8-23-2017

An Accuracy Study of RTK GNSS Positioning Applied to Comparing Maps for Bentgrass Habitat Modeling

Jinwon Chung
University of Connecticut - Storrs, jinwon.chung@uconn.edu

An Accuracy Study of RTK GNSS Positioning Applied to Comparing Maps for Bentgrass Habitat Modeling

Jinwon Chung, PhD

University of Connecticut, 2017

Theory and previous research suggest that sky obstructions and humidity can degrade global navigation satellite system (GNSS) positioning accuracy for static observation sessions. It is reasonable to suppose that these effects might be even worse for real-time kinematic (RTK) positioning, because the observer will likely collect fewer observations for RTK positioning than for a static occupation, so the statistics used to estimate the positions have fewer data to work with. These effects have not been thoroughly studied, which provided the motivation to conduct several experiments to quantity the effects. Temperature and humidity are variables of interest, so ***the first experiment establishes whether a digital weather station is an acceptable replacement for a sling psychrometer***. ***The second experiment quantifies RTK positioning accuracy affected by broad-leaf canopy conditions with the effect of ground-level absolute humidity and the effect of sky obstruction*** as determined using analysis of covariance**;** this is to study the RTK position-accuracy degradation caused by the water content in the atmosphere and by the possible signal blockage from physical structures around the occupation site. These results were then applied to extend previously published work about a work on habitat-suitability and environmental favorability maps for bentgrass species in Connecticut with logistic regression analysis from GNSS data. The information gained from the experiments was used to study four new biological habitat suitability and environmental favorability models by comparing their

strengths and weaknesses using GIS mapping and multiple comparison statistics of the created maps.

An Accuracy Study of RTK GNSS Positioning Applied to Comparing Maps for Bentgrass Habitat Modeling

Jinwon Chung

B.A., Korea University, 1999

M.N.R., North Carolina State University, 2005

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2017

APPROVAL PAGE

Doctor of Philosophy Dissertation

An Accuracy Study of RTK GNSS Positioning Applied to Comparing Maps for Bentgrass Habitat Modeling

Presented by
Jinwon Chung, B.A., M.N.R.

Major Advisor

_____
Thomas H. Meyer

Associate Advisor

_____
John C. Volin

Associate Advisor

_____
Jason Vokoun

Associate Advisor

_____
Chuanrong Zhang

University of Connecticut

2017

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF APPENDIX

ACKNOWLEDGEMENT

Studying and pursuing a doctoral degree at the University of Connecticut has been one of the most challenging periods of my life. Writing my dissertation was very difficult because English is not my native language. To dive into the world of geodesy and global navigation satellite systems (GNSS), I had to intensively study mathematics and statistics as well as new programming languages including *Mathematica* and *R*. Everything was quite new and it took an enormous amount of time to understand and absorb the knowledge needed for my research and education. For my successful completion of the doctoral program, I owe much to the help and support of many individuals who guided and supported me these past years.

Above all, I would like to express great gratitude to my advisor, Professor Thomas H. Meyer. My knowledge and accomplishments as a scientist were made possible by his enthusiastic teaching, guidance, and patience. He was my mentor who helped me past the obstacles and pushed me to reach my potential.

I would also like to express special thanks to my associate advisors, Professor John Volin, Professor Jason Vokoun, and Professor Chuanrong Zhang. They provided valuable advising and comments throughout my research. I have greatly appreciated the various perspectives of my associate advisors - ecology from Dr. Volin, research and quantitative methods from Dr. Vokoun, and spatial statistics and modeling from Dr. Zhang. I also would like to thank Dr. Daniel L. Civco, who is currently a Professor Emeritus and formerly my associate advisor. Even after his retirement, he remained active on my committee and his teachings about geographic information systems (GIS) were important for my research and my development as a scientist.

I am thankful for the Connecticut Department of Transportation (CTDOT) which, in collaboration with the Department of Natural Resources and the Environment, established the Advanced Continuously Operating Reference Network (ACORN). ACORN made my GNSS surveying work possible.

I would like to express thanks to Professor Carol Auer and my co-workers, Dr. Geoffrey Ecker and Dr. Collin Ahrens, who were my team members for the bentgrass project, which served as the basis for the final chapter of my dissertation. Geoffrey and Collin endured harsh field conditions with me to collect the data needed for my research.

I am grateful for the encouragement and support of my parents who helped me through the challenges of my studies - they have my deepest gratitude and love.

Finally, I thank God for always being with me and giving me the strength to get through difficult times.

Thank you very much!

Jinwon Chung

August 22, 2017.
From Storrs, Connecticut.

Chapter 1

Introduction

1.1 Overview

Prior to widespread civilian access to the US NAVSTAR Global Positioning System (GPS), land surveyors could not produce maps easily in formal coordinate systems, such as the State Plane Coordinate System (Stem, 1989). Maps were paper, and tax accessors determined taxes using rulers and protractors. Geographic information systems (GIS) offered the possibility of creating fully digital cadasters that could show entire towns – or even an entire State – at any scale as desired on-the-fly, creating maps on demand, but only if all the maps were in the same coordinate system. Even better, GIS allows users to analyze their data by combining the power of relational databases with custom-built spatial analysis tools. GIS created a pressing need for maps to be created using GPS.

For land surveying, the original global navigation satellite system (GNSS) positioning mode, called **static surveying**, was to erect a GPS receiver on a tripod or a range pole. The receivers recorded the phase shift of the carrier wave to each observed satellite. A static-occupation survey involves deploying several receivers and allowing them to collect observations for periods as short as 20 minutes or as long as several hours (Connecticut Association of Land Surveyors, 2008; Eckl, Snay, Soler, Cline, & Mader, 2001). The observations can be adjusted as a network, and the abundance of data allows the least-squares

routines to be highly robust and accurate (Crandall, 1906; Ghilani, 2010; Harvey, 2009; Mikhail & Gracie, 1981; Stigler, 1981; Van Sickle, 2008; Wolberg, 2006). Centimeter-level positioning accuracy is routinely achievable for a well-planned and correctly-executed static-occupation survey (U.S. Army Corps of Engineers, 2011; Zilkoski, D'Onofrio, & Frakes, 1997; Zilkoski, Carlson, & Smith, 2008). Positioning with phase requires at least two receivers to be collecting data simultaneously, so surveyors owned and deployed multiple receivers.

Although GPS surveying can be much more efficient than traditional survey methods, there are three drawbacks with surveying static positioning: first, it is slow; second, it requires several field personnel – at least two; and, third, it requires expensive equipment. These were remedied when real-time positioning was developed. This positioning mode is called **real-time kinematic** (RTK). RTK surveying requires deploying two receivers: a **base station** that is erected by the surveyor on a control marker near the survey area, plus a **rover** that is carried about by the surveyor to do work. The base station communicates with the rover via some kind of radio link, such as ultra-high frequency (UHF) or very high frequency (VHF) radios. The base station sends its position and observables to the rover, which allows the rover to calculate corrections on-the-fly resulting in centimeter-accuracy positions (Henning, 2011; Henning, Martin, Schrock, Thompson, & Snay, 2013). This remedied the first drawback: it allows surveyors to position features in only a few seconds. The second drawback was remedied because a surveyor can work without field-crew assistants. The third drawback was remedied when permanent networks of base stations were established and made accessible to the public. Such a network is called a **real-time network** (RTN). The University of Connecticut's Department of Natural Resources and the Environment (NRE) and the Connecticut Department of Transportation (CTDOT) have built and operate a RTN, named the Advanced Continuously

Operating Reference Network (ACORN), which provides surveying professionals statewide access to RTK via the Internet. A RTN eliminates the need for a surveyor to deploy a base station, as well as hiring someone to monitor it during the survey.

The primary vulnerability of RTK is that the receiver's position is determined from a single base station, which is called **radial surveying**. If there is a mistake, there is no check available to detect it. Deploying a network of receivers for a long static observation session is better (for example, see Zilkoski *et al.* (1997)) because the observables can be post-processed and adjusted using least squares methods, which are highly robust (Ghilani, 2010), but this cannot be done in real time. Another vulnerability of RTK is that usually the receiver collects fewer observations for RTK positioning than for a static occupation, so the statistics used to estimate the positions have fewer data. The accuracy of network surveying is well studied, but the accuracy of RTK, especially in the relatively unfavorable environment common in Connecticut's leaf-on season, has not been well studied. Throughout the study, positioning data were collected by multi-constellation receivers. The term **global navigation satellite system** is generic; for example, the GPS is the GNSS owned by the United States. Other GNSSs are *Globalnaya Navigatsionnaya Sputnikovaya Sistema* (*GLONASS*), which is operated by Russia, *Galileo*, which is owned by European Union, and *Compass*, which is owned by China. The receivers used in this research can accept signals from the GPS and GLONASS constellations (Galileo and Compass are not operational yet), so here the term **GNSS receiver** is preferred over **GPS receiver**.

The whole research is divided into two parts: GNSS practice and application. In the GNSS practice part, the main purpose is to quantify the accuracy degradation of real-time GNSS positioning under leaf-on and leaf-off tree canopy status controlling the ground humidity factor

and controlling sky obstruction factor related to the blockage of GNSS radio signal propagation. Temperature and relative humidity are the essential factors for determining absolute humidity, and a digital handheld weather station was used to collect these weather factors. So a comparison between a digital handheld weather station and a sling psychrometer for their temperatures and relative humidities was undertaken to test feasibility of using the digital handheld weather station. The GNSS application part was compares digital habitat modeling maps of bentgrass created from GNSS positioned point data (both static and real-time according to the field situations) with parametric and nonparametric multiple comparison analyses to find better predictive ability of the statistical models.

1.2 Background

1.2.1 Relationships of temperatures and relative humidities between a digital handheld weather meter and a sling psychrometer

A device measuring humidity is called a hygrometer, and there are various types of hygrometer according to their operational principles, such as the metal-paper coil hygrometer, the hair-tension hygrometer, and the psychrometer (Ackerman & Knox, 2006; Ahrens, 2012; Harold, 1952; Wiederhold, 1997). Among these types of hygrometers, a sling psychrometer is simple and portable unit that might be appropriate for field data collections. A sling psychrometer is combination of two glass tube thermometers with identical scale and filled with alcohol or some other chemicals fixed together, and these thermometers are linked to a handle via pivot or chain. A bulb of one of the thermometers is wrapped in a wick; the wick will be wet

while humidity measurement, and this thermometer is called a wet-bulb thermometer. The temperature measured by this wet-bulb thermometer is called a wet-bulb temperature. The other bulb of the sling psychrometer is just exposed to the air, and this side of thermometer is called a dry-bulb temperature. Also, the temperature measured by this dry-bulb thermometer is called dry-bulb temperature. To measure a relative humidity, the wick on wet-bulb has to soak in distilled water, and whirl the device at least one minute and no more than three minutes (Kilby, McManus, & Cumberland, 1993; Thomson, 1986). While whirling, water on the wick evaporates and deprives the heat from the wet-bulb (Bohren &Albrecht, 1998; Tsonis, 2007), thus the wet- and dry-bulb temperatures are different after the operation. Relative humidity can be obtained from a psychrometric charts with the dry-bulb temperature and the difference between dry- and wet-bulb temperatures, or it is possible to calculate relative humidity with a series of psychrometric equations which is a complicated alternative (Bohren and Albrecht, 1998; Tsonis, 2007; Wallace and Hobbs, 2006). A sling psychrometer is the choice as the standard instrument for comparison for this study because of a sling psychrometer is generally accepted by many peer researchers and professionals, and also because of the stability and portability of a sling psychrometer.

Modern electronic hygrometers often contain digital sensors. These sensors include thermal humidity sensor, gravimetric humidity sensor, optical humidity sensor, resistive humidity sensors, and capacitive sensor (Fraden, 2015; Kulwicki, 1991; Lee & Lee, 2005; Rittersma, 2002; Roveti, 2001; Yamazoe & Shimizu, 1986). Kestrel hand-held weather stations has a polymer capacitive sensor for humidity measurement and a thermistor to ambient temperature measurement (Nielsen-Kellerman, 2010). These hand-held weather stations are small, light-weight, easy to use, and, inexpensive, so they are considered ideal for surveyors to

determine meteorological conditions in the field. The research team tested to verify the credibility of the Kestrel 3000 for humidity and temperature, so to see whether the unit is accurate enough to support the surveying fieldwork. For this verification, the statistic relationship between the environmental data from Kestrel 3000 and the data from a sling psychrometer are quantified and compared.

1.2.2 Accuracy studies of real-time GNSS positioning

Positioning accuracy degrades (weakly) with increasing base station-rover separation, so there is a limit to how far the rover can move from the base station (Eckl *et al*., 2001). For example, the ACORN administrators recommend that users remain within the State borders or on Long Island Sound, which limits the baseline length to 20 km at most (Thomas Meyer, pers. comm. 2016). The ideal situation would be for the base station to follow the rover around during the survey, remaining always within a few meters of the rover. Obviously, this is impossible with a physical base station, which must remain statically on its control point. However, virtual reference station (VRS) positioning makes it possible to overcome this shortcoming related to the baseline distances, at least in principle. A VRS is a synthetic, virtual base station realized only several meters, or even less, from the rover, so the distance from the VRS to the rover is extremely short. Insofar as the VRS's synthetic observables faithfully mimic the observables a real, physical receiver would collect at that place, the very short distance from base to rover should result in more accurate positions than those derived from more distant RTK base stations (Hofmann-Wellenhof, Lichtenegger, & Wasle, 2008; Landau, Vollath, & Chen, 2009).

Compared to static positioning, real-time positioning is relatively new, and virtual real-time positioning newer still. The error budgets of real-time positioning are not well understood compared to those of static positioning. It is known that broadleaf canopies present severe challenges for GNSS positioning (Meyer, Bean, Ferguson, & Naismith, 2002). Therefore, an experiment that GNSS real-time positioning on control points beneath tree canopies of varying amounts of sky obstruction is conducted to assess the effects of humidity and sky obstruction.

1.2.3 Map comparisons with various creeping bentgrass habitat maps

GNSS positioning was used to map sampling locations in a creeping bentgrass study in cooperation with the Department of Plant Science. Herbicide-resistant creeping bentgrass (HRCB) is a genetically modified organism (GMO) which is favored by golf-course managers because of its cost benefits and simplification of landscape management. Ecologists investigate whether gene flow from GMOs into "feral" members of the same species is possible, and physical separation of the organisms is a barrier to gene flow (Abud *et al*., 2007; Armstrong, Fitzjohn, Newstrom, Wilton, & Lee, 2005; Chandler & Dunwell, 2008; Damgaard & Kjellsson, 2005). Therefore, Ahrens, Chung, Meyer, and Auer (2011) produced a habitat suitability map (HSM) and an environmental favorability map (EFM) of creeping bentgrass around a golf course in northern Connecticut to shed light on the locations and possible migrations of the GMO bentgrass into nearly populations. These maps were created using spatially explicit logistic regression, a statistical model enjoying global support. (A statistical model has global support when a single equation models the whole study area regardless of the spatial relationships between data points. In contrast, a statistical model has local support when each data point of the

model has its own equations with parameter-estimate values to do reflect the points' spatial relationships (Fotheringham & Wegener, 1999; Fotheringham, Brunsdon, & Charlton, 2003).)

Ahrens' maps were created using global-support logistic regression models, which begs the question whether local-support models might be superior for predicting actual bentgrass presence/absence. There are four maps to compare: global-support HSM, global-support EFM, local-support HSM and local-support EFM. A geographically weighted logistic regression (Atkinson, German, Sear, & Clark, 2003; Fotheringham *et al*., 2003; Rodrigues, de la Riva, & Fotheringham, 2014; Saefuddin, Setiabudi, & Fitrianto, 2012) is a reasonable choice for the new maps. Differences among these maps and relative predictive strength of each map for presence/absence of bentgrass are compared using methodologies presented by prior researchers (Wang, Zhang, & Li, 2012; Wang, Zhang, & Li, 2013; Wang, Zhang, Li, Lin, & Zhang, 2014) and by using a receiver operating characteristic analysis (Fawcett, 2006; Metz, 2000; Van Erkel & Peter, 1998), a χ^2 test (Agresti & Kateri, 2011; Zar, 2010), Cochran's Q test (Cochran, 1950), and Friedman's rank sum analysis (Friedman, 1937).

1.3 Problem Statements and Research Questions

Question 1: Temperature and relative humidity measured from a digital handheld weather station are statistically compared with temperature and relative humidity measured from a sling psychrometer. The temperature and relative humidity from the digital handheld weather station have a linear relationship having zero intercept and one slope with temperature and relative humidity acquired from sling psychrometer? In other words, are weather variables from digital handheld weather station credible and negligible for errors compared to its analog counterpart? Chapter 2 addresses the first question.

Question 2: *Controlling for ground-level absolute humidity, how does tree canopy status affect GNSS positioning accuracy*? Accuracy is assessed by statistically analyzing positioning errors observed at control markers with absolute humidity measured simultaneously with each GNSS observation in two seasons with leaf-on and leaf-off conditions respectively. Previous work (Meyer, Arifuzzaman, & Massalski, 2010) established high-accuracy control coordinates for survey markers throughout Connecticut. Twenty-three of these markers were occupied during leaf-on and leaf-off conditions, observing with GPS-only and with multi-constellations (GLONASS), and observing with RTN and VRN modes. Humidity was measured using a Kestrel meteorological instrument. Chapter 3 offers treatments and statistical analyses about the second question.

Question 3: *Controlling for percent sky obstruction, how does tree canopy status affect GNSS positioni*ng accuracy? Accuracy is assessed by statistically analyzing positioning errors observed at control markers with various levels of sky obstruction (Parent & Volin, 2014) during leaf-on and leaf-off conditions. Vertical photographs of the canopy at each survey marker, taken with camera fitted with a hemispherical lens, was analyzed to quantify the sky obstruction (Meyer *et al*., 2002). Chapter 3 also offers treatments and statistical analyses about the third question.

Question 4: How can continuous raster maps produced by statistical models be compared to find their differences, and how can categorical scale maps produced by equal models be compared to determine which model produces the best result predicting actual bentgrass presence/absence in the field? The life patterns of feral and GM bentgrass will probably be similar, so a habitat suitability model for GM bentgrass can be created with logistic regression using field observations of feral bentgrass. There are several methods to make a habitat

suitability map. Maps created from each method are quantitatively appraised with spatial statistics (Wang et al., 2012; Wang et al., 2013; Wang et al., 2014) and multiple comparison analyses. Each method for map creation requires random field samples for its input data, and the samples were located using GNSS positioning. Chapter 4 provides statistical tests for multiple comparisons and binary-classification analysis to answer the fourth question.

1.4 Bibliography

Abud, S., De Souza, P. I. M., Vianna, G. R., Leonardecz, E., Moreira, C. T., Faleiro, F. G., Júnior, J. N., Monteiro, P. M. F. O., Rech, E. L., & Aragão, F. J. L. (2007). Gene flow from transgenic to nontransgenic soybean plants in the Cerrado region of Brazil. *Genetics and Molecular Research*, *6*(2), 445-52.

Ackerman, S., & Knox, J. A. (2006). *Meteorology: understanding the atmosphere* (2nd ed.). Belmont, CA: Cengage Learning. 528 pp.

Agresti, A., & Kateri, M. (2011). Categorical Data Analysis. In M. Lovric (Eds.), *International Encyclopedia of Statistical Science* (pp. 206-208). Berlin, Germany: Springer. 1671 pp.

Ahrens, C., Chung, J., Meyer, T., & Auer, C. (2011). Bentgrass distribution surveys and habitat suitability maps support ecological risk assessment in cultural landscapes. *Weed Science*, *59*(2), 145-154.

Ahrens, C. D. (2012). *Meteorology today: an introduction to weather, climate, and the environment*. Belmont, CA: Cengage Learning. 640 pp.

