Agriculture spray wind velocity measurements and predictions

by

Matthew W. Schramm

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Agricultural and Biosystems Engineering

(Advanced Machinery Engineering)

Program of Study Committee:

Brian Steward, Major Professor
Matt Darr
Steven Hoff

Iowa State University

Ames, Iowa

2015

Dedication

To my friends and family

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

ACKNOWLEDGMENTS

I have many people that I would like to thank that have helped me throughout my time at Iowa State University. I, unfortunately, am not the best at expressing myself. I would like to express my sincere thanks to Dr. Hanna for his support, guidance, knowledge, and his many suggestions. He continuously drove me to do my best and has helped in countless ways throughout my research. I cannot thank him enough for being willing to help at a moment's notice.

I would also like to thank Dr. Steward for his many suggestions throughout my research and his comments on my thesis. I thank Dr. Darr for his numerous suggestions relating to my research and for his help filtering the first seasons wind data. I would also like to thank Dr. Hoff for his suggestions and feedback of my research and help with instrumentation selection.

I thank my many colleagues, (Katherine Hinkle, Jafni Jiken, Safal Kshetri, Zach Vanderleest, and Dillon Wirth), for allowing me to bounce ideas off of them, for helping me write, and for allowing a fun and interactive learning environment. I thank Brittany Dueker for dealing with me throughout my research and writing, her great patience and willingness to help, listening to my many ideas, (even if she was pretending to listen), and for understanding that sometimes I could not come out and watch Dr. Who with her.

Lastly, I would like to thank AGCO for their support in my research and allowed me the opportunity to continue my academic career.

ABSTRACT

During the spraying seasons of 2014 and 2015, wind velocity and solar radiation (2014 only) were collected at a one meter height above the ground to simulate conditions affecting droplets near a ground-based spray boom. This instrumentation was placed in a cross pattern with sensors at the four cardinal directions (north, south, east, and west) with a fifth sensor in the center (2015 only). Data were collected at 10 Hz to measure the turbulent properties of the wind near the ground.

Measurements of wind velocity profiles moving from upwind sensors to downwind sensors were used to evaluate correlation between the wind measurements. Two periods in which wind direction, on average, was collinear with multiple sensors were investigated. The first period contained five hours of data in which the average wind speed was 3.6 m/s (8 mi/h), while the second period contained 1.5 hours of data with an average wind speed of 1.5 m/s (3.4 mi/h). For the five hour dataset, correlation coefficients of 0.29 and 0.27 were found for wind direction and wind speed measured at two sensors respectively. This value fell when the five hours were broken up into multiple one minute periods. The correlation coefficients rose from less than 0.03 to greater than 0.14 once a lag term was introduced to the data. These results were not observed in the 1.5 hour dataset. Over the 1.5 hour period, the correlation coefficients were found to be less than 0.03. The introduction of a lag term had no clear effect.

The entirety of the datasets that were collected in 2014 and 2015 were investigated to see under what conditions large wind change events were more likely to occur. The datasets suggest that low wind speeds lead to higher probability of large wind changes. As solar radiation increased so did the probability of large changes in wind. As a tolerance

on the wind shift was tightened, the probability of wind changes became uniform.

In models that predict spray drift, a popular method to simulate turbulent wind conditions in which the droplet is entrained, is to update the current wind velocities with a random process to achieve new wind velocities. This type of process is known as a random walk. The random walk hypothesis was tested using data collected at 10 Hz, and the average of the collected data to simulate data recorded at 0.5 s, 1 s, 5 s, 10 s, 30 s, 1 min, 5 min, and 10 min. For all tests below five minute averages, the test rejected the hypothesis that wind velocity updates can be independent of previous measurements at greater than 95% confidence. Indicating that updates to the current wind velocity is dependent on previous velocities.

To help reduce the chances of spray drift, prediction models were developed and tested to predict wind direction 30 seconds into the future utilizing current and past measurements. The models tested included a kernel filter that is used for prediction of wind speeds for wind turbines, an autoregressive process (AR), a full ARIMA process, and a hybrid model that includes ideas from ARIMA and Taylor series expansions. The listed models were tested against a “No Model” model in which the predicted value was simply the current observed value. Models were trained over a one hour dataset and tested over a four hour data set. The AR and hybrid models lowered the RMS error value by 9% over the “No Model” model. The AR and hybrid models were outside of a 20 degree tolerance about 12% of the time.

The correlation values between an upwind and downwind sensors indicate that little correlation exists. Along with the predictive models yielding limited results indicate that the wind changes rather randomly. However, results from testing the time series against

the random walk hypothesis indicate that wind's random fluctuations are correlated with one another, but these correlations are not seen using linear correlations. Further effort is needed to better model the wind process.

CHAPTER 1. INTRODUCTION

Agricultural sprayers are used to provide agricultural chemicals that protect and improve crop plant health. As agricultural chemicals improved, so have yields (Stone, 2008). While these chemicals may protect and/or keep the crop plant healthy, they may be detrimental to adjacent crops or other forms of life by the off target spray drift. The Environmental Protection Agency (EPA) defines spray drift as; “the physical movement of pesticide through air at the time of application or soon thereafter, to any site other than that intended for application” (EPA, 2014). The EPA has set rules and regulations to which chemical manufacturers must adhere, i.e. the manufacturer must pass several risk assessments, such as aggregate risks, cumulative risks, and occupational risks (EPA, 2015). It is then up to the applicator to adhere to the manufacturer’s label when spraying. Not doing so results in fines, lawsuits, and suspensions. It is in the applicator’s best interest to understand what factors contribute most to spray drift.

Many factors affect how spray droplets are transported over a field surface such as topography of the surrounding land, current weather conditions (Nuyttens, et al., 2006; Stull, 2009), size droplet distribution, nozzle pressure, and the height that the droplets are released (Nordby & Skuterud, 1974; Smith, et al., 1982; Spray Drift Task Force, 1997). To mitigate spray drift, models are available that predict the movement of the droplet across the field. Popular models include Gaussian diffusion, plume, regression, random walk, and computational fluid dynamics (Baetens, et al., 2007; Frederic, et al., 2009; Holterman, et al., 1997; Teske, et al., 2002; Thompson & Ley, 1983; Zhu, et al., 1995). These models all have differing degrees of accuracy that depend on model assumptions and ability to measure or estimate parameters. An example is the plume model that has

poor performance at distances close to the nozzle but does agree with observations, once the plume of droplets has propagated some distance away from the point of release (Craig, 2004).

Regression Models

Regression models are statistical models developed from the data collected in the field or in a wind tunnel (Smith, et al., 1982). Smith developed regression models in which 95% and 99% of the spray was deposited as a function of multiple variables, such as horizontal wind velocity, nozzle height, nozzle tilt angle, and droplet kinetic energy. Smith's regression models produced high coefficients of determination, ranging from 65.6% to 90.2%. Regression can be very accurate when similar weather and operational conditions are present. If current conditions are not similar to when the model was developed, estimates/interpolation can be done to relate the current conditions to the conditions present during model development. However, accuracy can be limited when conditions fall outside of those associated with the original dataset.

Plume Models

Plume models approximate spray drift by treating individual droplets as a cloud of gas that follow Gaussian diffusion principles. Plume models were first introduced to describe air pollutants from factories in the 1930's (Bosanquet & Pearson, 1936). Plume models perform well for aerial applications and can also be used for ground sprayers. Models can predict spray drift up to 10 km away, but since models do not take into account conditions at the sprayer, they are not suitable for short range prediction (Kruckeberg, 2011).

Random Walk Models

Random walk models were first introduced when the botanist Brown described his *Brownian Motion* (Brown, 1828). From this beginning, random walk models have spread to multiple disciplines in biology, chemistry, physics, and economics (Codling, et al., 2008). A simple example of a one-dimension random walk is to stand in a hallway and flip a coin. If heads take a step forward, else take a step back. Random walks are used commonly for ground sprayers and track individual droplets until the water within the droplet completely evaporates or the droplet is deposited. A random walk model only incorporates a droplets' current conditions to update the droplets' position and speed. These droplets are tracked in a Lagrangian manner (meaning the droplets position and velocity are tracked) along short time steps in which physical phenomenon (gravity, air resistance, evaporation, etc.) act upon the droplet plus an additional statistical property that introduces random fluctuations. Random walks begin to be no longer useful when the time step is large and when there is high turbulence. The random walk model has been used to develop both AgDRIFT and DRIFTSIM (Teske, et al., 2002; Zhu, et al., 1995) which are commonly used models to predict or estimate spray drift.

Though the random walk model has been in use for some time, a method to test if a time series past values play no role in the updating process, by the variance of the series, was not available until the late 1980's (Charles & Darné, 2009; Lo & MacKinlay, 1988). Before this approach, time series models had to be simulated and compared to the original dataset (Lo & MacKinlay, 1988). The variance ratio test is a statistical test that has been prominently used in finance to test if financial markets could be predicted, specifically if current knowledge of the market can be useful to predict future changes.

The model is especially well-crafted to test whether a process is mean reverting, (if the process steps away from the mean path, it will eventually return to the mean path). Lo's and MacKinlay's (1988) variance ratio test can be used for both homoscedastic (the error terms are identically and independently distributed (the error terms come from the same distribution with same mean and variance)) and heteroscedastic (the error terms are independently distributed).

Computational Fluid Dynamics (CFD)

CFD models make use of the physical mathematical description of a droplet and the fluid in which it is entrained and integrates the first principle equations describing the two phase flow numerically. Examples of equations used are the Navier-Stokes equations and the Reynolds-averaged Navier-Stokes equations. These models can be very accurate dependent upon model assumptions, and the method used to solve the differential equation (Baetens, et al., 2007; Burden & Faires, 2011; Griffiths & Higham, 2010; Haberman, 2013). A computational problem exists when solving the equations at very high precisions. As the grid spacing is reduced, computational time increases. The integration time steps and spatial gridding must be small to insure all interactions are calculated properly. This can lead to a very high computational burden leading to long simulation times. To get around this challenge, simplifications of the mathematical description of the droplet/fluid can be made to simplify the computational complexity of the solver (Stull, 2009). A popular method used to simplify the integration step, is to use Forward Euler integration (Baetens, et al., 2007). To see how this affects the simulation, a popular program used to solve ordinary differential equations (ODE) is MATLAB (MathWorks). MATLAB's default ODE solver is "ode45" (MathWorks, 2015) which uses

the Dormand-Prince Runge-Kutta method (Griffiths & Higham, 2010). This method must evaluate the function(s) being integrated at least six times per step, whereas the Forward Euler method only evaluates the function once per step. This could reduce the computation time by a factor of six at the cost of more error being introduced.

Wind Turbulence Updating

Models that incorporate wind velocities do so either by collecting data at a single location and using its value throughout the simulation (Frederic, et al., 2009; Tsai, et al., 2005), or an averaged wind velocity is given random disturbances to simulate the turbulent nature of wind (Holterman, et al., 1997). These recordings are also done at intervals much greater than what the simulations use, (simulation at sub-second time steps using data recorded once per minute). This can introduce more error for the simulation. Little data currently exists that accurately represents what a droplet would experience at boom height (one-meter above the ground) and at sub-second sample times.

Prediction Methods

There is potential to use sprayer models and on-board computers on the sprayer, to perform simulations in real-time, and estimate spray drift potential. This gives the machine the means to mitigate spray drift (Brown, et al., 2004; Craig, 2004; Kruckeberg, 2011). The models listed have varying degrees of accuracy, but all rely on knowledge of current conditions. A sudden and unforeseen shift in wind velocity may defeat this strategy of reducing spray drift, if the shift occurs after the droplet has left the nozzle. However, if wind velocities could be predicted to some degree into the future from the time of droplet release, real-time predictive methods could be used to reduce spray drift

with a low probability of drift into sensitive areas.

Prediction methods are currently being used in the wind power industry to provide energy companies estimates of total power output. These methods are not directly transferrable to agricultural sprayers, because the prediction methods predict hours into the future (Giebel, 2003). For use in agriculture, prediction times less than one minute are desirable. One of the popular methods in the wind power industry is the Nadaraya-Watson kernel-weighted average (Chan, et al., 2010). This statistical method is used due to its ability to accept multiple inputs, such as wind speed, wind direction, solar radiation, pressure, and humidity, to predict the future wind velocity. Another time dependent model that is implemented in wind industries is the Autoregressive Integrated Moving Average (ARIMA) (Giebel, 2003). ARIMA suffers from the same flaws as a regression model, in that similar conditions for the model calibration and prediction are a tacit assumption of the ARIMA model. Wind velocity prediction using a Taylor series expansion, of wind direction, would not require similar conditions; however, the Taylor series technique requires an understanding of how the derivatives of wind velocity behave.

Research Objectives

Understandings of how wind behaves/interacts near an agricultural field surface over short time periods is limited, nor are the turbulent properties of wind well understood. This knowledge gap provided the motivation to pursue the following research objectives:

- Understand changes in transient wind velocity (direction and speed) at a typical ground sprayer boom height off the field surface,
- Evaluate under what conditions wind may be most likely to have a significant velocity change or be more turbulent,

- Evaluate the randomness of measured transient wind velocity changes for use in random walk models as applied to ground sprayers, and
- Determine the probability of a significant wind direction change up to 30 seconds into the future that impacts unforeseen spray drift in ground based sprayers

Thesis organization

Chapter Two describes some of the mathematical background for the methods used in the subsequent chapters. Specifically, variable interpolation methods are described, as is criteria for interpolation method viability. The development of ARIMA/ARIMAX models are outlined, starting with autoregressive and moving average methods, showing how they are complementary to each other. Finally, the Quasi-Newton's Method for unconstrained optimization is introduced by first sketching Newton's Method for finding a zero of a function with one variable. It is then shown the transition from finding the zero of a function to finding the maxima/minima of a function, and then how the method operates with multiple variables (higher dimensions).

Chapter Three which describes research focusing on the experimental design in which wind velocity measurements were acquired in an agricultural field at boom heights. Also included are investigations into the relationship between wind measured at an upwind sensor and a downwind sensor, and if the wind speed can indicate the amount of turbulence present. The chapter also describes the process that was used to investigate under what conditions wind changes may occur. Chapter Four describes the tests of whether wind velocity measurements have the statistical properties associated with a random walk. The random walk model is tested at time intervals ranging from 0.1 seconds

to 10 minutes to see how time steps affect the Random Walk Hypothesis. In Chapter Five, multiple prediction models are developed and tested. These models are compared against the default case in which the prediction value is simply the last recorded value. Chapter Six contains the overall conclusions of the thesis, along with recommendations for future research.

References

Baetens, K., D. Nuyttens, P. Verboven, M. De Schampheleire, B. Nicolai, and H. Ramon. 2007. "Predicting drift from field spraying by means of a 3D computational fluid dynamics model." *Computers and Electronics in Agriculture* 56: 161-173.

Bosanquet, C.H., and J.L. Pearson. 1936. "The Spread of Smoke and Gases from Chimneys." *Transactions of the Faraday Society* 32: 1249-1263.

Brown, R. 1828. "A brief account of microscopical observations made in the months of June, July, August, 1827, on the particles contained in the pollen of plants; and the general existence of active molecules in organic and inorganic bodies." *Philos. Mag.* 4: 161-173.

Brown, R.B., M.H. Carter, and G.R. Stephenson. 2004. "Buffer Zone and Windbreak Effects on Spray Drift Deposition in a Simulated Wetland." *Best Management Science* 60 (11): 1085-1090.

Burden, R. L., and J. D. Faires. 2011. "Numerical Differentiation and Integration: Numerical Differentiation." In *Numerical Analysis*, by R.L. Burden and J.D. Faires, 174-184. Delhi: Brooks/Cole.

Chan, C.P., J.R. Stalker, A. Edelman, and S.R. Connors. 2010. "Leveraging High Performance Computation for Statistical Wind Prediction." *American Wind Energy Association Wind Power Conference and Exibition.*

Charles, A., and O. Darné. 2009. "Variance Ratio Tests of Random Walk: An Overview." *Journal of Economic Surveys* 23 (3): 503-527.

Codling, E.A., M.J. Plank, and S. Benhamou. 2008. "Random Walk Models in Biology." *Journal of the Royal Society Interface* 5: 813-834.

Craig, I P. 2004. "The GDS Model - A Rapid Computational Technique for the Calculation of Aircraft Spray Drift Buffer Distances." *Computers and Electronics in Agriculture* 43 (3): 235-250.

EPA. 2014. "Glossary." *U.S. Environmental Protection Agency.* 02 04. http://www.epa.gov/pesticides/regulating/labels/pest-label-training/glossary/.

EPA. 2015. "Pesticide Registration." *U.S. Environmental Protection Agency.* 10 30. Accessed 11 20, 2015. http://www2.epa.gov/pesticide-registration/about-pesticide-registration.

Frederic, L., A. Verstraete, B. Schiffers, and M.F. Destain. 2009. "Evaluation of Realtime Spray Drift Using RTDrift Gaussian Advection-Diffusion Model." *Commun. Agric. Biol. Sci.* 74 (1): 11-24.

Giebel, G. 2003. *The State-Of-The-Art in Short-Term Prediction of Wind Power.* A Literature Overview, ANEMOS.

Griffiths, D.F., and D.J. Higham. 2010. *Numerical Methods for Ordinary Differential Equations: Initial Value Problems.* Springer.

Haberman, R. 2013. *Applied Partial Differential Equations with Fourier Series and Boundary Value Problems.* Person Education.

Holterman, H.J, J.C Van de Zande, H.A.J Porskamp, and J.F.M Huijsmans. 1997. "Modeling Spray Drift from Boom Sprayers." *Computers and Electronics in Agriculture* 19 (1): 1-22.

Kruckeberg, J.P. 2011. "An automated nozzle controller for self-propelled sprayers." Iowa State University Graduate Theses and Dissertations. Paper 12083. http://lib.dr.iastate.edu/etd/12083.

Lo, A.W., and A.C. MacKinlay. 1988. "Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test." *The Review of Financial Studies* 1 (1): 41-66.

MathWorks. 2015. "ode45." *MathWorks .* Accessed 10 3, 2015. http://www.mathworks.com/help/matlab/ref/ode45.html.

Nordby, A, and R Skuterud. 1974. "The effects of boom height, working pressure and wind speed on spray drift." *Weed Research* (Weed Research) 14 (6): 385-395. doi:10.1111/j.1365-3180.1974.tb01080.x.

Nuyttens, D., M.DE. Schampheleire, W. Steurbaut, K. Baetens, P. Verboven, B. Nicolai, H. Ramon, and B. Sonck. 2006. "Experimental Study of Factors Influencing the Risk of Drift from Field Sprayers, Part 1: Meteorological Conditions." *Aspects of Applied Biology* 77 (2): 321-329.

Smith, D B, F D Harris, and C E Goering. 1982. "Variables affecting drift from ground boom sprayers." *Transactions of the ASAE* 25 (6): 1499-1523.

Spray Drift Task Force. 1997. *A Summary of Ground Application Studies.* Macon, MO: Agricultural Research Services Inc.

Stone, D. 2008. "History of Pesticide Use." *Oregon State University.* Accessed 11 01, 2015. http://people.oregonstate.edu/~muirp/pesthist.htm.

Stull, R.B. 2009. *An Introduction to Boundary layer Meteorology.* 13ed. Springer.

Teske, M.E., S.L. Bird, D.M. Esterly, T.B. Curbishley, S.L. Ray, and S.G. Perry. 2002. "AgDrift: A Model for estimating near-field spray drift from aerial applications." *Environmental Toxicology and Chemistry* 21 (3): 659-671.

Thompson, N., and A.J. Ley. 1983. "Estimating Spray Drift using a Random-walk Model of Evaporating Drops." *Journal of Agricultural Engineering Research* 28: 419-435.

Tsai, M., K. Elgethun, J. Ramaprasad, M.G. Yost, A.S. Felsot, V.R. Hebert, and R.A. Fenske. 2005. "The Washington aerial spray drift study: Modeling pesticide spray drift deposition from an aerrial application." *Atmospheric Environment* 39: 6194-6203.

Zhu, H., D.L. Reichard, R.D. Fox, H.E. Ozkan, and R.D. Brazee. 1995. "DRIFTSIM, A Program to Estimate Drift Distances of Spray Droplets." *Applied Engineering in Agriculture* 11 (3): 365-369.

CHAPTER 2. MATHEMATICAL CONCEPTS

Interpolation

Interpolation is extensively used throughout the data analysis, and multiple methods are available to accomplish this task. Interpolation was needed for sensor synchronization and to create fixed time steps between observed values. The methods listed will interpolate the data set $Y(x)$ which is the image of X, where X is contained in the interval [a,b], which mathematically can be represented as:

($X \in [a,b] \in R = (a = x_0, x_1, \dots, x_{n-1}, x_n = b)$). X will contain as many points needed to carry out the interpolation method. For the research in this thesis, X represents time and Y is a wind property, such as wind direction and wind speed. There are several interpolation methods including Linear, PCHIP, and the Cubic Spline which are described in the following sections.

