Leaf Wetness Duration Modelling using Adaptive Neuro Fuzzy Inference System

Fatthy Amir

A thesis submitted to

Auckland University of Technology

in partial fulfilment of the requirement for the degree of

Master of Computer and Information Sciences (MCIS)

June, 2016

School of Engineering, Computer

& Mathematical Sciences

Co Supervisors:

Professor Philip Sallis

Associate Professor. Jacqueline Whalley

Akbar Ghobakhlou

ABSTRACT

Recent computational innovations in horticulture are geared towards giving farmers reliable tools with which to predict disease before it happens. Such disease warning systems should enable farmers to take preventive measures before their crops are infected. One of the main inputs for plant disease warning systems is leaf wetness. Leaf wetness refers to the presence of water drops on leaf surface, and is caused by rainfall, dew, or guttation. Leaf wetness can be determined either by measurement or estimation using meteorological and other plant variables.

Because leaf wetness or leaf wetness duration is so critical to the onset of many common fungal diseases in grapes there is a need, especially in wine growing countries like New Zealand, to find the most accurate way to determine leaf wetness. Leaf wetness sensors are not standard even in vineyards where sensors are deployed and these sensors are known to be costly to maintain and less than reliable.

Leaf wetness measurements are taken using sensors that are placed in crop canopy and the data logged usually for later use rather than real time processing. There are different types of leaf wetness sensors with no universally accepted measurement standard. This thesis presents a comparative analysis of various sensors that are commercially available. The sensors used in the experiment have different sizes, shapes, and working principles. Visual observations were made and a sensor response time test was performed to evaluate sensors’ performance. Using a dielectric constant based sensor to measure leaf wetness was shown to be more effective than resistance-based sensors. Paint application to the sensor also proved to increase sensitivity in larger sized sensors.

The rapid development and varied nature of these sensors have contributed to a lack of standardisation and the lack of a single accepted protocol for the use of sensors. An alternative to the use of sensors is the simulation or modelling of leaf surface wetness. Simulation enables surface wetness to be estimated from historical, forecast weather data, or both, rather than from monitoring and measurement using in-field leaf wetness sensors.

The primary focus of this thesis was to develop and evaluate a novel Adaptive Neuro-Fuzzy Inference Systems (ANFIS) as an approach to modelling leaf wetness duration in vineyards. A comparative analysis of ANFIS with Classification and Regression Tree/Stepwise Linear Discriminant (CART), Number of Hours Relative Humidity Greater than 90% (NHRH>90%) model, Fuzzy Logic System (FLS) model, an Artificial Neural Network (ANN), the Penman-Monteith (P-M) model, and the Surface Wetness Energy Balance (SWEB) model is presented. The ANFIS model was found to have a higher accuracy than other models investigated. These findings indicate that ANFIS is a useful and practical solution for estimating leaf wetness duration using meteorological predictor variables.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ATTESTATION OF AUTHORSHIP

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Auckland, 17 June 2016

____

Fatthy Amir

ACKNOWLEDGEMENTS

This thesis would not be possible to finish without the stupendous support from all my three supervisors. My additional supervisor, Akbar Ghobakhlou is bold, precise, and critical. Akbar has taught me a lot about academic writing, being precise and clear in every sentence, and has always challenged any ideas that I come up with and make it bold together. He has always welcomed me in his office even in his busiest times. I have had my ups and downs during these experiments and thesis writing, but Akbar has never shown less patience and motivation until submission date. His constant encouragements also inspired me to always keep doing what I have started and finish it. I have had the pleasant experience co-writing two journal papers with Akbar and published both on time. I am tremendously grateful that I can finally wrap up my thesis thanks to his supports.

It is also a great pleasure working with Associate Professor Jacqueline Whalley. Jacqui has helped me greatly in structuring the thesis, identifying key papers, and writing in general. English as my second language did not stop me from writing this thesis because of the kind support from Jacqui. In her busiest times, she can still squeeze in my work to be read and feedbacks are always quick and thorough. Jacqui always reminded me to look at thesis from larger point of view and widen my research from various perspectives. Without her guidance, it would be hard for me to start writing and finish up this thesis.

My first supervisor, Professor Philip Sallis has supported me throughout this thesis journey. Philip has always commenting my work with kind words and positivity. He is a charismatic leader and having him as my supervisor has given me the energy boost to write a good thesis.

My sincere thanks also goes to Jason Wang as my lab mate. Jason knows his way around the tools in the lab and has excellent programming skills. The experiments done in this thesis did not go far from Jason's help. I am also grateful for his companionship in the quiet times in the lab, especially in the weekends.

Besides everyone that was related to the work, I also like to thank other people that has integral importance in supporting me to finish this thesis. I would like to thank my proud single mother, dr. Neneng Faiza Said to have raised me until today. She has given me the true meaning of life and what to fight for. She was the one who encouraged me to get a master degree, but never once forced me into it. She makes my living overseas feels comfortable, along with my brothers, Rayhan Hafidz and Javier Malik. They are my peace of mind that gives me courage to go overseas in the first place.

The closest person I have to a father is my uncle, Farhat Said. I would like to thank him for all his support and motivation. His life advices and motivation makes me strive for a better future. His visions are the ones that I kept in my mind closely and soon will be a reality.

I would like to take this opportunity to also thank my friends that are kindly supporting me. Both academically and mentally. I would like to thank Luqman Bachtiar and Reggio Hartono. They were doing their PhD while I was writing this thesis. They are always there for me when I have to ask some academic related questions that I was too embarrassed about to ask other people. They have kindly passed their good knowledge to me and I am immensely grateful for that.

I would also like to thank Dyah Nindita Sahdarani and I Wayan Sathya Linggawa as my high school mates and my flatmate while I was finishing up this thesis. Their support may not be in academic form, but in food and comfort in our humble apartment.

Last but not least, I would like to thank Alenza Xinnia for being the greatest support in my critical times during thesis writing. Her positivity boosted my spirit to finally finish up this thesis.

CHAPTER 1

INTRODUCTION

An ongoing global problem for agriculture is increasing production costs, associated with the use of agrochemicals, fertilisers, and labour. The demand for agricultural products is always high and it is crucial to manage the cost of production while maintaining environmentally friendly and sustainable farming practices. In agricultural production, disease management is critical and over 13 billion dollars is spent on chemicals in the USA annually (Morton & Staub, 2008).

Grapes are one of the biggest commodities in New Zealand, wine is one of the country's largest global exports. In New Zealand's horticultural industry, wine makes the most income (NZ$1.1B), followed by kiwifruit (NZ$0.9B), and apples (NZ$0.4B), these three comprise the largest contributor to New Zealand's estimated annual 3.5 billion dollar exports (Aitken & Hewett, 2011).

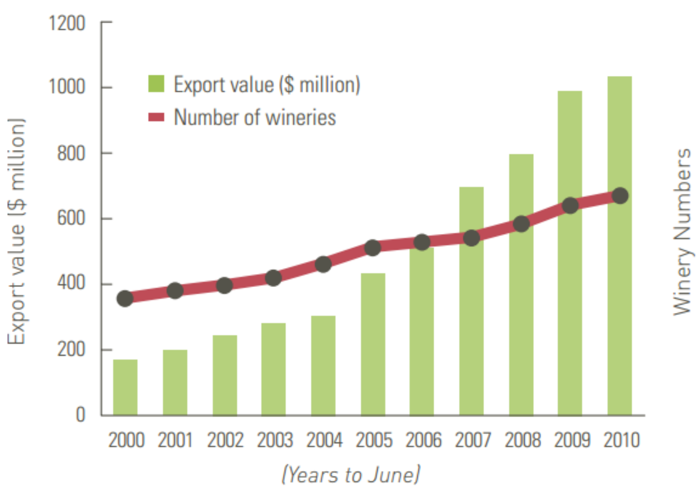

Figure 1. New Zealand wine export value 2000-2010 (Aitken & Hewett, 2011)

Figure 1 shows that the growth of exports from wineries in the last decade has increased by roughly 20%. The export value shown in the bar graph has increased from slightly under $ 200 million in the year 2000 to $ 1 billion in 2010. Export value has had a quintuple growth in a decade, suggesting that New Zealand's wine industry is growing rapidly. The growth of wine as an export and therefore the production of grapes is crucial to New Zealand's economy. Crop loss in vineyards could cause in a serious decline in New Zealand's income. Preventative measures to protect against yield loss and disease is a growing area of research both in New Zealand and Australia.

Because of the highly competitive wine industry any preventative measure must also contribute to reducing costs in grape production while maintaining industry standards. Research related to improving the economic value of grape production includes areas such as increasing fertiliser efficiency (Khurana et al., 2008) and yield loss minimisation (Gaunt, 1995).

Disease control is another important viticulture management activity and recent work includes plant disease prediction and modelling. There are a plethora of different diseases in plants. In grapes the problem mostly lies with diseases caused by fungal infection (Fisher & Wicks, 2003). Fungal infections (e.g., Figure 2) can only occur when certain weather conditions exist such as prolonged periods of rain with warm temperatures.

Figure 2. Fungal diseases (left to right): Botrytis cinerea, Downy Mildew, and Powdery Mildew. Image Credit: Laura Jones/UC Davis

To reduce the risk of infection, growers traditionally use scheduled and regular preventative fungicide spraying. This practice means that vines are often sprayed with agrichemicals unnecessarily. Spraying fungicide is a huge cost to the industry. Fungicide usage has to be moderated as the usage of chemicals on crop is not environmentally friendly and there are always questions about their affect on human health. However, recent technologies and research in this area provide farmers with ways of predicting disease before it happens. These disease warning systems enable farmers to take preventive measures before their crops are infected (Gleason et al., 2008). The way they work is by suggesting fungicide application only when certain weather conditions favourable for disease development are present.

An example of such as system is the Strawberry Advisory Systems (SAS) that has been developed and is currently used in Florida in the United States (Pavan et al., 2006). It is a web-based climate information and decision support system designed to help reduce the risks associated with climate variability and weather changes in the south-eastern region of the United States. SAS monitors temperature and leaf wetness data from six weather stations and automatically calculates disease infection potential and provides spray recommendations for diseases affecting strawberry crop.

One of the main inputs for any plant disease warning systems is leaf wetness. Leaf wetness refers to the presence of water drops on the leaf surface, and is caused by rainfall, dew, or guttation (the secretion of droplets of water from the pores of a plant). To predict plant disease especially in grapes, for which the major diseases are related to fungal infection, leaf wetness is an important variable to measure (Rowlandson et al., 2015). For centuries, relationships between leaf wetness and plant diseases have been studied. However, leaf wetness is still not a standard sensor/measurement found in weather stations. Unlike temperature which is commonly and reliably measured, leaf wetness measurement might not be as reliable. Leaf wetness sensors require very careful and regular calibration, specific deployment steps, and pre-installation treatment (Madeira et al., 2002). There are a number of commercially available leaf wetness sensors. The sensor technology and the recommendations for installation vary and there is no universal agreement or standard as to how leaf wetness is measured or how the sensors are deployed.

To overcome reliability issues associated with leaf wetness sensors, researchers turned to developing mathematical models that can predict leaf wetness using weather variables as inputs. Such models enable leaf surface wetness to be estimated from historical and/or forecast weather data (Huber & Gillespie, 1992). There are more than 20 models that have been developed to estimate leaf wetness (Rowlandson, 2015). Different models take different inputs but all aim to estimate leaf wetness with an accuracy such that the estimates can be used as an input for disease warning systems. Using models to estimate leaf wetness can be an alternative solution to using physical sensors. Modelling requires less maintenance and less man hours to deploy sensors at stations. However in industry physical sensors are more widely used than leaf wetness models due to models being localised and their performance varying in different areas. There is lack of comprehensive research both in evaluating and comparing existing leaf wetness models and there is little work towards developing new models that are more accurate and generalizable. Moreover, leaf wetness models are typically evaluated and developed using data obtained from leaf wetness sensors. This brings into question the reliability of the models which are developed using a data driven approach.

In order to obtain reliable leaf wetness measurement it is important to know which sensors give the most accurate and reliable leaf wetness readings. Therefore there is a need for a comprehensive comparative analysis of commercially available leaf wetness sensors as well as improved leaf wetness models.

This thesis aims to answer the following research questions:

1. Which of the commercially available leaf wetness sensors gives the most accurate measurement?
2. Which of the existing leaf wetness models give the most accurate estimation of leaf wetness for New Zealand vineyards?
3. Can an adaptive neuro-fuzzy inference system be used as a leaf wetness model?

1.1 Thesis Contribution

This thesis proposes a new approach to leaf wetness duration modelling using an adaptive neuro-fuzzy inference system (ANFIS). Previously researchers have used Artificial Neural Networks (ANN) and Fuzzy Inference Systems (FIS) to estimate leaf wetness with reasonably high accuracy. ANFIS is a hybrid model that combines the strengths of ANN and FIS and tends therefore to improve classification accuracy. ANFIS also has the advantage of requiring less explanatory variables than commonly used physical models. ANFIS has not previously been used to estimate either leaf wetness or leaf wetness duration. In this research for the first time ANFIS is used and evaluated against existing models to estimate leaf wetness duration.

1.2 Thesis Organisation

Chapter Two provides an overview of the literature related to leaf wetness. It provides a discussion of many aspects of leaf wetness duration from its relationship with plant disease to its measurement and modelling. This chapter also presents in detail the existing models which will be investigated in this research to determine their applicability to New Zealand and also as benchmarks for evaluating the usefulness of the novel ANFIS model developed as part of this research.

Chapter Three outlines the experimental methods used in this research. From experimental protocols for evaluating sensors to data acquisition, preparation and model evaluation.

Chapter Four presents and discusses the results of both the sensor evaluation experiments and the leaf wetness modelling.

Chapter Five draws conclusions and provides suggestions for future research.

CHAPTER 2

LEAF WETNESS DURATION OVERVIEW

2.1 Importance of Leaf Wetness Duration

Leaf wetness is an important agricultural variable to measure and is often used to predict occurrences of plant diseases. Mechanisms which cause plant disease such as sporulation and infection are commonly influenced by the presence of water on a plant's surface (Yarwood, 1978; Rotem et al., 1978). Studies have found that a sustained period of water presence on a plants surface is necessary for plant disease development (Melching et al., 1989). The length of time wetness remains visible or stays on surface of plants is referred to, in the literature, as Leaf Wetness Duration (LWD). It is important to monitor LWD in order to prevent plant disease outbreaks and to reduce the costs related to the use of agrochemicals for crop protection.

Leaf wetness is a difficult variable to measure and estimate because it is driven by both atmospheric conditions and their interaction with the structure, composition and physiology of the crop canopy (Kudinha, 2014). In large plantations the high spatial variability of LWD requires multiple sensors to acquire accurate measurements. Measuring LWD is also labour sensitive and time consuming in terms of sensor setup and maintenance. Mathematical models to estimate LWD have been studied as an alternative to directly measuring LWD. Using LWD models removes the need for the installation of specialist sensors on site.

2.1.1 Leaf Wetness Duration in Plant Disease

The exploration of the relationship between plant diseases and leaf wetness began as early as 1853 when DeBary became one of the first researchers to associate the infection of *Phytophthora infestans* on potato crop with canopy leaf wetness (Yarwood, 1978). Since then leaf wetness has been identified as a risk factor in the development of many bacterial and fungal diseases in many types of crops. Furthermore, incidences of infection have also been linked with air temperature during wet periods which is necessary for the germination of most phytopathogenic fungi to infect plants (Jones, 1986). Leaf wetness and relative humidity influence germination, sporulation, and infection during the production and transport stages of infection (Schuepp & Peter, 1989). Standard weather station variables that are readily available such as air temperature, relative humidity, and rainfall are important indicators for plant disease. Leaf wetness is not included as standard largely because there is no recognised standard method of measurement to verify the accuracy of leaf wetness sensors (WMO, 2008).

The length of LWD required for plant disease infection varies globally from 0.5 h to >100 h (Magarey, 2005). The three major plant diseases that affect vineyards in New Zealand, namely Powdery mildew (*Erysiphia necator*), Downy Mildew (*Plasmopara viticola*), and Botrytis (*Botrytis cinerea*), are all known to be dependent on leaf wetness. These diseases are all fungal diseases where the infection of crops is weather-driven. For example, a primary infection of Downy mildew requires at least 10mm rainfall and a temperature of above 10°C occurring for over more than a 24 hour period. For the oospores to germinate the soil needs to be wet at least 16 hours with 3-5mm of rainfall as well as air temperature of above 10°C. For secondary infection such as leaf to leaf, leaf to berries, or leaf to shoot to occur there must be at least 98% relative humidity, air temperature of above 13°C, more than four hours of darkness, and two to three hours of leaf wetness near dawn (Magarey, 2010). The same meteorological parameters are used to indicate the likelihood of occurrence of other fungal plant diseases but with varied threshold values and time periods. In New Zealand, these diseases can occur at any stage of the growing season (spring to summer) due to the cool maritime climate (Mundy, Agnew, & Wood, 2012). Leaf wetness is a measurable parameter that can be used as an early warning system for many plant diseases. To measure or estimate leaf wetness it is necessary to understand the meteorological and physical factors that influence the formation of moisture on a leaf surface. Different causes of leaf wetness are associated with different physical processes and meteorological variables.

2.1.2 Causes of Leaf Wetness

To measure leaf wetness, it is necessary to understand the process of moisture formation on a leaf surface. Leaf wetness can be caused by irrigation, guttation, rainfall, or dew. Of these four causes rainfall and dew are the most common.

When irrigation is applied by overhead or above-the-row sprinklers, it has direct effect on leaf wetness (Lomas, 1991). Other irrigation methods increase soil moisture and therefore indirectly cause leaf wetness by promoting the development of dew on the canopy. Guttation occurs when water from inside the leaf are extruded to the leaf surface as a result of an osmotic process produced by the hydathodes tissue (Hughes & Brimblecombe, 1994). Guttation occurs mostly during high relative humidity conditions. The contribution of guttation is minimal to the overall amount of water on a leaf, and is typically ignored when estimating leaf wetness (Jacobs et al., 1994). Wetness caused by irrigation is not easily measured with leaf wetness sensors. Irrigation is often measured in the same way as rainfall, however it requires calibration and different field placement of the sensor to measure overhead irrigation. Irrigation is often ignored in models as it is known to make a minimal contribution to overall leaf wetness.

Leaf wetness caused by rain can last for several days, whereas dew formations usually occur during the night or early morning. Rain has direct effect on leaf wetness as it falls down onto the leaves and surrounding area, increasing soil moisture and possibly promoting dew events. Rainfall as a predictor of leaf wetness has two problems, it can be quite spatially variable and water distribution in tall and/or dense crops tends to be indeterminate (Linacre, 1992; Sellers & Lockwood, 1981).

Dew formation is a result of radiative cooling when the leaf surface temperature drops under the atmospheric dew point. Moisture for dew formation arises from the atmosphere (dew fall) or from the soil (dew rise). Dew rise is most likely to occur when there is almost no wind and the soil is wet. Dew fall usually take place on cloudless nights with low wind speed (under 4 m/s) (Garratt & Segal, 1988). During cloudless nights the downward radiation is reduced and the canopy cools promoting dew deposition. Wind speed heavily influences the evaporation rate and hence dew deposition. When a crop canopy cools down to dew point temperature, slightly below the air temperature, the air around canopy area becomes saturated, forming dew on leaf surfaces. If the wind speed is substantially high (above 4 m/s), drier air moves along the canopy generating less humid condition and reducing the chance of dew deposition (Rowlandson, 2015).

2.1.3 Leaf Wetness Measurement

To measure leaf wetness there are three types of leaf wetness measurement instruments, static instruments, which only give an indication of wet or dry conditions; mechanical sensors, which records the changes of length, size, or weight of the sensor due to wetness; and electronic sensors, which measure the change of sensor impedance caused by wetness. Static leaf wetness instruments are the simplest of the three, with no mechanical or electronic parts. One example is the Duvdevani dew gauge, which is a wooden block that has standardised size and painted red (Duvdevani, 1947). Such devices are known to have a poor correlation with wetness caused by dew. Therefore, static leaf wetness instruments do not provide useful measurements for disease prediction as it does not provide a value for the period of dew duration (Rowlandson et al., 2015). Mechanical sensors have hygroscopic properties and work by measuring changes in the length and weight of the sensor. Mechanical sensors are able to provide a period of dew duration values, which makes superior static sensors (Rowlandson, et al., 2015). Mechanical sensors were used intensively from the 1950s (Hirst, 1954) until around 1970 (Lomas and Shashoua, 1970).

Measuring leaf wetness by employing flat plate electrical grids to record changes in electrical impedance was introduced by Davis and Hughes (1970). The development of these instruments was continued by Gillespie and Kidd (1978). They used mock leaf sensors constructed from electrical impedance grids using a light grey latex paint. They achieved an improvement in the accuracy of detection of leaf wetness onset and dry-off. Over the years more attempts at developing leaf wetness sensors have been attempted resulting in the development of various sensors with different principles and shape. The types of electronic leaf wetness sensors used in this study are discussed in Section 4.2.

2.1.4 Leaf Wetness Measurement Standardisation

According to the latest guide by the World Meteorological Organization (WMO) in 2008, leaf wetness may be described as light, moderate, or heavy and that it's most important measurement is the period of moisture onset or dry-off duration (WMO, 2008). The amount of dew depends mainly on properties of the leaf surface, such as its radiative properties, size, and shape. A leaf wetness measurement standard consists of a set of steps for leaf wetness measurement and sensor installation. For a standard to be useful it must be widely accepted. Having a standard of measurements would be particularly useful for leaf wetness data exchange and allow for wider adoption of models, such as disease warning systems, that are reliant on leaf wetness measurements.

In order to arrive at a widely adopted standard for the use of leaf wetness sensors more research must be undertaken in leaf wetness measurements and installation. To measure leaf wetness several methods are worth considering. The measurement of the amount of dew on a leaf sensor surface depends on properties of the sensor's surface such as its size, radiative properties of sensor materials, and aspect (horizontal or vertical) (WMO, 2008). Leaf wetness may be measured by exposing a plate or surface, assessing dew amount by weighing it or by measuring quantitative properties such as electrical conductivity of the surface. The problem lies in choosing the surface type and installation location/placement in order to get the best representation of the leaf wetness (Papastamati et al., 2004).

The problems associated with leaf wetness measurement are not related only to the sensors themselves, but also to how the sensors are used (Rowlandson, et al., 2015). To achieve the optimal accuracy of a leaf wetness sensor, the physical development (size, shape, materials of the sensors) as well as installation protocols (calibration, height, angle, and orientation) need to be considered and specified in order to establish an accepted standard (Gleason, 2007; Sentelhas et al., 2004). Some of these aspects have been proposed by researchers as standards for leaf wetness sensors, however the results of applying these standards shows that the data obtained from the sensors is not reliable, Therefore, these standards haves not been adopted in practice (Lau et al., 2000; Sentelhas et al., 2004).

Leaf wetness measurement standardisation can be categorised into three main steps (Figure 3); pre-calibration, calibration, and installation (Montone, 2013; Lau et al., 2000).

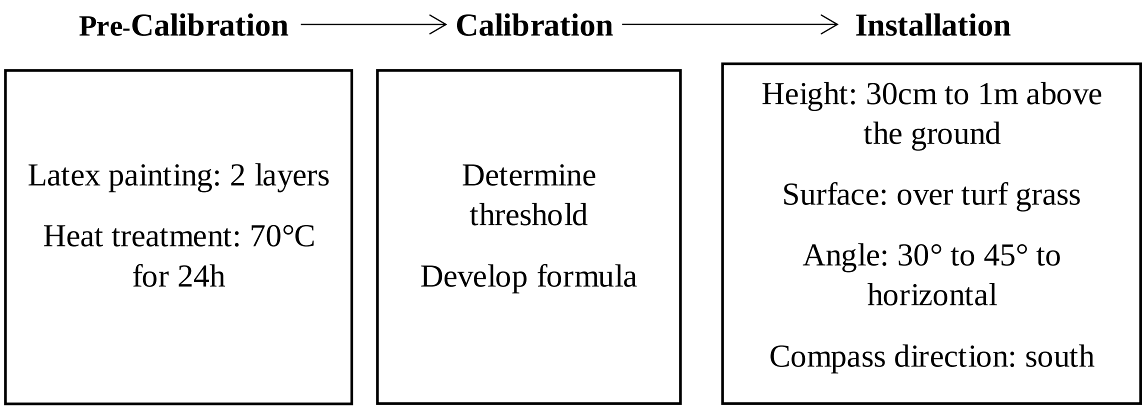

Figure 3. LWS measurement standardisation steps (Montone, 2013; Lau et al., 2000).