Armstrong, T. T., Fitzjohn, R. G., Newstrom, L. E., Wilton, A. D., & Lee, W. G. (2005). Transgene escape: what potential for crop–wild hybridization?. *Molecular Ecology*, *14*(7), 2111-2132.

Atkinson, P. M., German, S. E., Sear, D. A., & Clark, M. J. (2003). Exploring the relations between riverbank erosion and geomorphological controls using geographically weighted logistic regression. *Geographical Analysis*, *35*(1), 58-82.

Bohren, C. F., & Albrecht, B. A. (1998). *Atmospheric Thermodynamics*. Oxford, UK: Oxford University Press. 402 pp.

Chandler, S., & Dunwell, J. M. (2008). Gene flow, risk assessment and the environmental release of transgenic plants. *Critical Reviews in Plant Sciences*, *27*(1), 25-49.

Cochran, W. G. (1950). The comparison of percentages in matched samples. *Biometrika*, *37*(3/4), 256-266.

Connecticut Association of Land Surveyors. (2008). *Guidelines and Specifications for Global Navigation Satellite System Land Surveys in Connecticut*. Wethersfield, CT: Connecticut Association of Land Surveyors. Retrieved from http://ctsurveyors.org/wp-content/uploads/2010/12/GNSS_20080626_online.pdf

Crandall, C. L. (1906). *Text-book on Geodesy and Least Squares*. New York, NY: John Wiley & Sons. 329 pp.

Damgaard, C., & Kjellsson, G. (2005). Gene flow of oilseed rape (Brassica napus) according to isolation distance and buffer zone. *Agriculture, Ecosystems & Environment*, *108*(4), 291-301.

Eckl, M. C., Snay, R. A., Soler, T., Cline, M. W., & Mader, G. L. (2001). Accuracy of GPS-derived relative positions as a function of interstation distance and observing-session duration. *Journal of Geodesy*, *75*(12), 633-640.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861-874.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2003). *Geographically weighted regression: the analysis of spatially varying relationships*. Hoboken, NJ: John Wiley & Sons. 282 pp.

Fraden, J. (2015). *Handbook of Modern Sensors: Physics, Designs, and Applications* (5th ed.). Cham, Switzerland: Springer. 758 pp.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, *32*(200), 675-701.

Ghilani, C. D. (2010). *Adjustment Computations: Spatial Data Analysis* (5th ed.). Hoboken, NJ: John Wiley & Sons. 672 pp.

Harold, E., I. (1952, Jun 17). *U.S. Patent No. 2,600,396*. Washington, DC: U.S. Patent and Trademark Office.

Harvey, B. R. (2009). *Practical least squares and statistics for surveyors: a review of principles, problems and procedures* (3rd ed. Monograph 13). Sydney, Australia: School of Surveying and Spatial Information Systems, University of New South Wales. 332 pp.

Henning, W. (2011). *User Guidelines for Single Base Real Time GNSS Positioning* (ver. 2.1). NOAA: National Geodetic Survey. Retrieved from http://www.ngs.noaa.gov/PUBS_LIB/NGSRealTimeUserGuidelines.v2.1.pdf

Henning, W., Martin, D., Schrock, G., Thompson, G., & Snay, R. (2013). *National Geodetic Survey Guidelines for Real Time GNSS Networks* (ver. 2.2). NOAA: National Geodetic Survey. Retrieved from http://www.ngs.noaa.gov/PUBS_LIB/NGSGuidelinesForRealTimeGNSSNetworks.pdf

Hofmann-Wellenhof, B., Lichtenegger, H., & Wasle, E. (2008). *GNSS – Global Navigation Satellite Systems: GPS, GLONASS, Galileo and more*. Wien, Austria: Springer. 518 pp.

Kilby, V., McManus, E., & Cumberland, D. R., Jr. (1993, July). Using a psychrometer to measure relative humidity, *Conserve O Gram*. 3/1. Washington, DC: National Park Service. Retrieved from https://www.nps.gov/museum/publications/conserveogram/03-01.pdf

Kulwicki, B. M. (1991). Humidity sensors. *Journal of the American Ceramic Society*, *74*(4), 697-708.

Landau, H., Vollath, U., & Chen, X. (2009). Virtual reference station systems. *Journal of Global Positioning Systems*, *1*(2), 137-143.

Lee, C. Y., & Lee, G. B. (2005). Humidity sensors: a review. *Sensor Letters*, *3*(1-4), 1-15.

Metz, C. E. (2000). Fundamental ROC analysis. *Handbook of Medical Imaging*, *1*, 751-769.

Meyer, T. H., Bean, J. E., Ferguson, C. R., & Naismith, J. M. (2002). The Effect of Broadleaf Canopies on Survey-grade Horizontal GPS/GLONASS Measurements. *Surveying and Land Information Science. 62*(4), 215-224.

Meyer, T. H., Arifuzzaman, K., & Massalski, D. (2010). Assessing the Accuracy of GEOID03 GEOID09 in Connecticut. *Surveying and Land Information Science*, *70*(2), 89-101.

Mikhail, E. M., & Gracie, G. (1981). *Analysis and Adjustment of Survey Measurements*. New York, NY: Van Nostrand Reinhold. 340 pp.

Nielsen-Kellerman. (2010). *Kestrel 3000 Pocket Weather Meter Instruction Manual*. Boothwyn, PA: Nielsen-Kellerman. Retrieved from http://www.nkhome.com/pdfs/K3000_Instructions_7.23.10_WEB.pdf

Parent, J. R., & Volin, J. C. (2014). Assessing the potential for leaf-off LiDAR data to model canopy closure in temperate deciduous forests. *ISPRS Journal of Photogrammetry and Remote Sensing*, *95*, 134-145.

Rittersma, Z. M. (2002). Recent achievements in miniaturized humidity sensors—a review of transduction techniques. *Sensors and Actuators A: Physical*, *96*(2), 196-210.

Rodrigues, M., de la Riva, J., & Fotheringham, S. (2014). Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. *Applied Geography*, *48*, 52-63.

Roveti, D. K. (2001, July 1). Choosing a Humidity Sensor: A Review of Three Technologies. *Sensors Online*. Retrieved from http://www.sensorsmag.com/sensors/humidity-moisture/choosing-a-humidity-sensor-a-review-three-technologies-840

Saefuddin, A., Setiabudi, N. A., & Fitrianto, A. (2012). On comparison between logistic regression and geographically weighted logistic regression: with application to Indonesian poverty data. *World Applied Sciences Journal*, *19*(2), 205-210.

Stem, J. E. (1989). *State Plane Coordinate System of 1983 (Vol. 5)*. US Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Charting and Geodetic Services. Retrieved from https://www.ngs.noaa.gov/PUBS_LIB/Manual NOSNGS5.pdf

Stigler, S. M. (1981). Gauss and the invention of least squares. *The Annals of Statistics*, *9*(3), 465-474.

Thomson, G. (1986). *The Museum Environment* (2nd ed.). London, UK: Butterworth-Heinemann. 293 pp.

Tsonis, A. A. (2007). *An Introduction to Atmospheric Thermodynamics* (2nd ed.). Cambridge, UK: Cambridge University Press. 198 pp.

U.S. Army Corps of Engineers. (2011, February 28). *Engineering and Design: NAVSTAR Global Positioning System Surveying* (EM 1110-1-1003), Washington, DC: U.S. Army Corps of Engineers, Department of Army. Retrieved from http://www.publications.usace.army.mil/Portals/76/Publications/EngineerManuals/EM_1110-1-1003.pdf?ver=2013-09-04-075011-243

Van Erkel, A. R., & Peter, M. (1998). Receiver Operating Characteristic (ROC) Analysis: Basic Principles and Applications in Radiology. *European Journal of radiology*, *27*(2), 88-94.

Van Sickle, J. (2008). *GPS for Land Surveyors* (3rd ed.). Boca Raton, FL: CRC Press. 360 pp.

Wallace, J. M., & Hobbs, P. V. (2006). *Atmospheric Science, an Introductory Survey* (2nd ed.), Amsterdam, Netherlands: Elsevier Academic Press. 504 pp.

Wang, K., Zhang, C., & Li, W. (2012). Comparison of geographically weighted regression and regression kriging for estimating the spatial distribution of soil organic matter. *GIScience & Remote Sensing*, *49*(6), 915-932.

Wang, K., Zhang, C., & Li, W. (2013). Predictive mapping of soil total nitrogen at a regional scale: a comparison between geographically weighted regression and cokriging. *Applied Geography*, *42*, 73-85.

Wang, K., Zhang, C. R., Li, W. D., Lin, J., & Zhang, D. X. (2014). Mapping soil organic matter with limited sample data using geographically weighted regression. *Journal of Spatial Science*, *59*(1), 91-106.

Wiederhold, P. R. (1997). *Water vapor measurement: methods and instrumentation* (Vol. 1). New York, NY: Marcel Dekker. 384 pp.

Wolberg, J. (2006). *Data Analysis Using the Method of Least Squares: Extracting the Most Information from Experiments*. Berlin, Germany: Springer. 250 pp.

Yamazoe, N., & Shimizu, Y. (1986). Humidity Sensors: Principles and Applications. *Sensors and Actuators*, *10*(3-4), 379-398.

Zar, J. H. (2010). *Biostatistical Analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall. 944 pp.

Zilkoski, D. B., D'Onofrio, J. D., & Frakes, S. J. (1997). *NOAA Technical Memorandum NOS NGS-58: Guidelines for Establishing GPS-derived Ellipsoid Heights (standards: 2 cm and 5 cm),* Version 4.3. US Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, National Geodetic Survey. Retrieved from https://www.ngs.noaa.gov/PUBS_LIB/NGS-58.pdf

Zilkoski, D. B., Carlson, E. E., & Smith, C. L. (2008). *NOAA Technical Memorandum NOS NGS-59 Guidelines for Establishing GPS-Derived Orthometric Heights* (Vol. 20910). US Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, National Geodetic Survey. Retrieved from https://www.ngs.noaa.gov/PUBS_LIB/NGS592008069FINAL2.pdf

Chapter 2

Comparing Kestrel 3000 handheld weather meter measurements of temperature and relative humidity against those of the WEKSLER Sling Psychrometer (Model# 315-1)

Abstract

Relative humidity affects surveying instruments whose observations depend on optical or electromagnetic waves. A handheld digital weather meter is among the simplest and most convenient instruments to measure temperature and relative humidity compared to analog thermometers and sling psychrometers. Temperatures and relative humidities were collected using a Kestrel 3000, a digital handheld weather station and a sling psychrometer, and compared with a linear regression. The slope of the regression revealed that temperatures from the sling psychrometer increase 1.017 °C ($p \ll 0.05$) when the Kestrel indicates a temperature increase of 1 °C. Ideally, the intercept should not be statistically different from zero but the intercept was 0.565 °C ($p \ll 0.05$). However, this difference is close to the temperature accuracy of Kestrel (±0.5 °C) so it doesn't seem alarming. Relative humidity from the sling psychrometer increases 1.003% when the Kestrel indicates a relative humidity increase of 1%. The p-value of the intercept for relative-humidity model is not significant ($p = 0.92$), which indicates there is no measurement relative humidity bias for the Kestrel compared to the sling psychrometer. Our results show that a digital handheld weather station such as Kestrel 3000 will be acceptable for surveying procedures.

2.1 Introduction

Refraction occurs when light or microwaves pass through the interface between media of different refractive indexes. The refractive index, a.k.a. refraction coefficient, of a certain medium (n) is defined as the ratio of the speed of light in vacuum (c) and the phase velocity of light in the medium (v).

$$n = \frac{c}{v} \tag{2.1}$$

Equation (2.1) means that light propagates *n* times faster in a vacuum than in the medium. Refraction affects surveying observations by causing light to not follow a straight-line path, which, for example, affects the zenith angle of a theodolite observation. Refractive index is also necessary for leveling and reductions for precise electromagnetic or laser distance measurements (Baselga, García-Asenjo & Garrigues, 2013).

The refractive index of air (n_{air}) depends on temperature, pressure, relative humidity, and wavelength of the light/microwave source. There have been different versions of the refractive index of air published by different scientists (Birch & Downs 1993; Birch & Downs, 1994; Ciddor, 1996; Edlén, 1966) and scholars in surveying introduced the concept and applications of the refractive index of air for surveyors (Burnside, 1991; Duggal, 2013; Laurila & Harris 1983; Moffit & Bossler, 1998; Rüeger, 1996). For example, Stone and Zimmerman (2001) of the Physical Measurement Laboratory, the National Institute of Standards and Technology (NIST) provide a simple formula for calculating n_{air} for red He-Ne laser with 633-nm wavelength which are universally used for displacement interferometry:

$$n_{air} = 1 + \frac{7.86\times10^{-5}\times P_{air}}{273 + T_{air}} - 1.5\times10^{-11} \times RH \times (T_{air}^2 + 160) \quad (2.2)$$

where T_{air} is dry air temperature (°C), RH is relative humidity in percent (from zero to 100), and P_{air} is atmospheric pressure (hPa). Formulas that correct for refraction and light speed in the atmosphere depend on these meteorological quantities. For precise geodetic surveying, or for scientific/academic purposes, surveyors need to measure air temperature and humidity information during observation sessions to derive refractive index using for corrections. To calculate refractive index, Torge (2001) proposed using air temperature, atmospheric pressure, and relative humidity. Baselga *et al*. (2013) proposed a simpler model to calculate refractive index only with temperature and atmospheric pressure. Temperature and relative humidity are also required for modelling zenith wet delay (ZWD) (Davis, Herring, Shapiro, Rogers, & Elgered, 1985) to correct tropospheric delay of GNSS-radio signals caused by the wet components such as water vapor and condensed water (Leick, 2004; de Oliveira *et al*., 2017). For ZWD model, a partial pressure of water is required and this pressure can be derived from the relative humidity and air temperature (Leick, 2004).

Analog hygrometers, such as sling psychrometers, are generally accepted for measuring relative humidity by many professionals such as firefighters (Lemon & Mangan, 2000) and scientists because its structure is quite simple, and mercury thermometers have relatively slow response times. Some researchers believe that the accuracy of sling psychrometers is the most reliable (John C. Volin, personal communication, March 29, 2013). However, using psychrometers during surveying fieldwork is cumbersome, and the delicate glass thermometers are subject to damage, so a robust way to collect the meteorological conditions is preferred. Digital handheld weather stations are available that collect temperature and relative-humidity

measurements instantly; however, to our knowledge, whether they are accurate enough for surveying practice is an open question.

2.1.1 Background

A hygrometer is a device for measuring humidity, and there are various operational principles for hygrometers. Since the first introduction of hygrometer by Leonardo da Vinci in late 15th century, scientists and inventors devised various types of hygrometers such as the metal-paper coil hygrometer, the hair-tension hygrometer, the psychrometer, and the electronic psychrometer (Ackerman & Knox, 2006; Ahrens, 2012; Harold, 1952; Wiederhold, 1997). A sling psychrometer is one of simplest form among analogue hygrometers, and the advantage of this psychrometer is its small size so it is convenient to carry. A sling psychrometer consists of two glass thermometers linked to a handle via pivot or chain. One of the thermometers has its bulb wrapped in a wick, which is soaked in distilled water between measurements to keep it wet. A sling psychrometer measures wet- and dry-bulb temperatures after whirling at least one minute and no more than three minutes (Kilby, McManus, & Cumberland, 1993; Thomson, 1986). Water on the wick evaporates, which chills the wet-bulb while in operation (Bohren & Albrecht, 1998; Tsonis, 2007). Relative humidity is obtained using the wet- and dry-bulb temperatures as inputs to psychrometric charts or equations (Bohren & Albrecht, 1998; Tsonis, 2007; Wallace & Hobbs, 2006). Special caution is needed to keep the wick from contamination from dirt, chemical materials, or oils passed by bare hands, because those contaminations hinder the evaporation of the water so that the measured humidity values might be deteriorated. For our experiment, a WEKSLER® Sling Psychrometer (Model #315-1) was used (Figure 2.1). The thermometers on the sling psychrometer is in the Fahrenheit scale, and their advantages over the

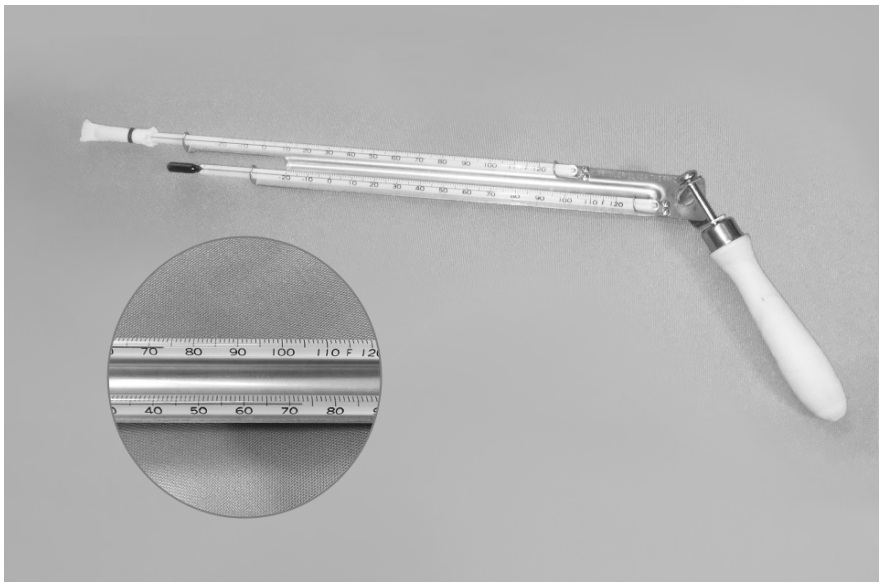


Figure 2.1: WEKSLER® Sling Psychrometer (Model #315-1)

Figure 2.2: Kestrel 3000 Pocket Weather Meter

Celsius scale is their finer graduations so more accurate measurement is possible.

Modern electronic hygrometers often contain digital sensors such as thermal, gravimetric, optical, resistive, capacitive hygrometers (Fraden, 2015; Kulwicki, 1991; Lee & Lee, 2005; Rittersma, 2002; Roveti, 2001). A thermal humidity sensor measures absolute humidity from the difference of thermal conductivity between humid and dry air detected by two thermistors (one encapsulated in dry air and the other exposed to humidity) (Roveti, 2001). A gravimetric humidity sensor compares the masses of two hygroscopic materials (wet and dry) simultaneously, then determines the humidity from the difference of the water vapor mass using a piezoelectric component that infers water mass difference from a change in electric charge, and calculates relative humidity from this electric charge (Lee & Lee, 2005; Rittersma, 2002). Gravimetric humidity sensors are highly sensitive and accurate so the U.S. national standard is determined by this method (Lee & Lee, 2005; National Institute of Standards and Technology, Thermodynamic Metrology Group, 2013). Optical humidity sensors use some kinds of materials that change their physical or mechanical properties when exposed to water vapor. Examples include optical fiber, such as fiber Bragg gratings (FBG), color-changing materials, volume-changing materials, fluorescent optode membranes, photoacoustic cells, and vapoluminescent salt (Bedoya, Orellana, & Moreno-Bondi, 2001; Bozóki, Szakáll, Mohácsi, Szabó, & Bor, 2003; Choi & Shuang, 2000; Correia *et al*., 2012; Drew, Mann, Marquardt, & Mann, 2004; Gupta & Ratnanjali, 2001; Shukla *et al*., 2004). When a certain physical property is changed by water vapor, these materials can change various optical variables of input light passing through the material. These optical changes can be converted to relative humidity. A chilled-mirror dew point hygrometer consists of a mirror linked to a temperature sensor and an optoelectronic device to monitor water vapor condensation on the mirror surface, plus an air temperature sensor, too. The instrument uses

electricity to cool down a metallic mirror, then measures dew point and surrounding temperatures at the moment water vapor is condensed. With these two temperatures, relative humidity can be derived (Wanielista, Kersten, & Eaglin, 1997). Resistive humidity sensors have materials such as salts, organic conductive polymers, or porous ceramics that are exposed to the air. The electrical resistances of these materials are affected by the amount of water vapor in the air, and these resistance changes are converted to the relative humidity (Lee & Lee, 2005; Rittersma, 2002; Yamazoe & Shimizu, 1986). A hygrometer with capacitive sensor contains a dielectric material, a plastic or a polymer, attached to a pair of electrodes (Roveti, 2001). The dielectric material absorbs water vapor in the air proportional to the relative humidity. The electric capacity of this system is also proportional to the relative humidity of the air (Islam, Khan, Akhtar, & Rahman, 2014; Rittersma, 2002). These capacitive sensors are widely used for handheld digital hygrometers because of their small size. The Kestrel 3000 Pocket Weather Meter (hereafter Kestrel 3000) measures temperature and relative humidity, so this unit was selected to the handheld weather station for this study (Figure 2.2). A Kestrel 3000 uses a polymer capacitive humidity sensor for its relative humidity measurement, and measures the atmospheric temperatures through the resistance change of its thermistor. The manufacturer's manual explains that, "The exposed thermistor of Kestrel 3000 responds to changes in temperature when air flows past it" (Nielsen-Kellerman, 2010). Other than ambient temperature and relative humidity, the Kestrel 3000 also measures wind speed, wind chill, maximum wind gust, average wind speed, heat stress and dew point. Kestrel handheld weather stations are lightweight, rugged, small, easy to use, inexpensive, and widely available, which makes them ideal for surveyors to determine meteorological conditions in the field.