Linear Interpolation

Linear Interpolation takes the form:

$$Y_{int}(x) = m(x - x_0) + b \tag{1}$$

where

- x_0 is the x-axis value from the set X,
- x is any value contained in the interval $[x_0, x_1]$,
- m is the slope between the points (x_0, y_o) and (x_1, y_1), and
- b is the y-intercept

Linear interpolation's constraint is that at the two endpoints, (x_0 and x_1), the interpolation must be equal to the y-values y_0 and y_1 respectively, ($Y_{int}(x_0) = Y(x_0)$ and $Y_{int}(x_1) =$

$Y(x_1)$). This constraint makes the method zeroth order continuous, C^0, indicating that the interpolant is continuous but not the interpolant's derivatives (MathWorks, 2015).

Piecewise Cubic Hermite Interpolating Polynomial (PCHIP)

PCHIP is a higher order method compared to linear interpolation. PCHIP interpolates between the points x_0 and x_1, same as linear interpolation, and requires at least two points with both Y-values and the first derivative of Y, ((x_0, y_0), (x_0, y_0'), (x_1, y_1), and (x_1, y_1')), or four points, so the derivatives can be calculated numerically using forward and backward differences, ((x_{-1}, y_{-1}), (x_0, y_0), (x_1, y_1), and (x_2, y_2)) to solve the equation:

$$Y_{int} = a + b(x - x_0) + c(x - x_0)^2 + d(x - x_0)^3 \quad (2)$$

Since information of the derivatives are available, PCHIP obtains C^1,(the interpolation and its derivative is continuous, but not necessarily its further derivatives), making it a smoother approximation than linear interpolation at the cost of being slightly more complicated (Burden & Faires, 2011; MathWorks, 2015).

Cubic Spline Interpolation

Cubic spline interpolation produces a smoother interpolant than PCHIP, with both continuous first and second derivatives making it C^2 (the interpolation, the interpolant's derivative and its second derivative are continuous but not necessarily its further derivatives are continuous). Unlike linear and PCHIP interpolants, cubic splines use all of the data (i.e., the entire range) supplied to produce the interpolation function (the Cubic Spline does not interpolate between only two values, like linear and PCHIP interpolation, but it interpolates every point at the same time). Cubic splines have the same form as PCHIP (Equation 2) but includes the extra requirement of continuous second derivatives. Calculating the cubic spline interpolation function involves solving a tridiagonal system of

linear equations. Because of increased complexity, cubic spline interpolation is more costly in terms of computer resources and computer time than PCHIP or linear methods. The cubic spline interpolant does not have quite enough information to compute a unique interpolant. The interpolant must be supplied two extra conditions supplied by the user. Popular methods used to supply the last two conditions (that are user assumed) are the Natural Spline which assumes that the second derivatives at the beginning and end of the data set are zero ($d^2Y_{int}(a) = d^2Y_{int}(b) = 0$), the Clamped Spline which sets the values of the first derivative at the beginning and end of the data set ($dY_{int}(a) = c, dY_{int}(b) = d$), and the Not-a-Knot Spline (function is continuous at $d^3Y_{int}(x_1)$ and $d^3Y_{int}(x_{n-1})$). This relationship assumes that the points $x_1 \text{and } x_{n-1}$ are not true points in the data set. Since a cubic polynomial's third derivative is a constant, there is no break in the spline in these locations. Cubic splines produce a smoother function of interpolated values, however there are limitations. Use of a cubic spline, assumes the second derivative of the data is continuous. Ideally, the third derivative needs to be bounded so that the error of the interpolating function can also be bounded (Burden & Faires, 2011; MathWorks, 2015).

Example using different interpolation methods

The function, $Y(x) = \sin(x)\, e^x$, as an example, is interpolated from 15 points in the interval [0,2] using four different interpolation methods (Linear, PCHIP, Clamped Cubic Spline and Not-a-Knot Cubic Spline Interpolants). Figure(1a) shows the function values and the sampled data points that the methods used for interpolation. The interpolants are not show in a figure because they lie on top of each other and are difficult to separate. The function is infinitely differentiable (derivatives of all orders) and thus all methods can

be used and the Not-a-Knot end conditions for the cubic spline are applicable.

Figure(1b) shows the error associated using each method with respect to the true function value. Linear interpolation yields the worst error. PCHIP yields a better approximation than the clamped cubic spline, while the Not-a-Knot yielded the least error. In this case, the clamped spline did poorly due to the numerical method used to calculate the derivative. As higher order methods to calculate the derivatives are used, better approximations can be made (Figure 2a and 2b). In figure 2b, it is shown that if the user knows the true derivatives at the end points, the clamped spline yields the best approximation. However, the user may not know these values. In these cases it is advised to use the Not-a-Knot method, if it can be assumed that the data being interpolated is C^2.

ARIMA/ARIMAX

In statistics, time dependent variables cannot use the same tools that are used for time independent variables because the currently observed value is often dependent on past values. Two popular methods that are used to analyze time dependent variables include the autoregressive and moving average functions. An autoregressive function takes a linear combination of constant coefficients multiplied by past observed values and a random white noise error term, to give an estimate of the next value (Equation 3).

$$Y_{t+1} = \phi_0 Y_t + \phi_1 Y_{t-1} + \cdots + \phi_n Y_{t-n} + \epsilon_t \quad (3)$$

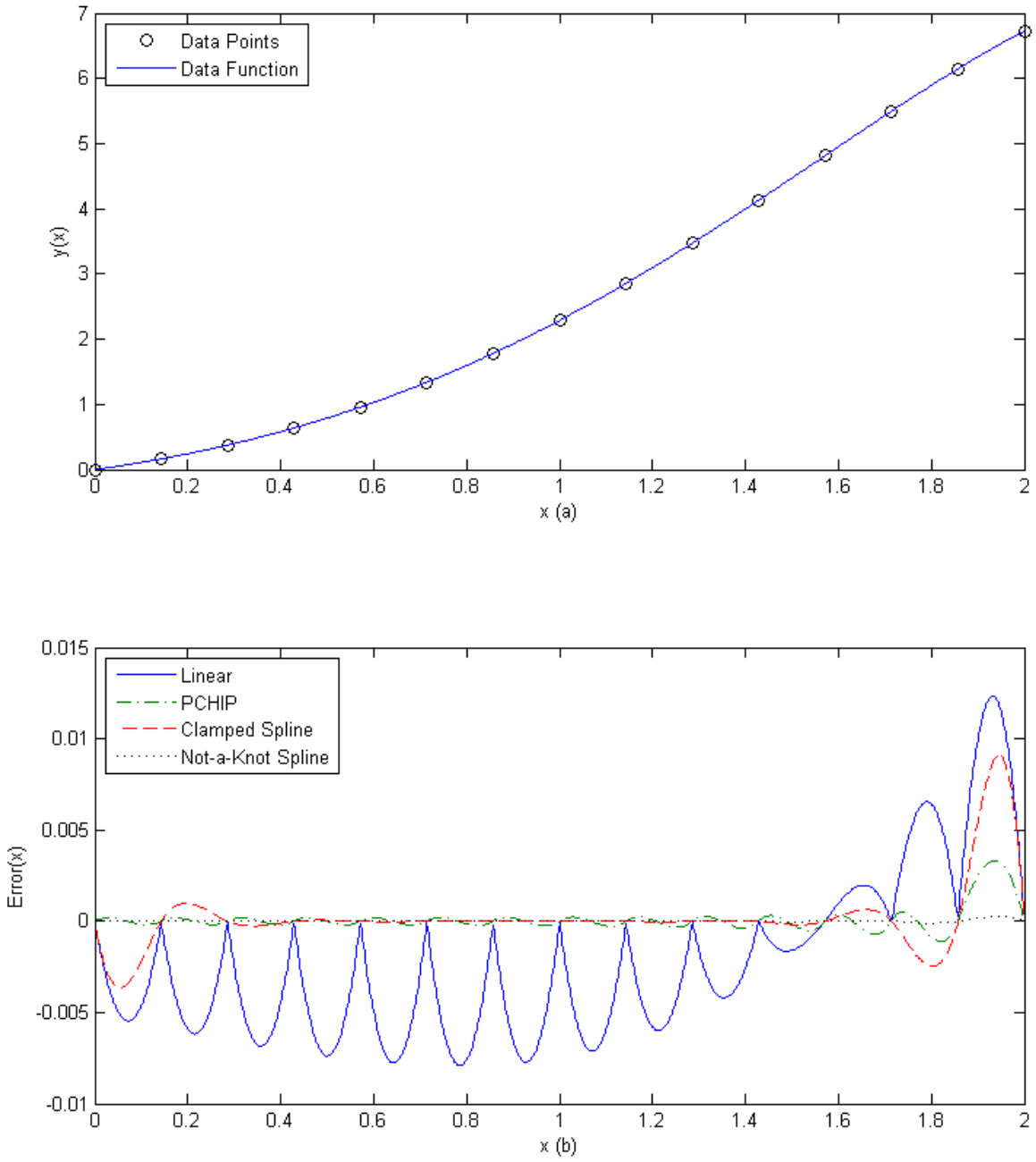

Figure 1: a) The function that is being interpolated and showing which points will be used for interpolation, b) the error of multiple interpolations

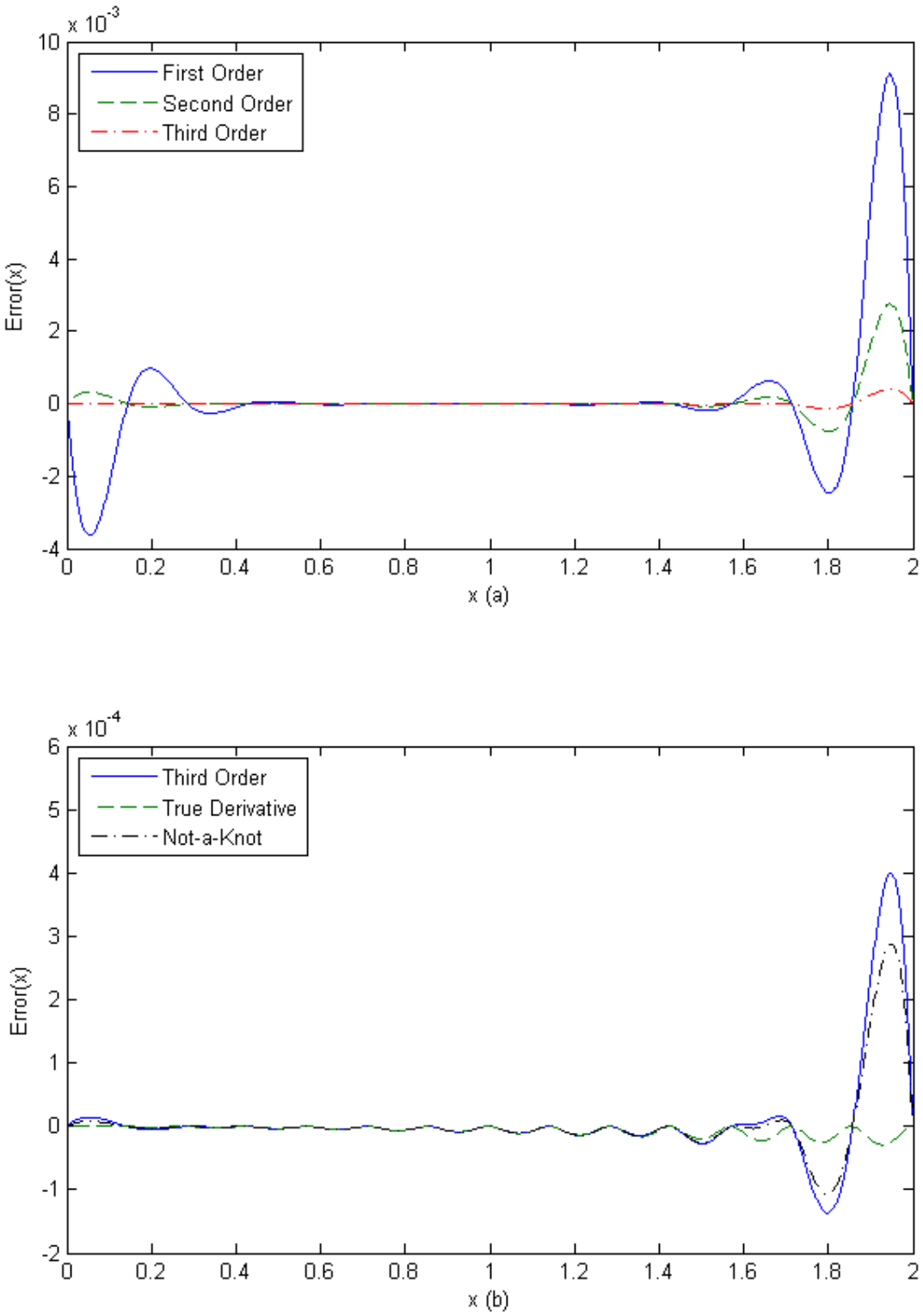

Figure 2: a) Comparison of first, second, and third order methods to find derivatives at the end points for use in clamped splines b) comparison of a third order method to find the derivative at the endpoints, and the true derivative at the endpoints for use in clamped splines versus the Not-a-Knot end conditions

A moving average uses the mean of a process and adds additional white noise error terms to predict the next value (Equation 4). The white noise error terms ($\theta_k \epsilon_k$) are the product of a constant coefficient (θ_k) and the random error in predictions of a series of terms in the immediate past.

$$Y_t = \delta + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_n \epsilon_{t-n} \quad (4)$$

Combining equations 3 and 4, an ARMA(P,Q) (Autoregressive Moving Average) model is formed (Equation 5).

$$\left(1 - \sum_{k=1}^{P} \phi_k L^k\right) Y_t = \delta + \left(1 + \sum_{k=1}^{Q} \theta_k L^k\right) \epsilon_t \quad (5)$$

where

- P and Q are the number of autoregressive and moving average coefficients,
- L^n is the nth order lag operator ($L^n Y_t \equiv Y_{t-n}$),
- δ is the drift term,
- ϵ_t is the time varying error term,
- ϕ_k are coefficients for previous observed data points for which it must be solved, and
- θ_k are coefficients for previous error terms for which it must be solved,

There are times when the time series cannot be modeled as is and the time series difference is needed to be taken, giving the Autoregressive Integrated Moving Average (ARIMA(p,d,q)) model (Equation 6). This differencing function can be thought of taking d^{th} derivative of the time series.

$$\left(1-\sum_{k=1}^{p}\phi_k L^k\right)(1-L)^d Y_t = \delta + \left(1+\sum_{k=1}^{q}\theta_k L^k\right)\epsilon_t \quad (6)$$

where

- Y_t is the time varying data,
- L^n is the nth order lag operator ($L^n Y_t \equiv Y_{t-n}$),
- δ is the drift term,
- ϵ_t is the time varying error term,
- ϕ_k are coefficients for previous observed data points for which it must be solved,
- θ_k are coefficients for previous error terms for which it must be solved,
- p is the number of autoregressive terms,
- q is the number of moving average terms, and
- d is the number of times the difference of the data is taken

ARMA(P,Q) models can be made from ARIMA(p,d,q) if the d in ARIMA is set to zero. Some examples of ARIMA(p,d,q), (that is an ARIMA model utilizing p autoregressive terms, d differencing terms, and q moving average terms), models include the

Random Walk with drift (ARIMA(0,1,0))

$$\left(1-\sum_{k=1}^{0}\phi_k L^k\right)(1-L)^1 Y_t = \delta + \left(1+\sum_{k=1}^{0}\theta_k L^k\right)\epsilon_t$$

$$(1-0)(1-L)Y_t = \delta + (1+0)\epsilon_t \quad (7)$$

$$Y_t = \delta + LY_t + \epsilon_t$$

$$Y_t = \delta + Y_{t-1} + \epsilon_t$$

ARIMA(1,0,1)

$$\left(1-\sum_{k=1}^{1}\phi_k L^k\right)(1-L)^0 Y_t = \delta + \left(1+\sum_{k=1}^{1}\theta_k L^k\right)\epsilon_t$$

$$(1-\phi_1 L)Y_t = \delta + (1+\theta_1 L)\epsilon_t \tag{8}$$

$$Y_t = \delta + \phi_1 L Y_t + \theta_1 L\epsilon_t + \epsilon_t$$

$$Y_t = \delta + \phi_1 Y_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t$$

In practice, the autoregressive, moving average, and drift coefficients are found by maximizing the log-likelihood function instead of minimizing the error norm. By maximizing the log-likelihood function using a training dataset, statistical properties of the coefficients can be found, including a t-statistic that represents if a certain coefficient is needed or not. The log-likelihood is maximized using the Quasi-Newton's Method. The ARIMA(p,d,q) model can incorporate other terms to create an ARIMAX model (Equation 9).

$$\left(1-\sum_{k=1}^{p}\phi_k L^k\right)(1-L)^d Y_t = \delta + \left(1+\sum_{k=1}^{q}\theta_k L^k\right)\epsilon_t + \alpha_0 X_0 + \cdots + \alpha_n X_n \tag{9}$$

This allows additional terms outside of the time series to be accounted for in the prediction of the next value in the time series (MathWorks, 2015). For example, utilizing an ARIMAX model for wind direction (W_d) could incorporate current wind speed (W_s) (Equation 10).

$$\left(1-\sum_{k=1}^{p}\phi_k L^k\right)(1-L)^d W_{d_t} = \delta + \left(1+\sum_{k=1}^{q}\theta_k L^k\right)\epsilon_t + \alpha_0 W_{s_t} \tag{10}$$

Quasi-Newton Method: Unconstrained Optimization

When solving for a variable, there are times in which no closed solution can be reached. An example of such an equation is:

$$\cos(x) = x^3 \tag{11}$$

In which the exact solution cannot be found. Such solutions can be found through the use of non-linear solvers, in particular, Newton's Method. Newton's Method searches for a function's zero, this implies that equation 11 will need to be altered to equation 12.

$$\begin{gathered} \cos(x) = G(x), x^3 = H(x) \\ F(x) = G(x) - H(x) = 0 \end{gathered} \tag{12}$$

The method's procedure consists of the following steps:

1. Guess an initial x-value (x_0) close to where $F(x) = 0$
2. Compute the tangent line at this point $\left(F'(x_0)\right)$
3. Update the solution (Equation 13)

$$x_1 = x_0 - \frac{F(x_0)}{F'(x_0)} \tag{13}$$

4. Repeat until zero is found

Equation 13 is derived from solving for a linear function to find the x-intercepts:

$$\begin{gathered} y = m(x - x_0) + b \\ m = F'(x_0), b = F(x_0) \\ 0 = F'(x_0)(x_1 - x_0) + F(x_0) \\ x_1 = x_0 - \frac{F(x_0)}{F'(x_0)} \end{gathered} \tag{14}$$

Using Newton's Method to find the x-value that satisfies equation 11, the method is repeated five times to find that $x \approx 0.865474033101614$, which is correct to 15 decimal places.

Finding a local minimum or maximum is finding where a functions first derivative equals zero. In the case of this thesis, the function that is being minimized is the RMS-error between a model and the observed data. These models utilize multiple constant coefficients (the variables of the model) that must be solved for. Newton's Method can now be applied on the first derivative of the function (Equation 15).

$$x_{n+1} = x_n - \frac{F'(x_n)}{F''(x_n)} \tag{15}$$

This can easily be altered to allow for multiple variable functions. The process shown in Equation 14 is altered to allow vectors of first derivative functions (Equation 16).

$$\begin{gathered} 0 = \vec{F}''(x_n)(x_{n+1} - x_n) + \vec{F}'(x_n) \\ -\left(\vec{F}''(x_n)\right)^{-1} \vec{F}'(x_n) = x_{n+1} - x_n \\ x_{n+1} = x_n - \left(\vec{F}''(x_n)\right)^{-1} \vec{F}'(x_n) \\ x_{n+1} = x_n - H(x_n)^{-1} J(x_n) \end{gathered} \tag{16}$$

The vector of first derivatives is called a Jacobian ($J(x_n)$) and the matrix of second derivatives is called the Hessian ($H(x_n)$). For Newton's Method to succeed, the Jacobian vector and Hessian matrix must be calculated for every iteration.

When the Jacobian and Hessian are not supplied, they are estimated using numerical approximations. Newton's Method is called the Quasi-Newton's Method when numerical derivatives are used. The methods used for finding these numerical derivatives are, in practice, limited to first and second order methods, i.e. forward,

backward, or central methods. This is done to reduce computational time. At times, computing numerical Hessians becomes too costly and iterative methods are needed (Burden & Faires, 2011; MathWorks, 2015).

References

Burden, R. L., and J. D. Faires. 2011. "Interpolation and Polynomial Approximation: Cubic Spline Interpolation." In *Numerical Analysis*, by R. L. Burden and J. D. Faires, 144-163. Delhi: Brooks/Cole.

Burden, R. L., and J. D. Faires. 2011. "Interpolation and Polynomial Approximation: Hermite Interpolation." In *Numerical Analysis*, by R. L. Burden and J. D. Faires, 136-143. Delhi: Brooks/Cole.

Burden, R. L., and J. D. Faires. 2011. "Numerical Differentiation and Integration: Numerical Differentiation." In *Numerical Analysis*, by R.L. Burden and J.D. Faires, 174-184. Delhi: Brooks/Cole.

Burden, R.L., and J.D. Faires. 2011. "Numerical Solutions of Nonlinear Systems of Equations: Quasi-Newton Method." In *Numerical Analysis*, by R.L. Burden and J.D. Faires, 647-653. Dalhi: Brooks/Cole.