The initial pre-calibration step is only applicable to resistance based LWS; and includes standards for the coating and heat treatment of the sensor. It has been suggested that two layers of latex paint should be applied to the sensor in order to provide a surface which allows for the even spread of dew droplets. It is also suggested that the latex coating will make the connection between the bridges more sensitive to small droplets (Davis & Hughes, 1970). After the coating process, the sensor is heat treated at 70°C for 24 hours. This step is recommended as a means of deactivating any hygroscopic substances contained in the paint formula. Applying these standards was shown to reduced variability between sensors from 67% (without coating) to 9% (with coating) (Lau et al., 2000).

The second step is calibration of the sensor. This calibration is performed in a laboratory and in the field. In the laboratory, calibration can be performed by observing the output values when 1 mm of water droplet applied on sensor's surface (Lau et al., 2000; Rao et al., 1998). The water droplet causes a change in the impedance of the sensor which is recorded as the sensor's threshold. To verify a laboratory calibration, field testing must to be conducted. Field testing involves visual observations of dew onset and dry-off to test the laboratory calibration in an outdoor environment (Royle & Butler, 1986). If the field testing does not agree with the laboratory calibration, then the sensors' response from the field test are used to recalibrate the sensor.

In the installation step the following deployment strategies are recommended (Montone, 2013):

- Sensor is placed over turf grass.
- Sensor height is at least 30 cm above the ground.
- Sensor is angled 30° to 45° horizontally.
- Sensor is facing south (in the southern hemisphere).

Instead of measuring leaf wetness above or under crop canopy, turf grass around particular crop has been suggested as a standard surface for leaf wetness measurements since the reading was found to strongly agree with the leaf wetness measured in the upper canopy of corn, cotton, and muskmelon crops (Lau et al., 2000; Rao et al., 1998). Sensors should be deployed at least 30 cm from the ground, angled 30° to 45°, and directed to south if the location is in the southern hemisphere to minimise the interception of solar radiation (Gillespie & Kidd, 1978).

Some of these standardisation steps have more influence on LWS measurement than the others. A sensor's height has more influence than its angle. The greater the sensor's installation height the higher the chance of false negatives being recorded, and the smaller the sensor's installation angle, the higher the likelihood of false positives occurring (Sentelhas et al., 2004; Van Der Wal, 1978). Paint coating the sensor makes a much more significant difference to LWS measurements that the angle and orientation of its deployment. This conclusion is based on research that indicates that unpainted sensors fail to respond to dew onset in up to 30.8% of the cases when it was deployed at a 45° angle (Madeira et al., 2002). When possible, it is best to adopt all of the sensor standardisation when installing a sensor, and when it is not possible paint coating should be first priority as it has the highest impact on the sensor's accuracy.

2.1.5 Leaf Wetness Duration Modelling

An alternative to the use of physical sensors is the use of mathematical models to simulate or estimate LWD using meteorological data. Simulation enables the leaf surface wetness to be estimated from historical and/or forecast weather data, rather than from monitoring and measurement using in-field leaf wetness sensors (Huber & Gillespie, 1992; Weiss, 1990).

The problems associated with leaf wetness measurement using sensors led to the development of models to estimate LWD. There are more than 20 models that have developed to estimate LWD (Rowlandson, 2015). These models range from simple (a model that uses only one input variable) to models that require calculations of leaf surface condensation and evaporation in order to determine leaf wetness. More recent LWD models have been compared with sensors as a ‘reference’ and some of these models have been found to perform favourably (less than 1 misclassification per hour) when compared with sensors (Magarey et al., 2005; Sentelhas et al., 2004).

Various approaches, such as fuzzy logic and neural networks, have been employed to model and characterise leaf wetness patterns. (Jang et al., 1997, Kim et al., 2004). Fuzzy logic and neural networks have proven to achieve higher accuracy and precision than classic statistical approaches (Weiland & Mirschel, 2008). Fuzzy logic and neural networks are known to be methods that are suitable for modelling complex nonlinear functions, dealing with prediction, classification, and pattern recognition problems (Zadeh, 1994, Mellit & Kalogirou, 2008) which make them good candidates for modelling LWD.

In summary, leaf wetness is a difficult variable to measure and cannot be considered a true atmospheric variable as it is related to structural and surface optical properties and microclimate (Sentelhas et al., 2004). Physical changes to the surrounding of a leaf play an important role in the formation of leaf wetness. LWD is a variable that has a close relationship with plant disease, therefore its measurement is essential if we wish to be able to predict possible occurrences of plant disease. However, physical sensor measurement is still unreliable and labour intensive in terms of both setup and maintenance. Mathematical and computational models, using readily available meteorological variables, are proving to provide a practical alternative to measurement of LWD using sensor estimated LWD.

2.2 Leaf Wetness Duration Models

Leaf wetness duration models can be categorised into three classes; empirical, physical, and hybrid (Kim, 2003). Empirical models use meteorological data as their input. Variables which are commonly used in empirical models include air temperature, relative humidity, wind speed, solar radiation, and rainfall. Physical models simulate the physical system of dew formation and evaporation on the leaf surface. For example such a model might incorporate theories of energy balance and/or heat transfer. Hybrid models combine empirical and physical models in order to draw on the strengths of each of those model types.

Empirical models are the most straightforward type of model to implement because they require the simplest inputs and generally these inputs are readily available either directly or as derived from other available inputs (Rowlandson, 2011). In contrast physical models typically require more complex variables but they are more generalizable in terms of geographical location because they rely solely on the physical energy balance system.

This research requires a comparative analysis of existing leaf wetness duration modelling approaches and their applicability to the estimation of leaf wetness in the New Zealand context. Additionally, because this work will entail investigating alternative and possibly novel approaches to estimating leaf wetness duration it is necessary to use existing models as benchmarks for comparison with any new model developed. To date a comprehensive comparative analysis of leaf wetness models has not been reported. Typically newly reported models, regardless of type, are compared with only one or two empirical models.

While there are many LWD models (for examples see Pedro and Gillespie, 1982a,b; Huber and Gillespie, 1992; Gleason et al., 1994; Rao et al., 1998; Sentelhas et al., 2004; Magarey et al., 2005) available many of these have not been widely adopted and are not currently in use. The majority of these models were developed as research models and as a consequence employ variables which are not readily available in practice (Montone, 2013). The simpler models are usually localised and not applicable to different regions or areas unless a complex calibration of the model is undertaken (Sentelhas et al., 2005). The models selected for analysis were chosen based on the following criteria; data availability, level of acceptance, and model type (for comprehensive coverage).

The following section discusses the selected models in detail.

2.2.1 Selected Models

Empirical Models

Two empirical models were selected: Number of Hours of Relative Humidity greater than or equal to 90% (NHRH≥ 90%) model and a Classification and Regression Tree (CART) based model. NHRH≥ 90% is the simplest possible model for estimating LWD. In NHRH≥ 90% relative humidity is essentially used as a proxy for leaf wetness. This model is founded on the assumption that when relative humidity is over a certain level or threshold (90% in most cases), then surface wetness is present (Sentelhas et al., 2008). This model is included because in a recent paper by Rowlandson et al. (2015) it was suggested that growers should adopt NHRH≥ 90% over leaf wetness sensors and other models due the simplicity and relatively low cost of implementation. This recommendation is made for US growers and the authors suggest that relative humidity measurements from the closest available national weather service station is adequate for the purposes of estimating LWD. For the purposes of this study NHRH≥ 90% is considered to be the benchmark standard. If any model does not outperform NHRH≥ 90% then it is considered to be of little use as an LWD estimation method. Applying Occam's razor it is logical to assume that the simplest possible model which in this case is NHRH≥ 90% is the best model, unless an alternate model is significantly more accurate.

Gleason and Koehler (1994) developed an empirical CART/SLD model for estimating the occurrence and duration of dew points. The first step involved using CART (Breiman et al., 1984) to eliminate periods when dew was unlikely to occur. Their model used hourly readings of wind speed, relative humidity and dew point depression (D) to build a classification and regression tree (CART).

D was a derived variable calculated as the difference between air temperature (T_{air}) and dew point temperature (T_{dew}).

$$\mathrm{D} = \mathrm{T}_{air} - T_{dew} \tag{1}$$

Unlike air temperature, dew temperature is not commonly available from a standard weather station so for this model, T_{dew} was calculated using formula:

$$T_{dew} = \left(\frac{RH}{100}\right)^{\left(\frac{1}{8}\right)} * (112 + 0.9\, T_{air}) + 0.1 * T_{air} - 112 \tag{2}$$

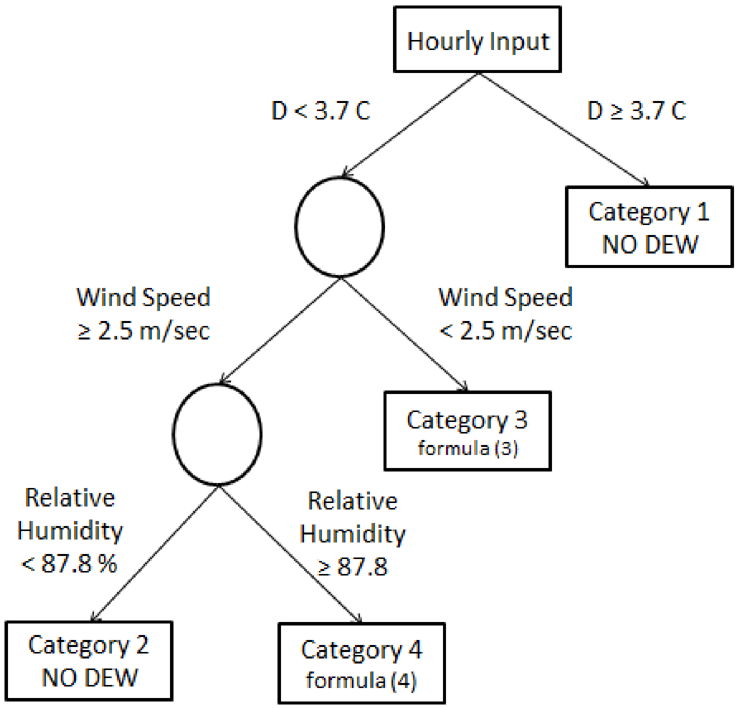

Figure 4. Classification and regression tree generated by CART model (Gleason and Koehler, 1994, p1013)

The resultant classification and regression tree (CART), shown in Figure 4, assigned the hourly data to one of four categories. A high proportion of the data was noted to have arisen in categories 1 and 2; these categories were classified as non-dew hours. However categories 3 and 4 were considered to be non-determinant because they could not be clearly classified into dew or non-dew hours using CART. Therefore, Gleason and Koehler undertook a stepwise linear discriminant (SLD) analysis step on the instances in categories 3 and 4 in order to establish discriminant functions (equations (3) and (4)) to classify these hours into dew and non-dew hours.

For category 3 dew was assumed to occur if equation (3) is satisfied and for category 4 if equation (4) is satisfied.

$$\left[\left(1.6064\sqrt{T_{air}}\right) + \left(0.0036\,{T_{air}}^2\right) + (0.1531\,RH) - (0.4599\,W * D) - (0.0035\,T_{air} * RH)\right] > 14.4674 \quad (3)$$

$$\left[\left(0.7921\sqrt{T_{air}}\right) + \left(0.0046\,RH^2\right) - (23889\,W) - (0.0390\,T_{air} * W) + (1.0613\,W * D)\right] > 37 \quad (4)$$

Where, W is wind speed and RH is relative humidity.

The authors compared the performance of CART/SLD with SLD alone and with NHRH≥ 90%. SLD on its own was found to not be sufficient and the authors noted significant improvement in the model performance when CART was used prior to SLD. They also noted that NHRH≥ 90% gave a much greater root mean square error than CART/SLD for the Iowa case study area. Therefore CART/SLD was found to be more reliable than NHRH≥ 90%. However, it should be noted that the CART/SLD model does not include rainfall and Gleason and Koehler (1994) at the time suggested that this model might need to be further refined in order to be used as an input for disease modelling. Moreover, given that this model is an empirical model it may limit its spatial and temporal portability meaning that it is not generalizable. Despite these limitations CART/SLD was selected for this research because it uses variables which can be obtained from standard weather stations and meteorological services. One reason for considering CART/SLD was that when compared with other empirical models CART/SLD requires relatively few input variables. This choice is also supported by the fact that CART/SLD is reported to outperform NHRH≥ 90% which is adopted in this study as the benchmark technique. More recently the CART/SLD model was reported to also outperform an Artificial Neural Network (ANN) with a backpropagation architecture (ANN is discussed later in this section) to predict wetness on wheat flag leaves using environmental variables (Francl and Panigrahi, 1997).

During the course of the study, after the modelling experiments had been completed, a paper was published which investigated the use of CART for LWD modelling in New Zealand (Henshall, Hill and Beresford 2015). The researchers compared the results of the model with that of several different types of leaf wetness sensors deployed at seven sites across New Zealand from Clyde in the south to Kerikeri in the north. They found that CART consistently underestimated leaf wetness. Their conclusion was that the use of modelled and measured wetness inputs into a grape botrytis bunch rot prediction model indicated that estimated leaf wetness using CART was unsuitable for use in New Zealand without being calibrated for local conditions.

Physical Models

Physical methods for measuring LWD are based on dew deposition and evaporation or intercepted rain and have been shown to have low spatial variability, good portability, and sufficient accuracy for operational use (Gillespie and Barr, 1984; Rao et al., 1998; Dalla Marta et al., 2005). Physical models of LWD include maize, soybean, and apple (Pedro and Gillespie, 1982a, b), onion (Gillespie and Barr, 1984), maize ears (Rao et al., 1998), and grapes (Magarey, 2006, and Dalla Marta et al., 2005). All of these models require net radiation, which is rarely measured, as an input variable (Madeira et al., 2002). Estimations of net radiation can be derived using a combination of incoming solar radiation, air temperature, relative humidity, cloud cover, and cloud height (Madeira et al., 2002). The two physical models selected for this research are Penman-Monteith (P-M) and Surface Wetness Energy Balance (SWEB).

The P-M model is based on an equation that estimates latent heat flux (LE) which is then used to classify surface wetness (SW). The P-M equation (Monteith and Unsworth, 1990; Monteith and Unsworth, 2015) is easier to implement than most physical models because it does not require an air temperature measurement at crop (leaf) level (Sentelhas et al., 2006). Instead the P-M model assumes that air temperature measured at a given height above turf grass, at a standard weather station, is equivalent to the temperature at the top of the crop canopy. An adjustment resistance (r_a) was added to the model in order to account for the air layer from air temperature measurement height to the level of the canopy (Rao et al., 1998).

Like other physical models, the P-M model requires net radiation as an input. Net radiation is almost always an estimated variable and is estimated from other more readily available weather parameters.

The P-M model has been tested on various climate conditions such as tropical area of Philippines (Lou and Goudriaan, 1999), Mediterranean region (Jacobs et al., 2002), southern Canada (Rao et al., 1998), and tropical region of Brazil (Sentelhas et al., 2004), the results have shown that P-M model performs well under different climate conditions. In a study by Sentelhas et al. (2005), P-M model was reported to overestimate LWD by 1.33 hours in average across three locations.

In the rare cases when the LE value is known, SW can be determined by using equations (5) & (6).

$$SW = 1, \text{ when } LE > 0 \text{ or rain begins} \quad (5)$$

$$SW = 0, \text{ when } LE < 0 \text{ and/or estimated rain} < \text{evaporation} \quad (6)$$

When LE is unknown it can be estimated for each interval of time using the P-M formula (7) for leaf wetness:

$$LE = -\frac{\{sR_n + [1200(e_s - e_a)/(r_a + r_b)]}{s + \gamma} \quad (7)$$

Where:

s is slope of the saturation vapour pressure (h Pa),

e_a is actual air vapour pressure (h Pa),

e_s is saturated vapour pressure at the weather station temperature (h Pa),

γ is psychrometric constant (0.64 kPa K-1 during dew or 1.28 kPa K-1 after rain),

sR_n is canopy net radiation (J min-1 cm-2),

r_a is additional resistance which is formulated by equation (8),

$$r_a = \frac{\ln[(Zs - d)/Z0]}{0.4u^*} \quad (8)$$

Where:

Z_s is height of wetness sensor (m),

d is displacement height (-0.65 Z_c),

Z_c is crop height (assumed 0.1 m),

Z_0 is roughness length (=0.13 Z_c),

u^* is friction velocity (m s^{-1}) which is calculated according to equation (9),

$$u^* = \frac{0.4uZT}{\ln[(ZT-d)/Z0]} \tag{9}$$

Where:

Z_T is weather station height (m), and

uZ_T is wind speed at Z_T.

And r_b is the boundary layer resistance for heat transfer (s m^{-1}), calculated by equation (8):

$$r_b = 307(\frac{D}{u})^{1/2} \tag{10}$$

Where:

D is effective dimension of mock leaf (= 0.07m), and

μ is wind speed (m s^{-1}).

In a study by Sentelhas et al. (2005), the P-M model was found to overestimate leaf wetness duration by 1.33 hours on average across three different locations. Despite this overestimation P-M was included in this research as part of the comparative analysis due to it being a model that can be implemented using widely available meteorological variables and because it has been shown to be relatively reliable and generalizable across various crop types, including grapes, and diverse geographic locations.

As an addendum, a paper was published (in May 2016 after this research was completed) which evaluated four LWD models (Montone et al, 2016). The models were CART, dew point depression (Gillespie et al. 1993), NHRH ≥ 90% and the P-M model. The aim of the evaluation was to establish which LWD estimation model provided the best input for the Strawberry Advisory System (SAS). SAS is a system used by strawberry growers in the state of Florida in the USA. Growers use the output of the model to assist in deciding when to spray strawberry fields to control diseases such as Botrytis fruit rot. The data used in the models included air temperature, RH, wind speed, solar radiation, and rainfall collected at heights of 1.5- to 2.0-m in the field. It was reported that the P-M model estimated LWD most accurately at a weather stations which had high precision RH, solar radiation and temperature sensors. However at sites where lower precision sensors were installed the P-M model was found to be the least accurate LWD estimator. The authors concluded that the CART model provided the most robust solution for estimating LWD.

The SWEB model was developed by Magarey (2006) and employs an energy balance principle. The SWEB model is essentially a canopy water budget model. This model is also called "Big Leaf" because it works on the assumption that the whole crop is one big leaf. SWEB consists of four modules, a water distribution module, a canopy water budget module, an energy balance module, and a transfer coefficient module which is calibrated to surface wetness. These later three modules are used to calculate and supply input variables to the first model (the water distribution module). In the first module, the index of the fraction of canopy wet surface area (W_{ind}) is compared with the surface wetness threshold (W_{th}) in order to classify surface wetness (SW) as given by equations (11) and (12):

$$SW = 1, \text{ for } W_{ind} > W_{th}, \tag{11}$$

$$SW = 0, \text{ for } W_{th} > W_{ind} \tag{12}$$

A value of 0.1 for W_{th} was selected based on observations in a grape canopy (Magarey, 2006) that indicated this value was appropriate to minimise the influence of large drops that dry several hours after the rest of the canopy has dried.

$$W_{ind} = \left(\frac{S}{C}\right)^{0.67} \tag{13}$$

Where:

S is canopy water storage (cm) and

C is maximum water storage (cm).

In the canopy water budget module, the maximum water storage (C) is calculated using equation (14):

$$C = LAI\ x\ C_1 \tag{14}$$

Where:

LAI is the leaf area index, and

C_1 is the maximum water storage for an average leaf (=0.02) (cm).

The LAI for a certain crop canopy can be measured or estimated from a model and the maximum water storage for leaf (C_1), is a function of the age and species of the plant. Different plants at different ages may have a different value. However, a commonly accepted value for C_1 is 0.02 cm (Noilhan and Planton, 1989). The canopy water storage (S) can be calculated by equation (15):

$$S = \left(I + D_p - E\right), for\ 0 < S < C \tag{15}$$

Where:

S is water storage (cm),

D_p is potential condensation of dew (cm),

E is evaporation (cm), and

I is intercepted precipitation (cm) and is calculated using equation (16):

$$I = (1 - \exp(-0.5LAI))P \tag{16}$$

Where P is the precipitation/rainfall (cm).

The interception of precipitation estimation formula (16) was derived from the work of Norman and Campbell (1983) and is one of the simplest interception models available.

The energy balance module is employed to compute the condensation and evaporation processes in the canopy water budget model. The formula used to calculate the potential condensation of dew (D_p) is:

$$D_p = \frac{\Delta}{\lambda(\Delta+\delta)} 0.5R_{nc} \tag{17}$$

Where:

R_{nc} is canopy net radiant flux density (J min^{-1} cm^{-2}),

λ is latent heat of vaporisation (J g^{-1}),

δ is psychrometric constant (mbar C^{-1}), and

Δ is slope of the saturation vapour pressure curve (mbar C^{-1}).

In the SWEB model it is assumed that daytime radiation does not contribute to evaporation in the shaded grape canopy. Therefore, the water-loss potential from a wet surface area (E_p) is calculated according to equation (18):

$$E_p = \frac{\Delta}{\lambda(\Delta+\delta)} \left\{\rho C_p \left(\frac{h}{\Delta}\right)(e_{a^*} - e_a)\right\} \tag{18}$$

Where:

R_n is net radiant flux density (J min^{-1} cm^{-2}),

ρ is density of air (g cm^{-3}),

C_p is specific heat of air (J g^{-1} C^{-1}),

h = transfer coefficient for heat and vapour from the surface to the atmosphere (cm min^{-1}),

e_a = water vapour pressure of the atmosphere (mbar), and

e_a* = saturated water vapour pressure of the atmosphere (mbar).

The calculation for evaporation, the total moisture loss from the entire canopy, (E) is given by equation (19):

$$E = E_p W \tag{19}$$

Where:

W is actual fraction of wet area to total canopy surface area, and

E_p is potential latent heat flux density (evaporation) (cm min^{-1}), and

W is actual fraction of canopy wet surface area obtained from formulas (20) and (21):

$$W = W_{ind}\, x \tag{20}$$

$$W_{max} = pW_f + (1 - p)W_d \tag{21}$$

Where

p is fraction of wettable leaves to total leaves in canopy,

W_f is average fraction of wet area to total area of wettable leaves, and

W_d is average fraction of wet area to total area of non-wettable leaves.

A W_{max} value of 0.5 was used in this study. This value is based on the observation that immature grape leaves are wettable and that mature leaves are non-wettable and a canopy is assumed to be comprised of 50% mature and 50% immature leaves (Magarey, 2006).

The transfer coefficient (h) in equation (18) originated from the basic transfer coefficient (Bird et al., 1960) which was later modified, equation (22), to take into account wind speed and an object's shape and size:

$$h = cU_c^{0.5} \tag{22}$$

Where:

U_c is canopy wind speed (cm min^{-1}), and

c is shape scale constant of an object ($cm^{0.5}$ $min^{-0.5}$).

In SWEB the shape scale constant (c) becomes a variable. When the leaf surface becomes wetter, the moisture transfer behaves as if the water forms a film. In contrast, when the surface becomes drier, it behaves as if the water is a droplet. To take these shape changes into account, a shape scale variable is calculated by means of equation (23):

$$c = Wc_f + (1 + W)c_d \tag{23}$$

Where:

c_f is shape scale constant for film ($cm^{0.5}$ $min^{-0.5}$), and

c_d is shape scale constant for drops ($cm^{0.5}$ $min^{-0.5}$).

The height and density of the canopy plays an important role in determining canopy wind speed (U_c). Canopy wind speed is calculated from a logarithmic wind profile equation and an analytical wind speed profile as shown in equation (24):

$$U_c = U_z \left(\frac{\ln[(Z_c - D_Z)/Z_0]}{\ln[(Z - D_Z)/Z_0]}\right) \left[1 + \alpha \left(\frac{1 - Z_c}{Z}\right)\right]^{-2} \tag{24}$$

Where:

U_c is wind speed at average height of the canopy (cm min^{-1}),

Z_c is the height of canopy (cm),

U_z is wind speed (cm min_{-1}),

D_z is zero plane displacement (= 2/3 Z_c) (cm),

Z_0 is roughness length (=1/10Z) (cm),

Z = reference height (cm), and

α = wind speed profile.

A value of 1.3 is selected for α because the wind speed profile is dependent upon whether or not the wind blowing parallel or perpendicular to the direction of vine rows (Heilman et al., 1994).

The water storage function in SWEB enables the simulation of both dew and rain events. Unlike other physical models, SWEB takes rainfall into account when classifying leaf wetness. SWEB has the ability to adapt to physical characteristics of particular plants through the adjustment of four crop parameters: leaf area index (LAI), crop height, maximum fraction of canopy allowed as wet surface area (W_{max}), and maximum water storage per unit area (C_1). SWEB's adaptability makes the model easily scalable and results in low spatial variability. This model was reported to give under one hour of error (such a degree of error is considered to be low for LWD estimation models) in two out of three sites tested and these two sites were located in different hemispheres. SWEB was validated with visual observations of grape canopies in the three sites. The mean absolute error (MAE) of the model varied between sites from 0.7 to 1.5 hours (Magarey, 2006).