2.1.2 Literature review

Swan (1935) gave some advice when measuring relative humidity with a pocket-type sling psychrometer, such as measuring in the shade, protecting from direct sun light, and protecting from rain. Waite (1971) appraised the relative-humidity error of both a sling psychrometer and meteorological measuring set (which was called an AN/TMQ-22), which measures the relative humidity with the dew-point-temperature method. In his study, the average relative-humidity error with the sling psychrometer was reported as +9%, whereas the error with AN/TMQ-22 was reported as +2%. Ahti, Sankola, and Heikkinen (1982) predicted relationships between the relative-humidity errors of an aspirated psychrometer and a thin-film humidity sensor (HUMICAP) installed outdoors inside a Stevenson screen (instrument shelter) and the actual relative humidity. The measurement of relative humidity was seasonal, and the data were calculated as 10-day moving averages of psychrometer-HUMICAP differences. The standard deviation of the differences was 2.5~3.0% RH, and the psychrometer-HUMICAP difference decreased when the relative humidity increases with slope of –0.09 and intercept of 7.54. Lemon and Mangan (2000) tested temperature and relative humidity accuracies for eight digital weather meters including a Kestrel 3000 against a single sling psychrometer reading in an uncontrolled environment. They reported the Kestrel's temperature accuracy as +0.1 °F, and the relative humidity accuracy as –4%. White (2011) empirically record the relative humidity readings of a sling psychrometer and a Kestrel 3000 from random field data without any statistical analysis, and found that the Kestrel relative humidity was approximately 5~6% lower than the relative humidity from sling psychrometer. Afterward their Kestrel meter was sent to the manufacturer for the calibration, and they double-checked the humidity by sending the instrument to a scientific instrument testing and calibration lab. The manufacturer and testing lab reported the

relative humidity of Kestrel was within design specifications, and White concluded the relative humidity from their sling psychrometer is relatively high.

2.1.3 Hypotheses and purpose of the study

The concern of this study is to test whether a Kestrel 3000 produces results that are acceptable to support fieldwork by comparison with the sling psychrometer. To verify the credibility of our Kestrel 3000 for humidity and temperature, there needs to be a quantification and comparison of the statistical relationship between our digital weather meter and the sling psychrometer. Linear regression analyses between temperatures of both instruments, and also between relative humidities of both instruments were selected as a quantification method. The following hypotheses are tested.

Hypothesis 1: Between temperatures of Kestrel 3000 and sling psychrometer:

H_0: The slope between Kestrel temperature and sling psychrometer is 1, and the intercept of the same model is 0.

H_A: Not H_0.

Hypothesis 2: Between relative humidities of Kestrel 3000 and sling psychrometer:

H_0: The slope between Kestrel relative humidity and sling-psychrometer relative humidity is 1, and the intercept of the same model is 0.

H_A: Not H_0.

2.2 Material and methods

2.2.1 Variables: wet- and dry-bulb temperatures and relative humidity

The variables for this experiment were temperature and relative humidity. The Kestrel 3000 can collect atmospheric temperatures and relative humidity; however, the sling psychrometer can only collect the dry-bulb temperature (ambient temperature) and the wet-bulb temperature. Thus four variables were collected: temperature ($T_{Kestrel}$) in Fahrenheit and relative humidity ($RH_{Kestrel}$) in percent acquired from the Kestrel 3000, and wet-bulb temperature (T_w) in Fahrenheit and dry-bulb temperature (T_d) in Fahrenheit from the sling psychrometer. T_w and T_d were used to compute the relative humidity (RH_{sling}) from the sling psychrometer measurements; see (2.4) – (2.7).

For completeness, the Fahrenheit to Celsius conversions is calculated for all the temperatures with following equation,

$$T_C = \frac{5}{9}(T_F - 32) \tag{2.3}$$

where T_C is a temperature in Celsius scale and T_F is a temperature in Fahrenheit scale.

Relative humidity is the ratio between the partial pressure of the water vapor present and the partial pressure of potential saturated water vapor at a given temperature and barometric pressure (Parish & Putnam, 1977). The unit of relative humidity is percent (%). The maximum capacity of water vapor amount for a given air temperature is fixed, and this capacity increases/decreases as air temperature increases/decreases. Relative humidity (RH_{Sling}) can be calculated with wet- and dry-bulb temperatures from (2.4), where E_w is the saturation vapor pressure in saturated air, E_d is the saturation vapor pressure in dry air, A is a conversion factor, P

is the mean atmospheric pressure, and ΔT is the difference between the dry- and wet-bulb temperatures (Huang, Zhang, Yang, & Jin, 2013; Parish & Putnam, 1977).

$$RH_{sling} = \frac{E_w - A \times P \times \Delta T}{E_d} \tag{2.4}$$

$$E_w = 6.112 \times e^{\frac{17.502 \times T_w}{T_w + 240.97}} \tag{2.5}$$

$$E_d = 6.112 \times e^{\frac{17.502 \times T_d}{T_d + 240.97}} \tag{2.6}$$

$$A = 0.00066 \times (1 + 0.00115 \times T_w) \tag{2.7}$$

The coefficients in (2.5) and (2.6) are empirical values (Sadeghi, Peters, Cobos, Loescher, & Campbell, 2013) so they vary slightly among researchers (Buck, 1981; Dilley, 1968; Huang *et al*., 2013; Singh, Singh, Singh, & Sawhney, 2002). In our research, the coefficients applied was the one suggested by Buck (1981). For the mean atmospheric pressure P, the value 1013 hPa was applied. The temperature difference ΔT was calculated by subtracting T_w from T_d after converted into Celsius scale (Butler & García-Suárez, 2012; Huang *et al*., 2013).

2.2.2 Devices for temperature and relative humidity measurement

The Kestrel 3000's temperature measuring accuracy is ±0.9°F (±0.5 °C) with a resolution of 0.1°F. The Kestrel 3000's temperature range is from –20.0°F to 158.0°F. Its measuring accuracy of relative humidity is ±3.0 %, the resolution is 0.1 %, and its specification range of relative humidity is 5% to 95% (Nielsen-Kellerman, 2014).

The WEKSLER® sling psychrometer has analogue, glass thermometers filled with kerosene, whose range is –20°F to 120°F. The measuring accuracy of temperature for the thermometers on the psychrometer is ±1.5°F, and the resolution is 1°F (Kevin Marks of Weksler Glass Thermometer Corporation, personal communication, January 18, 2017). The temperature-matched pair of glass thermometers are used for the sling psychrometer to minimize the relative-humidity errors. The wick of the wet-bulb thermometer was wetted only with distilled water ensure accuracy and consistency (Kilby, McManus, & Cumberland, 1993). A digital timer was used to keep the swinging time, which was 90 seconds in every trial.

2.2.3 Data collection environment

The data collection was performed indoors. Data were collected year round to generally produce a wide range of temperature and humidity conditions. However, the indoor environment was manipulated with heaters, air conditioners, humidifiers, and dehumidifiers to produce a wide range of temperature and humidity conditions. The temperature and humidity data were collected simultaneously with Kestrel 3000 and the sling psychrometer.

2.2.4 Data collection procedures

The sling psychrometer was swung simultaneously with the Kestrel 3000 running until the timer alarm went off. At the alarm, the measurer pressed hold button to freeze the record of the Kestrel 3000, then wrote down the wet- and dry-bulb temperatures first on the field book. Then the temperature and the relative humidity from Kestrel 3000 were recorded. After recording all the observations on the set, Kestrel 3000 was released from the frozen screen. This procedure

was repeated in various temperatures and humidity conditions, and the number of total observation was 175.

2.2.5 Data preparations

Because none of the equipment has the ability to record the measurements digitally, the data were recorded on paper and then transcribed into Microsoft Excel. To calculate relative humidity from psychrometric data, (2.4) – (2.7) were implemented using Wolfram *Mathematica*® (Wolfram, 1999). The final dataset has four variables: dry-bulb temperature and relative humidity derived from the sling psychrometer data, and temperature and relative humidity directly measured using Kestrel 3000. These data were used for linear regression analysis with the *R* statistical software system (Kabacoff, 2015; Teetor, 2011).

2.2.6 Statistical analyses

The two analyses were performed. The first was with dry-bulb temperature from sling psychrometer vs. temperature from Kestrel 3000. The second analysis was to check the relationship between the relative humidity derived from sling psychrometer data and the relative humidity measured by Kestrel 3000. The one-way linear regression models were performed to check the slope of the regression line is approximately one and the intercept of the regression line is approximately zero.

The diagnostic statistical analyses were followed to check independence of observations, linear relationship of dependent and independent variables, existence of outliers, homoscedasticity and normality of the residuals of the regression lines.

2.3 Results

2.3.1 Comparison of temperatures

2.3.1.1 Temperature data characteristics

Each observation of both dry-bulb temperatures from sling-psychrometer and temperature from Kestrel 3000 were collected simultaneously. Both variables are continuous numerical variables, and each data measurement was independent. Thus linear regression analysis with these variables were conducted to appraise their relationship.

2.3.1.2 Linear regression model for temperatures

One-way linear regression analysis between dry-temperature from sling psychrometer and Kestrel temperature were performed.

Table 2.1 shows the coefficients results, and there is a positive correlation between the two variables with $F(1, 173) = 3.26 \times 10^4$, $p \ll 0.01$, and $R^2 = 0.995$. The regression model is shown in (2.8), where T_{dry} is the dry-bulb temperature (°C) measured with the sling psychrometer and $T_{Kestrel}$ is the temperature measured with the Kestrel 3000 (°C). A unit change in Kestrel temperature produced a 1.017 °C degree change in the dry-bulb temperature.

$$T_{dry} = -0.565 + 1.017 \cdot T_{Kestrel} \tag{2.8}$$

In Table 2.1, the p-value of the constant term was significant ($p \ll 0.01$), so the hypothesis that the constant value is zero is rejected. However, the constant value of 0.565 °C is only slightly off from the machine error range of Kestrel 3000 weather meter, which is ±0.5 °C.

2.3.1.3 Diagnosis of the linear regression model for temperatures

The linearity between two temperatures (T_{dry} and $T_{Kestrel}$) and the existence of severe outliers were checked with a scatter plot with fitted line in Figure 2.3. The data appear to have a linear relationship, and there seems to be no severe outliers. A residual plot between Kestrel temperatures and dry-bulb temperatures is presented in Figure 2.4. The lack of a trend in the point cloud suggests that the residuals are uncorrelated, which confirms that the Kestrel temperatures follow a linear relationship with the sling psychrometer temperatures. The shape of the point cloud suggests the relationship is homoscedastic because the shape of the cloud seems to lack any trend. To confirm homoscedasticity, Non-constant Variance (NCV) Test ($\chi^2 = 4.89$, $p = 0.03$) and Studentized Breusch-Pagan test ($BP = 4.07$, $p = 0.04$) were performed. The result of both NCV and Breusch-Pagan test shows that the residuals of the linear regression of two temperatures can be considered as homoscedastic when $\alpha = 0.01$.

Normality of the residuals was diagnosed. A normal Q-Q plot for residuals is provided in Figure 2.5, and the result of Shapiro-Wilk's normality test ($W = 0.99$, $p = 0.14$) shows the residuals of the regression model is normally distributed. Skewness (–0.15, $S.E. = 0.18$) and kurtosis (0.45, $S.E. = 0.37$) confirms the normality of the data.

Table 2.1: Linear Regression for dry-bulb temperature acquired from a sling psychrometer.

Explanatory Variables	Coefficients Estimate	Standard Error	*t*-statistic	*p*-value
Constant	– 0.565	0.135	– 4.197	4.32×10^{-5}
Temperature from Kestrel 3000 Weather Station (°C, $T_{Kestrel}$)	1.017	0.006	180.425	$\ll 0.01$

Dependent Variable: Dry-bulb temperature (°C, T_{dry})
$R^2 = 0.995$, $F(1, 173) = 3.26 \times 10^4$, p-value << 0.01

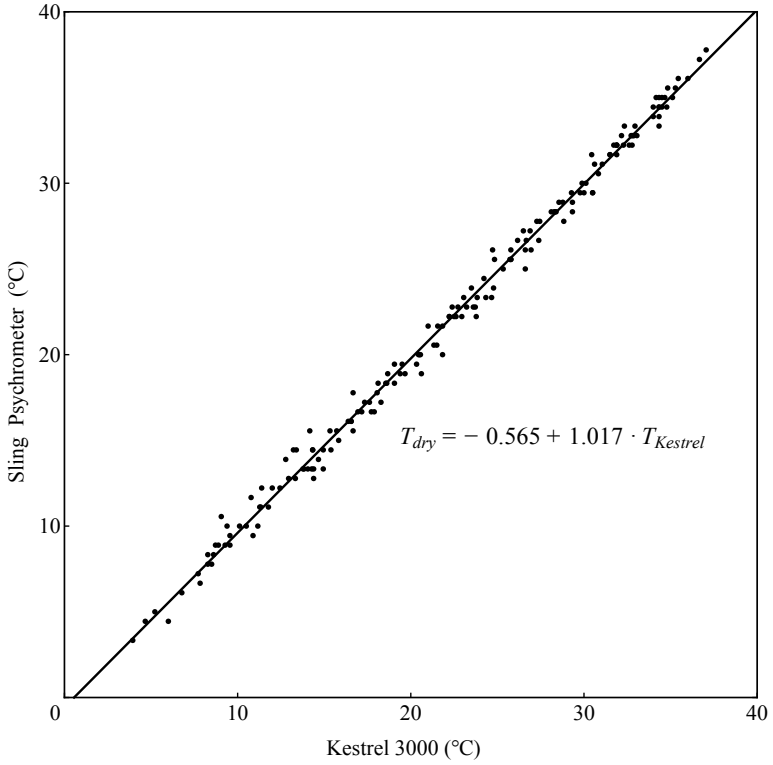


Figure 2.3: Scatterplot with linear regression line fitted. Temperature from a sling psychrometer as a function of temperature measured with digital handheld weather meter at the same time

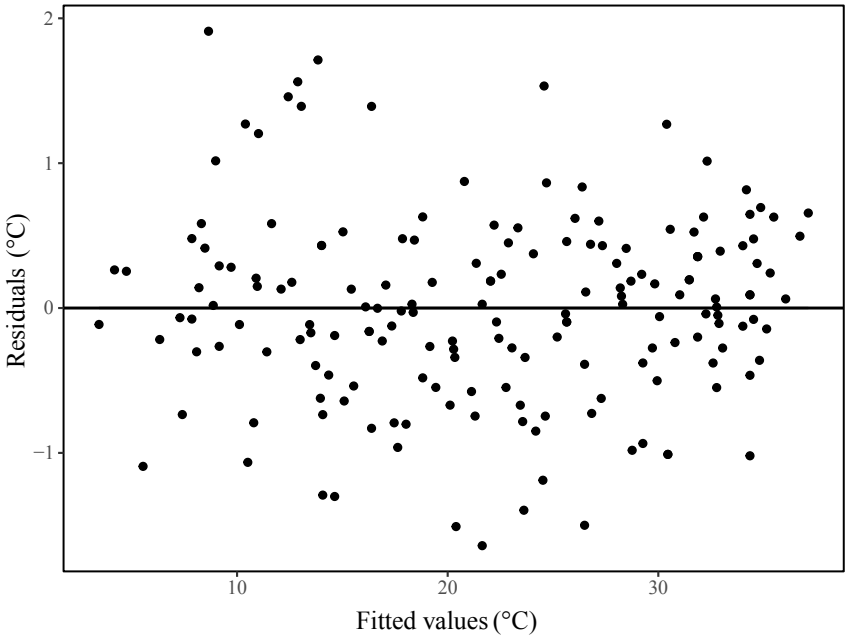


Figure 2.4: Residuals vs. fitted for temperature comparison

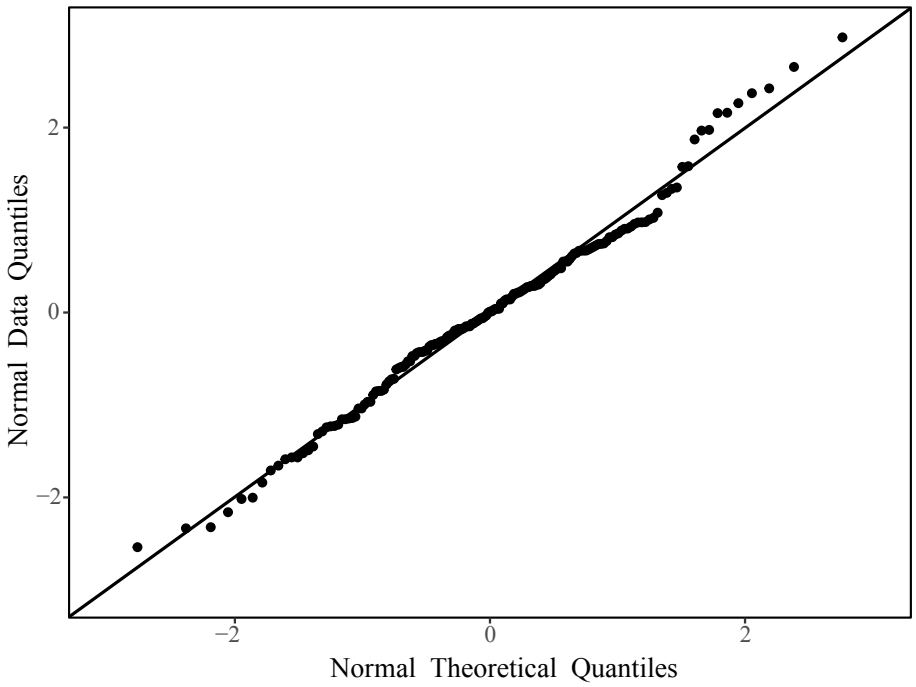


Figure 2.5: Normal Q-Q plot for temperature comparison

2.3.2 Comparison of relative humidities

2.3.2.1 Relative humidity data characteristics

Each observation of both relative humidities from the sling-psychrometer and from the Kestrel 3000 are continuous numerical variables, and each observation of the data measurement was independent. Thus linear regression analysis with these variables were conducted to appraise the relationship between relative humidity from sling psychrometer and relative humidity of Kestrel 3000 weather station.