MathWorks. 2015. "arima class." *MathWorks.* Accessed Oct 3, 2015. http://www.mathworks.com/help/econ/arima-class.html.

MathWorks. 2015. "fminunc." *MathWorks.* Accessed Oct 3, 2015. http://www.mathworks.com/help/optim/ug/fminunc.html.

MathWorks. 2015. "interp1." *MathWorks.* Accessed Oct 2, 2015. http://www.mathworks.com/help/matlab/ref/interp1.html.

MathWorks. 2015. "pchip." *MathWorks.* Accessed Oct 2, 2015. http://www.mathworks.com/help/matlab/ref/pchip.html.

MathWorks. 2015. "spline." *MathWorks.* Accessed Oct 2, 2015. http://www.mathworks.com/help/matlab/ref/spline.html.

CHAPTER 3. MEASURING SUB-SECOND WIND VELOCITY CHANGES AT ONE METER ABOVE THE GROUND

A paper to be submitted to Applied Engineering in Agriculture

Matt Schramm[1], Mark Hanna[1], Matt Darr[1], Steven Hoff[1], Brian Steward[1]

Agricultural and Biosystems Engineering, Iowa State University[1]

Abstract

Agricultural spray drift is affected by many factors including current weather conditions, topography of the surrounding area, fluid properties at the nozzle, and the height at which the spray is released. During the late spring/summer spray seasons of 2014 and 2015, wind direction, speed, and solar radiation (2014 only) were measured at 10 Hz. Instrumentation was placed at a one meter height above the ground to simulate conditions affecting droplets near a spray boom. Measurements of wind velocity as wind moved from an upwind sensor to a downwind sensor were used to evaluate under what conditions wind may be most likely to have a significant direction or speed change affecting droplet trajectory.

Linear correlations coefficients between the upwind and downwind sensors during a five hour period with 3.6 m/s (8 mi/h) mean wind speed were 0.29 for wind direction and 0.27 for wind speed. During a 1.5 hour period with 1.5 m/s (3.4 mi/h) mean wind speed, linear correlation coefficients fell to less than 0.03. Over one-minute time periods, correlation coefficients between upwind and downwind sensors were less than 0.03. The introduction of a lag value increased correlation coefficient for the data set with greater wind speed.

For the dataset with greater wind speed, over a 5 hour period, the probability that wind direction at the downwind sensor would be greater than 20 degrees different from the upwind sensor was 30%, and probability of greater than 1 m/s, (one quarter of the mean speed), different than the upwind sensor was 50% of the time. For the dataset with lesser wind speed, over a 1.5 hour period, probability of a reading of greater than 20 degrees different at the downwind sensor was 65%, and probability of a reading of greater than 0.25 m/s, (one quarter of the mean speed), was 80%.

Using data across of a range of dates, changes in wind direction of exceeding 25 degrees occurred most frequently at wind speeds less than 3 m/s (6.7 mi/h), rather than at higher wind speeds. As the day progressed and/or as the solar radiation increased, so did the chances of a wind change.

Keywords. sprayers, spray drift, data collection, wind effects, turbulence, simulation parameters

Introduction

Agricultural sprayers are used to provide agricultural chemicals that protect and improve crop plant health. However, off-site drift of these chemicals can be detrimental to adjacent crops or other forms of life. Many factors affect spray drift ranging from the topography of the land, current weather conditions, size droplet distrobution, and the height at which the droplets are released. The EPA defines spray drift as follows, “Pesticide spray drift is the physical movement of a pesticide through air at the time of application or soon thereafter, to any site other than that intended for application” (EPA, 2014). Spray technology includes capabilities to mitigate spray drift by affecting droplet size for wind conditions (Nordby & Skuterud, 1974; Smith, et al., 1982). Boom height, air temperature, relative humidity, droplet size, droplet release pressure, air pressure, and wind conditions are just some of the factors that control spray drift with droplet size being the biggest contributing factor (Spray Drift Task Force, 1997). The sprayer can control only some of these factors Instrumentation placed on a sprayer may be used to control droplet size dependent upon current weather conditions giving extra control over spray drift, but this strategy may be defeated by unforeseen changes in wind direction or speed after the droplet has left the nozzle.

To better understand how spray drift propagates down wind, simulation models have been developed. Popular methods include Lagrangian, Gaussian, Random Walk, Regression, and CFD models (Holterman, et al., 1997; Baetens, et al., 2007; Teske, et al., 2002; Tsai, et al., 2005; Frederic, et al., 2009; Smith, et al., 1982). Through the use of such models, applicators gain knowledge of when drift potential is high and can adjust buffer zones to minimize the risk for spray drift (Craig, 2004; Brown, et al., 2004). Attention

is given to the development and progression of the droplets, but less attention is devoted to the random nature of weather surrounding the droplet, such as the distribution of wind speed and direction with which the droplet interacts. The models include wind turbulence interactions by either using real data collected at one physical location that is used for the entire range of the simulation (Tsai, et al., 2005; Frederic, et al., 2009), or by using the averaged wind velocity and updating the velocity at every time step with a random fluctuation to simulate the turbulent nature of wind (Holterman, et al., 1997).

Previous research has attempted to incorporate the random nature of wind speed and wind direction, but little work was found seeking to characterize the wind conditions that a droplet experiences from release to deposition. Data recorded for short-term wind velocity changes near the ground's surface are needed to better understand their effects on spray droplet trajectories.

Objectives

The objective of this research was to:

- Understand changes in transient wind velocity (direction and speed) at a typical ground sprayer boom height off the field surface, and
- Evaluate under what conditions wind may be most likely to have a significant velocity change or be more turbulent

Methods and Materials

Experimental Design and Apparatus

Field measurements of wind speed and direction were collected during the late spring/summer spraying season of 2014 and throughout the 2015 season using

instrumentation set into a field of growing oats. The fields were located at the Iowa State University Research Farm's Bruner Farm fields F1 and F3 (respectively for 2014 and 2015) near Ames, Iowa (Figure 3). The field dimensions were 268 m long (north to south) by 105 m wide (east to west) (880 by 348 ft) for field F1 and 107 m long by 201 m wide (350 by 660 ft).

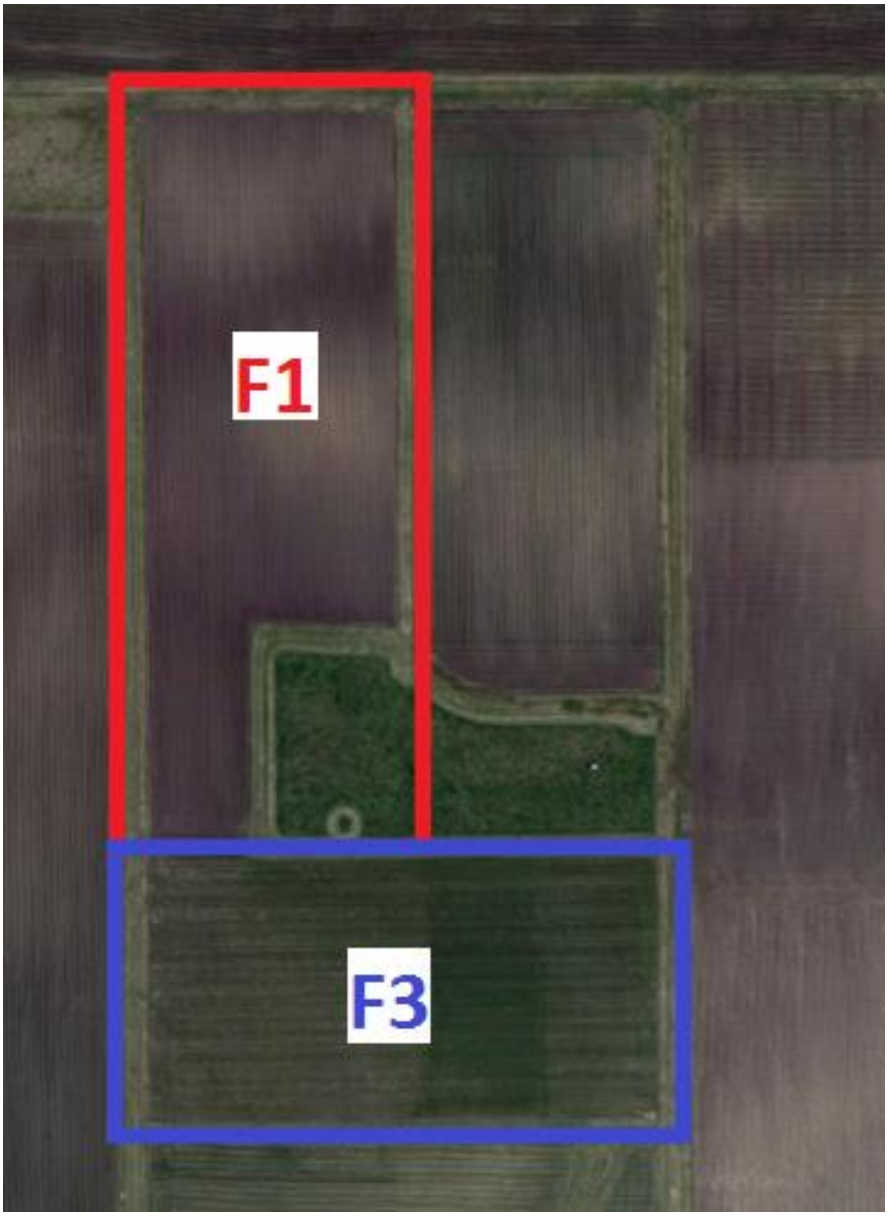

Figure 3: Bruner Farm field F1 (Northern most, 2014) and F3 (Southern most, 2015) (42.014911 N 93.731241 W) (Google, 2015)

Wind speed and solar radiation measurements were acquired at 10 samples per second using ultrasonic anemometers (model: WindMaster 3d, Gill Instruments, Lymington, Hampshire, UK) and a pyranometer (model: SP-212, Apogee Instruments, Logan, UT). The anemometers measured the wind speed in the north-south, east-west, and vertical directions (Figure 4). Open source microcontrollers equipped with a GPS module (model: Arduino Uno, Arduino Inc; Ultimate GPS Shield, Adafruit, New York, USA) were used to log data to micro SD data cards. Using the GPS's PPS (Pulse per Second) output, time correction was able to be done to ensure ok time synchronization of the wind velocity measurements among the microcontrollers. To reduce influences to wind speed, the microcontrollers and power supplies were located separate from the anemometers at a distance of approximately 2-3 meters away. Sensors were placed in a cross pattern with 15.2 m (50 ft) spacing from a center point (Figure 5). Sensor 5 was only used for the 2015 season. Anemometers were placed one meter above the ground's surface to collect wind measurements. The pyranometer was placed on the west most sensor's (Sensor 1) charging station near the anemometer.

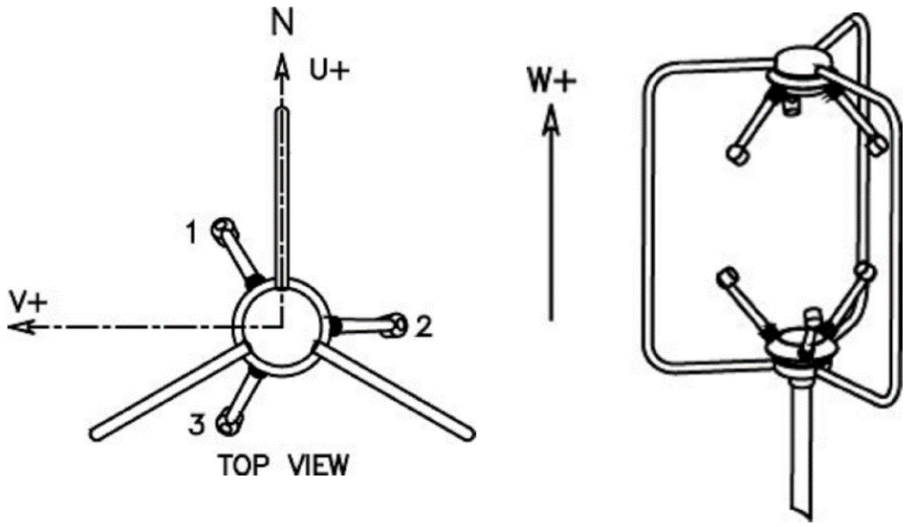

Figure 4: Ultrasonic Anemometer measuring velocity in U, V, and W component directions

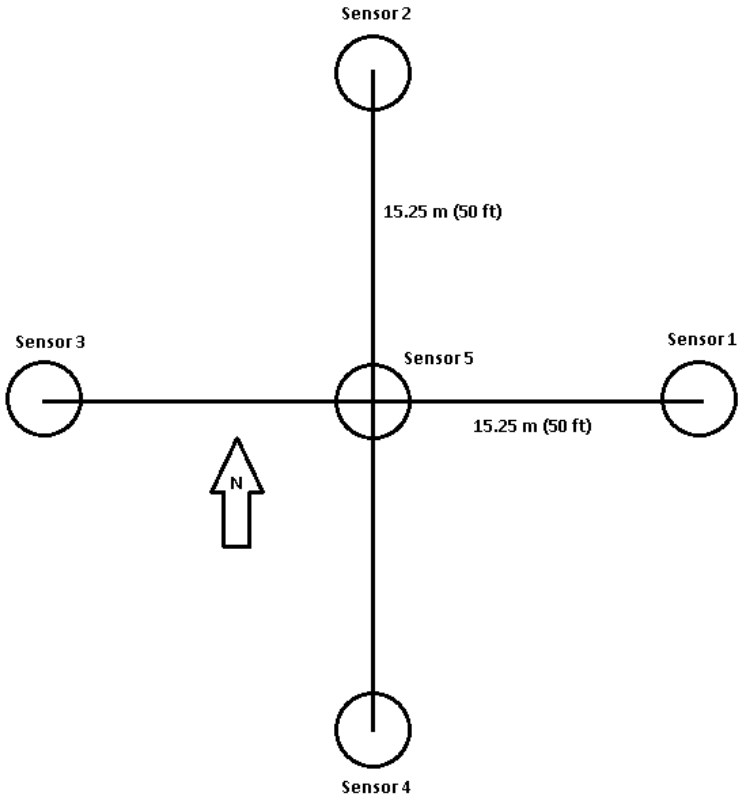

Figure 5: Sensor Positioning

Data Analysis

Wind speed was converted to wind direction using trigonometric relationships. A wraparound method was used to produce a semi-continuous wind direction data set. Wraparound refers to the convention of allowing the wind direction to go above 359 degrees and below 0 degrees, i.e. 12 degrees = 372 degrees.

Because of small amounts of drift in the actual recording time of wind velocity at each sensor location, a form of interpolation was used to estimate wind velocity at points between 10 Hz collection times. This need comes from the data recorders not staying exactly in sync with one another. MATLAB (Version 8.3 (R2014a), MathWorks), used for analysis in this project, offered many different schemes for interpolation ranging from discontinuous to second derivative continuity (C^2). MATLAB's cubic spline offered the

smoothest fit to the data, however the assumption that wind speed and direction have continuous first and second derivatives cannot be confirmed. For this reason MATLAB's pchip (Piecewise Cubic Hermite Interpolating Polynomial) was also not used due to its C^1 (first derivative) continuity. Ultimately, a piecewise linear polynomial was chosen. For linear interpolation, no assumption was needed on the derivatives of the data (Faires, 2011).

To find if wind measurements recorded at an upwind sensor were correlated to measurements at a downwind sensor, shorter length datasets were taken in which the wind direction, on average, would move a particle over multiple sensors. The first dataset consists of a 5 hour period (7:30 am to 12:30 pm) in which the average wind speed was 3.6 m/s (8 mi/h). The second dataset is from a 1.5 hour long period (12:50 pm to 2:20 pm) with a mean wind speed of 1.5 m/s (3.4 mi/h). The entire recorded dataset was then used to determine conditions more likely to be present during significant wind directional changes. This analysis was divided into 2014 and 2015 datasets.

Results and Discussion

Linear correlation

Relationship at individual 0.1 s intervals between sensors over 5 hour period (2014)

Initial analysis checked for a linear correlation between the upwind and downwind sensors. Figure 6 shows the data collected. The average wind speed during this time was 3.6 m/s (8 mi/h). Due to the large number of points, it is unclear graphically if there is a correlation (Figure 7) but the linear correlation of individual 10 Hz measurements between the sensors, for the five hour period, was 0.29 and 0.27 for wind direction and speed, respectively.

To graphically see the structure of the data better, two-dimensional natural logarithmic histograms were used (Figure 8). This was done by creating a two-dimensional mesh of cells of all possible combinations of upwind and downwind sensor data values and then counting the number of data points occurring within each cell of the mesh. Using the natural logarithm of counts inside the mesh reduced the wide range of individual counts in each cell and was done to see smaller structures in the data. To accommodate cell locations with a count of one (natural log of one equals zero), one was added to the value of all cells after the natural log transformation. This step allowed differentiation between cells with an original count of one and cells with no observed (zero) values. The bar graph to the right of the graphs show the shading equivalents to the natural log values. Data grouped around the mean with no apparent linear dependency structure.

Relationship at individual 0.1 second intervals between sensors over 1.5 hour period (2015)

This analysis was also done for a data sample with average wind speed of 1.5 m/s (3.4 mi/h) from 12:50 to 2:20 p.m. during which wind generally came from the south and passed over the southern, central, and northern sensors (Figure 9). The correlation coefficients for wind direction and wind speed for all combinations of two of the three sensors were all below 0.03. Figure 10 shows the logarithmic histograms for the central and northern sensors. Similar figures were attained at the other sensors. An absence of change in general wind direction and speed during the shorter time period in 2015 resulted in the area of highest frequency of values in the histograms being more circular (less elliptically elongated) than in 2014.

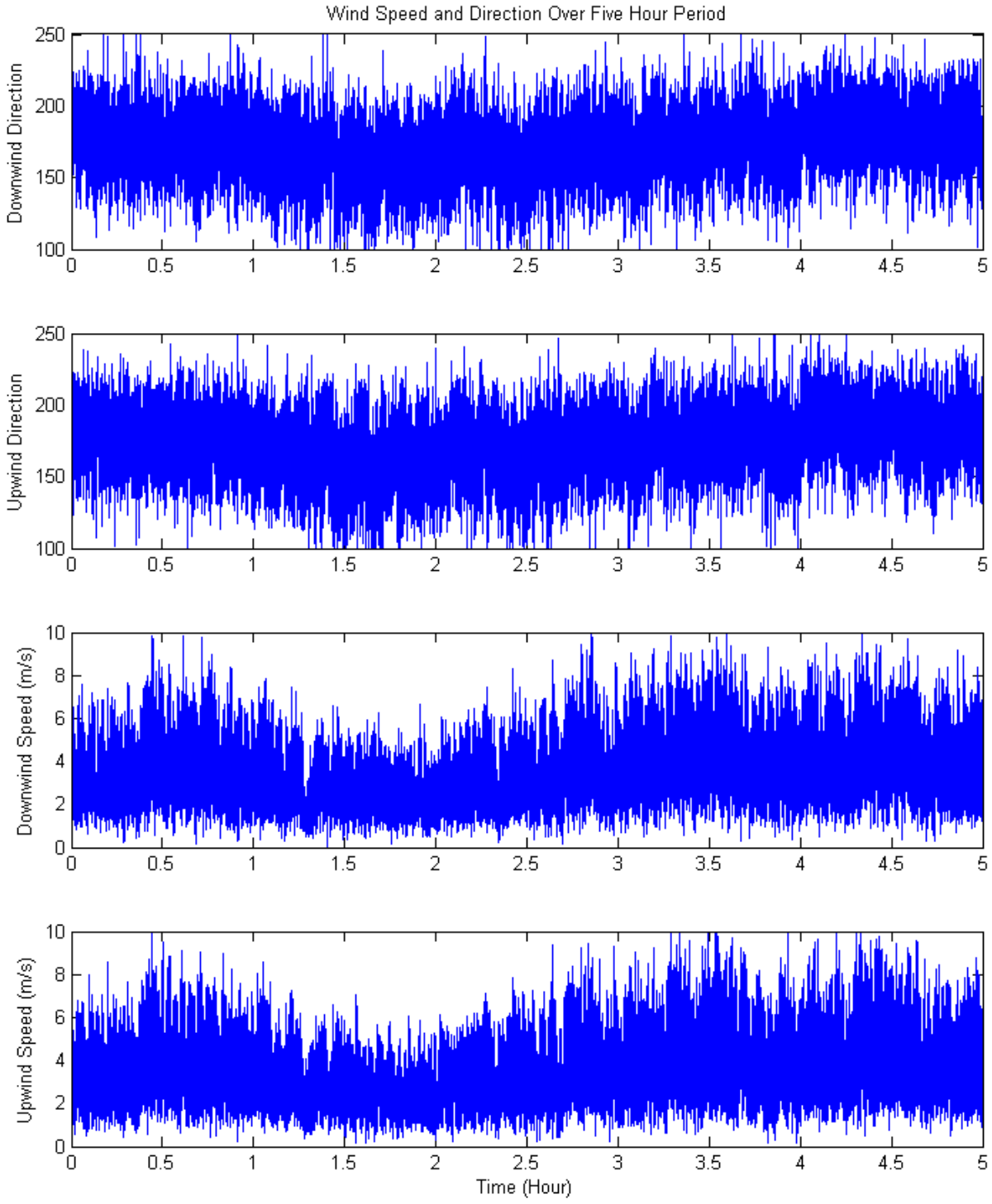

Figure 6: Five hour period of upwind/downwind sensors for both wind speed and wind direction