SWEB relies on input variables that are calculated from standard weather parameters. Some of the derived variables are calculated employing constant values that are assumed, such as net radiation, based on particular conditions. The use of these assumed constants can lead to uncertainty in the model Uncertainties of 5% or more are expected as a result of using these estimated variables (Oke, 1988). This uncertainty may lead to inaccuracy in the leaf surface wetness classification. Despite these limitations, SWEB was chosen as one of the physical models in this research because it is one of the most recently developed leaf wetness models, it has been validated in the southern hemisphere, and has been shown to have low spatial variability (Magarey, 2006).

Hybrid Models

The two hybrid models selected for this research are Fuzzy Logic System (FLS) and an Artificial Neural Network with a backpropagation architecture (ANN). Both models were categorised as hybrid models because they use both meteorological variables and energy balance principles as inputs to estimate LWD.

Both FLS and ANN were chosen because they will provide a good baseline for comparison with the novel implementation and adaptation of an Adaptive Neuro-Fuzzy Inference System (ANFIS) for LWD estimation which is proposed as part of this research. ANFIS is itself a hybrid model and combines ideas from Fuzzy logic and ANN (for further details on ANFIS please refer to Section 2.2.2).

In Boolean logic a given statement can only be either true (1) or false (0), it cannot be anything in between (this is known as the law of the excluded middle). This means that logical reasoning is exact, there is no uncertainty. On the other hand, fuzzy logic truth is presented in relative terms (non-exact) but may also include 0 and 1 as extreme cases of truth.

Truth in fuzzy logic can be presented in linguistic terms, such as “low”, “moderate”, or “high” to symbolise physical processes. These linguistic terms form a fuzzy set called a *membership function* (μ). A membership function assigns a membership degree from 0 to 1 to a given variable. The set of input values that gives valid degrees of membership function (> 0) is called a *domain*.

In a fuzzy logic approach to measuring leaf wetness, relationships between meteorological variables and wetness occurrence can be expressed in a form of rules. These rules can be used to define specific conditions in which wetness is likely to be present or absent. Once a set of rules has been created, the membership functions of each variable can be defined heuristically.

Fuzzy logic was used as a model to estimate leaf wetness by Kim et al. (2004), they called their method the Fuzzy Logic System (FLS). The model is essentially an empirical model that complies with energy balance principles. The FLS classifies leaf wetness utilising meteorological variables, membership functions for each of the variables, and rules to determine whether wetness is present or absent on leaf surface. Three variables were selected by Kim et al. in the study; vapour pressure deficit (VPD), wind speed, and net radiation (pR_n). The inputs were selected because of their high correlation with to leaf wetness occurrences. Wind speed is a direct variable that is available in most weather stations, pR_n may be available but can be estimated when it is not, and VPD is a derived variable that can be calculated using equation (25):

$$VPD = e_s - e_a \tag{25}$$

$$e_s = e_0 exp\left[\frac{1}{R}\left(\frac{1}{273} - \frac{1}{T+273}\right)\right] \tag{26}$$

$$e_a = RH\ e_s/100 \tag{27}$$

Where:

e_s is the saturated vapour pressure of the air (kPa),

e_a is the partial pressure of water vapour in the air (kPa),

e_0 is the vapour pressure constant (=0.611) (kPa),

R is the gas constant for water vapour (=461 J kg $_{-1}$ K$_{-1}$), and

T is the air temperature (°C), and RH is the percentage relative humidity (%).

When net radiation (pR_n) is not available from weather station, an estimation formula (28) can be used:

$$pR_n = \varepsilon_a \sigma (T + 273)^4 - [\varepsilon_a \sigma (T_{dew} + 273)^4 + (1 - \varepsilon_s)\varepsilon_a \sigma (T + 273)^4] \tag{28}$$

$$\varepsilon_a = 1 - 0.261 \exp(-7.77x10^{-4}T^2) \tag{29}$$

Where:

ε_a is the emissivity of the atmosphere,

ε_s is the emissivity of the sensor surface (=0.98),

σ is the Stefan Boltzmann constant (=5.67 x 10^{-8} W m^{-2} K^{-4}), and

T_{dew} is the dew point temperature (°C).

Surface temperature can be assumed to be the same as the dew point temperature when estimating outgoing radiation in this model. The value for ε_s was assumed to be 0.98 (Idso and Jackson, 1969).

After the inputs were obtained, the values were fuzzified into linguistic terms. The process of translating crisp values to linguistic terms with membership functions is called *fuzzification*. The fuzzified input was then compiled into rules and weights were assigned to each rule. For example "*if* VPD is high *then* wetness is absent" is a rule the weight of which can be assigned based on the opinion of domain expert, in the case of VPD a weight of 0.55 was assigned. In most cases there will be multiple statements, therefore logical operators are used to obtain a single degree membership. In the FLS developed by Kim et al. (2004), they used two logical operators, AND and OR to create a single degree membership function (equations (30) and (31)):

$$\mu A(x)\ AND\ \mu B(y) = \mu A(x)\mu B(y) \tag{30}$$

$$NOT\ \mu A(x) = 1 - \mu A(x) \tag{31}$$

Where $\mu A(x)$ is membership function of fuzzy set A for $x \in [-\infty, \infty]$.

Given that the degree of the antecedent has a value between 0 and 1 the result of applying the fuzzy rule was obtained in the fuzzy set form using the implication operator IF-then (32):

$$IF - Then\ \big(a, \mu C(x)\big) = \min\big(a, \mu C(x)\big) \tag{32}$$

Where a is the degree of antecedent, and

$\mu C(x)$ is a membership function of consequence over $x \in [0, 1]$.

According to Kim et al. (2004), LWD estimation error using the FLS model was less than one hour per day and this is comparable to sensor measurements in terms of overall accuracy. FLS tended to have a smaller error rate and a superior accuracy when compared to a CART model in terms of mean average error and hourly classification error across different sites (Kim et al., 2004).

The main challenge in implementing the FLS model is in determining the rules and the membership functions for each variable; the user needs to understand the characteristics of each variable. If they do not, often an arbitrary number is employed introducing a source of uncertainty to the model. This uncertainty may cause a suboptimal model to be developed for a particular dataset. However, for leaf wetness a set of membership functions and rules has already been established by domain experts (Kim et al., 2004) and therefore the likelihood of error in determining the rules and membership functions is minimal.

The final approach selected for comparative analysis was an ANN. There are many variants of ANN's reported in the literature and descriptions of these different forms of ANN's can be found in Bishop's seminal paper (1995). ANNs are inspired by biological neural networks and are used to estimate or approximate functions that depend on a large number of inputs, are generally unknown and often indeterminate. ANN's have been used successfully to estimate and predict other environmental variables using meteorological variables for example; particulate matter (McKendry, 2002), and rainfall (Cleofé, 2005). This makes ANNs a viable candidate for the estimation of LWD.

The ANN discussed here is one designed by Fancl and Panigrahi (1997) for the prediction of dew duration and is the most relevant. In later work, Chtioui, Panigrahi and Francl (1999) investigated the use of a generalised regression neural network (GRNN). In this work they compared the performance of GRNN with that of a standard Multiple Linear Regression (MLR) approach. They found that GRNN outperformed MLR but that GRNN required more computational time. Because most LWD methods reported in the literature are usually compared with NHRH ≥ 90% and because no other researchers have reported comparing MLR with other accepted dew duration (or LWD) models it was decided not to include the more recent GRNN method in this research.

Fancl and Panigrahi's ANN for leaf wetness prediction was a feed-forward, multilayer perceptron with back-propagated error Neural Network (NN). They used wheat flag leaves from a North Dakota (USA) plantation as their model. The variables included in their model were air temperature, relative humidity, wind speed, wind direction, solar radiation and rainfall (precipitation). These variables are normally measured by standard weather stations so data is readily available for input into the network. Fancl and Panigrahi also developed a second NN model which used as input these meteorological variables as well as 'remote wetness' measured by a leaf wetness sensor that was located above the crop canopy. They found that the ANN models predicted wetness with 82-96% accuracy and that all models improved when remote wetness was included. They also performed a sensitivity analysis and found that both remote wetness and relative humidity were highly significant inputs to the model. The best performing ANN model performed favourably (1 hour/day error) with previously developed models for the prediction of dew duration CART/SLD and other physical models. The authors also note that this ANN approach had the added advantage of predicting leaf wetness from both dew and rain which means that it is a particularly useful method for determining the LWD for input into disease forecasting models.

2.2.2 ANFIS Overview

ANFIS is essentially a hybrid model which combines two intelligent models, Artificial Neural Network (ANN) and Fuzzy Inference System (FIS) (Alves et al., 2011). ANN maps an input space to an output space through a collection of neurons that are interconnected. ANN have the ability to learn and are data driven, and thus benefit from being able to find and learn from patterns in the information presented to the network. FIS is based on fuzzy logic, If-THEN fuzzy rules, and fuzzy reasoning. These features allows the model to make inferences using the rules and known facts to obtain reasonable decision (Jang, 1993). The combination of both intelligent models to form ANFIS incorporates the individual advantages of the models and should result in higher leaf wetness estimation accuracy than previously developed models.

ANFIS is a class of adaptive networks that are functionally equivalent to fuzzy inference systems. An adaptive network is a network structure that consists of nodes and directional links through which the nodes are connected. ANFIS architecture consists of layers that involve nodes that are adaptive. The output of an adaptive node depends on the parameter(s) that are related to these nodes, and the learning rule stipulates how these parameters should be changed to minimise predicted error.

To understand the operation and structure of ANFIS an example of ANFIS is presented, which consists of a five layer feed-forward neural network with backpropagation and a Sugeno-type FIS (Sugeno, 1985). The example model is a 2-input–1-output ANFIS model. In general, an n-input-1 output ANFIS model is an n+1 dimensional input spaces. Therefore the example is a three dimensional input-output space.

Figure 5 shows the three dimensional ANFIS structure, and equations (33) and (34) provide the model's rule, where the IF part is referred to as the antecedent, and the THEN part as the consequent:

$$\text{Rule 1: If } x \text{ is } A_1 \text{ and } y \text{ is } B_1\text{, then } f_1 = p_1x + q_1y + r_1 \tag{33}$$

$$\text{Rule 2: If } x \text{ is } A_2 \text{ and } y \text{ is } B_2\text{, then } f_2 = p_2x + q_2y + r_2 \tag{34}$$

Where:

x and y are the inputs and

A and B are the fuzzy sets in the antecedent.

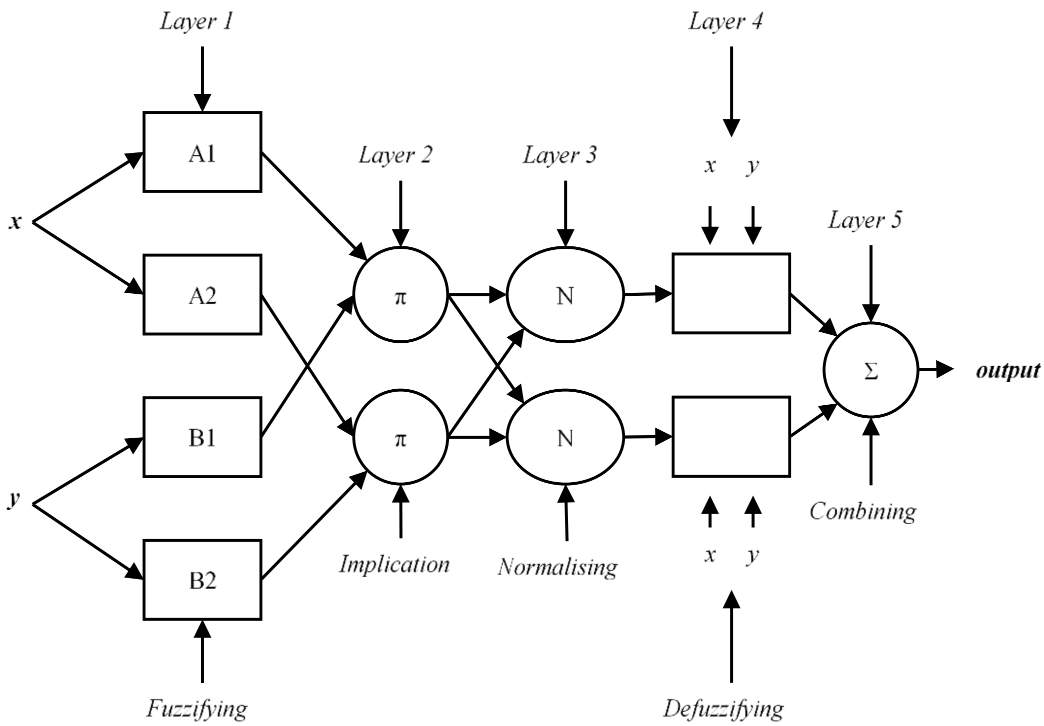

Figure 5. The architecture of 2 inputs ANFIS network (Jang, 1993 p4)

To model a system with ANFIS, representative data from the target system must be presented as input to the model. The crisp value of the data is entered into the system which then corresponds to the first layer, *fuzzifying* layer, to be fuzzified.

In Layer 1, the *fuzzifying* layer; the inputs are fuzzified or translated into linguistic labels, and a membership grade generated for each label.

Every node i in this layer node is a square node with a node function (35):

$$O_i^1 = \mu A_i(x) \tag{35}$$

Where:

x is the input to node i, and

A_i is the linguistic label associated with node i.

O_i^1 the membership function of A_i, specifies the degree to which x satisfies the quantifier A_i.

Usually the membership function is chosen to be a bell-shaped function with a maximum value of 1 and minimum value of 0, as presented in equation (36):

$$\mu A_i(x) = \frac{1}{1+\left[\left(\frac{x-c_i}{a_i}\right)^2\right]^{b_i}} \tag{36}$$

Where:

$\{a_i, b_i, c_i\}$ is the parameter set.

As the values of these parameters change, the bell curve membership function may vary accordingly. The parameters to the membership function are referred to as the *premise parameters*.

In layer 2, the *implication layer*, every node multiplies the incoming signals and returns the product, according to equation (37).

$$w_i = \mu A_i(x) \times \mu B_i(y), 1 = 1, 2 \tag{37}$$

This layer contains the fuzzy rules, and each node's firing strength for a given rule. There are A^b number of rules in ANFIS, where A is the number of membership functions of every input, and b is the number of inputs. The number of nodes in this layer is reliant on the number of rules, and is fixed.

In layer 3, the *normalising layer*, the ratio of each rule's firing strength is scaled according to the total of all the rules' firing strength, equation (38).

$$\overline{w}_i \frac{w_i}{w_1+w_2}, i = 1, 2 \tag{38}$$

In layer 4, the *defuzzifying layer,* the output of each node is weighted and computed towards the overall output. The output of this layer is a crisp value.

$$O_i^4 = \overline{w}_i f_i = \overline{w}_i(p_i x + q_i y + r_i) \tag{39}$$

Where $\overline{w}_i$ is the output of layer 3 and $\{p_i, q_i, r_i\}$ is the parameter set.

The parameters in this layer are referred to as *consequent parameters.*

Layer 5 is called the *combining layer* and in this layer the overall output from the previous layer is computed as the summation of every rule's contribution.

$$O_1^5 = overall\ output = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{40}$$

The learning algorithm proposed for ANFIS is a hybrid learning algorithm that minimises error between ANFIS model and target system. (Pabreja, 2011). The proposed hybrid learning algorithm consists of least squares estimate and gradient descent method. Once input-output data is presented to ANFIS model, this data is propagated forward from layer 1 to 3, and the least squares estimate is employed to update the consequent parameters. Then an error value is calculated and propagated backwards and a gradient descent is used to update the non-linear or premise parameters (Jang, 1993).

ANFIS has shown to be successful as a model for estimating a number of different plant diseases and environment factors but has not to date been evaluated as a means of predicting or estimating leaf wetness duration. Two of the most relevant applications of ANFIS are detailed below.

The ANFIS model has been shown to be capable of describing the severity of soybean rust, under the effects of leaf wetness, temperature, and days after fungi inoculation as input variables (Alves et al., 2011). Alves et al. developed a four dimensional ANFIS model input-output space with three inputs, 27 rules, and nine Gaussian membership functions. Their model was developed to characterise the relationship between leaf wetness, temperature and days after fungi inoculation in order to estimate the severity of soybean rust. The model was trained with three and 3000 epochs. They found that ANFIS was able to describe higher soybean rust severity occurrences between 20°C and 25°C, when leaf wetness persisted for over above six hours, with higher values 10 hours, and 15 days after fungi inoculation. A higher overall accuracy (81.6%) was achieved when the system was trained with 3000 epochs than training with only three epochs (72.5%) for all three case study sites. However, a maximum accuracy was reached with only three epochs in one of the sites (89.9%) (Alves et al., 2011).

ANFIS has also been used for forecasting solar radiation data from weather parameters. Solar radiation is used in many agricultural models, such as leaf wetness prediction, and also in renewable energy applications, such as the sizing of photovoltaic systems. Because solar radiation is not always available in weather stations, particularly in remote areas, Mellit and colleagues (2007) introduced an ANFIS approach to forecasting daily solar radiation by using ambient temperature and sunshine duration as inputs. In the study, nine years of solar radiation data were used to train ANFIS with the aim of developing a network that was ready to accept and could handle a number of rare/unusual or infrequent cases. Once satisfactory input-output mapping had been achieved, the trained ANFIS model was frozen and exposed to a test dataset for validation. This approach resulted in a high accuracy, of less than 1% mean relative error between the actual data and predicted values, with 98% of the correlation coefficient for the validation dataset. The trained ANFIS model which reached optimum solar radiation forecasting error included one hidden layer with 15 neurons (Mellit et al., 2007).

This chapter has outlined the two principle means of determining LWD, direct sensor measurement and modelling. Because of the wide availability of different sensors a comparative analysis of these instruments is useful in order to establish which commercially available instruments are the most accurate and with a longer term view of establishing a practical sensor calibration and installation standard. Additionally, because of the limitations of these instruments leaf LWD models were also discussed and six LWD models selected for a comparative analysis and of their suitability for use in modelling the LWD in New Zealand vineyards. The adaptation of ANFIS as a potential and novel model for predicting leaf wetness duration is proposed in this research.

CHAPTER 3

EXPERIMENT DESIGN

3.1 Research Methodology

This research is undertaken from a largely positivist perspective. This type of research is characterised by "*an emphasis on internal validity through tight experimental control and quantitative methods*" (Fitzgerald & Howcroft, 1998, p10).

The work undertaken involves two different methods.

For the leaf wetness sensor evaluation a systematic experimental approach will be employed in which empirical measurements and scientific observations are made. Sensor variability is controlled both by in laboratory calibrations and infield calibration verification steps. Sensor location and installation variables are controlled by employing a set of precise standards for installation. Experimental procedures are put in place, as detailed in section 3.2.3, to remove any potential error or bias in the measurements.

For the comparative evaluation of leaf wetness models a quantitative approach is taken in which the output of each of the models are compared with the real world measurements (aka sensor measurements) of the leaf wetness phenomena (Figure 6). With this type of methodological approach it is important to acknowledge any constraints, assumptions, uncertainties and potential sources of error which might be introduced to the model. Each leaf wetness model is then compared with the other using statistical techniques in order to establish the best performing model.

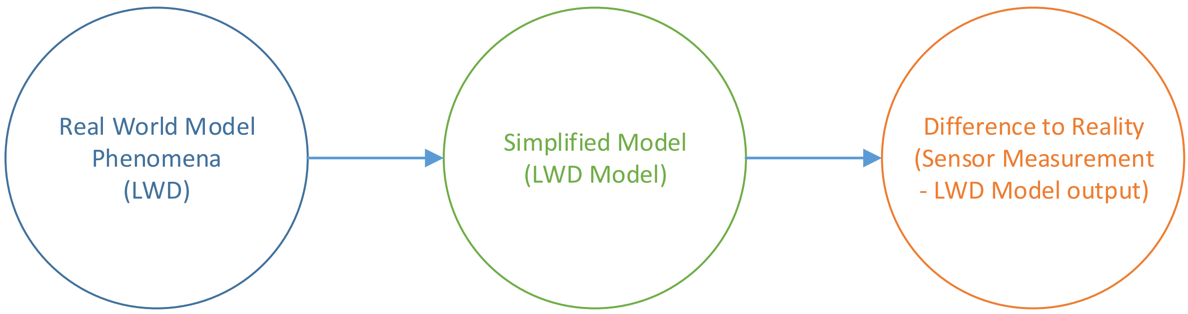

Figure 6. Modelling method

3.2 Materials and Methods

This section details the materials and methods of two separate experiments conducted in this study; leaf wetness sensors comparison and LWD model comparison. Since both have different requirements and objectives, both will be discussed in different sections. The scope covers what is required to do the experiments, experimental design, and the method of evaluation.

3.2.3 Leaf wetness sensor comparison experiments

One of the objectives of this experiment is to evaluate the performance of various commonly used commercially available leaf wetness sensors. To do so, several leaf wetness sensors were obtained, including three from well-known manufacturers: Decagon Devices, Inc. (DD), model 237 sensor from Campbell Scientific, Inc. (CS), Pessl Instruments GmbH (PI), and one from a generic manufacturer Hobby Boards (HB).

Both CS and HB are flat plate resistance based sensors with dimension of 102mm x 58mm x 58mm with 0.5mm electrode gap and 40mm x 20mm x 20mm with 0.25mm electrode gap, respectively. These sensors work based on the measure of impedance change on the sensor's surface. Conductive electrode gaps on the sensor's surface lowers the impedance when exposed to water. The CS and HB resistance based sensors are supplied unpainted. In this experiment they were used in both unpainted and painted form.

The DD sensor is a flat plate dielectric based sensor with a dimension of 112mm x 58mm x 0.75mm. The DD sensor is made from fibreglass and comes painted with white latex paint. It is also the only sensor which has a shape that is similar to a real leaf.

The PI sensor is a filter paper conductivity based sensor with acrylic housing in the dimension of 127mm x 254mm x 508mm. PI sensor works by measuring the impedance change based on two metal bars connection that is bridged by a single absorptive filter paper.

Experimental setup

In order to conduct the experiments the sensors must be installed and calibrated. In order to ensure that this process is consistent across the sensors a set of standards, discussed in the chapter 2, will be applied. The core steps are firstly a laboratory based sensor calibration step, then the sensors are installed in the field, and finally sensor calibration is verified in the field.

In lab sensor calibration

All sensors in this research were calibrated using the same method. Each sensor was connected to a digital multimeter (Digitech QM1539) and a baseline reading taken. The sensor was then exposed to a 1mm diameter droplet of water. The voltage change was measured, the sensor surface was wiped clean and then left to dry. The sensor was considered to be dry when the multimeter gave a reading that was equal to the baseline voltage reading. This process was repeated three times for each sensor and the average value was taken to be the sensors threshold.

In Field Sensor Installation

The factors which need to be considered when installing the leaf wetness sensors in the field are deployment angle, compass orientation, and height (Lau et al, 2000; Sentelhas et al., 2006).

Research has shown that the deployment angle should be between 30 and 45 degrees to horizontal. A protractor was used to measure the angle when the sensor was installed. Because the study area is in the southern hemisphere the sensor was oriented due south in order to minimise solar radiation interception (Madeira et al., 2002). The sensor was located 1 meter above the turf grass. This height is within the tolerances recommended by Lau and colleagues (2000).

One of each type of flat plate resistance based sensor (HB and CS) was coated with two layers of off-white water based latex paint (Lau et al., 2000). The paint used was DULUX *prepcoat*® water based acrylic primer in white (colour: 630-01851). Off white is recommended because it gives a similar evaporation rate to that of real leaves (Gillespie and Kidd, 1978; Sentelhas et al., 2004). After the paint is applied the sensor was treated with heat at 70 degrees Celsius for 24 hours as recommended by Sentelhas et al. (2006). This heat treatment is required to reduce the hygroscopic materials from the latex paint in order to allow water to be detected on the sensor surface.

The six sensors were then connected to a custom weather station node which was in turn connected to a wireless base station. This node was produced as part of this research and used because commercially available nodes do not have the capabilities to integrate non-proprietary sensors and only support the manufacturer’s sensors. The weather station node was built on a libelium Wasp Mote which was attached to a libelium Agriculture Board which extended the Wasp Mote which allowed for the installation of all six sensors in a single node. The connection of the node to the base station used a ZigBee network protocol (XBeePro™ Series 2).

An Intel® NUC (NUC5i3RYH) stand-alone computer was used as the base station, connected to a ZigBee network receiver to retrieve information from the node. Base station was tasked to receive the logged data from the node and store it. The base station is also connected to the internet via Wi-Fi connection to enable real-time monitoring and remote maintenance. Sensor node management software (SeNoMa) was installed in the base station to record the data from nodes (Ghobakhlou, 2014). Variable names and calibration was modified to fit the 6 leaf wetness sensors included in the experiment.

The overall configuration of the sensor network is shown in Figure 7.