2.3.2.2 Linear regression model for relative humidities

One-way linear regression analysis between relative humidity from sling psychrometer (RH_{Sling}) and relative humidity from Kestrel 3000 ($RH_{Kestrel}$) were accomplished. Table 2.2: (a) shows the coefficients results, and there is a positive correlation between the two variables with $F(1,173) = 3147$, $p \ll 0.01$, and $R^2 = 0.948$. The regression model is shown in (2.9), where RH_{Sling} is the relative humidity (%) measured with the sling psychrometer and $RH_{Kestrel}$ is the relative humidity measured with the Kestrel 3000 (%). A unit change in relative humidity from Kestrel produced a 1.003% change in the relative humidity from sling psychrometer.

$$RH_{Sling} = -0.094 + 1.003 \cdot RH_{Kestrel} \tag{2.9}$$

In Table 2.2 (a), the , p-value of constant was greater than 0.05, so that the constant value can be assumed to be zero. The linear regression model suppressing intercept, which is expressed in (2.10), was also fitted, and the result is shown in Table 2.2 (b). For the intercept-suppressed regression model, $F(1, 174) = 7.925 \times 10^4$, $p \ll 0.01$, and $R^2 = 0.998$.

$$RH_{Sling} = 1.001 \cdot RH_{Kestrel} \tag{2.10}$$

2.3.2.3 Diagnosis of the linear regression model for relative humidities

The linearity between two relative humidities (RH_{Sling} and $RH_{Kestrel}$) and the existence of severe outliers were checked with a scatter plot with fitted line in Figure 2.6. Also, a residual vs. fitted plot between relative humidity from sling psychrometer and the relative humidity from Kestrel is presented in Figure 2.7. The pattern of point cloud of the residuals is uncorrelated and no trend is apparent. This implies that the relative humidity from the Kestrel follows a linear

relationship with the relative humidity from the sling psychrometer, and the zero mean suggests the relationship is unbiased. To check homoscedasticity of the residuals, Non-constant Variance (NCV) Test ($\chi^2 = 4.46, p = 0.03$) and Studentized Breusch-Pagan test ($BP = 6.52, p > 0.01$) were performed. The result of both NCV and Breusch-Pagan test shows that the residuals of the linear regression of two relative humidities can be considered as homoscedastic when $\alpha = 0.01$. The shape of the point cloud looks mostly homoscedastic because it is hard to find a trend from the shape of the cloud (Figure 2.7).

Table 2.2: Linear Regression for relative humidities acquired from a sling psychrometer and Kestrel 3000

(a) Non-suppressed Regression model with constant value

Explanatory Variables	Coefficients Estimate	Standard Error	*t*-statistic	*p*-value
Constant	–0.094	0.944	–0.100	0.921
Relative humidity from Kestrel 3000 Weather Station (%, $RH_{Kestrel}$)	1.003	0.018	56.100	$\ll 0.01$

Dependent Variable: Relative humidity from sling psychrometer (%, RH_{Sling})
$R^2 = 0.948$, $F(1,173) = 3147$, p-value $\ll 0.01$

(b) Suppressed Regression model without constant value

Explanatory Variables	Coefficients Estimate	Standard Error	*t*-statistic	*p*-value
Relative humidity from Kestrel 3000 Weather Station (%, $RH_{Kestrel}$)	1.001	0.004	281.519	$\ll 0.01$

Dependent Variable: Relative humidity from sling psychrometer (%, RH_{Sling})
$R^2 = 0.998$, $F(1,174) = 7.925 \times 10^4$, p-value $\ll 0.01$

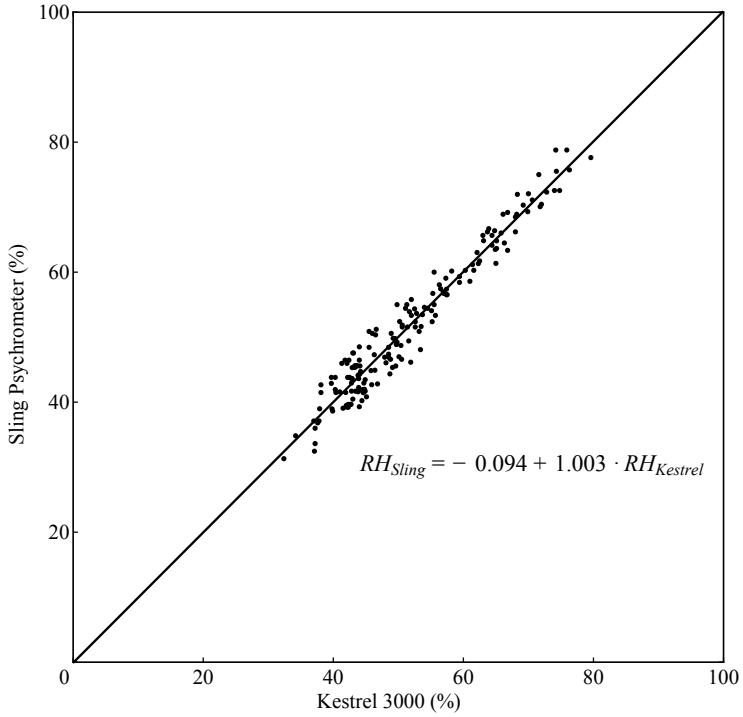


Figure 2.6: Scatterplot with linear regression line fitted. Relative humidity from a sling psychrometer as a function of relative humidity measured with digital handheld weather meter at the same time.

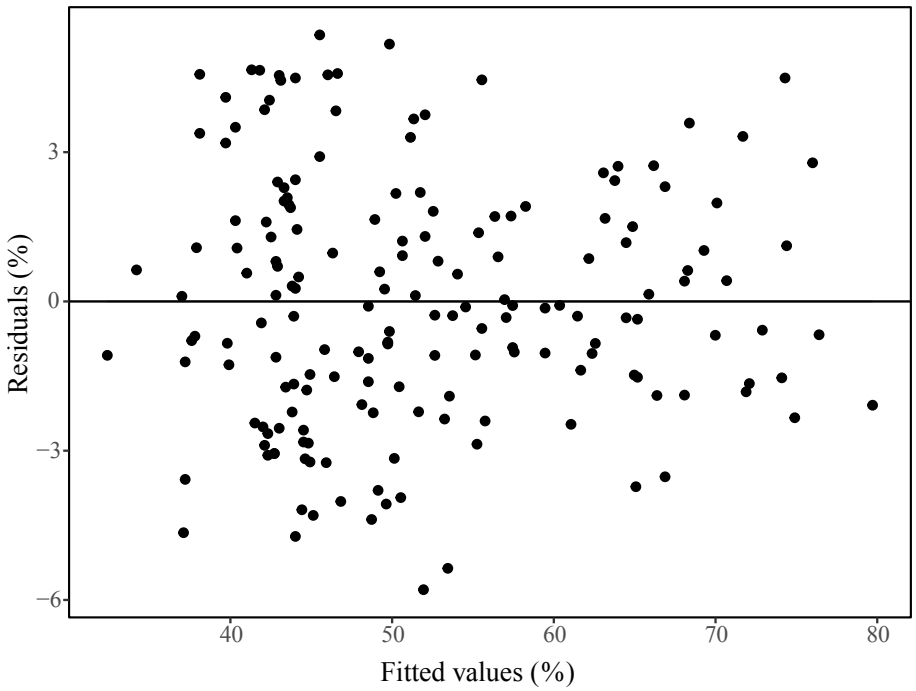


Figure 2.7: Residuals vs. fitted for relative humidity comparison

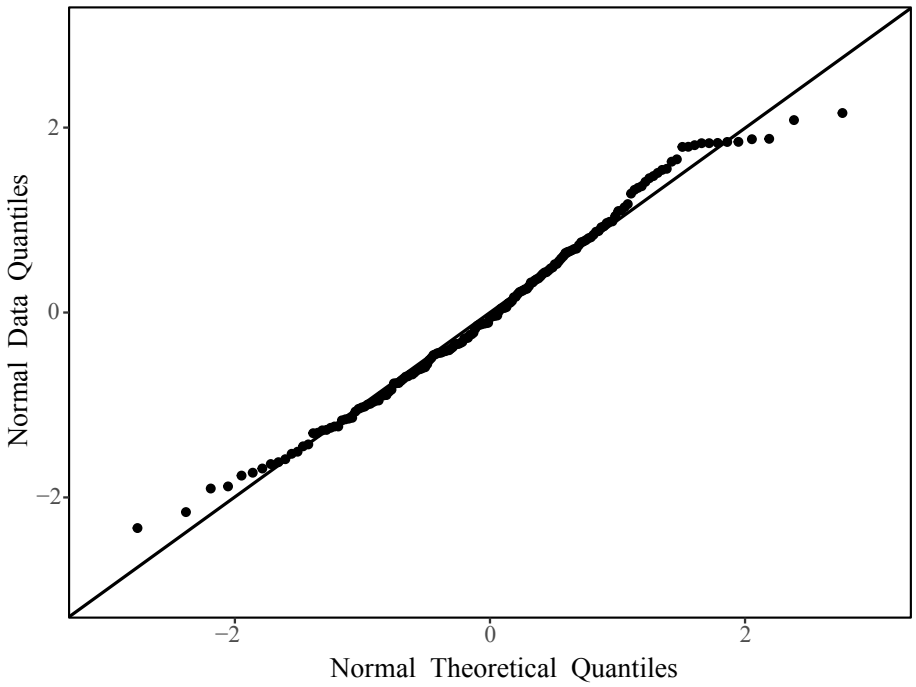


Figure 2.8: Normal Q-Q plot for relative humidity comparison

Normality of the residuals were diagnosed with a normal Q-Q plot for residuals (Figure 2.8) and the Shapiro-Wilk's normality test. The result of Shapiro-Wilk's normality test (W = 0.99, p = 0.84) shows the residual of the regression model is normally distributed. Skewness (– 0.13, $S.E.$ = 0.18) and kurtosis (– 0.62, $S.E.$ = 0.37) confirms the normality of the data.

2.4 Discussion

What was found is that the Kestrel 3000 is suitable for making meteorological measurements for land-surveying purposes, especially in temperate climate zone, e.g. northeastern U.S. Initially, one of the expectation is that the temperature from the Kestrel weather meter and dry-bulb temperature from sling psychrometer would have linear relationship close to a slope of 1 and an intercept of 0. In other words, the temperature increment/decrement

levels measured by the Kestrel weather meter and the sling psychrometer are supposed to be identical, and the temperature errors of the Kestrel weather meter is negligible when compared to the measurement of the sling psychrometer. Also the other expectation is that the humidity from the Kestrel and the humidity from sling psychrometer would have identical linear relationship of slope 1 and intercept 0, meaning relative-humidity increment/decrement levels of the Kestrel weather meter and the sling psychrometer are identical and the relative-humidity error of the Kestrel weather meter compared to the sling psychrometer is negligible.

In the case of the temperature comparison, the slope was 1.017, which means a unit change of temperature from Kestrel corresponds nearly identically to a unit change of temperature from sling psychrometer. The intercept was –0.565 ($p < 0.01$), which means the sling psychrometer's dry-bulb temperature thermometer is about 0.565 °C lower than the temperature from Kestrel weather station. However, this is essentially the accuracy of Kestrel 3000 thermometer, which is ±0.5°C (±0.9°F) (Nielsen-Kellerman; 2014). A sensitivity analysis for the ratio between the partial derivative (Blanchard, Devaney, & Hall, 2012; Caracotsios & Stewart, 1985; Dickinson & Gelinas, 1976; Tenenbaum & Pollard, 1963) of the Air Refractive Index according to temperature and the Air Refractive Index in (2.2) shows this is a negligible error. For example, the ratio was -1.06×10^{-4} when the environmental condition is 15 °C of temperature, 1013.25 hPa of air pressure, and 50% of RH.

For the relative humidity, the linear regression model gives –0.094 ($p = 0.921$) for the intercept value, which was not significantly different from zero. This means the kestrel measurements of relative humidity have a negligible bias compared to those of the sling psychrometer. So, adopting the linear model with a forced-zero intercept is reasonable (Casella, 1983; Hahn, 1977; Turner, 1960). The slope was 1.003 for the non-intercept-suppressed linear

model, and 1.001 for intercept-suppressed linear model. Both cases, the slope is fairly close to 1. This means a unit change in relative humidity from Kestrel and unit change in relative humidity from sling psychrometer can be considered identical. Relative humidity is computed from temperature (Huang, Zhang, Yang, & Jin, 2013; Parish & Putnam, 1977), and the relative humidity model is almost perfect in spite of using temperatures with the aforementioned bias. This supports our assertion that the temperature bias is negligible.

Kestrel 3000 model was the only digital handheld weather station used in this study, and WEKSLER® Model #315-1 was the only sling psychrometer that was used. Thus, it would be inappropriate to conclude that the result of this paper will represent all kinds of digital handheld weather stations and all kinds of sling psychrometers. Indeed, it is unable to show that all Kestrel 3000s work like the one used in this study. Nonetheless, the Kestrel 3000 used in this study performed as advertised according to the demonstration of this study, and it might be reasonable to generalize that the result applies to all units of that model unless a specific unit's performance becomes suspect.

The use of a sling psychrometer doesn't imply that the psychrometer is considered as an absolute standard for temperature and relative humidity. The critical reason to compare the Kestrel 3000 to a sling psychrometer is because sling psychrometers have a well-established history of being used to collect weather variables in the field, and many experts, professionals, and scientists, for example firefighters, museum technicians, and ecologists, have believed the use of a sling psychrometer is more reliable (Blackman & Tansley, 1905; Lemon & Mangan, 2000; Roberts & Smith, 1979; Thomson, 1986). This is because the measurements from the two analog thermometers in a sling psychrometer can be read directly without any other processes (American Warehousemen's Association, 1922), the measurements are consistent, and also

calibration is not required. Also, the difference of the accuracy from sling psychrometers made by other manufacturers were not considered in our study.

The experiment was performed indoors to provide a somewhat controlled environment. Lacking an environmental testing chamber, it was unable to collect data of all combinations of temperatures and humidities. The data range was around 3 °C – 38 °C for temperature, and around 30% – 80% for relative humidity. The temperature range is approximately similar to actual surveying conditions in northeastern U.S. but extremely hot or cold were not covered by our data. Also, the relative-humidity range did not cover all possible field conditions, such as arid climate zones or tropical rainforest climate zones.

A Kestrel weather meter uses a polymer capacitive humidity sensor, which is known to have certain kinds of errors (most of which are common to all sensors): they can be slightly nonlinear, biased, possess hysteresis, and might be miscalibrated (Delapierre, Grange, Chambaz, & Destannes, 1983; Gu, Huang, & Qin, 2004; Kang & Wise, 2000; Kulwicki, 1991; Ralston, 1995; Rotronic Instrument, n.d.). Although the manufacturer should have its own algorithm to correct these errors, the capacitive humidity sensors tend to de-calibrate as times goes by so periodic calibration of the weather meter is necessary (Griesel, Theel, Niemand, & Lanzinger, 2012; Wang *et al*., 2002).

Assuming a periodic calibration of the Kestrel 3000 is assured, the temperature and relative humidity measurements for surveying purpose with the Kestrel weather meter appear to adequate. However, a verification with broader range of data using an environmental testing chamber and another verification in outdoor conditions with all the environmental noises could strengthen the reliability of using Kestrel weather meter in surveying fields.

2.5 Bibliography

Ackerman, S., & Knox, J. A. (2006). *Meteorology: Understanding the Atmosphere* (2nd ed.). Belmont, CA: Cengage Learning. 528 pp.

Ahrens, C. D. (2012). *Meteorology Today: An Introduction to Weather, Climate, and the Environment*. Belmont, CA: Cengage Learning. 640 pp.

Ahti, K., Sankola, T., & Heikkinen, J. (1982). *A comparison of humidity measurements made with a psychrometer, a hair hygrometer and a thin film humidity sensor* (Technical Report. No. 27). Helsinki, Finland: Finnish Meteorological Institute. 27 pp.

American Warehousemen's Association. (1922). Thursday Afternoon Session. In *Proceedings of the Thirty-first Annual Meeting of the American Warehousemen's Association and the Fourteen's Annual Meeting of the Central Warehousemen's Club*, St. Louis, MO, December 6–10, 1921. (pp. 315–327). Retrieved from https://play.google.com/books/reader?id=zfY-AQAAMAAJ&printsec=frontcover&output=reader&hl=en&pg=GBS.PP7

Baselga, S., García-Asenjo, L., & Garrigues, P. (2013). Practical Formulas for the Refraction Coefficient. *Journal of Surveying Engineering*, *140*(2), 06014001.

Bedoya, M., Orellana, G., & Moreno-Bondi, M. C. (2001). Fluorescent optosensor for humidity measurements in air. *Helvetica Chimica Acta*, *84*(9), 2628-2639.

Birch, K. P., & Downs, M. J. (1993). An updated Edlén equation for the refractive index of air. *Metrologia*, *30*(3), 155-162.

Birch, K. P., & Downs, M. J. (1994). Correction to the updated Edlén equation for the refractive index of air. *Metrologia*, *31*(4), 315-316.

Blackman, F. F., & Tansley, A. G. (1905). Ecology in its Physiological and Phyto-topographical Aspects – A Review. *New Phytologist*, *4*(9), 232-253. Retrieved from http://onlinelibrary.wiley.com/store/10.1111/j.1469-8137.1905.tb05910.x/asset/j.14698137.1905.tb05910.x.pdf?v=1&t=j6cjcrgy&s=b684a687e9f8917d567fc78f7 37f3dcb844c32b7

Blanchard, P., Devaney, R. L., & Hall, G. R. (2012). Differential Equations (4th ed.) Boston, MA: Cengage Learning. 864 pp.

Bohren, C. F., & Albrecht, B. A. (1998). *Atmospheric Thermodynamics*. Oxford, UK: Oxford University Press. 402 pp.

Bozóki, Z., Szakáll, M., Mohácsi, Á., Szabó, G., & Bor, Z. (2003). Diode laser based photoacoustic humidity sensors. *Sensors and Actuators B: Chemical*, *91*(1), 219-226.

Buck, A. L. (1981). New equations for computing vapor pressure and enhancement factor. *Journal of applied meteorology*, *20*(12), 1527-1532.

Burnside, C. D. (1991). *Electromagnetic distance measurement* (3rd ed.). Oxford, UK: BSP Professional Books. 288 pp.

Butler, C. J., & García-Suárez, A. M. (2012). Relative humidity at Armagh Observatory, 1838–2008. *International Journal of Climatology*, *32*(5), 657-668.

Caracotsios, M., & Stewart, W. E. (1985). Sensitivity analysis of initial value problems with mixed ODEs and algebraic equations. *Computers & Chemical Engineering*, *9*(4), 359-365.

Casella, G. (1983). Leverage and regression through the origin. *The American Statistician*, *37*(2), 147–152.

Choi, M. M., & Shuang, S. (2000). Fluorescent optode membrane based on organogel for humidity sensing. *Analyst*, *125*(2), 301-305.

Ciddor, P. E. (1996). Refractive index of air: new equations for the visible and near infrared. *Applied Optics*, *35*(9), 1566-1573.

Correia, S. F., Antunes, P., Pecoraro, E., Lima, P. P., Varum, H., Carlos, L. D., Ferreira, R. A., & André, P. S. (2012). Optical fiber relative humidity sensor based on a FBG with a di-ureasil coating. *Sensors*, *12*(7), 8847-8860.

Davis, J. L., Herring, T. A., Shapiro, I. I., Rogers, A. E. E., & Elgered, G. (1985). Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimates of baseline length. *Radio Science*, *20*(6), 1593-1607.

de Oliveira, P. S., Morel, L., Fund, F., Legros, R., Monico, J. F. G., Durand, S., & Durand, F. (2016). Modeling tropospheric wet delays with dense and sparse network configurations for PPP-RTK. *GPS Solutions*, *21*(1), 237–250.

Delapierre, G., Grange, H., Chambaz, B., & Destannes, L. (1983). Polymer-based capacitive humidity sensor: characteristics and experimental results. *Sensors and Actuators*, *4*, 97-104.