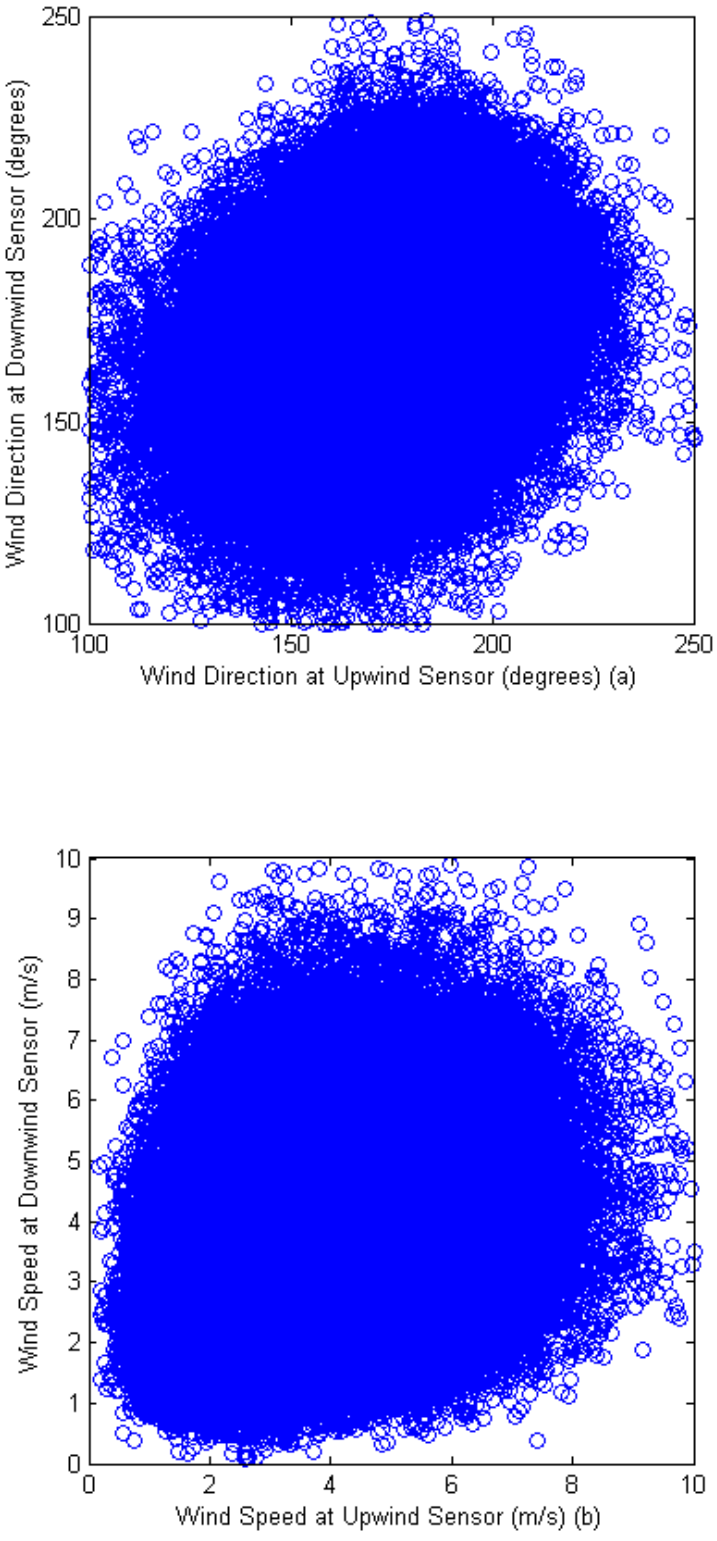

Figure 7: Scatter plot showing downwind direction (a) and downwind velocity (b) as a function of upwind conditions

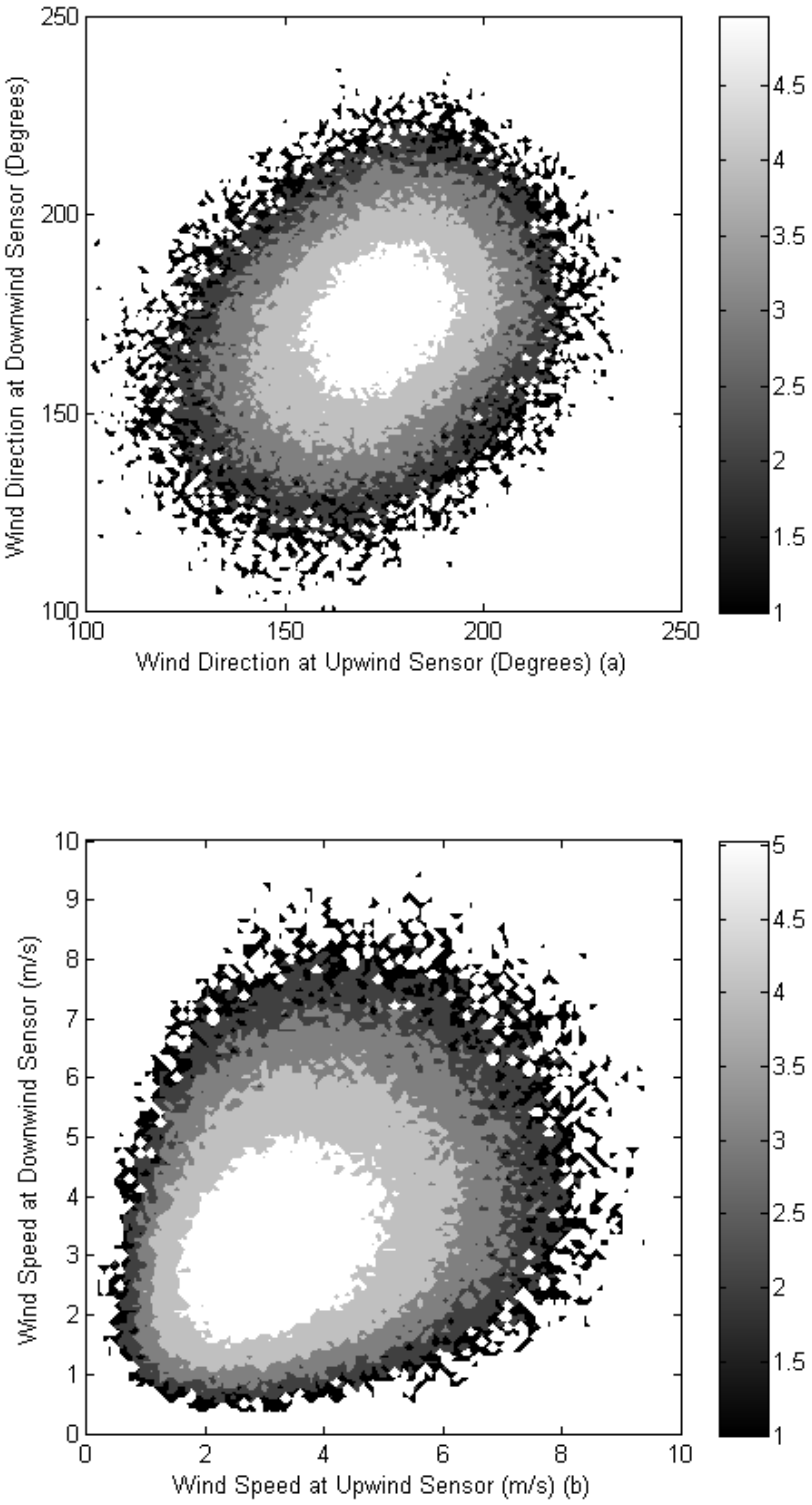

Figure 8: Density plots of point grouping for downwind direction (a) and downwind speed (b) as a function of upwind conditions. The range was divided into sectors (e.g., for wind direction, grid spacing of 1x1 degree were used) and the number of points were then counted for each sector. The natural log was then applied to the counts to show underlying structure. Since ln(1) = 0, a value of 1 was added to all sectors that originally had a value of at least one.

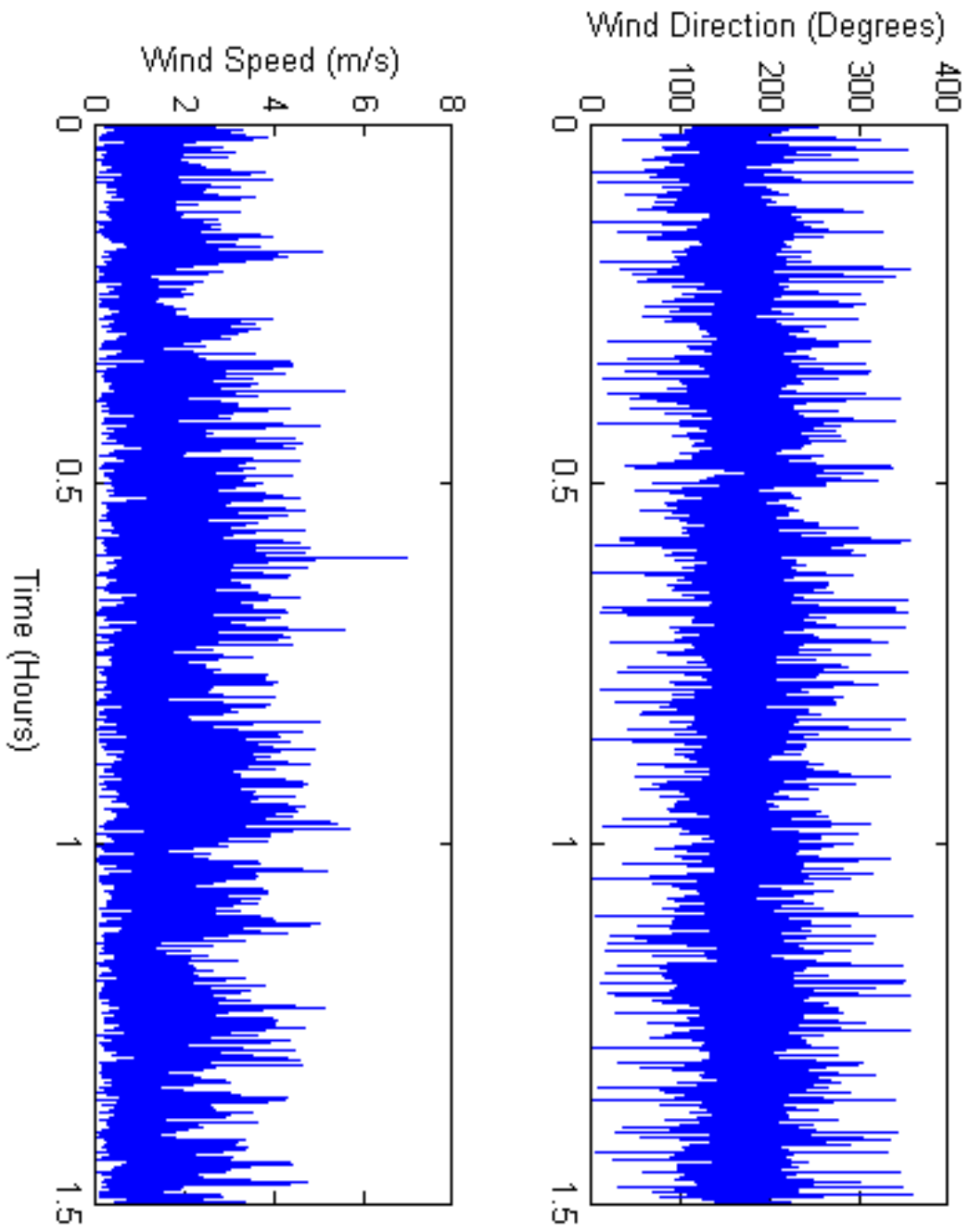

Figure 9: One and a half hour period of wind speed and wind direction at the center sensor

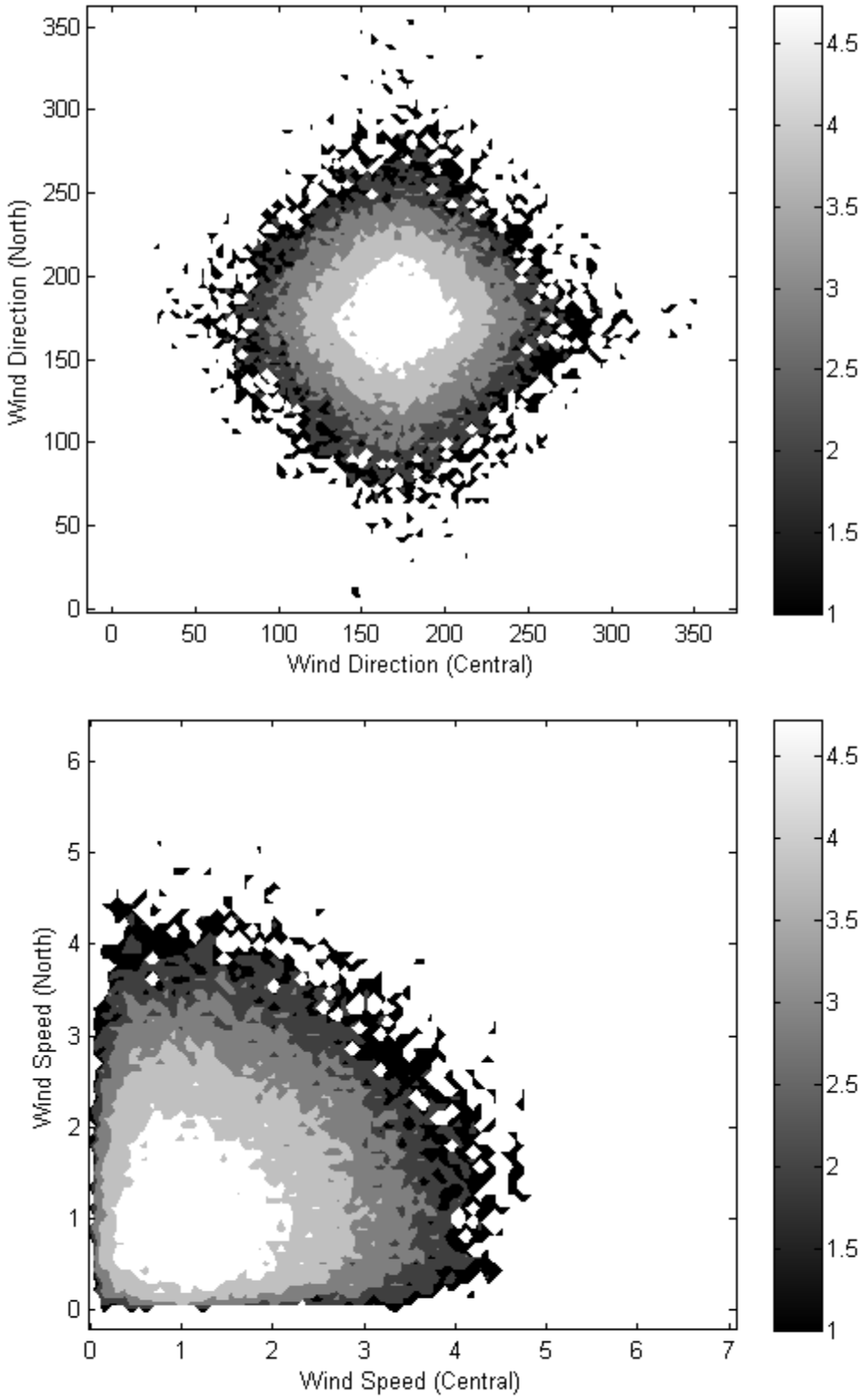

Figure 10: Histograms of wind direction (degrees) and wind speed (m/s) at the central and northern sensors

Short Term Relationships during one minute intervals (2014)

Significant flight path changes of small, drift-prone spray droplets take place over much shorter time periods, on the order of a minute rather than five hours. The five hour period was broken into shorter, one minute long data sets. Linear correlation coefficients were then calculated for each one minute period in the five hour (300 minute) long data set using the 10 Hz data. Averaging these 300 linear correlation coefficients yielded adjusted correlation coefficient values to compare velocity relationships between upwind and downwind sensors during one-minute time periods. For wind direction and wind speed, the adjusted average correlation coefficients were 0.01 and 0.02 respectively. This result gives insight of the highly variable nature of wind during short term measurements. When the correlation was calculated for the entire five hour period, there was time for the wind speed and direction to shift sufficiently to show long term correlation. This is not the case during short time periods when the droplet is in the air.

Short Term Relationships during one minute intervals (2015)

The one and a half hour period was split into multiple one minute long data sets. The linear correlation coefficients were then calculated as before. To reduce confusion, when discussing interactions between two sensors, the first letter of the sensor location is used with the second letter denoting the upwind sensor. For example, if discussing a process at the central sensor that came from the south sensor, this process would be given the name CS. The linear correlation coefficient for wind speed for NC was 0.02, and the coefficient for CS was 0.01. The wind direction linear correlation coefficients for NC and CS were 0.003 and 0.005 respectively.

Max Correlation using lag adjustment (2014)

Using smaller one-minute segments of data, the time it takes for the wind change to travel to the next sensor can be included. During the five hour period, there was a 3.6 m/s (8 mi/h) average wind speed and the sensors were spaced 30.5 m (100 ft) apart. Average wind speed implies a lag of about 8.5 seconds, (taking the distance separated by the sensors and dividing by the average wind speed), before similar conditions would be observed downwind if the wind moved smoothly over the field. However, maximum correlation occurs at 12 seconds (Figure 11). This could be caused by turbulent interactions across the top of the oat crop canopy.

From investigating the five hour segment of data, wind direction and wind speed cannot be considered to remain constant at a downwind location after leaving the upwind sensor. Taylors Hypothesis of Frozen Turbulence states that turbulence can be considered “frozen” as eddies (turbulent motion of wind) advects pass a sensor(s) (Stull, 2009). The lack of correlation between upwind and downwind sensors, even after adjusting for lag in wind conditions shown in Figure 11 implies that turbulence is not frozen on these small length and time scales.

Lag was also checked for the 2015 data. There was no set lag time that gave the highest correlation value unlike the 2014 data. Mean wind speed of 2015 data was slower than 2014 data, and the inability to identify a lag period may have been related to the light and variable nature of slower wind speeds.