Post installation calibration

Verification of the calibration of the sensors was undertaken in the field after installation. This verification process was carried out on a rainy day and involved visual observation of each sensor's response to exposure to moisture. The check verified that the sensor was reading as wet when it was wet and reading dry when it had not been exposed to rain. If a sensor is not reading as expected it must be uninstalled and returned to the lab for recalibration, then reinstalled in the field and a post installation calibration undertaken again.

Experimental Design: Sensor performance evaluation

To observe the sensor's performance, visual observation was undertaken at two different times of the day. Firstly, the time from an hour before sunrise until the sensors are completely dry to measure the dew dry-off period, and secondly an hour before sunset until the sensors are wet to observe dew onset (Lau et al., 2000). Both must be done on a day which has little or no-rain. Because the identification of a dry sensor is done by eye, and therefore not entirely accurate the dew dry-off and dew onset measurements were undertaken 8 times to minimise any potential error. The visual observation was undertaken every five minutes increments, the interval for the sensor readings, in order to match the visual observation with the sensor's reading. A visual observation of the leaf surface, on the grape vine canopy, was also recorded with the same time interval.

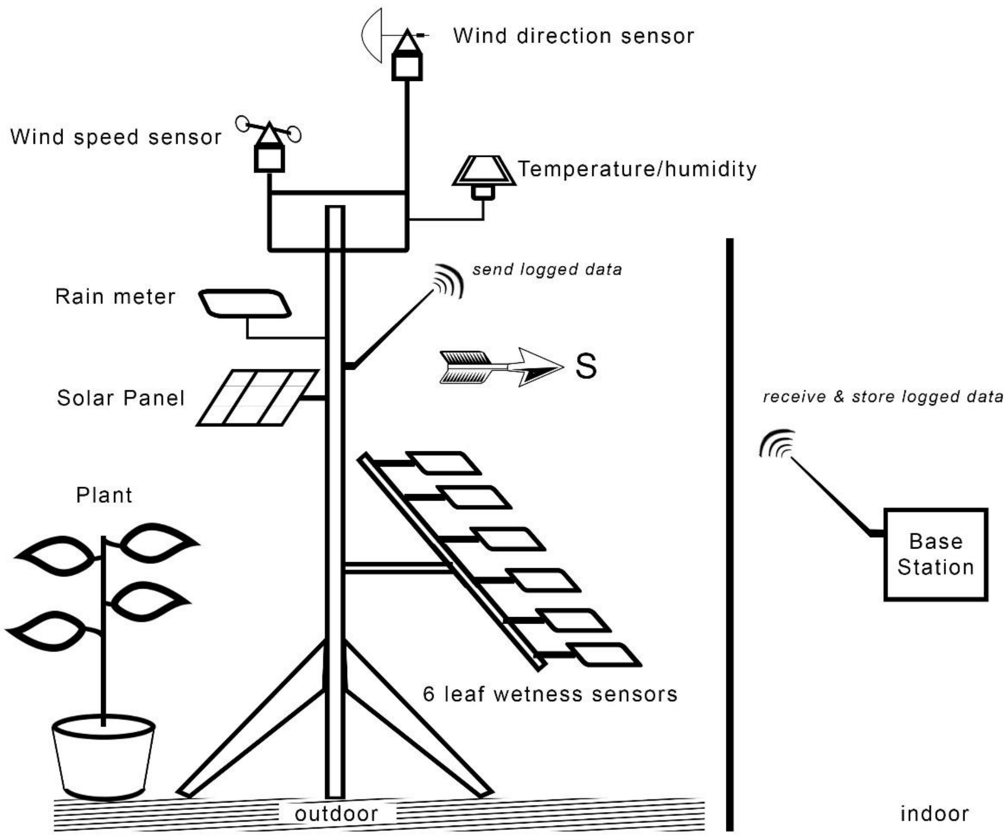

Figure 7. Leaf wetness sensor experiment setup

There are two main objectives of the visual observation, the first is to compare the wetness between the sensor and actual grapevine leaf surfaces, and the second is to compare the sensor surfaces wetness and the sensor's logged reading. The first objective's purpose is to observe which of the sensors evaluated best represents real leaves in terms of the wetness of the actual leaf. This first evaluation considers the physical appearance, size, composition, and shape of the vine leaf and therefore evaluates how well the sensor represents a real leaf. The second objective is to observe the level of agreement between a sensor's physical wetness and its own reading.

Another performance evaluation test is a measure of the sensor's response delay. This test was conducted by comparing the sensors readings when dew was present. All of the sensors were placed next to each other, it is assumed that by placing the sensors close to each other the timing for dew formation should be similar. The first sensor to log a wet response was used as the benchmark response time (t_0). The other sensors response delays were then calculated as the difference between the sensor recording a response and the response time (t_s) according to equation (41). This test was undertaken for both the dew onset and the dew dry-off.

Evaluation method

A quantitative approach is used to evaluate and compare sensor performance. For each sensor, the reading is compared with a visual observation and the mean error (ME) and mean average error (MAE) were calculated. ME was computed by averaging the differences between measured and estimated LWD for a 24 hour period. ME determines the tendency of a sensor to overestimate or underestimate LWD. MAE was calculated by averaging the absolute values of hourly errors. MAE determines the overall accuracy of the sensor.

To evaluate the LWS response delay time, the response times of wetness onset and wetness dry-off period were calculated. The first sensor's onset time t_o denotes the water presence and the response time delay for this sensor is set to zero. The onset response for all the other sensors are calculated according the equation 41.

$$\Delta_o = t_s\text{-}t_o \tag{41}$$

Where:

t_0 is the first sensor's onset time,

t_s is sensor onset time,

Δ_o is onset response time.

The first sensor's dry-off time (t_d) denote the water absence and the response time delay for this sensor set to be zero. The dry-off response for all the other sensors are calculated according the following equation 2.

$$\Delta_d = t_s\text{-}t_o \tag{42}$$

Where:

t_d is the first sensor's dry-off time,

t_s is sensor dry-off time,

Δ_d is dry-off response time.

3.2.4. Leaf wetness model comparison

This research required weather data, see Appendix A for a full list of the required variables and maps of the stations for the two datasets used (see Chapter 4.2 for further details) for the development of the LWD models. The case study area included a range of grape growing regions in New Zealand from Crownthorpe to Pukekohe in the north (Figure 8). These stations are located at least 150km apart to ensure that each station has different micrometeorological characteristics in order to be able to evaluate the spatial variability of each of the LWD models. Additionally, the stations chosen were selected due to the comprehensiveness of the data. Many of the sixteen weather stations, which log both meteorological variables and leaf wetness, lack a continuous set of data for the time period(s) investigated in this research.

Figure 8. Map of all stations used in the experiment

Data Acquisition

The data used in this experiment was obtained from Plant & Food Research New Zealand (P&F) in the form of an excel spreadsheet. The variables in the data consisted of hourly measurements including: air temperature, relative humidity, rainfall, and leaf wetness. Three stations were chosen for this experiment they were: Pukekohe Research Station (PKE), Hexton (HXT), and Cornwall (CRN), in the selected period of time in the year 2012 with months as follows: January, April, July, and October. The selected months were chosen in order to have one month from each season in New Zealand, starting in summer 2012. The year, 2012 was chosen as the most recent year with the most complete data from the available P&F. The data was chosen to include all the seasons to make neural network approaches to be able to generalise in various conditions.

Most models require one crucial variable which is missing from the P&F data, namely wind speed. In this study, wind speed was obtained from adjacent weather stations data provided by the National Institute of Water and Atmospheric Research (NIWA). The data was downloaded from the CliFlo website (cliflo.niwa.co.nz) as a comma delimited file. The NIWA stations were chosen based on their proximity to the P&F stations in order to provide the closest possible 'reference' wind speed measurements. Table 1 shows the distance between P&F and NIWA stations. The wind speed measured by NIWA is provided in meters per second (ms^{-1}), which matches the requirements of the leaf wetness models used in this research, therefore no data conversion is required

Table 1. Plant & Food Research and Adjacent NIWA stations

Plant & Food station name	**Coordinates**	**NIWA station name**	**Coordinates**	**Distance (km)**	**Area**
Pukekohe Research Station (PKE)	-37.20° N 174.86° W	Pukekohe Ews	-37.20°N 174.86° W	0.8	Auckland
Hexton (HXT)	-38.62° N 177.97° W	Gisborne Ews	-38.62° N 177.92° W	4.3	Gisborne
Crownthorpe (CRN)	-39.57° N 176.55° W	Napier Aero Aws	-39.46° N 176.85° W	29.1	Hawke's Bay

Data preparation

The P&F leaf wetness is recorded in a range between 0 (driest) and 100 (wettest) but LWD models are only interested in if it is wet or dry; therefore the P&F leaf wetness data was classified into either wet or dry hours using a threshold of 50 if the reading was >50 then it was classified as wet. For any day where the day was missing more than 2 hourly readings the day's data was removed.

Model evaluation method

For each model ME, MAE and Estimation Accuracy (EA) were calculated. ME and MAE were described in leaf wetness sensors evaluation method. The EA represents the degree of closeness of the estimated and measured LWD as a percentage, calculated according to equation 43.

$$\mathrm{EA} = \left(1 - \frac{\Sigma\,|\mathrm{Actual} - \mathrm{Estimated}|}{N}\right) \times 100 \tag{43}$$

Where N is the total number of data points.

CHAPTER 4

RESULTS AND DISCUSSIONS

In this chapter two sets of experimental results are presented. The first section presents a leaf wetness sensor comparison, and the second an implementation and evaluation of leaf wetness modelling. To complement the second set of experiments, this chapter also includes an exploration of the data, used for leaf wetness modelling, using statistics in order to understand the relationships between the explanatory variables and to explore the data distribution.

4.1 Leaf Wetness Sensors Comparison

Visual observation of four different leaf wetness sensors, with two of the sensors treated with paint and heat, totalling six sensors was undertaken by comparing real leaves with sensor readings (see photo's in Figure 9 which show the experimental and deployment of the sensors for this research).

Figure 9. Experimental deployment of the sensors for this research

Visual observation of sensor's surface and real leaves' surface was done and the result is presented in Figure 10. The purpose of the result presented in Figure 10a) is to examine whether the LWS physical attributes such as size, shape, and thickness are affecting the amount of time free water stay on the sensor's surface. Figure 10a) shows the average delay in minutes observed for each of the tested sensors. Three real leaves on a living plant that are leaning on the same angle as the sensor are flagged as observed.

According to Figure 10 a) among the sensors, the DD sensor (for sensor name references see section 3.2) has the shortest delay time. The DD sensor has a mean delay time of 7.5 minutes measured from the point that the real leaves are dry. The PI sensor is the worst performer in this experiment, with an average 60 minutes delay.

The four remaining sensors dried quicker than the real leaves. The behaviour of the HB and CS sensor's painted sensors was found to be quite different. Painting gave an improvement for the CS sensor but actually resulted in less accuracy in the HB sensor. This lower accuracy of the HB sensor may be because it holds less water on the surface due to the size of the sensor. Sensor size influences how much water a sensor can possibly hold. The other sensors evaluated have a size that is closer to the size of a real leaf.

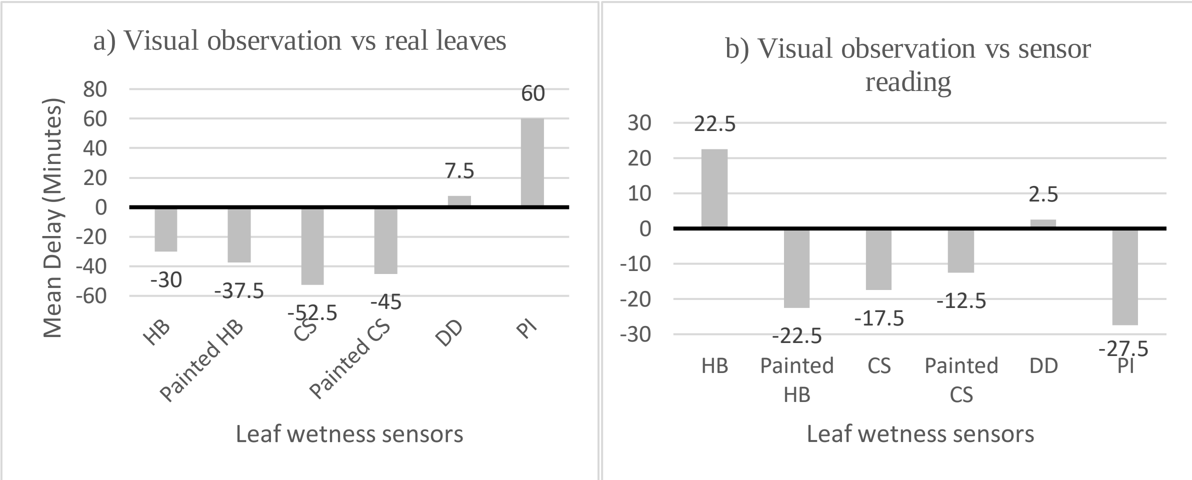

Figure 10. Two days dry-off visual observation on 6 LWS, a) against real leaves, and b) against sensor measurements

A comparison of visual observation with sensor reading was also undertaken to examine the delay between when actual water droplets dry on a sensor's surface and the sensor reading log time. Figure 10b) presents the result of the visual observation, the closer the number to zero the shorter the sensor delay. The DD sensor is the best performer, in this test, by an average delay of 2.5 minutes between the water having dried on its surface and the time recorded in its log.

The painted and non-painted HB sensor, in this test, gave an inverse result, the painted HB sensor still detects wetness state on average 22.5 minutes after the leaves dried. This delay in recording dry-off tends to happens if the sensor is too sensitive to small water droplets. In general sensors that are painted have higher hygroscopic materials which mean that they detect smaller water droplets than unpainted sensors. The unpainted HB sensor record log shows a leaf wet state up to on average 22.5 minutes before it was dry. This suggests that the unpainted sensor is less sensitive than the painted one.

The same was observed for the CS sensor, although not as significant as HB sensor, the result for CS sensor indicates that the painted sensor is more sensitive, it records wetness with a five minute average delay. However, Figure 10b) shows both painted and unpainted CS sensor gave false readings of wetness up to on average 17.5 minutes after the sensor was actually dry.

The PI sensor was the worst performer in both tests shown in Figure 10b). It has an average 60 minutes wetness detection delay compared to real leaves. The PI sensor was also found to record wetness with average 27.5 minutes delay after the sensors are visually dry.

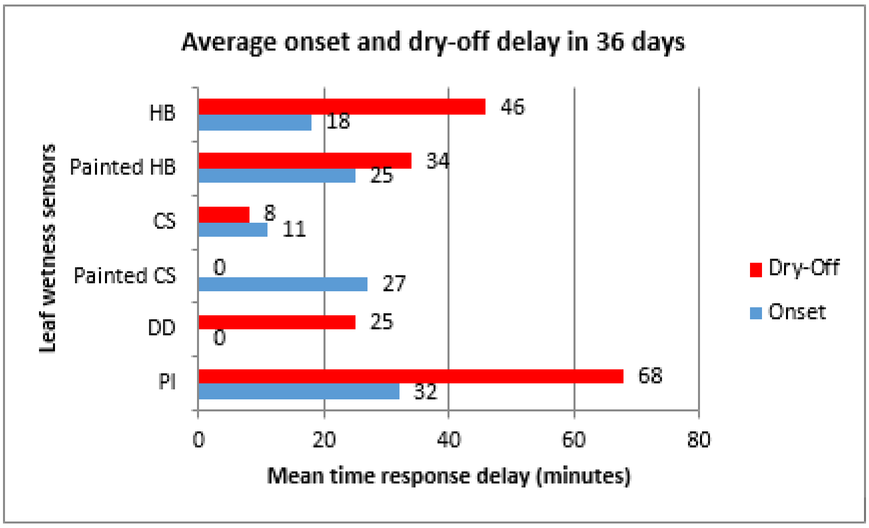

Figure 11. Average onset and dry-off delays of 6 sensors observed in 36 days

Figure 11 shows the average onset and dry-off response time delay for six leaf wetness sensors. The figure shows a result from 36 days of observation on days that either rain, dew, or both occurred. This test undertaken to determine which sensor reading responds the quickest when water droplets fall on its surface (onset) and when the sensor' surface is dry (dry-off).

The DD sensor has always been the initial onset responder in this experiment, thus the "0" value in average onset. The rest of the sensors followed behind the DD sensor in onset response time. The CS and HB sensors were the second and third to record onset, at 11 and 18 minutes average delay respectively. The painted version of the CS and HB sensors had longer delays than their unpainted counterparts. The painted HB recorded on average 25 minutes delay while the painted CS sensor was delayed by 27 minutes. The PI sensor was the last to detect onset by an average of 32 minutes response time.

Dry-off period evaluation is represented by the red bars Figure 11, with the painted CS sensor as the initial sensor dry-off to respond hence the "0" value. The unpainted CS sensor was the second with an 8 minute response to drying of the surface. The CS sensor has the largest surface size, compared to the rest of the sensors, which means more surface for sunlight and wind to help evaporate water drops. On the contrary the HB sensor, which has the smallest surface size in this group, recorded a longer delay compared to unpainted

HB sensor, 46 and 34, respectively. The DD sensor on the other hand recorded 25 minutes average delay, putting itself in between CS and HB sensors. The PI sensor recorded the longest average delay, 68 minutes. When this test was performed, a wind speed of between 2 to 9 m/s was recorded at all initial dry-off stages. This may contribute to the quick dry-off responses of the flat plate based sensors. Wind speed doesn't tend to affect the PI sensor to the same degree. This is due to PI sensors being filter based and protected by the enclosure from wind speed.

The results presented above indicate that the DD sensor is more accurate for detecting leaf wetness presence when compared to the other sensors used in this study. The key differences between the DD sensor and the other five sensors tested are that it has a dielectric measurement principle and came from the supplier already fabricated with painted surface that is also rust free. However, the CS sensors, both painted and unpainted, were better for dry-off detection than the DD sensor.

The painted HB sensor underestimates water presence more than unpainted HB sensor, while the painted CS underestimated wetness to a lesser degree than the unpainted CS sensor. Painting the sensor did not show any improvement for the HB sensor, however CS sensor did benefit from sensor painting. The only difference between both sensors is the surface size, the CS sensor is more than twice the size of HB sensor. This suggests that sensor dimensions and surface size needs to be considered prior paint application.

In both visual observations, the PI sensor was the least accurate to detect wetness presence and dry-off. It holds water for more than an hour longer than real leaves, and did not agree with its visual water presence (see Figure 3). PI sensors use filter paper to absorb water and tend to hold wetness longer and this tends to lead to false positives. In practice, this sensor needs close monitoring and maintenance to ensure proper filter placement and it is in good condition.

The DD sensor was found to be the most accurate leaf wetness sensor. It has the closest representation to real leaves in terms of detecting wetness presence. It is also has the physical attributes that are required for detecting leaf wetness better. DD sensors are the most recent leaf wetness sensor developed, and were first announced in 2006 (Decagon Devices, 2006). It was built based previous research on leaf wetness sensors, which points to that fact that a material of fibreboard coated with latex paint using a heat treatment leads to better accuracy (Lau et al., 2000).

To evaluate the mathematical models for estimating leaf wetness, there is a need to use one of the leaf wetness sensor as reference. The sensor data that was used as the reference data for all the models in this thesis was recorded by a CS sensor. CS sensors were available decades before DD sensor was introduced. Past data was used to test and evaluate the leaf wetness estimation models and CS sensors were used to collect that data. CS sensors have been used in industry and research facilities for more than 10 years. Moreover, CS sensors have survived as the sensor of choice prior to the introduction of DD sensors despite the introduction of other leaf wetness sensors. The results from the experiments performed as part of this research shows that CS sensors are an appropriate second choice and also suggest that any future data logging of leaf wetness should make use of DD sensor technology.

To our knowledge this is the first comprehensive leaf wetness comparison study involving commonly used leaf wetness sensors. This study indicates DD sensor is more accurate than other commonly used leaf wetness sensor.

4.2 Data Exploration

This section presents an exploratory data analysis of the variables available in the data used in this research for developing leaf wetness estimation models.

In the second experiment, there are two dataset included:

A) Two months of data from five stations –January – February 2012
B) Four months of data from three stations – January, April, August, October 2012

The primary variables included in these datasets are:

- Temperature
- Relative humidity
- Rainfall
- Wind speed
- Leaf Wetness (CS sensor readings)

The initial dataset A only included two summer months and during this period there was less rainfall but more complete station data was available than for dataset B. Dataset B has more months of data and includes months with more rain but data was available for only three stations.

Dataset B contains data from January, April, August and October; these months are, in a typical year, times at which the vines have not yet lost their leaves. Grapevine annual cycle includes a period when new leaves, shoots, flowers and fruit are produced, and a dormant period when they lose their leaves. In winter (June–August), the vines are bare but at some point in August buds begin to appear. In spring (September–November) buds on the vine swell and turn into a leafy canopy. Bunches of grapes develop in summer (December–February) and are picked in autumn (March–May). During autumn the leaves being to colour and by late autumn the leaves have fallen.

4.2.1 Feature Evaluation

In order to explore which of the explanatory variables have the strongest predictive power a simple Correlation-based Feature Subset Selection (CfsSubsetEval) attribute evaluator method was used (Hall, 1999) which is available in Weka (Hall et al., 2009). This method evaluates the value of potential explanatory variables by measuring the individual predictive power ability of each of the variables (or features). Features that are highly correlated with the leaf wetness binary class (dry or wet) that also have a low correlation (with the other explanatory variables) are favoured. The search method used was a greedy hill climbing search with back tracking (BestFirst) and a 10 fold cross validation method was used. The explanatory variables considered were station, temperature, relative humidity, rainfall, and wind speed.

Table 2. Feature selection 10 fold cross-validation results for the data sets

	Dataset A	**Dataset B**
Feature	**number of folds (%) attribute**	
Station	10(100%)	0(0%)
Temperature	10(100%)	0(0%)
Relative Humidity	10(100%)	10(100%)
Rainfall	10(100%)	10(100%)
Wind speed	1(10%)	0(0%)

The results for datasets A and B are given in Table 2. For dataset A, all the features were found to have predictive power but wind speed was the least predictive. Interestingly the station ID is a relevant feature which suggests that the location of the stations where the data is collected is important and that models might need to be determined or tuned for each location in order to find the best performing leaf wetness estimation model. The reason both dataset are included was to allow models that requires training to generalise better.

In order further explore the relationship between the variables in each dataset a simple Kendal's tau_b (two tailed) correlation was performed. Kendal's tau_b is suitable for data that is not normally distributed and is continuous. Correlation coefficients range in value from –1 (a perfect negative relationship) and +1 (a perfect positive relationship) with zero indicating no linear relationship. For dataset A most of the coefficients are small but significant (Table 3. Kendal's tau_b correlation coefficients for dataset A). For large samples such as this (N = 7128), it is easy to achieve significance, and the strength of the correlation is important. Relative Humidity (RH) had the strongest correlation with sensor wetness readings (Wet) which was statistically significant ($\tau b = .531$, $p < 0.001$). The next most significant relationship was that of RH ($\tau b = .507$, $p < 0.001$) and temperature (Temp).

Table 3. Kendal's tau_b correlation coefficients for dataset A

	Temp	RH	Rainfall	Wind speed	Wet
RH	-.507**	1.000	.265**	-.429**	.532**
Rainfall	-.084**	.265**	1.000	-.037**	.267**
Windspeed	.320**	-.429**	-.037**	1.000	-.371**
Wet	-.365**	.532**	.267**	-.371**	1.000

**. Correlation is significant at the 0.01 level (2-tailed), N = 7128.

For dataset B, the strongest correlation is again the relationship between RH and Wet ($\tau b = 0.550$, $p < 0.001$). The second strongest is a negative correlation between RH and Temp. This relationship is the reverse of that for dataset A where a moderate positive relationship was observed between RH and temperature.

Table 4. Kendal's tau_b correlation coefficients for dataset B

	Temp	RH	Rainfall	Wind speed	Wet
Temp	1.000	-.425**	-.128**	.162**	-.319**
RH	-.425**	1.000	.269**	-.321**	.550**
Rainfall	-.128**	.269**	1.000	.084**	.313**
Wind speed	.162**	-.321**	.084**	1.000	-.210**
Wet	-.319**	.550**	.313**	-.210**	1.000

**. Correlation is significant at the 0.01 level (2-tailed), N = 8679.

4.2.2. Dataset A: Exploration

Dataset A contains five recorded variables at five different stations over two months from January to February 2012.