Dickinson, R. P., & Gelinas, R. J. (1976). Sensitivity analysis of ordinary differential equation systems–a direct method. *Journal of computational physics*, *21*(2), 123-143.

Dilley, A. C. (1968). On the computer calculation of vapor pressure and specific humidity gradients from psychrometric data. *Journal of Applied meteorology*, *7*(4), 717-719.

Drew, S. M., Mann, J. E., Marquardt, B. J., & Mann, K. R. (2004). A humidity sensor based on vapoluminescent platinum (II) double salt materials. *Sensors and Actuators B: Chemical*, *97*(2), 307-312.

Duggal, S. K. (2013). *Surveying* (Vol. 2. 4th ed.). New Delhi, India: Tata McGraw-Hill Education. 520 pp.

Edlén, B. (1966). The Refractive Index of Air. *Metrologia*, *2*(2), 71.

Fraden, J. (2015). *Handbook of Modern Sensors: Physics, Designs, and Applications* (5th ed.). Cham, Switzerland: Springer. 758 pp.

Griesel, S., Theel, M., Niemand, H., & Lanzinger, E. (2012). Acceptance Test Procedure for Capacitive Humidity Sensors in Saturated Conditions. In *WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (TECO-2012)*, Brussels, Belgium, October 16–18, 2012. Retrieved from http://www.wmo.int/pages/prog/www/IMOP/publications/IOM-109_TECO-2012/Session1/P1_09_Griesel_Acceptance_test_procedure_humidity_sensors.pdf

Gu, L., Huang, Q. A., & Qin, M. (2004). A novel capacitive-type humidity sensor using CMOS fabrication technology. *Sensors and Actuators B: Chemical*, *99*(2), 491-498.

Gupta, B. D., & Ratnanjali. (2001) A novel probe for a fiber optic humidity sensor. *Sensors and Actuators B: Chemical*, *80*(2), 132-135.

Hahn, G. J. (1977). Fitting regression models with no intercept term. *Journal of Quality Technology*, *9*(2), 56–61.

Harold, E., I. (1952, Jun 17). *U.S. Patent No. 2,600,396*. Washington, DC: U.S. Patent and Trademark Office.

Huang, Y., Zhang, K., Yang, S., & Jin, Y. (2013). A method to measure humidity based on dry-bulb and wet-bulb temperatures. *Research Journal of Applied Sciences, Engineering and Technology*, *6*(16), 2984-2987.

Islam, T., Khan, A. U., Akhtar, J., & Rahman, M. Z. U. (2014). A digital hygrometer for trace moisture measurement. *IEEE Transactions on Industrial Electronics*, *61*(10), 5599-5605.

Kabacoff, R. (2015). *R in Action: Data Analysis and Graphics with R* (2nd ed.). Shelter Island, NY: Manning Publications. 608 pp.

Kang, U., & Wise, K. D. (2000). A high-speed capacitive humidity sensor with on-chip thermal reset. *IEEE Transactions on Electron Devices*, *47*(4), 702-710.

Kilby, V., McManus, E., & Cumberland, D. R., Jr. (1993, July). Using a psychrometer to measure relative humidity, *Conserve O Gram*. 3/1. Washington, DC: National Park Service. Retrieved from https://www.nps.gov/museum/publications/conserveogram/03-01.pdf

Kulwicki, B. M. (1991). Humidity Sensors. *Journal of the American Ceramic Society*, *74*(4), 697-708.

Laurila, S. H., & Harris, D. W. (1983). *Electronic Surveying in Practice*. New York, NY: John Wiley & Sons. 388 pp.

Lee, C. Y., & Lee, G. B. (2005). Humidity Sensors: A Review. *Sensor Letters*, *3*(1-4), 1-15.

Leick, A. (2004). *GPS Satellite Surveying*. Hoboken, NJ: John Wiley & Sons. 435 pp.

Lemon, G., & Mangan, D. (2000, June). Evaluating Digital Meters for Fire Weather Observations. *Fire Tech Tips*, 0051-2315-MTDC. USDA Forest Service, Technology and Development Center. 1-8.

Moffitt, F. H., & Bossler, J. D. (1998). *Surveying* (10th ed.). Menlo Park, CA: Addison-Wesley. 738 pp.

National Institute of Standards and Technology, Thermodynamic Metrology Group. (2013). *A New Gravimetric Hygrometer for Verification of NIST Humidity Standards*. Gaithersburg, MD: National Institute of Standards and Technology (NIST), U.S. Department of Commerce. Retrieved from https://www.nist.gov/pml/sensor-science/thermodynamic-metrology/new-gravimetric-hygrometer-verification-nist-humidity

Nielsen-Kellerman. (2010). *Kestrel 3000 Pocket Weather Meter Instruction Manual*. Boothwyn, PA: Nielsen-Kellerman. Retrieved from http://www.nkhome.com/pdfs/K3000_Instructions_7.23.10_WEB.pdf

Nielsen-Kellerman. (2014). *Kestrel 1000-3500DT Sensors & Specs*. Boothwyn, PA: Nielsen-Kellerman. Retrieved from https://kestrelmeters.com/pages/sensors-specs

Parish, O. O., & Putnam, T. W. (1977). *Equations for the determination of humidity from dewpoint and psychrometric data* (NASA-TN-D-8401, H-937). Washington, DC: National Aeronautics and Space Administration (NASA). Retrieved from https://ntrs.nasa.gov/search.jsp?R=19770009916

Ralston, A. R. K. (1995). *Capacitive polymer relative humidity sensors* (Doctoral dissertation, University of Wisconsin-Madison).

Rittersma, Z. M. (2002). Recent achievements in miniaturised humidity sensors – a review of transduction techniques. *Sensors and Actuators A: Physical*, *96*(2), 196-210.

Roberts, S. W., & Smith, R. L. (1979). Aspects of Water Relations in Coast Live Oaks and Valley Oaks Subjected to Root Damage from Land Development Operations. In *Proceedings of the Symposium on the Ecology, Management, and Utilization of California Oaks*, Claremont, CA, June 26–28, 1979. (pp. 171-175).

Rotronic Instrument. (n.d.). *Rotronic Technical Note: The Capacitive Humidity Sensor–How it Works & Attributes of the Uncertainty Budget.* Retrieved from https://www.rotronic.com/en-us/humidity_measurement-feuchtemessung-mesure_de_l_humidite/capacitive-sensors-technical-notes-mr

Roveti, D. K. (2001, July 1). Choosing a Humidity Sensor: A Review of Three Technologies. *Sensors Online*. Retrieved from http://www.sensorsmag.com/sensors/humidity-moisture/choosing-a-humidity-sensor-a-review-three-technologies-840

Rüeger, J. M. (1996). *Electronic Distance Measurement: An Introduction* (4th ed.). Berlin, Germany: Springer-Verlag. 276 pp.

Sadeghi, S. H., Peters, T. R., Cobos, D. R., Loescher, H. W., & Campbell, C. S. (2013). Direct calculation of thermodynamic wet-bulb temperature as a function of pressure and elevation. *Journal of Atmospheric and Oceanic Technology*, *30*(8), 1757-1765.

Shukla, S. K., Parashar, G. K., Mishra, A. P., Misra, P., Yadav, B. C., Shukla, R. K., Bali, L. M., & Dubey, G. C. (2004). Nano-like magnesium oxide films and its significance in optical fiber humidity sensor. *Sensors and Actuators B: Chemical*, *98*(1), 5-11.

Singh, A. K., Singh, H., Singh, S. P., & Sawhney, R. L. (2002). Numerical calculation of psychrometric properties on a calculator. *Building and Environment*, *37*(4), 415-419.

Stone, J. A., & Zimmerman, J. H. (2001, February 16). *Engineering Metrology Toolbox: Refractive Index of Air Calculator.* Gaithersburg, MD: Physical Measurement Laboratory, National Institute of Standards and Technology (NIST), U.S. Department of Commerce. Retrieved from http://emtoolbox.nist.gov/Wavelength/Documentation.asp

Swan, K. D. (1935). *Measuring relative humidity by use of a 'pocket type' sling psychrometer*. Moscow, ID: The Experimental Forest and Savenac Nursery Photo Archive, University of Idaho Library. Retrieved from http://digital.lib.uidaho.edu/cdm/ref/collection/expforsav/id/350

Teetor, P. (2011). *R Cookbook: Proven Recipes for Data Analysis, Statistics, and Graphics*. Sebastopol, CA: O'Reilly. 438 pp.

Tenenbaum, M., & Pollard, H. (1963). Ordinary Differential Equations: An Elementary Textbook for Students of Mathematics, Engineering, and the Sciences. New York, NY: Dover Publications. 832 pp.

Thomson, G. (1986). *The Museum Environment* (2nd ed.). London, UK: Butterworth-Heinemann. 293 pp.

Torge, W. (2001). *Geodesy* (3rd ed.). Berlin, Germany: Walter de Gruyter. 431 pp.

Tsonis, A. A. (2007). *An Introduction to Atmospheric Thermodynamics* (2nd ed.). Cambridge, UK: Cambridge University Press. 198 pp.

Turner, M. E. (1960). Straight line regression through the origin. *Biometrics*, *16*(3), 483–485.

Waite, R. W. (1971). *Field Comparison between Sling Psychrometer and Meteorological Measuring Set AN/TMQ-22* (Technical Report ECOM-5399). White Sands Missile Range, NM: Atmospheric Sciences Laboratory, U.S. Army Electronics Command. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/733293.pdf

Wallace, J. M., & Hobbs, P. V. (2006). *Atmospheric Science, an Introductory Survey* (2nd ed.), Amsterdam, Netherlands: Elsevier Academic Press. 504 pp.

Wang, J., Cole, H. L., Carlson, D. J., Miller, E. R., Beierle, K., Paukkunen, A., & Laine, T. K. (2002). Corrections of humidity measurement errors from the Vaisala RS80 radiosonde—Application to TOGA COARE data. *Journal of Atmospheric and Oceanic Technology*, *19*(7), 981-1002.

Wanielista, M. P., Kersten, R., & Eaglin, R. (1997). *Hydrology: Water Quantity and Quality Control* (2nd ed.), New York, NY: John Wiley & Sons. 592 pp.

White, G. L. (2011). How Accurate Is Your Kestrel?. *Fire Management Today*, *71*(1), 33-35. Retrieved from https://www.r5.fs.fed.us/fire/fmt/fmt_pdfs/FMT71-1.pdf

Wiederhold, P. R. (1997). *Water Vapor Measurement: Methods and Instrumentation* (Vol. 1). New York, NY: Marcel Dekker. 384 pp.

Wolfram, S. (1999). *The Mathematica*. Cambridge, UK: Cambridge University Press. 1470 pp.

Yamazoe, N., & Shimizu, Y. (1986). Humidity Sensors: Principles and Applications. *Sensors and Actuators*, *10*(3-4), 379-398.

Appendix 2A: Raw observations from Kestrel 3000 and WEKSLER® Sling Psychrometer (Model #315-1)

Sling Psychrometer		Kestrel 3000		Sling Psychrometer		Kestrel 3000	
Wet-bulb Temp. (°F)	Dry-bulb Temp. (°F)	Ambient Temp. (°F)	Relative Humidity (%)	Wet-bulb Temp. (°F)	Dry-bulb Temp. (°F)	Ambient Temp. (°F)	Relative Humidity (%)
68	80	80.0	55.1	43	52	52.3	42.1
85	93	93.2	70.0	74	92	90.2	46.8
52	63	63.2	44.0	58	68	68.9	54.5
64	79	80.5	42.9	75	85	86.0	62.1
41	50	50.9	40.3	54	62	62.9	59.4
58	68	71.3	51.1	71	80	79.1	66.3
62	73	74.7	52.8	62	71	69.8	61.6
74	87	87.5	52.5	36	43	44.2	44.0
40	49	51.6	38.1	72	90	91.0	42.8
71	82	83.9	59.4	74	91	89.9	46.4
62	74	74.9	46.0	37	45	45.9	43.8
39	47	46.9	41.8	69	76	75.6	72.0
52	61	61.5	51.7	59	65	64.6	69.2
33	38	39.1	56.3	87	95	93.8	74.8
63	73	74.5	56.5	88	95	94.4	76.3
68	75	74.3	71.8	38	46	47.3	42.9
49	57	58.4	52.0	73	91	90.9	43.9
39	48	49.2	40.4	49	58	57.8	52.6
65	82	81.1	42.0	69	83	82.9	49.8
37	46	46.9	39.8	43	52	52.4	41.3
78	88	87.9	64.4	54	61	61.8	64.9
78	96	94.7	45.8	34	40	40.4	52.6
82	99	98.0	50.4	60	67	67.1	63.9
71	90	89.1	41.5	56	65	65.5	55.3
80	100	98.7	44.8	79	92	93.8	57.4
40	48	48.7	43.0	74	83	83.0	64.4
50	61	61.8	44.2	76	85	86.9	64.8
59	68	69.0	61.0	51	62	62.5	43.5
63	78	76.7	42.8	73	89	89.4	48.8
86	94	93.8	72.8	73	89	86.8	50.5
36	44	46.1	39.7	74	83	84.8	63.0
41	50	48.9	42.5	51	56	58.9	70.6
67	81	80.4	53.4	72	90	89.4	44.5
46	54	54.4	52.0	53	63	64.9	53.2

(continue to the next page)

(continued from the previous page)

Sling Psychrometer		Kestrel 3000		Sling Psychrometer		Kestrel 3000	
Wet-bulb Temp. (°F)	Dry-bulb Temp. (°F)	Ambient Temp. (°F)	Relative Humidity (%)	Wet-bulb Temp. (°F)	Dry-bulb Temp. (°F)	Ambient Temp. (°F)	Relative Humidity (%)
64	73	73.8	62.3	44	55	56.0	39.9
59	72	73.3	43.6	50	60	59.6	48.5
53	64	64.5	46.3	48	60	57.5	42.1
61	74	76.4	50.1	70	84	83.4	49.5
59	72	72.5	49.6	48	56	56.9	55.5
72	88	87.1	48.1	72	91	91.2	42.3
54	64	62.0	53.5	55	67	68.6	44.1
48	60	62.0	42.3	88	94	94.6	75.9
39	49	49.2	37.6	39	47	47.5	42.4
46	55	57.9	49.7	45	57	55.0	37.2
69	85	85.6	48.7	79	96	95.5	48.5
62	71	71.3	60.3	66	73	72.9	69.9
62	74	75.8	48.9	73	85	84.7	57.5
79	90	89.4	62.5	47	56	57.3	49.2
57	68	68.8	46.5	72	78	78.3	71.6
74	86	86.2	56.9	41	50	50.2	42.2
51	62	63.9	43.3	80	90	90.1	63.1
64	79	78.4	42.8	42	52	53.2	43.0
39	48	48.0	38.1	46	54	53.6	55.7
78	95	95.2	48.5	53	63	63.7	45.5
44	56	56.8	34.2	69	82	81.4	50.6
73	92	91.3	44.4	59	66	65.6	68.0
75	86	85.8	58.2	44	54	52.5	45.9
49	58	59.7	50.6	57	66	67.4	57.4
35	40	42.8	55.5	41	50	52.1	39.7
67	84	83.8	45.1	59	66	66.9	63.7
61	69	70.7	66.8	46	58	56.1	37.8
58	72	72.0	44.5	69	81	79.7	54.0
49	59	60.5	43.1	50	60	60.3	45.5
55	66	69.1	49.7	57	65	66.3	61.4
79	97	96.8	43.3	85	94	94.1	66.8
56	65	65.4	57.0	61	77	77.6	37.9
65	77	79.9	50.2	78	95	93.5	47.9
65	80	81.3	44.0	64	79	79.9	43.9
68	83	82.6	51.9	84	93	93.8	66.1
87	94	93.8	74.3	43	55	55.3	37.2
79	97	95.8	49.1	75	84	84.8	65.8
68	85	86.9	44.6	49	58	58.9	51.4

(continue to the next page)

(continued from the previous page)

Sling Psychrometer		Kestrel 3000		Sling Psychrometer		Kestrel 3000	
Wet-bulb Temp. (°F)	Dry-bulb Temp. (°F)	Ambient Temp. (°F)	Relative Humidity (%)	Wet-bulb Temp. (°F)	Dry-bulb Temp. (°F)	Ambient Temp. (°F)	Relative Humidity (%)
69	78	78.4	65.1	87	95	94.1	74.0
63	78	78.4	44.7	39	48	47.7	41.9
65	72	72.7	68.2	70	83	83.1	55.2
74	89	88.7	51.6	78	89	88.7	65.0
45	58	55.8	37.1	60	79	76.5	32.4
48	56	57.7	49.8	63	69	70.4	68.3
41	51	48.3	44.0	72	91	91.0	42.7
68	73	72.3	79.6	72	91	91.1	42.7
64	72	72.0	65.1	51	62	64.2	43.7
64	71	70.8	68.0	54	67	66.3	41.0
58	72	74.8	40.3	72	90	90.7	44.9
35	41	41.4	53.7	88	94	93.2	74.2
63	75	76.6	46.6	79	91	91.5	57.3
48	56	57.9	51.3	43	53	51.4	43.8
72	90	89.4	43.4	46	58	57.8	37.0
60	74	73.5	44.9				

(End of the table)

Chapter 3

The effect of sky obstruction and humidity on the accuracy of RTK positioning in a broadleaf forest in the Northeastern U.S.

Abstract

Broadleaf canopies inject water vapor into the atmosphere during the leaf-on season (evapotranspiration), and water vapor is a component of the error budget for satellite-based positioning, such as with global navigation satellite system (GNSS) receivers. The leaves in broadleaf canopies can also create multipath, attenuate the GNSS signals, and cause cycle slips. Therefore, GNSS positioning accuracy could be negatively correlated with the presence of leaves. Real-time GNSS positioning, such as real-time kinematic (RTK), can use as few as two observations to determine a position, so it is reasonable to study whether tree canopies affect real-time positioning more than longer occupations. Connecticut is well-suited for such a study because it is widely covered with a temperate broadleaf forest, has a largely intact First-Order benchmark control network, and also has a real-time network. More than 500 RTK-determined positions were collected in leaf-on (241) and leaf-off (261) conditions, along with concurrent relative humidity and temperature measurements. Also, percent sky obstruction above each benchmark was derived from a hemispherical photograph at each site. The means of the random distributions of the individual coordinates (eastings, northings, ellipsoid heights) were compared using analysis of covariance (ANCOVA) and rank analysis of covariance (RANCOVA). When

surface absolute humidity was a covariate, contrary to previous findings, there was no significant correlation between canopy status (leaf-on/leaf-off) and positioning accuracy. When sky obstruction was a covariate, also there was no significant correlation between canopy status and positioning accuracy for easting, northing, and ellipsoid-height errors. The humidity measured only at the Earth's surface level might not fully represent the actual wet delay, which results from summing the absolute humidity encountered by the signal between the transmitter to the receiver. Certainly the measured humidity is an underestimate of the total, so the accuracy scatterplots' abscissae are compressed, but it is hard to see how the full answer would change the results. These results might imply that the GNSS receivers used in this study might have firmware to correct for multipath and humidity and they might operate better than conventional wisdom suggests.

3.1 Introduction

3.1.1 Overview

Degradation of global navigation satellite system (GNSS) positioning is unavoidable whenever GNSS surveying is performed close to a tree canopy because the signals coming from the satellites are blocked, affected by multipath, or attenuated by tree leaves and woody materials in the canopy. Broadleaf trees inject water vapor into the atmosphere through evapotranspiration, and water vapor is also a component of the GNSS error budget. This was studied for static occupations in networks by peer researchers (e.g. Hasegawa & Yoshimura, 2003; Meyer, Bean, Ferguson, & Naismith, 2002; Sigrist, Coppin, & Hermy, 1999) who found that accuracy decreased with increasing canopy coverage. Real-time kinematic (RTK) accuracy under broadleaf canopy has not been carefully studied. Although the fundamental positioning

processing – relative positioning with double-differenced phase observables – is the same for RTK and for static, simultaneously processing the observations from a network of receivers allows for a least-squares adjustment of all the unknowns, which greatly enhances the robustness of the solution. With RTK, the survey is between only two receivers, no network is available to mitigate errors. Given the commonplace usage of RTK nowadays, it is worthwhile to examine RTK performance *in situ*, applying a statistical analysis to control environmental factors. In this study we focused on the quantification of differences of positioning errors between leaf-on and leaf-off conditions with humidity and sky obstruction as covariates using an analysis of covariance (ANCOVA) and a rank analysis of covariance (RANCOVA or rank ANCOVA).

