples. *Transactions of the ASAE*, 46(6):1721–1731.

Peirs, A., Lammertyn, J., Ooms, K., and Nicolaï, B. M. (2001). Prediction of the optimal picking date of different apple cultivars by means of VIS/NIR-spectroscopy. *Postharvest Biology and Technology*, 21(2):189–199.

Pirouz, D. M. (2006). An overview of partial least squares. SSRN Working Paper No. 1631359.

Plotto, A., Azarenko, A. N., McDaniel, M. R., Crockett, P. W., and Mattheis, J. P. (1997). Eating quality of 'Gala' and 'Fuji' apples from multiple harvests and storage durations. *HortScience*, 32(5):903–908.

Porep, J. U., Kammerer, D. R., and Carle, R. (2015). On-line application of near infrared (NIR) spectroscopy in food production. *Trends in Food Science & Technology*, 46(2):211–230.

Pourdarbani, R., Sabzi, S., Kalantari, D., and Arribas, J. I. (2020). Non-destructive visible and short-wave near-infrared spectroscopic data estimation of various physicochemical properties of Fuji apple (Malus pumila) fruits at different maturation stages. *Chemometrics and Intelligent Laboratory Systems*, 206:104147.

Priyadarshan, P. and Jain, S. M. (2009). *Breeding plantation tree crops: Temperate species*. Springer.

Qi, X.-d., Wei, J.-m., Li, H., and Zhao, D. (2015). Cell wall metabolism and related gene expression in Malus domestica Borkh. during fruit growth and softening. *Fruits*, 70(3):153–161.

Redgwell, R. J. and Fischer, M. (2002). Fruit texture, cell wall metabolism and consumer perceptions. In Knee, M., editor, *Fruit Quality and Its Biological Basis*, chapter 3, pages 46–88. Sheffield Academic Press, Sheffield, UK.

Rinnan, Å., Van Den Berg, F., and Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry*, 28(10):1201–1222.

Róth, E., Hertog, M. L., Kovács, E., and Nicolaï, B. (2008). Modelling the enzymatic softening of apples in relation to cultivar, growing system, picking date and season. *International Journal of Food Science and Technology*, 43(4):620–628.

Saei, A., Tustin, D., Zamani, Z., Talaie, A., and Hall, A. (2011). Cropping effects on the loss of apple fruit firmness during storage: The relationship between texture retention and fruit dry matter concentration. *Scientia Horticulturae*, 130(1):256–265.

Sams, C. E. (1999). Preharvest factors affecting postharvest texture. *Postharvest Biology and Technology*, 15(3):249–254.

Savitzky, A. and Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8):1627–1639.

Shenk, J. S., Workman Jr, J. J., and Westerhaus, M. O. (2007). Application of NIR spectroscopy to agricultural products. In *Handbook of near-infrared analysis*, pages 365–404. CRC Press.

Steudle, E. and Wieneke, J. (1985). Changes in water relations and elastic properties of apple fruit cells during growth and development. *Journal of the American Society for Horticultural Science*, 110(6):824–829.

Stevens, A. and Ramirez-Lopez, L. (2014). *An Introduction to the* `prospectr` *Package*. R package vignette, version 0.1.3.

Su, Q., Li, X., Wang, L., Wang, B., Feng, Y., Yang, H., and Zhao, Z. (2022). Variation in cell wall metabolism and flesh firmness of four apple cultivars during fruit development. *Foods*, 11(21):3518.

Sulistya, Y. I., Istighosah, M., Septiara, M., Septiadi, A. D., and Amrullah, A. (2024). Classification of Noni fruit ripeness using support vector machine (SVM) method. *Indonesian Journal of Data and Science*, 5(3):206–215.

Sun, Y., Shi, Z., Jiang, Y., Zhang, X., Li, X., and Li, F. (2021). Effects of preharvest regulation of ethylene on carbohydrate metabolism of apple (Malus domestica Borkh cv. Starkrimson) fruit at harvest and during storage. *Scientia Horticulturae*, 276:109748.

Szymanska-Chargot, M., Adamiak, A., and Zdunek, A. (2012). Pre-harvest monitoring of apple fruits development with the use of biospeckle method. *Scientia Horticulturae*, 145:23–28.

Tacken, E., Ireland, H., Gunaseelan, K., Karunairetnam, S., Wang, D., Schultz, K., Bowen, J., Atkinson, R. G., Johnston, J. W., Putterill, J., et al. (2010). The role of ethylene and cold temperature in the regulation of the apple POLYGALACTURONASE1 gene and fruit softening. *Plant Physiology*, 153(1):294–305.