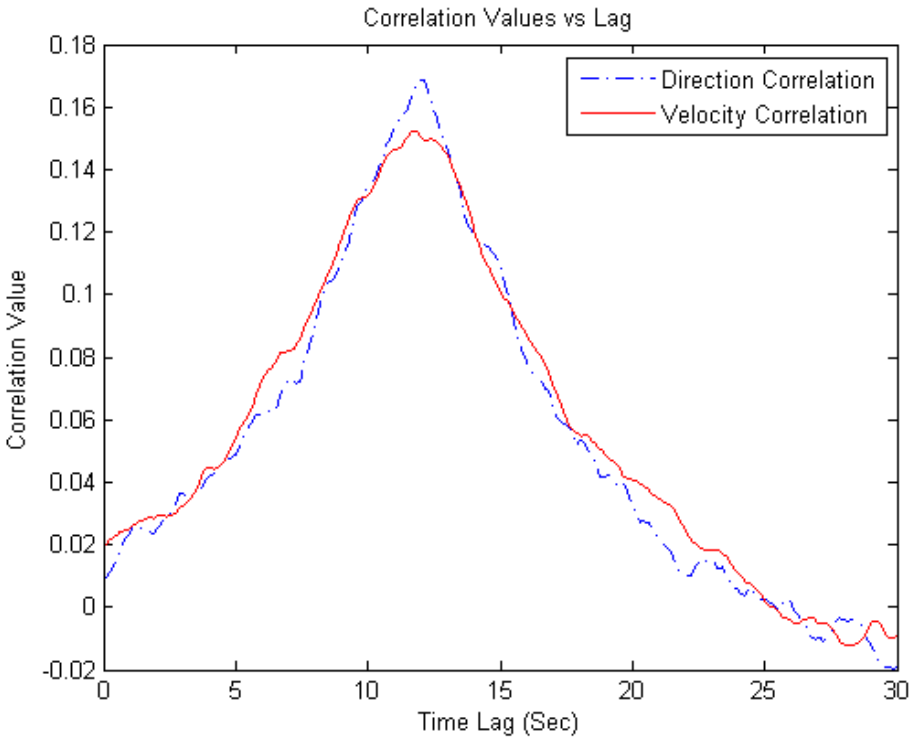

Figure 11: Correlation plot as a function of lag between the sensors for both wind direction and speed with 2014 data

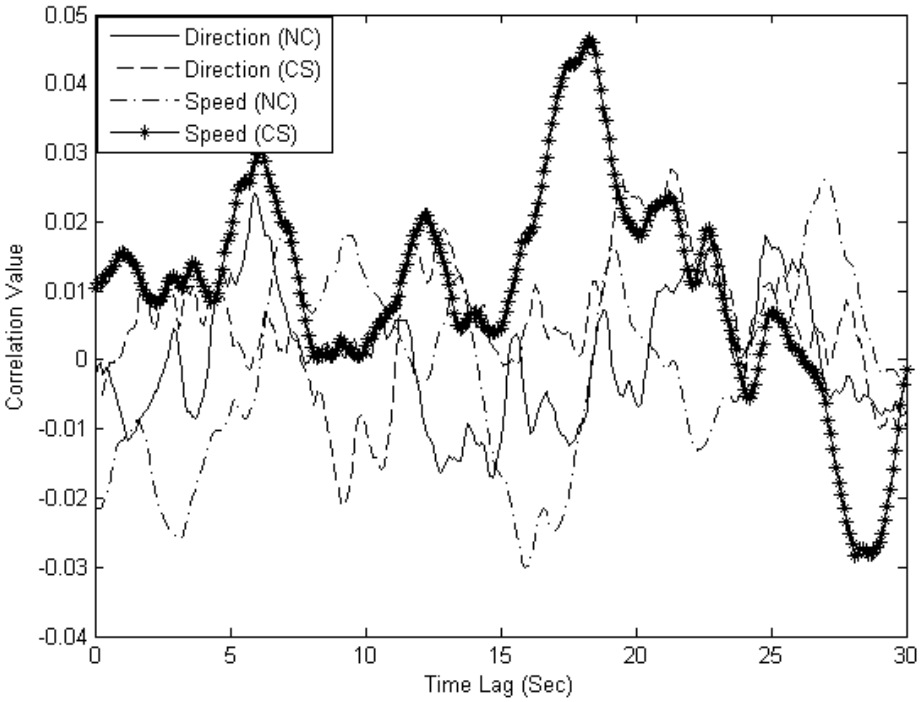

Figure 12: Correlation plot as a function of lag between the different sensors for both wind direction and speed with 2015 data

Probability that downwind velocity is within range of upwind velocity

To anticipate an unexpected change in wind direction or speed after spray is released from the nozzle, it would be desirable to know what the probability is of the downwind sensor being within a given tolerance range of an upwind sensor. Multiple sensors were used again utilizing the higher wind speed dataset. From Figure 11, an adjusted lag value of 12 seconds was used in this analysis of the 10 Hz data. The percentage of time that the downwind sensor is outside a given tolerance of the upwind sensor is shown in Figure 13. This probability analysis was done by comparing the difference between the upwind and downwind sensors and counting how many data points fell outside of a tolerance. This count was then divided by the total number of points. For example, if the speed tolerance was set to $\pm$ 1 m/s, then about 50% of the time the absolute value of the difference between the upwind and downwind sensor would be greater than 1 m/s. If the directional tolerance was set at $\pm$ 20 degrees, then about 30% of the time the absolute value of the difference between the sensors would be greater than the set tolerance. Figure 13 gives insight on the probability of random finite fluctuations in both wind velocity and direction.

Figurer 14 shows this analysis for the low wind speed dataset for wind direction and wind speed. The wind speed was scaled to 1 m/s to better match the scaling of the 2015 data. Due to the lack of a peak lag for all sensors, no lag time was used in figure 14. Figure 14 shows that the sensors followed similar curves with one another. This result shows that the probability that a downwind measurement will be within an upwind measurement is similar within at least 30 meters of the upwind sensor. Due to the low wind speed and the lack of a group lag time, the variability at downwind sensors is higher

leading to higher degrees of uncertainty. As an example, if a tolerance of 20 degrees was set, the downwind sensor would be outside of the tolerance about 65% of the time.

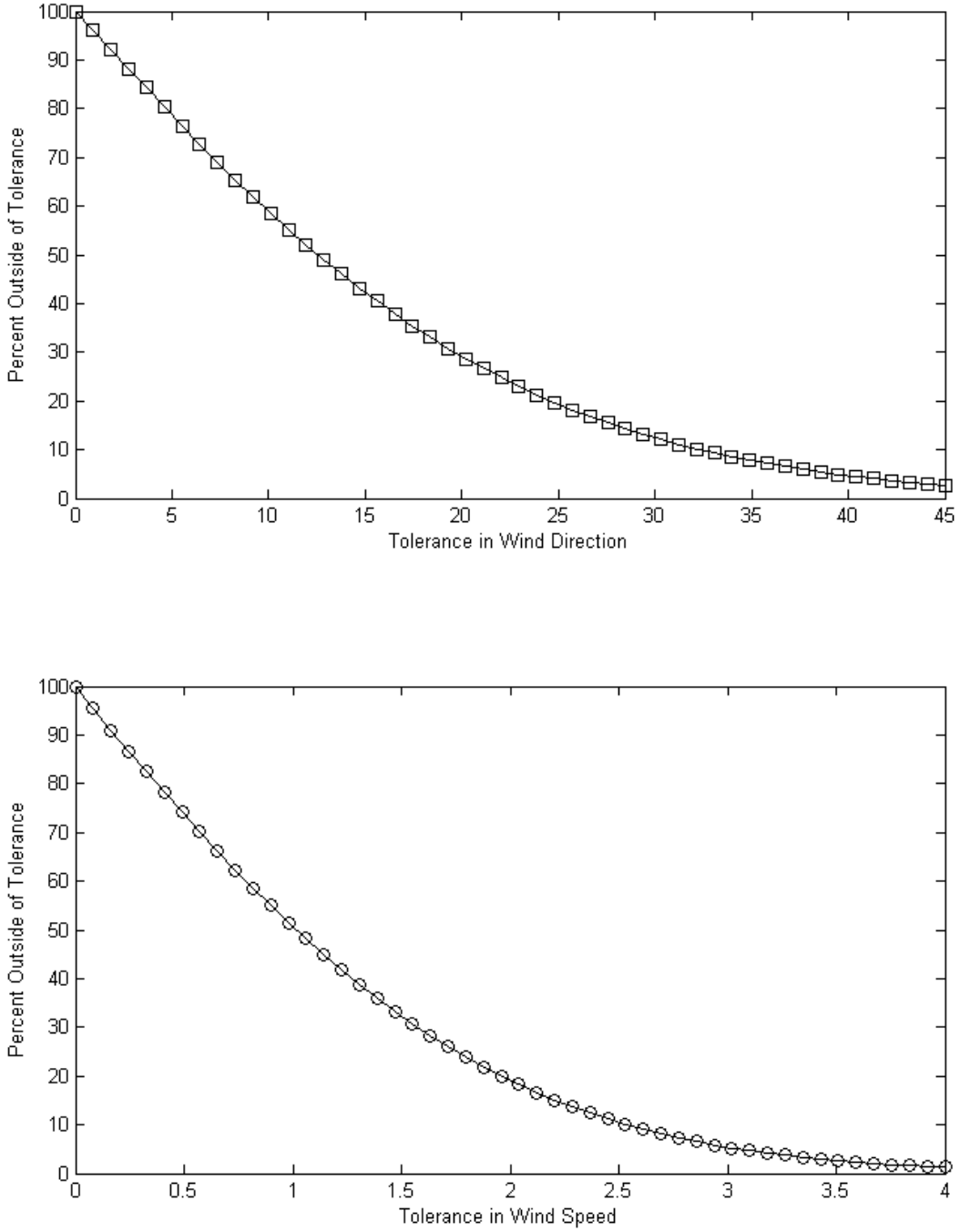

Figure 13: Probability that the downwind sensor will be outside a given tolerance to the upwind sensor for wind direction and wind speed

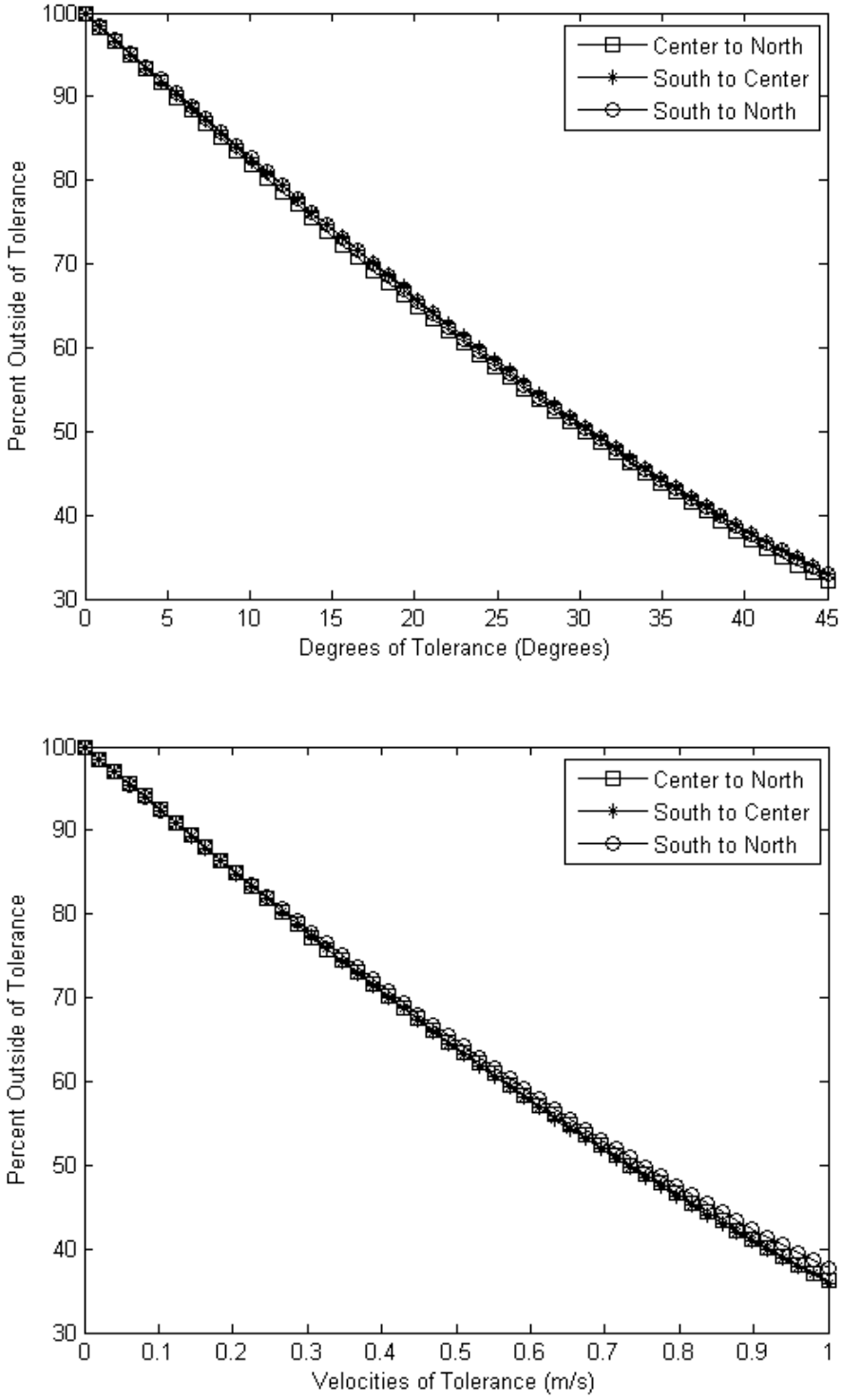

Figure 14: Probability that the downwind sensor is within a tolerance of an upwind sensor

Parameters related to wind changes

To gain greater insight as to what meteorological weather conditions are present when transient wind velocity changes occur that may affect trajectory after the droplet has left the spray boom, data collected over a range of 36 days (greater than 100 million data points) in the late spring/summer spraying seasons of 2014 and 2015 were analyzed to see when it was more likely to see a change in the wind direction 30 seconds into the future. As data were processed, if the value for wind direction at a sensor, thirty seconds into the future, was greater than a given tolerance (45, 25, or 5 degrees), the current wind direction, speed, solar radiation, and time of day were recorded. Figures 15, 16, and 17 show the percentage of occurrences when wind direction changed by more than 45, 25, or 5 degrees for each individual finite segment of solar radiation, time of day, and wind speed, respectively, when comparing wind direction at a specific time and 30 seconds later at individual wind sensors. For example, at a tolerance of 45 degrees, out of the total time in which the day was between the hours of 1:30 am and 4:30 am, about 5% of that time had a large wind change that was greater than the allowed tolerance.

The figures 15, 16, and 17 show how wind change events are distributed. As the sun heats the earth, increasing the surface energy, the weather becomes turbulent (Stull, 2009) as is seen in Figure 15. As the tolerance is tightened, to 25 degrees and 5 degrees in Figures 16 and 17 the distributions for Time of Day and Solar Radiation become more uniform. However, the tendency for wind directional shifts to occur more frequently at lower wind speeds remains relatively unchanged. This shows that most unstable events occur below 1 m/s (2.2 mi/h) winds. Application at these wind speeds is often recommended and may limit the distance of off-target drift even if the likelihood of

unforeseen wind directional shift is greater.

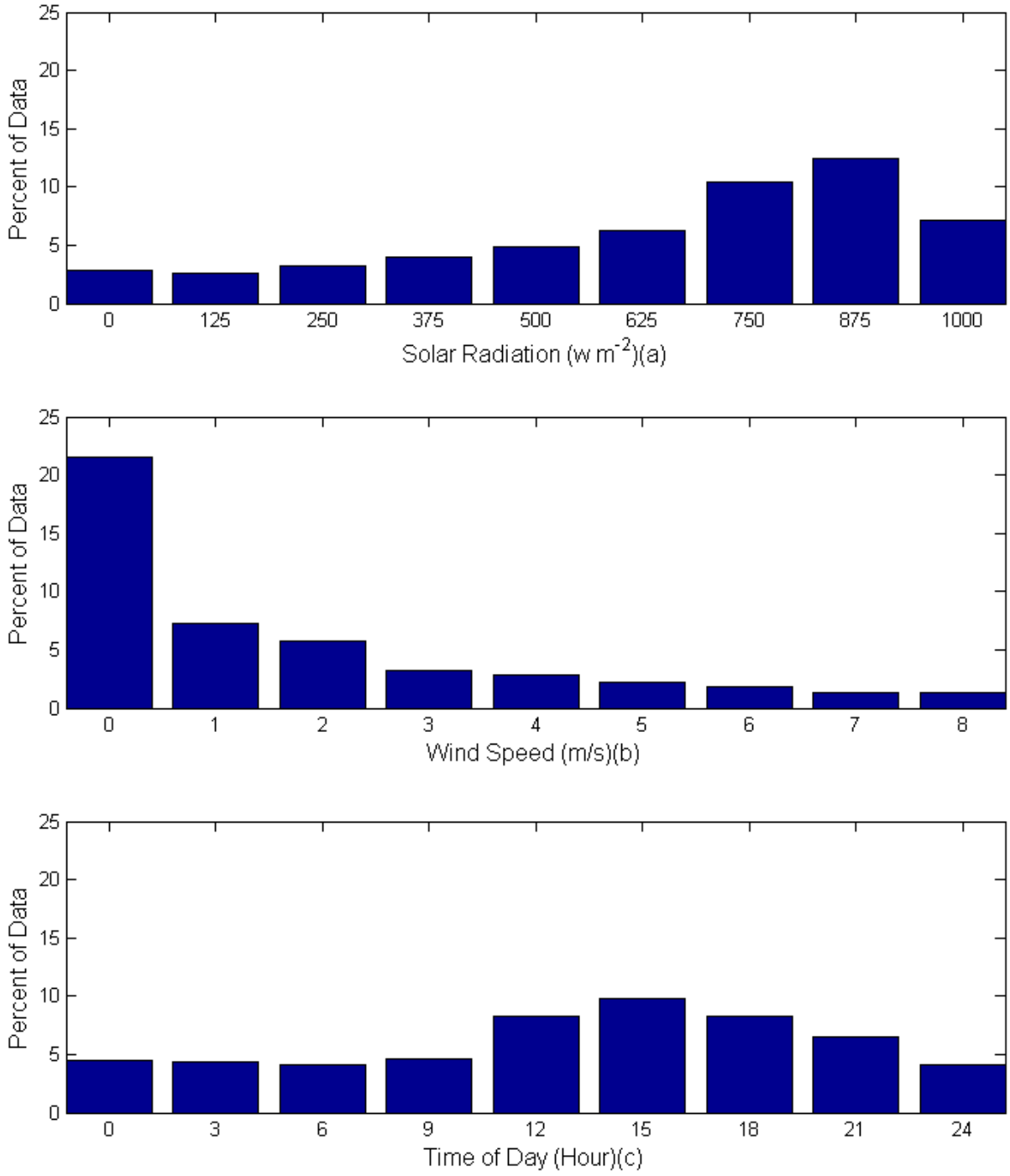

Figure 15: Percentage of data within a certain range in which a wind change 30 seconds in the future was greater than 45 degrees

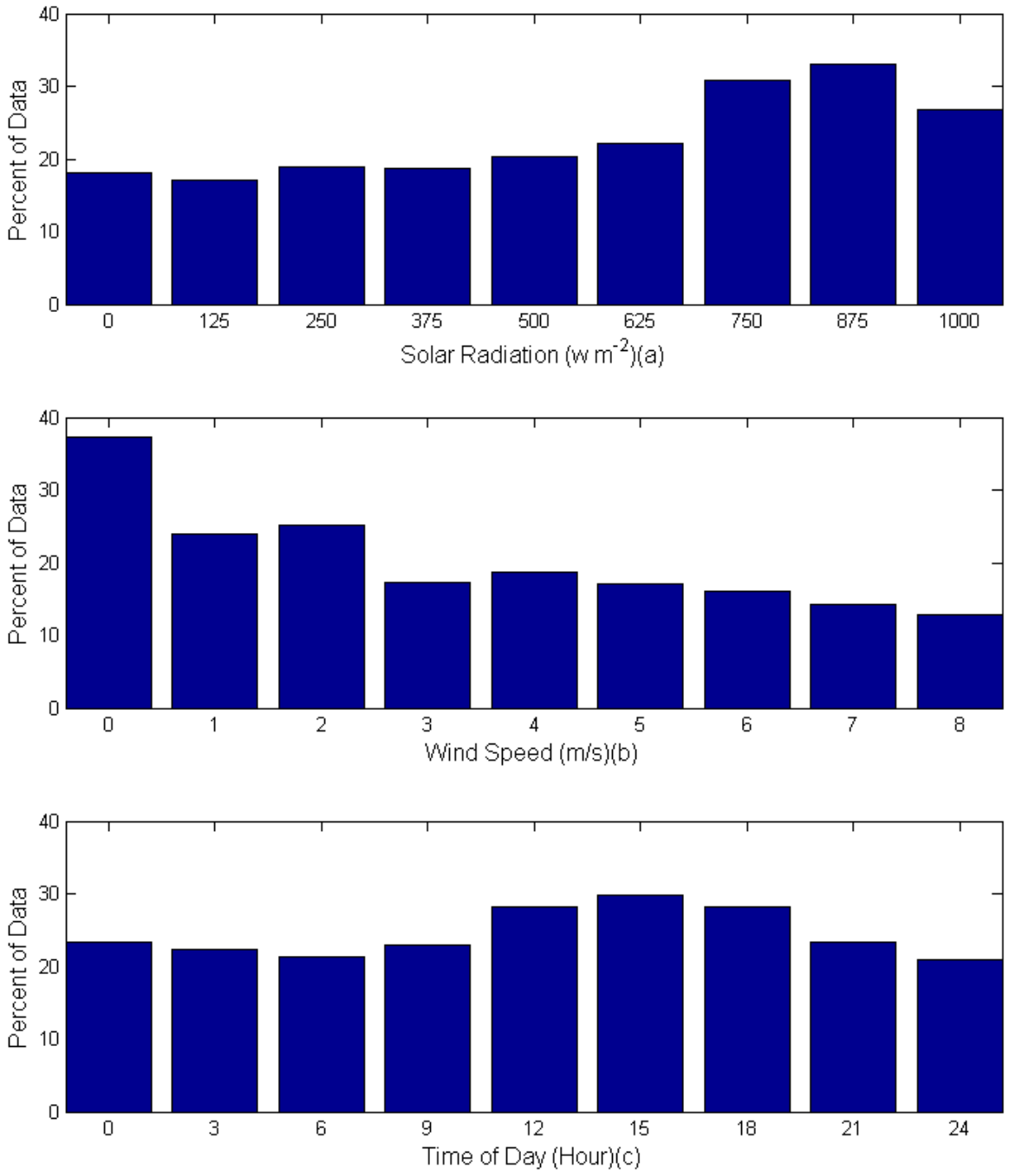

Figure 16: Percentage of data within a certain range in which a wind change 30 seconds in the future was greater than 25 degrees