3.1.2 Background

3.1.2.1 Positioning methods

Static positioning: The term **static positioning** refers to an occupation strategy in which GNSS receivers are erected atop tripods and left to collect data for relatively long periods. The Texas Department of Transportation (2016) guidelines indicate that their static-positioning occupation time must be minimum of 12 minutes for baselines up to 30 km, and others recommend occupation periods as short as 20 minutes (Connecticut Association of Land Surveyors, 2008; Eckl, Snay, Soler, Cline, & Mader, 2001). The usual occupation time to achieve survey-level accuracy is five hours (Zilkoski, D'Onofrio, & Frakes, 1997), or even longer. Static positioning also implies that at least two receivers are collecting observations simultaneously so that phase-differencing error mitigation can be applied (Leick, Rapoport, & Tatarnikov, 2015; Teunissen & Kleusberg, 1998). Some of these receivers might occupy control points, and such are called **base stations** or **reference stations (**Van Sickle, 2008). Base stations

might occupy passive survey markers, such as National Geodetic Survey (NGS) benchmarks and monuments, and they might also be permanent installations atop buildings or towers. Permanent base stations are called **continuously operating reference stations** (CORS) (Snay & Soler, 2008). After the survey, the raw observations are post-processed with a least squares adjustment to estimate the unknowns, quantify the unknowns' uncertainties, and to detect blunders or outliers. Phase differencing produces high-accuracy baselines between the receivers, which establishes their relative locations very accurately. Static positioning generally produces the most accurate results because it employs phase differencing and it has long observation times, which provides the statistics ample data to work with.

Real-time kinematic: Real-time kinematic (RTK) is another type of occupation strategy that also uses phase-differencing between a (single) base station and another receiver, which for RTK is called a **rover**. The rover might or might not be moving. If it is moving, like in an aircraft providing photogrammetry control, it is performing a **kinematic survey**, which is outside the scope of this investigation. If it is not moving, it is performing a **static RTK survey**. The fundamental difference between static RTK and static surveying is that, with RTK, the rover performs the phase differencing in real time, which produces the positions in the field – not post processed. RTK solutions can be available with as few as two epochs of data (Leick *et al*., 2015; Teunissen & Kleusberg, 1998), and they can be just as accurate as post-processed solutions, at least in principle. The rover receives the base-station's observables via some kind of communication link, like ultra-high frequency (UHF) or very high frequency (VHF) radios, broad spectrum radios, or cellular modems over the Internet. The rover's phase-differencing process computes the baseline from the base station (Henning, 2011; Wegener & Wanninger, 2005). Adding the baseline, which is a geocentric Cartesian vector, to the base station's

geocentric Cartesian coordinates produces the geocentric Cartesian coordinates of the rover. The accuracy of the solution depends on how successfully the phase differencing eliminated the errors in the observables as well as whether the processing successfully chose the proper values for the range integer ambiguities (Leick *et al*., 2015; Strang & Borre, 1997). The error cancelation in phase differencing works on the assumption that the errors at the base station will be similar to those at the rover, which is likely to be true so long as they are relatively near one another. The industry standard for this distance isn't consistent, but 20 km seems to be a reasonable limit. (El-Rabbany, 2006; Henning, 2011; Leick *et al*., 2015).

Some authors have suggested that the phrase "classic RTK" means that the base station is not a CORS and the communication link would typically be a radio (Henning, 2011). The more modern approach is for the rover to communicate over the Internet to a network of multiple permanent reference stations, such as CORS network, via cellphone technology (Snay & Soler, 2008), which would be called **real-time network** (RTN) or network RTK. Rovers using a RTN communicate with digital computers that serve them the base stations' observables rather than communicating with the base stations directly as with (classical) RTK. When working with a single-base RTK method, the RTN tests multiple baseline lengths and finds the nearest CORS to the rover (Henning, 2011). The corrections are transferred to the rover from this nearest CORS through the Internet using the Networked Transport of RTCM via Internet Protocol (NTRIP) (Weber, Dettmering, & Gebhard, 2005) – RTCM is the acronym of Radio Technical Commission for Maritime Services. The benefit of using a RTN is that surveyors don't need to deploy their own base station, which made RTK surveying more simple and a time saver, and there are no security issues from having one's own base station stolen or tampered with (Henning, 2011). RTN enables surveyors to move further than working with traditional RTK,

because the communication distance through cellular modem is much longer than the radio-link coverage of classical RTK (El-Mowafy, Fashir, Al Marzooqi, Al Habbai, & Babiker, 2003; Hu, Khoo, Goh, & Law, 2003).

A RTN can operate with multiple base stations simultaneously. There are several multi-base RTN positioning methods such as virtual reference station (VRS), master-auxiliary, and reverse processing (Henning, Martin, Schrock, Thompson, & Snay, 2013). In this study VRS is the only method considered because it is the only one offered by the Connecticut RTN, the Advanced Continuous Operating Reference Network (ACORN), which is managed by University of Connecticut's Department of Natural Resources and the Environment and the Connecticut Department of Transportation (CTDOT). In the VRS method, a non-physical imaginary reference station is realized by interpolating the observations from the physical base stations, providing the fiction that the surveyor is doing RTK with a base station that is only a few meters away (Hu *et al*., 2003; Landau, Vollath, & Chen, 2002; Petovello, 2011; Talbot, Lu, Allison, & Vollath, 2002). In principle the errors related to the baseline length are removed with the VRS method because the virtual base station is extremely close to the rover (Landau *et al*., 2002; Retscher, 2002). In this study, single-base RTK and VRS are used to collect the positioning data.

3.1.2.2 Error sources of GNSS positioning

Signal attenuation: Signal attenuation is a loss of signal strength due to environmental objects absorbing or dispersing some of the energy. Firth and Brownlie (1998) reported that tree leaves attenuate GNSS signals, which can cause a receiver to lose lock on the signal (a **cycle slip**). Cycle slips force a receiver to abandon its current differencing process and to restart it

from scratch. Excessive cycle slips cause a receiver to fail to find a fixed-integer solution, defeating the positioning altogether. Attenuation can also cause a receiver to incorrectly interpret the digital signals, causing cyclic redundancy check (CRC) faults, parity faults, or even misinterpretations of the binary data (Lachapelle, Henriksen, & Melgara, 1994; Pirti, 2005; Savage, Ndzi, Seville, Vilar, & Austin 2003; Spilker, 1996).

Multipath: When a GNSS signal arrives at the receiver's antenna via more than one path, the multiple signals interact electromagnetically creating constructive and destructive interference patterns, which include (possibly substantial) phase shifts. This phenomenon is called **multipath**, and it can introduce meters of error into positions. In precise GNSS positioning, multipath has to be carefully considered (Braasch, 1996; Langley, 1998a; Hofmann-Wellenhof, Lichtenegger, & Wasle, 2008). Regrettably, there appears to be no way to measure multipath *in situ* so it cannot be accounted for directly in this study; it remains a confounding variable.

Humidity: GNSS-signal delays occur when the signals pass through the troposphere. About 90% of the tropospheric delay is caused by dry gases and is a function of atmospheric pressure and temperature (Conley *et al*., 2005; Pullen & Rife, 2009). Dry delay is relatively stable and can be removed by processing, whereas other 10% of the delay is caused by water vapor (Groves, 2008; Langley, 1998b), which is called the wet delay. The spatial distribution of atmospheric water vapor literally depends on the weather, and it is not possibly to model it today. Relative humidity is the ratio of current water-vapor amount in an area and the potential saturated water-vapor amount in the same area at certain temperature (Ahrens & Henson, 2015; Grenci & Nese, 2001). Absolute humidity (g/m^3) is the actual amount of water vapor in a unit volume of air (regardless of the air temperature) (Ahrens & Henson, 2015). For this study,

absolute humidity was the proper choice because it is the amount of water vapor in the air. Because our team doesn't have a weather station that can measure absolute humidity, we measured relative humidity and ambient temperatures, and converted them to absolute humidity as follows (Hall *et al*., 2015; Mander, 2012):

$$AH = \frac{6.112 \times e^{\left[\frac{17.67+T}{T+243.5}\right]} \times RH \times 2.1674}{273.15 + T} \quad , \tag{3.1}$$

where AH is absolute humidity (g/m^3), RH is the relative humidity (%), T is the ambient temperature in Celsius scale (°C), and e is the base of natural logarithm, a.k.a. Euler's number.

Sky obstruction (%): The degree of signal blockage by visible components of the error budget can be quantified with sky obstruction. If a hemispherical photograph is taken toward the zenith, all but sky pixels should be these signal-blocking features. Sky obstruction is the number of obstructing pixels divided by the total number of the pixels in the circular area captured by a vertical hemispherical photograph. The GNSS receivers collected data only from SVs 13° above the horizon, so the pixels below this threshold were masked out. The calculation of sky obstruction was based on Parent and Volin's (2014) method.

3.1.2.3 Statistical analysis to compare multiple groups controlling unpredictable effects.

ANCOVA: Lacking anechoic chambers and other infrastructure necessary to perform a controlled study, we sought to quantify RTK positioning degradation due to tree canopy and humidity by applying a statistical analysis of *in situ* data. Consequently, this study is not manipulative, it is mensurative (Finn, Maxwell, & Calver, 2002; Hurlbert, 1984; Krebs, 1999). Without control to provide a comparative standard, there is high probability of confusing errors

caused by unknown nuisance variables with errors from our independent variable (canopy occlusion and humidity). ANCOVA is widely used in such situations because it is able to control for the effects of variables that we do not want to examine (Steel, Torrie, & Dickey, 1997; Vogt, 2005). These variables are called **covariates** or **control variables**, and they should be in interval or ratio scale (*ibid*.). In this study, ANCOVA is used to quantify the effect caused by tree-canopy status with eliminating confounders which are the surface absolute humidity effect or sky obstruction.

ANCOVA is known as a combined model of analysis of variance (ANOVA) and linear regression (Milliken & Johnson, 2002). ANCOVA consists of a dependent variable in continuous scale, one or more independent variables in categorical or ordinal scale, and one or more covariates in continuous scale. The procedure of ANCOVA is (1) to verify whether the slopes of the regression lines for each group are identical (assumption of parallel) and (2) to test whether the group means are identical when a covariate is able to explain some portion of errors or variations inside the dependent variable (Field, 2009; Rutherford, 2001). If there are differences among the group means, then a *post hoc* test can follow. In this study, *post hoc* tests aren't needed because there are only two groups (leaf-on/leaf-off). Compared to ANOVA, ANCOVA uses continuous covariate to control unwanted effects. Tabachnick and Fidell (2012) mentioned that the statistical power of ANCOVA is increased by this process, which reduces the within-group error variance in dependent variable.

RANCOVA: ANOVA and ANCOVA assume that the residuals of dependent variables are distributed normally (among other assumptions). This assumption can be explored using statistical tests such as Anderson–Darling test, Lilliefors test, Kolmogorov–Smirnov test and Shapiro–Wilk test. The data here are the distributions of the easting, northing, and height

residuals (difference from control value). ANCOVA can be performed when the normality assumption of the dependent variable is violated, but the statistical power will be weakened. Quade (1967) introduced RANCOVA, a nonparametric counterpart of ANCOVA (Conover & Iman, 1982; Lawson, 1983). In Quade's method, both the dependent variable and all covariates are rank-transformed regardless of the groups. With this rank-transformed data, the linear regression model has to be performed, and residuals come from that linear regression analysis. With these residuals and the grouping variable, ANOVA model can be made. The Quade's RANCOVA was performed to double check and compensate the statistical power of the ANCOVA model of this study.

3.1.3 Literature Review

Static positioning is the method for survey-level positioning. Peer scientists have studied the accuracy of static positioning under canopy environments. Meyer *et al*. (2002) investigated the effect of broadleaf canopies (independent variable) on horizontal GNSS measurements (dependent variable) with 20-minute static occupations. They occupied stations with a broad range of sky obstructions, and each station was occupied twice: GPS-only and GPS/GLONASS. Simple linear regression model was applied to each occupation. The slope of regression line for the GPS-only setup was 2.32 mm per percent sky obstruction (R^2 = 0.82), and slope for the GPS/GLONASS setup was 2.87 mm per percent sky obstruction (R^2 = 0.77), meaning the accuracies of both setups deteriorate as percent sky obstruction increased. Also they found no evidence that the GLONASS constellation either helped or hurt the GNSS horizontal accuracy under broadleaf canopies. Hasegawa and Yoshimura (2003) confirmed that positional accuracies of a static GNSS survey was decreased more under tree canopies than in the open. The mean

horizontal error with 30-minute, dual-frequency, and carrier-phase GNSS observation was 0.029 m for treeless conditions, and the error increased to 0.415 m with 17.6% of sky openness. For single-frequency static observations, the error was 0.226 m for the treeless condition, but the error increased up to 0.894 m at 25.7% sky openness.

There was some research about the effect of forest canopy on RTK positioning accuracy. Morales and Tsubouchi (2007) performed a statistical comparison of the performance between RTK, differential global positioning system (DGPS), and the variation of DGPS called StarFire for kinematic mode under tree-shading environments. They reported that (1) DGPS is less accurate than other two, but the most robust with the existence of obstacles; (2) StarFire DGPS has close accuracy to RTK in open sky, but it does not provide precise solution when signals are blocked; and (3) RTK was most accurate, but frequently failed to find fixed-integer solutions when nearby obstacles caused cyclic slips. Argiropoulou and Doucas (2015) investigated positioning accuracy under broadleaf canopy in the conditions with and without leaves using single-base RTK, VRS, master-auxiliary concept (MAC), and network DGPS. The horizontal root mean square errors (RMSE) were simply compared according to positioning solutions and canopy conditions (leaf-on/leaf-off). They reported the RMSE in leaf-on condition is almost double of RMSE in leaf-off condition for all RTN solutions, but RMSEs are about same in network DGPS solution. Overall, the data from leaf-off condition are more accurate than from leaf-on condition. The authors ranked the approaches but their research doesn't compare the results with statistical tests, basing the conclusions only on a simple comparison of descriptive statistics. Also RMSE cannot show the positive or negative tendencies of the positioning errors. Pirti (2016) examined positional accuracies of VRS, area correction parameters (FKP, Flächenkorrekturparameter in German), and static positioning in forest area. The data of that

study was the differences between positions from each solution and from a total station. Accuracies of both VRS and FKP are not as good as the accuracy of static positioning. They commented that positions computed from attenuated signals tend to be less accurate, but static positioning produced better accuracy than RTN. The data of their study cannot determine which method, VRS vs. FKP, is more accurate. They also didn't provide a statistical test, so statistical power cannot be estimated. Generally, there is a gap on accuracy assessment about both RTK under deciduous canopy conditions with and without foliage, which is the main concern of our study.

3.1.4 Research Questions and Hypotheses

Hypothesis 1: After controlling the effect of surface absolute humidity, are the RTK accuracies of leaf-on season and leaf-off season identical? For this question the hypothesis to test is established as follows.

H_0: Eliminating the effect of covariate, which is the surface absolute humidity, the RTK errors of both leaf-on and leaf-off status are not significantly different.

H_A: Eliminating the effect of covariate, which is the surface absolute humidity, the RTK errors of both leaf-on and leaf-off status are significantly different.

Hypothesis 2: After controlling the effect of percent sky obstruction, are the RTK accuracies of leaf-on season and leaf-off season identical? For this question the hypothesis to test is established as follows.

$\begin{cases} H_0: \text{ Eliminating the effect of covariate, which is the sky obstruction, the RTK errors of both leaf-on and leaf-off status are not significantly different.} \\ H_A: \text{ Eliminating the effect of covariate, which is the sky obstruction, the RTK errors of both leaf-on and leaf-off status are significantly different..} \end{cases}$

3.2 Material and methods

3.2.1 First-order benchmarks and study locations

Survey markers are monuments with an identifiable point established on a permanent and stable object such as bedrock or a concrete column. Although the terms survey marker and benchmark are frequently used interchangeably, a benchmark is a specific type of a survey marker whose elevation is known and refers to a vertical datum. In Connecticut, the vertical control network was established and is maintained by NGS and CTDOT. The next-generation Federal reference frame will shift vertical control from passive monuments to GNSS methods, which underlines the importance of this study. Federal Geodetic Control Committee (1984) provides the horizontal and vertical standards of accuracy for survey markers. They are coarsely categorized as First-, Second-, and Third-Order, and refined with sub-classes; the First-Order is the most accurate class. This study occupied First-Order benchmarks only.

The benchmarks are in central and eastern Connecticut (Figure 3.1). These benchmarks are near roads and highways, and their surroundings vary in landscape: under heavy forest canopy cover, in the open space, in the urban area, and in the rural area. Tree species around the benchmarks are mostly broadleaf trees including oaks (*Quercus spp*.) and maples (*Acer spp*.).

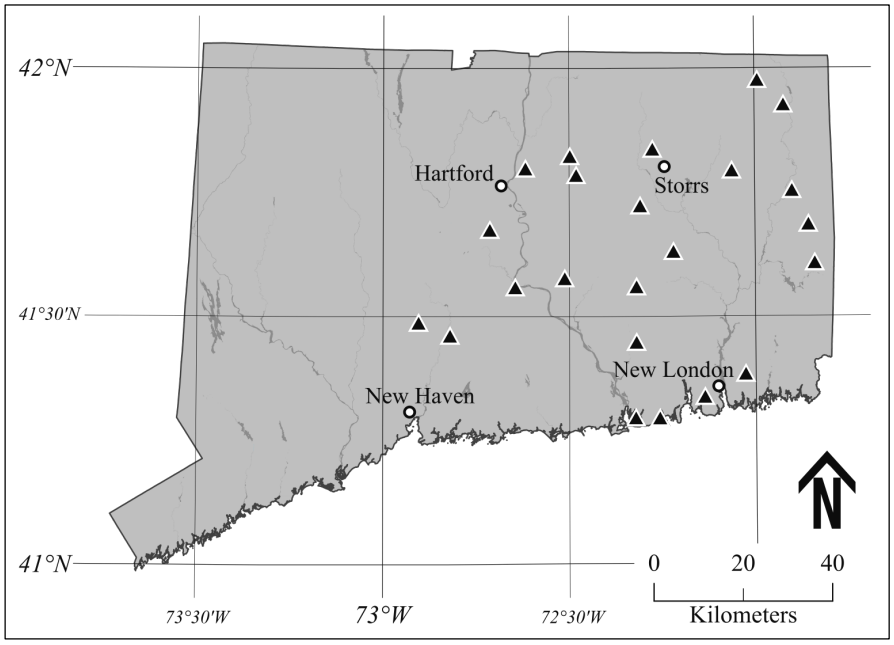


Figure 3.1: The locations of the occupied NGS benchmarks in Connecticut.

3.2.2 Field data collection and variables

GNSS data collection: The data collection was performed for two seasons: the summer of 2012 with leaves on the trees (leaf-on) and the late fall of 2012 to early spring of 2013 while the leaves were off the trees (leaf-off). The survey team occupied each benchmark multiple times and with two different GNSS receivers (Leica GS15, Topcon HiperLite+). **However, the means of the positions produced by the different receivers were not statistically different, so no distinction was made regarding receiver manufacturer – the data from all the receivers were pooled into a single data set.** Both GPS-only and GPS+GLONASS were observed for every occupation using both single-base RTK and VRS modes, a total of four sets of coordinates per session. Again, no difference was found for GPS and GPS+GLONASS, so these data were likewise pooled into a single data set.