Table 5. Basic statistics for Wind Speed in 5 stations

Wind Speed (m/s)					
Statistic \ Stations	TRI	RPU	HXT	PKE	MTB
Min value	0.50	0.00	0.00	0.00	0.20
Max value	9.90	13.10	10.80	7.10	19.30
Mean	2.86	3.91	2.95	2.22	4.19
Variance	2.79	6.45	3.16	2.34	11.62
Standard deviation	1.67	2.54	1.78	1.53	3.41
Total recorded hours	1440	1436	1439	1368	1440

Inspecting Table 5, MTB station had the highest recorded wind speed (19.30 m/s). MTB station also scored the highest mean wind speed (4.19 m/s), which indicates that MTB station is the windiest station in dataset A. Wind speeds higher than 4 m/s are known to directly affect moisture evaporation on leaf surfaces (Rowlandson et al., 2015).

Table 6.Basic statistics for Relative Humidity in 5 stations

Relative Humidity (%)					
Statistic \ Stations	TRI	RPU	HXT	PKE	MTB
Min value	28.45	25.10	30.85	36.89	32.00
Max value	96.30	103.50	100.00	95.00	98.00
Mean	70.44	68.84	72.71	76.80	73.22
Variance	357.71	273.43	280.73	157.20	219.42
Standard deviation	18.91	16.54	16.75	12.54	14.81
Total recorded hours	1440	1436	1439	1368	1440

Table 6 shows the relative humidity basic statistics for the five stations of dataset A. Average humidity in four stations scored above 70%. RPU station recorded the lowest mean relative humidity with 68.84%. PKE station had the highest mean relative humidity and the lowest variance. Suggesting that at the PKE station relative humidity shifts within a smaller range of values than the other stations. A score above 100% in relative humidity is possible when supersaturation occurs. 100 percent of relative humidity is produced at a certain temperature and air pressure, given the maximum amount of water vapour in the air. In supersaturation, the air contains more water vapour than is needed to cause saturation.

Table 7. Basic statistics for Temperature in 5 stations

Temperature (°C)					
Statistic \ Stations	TRI	RPU	HXT	PKE	MTB
Min value	5.24	5.24	7.50	10.39	2.30
Max value	29.95	29.07	29.60	25.68	26.50
Mean	17.80	16.79	18.20	17.96	15.91
Variance	20.38	18.08	16.73	7.73	19.61
Standard deviation	4.51	4.25	4.09	2.78	4.43
Total recorded hours	1440	1436	1439	1368	1440

Looking at Table 7, the highest temperature value was recorded at TRI station with 29.95°C, and the lowest temperature recorded at MTB station with 2.30°C. In terms of temperature, PKE station had the lowest variance and thus the smallest range of temperatures. This indicates the temperature change in particularly PKE station is not as volatile as for the other stations.

Table 8. Basic statistics for Rainfall in 5 stations

Rainfall (mm)					
Statistic \ Stations	TRI	RPU	HXT	PKE	MTB
Min value	0.00	0.00	0.00	0.00	0.00
Max value	12.10	6.80	17.40	13.20	7.40
Total Rainfall	216.60	65.80	207.70	148.49	95.00
Variance	0.75	0.15	0.70	0.47	0.20
Standard deviation	0.87	0.38	0.84	0.68	0.45
Total recorded hours	1440	1436	1439	1368	1440

Table 8 shows the rainfall statistics for all five stations. The wettest station based on rainfall readings is TRI station with 216.60 mm of rainfall in 2 months. This dataset was taken in New Zealand summer time, hence the relatively low rainfall readings across all the stations. The driest station according to the total rainfall is RPU station with 65.80 mm rainfall. HXT station is the second wettest station with the highest single hour rainfall recorded, 17.40 mm. This suggests that HXT can also be considered a wet station during

January and February 2012. To better see the recorded rainfall comparison a monthly view is presented in Figure 12.

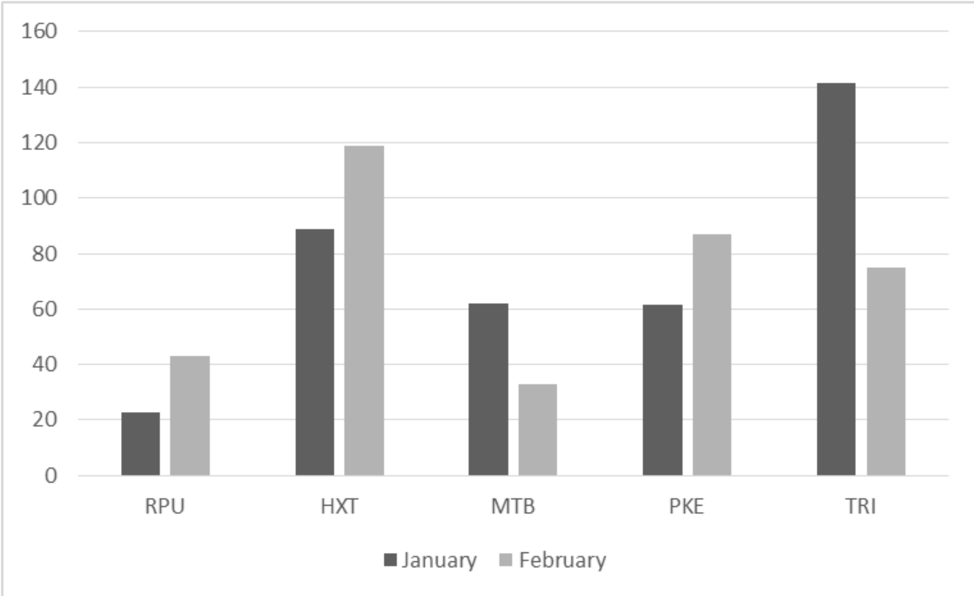

Figure 12. Monthly total rainfall in 5 stations for dataset A

In the monthly view, TRI had the highest rainfall in January and HXT in February. This makes both stations considered wet among the rest of the other stations. RPU and MTB are considered dry with total rainfall less than 65 mm each month, which is a quarter of the maximum rainfall in the wettest station.

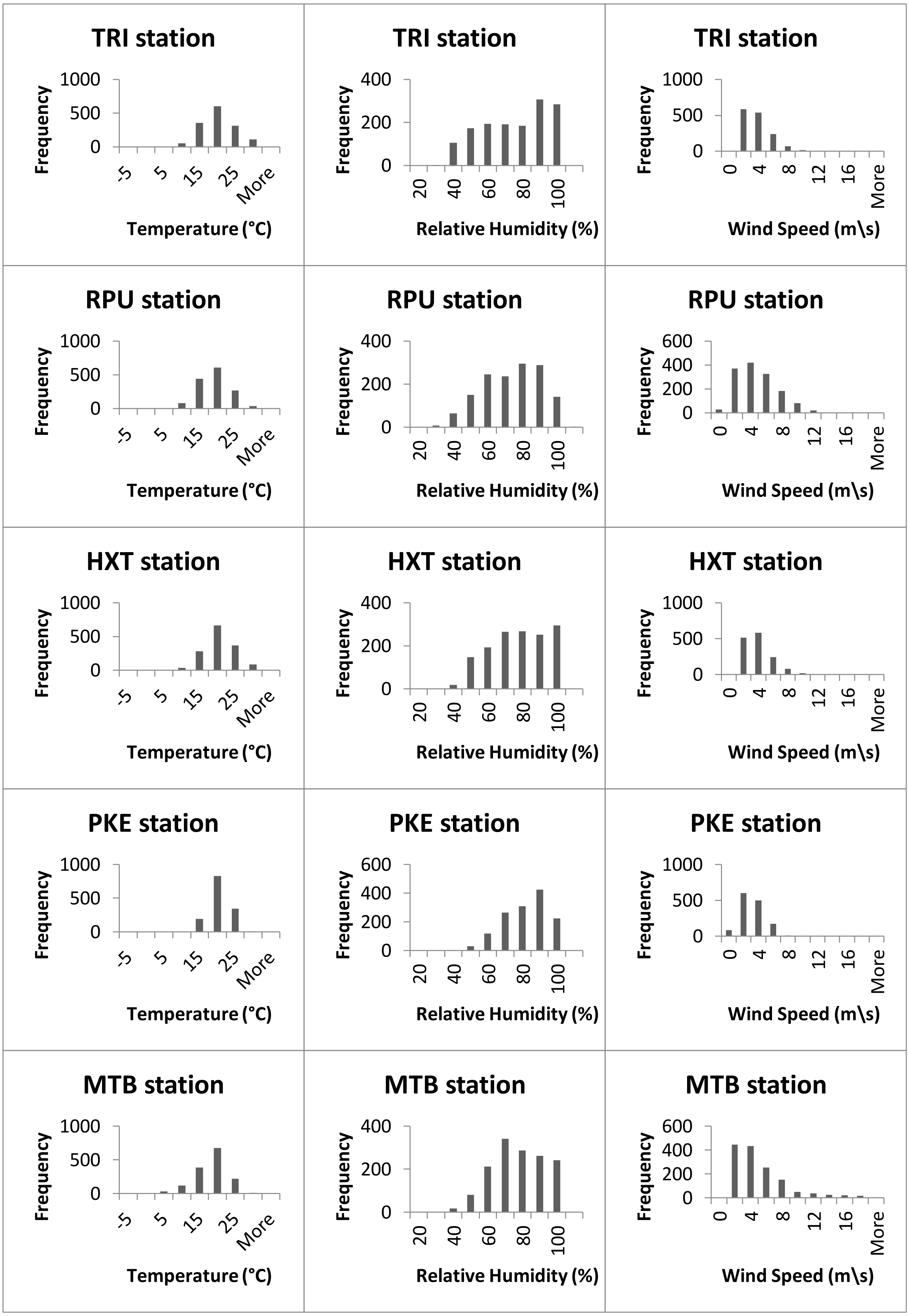

Figure 13. Histogram of temperature, wind speed, and relative humidity in 5 stations

Figure 13 gives histograms showing the distribution of all variables by station. The temperature appears to have a normal distribution for all five stations. Temperature has a peak frequency of around 20°C. However, a Kolmogorov-Smirnov test for normality the combined dataset found that none of the variables were normally distributed, see Appendix B.

Relative humidity variable at TRI, RPU, and HXT stations had a bimodal distribution. In this case the mean relative humidity of three stations is around 70%, but there are a number of 50-60% relative humidity readings as well. On the other hand, relative humidity in PKE and MTB stations are left skewed. Left skewness indicates the mean of all these stations are less than the median. In this case, both PKE and MTB stations' mean relative humidity is less than the median. This suggests that there are higher relative humidity readings in both stations.

Wind speed in all stations except MTB stations are left skewed with outliers. In MTB station the mean wind speed is greater than the median. This caused by the large number of low wind speed readings in the station. Wind speed mean value in 5 stations are ranging from 2.22 to 4.19 m/s, while MTB reading shows up to 19.1 m/s. MTB station's highest recorded wind speed was 6 m/s higher than the next station's maximum reading.

A closer look at rainfall reading correlation with leaf wetness duration in dataset A.

In the next figures the correlation between rainfall and leaf wetness readings is presented. In the graphs, the hourly leaf wetness are denoted by black dots and the left y-axis as hour in the day. Daily total rainfall is presented with bar graphs with the right y-axis as the millimetres of rain. The blue line, with numeric labels, indicates the trend in total number of wetness hours per day. Grey areas in the graph indicates night and the white area indicates daylight hours. The graphs (Figure 14 to Figure 19) show data logged for a period of 10 days. The days shown are selected to illustrate periods that have some wetness activity. For more leaf wetness and rainfall graphs, see Appendix B

1-10 January period

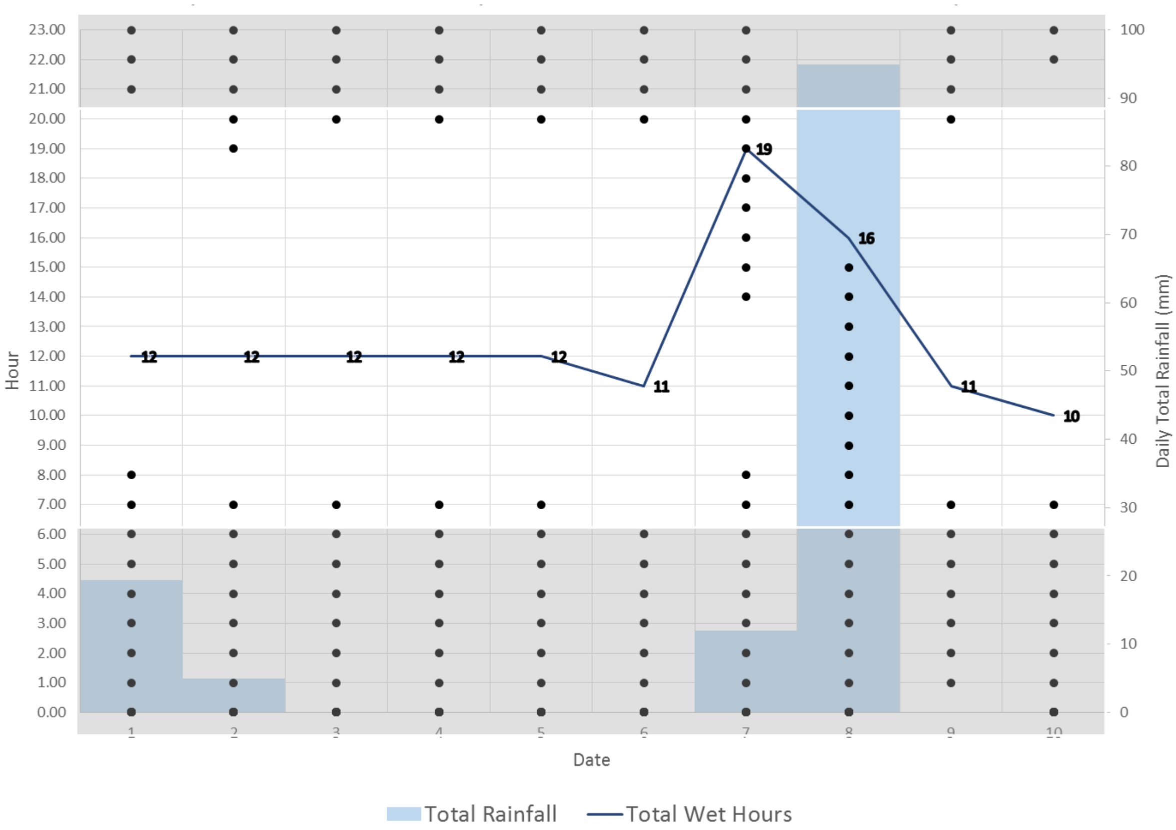

Figure 14. Daily total rainfall and hourly leaf wetness reading in TRI station 1-10 January 2012

In the first ten days of January at TRI station, leaf wetness was observed every night, even on days without rainfall. TRI station had weather highly conducive to dew formation in this period of time. The highest total wet hours were observed when rainfall was occurring. Daylight leaf wetness was only promoted by rainfall, as shown for the 7th and 8th of January.

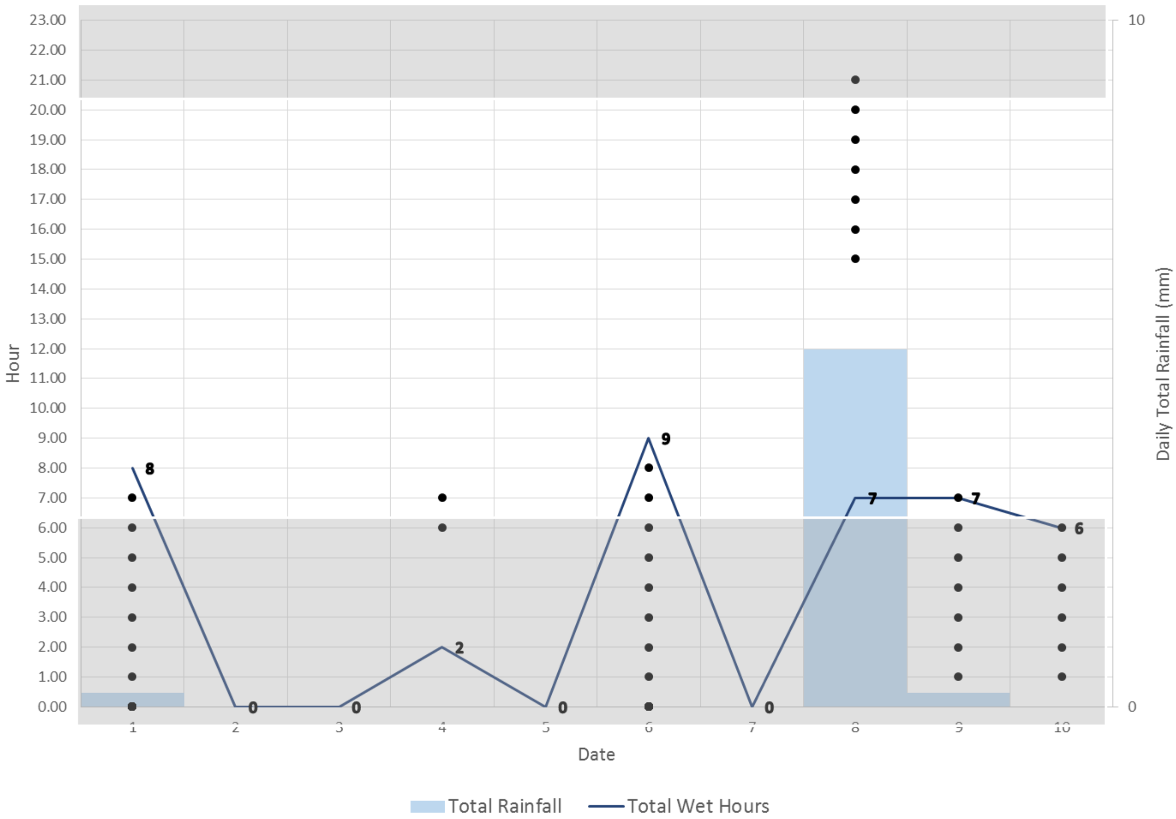

Figure 15.Daily total rainfall and hourly leaf wetness in RPU station 1-10 January

Figure 15 shows what is considered to be a 10 day dry period at RPU station. There were 2 days that had rainfall under 1mm, and the rainfall doesn't appear to directly affect daylight leaf wetness. The small amount of rainfall which fell during the night only contributed to one additional hour of wetness after sunrise. The only day with rainfall above 5 mm was on the 8th of January, same as the peak rainfall in Figure 14. At RPU station, prior to the 8th there was no recorded wetness. The rain on 8th of January lead to 6 hours of daylight wetness and 1 hour of wetness when the sun had set.

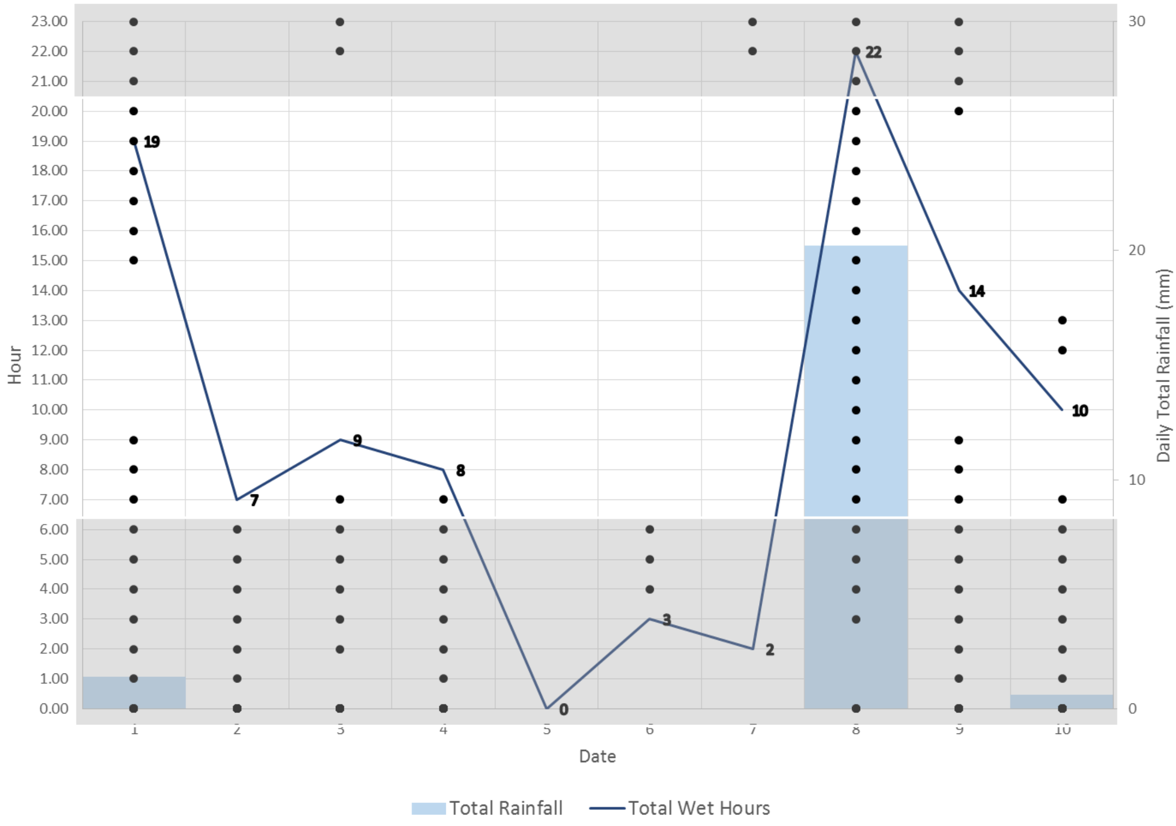

Figure 16. Daily total rainfall and hourly leaf wetness in MTB station 1-10 January

Figure 16 shows 3 days of rainfall with the peak on the 8th January. On the 8th there is a spike of total wetness from 2 hours on the previous day until 10 in the evening. This is due to total rainfall of 20mm that was spread throughout the day. The other days that don't have rainfall shows daylight wetness one to three hours onset due to the previous night's wetness.

21-31 January period

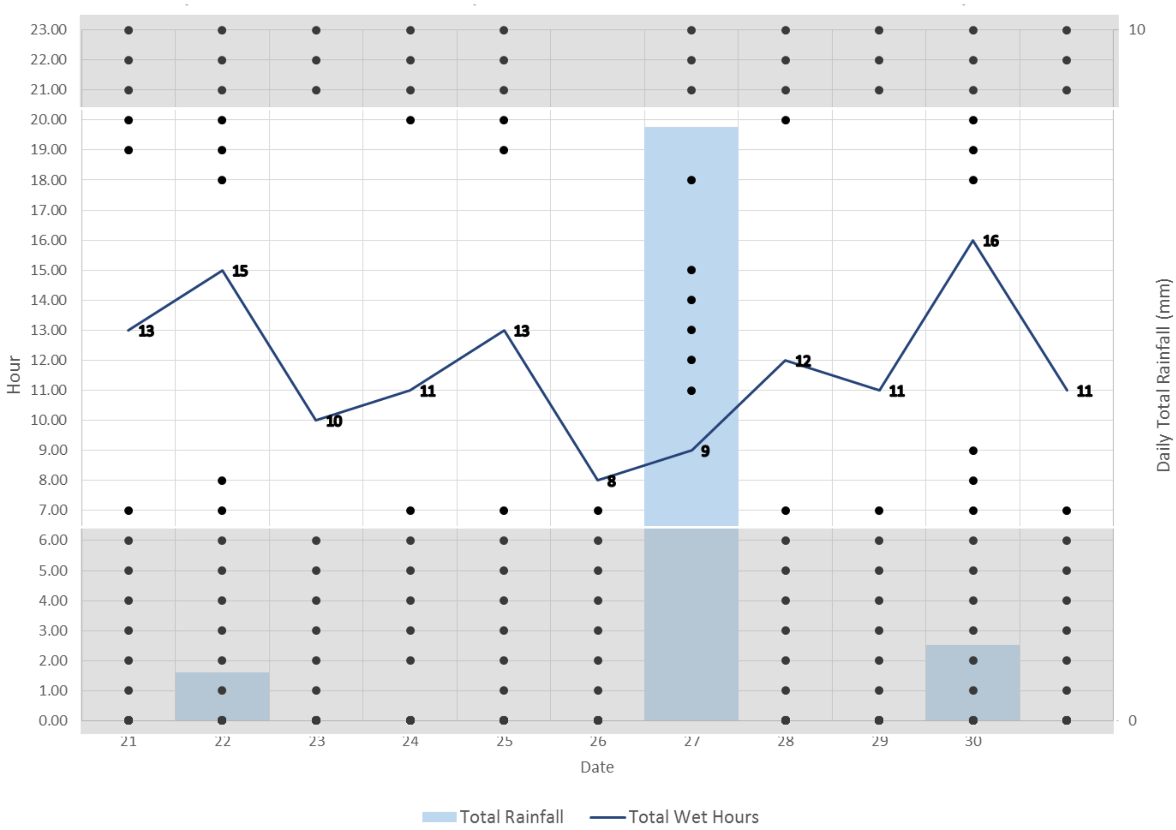

Figure 17. Daily total rainfall and hourly leaf wetness in TRI station 21-31 January

In the last 10 days of January at TRI station, 6 hours of daylight wetness was observed on the 27th which was the highest rainfall reading of the week. However, on the 27th the total rainfall was not the highest (9 hours). The two highest daylight total wetness values were recorded on the 22nd and 30th, both days were rainy. The other days had no rainfall and night time wetness was most likely caused by dew formation.