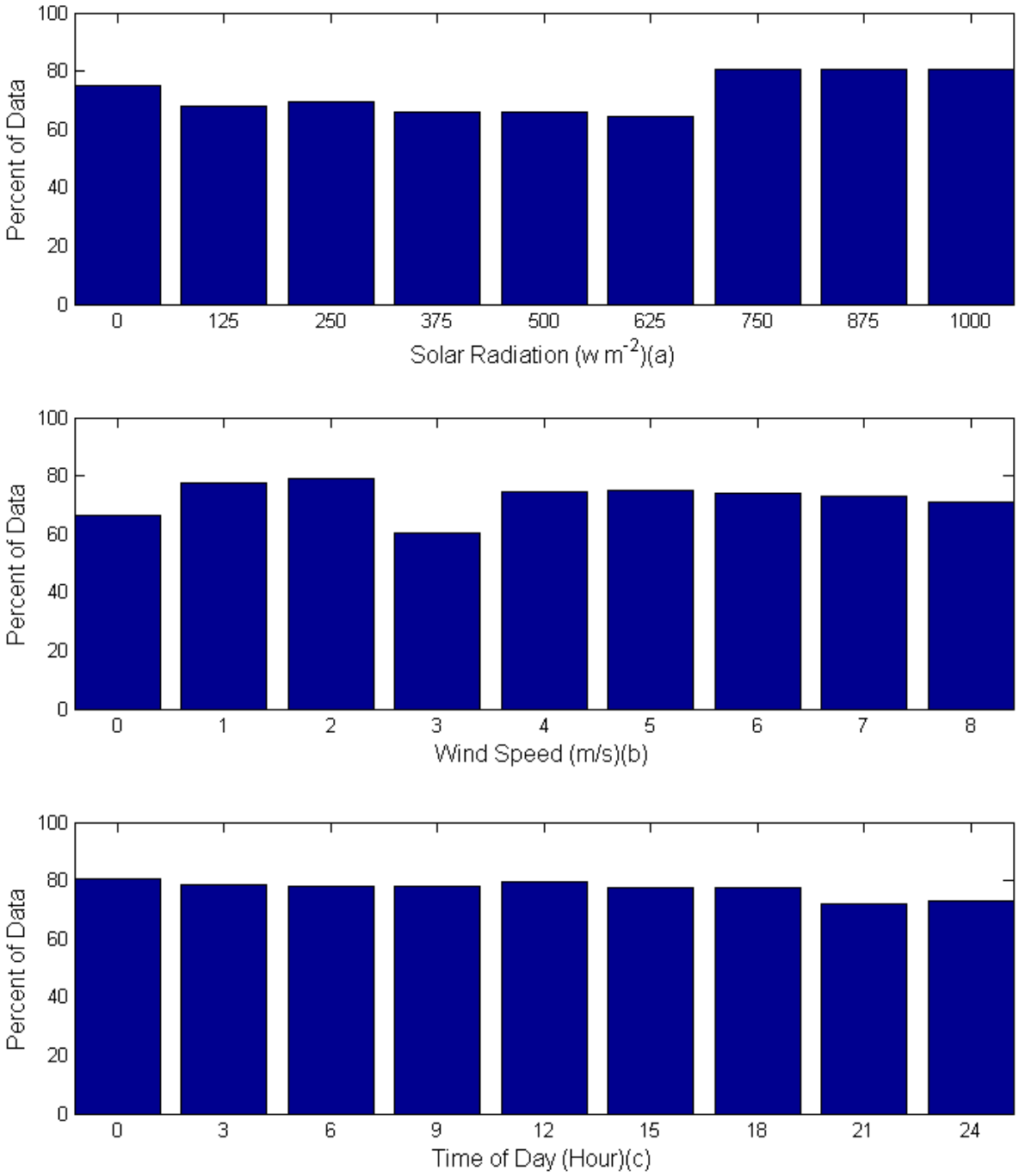

Figure 17: Percentage of data within a certain range in which a wind change 30 seconds in the future was greater than 5 degrees

Conclusion

Wind velocity data taken during late spring/summer conditions that are suitable for ground based spraying support the following conclusions:

- During a five hour period (7:30 am to 12:30 pm) with an average wind speed of 3.6 m/s (8 mi/h) and general wind direction of due north
    - Random fluctuation of wind direction and speed at upwind and downwind sensors, 30 m (100 ft) apart, had correlation coefficients of 0.29 and 0.27
    - Correlation was less during shorter one-minute periods in which a spray droplet may travel, but improved to a coefficient of 0.15 if a lag time was used between the two sensors
    - Using a lag time, downwind direction was greater than 20 degrees different than the upwind sensor 30% of the time while wind speed was greater than 1 m/s (about a quarter of the mean wind speed) different than the upwind speed about 50% of the time
- During a 1.5 hour period (12:50 pm to 2:20 pm) with an average wind speed of 1.5 m/s (3.3 mi/h) and general wind direction of due north
    - Correlation coefficients between the downwind and upwind sensors either 15 or 30 m (50 or 100 ft) apart had lower values < 0.03
    - Downwind direction was greater than 20 degrees different than the upwind direction 65% of the time, while the downwind speed was greater than 0.25 m/s (about a quarter of the mean wind speed) different than the upwind speed 80% of the time
- Across a range of late spring/summer days in which suitable conditions for ground

spraying were present, significant change in wind direction 30 seconds later was more likely to occur during wind speeds in the range of 0-3 m/s (0-6.7 mi/h) with all tolerances

- Solar radiation and time of day seemed to have a greater effect on wind changes when investigating tolerances of 45 degrees, as the tolerance tightened to 5 degrees, solar radiation and time of day became more uniformly distributed

References

Baetens, K., D. Nuyttens, P. Verboven, M. De Schampheleire, B. Nicolai, and H. Ramon. 2007. "Predicting drift from field spraying by means of a 3D computational fluid dynamics model." *Computers and Electronics in Agriculture* 56: 161-173.

Brown, Ralph B, Margaret H Carter, and Gerald R Stephenson. 2004. "Buffer Zone and Windbreak Effects on Spray Drift Deposition in a Simulated Wetland." *Best Management Science* 60 (11): 1085-1090.

Craig, I P. 2004. "The GDS Model - A Rapid Computational Technique for the Calculation of Aircraft Spray Drift Buffer Distances." *Computers and Electronics in Agriculture* 43 (3): 235-250.

EPA. 2014. "Glossary." *U.S. Enbironmental Protection Agency.* 02 04. http://www.epa.gov/pesticides/regulating/labels/pest-label-training/glossary/.

Faires, Richard L. Burden and J.Douglas. 2011. *Numerical Analysis.* Cengage Learning.

Frederic, L., A. Verstraete, B. Schiffers, and M.F. Destain. 2009. "Evaluation of Reatime Spray Drift Using RTDrift Gaussian Advection-Diffusion Model." *Commun. Agric. Biol. Sci.* 74 (1): 11-24.

Google. 2015. *Google Maps.* https://www.google.com/maps/@42.0147298,-93.731401,396m/data=!3m1!1e3.

Holterman, H.J, J.C Van de Zande, H.A.J Porskamp, and J.F.M Huijsmans. 1997. "Modeling Spray Drift from Boom Sprayers." *Computers and Electronics in Agriculture* 19 (1): 1-22.

Nordby, A, and R Skuterud. 1974. "The effects of boom height, working pressure and wind speed on spray drift." *Weed Research* (Weed Research) 14 (6): 385-395. doi:10.1111/j.1365-3180.1974.tb01080.x.

Smith, D B, F D Harris, and C E Goering. 1982. "Variables affecting drift from ground boom sprayers." *Transactions of the ASAE* 25 (6): 1499-1523.

Spray Drift Task Force. 1997. *A Summary of Ground Application Studies.* Macon, MO: Agricultural Research Services Inc.

Stull, R.B. 2009. *An Introduction to Boundary layer Meteorology.* 13. Springer.

Teske, M.E., S.L. Bird, D.M. Esterly, T.B. Curbishley, S.L. Ray, and S.G. Perry. 2002. "AgDrift: A Model for estimating near-field spray drift from aerial applications." *Environmental Toxicology and Chemistry* 21 (3): 659-671.

Tsai, M., K. Elgethun, J. Ramaprasad, M.G. Yost, A.S. Felsot, V.R. Hebert, and R.A. Fenske. 2005. "The Washington aerial spray drift study: Modeling pesticide spray drift deposition from an aerrial application." *Atmospheric Environment* 39: 6194-6203.

CHAPTER 4. A NOT-SO-RANDOM WALK WITH WIND: A LOOK AT RANDOM FLUCTUATIONS IN WIND VELOCITIES FOR USE IN MODELS OF AGRICULTURAL SPRAY DRIFT

A Paper to be submitted to Transactions of the ASABE

Matt Schramm[1], Mark Hanna[1], Matt Darr[1], Steven Hoff[1], Brian Steward[1]

Agricultural and Biosystems Engineering, Iowa State University[1]

Abstract

The notion that wind speed and direction can be approximated by adding a random fluctuation to the previous value was investigated. The data were recorded at one meter above a field to simulate conditions that are present at a ground sprayer's boom. Variance ratio tests were carried out to test the null hypothesis that wind process similar properties to a random walk versus the alternative that wind does not. More specifically, that the random fluctuations are auto correlated with one another in time. This process was done to a 10 Hz sample and averages of the measured wind data at 1/2 s, 1 s, 5 s, 10 s, 30 s, 1 min, 5 min, and 10 min. It was found that for all tests, except for the 5 and 10 min data samples, the null hypothesis was rejected at greater than 99.9% certainty. Indicating that there is evidence of autocorrelation in the measurements of wind speed and direction, associated with each other in time.

Keywords: *Random Walk, Spray Drift Model Validation, Spray Drift*

Introduction

Agricultural sprayers provide agricultural chemicals that protect and improve crop plant health, but the chemicals can be detrimental to adjacent crops or other forms of life. Understanding how spray propagates (through the use of models) is essential to provide the industry with the necessary information to keep the chemicals in the field. The EPA defines spray drift as follows, “Pesticide spray drift is the physical movement of a pesticide through air at the time of application or soon thereafter, to any site other than that intended for application” (EPA, 2014). Multiple models are available to calculate the physics of droplet movement including Lagrangian, Gaussian, Random Walk, Regression, and Computational Fluid Dynamic (CFD) models (Frederic, et al., 2009; Holterman, et al., 1997; Baetens, et al., 2007; Milton E. Teske, 2002; Ming-Yi Tsai, 2005; Frederic, et al., 2009; Smith, et al., 1982). Models apply random fluctuations to the mean wind to simulate the wind’s turbulent nature. Essentially, the wind takes a random walk. A popular application of random walks, is to use averaged wind velocities acquired from measurements and introduce random fluctuations to simulate turbulence (Holterman, et al., 1997; Thompson & Ley, 1983). Wind fluctuations may appear random but little information exists of short term, transient wind velocity changes that can confirm this property accurately near the surface in the vicinity of a ground sprayer boom.

A statistical method to test the random walk hypothesis that is used primarily for financial predictions is the variance ratio test (MathWorks, 2015; Ostasiewicz, 2000). The variance ratio test investigates the “random fluctuations” of a time series dataset and tests if changes in the time series are statistically independent or if these changes are correlated with one another. The variance ratio test has been primarily used to test the

random walk hypothesis for market efficiencies in finance (Charles & Darné, 2003). The test is particularly useful for testing if the process eventually returns to the average (mean reversion) (Charles & Darné, 2003). The random walk model was first introduced in 1828 when the botanist Brown described his Brownian motion. Since then the model has gained ground in multiple fields from biology, physics, and finance (Codling, et al., 2008). The test devised by Lo and MacKinlay (1988) was directed at financial markets, though no assumptions were ever made that limit the test to only financial random walks (Lo & MacKinlay, 1988).

Objectives

The objectives of this research were to:

- Investigate changes in transient wind velocity near the ground surface at a typical ground sprayer boom height, and
- Evaluate the randomness of measured transient wind velocity changes for use in random walk models as applied to ground sprayers

Methods and Materials

Experimental Design and Apparatus

Field measurements of wind speed and direction were collected during the late spring/summer spraying season of 2014 using instrumentation set into a field of growing oats. The field was located at the Iowa State University Research Farm's Bruner Farm field F1 near Ames, Iowa (Figure 18). The field dimensions were 268 m long by 105 m wide (880 by 348 ft).

Figure 18: Bruner Farm field F1 (42.014911 N 93.731241 W) (Google, 2015)

Wind speed and solar radiation measurements were acquired at 10 samples per second using ultrasonic anemometers (model: WindMaster 3d, Gill Instruments, Lymington, Hampshire, UK) and a pyranometer (model: SP-212, Apogee Instruments, Logan, UT). The anemometers measured the wind speed in the north-south, east-west, and vertical directions (Figure 19). Open source microcontrollers equipped with a GPS module (model: Arduino Uno, Arduino Inc; Ultimate GPS Shield, Adafruit, New York, USA) were used to log data to a micro SD data card. Using the GPS's PPS (Pulse per Second) output, time correction was done to ensure accurate time recording of wind velocity measurements on the microcontroller. To reduce influences to wind velocity, the

microcontrollers and their power supplies were located separate from the anemometers at a distance of approximately 2-3 meters. Anemometers were placed one meter above the ground's surface to collect wind measurements. The pyranometer was placed on the charging station near the anemometer.

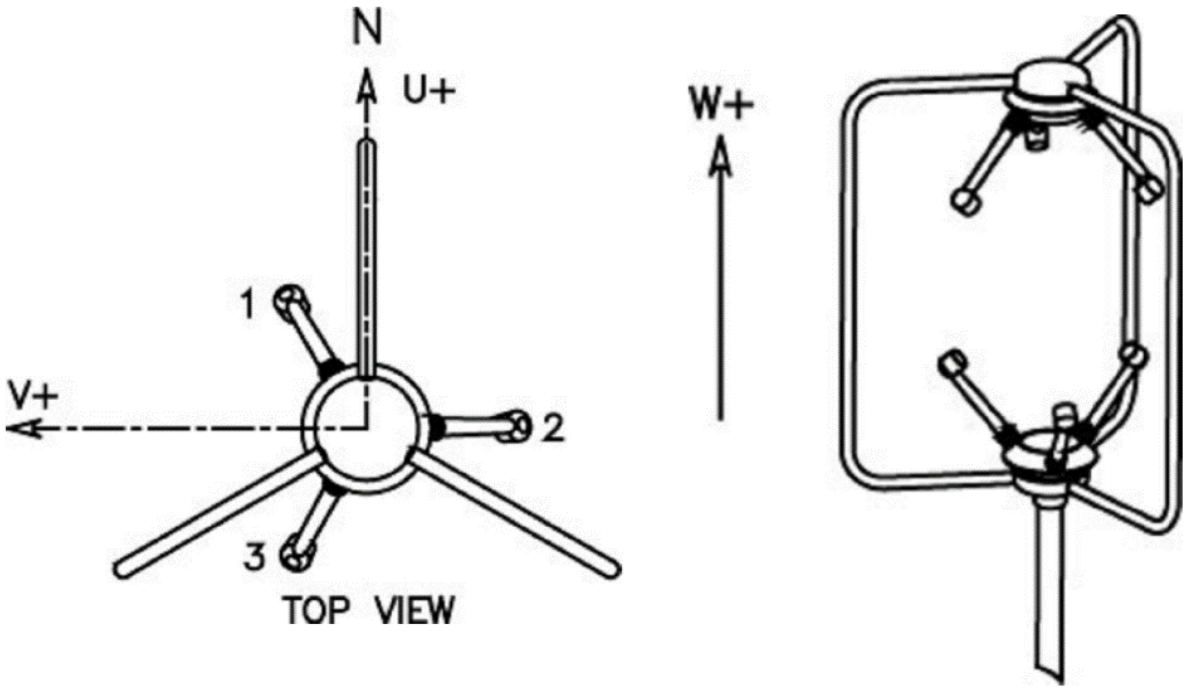

Figure 19: Ultrasonic Anemometer measuring velocity in U, V, and W component directions

Data Analysis

The data from a single sensor shown in figure 20 is from a five hour period, between 7:30 am to 12:30 pm during typical daytime spraying conditions, and will be used for the study. Wind direction was calculated from the wind speeds using trigonometric relationships. A wraparound method was used to produce a semi-continuous wind direction dataset. The wraparound method allowed wind direction to go above 359 degrees and below 0 degrees, e.g. 12 degrees = 372 degrees.

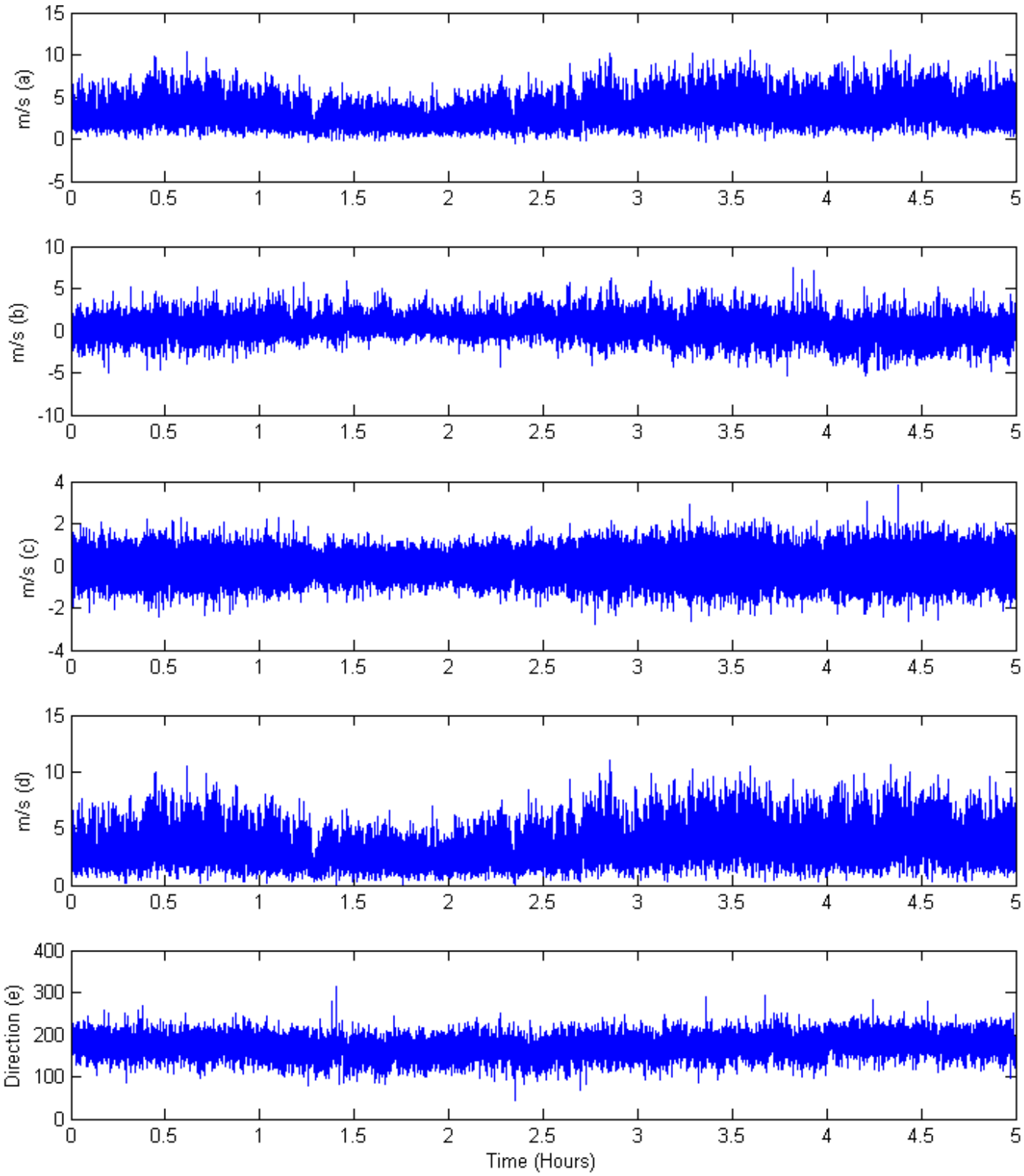

Figure 20: Data collected over a five hour period. (a) North-South, (b) East-West, (c) Vertical Wind, (d) Wind Magnitude, (e) Wind Direction

The variance ratio test was used to test the null hypothesis (h_0) that a random walk (Equation 17) faithfully represents the statistical properties of the measured wind fluctuations. Specifically, the test is used to measure if the values of ϵ_t are correlated.

$$y_t = c + y_{t-1} + \epsilon_t \tag{17}$$

where

- y_t are the time series process
- c is a drift constant for the random walk model
- ϵ_t are random independent processes and are distributed with mean zero and variance σ^2 (Box, et al., 1994; Enders, 1995; MathWorks, 2015; Ostasiewicz, 2000)

This test was done for the wind speeds traveling in the north-south direction (Uwind), the east-west direction (Vwind), the vertical direction (Wwind), the magnitude of the north-south and east-west winds (MWind), and wind direction (DWind). In addition to testing individual readings each 0.1 s, multiple averages (1/2 s, 1 s, 5 s, 10 s, 30 s, 1 min, 5 min, and 10 min) were tested to determine if longer-term averaged readings were eventually possessed properties of the random walk.

Variance ratio test statistics are calculated based upon the ratio of variance estimates from the time series' m-period returns (MathWorks, 2015; Box, et al., 1994; Enders, 1995; Ostasiewicz, 2000). The returns are calculated from the difference of terms in the series (Equation 18)

$$r_m = y_t - y_{t-m-1} \tag{18}$$

The variance ratio for period m is:

$$VarRatio_m = \frac{1}{m+1}\frac{\mathrm{var}(r_m)}{\mathrm{var}(r_0)} \tag{19}$$

The m parameter determines the period of the time difference. For example if the model that was being tested used the third previous value to predict the next value, m would equal 3. For the purposes of this paper, m will equal one to signify that the use of the previous step is used to predict the next step. This matches the form of the wind model shown in equation 17. This ratio indicates how the individual measurements behave, either by propagating to the mean (ratio less than one), away from the mean (ratio greater than one).