Dependent variables: The GNSS positioning data of this project were reported in the State Plane Coordinate System 1983, Connecticut zone (FIPS 0600), in meters. This coordinate system uses a Lambert Conformal Conic map projection producing Cartesian (i.e. grid) coordinates called **eastings** and **northings**. Grid coordinate systems lack any vertical information altogether, so the vertical coordinates refer to the Geodetic Reference System 1980 (GRS 80) reference ellipsoid as placed by the North American Datum of 1983 (NAD 83) 2000 realization. Control coordinates were provided by Meyer, Arifuzzaman, & Massalski (2010) rather than using the NGS coordinates because Meyer's coordinates were quite recent compared to the last time NGS occupied the stations, which was more than 30 years ago in some cases. Using ellipsoid heights also allows geoid models to not be used, which would unhelpfully add another uncertainty to the results. Thus, the dependent variables were three directional errors: easting error (ΔE), northing error (ΔN), and ellipsoid-height error (ΔU). These errors were calculated from:

$$Directional\ Errors = Observed\ Coordinates - Control\ Coordinates$$
$$\text{or}$$
$$\{\Delta E, \Delta N, \Delta U\} = \{E_{obs}, N_{obs}, U_{obs}\} - \{E_{control}, N_{control}, U_{control}\} \tag{3.2}$$

where, E_{obs}, N_{obs}, and U_{obs} are the easting, northing, and ellipsoid-height coordinates of positioning data from the field, and $E_{control}$, $N_{control}$ and $U_{control}$ are precise easting, northing, and ellipsoid-height coordinates from Meyer *et al*. (2010). These three directional errors were used as dependent variables for ANCOVA and RANCOVA.

Independent variable: Tree-canopy status is the independent variable for ANCOVA and RANCOVA. The canopy status has two levels (leaf-on and leaf-off).

Table 3.1: Variables and their descriptive statistics for ANCOVA and RANCOVA model.

(a) Variables in nominal scale (overall and split by tree canopy status)

Variable Name	*Abbr	*n*	Percentage	Comments
Independent Variable:				
Tree canopy status	*Canopy*	502	100%	Two-level categorical variable
Leaf-on condition		241	48%	Reference level of tree-canopy status. Data collected during leaf-on.
Leaf-off condition		261	52%	Data collected during late autumn to very early spring seasons during leaf-off.

(b) Variables in interval scale (overall and split by tree canopy status)

Variable Name	*Abbr	*n*	*Mean* ± *S.D.*	*95% C.I.*	*S.E.M.*	Comments
Covariate:						
Absolute humidity (g/m^3)	*AH*	502	10.96 ± 7.00	(10.35, 11.57)	0.31	Grams of water vapor in a cubic meter. Calculated from relative humidity and temperatures.
Leaf-on condition		241	17.17 ± 4.21	(16.63, 17.70)	0.27	
Leaf-off condition		261	5.23 ± 3.07	(4.85, 5.60)	0.19	
Sky obstruction (%)	*SkO*	502	22.20 ± 18.16	(20.61, 23.79)	0.81	Percentage of sky blocked by trees and structures. Calculated from field photographs taken by fish-eye lens.
Leaf-on condition		241	28.34 ± 21.22	(25.64, 31.03)	1.37	
Leaf-off condition		261	16.53 ± 12.37	(15.02, 18.04)	0.77	
Dependent Variables:						
Easting errors (cm)	*ΔE*	502	0.42 ± 2.61	(0.25, 0.60)	0.09	Horizontal RTK errors represented by east (+) - west (–) direction.
Leaf-on condition		241	0.38 ± 2.21	(0.10, 0.66)	0.14	
Leaf-off condition		261	0.46 ± 1.82	(0.24, 0.69)	0.11	
Northing errors (cm)	*ΔN*	502	– 0.40 ± 2.07	(– 0.58, – 0.22)	0.09	Horizontal RTK errors represented by north (+) - south (–) direction.
Leaf-on condition		241	– 0.26 ± 2.18	(– 0.53, 0.02)	0.14	
Leaf-off condition		261	– 0.53 ± 1.96	(– 0.77. – 0.29)	0.12	
Ellipsoid-height errors (cm)	*ΔU*	502	– 1.02 ± 4.07	(– 1.38, – 0.66)	0.18	Ellipsoid-height RTK errors represented by up (+) - down (–) direction.
Leaf-on condition		241	– 1.45 ± 4.65	(– 2.04, – 0.86)	0.30	
Leaf-off condition		261	– 0.62 ± 3.41	(– 1.03, – 0.20)	0.21	

* Abbr: abbreviations, *n*: sample size, *S.D.*: standard deviation, *C.I.*: confidence interval, *S.E.M.*: standard error of mean.

Covariate: For ANCOVA and RANCOVA of this study, there were two covariates: absolute humidity and sky obstruction. The absolute humidity was derived from relative humidity (%) and ambient temperature (°C). The relative humidity and the air temperature were collected with a Kestrel 3000 handheld weather station described in Chapter 2. The weather

variables were collected from 1 m to 1.5 m above the ground and no more than 3 m from the GNSS receiver during the occupation. The absolute humidity (g/m^3) was computed with (3.1).

Sky obstruction was determined using a *Nikon Coolpix 5000* digital camera with an *FC-E8* fisheye-converter lens. It was mounted on a tripod and leveled using a two-axis camera-mounted bubble level. This camera system was installed atop each benchmark and oriented towards the zenith direction, and the height of the lens was approximately 1.9 m above ground level. The exposure time was automatic and every image was taken with overcast skies and no direct sunlight. Using *ESRI*'s *ArcGIS 10.1* semi-automated approach (ISODATA algorithm), the hemispherical images were classified as two classes (sky, non-sky). Details of the image classification are in Parent and Volin (2014).

Summary of the variables and reason to select ANCOVA: The total sample size of the directional errors of RTK was 502. There were 241 observations for leaf-on condition and 261 observations for leaf-off condition. Table 3.1 describes the dependent variables, independent variable, and covariates, with descriptive statistics about these variables.

3.2.3 Statistical Analyses

Each directional error set ($\Delta E, \Delta N,$ and ΔU) was separately analyzed with ANCOVA and RANCOVA. Before these analyses, the directional error set was tested for the ANCOVA assumptions to verify the data structures.

3.2.3.1 Assumptions for ANCOVA

ANCOVA has several assumptions: no outliers, normality test for dependent variable in continuous scale, homogeneity of variance for dependent variable, linearity between dependent variable and covariance both in continuous scale, homoscedasticity between dependent variable and covariate, and homogeneity of regression slope between groups in independent variable.

Outlier detection: There were two tests for outlier detection: the outlier labeling method with boxplot (Tukey, 1977; Hoaglin, Iglewics, & Tukey, 1986; Seo, 2006) and outlier labeling by standard deviation ($\mu \pm 3\times\sigma$, out of 99% of the data) (Shiffler, 1988; Seo, 2006; Sincich, 2011). The conventional outlier labeling method by boxplot considers data points out of the inner fence, $Q_1 - 1.5\times IQR$ to $Q_3 + 1.5\times IQR$, as outliers, where Q_1 is the first quartile, Q_3 is the third quartile, and IQR is the interquartile range. However, we used the inner-fence range of $Q_1 - 2.2\times IQR$ to $Q_3 + 2.2\times IQR$, which is more conservative (Hoaglin & Iglewicz, 1987). These authors empirically found that approximately half of the detected outliers with 1.5 multiplier actually were not outliers, so they proposed to use 2.2 as a multiplier. The outlier detection by standard deviations considers the data points under -3σ or over $+3\sigma$ as outliers, after standardizing the data.

Normality test: The methods to test the normality were Shapiro-Wilk test (Shapiro & Wilk, 1965) with residuals of $\Delta E, \Delta N,$ and ΔU after fitting ANCOVA. Because there are two separate ANCOVA with different covariates (one with absolute humidity, the other with sky obstruction), the Shapiro-Wilk test was run twice with both residual sets.

Homogeneity of variance: Levene's test of homogeneity of variance (Levene, 1960; Fox, 2008; Fox & Weisberg, 2011) was applied to validate the assumption of homogeneity of variance for three dependent variables ($\Delta E, \Delta N,$ and ΔU) according to the canopy status (leaf-

on/leaf-off). The null hypothesis of the Levene's test was the variances of each group (leaf-on/leaf-off) are equal.

Linearity between dependent variable and covariate: Linear regression analysis was the method to test the linearity between covariates (absolute humidity or sky obstruction) and each dependent variable ($\Delta E, \Delta N,$ and ΔU). The fitted equation is $DV = a + b \cdot Cov$, where DV is dependent variable, Cov is covariate, a is estimated constant, and b is estimated coefficient for Cov. The linearity between two variables can be evaluated with the p-value of b.

Homogeneity of regression slope: Homogeneity of regression slopes between each level (leaf-on/leaf-off) of canopy status were tested with ANCOVA including interaction term between canopy status and each covariate (absolute humidity or sky obstruction) (García-Berthou, 2001; Field, Miles, & Field, 2012; Miller & Chapman, 2001). The null hypotheses for estimated coefficient of this interaction term is that the estimated coefficient equals zero, which means there is no interaction or no effect. If the p-value of the estimated coefficient for the interaction term is greater than selected significance level, than it is estimated that there is no interaction effect.

Homoscedasticity between dependent variable and covariate: Non-constant variance (NCV) score test (Breusch & Pagan, 1979; Cook & Weisberg, 1983; Weisberg, 2014) and Studentized Breusch-Pagan test (Breusch & Pagan, 1979; Koenker, 1981; Krämer & Sonnberger, 1986) were performed to evaluate the homoscedasticity of $\Delta E, \Delta N,$ and ΔU to the absolute humidity or sky obstruction covariates. The null hypothesis of both the NCV score test and the Breusch-Pagan test is that the variance of the residuals of each directional errors is constant.

3.2.3.2 Analysis of Covariance model

ANCOVA was used to verify whether there was a statistically significant difference between each directional error ($\Delta E, \Delta N,$ and ΔU) for leaf-on and leaf-off conditions, once their means had been adjusted for surface absolute humidity and for sky obstruction. The absolute humidity was used as covariate in the first analysis, and sky obstruction for the second analysis. Because $\Delta E, \Delta N,$ and ΔU are independent, a multiple analysis of covariance (MANCOVA) was unnecessary. The *R*-statistical software system was used for the ANCOVA test, and the results were verified with *SPSS*.

3.2.3.3 Quade's RANCOVA model

When data violate the assumptions, a parametric test can be abandoned in favor of a ranked test. In this study, a nonparametric counterpart of ANCOVA, which is called the rank ANCOVA or RANCOVA (Quade 1967), was applied. RANCOVA was applied for each of three directional errors ($\Delta E, \Delta N,$ and ΔU) as dependent variables, canopy status with two levels (leaf-on and leaf-off) for independent variable, and the absolute humidity or sky obstruction was the covariate. The analyses with absolute humidity and with sky obstruction were performed separately.

3.3 Results

3.3.1 Data description and diagnosis for ANCOVA

Figure 3.2 shows histograms overlapped with kernel density plots and theoretical normal curves for directional errors at leaf-on, leaf-off. Figure 3.3 shows how directional errors are

scattered by easting vs. northing, easting vs. ellipsoid-height, and northing vs. ellipsoid-height axes.

The mean value of absolute humidity in leaf-on (17.17 g/m^3) is greater than the mean absolute humidity in leaf-off (5.23 g/m^3). The mean sky obstruction in leaf-on (28.34%) was greater than the mean sky obstruction in leaf-off (16.53%).

3.3.2 Assumption tests for ANCOVA

The results of ANCOVA-assumption tests are reported in Appendices 3A–3F. With outlier labeling methods by boxplot with multiplier of 2.2 (Figure 3.4) (Tukey, 1977, Hoaglin *et al*., 1986; Cousineau & Chartier, 2015) and methods by standard deviation ($\mu \pm 3\times\sigma$) (Shiffler, 1988; Seo, 2006; Sincich, 2011), we found outliers from the directional errors (ΔE, ΔN, and ΔU in leaf-on and leaf-off). For each directional error in leaf-on and leaf-off conditions respectively, there are at most two outliers (Appendix 3A). The horizontal outliers are in the range of about 5 cm to 9 cm, and the vertical outliers are in the range of about 11 cm to 16 cm. However, we decided not to exclude these outliers because (1) the number of outliers is small, and (2) there were no differences between the pre-tests of ANCOVA with and without these outliers.

Shapiro-Wilk's test (Razali & Wah, 2011; Shapiro & Wilk, 1965; Verzani, 2014) was selected for the residuals' normality tests (Appendix 3B). (There were two sets of residuals: one for the residuals produced from ANCOVA with absolute-humidity covariate and the other for residuals from ANCOVA with sky-obstruction covariate). The tests with these two sets produced identical results in leaf-on and leaf-off conditions. The leaf-on residuals are normally distributed,

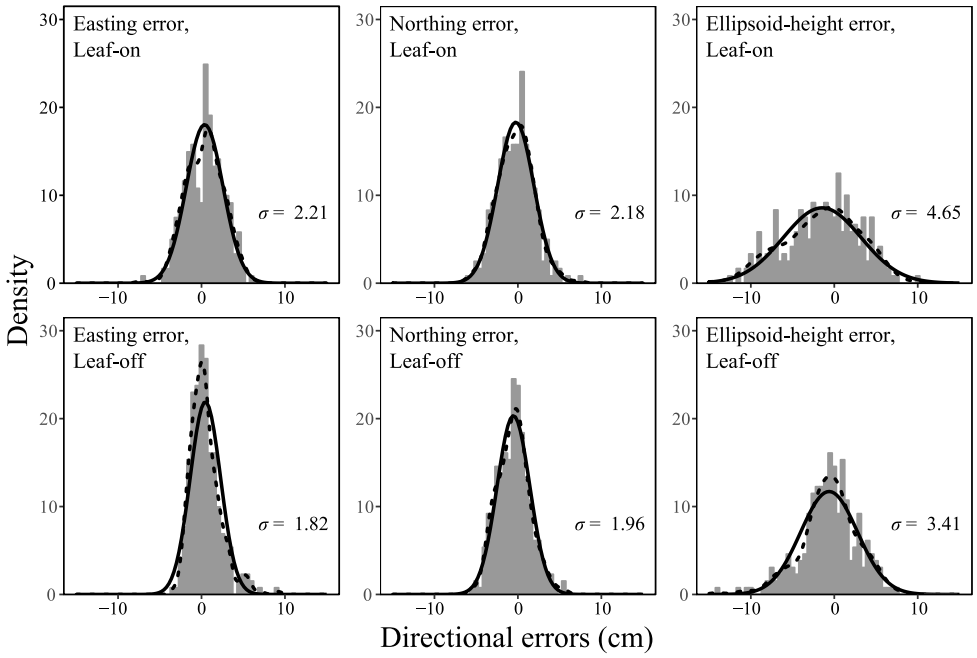


Figure 3.2: Histograms, kernel density plots (dashed lines), and theoretical normal curves (solid lines) for directional errors of RTK data. The unit of each directional error is centimeter and the bin-width bins is 0.5 cm.

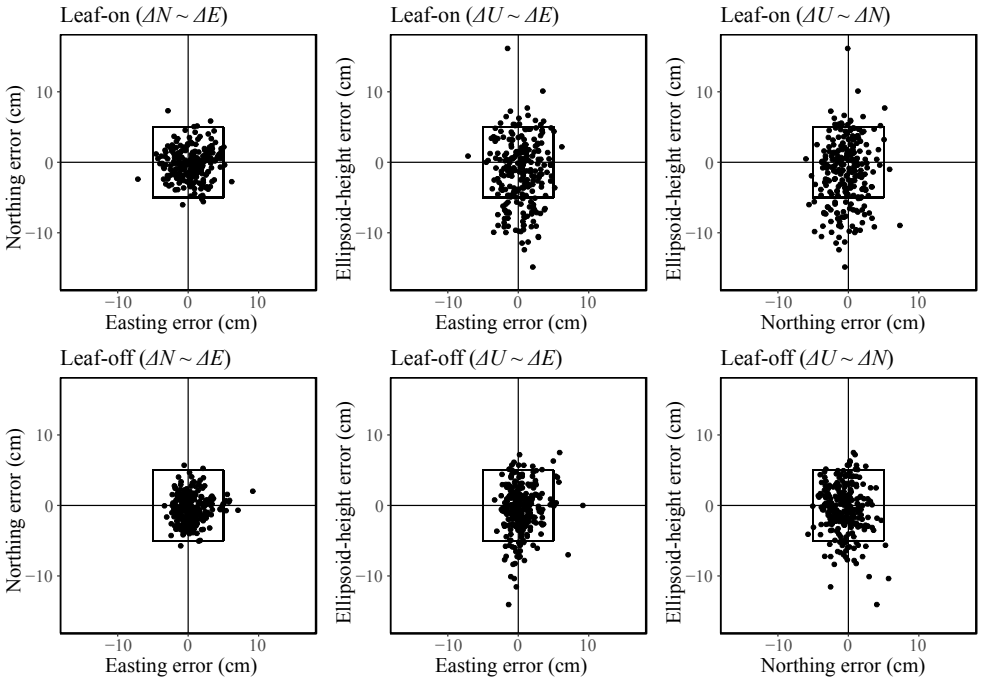


Figure 3.3: Directional scatter plots for the RTK errors. Scatter plots in the first row show the northing ~ easting, ellipsoid-height ~ easting, and ellipsoid-height ~ northing relationships in leaf-on. Scatter plots in the second row show the relationships of each directional error in leaf-off.

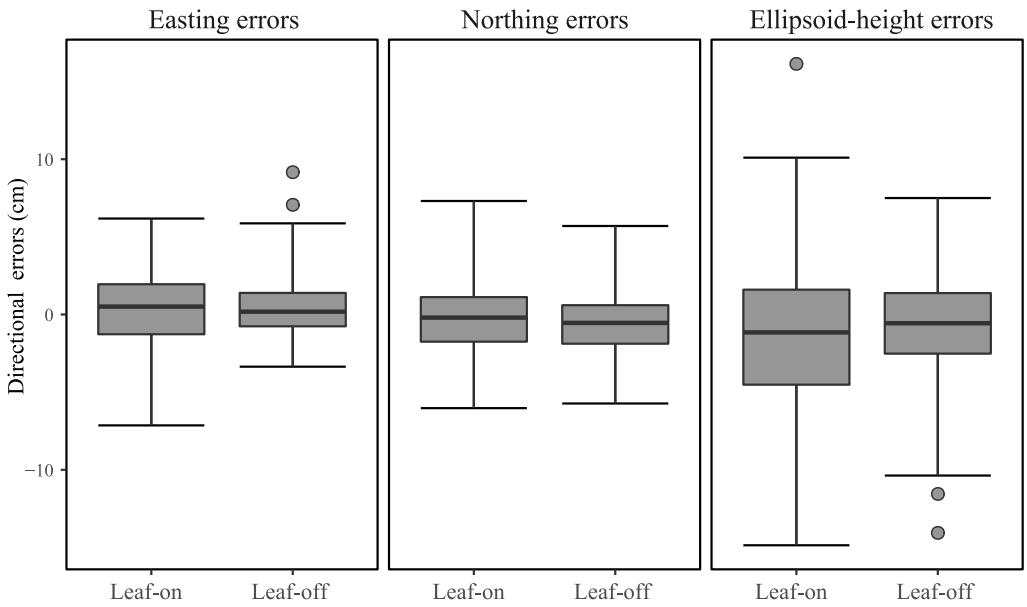


Figure 3.4: Boxplots with whiskers of conservative 2.2 × IQR which is proposed by Hoaglin and Iglewicz (1987).

but only the northing error in leaf-off condition is normally distributed. According to Glass, Peckham, and Sanders (1972), Harwell, Rubinstein, Hayes, and Olds (1992), and Lix, Keselman, and Keselman (1996), ANOVA and ANCOVA are not very sensitive to small deviations from normality. So, parametric tests of ANCOVA can be applied, because the observation number of our data are 241 for leaf-on and 261 for leaf-off.