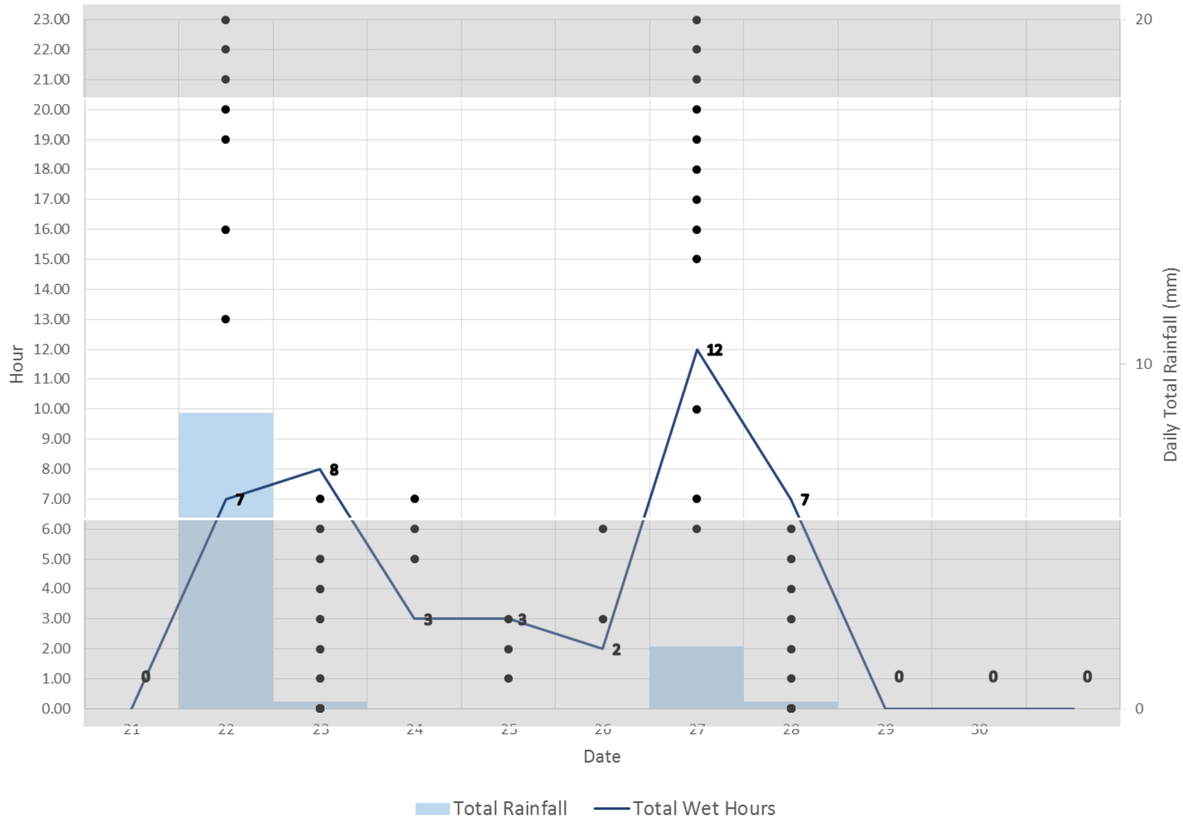

Figure 18. Daily total rainfall and hourly leaf wetness in RPU station 21-31 January

RPU station is considered located in a relatively dry location, see Figure 12. The graph above shows the last 10 days of January 2012 at the RPU station. Two days out of ten had significant rainfall, namely the 22^{nd} and 27^{th}. This rainfall assisted leaf wetness in the daylight totalling 7 and 12 hours of wetness, respectively. Rainy days tend to have continuous leaf wetness duration. Days that have wetness without rainfall do not show continuous leaf wetness durations of more than 3 hours. This is a good example of station where leaf wetness is dependent on rainfall.

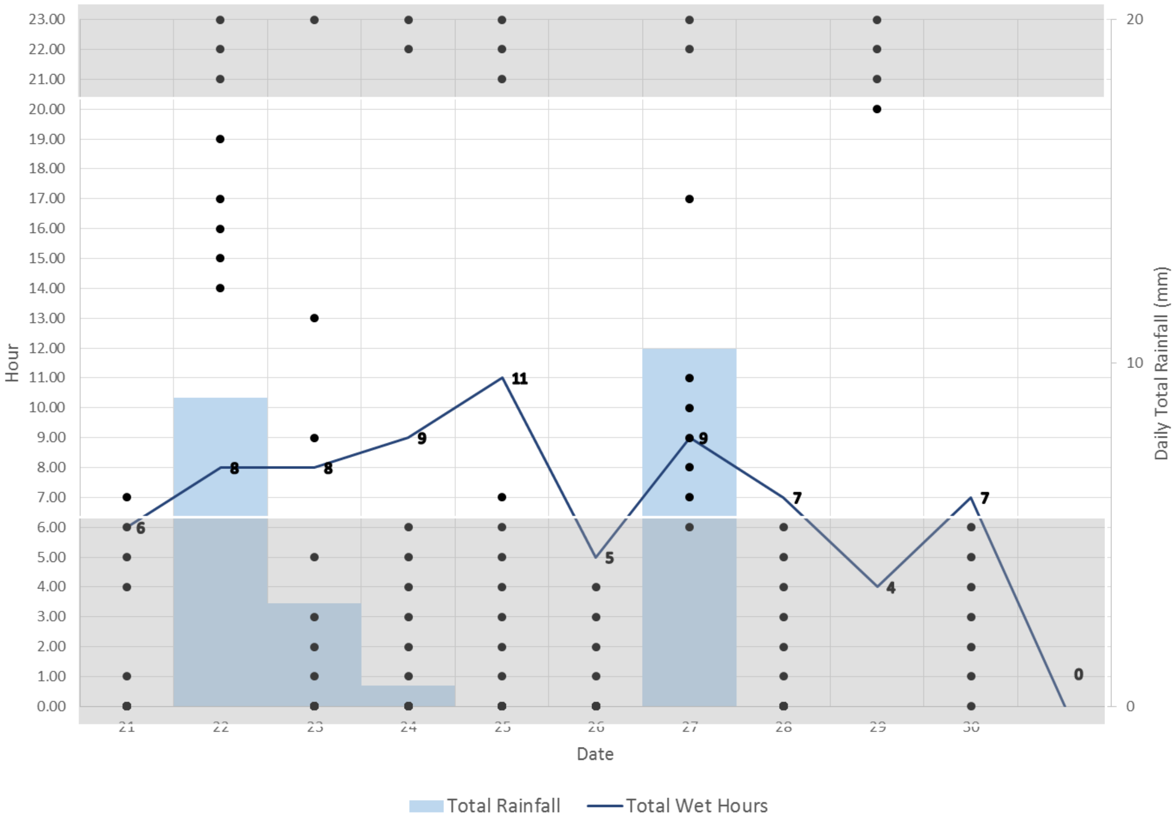

Figure 19. Daily total rainfall and hourly leaf wetness in MTB station 21-31 January

Figure 19 shows the last ten days of January 2012 at the MTB station. The rainfall pattern follows that of the TRI and RPU stations with peak total rainfall on the 22nd and 27th. The rainfall on the 22nd was responsible for all the daylight wetness of the day, as well as on the 27th. In this station, both days had no dew formation in the early morning. On the no-rain-days there is very little daylight leaf wetness observed.

4.2.3. Dataset B: Exploration

Table 9. Basic statistics for Wind Speed in 3 stations

Wind Speed (m/s)			
Statistic \ Stations	PKE	HXT	CRN
Min value	0.00	0.00	0.00
Max value	10.30	10.80	17.00
Mean	2.44	3.28	4.12
Variance	3.32	4.40	5.77
Standard deviation	1.82	2.10	2.40
Total recorded hours	2751	2952	2976

Looking at Table 9, CRN scored the highest recorded wind speed in the four months covered by this dataset. CRN station has the highest mean wind speed value (4.12 m/s). This suggests that CRN station is in a windy location.

Table 10. Basic statistics for Relative Humidity in 3 stations

Relative Humidity (%)			
Statistic \ Stations	PKE	HXT	CRN
Min value	31.91	24.02	18.52
Max value	95.90	109.00	100.00
Mean	77.63	77.17	71.30
Variance	160.94	301.77	319.62
Standard deviation	12.69	17.37	17.88
Total recorded hours	2751	2952	2976

Table 10 shows that CRN recorded the lowest relative humidity (18.52%). While the highest recorded relative humidity is that of the HXT station (109%). For this dataset the highest mean relative humidity is scored at the PKE station (77.63%).

Table 11. Basic statistics for Temperature in 3 stations

Temperature (°C)			
Statistic \ Stations	PKE	HXT	CRN
Min value	4.29	2.13	-3.80
Max value	25.68	33.88	33.63
Mean	14.43	13.65	12.30
Variance	15.88	34.17	41.15
Standard deviation	3.99	5.85	6.42
Total recorded hours	2751	2952	2976

Inspecting Table 11, CRN station has the lowest recorded temperature in the period (-3.8°C). The highest recorded temperature was at HXT station (33.88°C), very similar to CRN station's maximum reading (33.63°C). Both CRN and HXT exhibited a high variance in temperature. PKE station, on the other hand, had low variance (15.88) and the highest mean temperature of 14.43°C. Low variance with high minimum and low maximum means that temperature distribution in PKE station is more to the centre, see Figure 21.

Table 12. Basic statistics for Rainfall in 3 stations

Rainfall (mm)			
Statistic \ Stations	PKE	HXT	CRN
Min value	0.00	0.00	0.00
Max value	10.35	12.70	17.00
Total Rainfall	342.29	585.80	237.2
Variance	0.38	0.80	5.77
Standard deviation	0.62	0.89	2.40
Total recorded hours	2751	2952	2976

Table 12 shows the highest total rainfall was scored in HXT station (585.8 mm). Suggesting HXT station to be the wettest station in dataset B. The driest station according to the table above is CRN station with 237.2 mm of rainfall in 4 months. More detailed rainfall reading per month is presented in Figure 20.

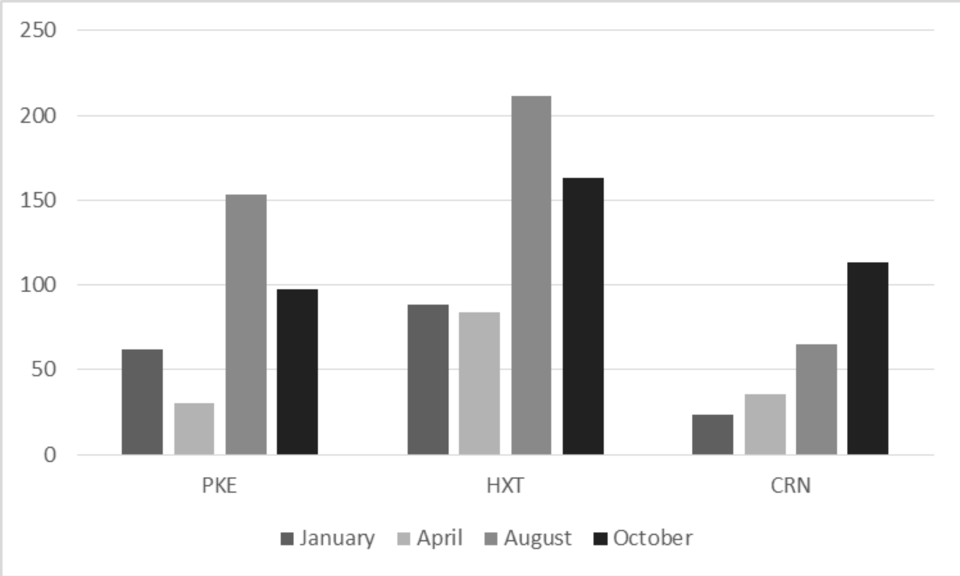

Figure 20. Monthly total rainfall in 3 stations for dataset B

In the figure above, it is clear that HXT station is the wettest station in dataset B. HXT station had a higher total rainfall every month when compared to the other stations. The rainfall in CRN is incrementally increasing over the months closer to winter. The peak total rainfall for PKE and HXT stations was in August.

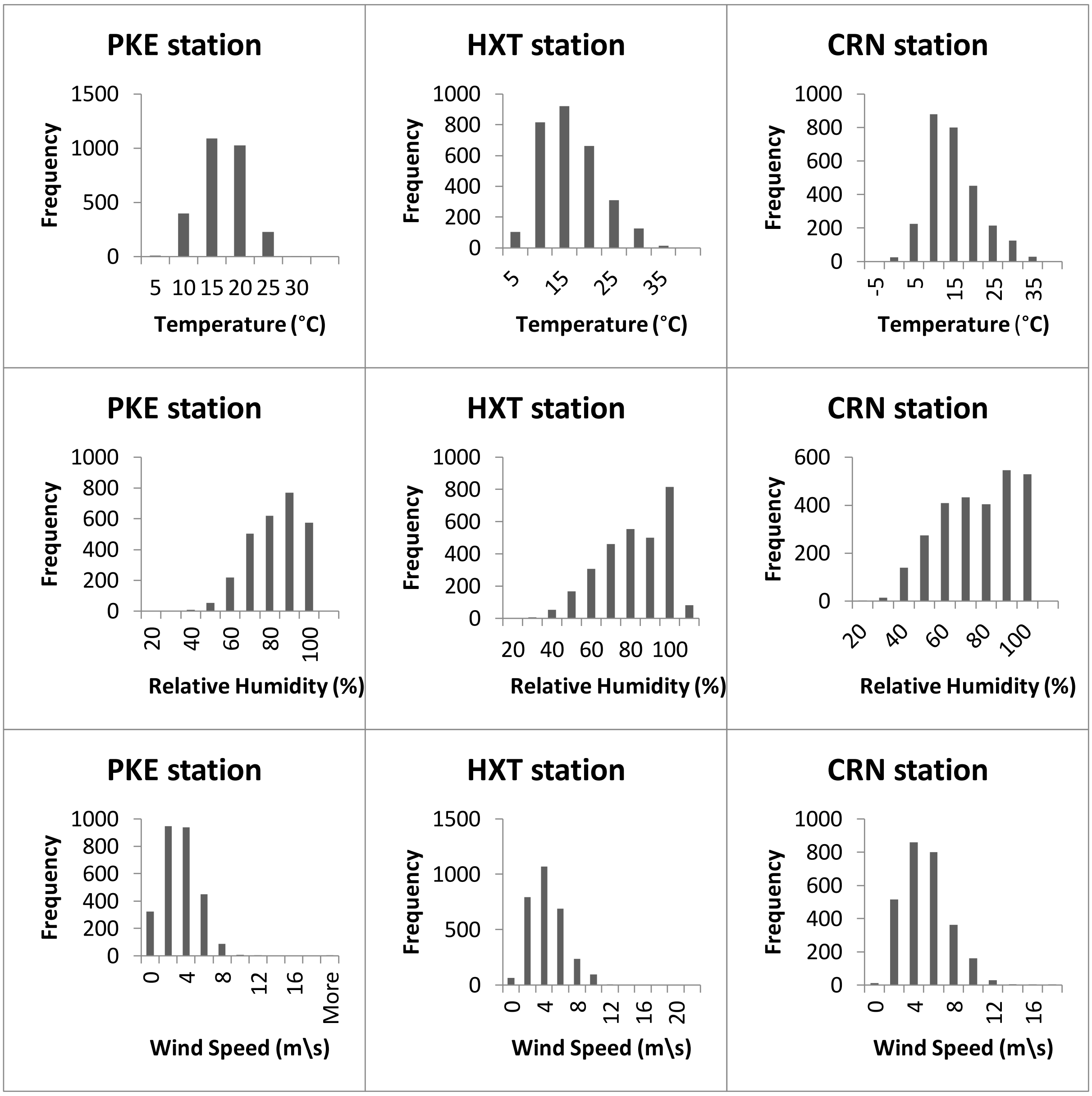

Figure 21. Histogram of temperature, wind speed, and relative humidity in 3 stations

Inspecting Figure 21, histogram for temperature, wind speed, and relative humidity for dataset B in all 3 stations. Temperature variable shows the closest to normal distribution compared to other variables. However, according to the normality test none of the variables had normal distribution (see Appendix B). Temperature frequency in all three stations peaked in the range of 10 to 20°C.

Relative humidity in all three stations is left skewed. This concurs with the previous histogram in Figure 13. In all three stations the mean relative humidity is higher than the median indicating that there are more high readings. The histogram above shows that wind speed readings are mostly focused around 2-6 m/s. There are a lot of outliers, which are caused by isolated wind gusts. The figure shows a mostly right skewed histogram, suggesting there are lower wind speeds recorded.

A closer look at rainfall reading correlation with leaf wetness duration in dataset B.

1-10 January period

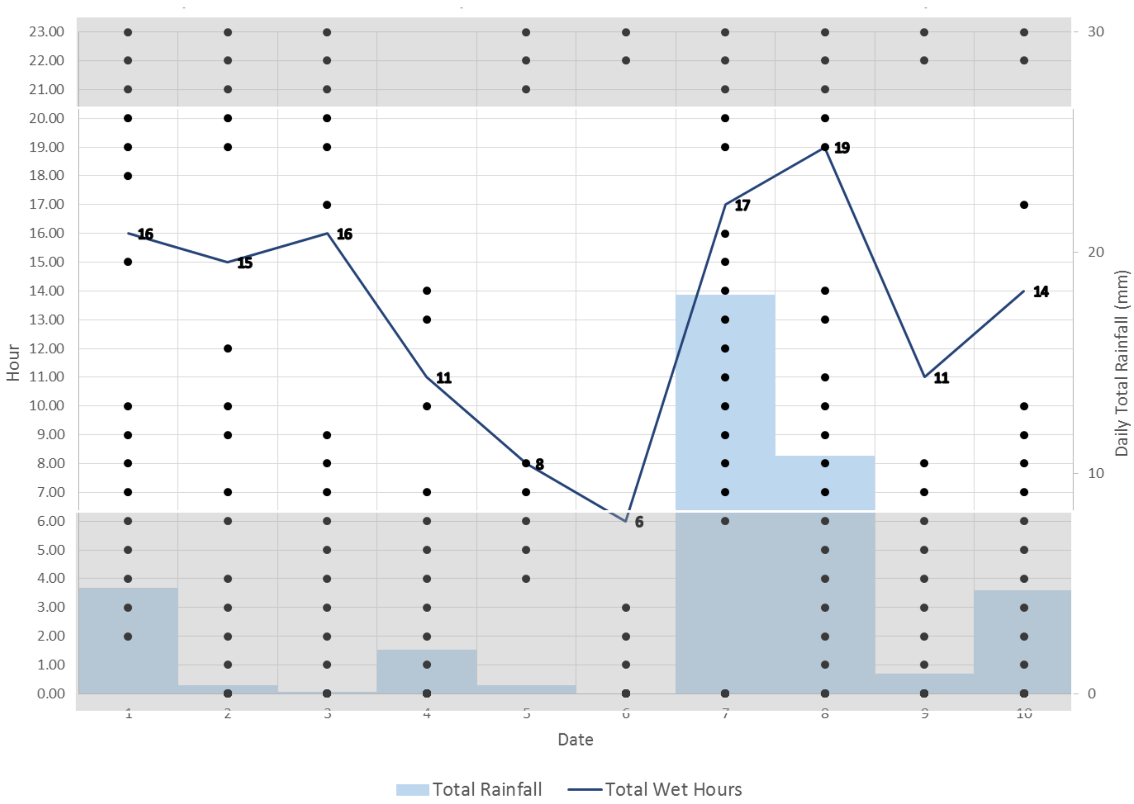

Figure 22. Daily total rainfall and hourly leaf wetness in PKE station 1-10 January

PKE station had highest mean relative humidity and also mean temperature, which makes this station is more dynamic than the other two. Relative humidity promotes dew fall, but high temperature responsible for evaporation of moisture on leaf surface. Rainfall activity in PKE station on January 2012 is substantial. In the first ten days period there was only one no-rain-day. Dew keeps forming throughout the dusk until dawn at PKE station. However, the daylight meteorological conditions are such that any surface wetness on the leaf will quickly dry off.

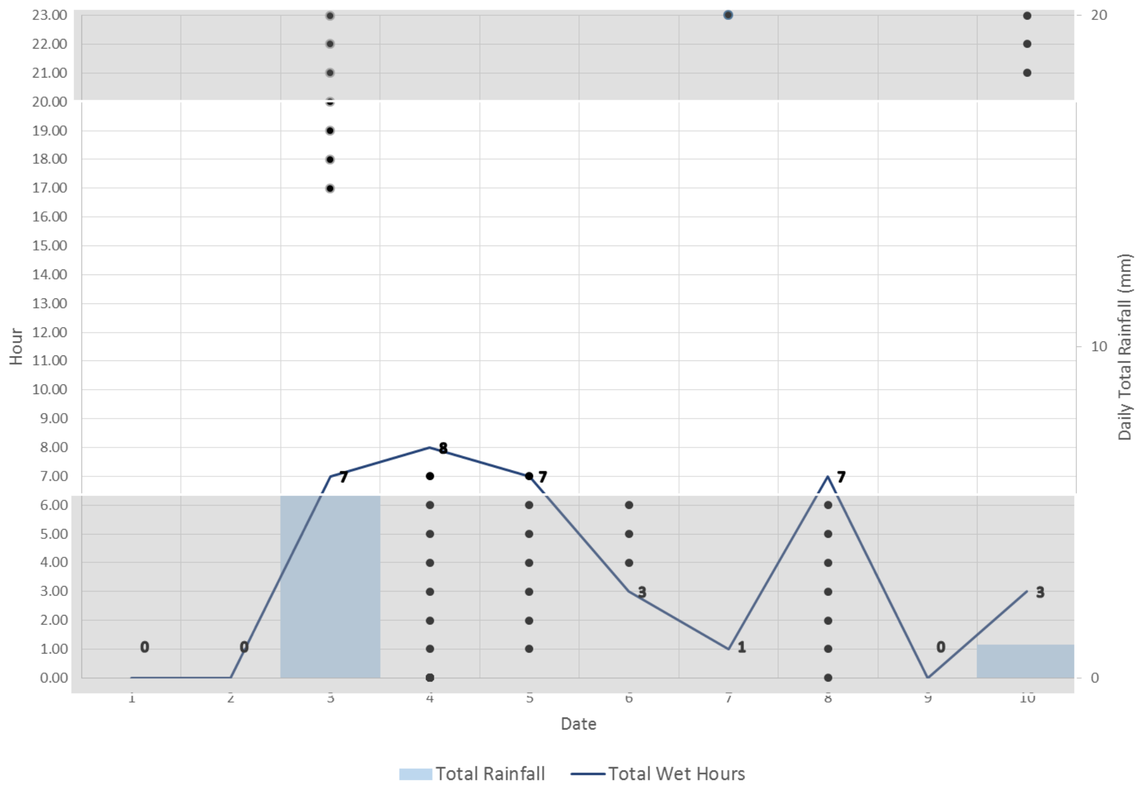

Figure 23. Daily total rainfall and hourly leaf wetness in HXT station 1-10 January

HXT station can be considered a relative wet station location, according to Figure 20. However the first ten days of January were mostly dry. With the exception of rainfall on the 3rd of January. There are no leaf wetness readings in the daylight at this station. Rainfall in the 10th was a small amount (<1 mm) and occurred during the night resulting in the night wetness reading.