With the variance ratio calculated, the test statistic is then calculated using Lo and MacKinlay's heteroscedasticity method (Charles & Darné, 2003) (Equation 20). Heteroscedasticity is needed because the assumption that ϵ_t are independently and identically distributed is not required for this test.

$$Tvr = \frac{VarRatio_m - 1}{\sqrt{\phi_m^*}}$$

$$\phi_m^* = \sum_{j=1}^{m}\left[\frac{2(m+1-j)}{m+1}\right]^2 \delta(j), \qquad \phi_1^* = \delta(1) \tag{20}$$

$$\delta(j) = \frac{\sum_{i=j+1}^{T}(y_i - \hat{y})^2\left(y_{i-j} - \hat{y}\right)^2}{(\sum_{i=1}^{T}(y_i - \hat{y})^2)^2}$$

Where y_i is the ith observed value in the time series, $\hat{y}$ is the average value in the time series, and T is the length of the time series, and as T approaches infinity, this test statistic approaches a normal distribution (Charles & Darné, 2003). Test statistics are used to calculate the p-values.

Results and Discussion

Table 1 summarizes the test results of all tests done. The statistical method was done at a 95% confidence level to determine if the null hypothesis was accepted or rejected.

Table 1: Variance Ratio Test results at multiple averaging values in which the null hypothesis is accepted or rejected. (Uwind is the north-south wind speed, Vwind is the east-west wind speed, Wwind is the vertical wind speed, Mwind is the magnitude of the wind speed, and Dwind is the direction of wind)

		Sample Time								
		1/10 sec	1/2 Sec	1 Sec	5 Sec	10 Sec	30 Sec	1 Min	5 Min	10 Min
Test Results	Uwind	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Plausable	Reject
	Vwind	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Plausable
	Wwind	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject
	Mwind	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Plausable	Reject
	Dwind	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Plausable
P-Value	Uwind	0	0	0	0	0	1.20E-14	6.50E-09	0.132	0.006
	Vwind	0	0	0	0	0	6.00E-07	2.70E-08	0.015	0.205
	Wwind	0	0	0	0	0	5.90E-20	1.80E-10	0.003	0.027
	Mwind	0	0	0	0	0	8.50E-15	4.80E-09	0.069	0.007
	Dwind	0	0	0	0	0	3.60E-07	4.60E-08	0.028	0.429
Variance Ratio	Uwind	0.94	0.86	0.75	0.71	0.67	0.61	0.59	0.82	1.40
	Vwind	0.84	0.75	0.69	0.68	0.72	0.74	0.58	0.60	0.74
	Wwind	0.75	0.60	0.54	0.52	0.54	0.45	0.57	0.48	0.62
	Mwind	0.94	0.86	0.75	0.71	0.68	0.61	0.59	0.76	1.35
	Dwind	0.86	0.73	0.67	0.68	0.72	0.75	0.58	0.61	0.85

From table 1 the null hypothesis of random data is rejected with greater than 99.9% certainty, except for longer periods when data is averaged over 5 or 10 minutes. This shows that the wind fluctuations are correlated with one another at short time steps, at one meter height near the ground surface, but may be considered to be randomly independent as the time between each step increases. Droplet sizes greater than 50 µm contain a larger total fraction of driftable spray volume than tinier droplets and unless displaced vertically upward a significant distance would be expected to deposit in these shorter time periods with non-random wind velocity. For models that incorporate random fluctuations, greater care is needed and simply adding a random fluctuation to the

previous wind conditions is not completely correct. The exact relationship that should be used for modeling turbulent fluctuations is currently unknown.

Conclusions

It was found that wind velocity changes at one meter (near the height of a ground sprayer's boom) are not purely random when using sampling periods of less than five minutes. Greater care is needed in models that implement random numbers to simulate turbulence for ground spray applications. All tests below five minute averages yielded greater than a 99.9% confidence level that the pattern of wind velocity do not follow the model of a Random Walk. Exactly how these wind velocity values are correlated are currently unknown.

References

Baetens, K., D. Nuyttens, P. Verboven, M. De Schampheleire, B. Nicolai, and H. Ramon. 2007. "Predicting drift from field spraying by means of a 3D computational fluid dynamics model." *Computers and Electronics in Agriculture* 56: 161-173.

Box, G.E.P., G.M. Jenkins, and G.C. Reinsel. 1994. *Time Series Analysis: Forecasting and Control. 3rd ed.* Englewood Cliffs, NJ: Prentice Hall.

Charles, A., and O. Darné. 2009. "Variance Ratio Tests of Random Walk: An Overview." *Journal of Economic Surveys* 23 (3): 503-527.

Codling, E.A., M.J. Plank, and S. Benhamou. 2008. "Random Walk Models in Biology." *Journal of the Royal Society Interface* 5: 813-834.

Enders, W. 1995. *Applied Econometric Time Series .* Hoboken, NJ: John Wiley & Sons, Inc.

EPA. 2014. "Glossary." *U.S. Environmental Protection Agency.* 02 04. http://www.epa.gov/pesticides/regulating/labels/pest-label-training/glossary/.

Frederic, L., A. Verstraete, B. Schiffers, and M.F. Destain. 2009. "Evaluation of Reatime Spray Drift Using RTDrift Gaussian Advection-Diffusion Model." *Commun. Agric. Biol. Sci.* 74 (1): 11-24.

Google. 2015. *Google Maps.* https://www.google.com/maps/@42.0147298,-93.731401,396m/data=!3m1!1e3.

Holterman, H.J, J.C Van de Zande, H.A.J Porskamp, and J.F.M Huijsmans. 1997. "Modeling Spray Drift from Boom Sprayers." *Computers and Electronics in Agriculture* 19 (1): 1-22.

Lo, A.W., and A.C. MacKinlay. 1988. "Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test." *The Review of Financial Studies* 1 (1): 41-66.

MathWorks. 2015. "vratiotest." *mathworks.* Accessed 10 2, 2015. http://www.mathworks.com/help/econ/vratiotest.html.

Milton E. Teske, Sandra L. Bird, David M. Esterly, Thomas B. Curbishley, Scott L. Ray, Steven G. Perry. 2002. "AgDrift: A Model for estimating near-field spray drift from aerial applications." *Environmental Toxicology and Chemistry* 21 (3): 659-671.

Ostasiewicz, Walenty. 2000. *Socio-economical applications of statistical methods.* Wroclaw: Wroclaw University of Economics.

Smith, D B, F D Harris, and C E Goering. 1982. "Variables affecting drift from ground boom sprayers." *Transactions of the ASAE* 25 (6): 1499-1523.

Thompson, N., and A.J. Ley. 1983. "Estimating Spray Drift using a Random-walk Model of Evaporating Drops." *Journal of Agricultural Engineering Research* 28: 419-435.

Tsai, M., K. Elgethun, J. Ramaprasad, M.G. Yost, A.S. Felsot, V.R. Hebert, and R.A. Fenske. 2005. "The Washington aerial spray drift study: Modeling pesticide spray drift deposition from an aerrial application." *Atmospheric Environment* 39: 6194-6203.

CHAPTER 5. PREDICTING WIND DIRECTION FOR AGRICULTURAL GROUND SPRAYERS

A Paper to be submitted to Transactions of the ASABE

Matt Schramm[1], Mark Hanna[1], Matt Darr[1], Steven Hoff[1], Brian Steward[1]

Agricultural and Biosystems Engineering, Iowa State University[1]

Abstract

Multiple prediction schemes to predict wind direction 30 seconds into the future are presented for use with ground sprayers. Data were recorded during the late spring/summer spraying season of 2014. A five hour sample was taken from this to provide a one hour training and four hour testing dataset. A kernel filter, (commonly used to predict future wind changes for wind turbines), autoregressive, ARIMA, and hybrid models, (containing aspects of ARIMA and Taylor Series expansion), were compared to a default model (the prediction value was the value that was last observed) in their performance of predicting the win direction 30 seconds into the future. The autoregressive model yielded the lowest RMS error, when used on test data, followed by the hybrid model. These models improved prediction 9% over the default.

Keywords: *Prediction methods, spray drift reduction, ground sprayer application, turbulence model*

Introduction

Agricultural sprayers are used to provide agricultural chemicals that protect and improve crop plant health. The off-site drift of these chemicals is an environmental hazard and can be detrimental to adjacent crops or other forms of life. Many factors affect spray drift ranging from the topography of the land, current weather conditions, droplet size distribution, and the height in which the droplets are released. The EPA defines spray drift as the physical movement of a pesticide through air at the time of application or soon thereafter, to any site other than that intended for application (EPA, 2014). Spray technologies include capabilities to mitigate spray drift by altering droplet size dependent on wind conditions (Nordby & Skuterud, 1974; Smith, et al., 1982). A popular method to reduce spray drift is through the use of drift reduction nozzles which aim to decrease the number of droplets smaller than 150 µm. This is typically done either by a pre-orifice which reduces the exit velocity at the nozzle, or by an air induction nozzle that mixes air into the droplet thus making the diameter of said droplet larger (Kruckeberg, 2011).

To mitigate spray drift, models are available that predict the movement of the droplets across the field. Popular methods to model the physics of droplets or understand the distribution of spray deposition, include Gaussian, Plume, Random Walk, Regression and Computational Fluid Dynamic (CFD) models (Baetens, et al., 2007; Frederic, et al., 2009; Holterman, et al., 1997; Kruckeberg, 2011; Smith, et al., 1982; Teske, et al., 2002; Tsai, et al., 2005). Through the use of a model, real time buffer zones can be calculated to reduce the risk of spray drift (Brown, et al., 2004; Craig, 2004).

There is potential to use sprayer models and on-board computers on the sprayer, to perform simulations in real-time, and estimate spray drift potential. Kruckeberg (2011)

has shown that spray drift can be mitigated by changing to different nozzles automatically, utilizing simulation (DRIFTSIM) models (Zhu, et al., 1995). This strategy can be defeated by unforeseen changes in wind direction and wind speed after the droplet has left the nozzle. The wind power industry currently uses wind velocity predictions to forecast wind power outputs (Giebel, 2003; Zhu & Genton, 2012). A popular method used for these predictions is the kernel filter (Chan, et al., 2010). Wind power predictions are usually predicated hours in advance (Giebel, 2003).

Objective

The objective of this research was:

- Investigate the performance of statistical models in predicting wind direction in the future based on current measurements

Methods and Materials

Experimental Design and Apparatus

Field measurements of wind speed and direction were collected during the late spring/summer spraying season of 2014 using instrumentation set into a field growing oats. The field was located at the Iowa State University Research Farm's Bruner Farm field F1 near Ames, Iowa (Figure 21). The field dimensions were 268 m long by 105 m wide (880 by 348 ft).

Figure 21: Bruner Farm field F1 (42.014911 N 93.731241 W) (Google, 2015)

Wind speed and solar radiation measurements were acquired at 10 samples per second using ultrasonic anemometers (model: WindMaster 3d, Gill Instruments, Lymington, Hampshire, UK) and a pyranometer (model: SP-212, Apogee Instruments, Logan, UT). The anemometers measured the wind speed in the north-south, east-west, and vertical directions (Figure 22). Open source microcontrollers equipped with GPS modules (model: Arduino Uno, Arduino Inc; Ultimate GPS Shield, Adafruit, New York, USA) were used to log data to a micro SD. Using the GPS's PPS (Pulse per Second) output, time correction was able to be done to ensure accurate time recording of wind velocity measurements among the microcontrollers. To reduce influences to wind speed, the microcontrollers and power supplies were located separate from the anemometers at

a distance of approximately 2-3 meters. Anemometers were placed one meter above the ground's surface to collect wind measurements. The pyranometer was placed on the charging station near the anemometer.

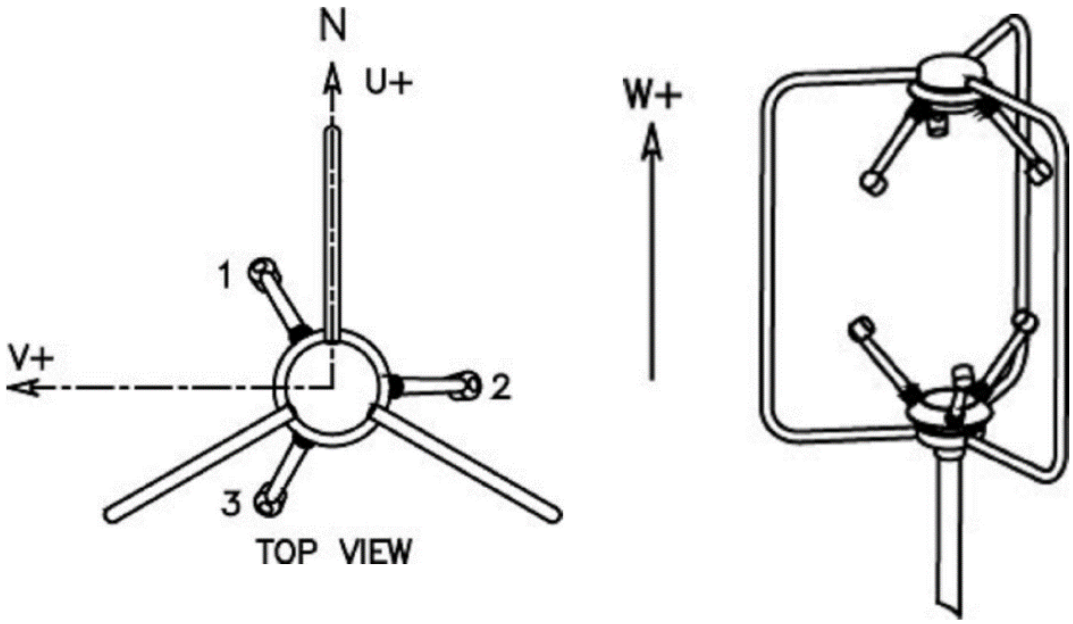

Figure 22: Ultrasonic Anemometer measuring velocity in U, V, and W component directions

Data Analysis

Wind direction was calculated using the north-south and east-west wind speeds with trigonometric relationships. A wraparound method was used to produce a semi-continuous wind direction dataset. Wraparound refers to the convention of allowing wind direction to go above 359 degrees and below 0 degrees, i.e. $12° = 372°$.

For the use of simplifying the modeling, a fixed time step (0.1 s unless otherwise noted) was assumed. From this assumption, an interpolation method was used to estimate these interior values to provide a fixed time step dataset. For example, the microcontroller may have data at 1.14, 1.23, and 1.34, but for the models interpolation is needed to give data at the times 1.1, 1.2, and 1.3. MATLAB (MathWorks), used for

analysis, offered multiple interpolation schemes ranging from discontinuous (C^0) to second order continuity (C^2) methods. MATLAB's cubic spline (C^2) offers the smoothest fit to the data, however the assumption that first and second derivatives are continuous in the wind speed data could not be confirmed. For this reason MATLAB's Piecewise Cubic Hermite Interpolating Polynomial (pchip) was not used. Ultimately, a piecewise linear interpolating spline was chosen.

A mean value smoothing function was carried out on the wind speeds to smooth out fluctuations leaving the lower frequency data that has more of an effect on spray droplets (Kruckeberg, 2011). Figure 23 shows this process during a 10 minute period on the north-south wind speed. This process is done by utilizing the nine values centered on the current data point and taking an average of the data present (Equation 21).

$$Y_m = \frac{X_{m-\frac{n-1}{2}} + \cdots + X_m + \cdots + X_{m+\frac{n-1}{2}}}{n} \tag{21}$$

where

- *n* is the number of data points being used to smooth the data,
- X_m is the mth observed data point, and
- Y_m is the mth averaged value

After smoothing the data, the data was averaged at one second intervals. This dataset was then split into a training dataset and a testing dataset. The training dataset contains the first hour of the original data set and the testing dataset is the last four hours.

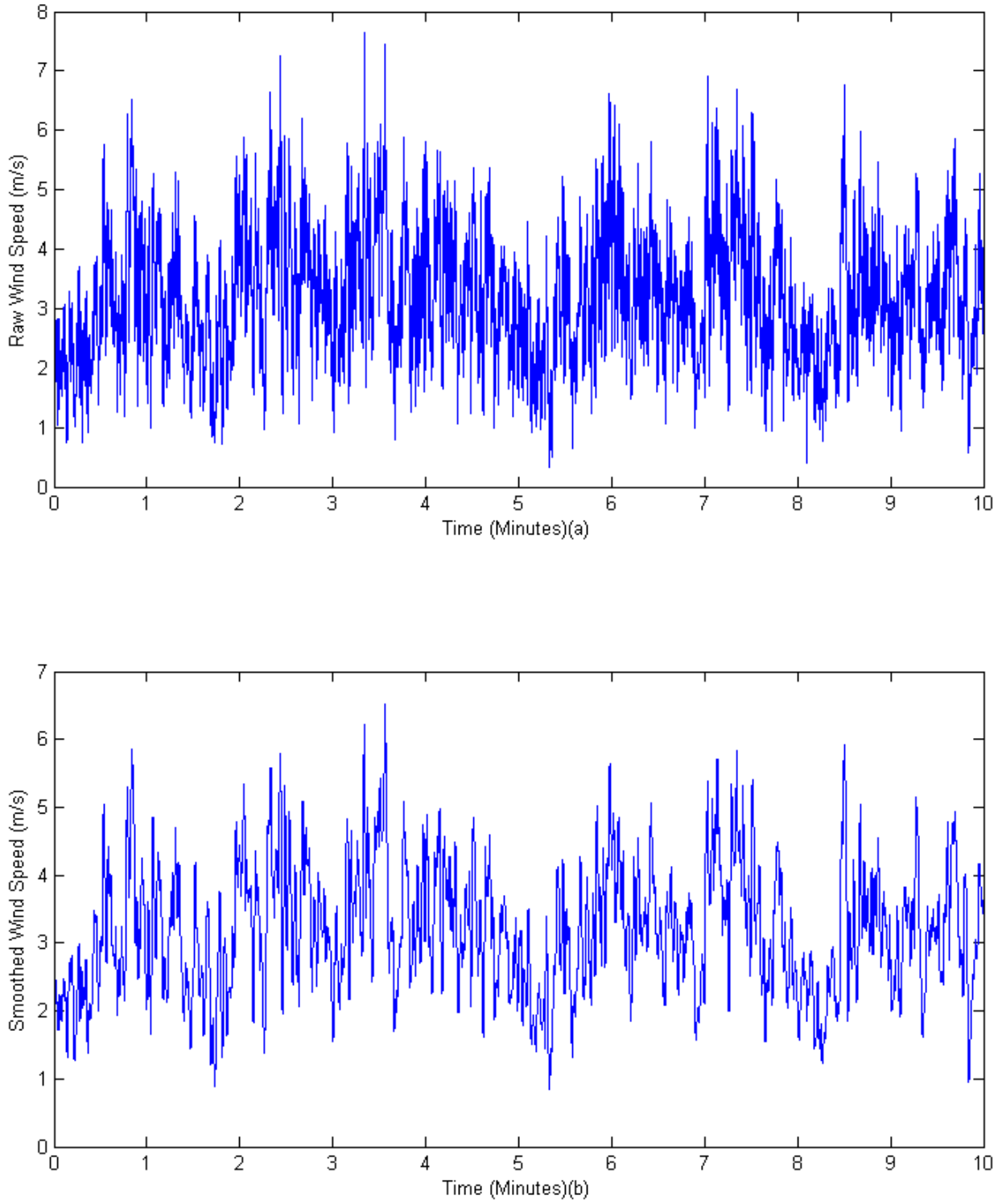

Figure 23: Smoothing function example in which the raw data (a) is smoothed using the 9-nearest neighbors to generate a new smoothed dataset (b)

Model Development

Multiple models were developed and tested. Differing models included kernel filtering, auto regressive, autoregressive integrated moving average (ARIMA), and a hybrid model using aspects of median filtering, Taylor Series Expansion, and ARIMA. The models will be compared to a default model (the “No Model”) in which the current recorded value for wind direction was used as the prediction value (Equation 22).

$$\hat{Y}_{n+30} = Y_n \tag{22}$$

Kernel Filter

For a kernel filter, the predicted value is calculated by fitting a probability of existing values that is dependent on past observed values. As the process is carried out, the current observations are compared to past values to determine how different these observations are. A weight is then calculated that depends on this difference of observed values (Equations 24 and 25). Finally, to calculate the variable that is being predicted, past values are summed with corresponding weights (Equation 23). This sum is then divided by the sum of the weights. The process used is known as the Nadaraya-Watson kernel-weighted average (Hastie, et al., 2009; Chan, et al., 2010) (Equations 23-25). Chan gives an excellent description of this process using multiple inputs, but the basis of the model is given as follows.

$$\hat{Y}(x_1, x_2, \dots, x_k) = \frac{\sum_{j=1}^{N} \left(y_j \left(\prod_{i=1}^{k} K_{\lambda_i}(x_i - x_{i,j}) \right) \right)}{\sum_{j=1}^{N} \left(\prod_{i=1}^{k} K_{\lambda_i}(x_i - x_{i,j}) \right)} \tag{23}$$

$$K_{\lambda_i}(x_i, x_{i,j}) = D\left(\frac{|x_i - x_{i,j}|}{\lambda_i} \right) \tag{24}$$

$$D(\xi) = \begin{cases} \frac{3}{4}(1 - \xi^2), & if\ |\xi| \leq 1 \\ 0 \quad , & \text{Otherwise} \end{cases} \tag{25}$$

where

- $\hat{Y}(X)$ is the new predicted value based on the X observations,
- y_i are the past observed target values that $\hat{Y}$ tries to predict,
- X contains the inputs to the kernel ($X = [x_1, x_2, \dots, x_k]$),
- The summation goes over the N past observations for all inputs X with observed target value y_j, and
- $K_\lambda(x_i, x_{i,j})$ is the kernel being used with constant λ and weighting function *D*

For kernel filtering, past wind speed, wind direction, solar radiation, changes of wind direction, and changes in wind speed were given their own λ. The λ values were changed to find the minimum RMS error (Equation 26) to optimize the model's predictive ability using the training dataset.

$$RMS_{error} = \sqrt{\frac{\sum_{i=1}^{N} Error_{Wd}^2}{N}} \tag{26}$$

where

- $Error_{Wd} = Wd_p - Wd_a$
- Wd_p is the predicted value of wind direction

- Wd_a is the actual wind direction

This minimization was done using MATLAB's fminunc function (unconstrained minimization) using the Quasi-Newton's Method.