Tao, Z., Chen, G., Wang, X., and Siddique, K. H. (2023). Root water uptake model shows age-related water uptake patterns of apple trees on the Chinese Loess Plateau. *Journal of Hydrology: Regional Studies*, 50:101594.

Tijero, V., Girardi, F., and Botton, A. (2021). Fruit development and primary metabolism in apple. *Agronomy*, 11(6):1160.

Toivonen, P. M. and Lannard, B. (2020). Dry matter content association with time of on-tree maturation, quality at harvest, and changes in quality after controlled atmosphere storage for 'Royal Gala' apples. *Canadian Journal of Plant Science*, 101(1):98–106.

Van Der Sman, R. and Sanders, M. (2012). Prediction of postharvest firmness of apple using biological switch model. *Journal of Theoretical Biology*, 310:239–248.

Vicente, A. R., Saladié, M., Rose, J. K., and Labavitch, J. M. (2007). The linkage between cell wall metabolism and fruit softening: Looking to the future. *Journal of the Science of Food and Agriculture*, 87(8):1435–1448.

Volz, R., Harker, F., Hallet, I., and Lang, A. (2002). Development of texture in apple fruit – A biophysical perspective. In *XXVI International Horticultural Congress: Key Processes in the Growth and Cropping of Deciduous Fruit and Nut Trees 636*, pages 473–479.

Volz, R. K., Harker, F. R., and Lang, S. (2003). Firmness decline in 'Gala' apple during fruit development. *Journal of the American Society for Horticultural Science*, 128(6):797–802.

Walsh, K. B., Blasco, J., Zude-Sasse, M., and Sun, X. (2020). Visible-NIR 'point' spectroscopy in postharvest fruit and vegetable assessment: The science behind three decades of commercial use. *Postharvest Biology and Technology*, 168:111246.

Wang, H., Peng, J., Xie, C., Bao, Y., and He, Y. (2015). Fruit quality evaluation using spectroscopy technology: A review. *Sensors*, 15(5):11889–11927.

Warrington, I. J. (2000). New Zealand horticulture–success of an export-led industry. In *Annual IDFTA Conference report, Napier, New Zealand. http://www. virtualorchard. net/idfta/cft/2000/july/cftjuly2000p66. pdf.*

Wei, J., Ma, F., Shi, S., Qi, X., Zhu, X., and Yuan, J. (2010). Changes and postharvest regulation of activity and gene expression of enzymes related to cell wall degradation in ripening apple fruit. *Postharvest Biology and Technology*, 56(2):147–154.

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R.* Chapman and Hall/CRC, 2 edition.

Xiao, P. and Chen, D. (2022). Near infrared spectra data analysis by using machine learning algorithms. In *Science and Information Conference*, pages 532–544. Springer.

Xiao, X. and Li, M. (2022). Fusion of data-driven model and mechanistic model for kiwifruit flesh firmness prediction. *Computers and Electronics in Agriculture*, 193:106651.

Xu, X., Xie, L., and Ying, Y. (2019). Factors influencing near infrared spectroscopy analysis of agro-products: A review. *Frontiers of Agricultural Science and Engineering*, 6(2):105–115.

Yan, C. (2025). A review on spectral data preprocessing techniques for machine learning and quantitative analysis. *iScience*, 28(7):112759.

Zanotelli, D., Montagnani, L., Andreotti, C., and Tagliavini, M. (2019). Evapotranspiration and crop coefficient patterns of an apple orchard in a sub-humid environment. *Agricultural Water Management*, 226:105756.

Zeng, S., Zhang, Z., Cheng, X., Cai, X., Cao, M., and Guo, W. (2024). Prediction of soluble solids content using near-infrared spectra and optical properties of intact apple and pulp applying PLSR and CNN. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 304:123402.

Zhang, L. and Wang, Y. (2023). Analysis of soil water potential characteristics of wheat croplands and apple orchards in an agroforestry ecosystem based on the van Genuchten model. *Journal of Soil and Water Conservation*, 78(1):33–43.

Zhang, Y., Nock, J. F., Al Shoffe, Y., and Watkins, C. B. (2019). Non-destructive prediction of soluble solids and dry matter contents in eight apple cultivars using near-infrared spectroscopy. *Postharvest Biology and Technology*, 151:111–118.

Zude, M., Herold, B., Roger, J.-M., Bellon-Maurel, V., and Landahl, S. (2006). Non-destructive tests on the prediction of apple fruit flesh firmness and soluble solids content on tree and in shelf life. *Journal of Food Engineering*, 77(2):254–260.

Zude-Sasse, M., Truppel, I., and Herold, B. (2002). An approach to non-destructive apple fruit chlorophyll determination. *Postharvest Biology and Technology*, 25(2):123–133.