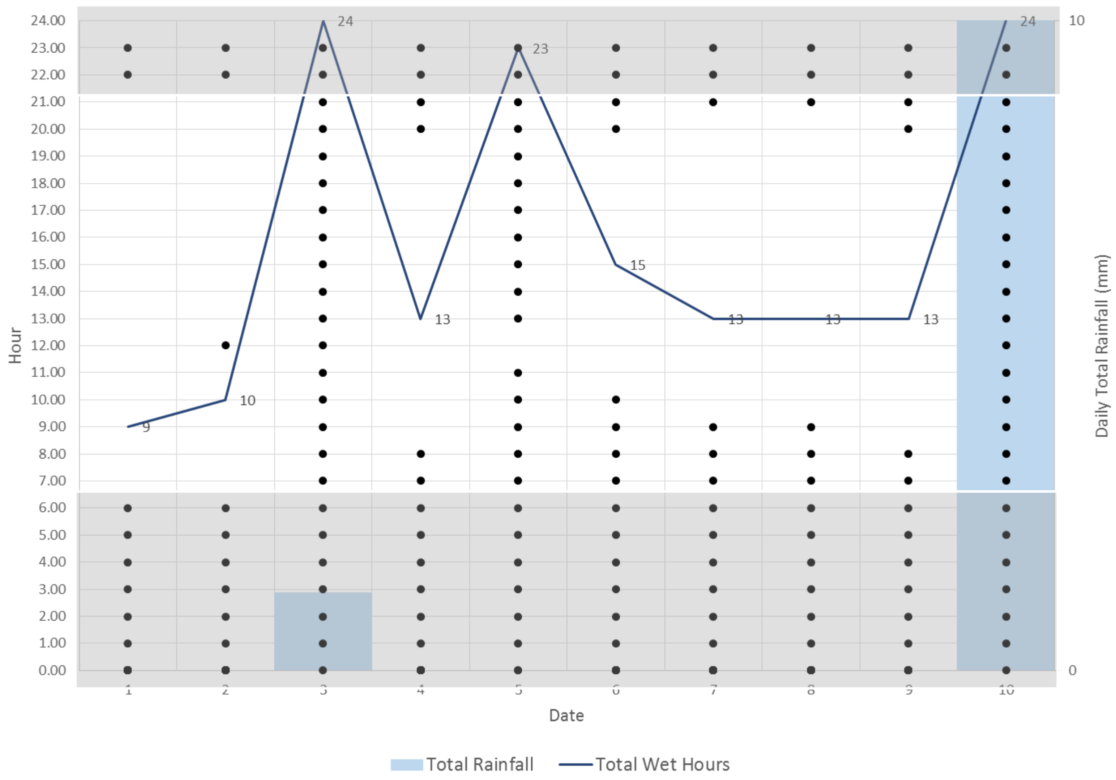

Figure 24. Daily Total Rainfall and Hourly Leaf Wetness in CRN Station 1-10 January 2012

CRN station scored the lowest monthly total rainfall. However, CRN also had the lowest minimum and mean temperature and there is a higher chance of dew formation when temperature readings are low. In Figure 24 above, dew forms ten days straight from dusk until dawn. Leaf wetness duration peaks on the 3rd and 10th, with 24 hours of wetness readings. This is regardless of the amount of rain which fell on those days. This suggests the weather conditions are more suitable for dew formation at this station when compared to the other two stations. An extreme example is on the 5th of January even without rainfall up to 23 hours leaf wetness duration was reached.

21-31 January period

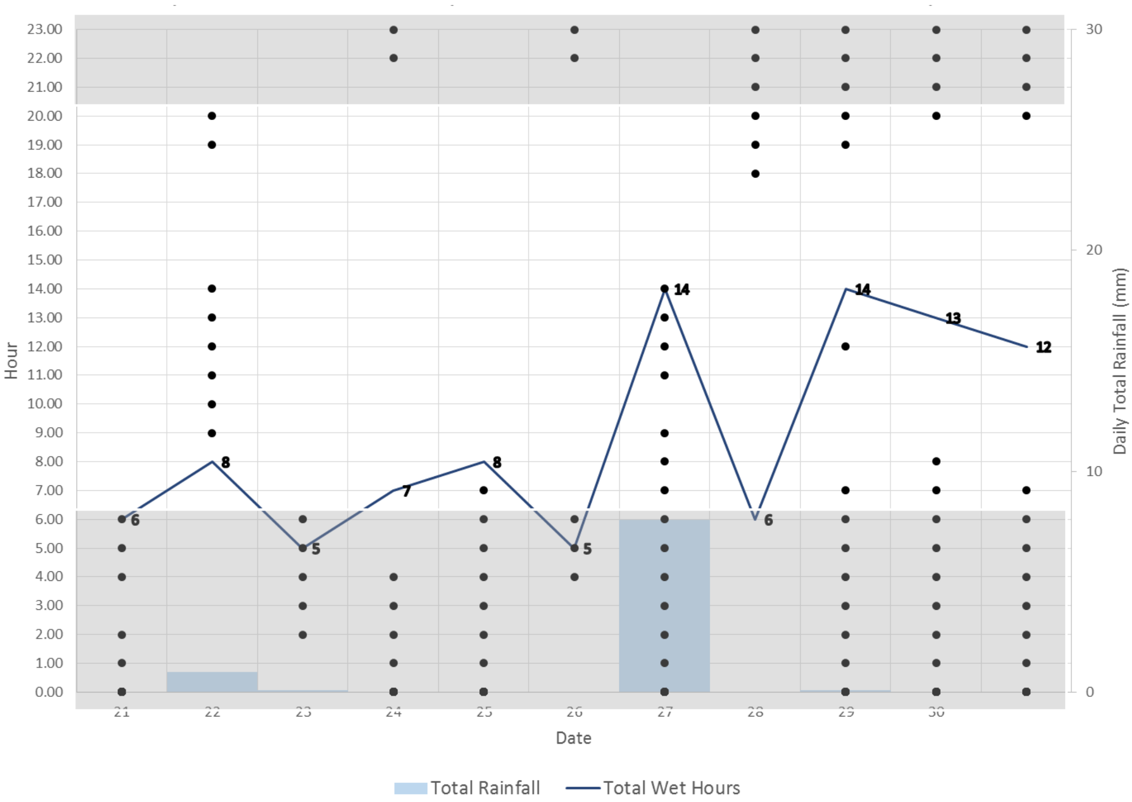

Figure 25. Daily total rainfall and hourly leaf wetness in PKE station 21-31 January

Figure 25. *Daily total rainfall and hourly leaf wetness in PKE station 21-31 January* shows leaf wetness and rainfall activity for the last 10 days of January at the PKE station. Two rainy days were recorded on the 22nd and 27th. Both rainy days promoted longer daylight wetness. On no rain days, leaf wetness was present at night.

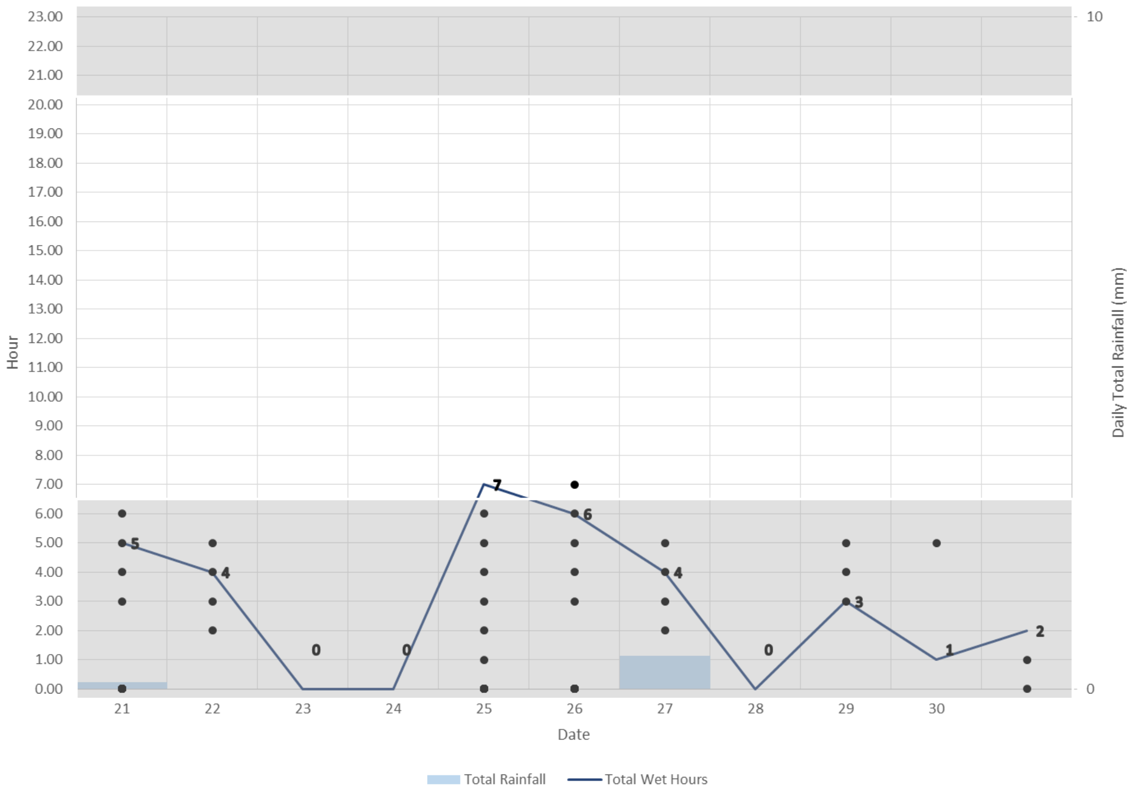

Figure 26. Daily total rainfall and hourly leaf wetness in HXT station 21-30 January

Leaf wetness and rainfall activity of the last ten days in HXT station are shown in Figure 26. There are two days where the amount of rain fall is less than 1 mm. On both days the wetness continued until sunrise and evaporated once the sun was up. This is considered to be a dry period because there was almost no daylight leaf wetness. Dew formed on a number of days in this period, with a maximum of 7 hours of leaf wetness during the hours of daylight.

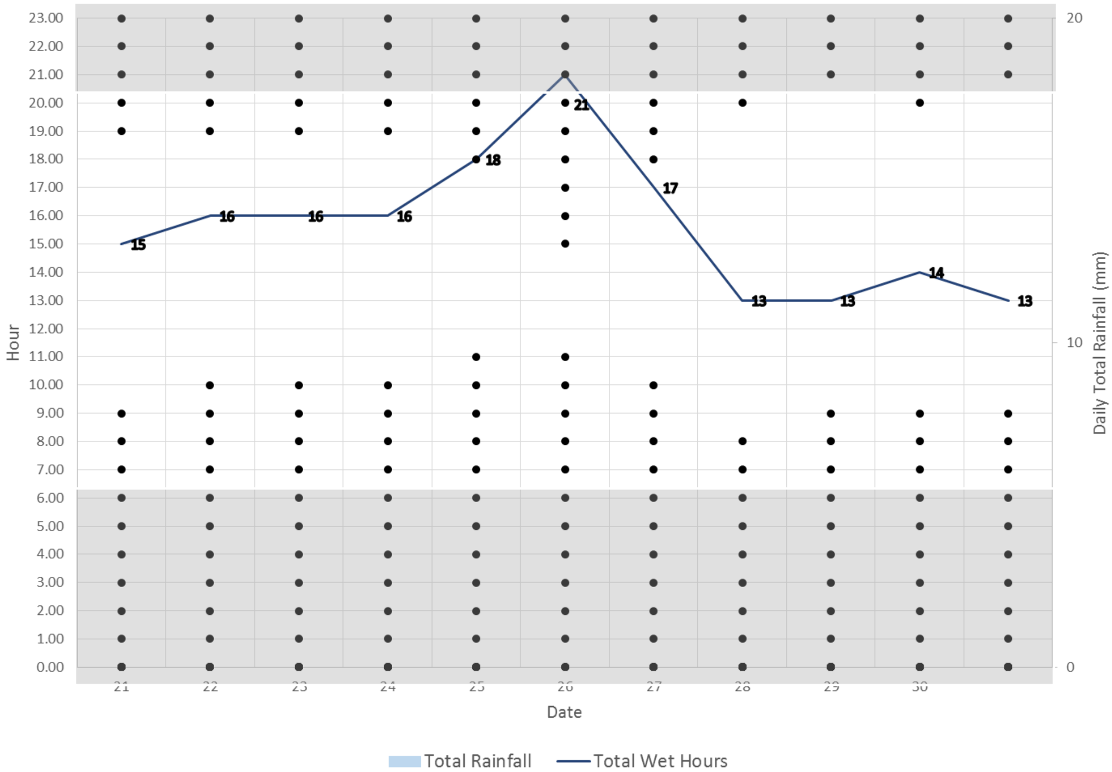

Figure 27. Daily Total Rainfall and Hourly Leaf Wetness in CRN Station 21-31 January 2012

HXT station is the station with the lowest temperature reading. The graph in Figure 27 ten no rain days in CRN station. There is a large amount of wetness during this period. This is presumably due to the low temperature that leads to dew formation. The peak number of wetness hours is 22 recorded on the 26th of January. Humid conditions that last for this long can lead to the formation fungi on the leaf surface that eventually leads to plant (leaf and fruit) diseases.

4.3 Leaf Wetness Model Evaluation

The datasets, described in the previous section, were used to generate leaf wetness models for each station. All of the models were validated with actual leaf wetness sensor readings and their performance was compared. For each model the Mean Error (ME), Mean Absolute Error (MAE), and Accuracy was calculated. ME was computed by averaging the differences between measured and estimated LWD for 24 hours period. ME determines the tendency of a model to overestimate or underestimate LWD. MAE was calculated by averaging the absolute values of hourly errors. MAE determines absolute accuracy of the model. The degree of closeness of estimated and measured LWD as a percentage was also calculated according to equation (1).

$$\text{EA} = \left(1 - \frac{\Sigma\,|\text{Actual} - \text{Estimated}|}{N}\right) \times 100 \qquad (1)$$

Dataset A) results

Inspecting the results in Table 13, ANFIS (Adaptive Neuro Fuzzy Inference System) model tended to underestimate LWD on average 0.06 h/day. Other models that that also tended to underestimate LWD were FLS (Fuzzy Logic System) model and P-M (Penman-Monteith) model. Most of the models showed the same trend in all five stations, except for SWEB (Surface Wetness Energy Balance) model that was found to underestimate for three stations (RPU, HXT and PKE), but overestimated for the other stations (up to 0.19 h/day). This result suggests that the portability of SWEB model is low when the physical attributes of canopy as inputs the model are not directly measured. In this research the canopy attributes (leaf size etc.) were set the SWEB models defaults. These default values were measured for a single vineyard in the USA. The NHRH≥ 90% (Number of Hours Relative Humidity ≥ 90%) model overestimated LWD as did CART (Classification and Regression Tree) model, and NHRH≥ 90% performed the worst of all the models overestimating by an average of 0.22 h/day.

Table 13. Mean Error of 7 models in 5 stations

Sites	ME[a] (h/day)						
	ANFIS	FLS	ANN	CART	NHRH≥ 90%	SWEB	P-M
RPU	-0.13	-0.13	0.08	0.18	0.16	-0.05	-0.12
HXT	0.00	-0.03	0.03	0.16	0.12	-0.05	-0.07
MTB	0.00	-0.18	0.02	0.15	0.19	0.02	-0.09
PKE	-0.11	-0.09	0.01	0.14	0.29	-0.05	-0.11
TRI	-0.03	0.03	0.01	0.16	0.31	0.19	0.09
Overall	-0.06	-0.08	0.03	0.16	0.22	0.01	-0.06

[a] ME = Mean error

Table 14 shows the MAE for the models developed using dataset A, it shows that ANFIS has the best overall MAE followed by P-M. The best MAE recorded is also given by ANFIS at the MTB station. FLS model with a predefined membership function performed poorly on this dataset, with a MAE of 0.28 h/day and the lowest MAE score 0.35 h/day at the RPU station. ANN on the other hand has a similar result (0.17 h/day) to that of the ANFIS model (0.14 h/day) both overall and by station. The overall mean absolute errors were similar for both physical models − 0.16 and 0.15 h/day for SWEB and P-M, respectively. For the empirical models there is a gap in the performance of NHRH≥ 90% and CART which scored 0.24 and 0.16 h/day. NHRH≥ 90% model being the simplest model outperformed the FLS model based on the overall MAE for this dataset. This is possibly due to training of the FLS model. This dataset was relatively small and training usually requires more data to generalise the model.

Table 14. Mean Absolute Error of 7 models in 5 stations

Sites	MAE[a] (h/day)						
	ANFIS	FLS	ANN	NHRH≥ 90%	CART	SWEB	P-M
RPU	0.20	0.35	0.22	0.20	0.18	0.11	0.13
HXT	0.16	0.29	0.18	0.16	0.16	0.19	0.18
MTB	0.05	0.30	0.09	0.24	0.15	0.09	0.09
PKE	0.18	0.31	0.16	0.30	0.14	0.22	0.23
TRI	0.10	0.17	0.20	0.31	0.16	0.20	0.11
Overall	0.14	0.28	0.17	0.24	0.16	0.16	0.15

[a] MAE = Mean absolute error

Table 15 shows the performance evaluation based on percentage accuracy. The results portrayed in this table show that ANFIS gave the highest overall accuracy of all the models (86.04%). ANFIS's lowest accuracy was recorded for the RPU station (79.55%) and the highest accuracy at the MTB station (94.76%).

Both physical models gave greater than 80% accuracy, SWEB with 81.51% and P-M with 80.79% accuracy. CART performed in line with the physical models with 84.26% accuracy and remarkably stable across stations suggesting that CART in general has less spatial variability. The NHRH≥ 90% model outperformed the FLS model by 4%, making FLS the least accurate model for estimation of LWD. However, the FLS model had the best accuracy for the TRI station in both Table 14 and Table 15. TRI station had the wettest location according to total rainfall records Considering that the dataset was imbalanced, it might be that the higher rainfall readings of the TRI station provided a more balanced data for training and therefore the FLS model was more accurate for that station.

Table 15. Accuracy of 7 models in 5 stations

Sites	Accuracy (%)						
	ANFIS	FLS	ANN	NHRH≥ 90%	CART	SWEB	PM
RPU	79.55	65.45	78.65	80.00	82.27	76.46	75.12
HXT	83.61	70.83	80.33	84.44	84.17	81.39	82.50
MTB	94.76	70.48	88.43	76.43	85.00	91.45	80.56
PKE	82.28	69.47	76.82	69.65	85.61	77.89	76.67
TRI	90.00	82.56	86.85	68.59	84.23	80.38	89.10
Overall	86.04	71.76	82.22	75.82	84.26	81.51	80.79

Table 16 gives the overall summary of results for the evaluation of models for estimating LWD based on dataset-A. These results show that ANFIS was the most precise model for predicting leaf wetness. It had an accuracy of well over 80% and the lowest overall absolute error of 0.14 h/day. Integrating the aspects of FLS and ANN model into one model therefore has proven to lead to better accuracy for leaf wetness prediction. However ANFIS outcome was only slightly better than CART, ANN, and the two physical models.

Table 16. Overall result of 2 months dataset in all stations

Models	Overall		
	ME[a] (h/day)	MAE[b] (h/day)	Accuracy (%)
ANFIS	-0.06	0.14	86.04
FUZZY	-0.08	0.28	71.76
ANN	0.03	0.17	82.22
NHRH≥90%	0.16	0.24	75.82
CART	0.22	0.16	84.26
SWEB	0.01	0.16	81.51
P-M	-0.06	0.15	80.79

[a] ME = Mean error
[b] MAE = Mean absolute error

ANN, CART, SWEB, and P-M also scored under 0.2 h/day MAE and accuracy above 80%. Both physical models show good portability in this dataset, with the tendency of overestimation, agreeing with the results from Magarey (2006) for SWEB and Sentelhas (2006) for the P-M model. In both research, both models are overestimating up to 0.7 and 1.3 h/day, respectively. CART model was developed in Ames, Iowa, which was wetter than the 7 locations tested New Zealand, this might explain the overestimation of CART in this study (Gleason et al, 1994). Nonetheless, CART model shows low variability between stations that suggests that this model have good portability.

NHRH≥ 90% model had the second to worst result overall, as low as 75% accuracy and with a tendency to overestimate LWD. FLS performed the worst. In this study the results for the NHRH≥ 90% and the FLS model did not agree with the results obtained by Kim (2006). Kim's study showed that the NHRH≥ 90% model underestimated leaf wetness and was also less precise than the FLS model (Kim et al, 2006). Of course Kim used a different data set and this disagreement of results illustrated the localisation issues which may be associated with estimation models. This discrepancy also indicates the need for an evaluation with a larger and a more balanced leaf wetness class in order to achieve optimal accuracy.

Evaluation results for dataset-A indicate that the ANFIS, FLS, and ANN models, which require training and tuning of parameters may benefit from a larger and more balanced dataset, hence there a second evaluation was performed using a different dataset.

Dataset B) results

Table 17 indicates that the ANFIS model generated for dataset-B is mainly underestimated LWD, in agreement with the results for dataset-A (Table 13), but with a value that is closer to zero (-0.05 h/day). FLS and ANN model both also underestimated LWD by 0.1 and 0.07 h/day respectively. The FLS model actually performed more poorly on this larger dataset consistently underestimating leaf wetness regardless of location. ANN was consistently overestimating leaf wetness for dataset-A (Table 13). Dataset-B resulted in ANN underestimating leaf wetness. Dataset-B is larger than dataset-A and overall has a higher total rainfall and more leaf wetness hours, this change in data explains the changes result magnitude in some models such as ANN where the larger dataset allows for a larger training set.

CART again gave a similar outcome to ANFIS and ANN (0.08 h/day) with a better result than for dataset-A (Table 13), which where the mean error was 0.16 h/day. NHRH≥90% was found to overestimate LWD by 0.18h/day overall. SWEB was the worst performing model in terms of magnitude of error, and overestimated leaf wetness by 0.19 h/day.

Table 17. Mean Error of 7 models in 3 stations within 4 months

Sites	*N*[a]	AWD[b]	ME[c] (h/day)						
		(h/day)	ANFIS	FLS	ANN	NHRH≥ 90%	CART	SWEB	P-M
PKE	114	3	-0.09	-0.10	-0.09	0.21	0.10	0.14	-0.07
HXT	123	3	0.00	-0.13	-0.03	0.18	0.05	0.19	-0.13
CRN	124	4	-0.05	-0.07	-0.08	0.15	0.08	0.25	-0.11
Overall	361	3	-0.05	-0.10	-0.07	0.18	0.08	0.19	-0.10

[a] N = Number of data
[b] AWD = Average wetness / day
[c] ME = Mean error

Table 18 shows the overall absolute errors of the models generated using dataset-B. ANFIS had the lowest MAE. ANFIS model result in dataset-B was also improved from dataset-A, from 0.14 to 0.10 h/day. ANN was a close second with 0.14 h/day while FLS had 0.21 h/day. Both physical models gave close results with low variability between stations. MAE for SWEB and P-M was slightly better than the MAE of the CART model, 0.16, 0.17, and 0.18, respectively. NHRH≥ 90% had the worst MAE for estimating LWD, 0.29 h/day in dataset-B. Among the approaches investigated, only NHRH≥ 90% model gave a worse performance when introduced to the larger dataset.

Table 18. Mean Absolute Error of 7 models in 3 stations within 4 months

Sites	N^a	AWD[b] (h/day)	MAE[c] (h/day)						
			ANFIS	FLS	ANN	NHRH≥ 90%	CART	SWEB	P-M
PKE	114	3	0.10	0.18	0.15	0.24	0.11	0.18	0.18
HXT	123	3	0.05	0.28	0.09	0.30	0.22	0.15	0.22
CRN	124	4	0.16	0.17	0.19	0.34	0.20	0.16	0.11
Overall	361	3	0.10	0.21	0.14	0.29	0.18	0.16	0.17

[a] N = Number of data

[b] AWD = Average wetness / day

[c] MAE = Mean absolute error

Table 19 shows the performance evaluation in percentage of accuracy of dataset-B. The ANFIS model was the only model which had more than 90% overall accuracy. ANFIS had improved performance on the new dataset (86% to 90%). The same improvement was seen for the FLS and ANN models, 71% to 79%, and 82% to 86%, respectively. The larger dataset improved the outcomes of these models. Both ANFIS and ANN clearly benefited from using a larger training set and the FLS model can better generalise with larger dataset.

NHRH≥ 90% accuracy was just over 70%, this result is worse than that of the smaller dataset. This high degree of error suggests that the NHRH≥ 90% model needs to be calibrated in order to find an appropriate relative humidity threshold based on local conditions.

Table 19. Accuracy of 7 models in 3 stations within 4 months

Sites	Accuracy (%)						
	ANFIS	FLS	ANN	NHRH≥ 90%	CART	SWEB	PM
PKE	90.71	82.15	85.63	76.25	89.20	82.43	82.12
HXT	95.23	72.42	91.44	70.16	78.74	85.20	78.55
CRN	84.24	83.53	81.85	66.00	80.59	84.62	89.93
Overall	90.06	79.37	86.31	70.80	82.84	84.08	83.53

Table 20 shows overall result of MAE, ME and ACCURACY for all the models. The ANFIS model is a definite winner in the group with accuracy over 90%. FLS and ANN models benefit from larger dataset with better accuracy, but they were both still outperformed by ANFIS.

Physical models show the best consistency of performance for different datasets with only slight differences in the overall results between dataset-A and B and also low variability in between stations. The CART model outcome was slightly worse than for dataset-A, together with NHRH≥ 90% which was the worst performer in the group with 70.8% accuracy.