Autoregressive

The prediction utilizing autoregressive method took a simpler form to predict wind direction 30 seconds into the future (Equation 27). With this method, the past 30 seconds of data were used to predict the next value.

$$\hat{Y}_{n+30} = a_0 Y_n + a_1 Y_{n-1} + \cdots + a_N Y_{n-N} + a_{N+1} \quad (27)$$

where

- Y_n is the observed wind direction at time n,
- $a_0, a_1, \ldots, a_{N+1}$ are constants used to weight the observations, and
- $N = 30$

For the auto-regressive prediction scheme, the constants were optimized to reduce the RMS error on the training dataset.

ARIMA

ARIMA takes the previous prediction scheme (auto-regressive) and expands onto it (Equation 28).

$$\left(1 - \sum_{k=1}^{p} \phi_k L^k\right)(1 - L)^d Y_t = \delta + \left(1 + \sum_{k=1}^{q} \theta_k L^k\right)\epsilon_t \quad (28)$$

where

- Y_t is the next observed value,
- δ is a drift constant which was set to zero,
- ϕ_k is an auto-regressive constant that acts on the observed value Y_{t-k},
- θ_k is a moving average constant that acts on ϵ_t,

- L^k is a lag operator, e.g. $L^k Y_t \equiv Y_{t-k}$,
- ϵ_t the error of the prediction value and the observed value and is distributed with mean zero and standard deviation of σ,
- p and q are the number of auto-regressive and moving average terms (p = q = 40), and
- d is the number of times the difference of the data is taken (d = 0)

For the prediction scheme described, the first 30 constants for the auto-regressive and the moving average terms were taken to be zero. This insured the next observed value would be the prediction of wind direction, 30 seconds into the future. For ARIMA, the error was not minimized to find the constants $\phi_k \text{ and } \theta_k$, but the log-likelihood function was maximized (Box, et al., 1994; Enders, 1995; MathWorks, 2015) while using the training dataset.

Hybrid

The hybrid method takes the form (Equation 29). This method is the result of trial and error, incorporating different ideas of modeling.

$$\hat{Y}_{n+30} = W_1^{a_1} W_2^{a_2} W_3^{a_3} \tag{29}$$

Where W_i have the following structures (Equation 29-32)

$$W_1 = a_4 Y_n + a_5 Y_{n-1} + \cdots + a_{10} Y_{n-6} + a_{11} X_n \tag{30}$$

$$W_2 = Y_n + a_{12}(Y_n - Y_{n-1}) + a_{13}(2Y_{n-1} - Y_n - Y_{n-2}) + a_{14}\left(Y_n - Y_{n-3} + 3(Y_{n-2} - Y_{n-1})\right) \tag{31}$$

$$W_3 = a_{15} median(Y_n : Y_{n-30}) \tag{32}$$

Where equation 26 takes the form of an ARIMAX model using the auto-regressive terms of the standard ARIMA model with the addition of an *X* term (Equation 33) (Box, et al.,

1994; MathWorks, 2015).

$$X_n = \frac{Ws_n - Ws_{n-1}}{mean(Ws_n : Ws_{n-30})} \tag{33}$$

where

- Ws_n is the wind speed at time n, and
- The mean is taken from time n to time n-30

Equation 31 contains the first four terms of a Taylor Series expansion utilizing numerical backward differences to estimate the first, second, and third derivatives with differing constants to better estimate the future wind direction. Equation 32 is a median filtering function over the past 30 seconds with an arbitrary constant. The constants were found by minimizing the RMS error using quasi-newton unconstrained minimization on the training dataset.

All models were developed on the training data set by running multiple times starting at different initial values. The training dataset contains the first hour of data of the five hour dataset shown in Chapter Three. These starting values were chosen at random from a uniform distribution to find the lowest RMS error values. The process searched for the coefficients that would minimize RMS error. After coefficients were found, the models were applied to the testing dataset and the RMS error was calculated. The testing dataset contained the remaining four hours of the five hour dataset.

Results and Discussion

Table 1 shows these RMS errors along with the 2-norm errors (Equation 34).

$$L^2 = \sqrt{\sum Error_{Wd}^2} \tag{34}$$

The 2-norm is used as an addition metric of the error that is often used in numerical

analysis. Table 1 shows No Model performed the worst, with other methods, notably Autoregressive and Hybrid, having lower errors.

Table 2: RMS and Norm 2 Errors for the differing prediction methods

Error	Method				
	Kernel	Autoregressive	ARIMAX	Hybrid	Default
RMS (Degrees)	14.19	12.97	17.10	13.01	14.30
2-Norm (Degrees)	1,681	1,552	3,248	1,556	1,692
Percent Out of Tolerance (20°)	15.15%	11.73%	12.28%	11.90%	15.65%

From table 2, the autoregressive model yielded the lowest RMS and 2-Norm errors with the hybrid model following close behind. The ARIMAX model does do poorly for the measures of RMS and Norm error but does do better using the percent out of tolerance metric, however it does not do better than the Autoregressive nor Hybrid models. Figure 24 shows how each method did predicting values within a certain tolerance of the testing dataset. The ARIMAX model is not shown due to the model needing many simulations to get an estimate of the average path. Here there is tight grouping for both the autoregressive and hybrid model to the point it is hard to distinguish between the two. This is also the case for the no model and kernel models. The hybrid model is the most general model due to the constants not depending on an average wind direction.

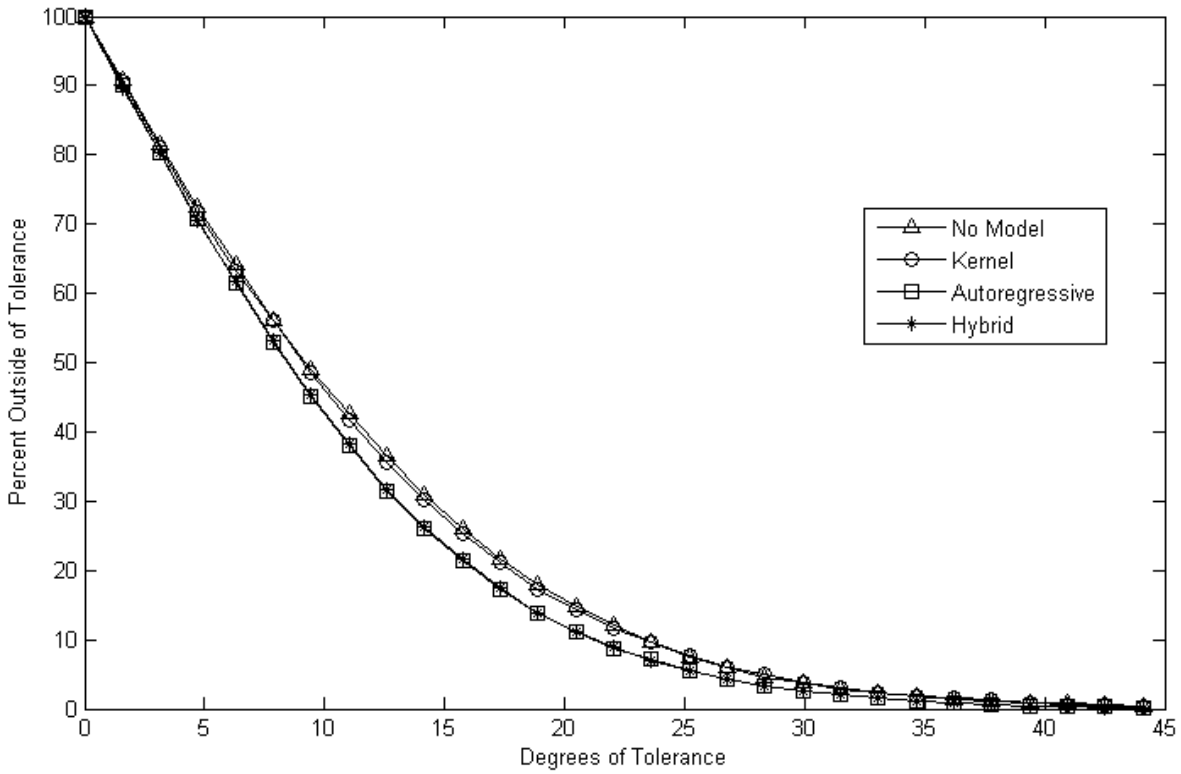

Figure 24: Percent outside of tolerance for differing methods

Figure 25 shows the results of the ARIMAX model after running 1,000 simulations and finding the average path taken. Figure 25 shows the last 10 minutes of training data along with the first 10 minutes of the simulations and testing data. The ARIMAX method that was used was a ARIMAX(40,0,40) method with two additional x terms (the change in wind direction and the change in wind speed). Note, the first 30 coefficients for both the autoregressive and moving average are zero. Like the hybrid model, the ARIMAX model did not contain a constant value.

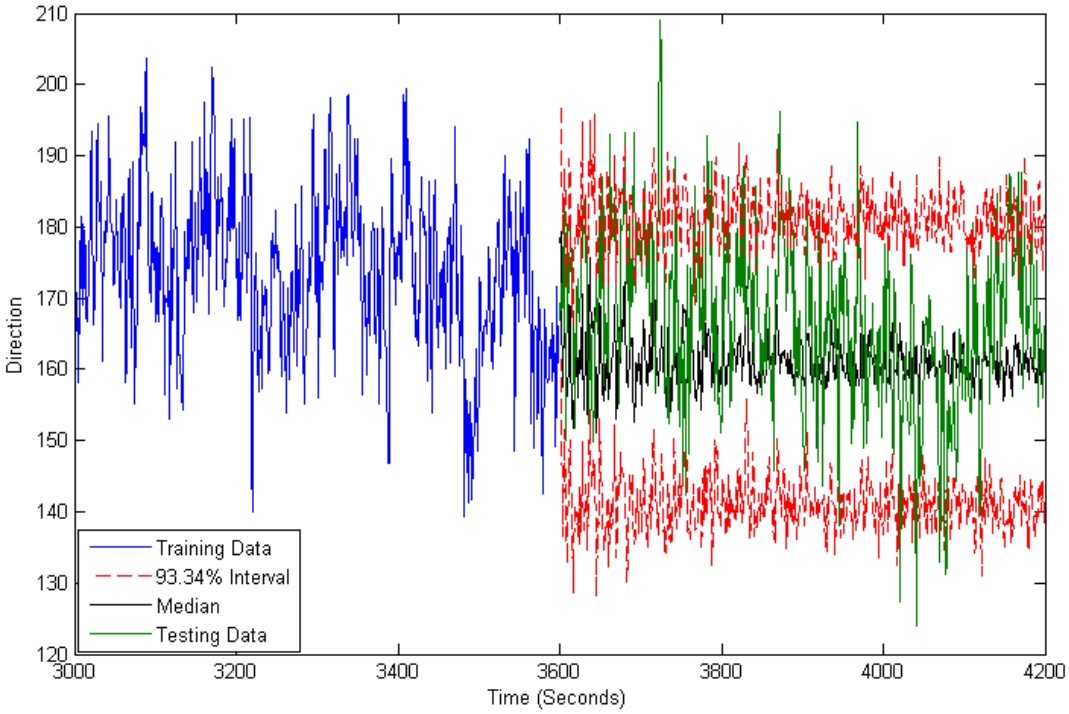

Figure 25: 1000 Simulations of an ARIMAX(40,0,40) model with x containing information on changes in wind direction and wind speed. The simulation holds the average path taken (black line) contained by a 93.34% of possible paths (red dashed lines).

The bounds that are in figure 25 (dashed lines) represent a $20°$ spread in the data. With this spread a comparison can be made with the other methods. As seen in figure 25, the method bounds the testing data well and contained the testing dataset all but 12.28% of the time. However, if each path is tracked separately, the ARIMAX model does do poorly as reflected in table 2.

Conclusion

Four prediction schemes (Kernel, Autoregressive, ARIMAX, and Hybrid) were investigated and compared to a default. All models did achieve RMS errors lower than the default with the Autoregressive model yielding the lowest error, and the ARIMAX

model also reducing the number of times predictions varied from actual values by more than 20 degrees. The hybrid model yielded the most universal results due to the model needing no information of the averaged wind speed. Models were trained and tested on a dataset during a period of wind moving generally south to north. It is unknown how models would react to more turbulent conditions.

References

Baetens, K., D. Nuyttens, P. Verboven, M. De Schampheleire, B. Nicolai, and H. Ramon. 2007. "Predicting drift from field spraying by means of a 3D computational fluid dynamics model." *Computers and Electronics in Agriculture* 56: 161-173.

Box, G.E.P., G.M. Jenkins, and G.C. Reinsel. 1994. *Time Series Analysis: Forecasting and Control. 3rd ed.* Englewood Cliffs, NJ: Prentice Hall.

Brown, R.B., M.H. Carter, and G.R. Stephenson. 2004. "Buffer Zone and Windbreak Effects on Spray Drift Deposition in a Simulated Wetland." *Best Management Science* 60 (11): 1085-1090.

Chan, C.P., J.R. Stalker, A. Edelman, and S.R. Connors. 2010. "Leveraging High Performance Computation for Statistical Wind Prediction." *American Wind Energy Association Wind Power Conference and Exibition* Dallas, Texas, USA, May 23-26.

Craig, I P. 2004. "The GDS Model - A Rapid Computational Technique for the Calculation of Aircraft Spray Drift Buffer Distances." *Computers and Electronics in Agriculture* 43 (3): 235-250.

Enders, W. 1995. *Applied Econometric Time Series .* Hoboken, NJ: John Wiley & Sons, Inc.

EPA. 2014. "Glossary." *U.S. Enbironmental Protection Agency.* 02 04. http://www.epa.gov/pesticides/regulating/labels/pest-label-training/glossary/.

Frederic, L., A. Verstraete, B. Schiffers, and M.F. Destain. 2009. "Evaluation of Reatime Spray Drift Using RTDrift Gaussian Advection-Diffusion Model." *Commun. Agric. Biol. Sci.* 74 (1): 11-24.

Giebel, G. 2003. *The State-Of-The-Art in Short-Term Prediction of Wind Power.* A Literature Overview, ANEMOS.

Google. 2015. *Google Maps.* https://www.google.com/maps/@42.0147298,-93.731401,396m/data=!3m1!1e3.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* New York: Springer.

Holterman, H.J, J.C Van de Zande, H.A.J Porskamp, and J.F.M Huijsmans. 1997. "Modeling Spray Drift from Boom Sprayers." *Computers and Electronics in Agriculture* 19 (1): 1-22.

Kruckeberg, J.P. 2011. "An automated nozzle controller for self-propelled sprayers." Graduate Theses and Dissertations. Paper 12083.

MathWorks. 2015. "arima class." *MathWorks.* Accessed Oct 3, 2015. http://www.mathworks.com/help/econ/arima-class.html.

Nordby, A, and R Skuterud. 1974. "The effects of boom height, working pressure and wind speed on spray drift." *Weed Research* (Weed Research) 14 (6): 385-395. doi:10.1111/j.1365-3180.1974.tb01080.x.

Smith, D.B., F.D. Harris, and C.E. Goering. 1982. "Variables Affecting Drift from Ground Boom Sprayers." *Transactions of the ASAE* 25 (6): 1499-1523.

Teske, M.E., S.L. Bird, D.M. Esterly, T.B. Curbishley, S.L. Ray, and S.G. Perry. 2002. "AgDrift: A Model for estimating near-field spray drift from aerial applications." *Environmental Toxicology and Chemistry* 21 (3): 659-671.

Tsai, M., K. Elgethun, J. Ramaprasad, M.G. Yost, A.S. Felsot, V.R. Hebert, and R.A. Fenske. 2005. "The Washington aerial spray drift study: Modeling pesticide spray drift deposition from an aerrial application." *Atmospheric Environment* 39: 6194-6203.

Zhu, H., D.L. Reichard, R.D. Fox, H.E. Ozkan, and R.D. Brazee. 1995. "DRIFTSIM, A Program to Estimate Drift Distances of Spray Droplets." *Applied Engineering in Agriculture* 11 (3): 365-369.

Zhu, X., and M.G. Genton. 2012. "Short-Term Wind Speed Forecasting for Power System Operations." *International Statistical Review* 80 (1): 2-23.

CHAPTER 6. CONCLUSIONS

From the research described in this thesis, we can draw the following conclusions:

- Chapter 3. Measuring Sub-Second Wind Velocity Changes at One Meter
    - During a five hour period with an average wind speed of 3.6 m/s
        - Correlation Coefficients of 0.29 (0.27) between an upwind and downwind sensor, separated by 30m), were found for wind direction (speed).
        - Correlation was less during shorter one-minute periods in which a spray droplet may travel, but improved to a coefficient of 0.15 if a lag time was used between the two sensors.
        - Using a lag time, downwind direction was greater than 20 degrees different than the upwind sensor 30% of the time while wind speed was greater than 1 m/s (about a quarter of the mean wind speed) different than the upwind speed about 50% of the time.
    - During a 1.5 hour period with an average wind speed of 1.5 m/s
        - Correlation Coefficients less than 0.03 were found between upwind and downwind sensors, separated by 15 or 30 m.
        - Downwind direction was greater than 20 degrees different than the upwind direction 65% of the time, while the downwind speed was greater than 0.25 m/s (about a quarter of the mean wind speed) different than the upwind speed 80% of the time.
    - Across a range of late spring/summer days in which suitable conditions for ground spraying were present, significant change in wind direction 30

seconds later was more likely to occur during wind speeds in the range of 0-3 m/s (0 -6.7 mi/h)

- Chapter 4. A Not-so-Random-Walk with Wind: A Look at Random Fluctuations in Wind Velocities for use in Models of Agricultural Spray Drift
    - Wind speed in the north-south, east-west, and vertical directions, wind magnitude, and wind direction were found to not follow the random walk model at 99.9% certainty for time averaged values less than five minutes.
    - Model that accurately describes wind is currently unknown.
- Chapter 5. Predicting Wind Direction for Agricultural Ground Sprayers
    - An autoregressive model, utilizing 30 seconds of past values, and a hybrid model, utilizing ideas from ARIMA and Taylor series expansions, produced the lowest RMS values when compared over a four hour testing period.
    - The above models lowered the error RMS value by 9% over the “No Model” model.

Correlation values between sensors were low, indicating little linear correlation, and the prediction models had limited success. Although lack of correlation and difficulty with prediction suggests random variation, results of chapter 4, indicate the changes in wind velocity are not purely random and that changes in the velocity are related. Chapter 3 showed that a greater likelihood of significant wind directional change occurs during periods of lower wind speeds. Wind speeds seem to have greater effect on wind direction change than time of day or solar radiation.

Recommendations for further research

- Reevaluate existing predictive spray drift models utilizing the updated information about downwind velocity distributions. This will provide better approximations the droplet sees during a simulation, thus the model should have better agreement with observations.

In chapter 4, it was shown than an assumption that the next wind velocity measurement can be predicted from the current measurement plus some random fluctuation, did not agree with observations. It is suggested that differing models/tests in time series statistics be investigated. MATLAB and R, an open source statistical package, offer many time series statistical functions that test many aspects of said time series. This may provide a deeper understanding into how to better approximate wind velocities. An example of a test is the Leybourne-McCabe stationarity test that tests the null hypothesis that the time series can be described by some autoregressive process (AR(p)), versus the alternative hypothesis that the time series is some nonstationary ARIMA(p,1,1) process. It is also suggested that other ARIMA models be tried in place of the Random Walk model, and tested. It is currently beyond the scope of this work to determine if there exists a way to test if a time series can be modeled by a general ARIMA model. An example of a model to test is the ARIMA(0,0,0), in which the wind velocities are modeled by some mean wind speed plus some random process.