Table 20. Overall results of 4 months dataset in 3 stations

Models	Overall		
	ME[a] (h/day)	MAE[b] (h/day)	Accuracy (%)
ANFIS	-0.05	0.10	90.06
FLS	-0.10	0.21	79.37
ANN	-0.07	0.14	86.31
NHRH≥ 90%	0.18	0.29	70.80
CART	0.08	0.18	82.84
SWEB	0.19	0.16	84.08
P-M	-0.10	0.17	83.53

[a] ME = Mean error
[b] MAE = Mean absolute error

In these experiments, ANFIS, FLS, and ANN showed significant improvement for dataset-B. ANFIS and ANN performed with good results on the HXT station. In dataset-A the best result using ANN and ANFIS methods was with data from the TRI station. HXT and TRI stations are the stations with most rainfall in each dataset; other stations had little rainfall during the same period.

It is important for the FLS logic model to generalise to the dataset in order to achieve an optimal result. Defining the membership function is an important step in FLS model development. In this study the membership function established by Kim (2004) for LWD was used in the FLS model. This membership function was used in order to provide a benchmark comparison with ANFIS. In ANFIS the membership function is generated using an ANN.

Physical models, SWEB, and P-M performed with slightly better result (overall above 80%) with the same low variance between stations. This indicates that physical models have higher portability than empirical and hybrid models. PKE station had the highest mean relative humidity that helps dew formation. It also prolongs moisture presence on leaf surface. PKE station also had the highest mean temperature, which responsible for water evaporation. This balanced variable dynamic makes PKE station data suitable for optimal models result.

Both empirical models, CART and NHRH≥ 90% suffer from a loss of accuracy when introduced to the larger dataset. NHRH≥ 90% model gave the lowest result for the CRN station which has the lowest average relative humidity. LWD is not as simple as considering only RH. Based on the feature evaluation given in section 5.2 there are other explanatory variables which also have leaf wetness predictive power.

CART is also developed using thresholds for variables such as relative humidity, wind speed, and dew point depression. These thresholds might have to be adjusted when the model is applied in a different region. Therefore in the case of empirical models, when applied to New Zealand region datasets, it appears that it is important for empirical models to be calibrated locally. This is supported by a recent paper published in New Zealand using CART model as inputs for botrytis bunch rot prediction system. The study found that CART consistently underestimated leaf wetness across seven stations and concluded that CART model is not suitable for use without local calibration (Henshall, 2015).

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

Monitoring leaf wetness is a crucial component to the management of many plant diseases. Leaf wetness measurement or estimation has its own limitations that needs to be acknowledged. Leaf wetness sensors are labour sensitive. They require maintenance to some extent, some types need to be treated to increase sensitivity, particular placement in crop canopy, and also calibrations variability for different sensors. To date, there has been no standardisation for leaf wetness sensor production, usage, and measurement. This fact leads to inability to compare results among research. There was a need for a comprehensive comparative analysis on commercially available leaf wetness sensors. To the author's knowledge this is the first research that includes four different sensor types and two of which are treated with paint and heat, totalling six leaf wetness sensors involved. This research found that of the commonly used commercial leaf wetness sensors DD (Decagon Devices) sensor is superior to the three other sensor types evaluated, regardless of any treatment of the other sensors. Thus answering research question one "Which of the commercially available leaf wetness sensors gives the most accurate measurement?".

DD sensors were found to have the highest accuracy when compared with other sensors that have been used for decades. It benefits from its dielectric principle design, fibreboard surface, and also latex paint coating. Previous research has suggested that the physical features of the sensor are beneficial (Sentelhas, 2004). Painting sensors has also proven to improve sensitivity in larger sensor sizes such as CS (Campbell Scientific) sensors. However, smaller sensors such as HB (Hobby Board) sensor tend to suffer from overestimation when painting was applied to the surface. This is likely to be due to the sensor surface to water droplet ratio being high resulting in water droplets joining to form larger droplets. Painting assisted the grouping of water droplets which makes it too sensitive for the sensor. Rust was also prevented or at least slowed down when painting was applied which slows deterioration of the sensor in the field.

Filter paper based sensor such as PI (Pessl Instrument) sensor is strongly recommended that this type of sensor is not used in industry. It holds water for a longer time as it is absorbed in the paper. The shape of the sensor also intercepted wind blowing to the filter paper, leading to overestimation of leaf wetness. This research indicates that a DD sensor is a stellar choice for industry use. Industries that are already using other sensors such as

CS sensor should consider using DD sensor. The next time they need to change their deployed sensor due to rust they should consider a DD sensor as a replacement. DD sensors come painted which means that they typically have a longer life span than unpainted sensors and require less maintenance.

It should be noted that in this research the sensor calibration experiments, to determine which type of sensor measures leaf wetness the most accurately, were undertaken independently of the LWD modelling experiments. The modelling was undertaken using historical data and the leaf wetness was measured using CS sensors.

Simple mathematical models have been suggested in some studies to be a preferable alternative to sensor measurements for determining LWD. Empirical models such as NHRH≥ 90% require one variable, relative humidity, and are often used in practice to estimate LWD but the accuracy of the model can be compromised by considering only one explanatory variable. NHRH≥ 90% model only takes into account one variable, in a complex system and arguably does not represent the system sufficiently. Portability is also an issue for empirical models, when the model is applied to a region with different climates, it tends to give false estimates. These problems led to the development of physical models that take into account the physical attributes of the crop canopy as an additional variable.

Physical models are developed around an energy balance principle and evaporation rate. Some models also take into account variables such as leaf physical shape, water threshold, and size as calibration features. In this research physical models turned out to be better in terms of portability when compared with empirical models. However, they require a large number of inputs and often such information will not be available. To reduce the amount of variables needed as input but still considering important physical variables, a hybrid model was proposed as a preferable alternative. Hybrid models combine empirical and physical models in order to draw on the strengths of each of those model types. One of the hybrid models, FLS model takes vapour pressure deficit, wind speed, and net radiation as its inputs. These inputs were chosen because of their high correlation with leaf wetness. FLS model in the latest reported research is claimed to estimate leaf wetness with high accuracy (Kim, Taylor, & Gleason, 2004). However, membership functions to address the values of the inputs need to be defined locally.

This study has found that an FLS model without locally defined membership functions is not optimal (see Chapter 4). This site dependency is known to be an issue for LWD modelling. This leads to difficulty in applying the model to multiple stations in different areas due to the different meteorological data collected. Different weather stations use different sensors and collect different variables and datasets are often not complete. It is difficult to compare locally specific models or to come up with a more generalised model due to such inconsistencies in the data.

In this study, ANFIS was proposed for estimating leaf wetness. The study points out that ANFIS system benefits from the training process to generate membership functions. In this research ANFIS was developed and the results of estimating leaf wetness was evaluated against six different existing leaf wetness models. The first comparison, using Dataset A resulted in only slightly higher accuracy than the other models. The insignificant improvement found in the first comparative analysis was believed to be mainly caused by the small amount of data which was available. It prevents ANFIS model to train the data optimally. The second comparative analysis shows ANFIS model has higher accuracy compared to all six models. ANFIS system allows any kind of variable to be taken as input. ANFIS also allows any dataset to have tailored membership functions thus providing a portable model. ANFIS omits the need of defining membership functions manually and trains using a neural network instead.

In this research the dataset used was small and consisted on data from one growing season and for only five sites. While the researcher had access to historical data covering up to a maximum of ten years for 25 stations the data was not continuous and was incomplete. While the variables were consistent, different stations had data for different periods of time within that ten years. The data was often missing for several weeks or even months due to sensor failure and time for repair. Thus it was difficult to find complete data from the same time period. This does to some extent limit the findings of this research because only one growing season was involved. However the data was sufficient to undertake validation of the models using ten-fold cross validation and ANIFS was found to give the most accurate LWD estimates for four of five stations. A physical model which required more variables gave better estimation for the other station. However these variables are often not available and therefore it is more practical to adopt ANFIS which was the second best model at this station.

Further research on ANFIS should be conducted in order to optimise the models. Investigation into alternative methods to determine the membership functions and further tuning of the parameters should be undertaken. Empirical models are known for only being valid at a local scale and require a large amount of data. In order to test the ANFIS model's portability further different sites with different climates should be included. A larger dataset would also allow for improved training in order to establish the membership functions.

Although NHRH$\geq$ 90% is simpler to use with adequate results, the model does not consider other equally important variables affecting leaf wetness and is not entirely reliable. Therefore, it is not suitable to use as a leaf wetness estimator for creating input for disease warning systems. On the other hand, ANFIS appears to be a practical alternative to existing models or the use of leaf wetness sensors. The outcome of this research suggests that ANFIS offers the potential to be a robust leaf wetness estimation system with high accuracy answering research question three "Can an adaptive neuro-fuzzy inference system be used as a leaf wetness model?".

Regardless of how leaf wetness is determined, either measured or estimated, the limitations has to be acknowledged. For smaller crops and when a sensor installation is possible, measurement would be more suitable. Larger crops or crops that require frequent maintenance would benefit more from models to estimate leaf wetness.

In conclusion this study has found that for measurement DD sensor is superior to other sensors and for estimation ANFIS model provides higher accuracy than previously developed leaf wetness models.

REFERENCES

Aitken, A., & Hewett, E. (2011). FreshFacts New Zealand Horticulture. *Plant & Food Research*, 33p. http://www.plantandfood.co.nz/file/freshfacts-brochure-2010.pdf

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). Classification and regression trees: CRC press.

Dalla Marta, A., De Vincenzi, M., Dietrich, S., & Orlandini, S. (2005). Neural network for the estimation of leaf wetness duration: application to a Plasmopara viticola infection forecasting. Physics and Chemistry of the Earth, Parts A/B/C, 30(1), 91-96.

Davis, D., & Hughes, J. E. (1970). A new approach to recording the wetting parameter by the use of electrical resistance sensors. Plant Disease Reporter, 54(6), 474-479.

de Carvalho Alves, M., Pozza, E. A., do Bonfim Costa, J. d. C., de Carvalho, L. G., & Alves, L. S. (2011). Adaptive neuro-fuzzy inference systems for epidemiological analysis of soybean rust. Environmental Modelling & Software, 26(9), 1089-1096.

Devices, D. (2006). LWS Leaf Wetness Sensor. Retrieved June 16, 2016, from https://www.decagon.com/en/canopy/canopy-measurements/lws-leaf-wetness-sensor/

Francl, L., & Panigrahi, S. (1997). Artificial neural network models of wheat leaf wetness. Agricultural and forest meteorology, 88(1), 57-65.

Fisher, D., & Wicks, T. (2003). Powdery Mildew in Wine Grapes in Western Australia. (4575).

Garratt, J., & Segal, M. (1988). On the contribution of atmospheric moisture to dew formation. Boundary-Layer Meteorology, 45(3), 209-236.

Gillespie, T., & Barr, A. (1984). Adaptation of a dew estimation scheme to a new crop and site. Agricultural and forest meteorology, 31(3), 289-295.

Gillespie, T., & Kidd, G. (1978). Sensing duration of leaf moisture retention using electrical impedance grids. Canadian Journal of Plant Science, 58(1), 179-187.

Gillespie TJ, Srivastava B, Pitblado RE (1993) Using operational weather data to schedule fungicide sprays on tomatoes in southern Ontario, Canada. J Appl Meteorol 32:567–573

Gleason, M., Taylor, S., Loughin, T., & Koehler, K. (1994). Development and validation of an empirical model to estimate the duration of dew periods. Plant disease (USA), 78, 1011-1016.

Gleason, M.L., Duttweiler, K.B., Batzer, J.C., Taylor, S.E., Sentelhas, P.C., Boffino Almeida Monteiro, J.E., Gillespie, T.J., 2008. Obtaining Weather Data for Input to Crop Disease-Warning Systems: Leaf Wetness Duration as a Case Study. *Sci.Agric*. 65, 76–87.

Hall, M. A. (1999). Correlation-based feature selection for machine learning (Doctoral dissertation, The University of Waikato).

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. ACM SIGKDD explorations newsletter, 11(1), 10-18.

Heilman, J., McInnes, K., Savage, M., Gesch, R., & Lascano, R. (1994). Soil and canopy energy balances in a west Texas vineyard. *Agricultural and Forest Meteorology*, 71(1), 99-114.

Henshall, W. R., Hill, G. N., Beresford, R. M., & Zydenbos, S. M. (2015). Comparison of measured and modelled wetness duration as inputs to a grape disease model. New Zealand Plant Protection, 68, 405-410.

Hirst, J. (1954). A method for recording the formation and persistence of water deposits on plant shoots. *Quarterly Journal of the Royal Meteorological Society*, 80(344), 227-231.

Huber, L., & Gillespie, T. (1992). Modeling leaf wetness in relation to plant disease epidemiology. *Annual review of Phytopathology*, 30(1), 553-577.

Hughes, R., & Brimblecombe, P. (1994). Dew and guttation: formation and environmental significance. *Agricultural and Forest Meteorology*, 67(3), 173-190.

Idso, S. B., & Jackson, R. D. (1969). Thermal radiation from the atmosphere. *Journal of Geophysical Research*, 74(23), 5397-5403. doi:10.1029/JC074i023p05397

Jacobs, A. F., Heusinkveld, B. G., & Berkowicz, S. M. (2002). A simple model for potential dewfall in an arid region. Atmospheric Research, 64(1), 285-295.

Jang, J.-S. R. (1993). ANFIS: adaptive-network-based fuzzy inference system. *Systems, Man and Cybernetics, IEEE Transactions on*, 23(3), 665-685.

Jang, J.-S. R., Sun, C.-T., & Mizutani, E. (1997). Neuro-fuzzy and soft computing; a computational approach to learning and machine intelligence. New Jersey, United States: Prentice Hall.

Jones, A. (1986). Role of wet periods in predicting foliar diseases. *Plant disease epidemiology*, 1, 87-100.

Kim, K., Taylor, S., & Gleason, M. (2004). Development and validation of a leaf wetness duration model using a fuzzy logic system. *Agricultural and Forest Meteorology*, 127(1), 53-64.

Kim, K., Taylor, S., Gleason, M., & Koehler, K. (2002). Model to enhance site-specific estimation of leaf wetness duration. *Plant disease*, 86(2), 179-185.

Kim, K.-S. (2003). Prediction of leaf wetness duration using a fuzzy logic system. Retreived 17 June, 2016, from http://lib.dr.iastate.edu/rtd/723/

Klir, G., & Yuan, B. (1995). *Fuzzy sets and fuzzy logic* (Vol. 4): Prentice Hall New Jersey.

Khurana, H. S., Phillips, S. B., Alley, M. M., Dobermann, A., Sidhu, A. S., & Peng, S. (2008). Agronomic and economic evaluation of site-specific nutrient management for irrigated wheat in northwest India. *Nutrient Cycling in Agroecosystems, 82*(1), 15-31.

Kudinha, M. T. (2014). Leaf Wetness Duration Measurements in a Citrus Canopy. University of the Free State Bloemfontein, South Africa.

Lau, Y. F., Gleason, M. L., Zriba, N., Taylor, S. E., & Hinz, P. N. (2000). Effects of coating, deployment angle, and compass orientation on performance of electronic wetness sensors during dew periods. *Plant Disease*, 84(2), 192-197.

Linacre, E. (1992). *Climate data and resources: a reference and guide*: Psychology Press.

Lomas, J. (1991). Sprinkler irrigation and plant disease under semi-arid climatic conditions1. *EPPO Bulletin*, 21(3), 365-370.

Lomas, J., & Shashoua, Y. (1970). The performance of three types of leaf-wetness recorders. *Agricultural Meteorology*, 7, 159-166.

Luo, W., & Goudriaan, J. (2000). Dew formation on rice under varying durations of nocturnal radiative loss. *Agricultural and Forest Meteorology*, 104(4), 303-313.

Madeira, A., Kim, K., Taylor, S., & Gleason, M. (2002). A simple cloud-based energy balance model to estimate dew. *Agricultural and Forest Meteorology*, 111(1), 55-63.

Magarey, P. A. (2010). Managing Downy Mildew Fact Sheet. Retreived November 10, 2016, from http://research.wineaustralia.com/wp-content/uploads/2012/09/2010-03-FS-Managing-Downy-Mildew.pdf

Magarey, R., Russo, J., & Seem, R. (2006). Simulation of surface wetness with a water budget and energy balance approach. *Agricultural and Forest Meteorology*, 139(3), 373-381.

Magarey, R., Russo, J., Seem, R., & Gadoury, D. (2005). Surface wetness duration under controlled environmental conditions. *Agricultural and Forest Meteorology,* 128(1), 111-122.

Magarey, R., Seem, R., Weiss, A., Gillespie, T., & Huber, L. (2005). Estimating surface wetness on plants. Retreived June 17, 2016, from http://digitalcommons.unl.edu/agronomyfacpub/696/

Montone, V. O., Fraisse, C. W., Peres, N. A., Sentelhas, P. C., Gleason, M., Ellis, M., & Schnabel, G. (2016). Evaluation of leaf wetness duration models for operational use in strawberry disease-warning systems in four US states [journal article]. International Journal of Biometeorology, 1-14. doi:10.1007/s00484-016-1165-4

McKendry, I. G. (2002). Evaluation of artificial neural networks for fine particulate pollution (PM10 and PM2. 5) forecasting. *Journal of the Air & Waste Management Association*, 52(9), 1096-1101.

Melching, J., Dowler, W., Koogle, D., & Royer, M. (1989). Effects of duration, frequency, and temperature of leaf wetness periods on soybean rust. *Plant Disease*, 73(2), 117-122.

Mellit, A., Arab, A. H., Khorissi, N., & Salhi, H. (2007). An ANFIS-based forecasting for solar radiation data from sunshine duration and ambient temperature. In Power Engineering Society General Meeting, 2007. IEEE (pp. 1-6). IEEE.

Monteith, J., & Unsworth, M. (2013). P*rinciples of Environmental Physics: Plants, Animals, and the Atmosphere*: Academic Press.

Morton, V., & Staub, T. (2008). A Short History of Fungicides. APSnet Features. Retrieved, June 17, 2016, from http://www.apsnet.org/publications/apsnetfeatures/Pages/Fungicides.aspx

Noilhan, J., & Planton, S. (1989). A simple parameterization of land surface processes for meteorological models. *Monthly Weather Review*, 117(3), 536-549.

Norman, J., & Campbell, G. (1983). Application of a plant-environment model to problems in irrigation. *Advances in irrigation*, 2 (1983), pp. 155–188.

Oke, T. R. (1988). Street design and urban canopy layer climate. *Energy and Buildings*, 11(1), 103-113.

Pabreja, K. (2011). An Adaptive Neuro-Fuzzy Inference System based on Vorticity and Divergence for Rainfall forecasting. I*nternational Journal of Computer Science and Information Security*, 9(12), 45.

Papastamati, K., McCartney, H., & Van den Bosch, F. (2004). Modelling leaf wetness duration during the rosette stage of oilseed rape. *Agricultural and Forest Meteorology*, 123(1), 69-78.

Pavan, W., Fraisse, C.W., Peres, N.A., 2011. Development of a web-based disease forecasting system for strawberries. *Comput. Electron. Agric.* 75, 169–175.

Pedro, M., & Gillespie, T. (1982 (a)). Estimating dew duration. I. Utilizing micrometeorological data. *Agricultural Meteorology*, 25, 283-296.

Pedro, M., & Gillespie, T. (1982 (b)). Estimating dew duration. II. Utilizing standard weather station data. *Agricultural Meteorology*, 25, 297-310.

Ramirez, M. C. V., de Campos Velho, H. F., & Ferreira, N. J. (2005). Artificial neural network technique for rainfall forecasting applied to the Sao Paulo region. Journal of Hydrology, 301(1), 146-162.

Rao, P., Gillespie, T., & Schaafsma, A. (1998). Estimating wetness duration on maize ears from meteorological observations. *Canadian Journal of Soil Science*, 78(1), 149-154.

Rotem, J., Cohen, Y., & Bashi, E. (1978). Host and environmental influences on sporulation in vivo. *Annual review of Phytopathology*, 16(1), 83-101.

Rowlandson, T., Gleason, M., Sentelhas, P., Gillespie, T., Thomas, C., & Hornbuckle, B. (2015). Reconsidering Leaf Wetness Duration Determination for Plant Disease Management. *Plant Disease*, 99(3), 310-319.

Rowlandson, T. L. (2011). Leaf wetness: implications for agriculture and remote sensing. Graduate Theses and Dissertations. Paper 12054. http://lib.dr.iastate.edu/etd/12054

Royle, D., & Butler, D. (1986). Liquid water in crop canopies and its role in disease forecasting. *Water, fungi, and plant*s, 11, 139.

Schuepp, P. H. (1989). Microstructure, density and wetness effects on dry deposition to foliage. *Agricultural and Forest Meteorology*, 47(2), 179-198.

Sellers, P., & Lockwood, J. (1981). A computer simulation of the effects of differing crop types on the water balance of small catchments over long time periods. *Quarterly Journal of the Royal Meteorological Society*, 107(452), 395-414.

Sentelhas, P., Monteiro, J., & Gillespie, T. (2004). Electronic leaf wetness duration sensor: why it should be painted. *International Journal of Biometeorology*, 48(4), 202-205.

Sentelhas, P. C., & Gillespie, T. J. (2008). Estimating hourly net radiation for leaf wetness duration using the Penman-Monteith equation. *Theoretical and Applied Climatology*, 91(1-4), 205-215.

Sentelhas, P. C., Gillespie, T. J., Batzer, J. C., Gleason, M. L., Monteiro, J. E. B., Pezzopane, J. R. M., & Pedro Jr, M. J. (2005). Spatial variability of leaf wetness duration in different crop canopies. *International Journal of Biometeorology*, 49(6), 363-370.

Sentelhas, P. C., Gillespie, T. J., Gleason, M. L., Monteiro, J. E. B., Pezzopane, J. R. M., & Pedro, M. J. (2006). Evaluation of a Penman–Monteith approach to provide "reference" and crop canopy leaf wetness duration estimates. *Agricultural and Forest Meteorology*, 141(2), 105-117.

Van der Wal, A. (1978). Moisture as a factor in epidemiology and forecasting. *Water Deficit and Plant Growth*, 5, 253-295.

Weiss, A. (1990). Leaf wetness: Measurements and models. *Remote Sensing Reviews*, 5(1), 215-224. doi:10.1080/02757259009532130

Wieland, R., Mirschel, W., Groth, K., Pechenick, A., & Fukuda, K. (2011). A new method for semi-automatic fuzzy training and its application in environmental modeling. *Environmental Modelling & Software*, 26(12), 1568-1573.

WMO. (2008). Guide to Meteorological Instruments and Methods of Observation (Seventh Edition) WMO No. 8. Geneva, Switzerland: Secretariat of the World Meteorological Organization. Retreived June 17, 2016, from https://www.wmo.int/pages/prog/.../CIMO_Guide-7th_Edition-2008.pdf

Yarwood, C. (1978). Water and the infection process. *Water Deficit and Plant Growth*, 5, 141-165.

Zadeh, L. A. (1994). Fuzzy logic, neural networks, and soft computing. *Communications of the ACM*, 37(3), 77-84.

Zadeh, L. A., & Yager, R. R. (2012). *An introduction to fuzzy logic applications in intelligent systems,* Vol. 165: Springer Science & Business Media.

APPENDIX A

Variables and stations

Here is a list of models used in this research and the variables we used in the experiments as the model's input. Please refer to section 2.2.1 for model naming convention.

Table XXI. Leaf wetness models variable requirements

Model	Variable						
	T	RH	WS	Rainfall	DPD	Net Rad.	VPD
ANFIS			✓			✓	✓
FLS			✓			✓	✓
ANN			✓			✓	✓
NHRH>90%		✓					
CART	✓	✓	✓		✓		
SWEB	✓	✓	✓	✓		✓	
P-M	✓					✓	✓

Variables included: Temperature (T), relative humidity (RH), wind speed (WS), rainfall, net radiation (Net Rad.), and vapour pressure deficit (VPD). These are the variables that are used as input for leaf wetness models. The first four were available as measurements/observations from our dataset. The remaining variables were derived from available meteorological variables.

Here is the map of all the stations that are in dataset A:

Figure XXVIII. Maps showing station locations: (left) five stations in Dataset A, (right) three stations in Dataset B.

APPENDIX B

Normality tests for the combined stations for datasets A and B. Using Kolmogorov-Smirnov test because the combined datasets are both larger than $N = 2000$. The null hypothesis is normality if you accept the null hypothesis then you assume normality. Guidelines suggest that of $p <$ then 0.05 then you should reject the null hypothesis and the data is not normally distributed. It can be seen that for all the variables in each dataset the significant $p < 0.0001$ and therefore if the data from the stations is combined none of the variables have a normal distribution.

	Kolmogorov-Smirnov[a]		
	Statistic	df	Sig.
Temp	.024	7128	.000
RH	.068	7128	.000
Rainfall	.480	7128	.000
Wind speed	.100	7128	.000

a. Lilliefors Significance Correction

	Kolmogorov-Smirnov[a]		
	Statistic	df	Sig.
Temp	.031	8679	.000
RH	.079	8679	.000
Rainfall	.466	8679	.000
Wind speed	.078	8679	.000

a. Lilliefors Significance Correction