As shown in Appendix 3C, homogeneity of variance between leaf-on and leaf-off conditions for each directional error were tested using Bartlett's test (Bartlett, 1950), Levene's test (Levene, 1960; Martin & Bridgmon, 2012), and Fligner-Killeen's test (Fligner & Killeen, 1976), and these tests produced identical results. The northing error had homogeneous variances between leaf-on and leaf-off ($p = 0.10$ for Bartlett's test and $p = 0.11$ for the others), but other two directional errors did not.

Linearity between covariate (absolute humidity or sky obstruction) and directional errors are reported in Appendix 3D. Absolute humidity has a linear correlation with horizontal errors in

leaf-on condition and ellipsoid-height error of leaf-off condition, but others do not. Sky obstruction has a linear correlation with easting errors in leaf-on and ellipsoid-height errors in leaf-off. The others don't have a linear relationship with sky obstruction.

Homogeneity of regression slopes between each level (leaf-on/leaf-off) of canopy status were tested (Appendix 3E). Only canopy status and absolute humidity of northing errors ($p = 0.02$) have an interaction effect. The others seem to have no interactions.

Homoscedasticity of residuals for each level of canopy status (leaf-on/leaf-off) were tested and the results of NCV score test and Breusch-Pagan's test were nearly identical. First, the absolute-humidity residuals were tested. The northing errors in leaf-off and ellipsoid-height errors in leaf-on have homoscedastic absolute-humidity residuals. However, the easting errors in leaf-on and leaf-off, the northing errors in leaf-on, and the ellipsoid-height errors in leaf-off have heteroscedastic absolute-humidity residuals (Appendix 3F, (a)). Second, the sky-obstruction residuals were tested. The easting errors in leaf-on, the northing errors in leaf-on, and the ellipsoid-height errors in leaf-off have homoscedastic sky-obstruction residuals. However, the easting errors in leaf-off, the northing errors in leaf-off, and the ellipsoid-height errors in leaf-on have heteroscedastic sky-obstruction residuals (Appendix 3F, (b)).

Our data did not uniformly satisfy the tests' assumptions so a nonparametric version of ANCOVA, which called rank ANCOVA (Conover & Iman, 1982; Lawson, 1983; Quade, 1967), was also applied to complement and to confirm the results from parametric ANCOVA.

Table 3.2: Results of ANCOVA and RANCOVA models for directional errors (ΔE, ΔN, and ΔU) with surface absolute humidity or sky obstruction as covariates. Adjusted means with 95% confidence interval (for ANCOVA) or adjusted medians with the first (Q_1)and the third (Q_3) quartiles (for RANCOVA) are reported.

Error type	Covariate	Model type	*Adjusted mean / adjusted median		*F*-ratio	*p*-value
			Leaf-on	Leaf-off		
ΔE	AH	ANCOVA	–0.22 (–0.61, 0.17)	1.02 (0.65, 1.38)	13.16	*< 0.00*
		RANCOVA	29.10 (–155.40, 145.40)	–15.14 (–102.40, 113.60)	0.23	0.63
	SkO	ANCOVA	0.49 (0.23, 0.75)	0.36 (0.11, 0.61)	0.50	0.48
		RANCOVA	20.50 (–151.30, 145.20)	–15.64 (–108.70, 114.60)	0.29	0.59
ΔN	AH	ANCOVA	–0.79 (–1.19, –0.39)	–0.04 (–0.41, 0.34)	4.63	*0.03*
		RANCOVA	8.66 (–127.00, 122.00)	1.34 (–117.00, 107.70)	1.58	0.21
	SkO	ANCOVA	–0.28 (–0.55, –0.01)	–0.51 (–0.77, –0.25)	1.44	0.23
		RANCOVA	8.29 (–127.20, 132.50)	–3.73 (–121.00, 114.60)	1.64	0.20
ΔU	AH	ANCOVA	–0.71 (–1.50, 0.08)	–1.31 (–2.05, –0.56)	0.75	0.38
		RANCOVA	–14.04 (–140.30, 144.80)	1.09 (–103.10, 103.20)	0.80	0.37
	SkO	ANCOVA	–1.40 (–1.93, –0.87)	–0.67 (–1.17, –0.16)	3.64	0.06
		RANCOVA	–6.50 (–146.00, 146.20)	0.25 (–102.50, 108.30)	2.75	0.10

*Adjusted means of directional errors by covariates with their 95% confidence intervals in parentheses are for ANCOVA. But in case of RANCOVA, adjusted median of residuals calculated from rank-transferred directional errors were reported with the first (Q_1) and the third quartiles (Q_3) in parentheses.

3.3.3 ANCOVA and RANCOVA for the directional errors adjusting for confounding effects from surface absolute humidity and sky obstruction

The overall results of ANCOVA and RANCOVA were identical as reported in Table 3.2, except the easting and the northing errors between leaf-on and leaf-off conditions after

adjustment for the confounding effects of absolute humidity. The scatterplots for ANCOVA for each directional error are presented in Appendix 3G.

Easting error (ΔE): Taking absolute humidity as a covariate, ANCOVA showed that easting errors of leaf-on and leaf-off were significantly different ($p < 0.00$). However, the result of RANCOVA reported a different result ($p = 0.63$) that these two easting errors are not significantly different. Although the ANCOVA result shows a difference of easting errors between leaf-on and leaf-off conditions, we cannot neglect that there were several violations during the assumption tests. Because of these violations, we put more weight on the RANCOVA results. Only when the ANCOVA and RANCOVA results are the same, it is more safe and reasonable to conclude there is a difference between the two easting-error groups. When sky obstruction was a covariate, ANCOVA ($p = 0.48$) and RANCOVA ($p = 0.59$) showed identical results that there was no difference between easting errors in the two conditions. Overall, it might be hard to conclude there is a significant difference between easting errors of two canopy conditions after adjusted the ground humidity effect and sky obstruction.

Northing error (ΔN): Similar to the easting-error case, the results of ANCOVA and RANCOVA of the northing error taking absolute humidity as a covariate produced different results. ANCOVA showed that northing errors of leaf-on and leaf-off were significantly different ($p = 0.03$), whereas RANCOVA showed northing errors of leaf-on and leaf-off were not significantly different ($p = 0.21$). For the sky-obstruction covariate, the northing errors were not significantly different between leaf-on and leaf-off. Overall, the adjusted means of horizontal errors (ΔE and ΔN) are not significantly different between two seasons regardless of the covariates.

Ellipsoid-height error (ΔU): For ANCOVA and RANCOVA, the ellipsoid-height errors were not significantly different between leaf-on and leaf-off with surface absolute humidity (p= 0.38 in ANCOVA, p = 0.37 in RANCOVA) and sky obstruction (p= 0.06 in ANCOVA, p = 0.10 in RANCOVA) as confounding factors.

3.4 Discussion

The initial expectation was that RTK accuracies in leaf-on and leaf-off seasons should not be identical after controlling the surface absolute humidity or sky obstruction effects. Our most surprising result, however, is that this was not the case. This result is quite different to previous reports about the negative effect of water vapor and canopy to the accuracy described in research with static positioning and theories written in textbooks (Meyer *et al*., 2010; Hofmann-Wellenhof *et al.*, 2008; Leick *et al.*, 2015). Also this result is different to the previous studies about the effect of negative canopy effect on positioning accuracy performed with code-based GNSS (Danskin, Bettinger, Jordan, & Cieszewski, 2009; DeCesare, Squires, & Kolbe, 2005; Klimánek, 2010; Rempel, Rodgers, & Abraham; 1995, Rodríguez-Pérez, Alvarez, & Sanz-Ablanedo, 2007; Tuček & Ligoš, 2002; Wing & Eklund, 2007) and DGPS (Liu & Brantigan, 1995; Næsset, Bjerke, Øvstedal, & Ryan, 2000; Næsset, 2001; Holden, Martin, Owende, & Ward, 2001). It is especially surprising given we used only real-time positioning so no error mitigation from least-squares post-processing happened.

There might be several explanations for this. (1) The GNSS receivers used in this study have firmware to correct for multipath and humidity. Antennas are constantly being improved, hardware noise floors are dropping, and multipath- rejection firmware is being refined. (2) The humidity measurements occurred only at the Earth's surface, and that beneath tree canopies

much of the time. But the signals propagate through all of the atmosphere, so our humidity measurements might not accurately represent the actual wet delay. Nonetheless, the atmosphere's water vapor is in the troposphere and that mostly in the boundary layer, which extends only a few tree lengths above the ground. Our measurements might well reflect the humidity in the boundary layer, so this point is unclear. (3) The boundary layer is but a tiny portion of the total thickness of the troposphere, so even if the humidity measurement does capture the water vapor in the boundary layer, it might not reflect the entire wet delay, which is the sum of the delays along the whole path. The situation with leaf-on/leaf- off, however, seems clear cut. There are no multipath sources above the trees and the hemispherical photographs should accurately capture the situation – yet there was likewise no statistically significant difference in accuracies. It seems that the hardware and firmware might be operating better than the industry believes.

Unlike the present study, many studies use root mean square error or a displacement (magnitude only) from a control point for GNSS-accuracy assessment (Argiropoulou & Doucas, 2015; Deckert & Bolstad, 1996). There are two problems with this. First, the directional information from positive or negative values is lost – the distributions of the data are, in fact, two-tailed so combining them into a magnitude-only metric unnecessarily throws away information. Second, the horizontal magnitude is not normally distributed in principle: it follows a Rayleigh distribution (Rayleigh, 1919; Siddiqui, 1962) and there are very few statistical tests available for Rayleigh distributions. The Rayleigh distribution requires that its components follow a normal distribution, have the same variance, and are independent. However, real data seldom meet these requirements.

The data were split into three directions (easting, northing, and ellipsoid-height) because it was hard to find an appropriate hypothesis test handling the non-normal and positive-value-only magnitude data at once. So the interpretations of the result of current study are somewhat complicated. To have a meaningful result, there are two scenarios: (1) leaf-on and leaf-off errors of all three directions are significantly different, (2) leaf-on and leaf-off errors of horizontal (easting and northing) directions are significantly different, or (3) leaf-on and leaf-off errors of ellipsoid-height error are significantly different.

The means of the leaf-on and leaf-off ellipsoid height errors are statistically different than zero and the horizontal coordinates are not, which is consistent with Dussault, Courtois, Ouellet, and Huot (2001), and Sigrist *et al*. (1999). Dussault *et al*. (2001) who suggested the reason is that satellite geometry always better for horizontal positioning than vertical positioning because the GNSS radio signals do not penetrate through the Earth. Meyer *et al.* (2010) concluded that the Connecticut vertical control coordinates (NGS) have a systematic bias, but that is not what the current results indicate because they were compared to Meyer's values. Instead, these results suggest that RTK heighting is not as accurate as static occupations adjusted in a network, which is hardly surprising. The standard deviations of the ellipsoid-height error are approximately two times greater than horizontal errors in leaf-on, which is perhaps slightly smaller than would be expected. (The scatterplots in Figure 3.3 show that the ellipsoid-height error is more widely dispersed than the horizontal errors.) Horizontal coordinates can be controlled from all compass points, but vertical coordinates can only be controlled from one direction, from above.

This study enjoyed a very large data set and yet some of our distributions failed the normality tests, mostly due to too-thick tails (see Q-Q plots in Appendix 3H). There is no clear cause for this, but multipath is a likely suspect. However, given that there was no clear

correlation between the presence of leaves and a reduction in accuracy, it is tempting to conclude that the multipath-rejection hardware and firmware in the receivers were highly effective, which contradicts the previous supposition. Only fixed solutions were accepted, and thick tails might imply incorrectly fixed integer ambiguities. This is consistent with cycle-slip rich and poor signal-to-noise ratio environments because the correlation circuits are working with very noisy data and are being reset frequently. We imagine that the integer-fixing processing usually works as desired, which results in positions accurate to a few centimeters. It, thus, acts as a filter and eliminates erroneous data but this mental model is quite wrong if the integer ambiguities are fixed incorrectly. Such positions would be incorrect at the decimeter level, or even more, but the receiver would be completely unware of the situation. Such erroneous positions would appear always in the error distributions' tails because an incorrectly fixed integer imparts offsets commensurate with the carrier's wavelength (~ 20 cm).

In this study, RTK generally yields a fixed solution in canopy coverage up to around 74%. In the sites with higher canopy coverages, it was unable to have fixed solutions. As a test, the field survey team surveyed under dense canopy coverage (> 74%) with RTK for about one and half hours, and the receiver didn't provide a fixed solution. The reason for not having a fixed solution might be the failure of finding enough number of satellites, weak cellphone coverages, and signal attenuation caused by heavy canopy coverages.

The findings of this study might not be very interesting to the general public because error-level of this study is too small for people who use recreational-level GNSS such as a car-navigation GPS. However, these findings might challenge peer scientists to find exact explanations why the result with RTK positioning is quite different to the results with other

modes. Also, these findings are likely to give credibility to surveying professionals who work with RTK in the field with a lot of trees.

3.5 Bibliography

Ahrens, C. D., & Henson, R. (2015). *Meteorology Today: An Introduction to Weather, Climate, and the Environment* (11th ed.). Boston, MA: Cengage Learning. 640 pp.

Argiropoulou, C., & Doucas, K. A. (2015). The Effect of the Canopy of Scots Pines (*P. Sylvestris*) in Positioning Accuracy Utilizing the Network of Permanent GPS Reference Stations of the Hellenic Positioning System (HEPOS). *Global Journal for Research Analysis*, *4*(9), 97-102.

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Mathematical and Statistical Psychology*, *3*(2), 77-85.

Braasch, M. S. (1996). Multipath effects. In B. W. Parkinson & J. J. Spilker (Eds.), *Global Positioning System: Theory and Applications*, *1* (pp. 547-568). Washington, DC: American Institute of Aeronautics and Astronautics. 632 pp.

Breusch, T. S., & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica: Journal of the Econometric Society*, *47*(5), 1287-1294.

Conley, R., Cosentino, R., Hegarty, C. J., Kaplan, E. D., Leva, J. L., de Haag, M. U., & Van Dyke, K. (2006). Performance of Stand-Alone GPS. In E. Kaplan & C. Hegarty (Eds.), *Understanding GPS: Principles and Applications* (2nd ed.) (pp. 301-378). Boston, MA: Artech House. 723 pp.

Connecticut Association of Land Surveyors. (2008). *Guidelines and Specifications for Global Navigation Satellite System Land Surveys in Connecticut*. Wethersfield, CT: Connecticut Association of Land Surveyors. Retrieved from http://ctsurveyors.org/wp-content/uploads/2010/12/GNSS_20080626_online.pdf

Conover, W. J., & Iman, R. L. (1982). Analysis of Covariance using the Rank Transformation. *Biometrics*, *38*(3), 715-724.

Cook, R. D., & Weisberg, S. (1983). Diagnostics for Heteroscedasticity in Regression. *Biometrika*, *70*(1), 1-10.

Cousineau, D., & Chartier, S. (2015). Outliers Detection and Treatment: A Review. *International Journal of Psychological Research*, *3*(1), 58-67.

Danskin, S. D., Bettinger, P., Jordan, T. R., & Cieszewski, C. (2009). A comparison of GPS performance in a southern hardwood forest: Exploring low-cost solutions for forestry applications. *Southern Journal of Applied Forestry*, *33*(1), 9-16.

DeCesare, N. J., Squires, J. R., & Kolbe, J. A. (2005). Effect of forest canopy on GPS-based movement data. *Wildlife Society Bulletin*, *33*(3), 935-941.

Deckert, C., & Bolstad, P. V. (1996). Forest canopy, terrain, and distance effects on global positioning system point accuracy. *Photogrammetric Engineering and Remote Sensing*, *62*(3), 317-321.

Dussault, C., Courtois, R., Ouellet, J. P., & Huot, J. (2001). Influence of Satellite Geometry and Differential Correction on GPS Location Accuracy. *Wildlife Society Bulletin*, *29*(1), 171-179.

Eckl, M. C., Snay, R. A., Soler, T., Cline, M. W., & Mader, G. L. (2001). Accuracy of GPS-derived relative positions as a function of interstation distance and observing-session duration. *Journal of Geodesy*, *75*(12), 633-640.

El-Mowafy, A., Fashir, H., Al Marzooqi, Y., Al Habbai, A., & Babiker, T. (2003). Testing the Dubai virtual reference system (DVRS) National GPS-RTK network. In M. J. Rycroft (Eds.), *Satellite Navigation Systems: Policy, Commercial and Technical Interaction* (pp. 141-150). Dordrecht, Netherlands: Springer. 266 pp.

El-Rabbany, A. (2006). *Introduction to GPS: the global positioning system*. Boston, MA: Artech House. 210 pp.

Federal Geodetic Control Committee. (1984). *Standards and Specifications for Geodetic Control Networks*. Rockville, MD: Federal Geodetic Control Committee, National Oceanic and Atmospheric Administration (NOAA), U.S. Department of Commerce. Retrieved from https://www.ngs.noaa.gov/FGCS/tech_pub/1984-stds-specs-geodetic-control-networks.htm

Field, A. (2009). *Discovering statistics using SPSS*. London, UK: SAGE Publications. 821 pp.

Field, A., Miles, J. & Field, Z. (2012). *Discovering Statistics Using R*. London, UK: SAGE Publications. 992 pp.

Finn, H., Maxwell, M., & Calver, M. (2002). Why does experimentation matter in teaching ecology?. *Journal of Biological Education*, *36*(4), 158-162.

Firth, J., & Brownlie, R. (1998). An Efficiency Evaluation of the Global Positioning System under Forest Canopies. *New Zealand Journal of Forestry*, *43*(1), 19-25.

Fligner, M. A., & Killeen, T. J. (1976). Distribution-free two-sample tests for scale. *Journal of the American Statistical Association*, *71*(353), 210-213.

Fox, J. (2008). *Applied Regression Analysis and Generalized Linear Models* (2nd ed.). Los Angeles, CA: SAGE Publications. 688 pp.

Fox, J. & Weisberg, S. (2011). *An R Companion to Applied Regression* (2nd ed.). Los Angeles, CA: SAGE Publications. 449 pp.

García-Berthou, E. (2001). On the Misuse of Residuals in Ecology: Testing Regression Residuals vs. the Analysis of Covariance. *Journal of Animal Ecology*, *70*(4), 708-711.

Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research*, *42*(3), 237-288.

Grenci, L. M., & Nese, J. M. (2001). *A World of Weather: Fundamentals of Meteorology* (3rd ed.). Dubuque, IA: Kendall Hunt. 550 pp.

Groves, P. D. (2008). *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*. Boston, MA: Artech House. 518 pp.

Hall, S. J., Learned, J., Ruddell, B., Larson, K. L., Cavender-Bares, J., Bettez, N., Groffman, P. M., Grove, J. M., Heffernan, J. B., Hobbie, S. E. & Morse, J. L. (2015). Convergence of microclimate in residential landscapes across diverse cities in the United States. *Landscape Ecology*, *31*(1), 101-117.

Harwell, M. R., Rubinstein, E. N., Hayes, W. S., & Olds, C. C. (1992). Summarizing Monte Carlo results in methodological research: The one-and two-factor fixed effects ANOVA cases. *Journal of educational statistics*, *17*(4), 315-339.

Hasegawa, H., & Yoshimura, T. (2003). Application of Dual-Frequency GPS Receivers for Static Surveying under Tree Canopies. *Journal of Forest Research*, *8*(2), 103-110.

Henning, W. (2011). *User Guidelines for Single Base Real Time GNSS Positioning* (ver. 2.1). National Geodetic Survey (NGS), National Oceanic and Atmospheric Administration (NOAA). Retrieved from https://www.ngs.noaa.gov/PUBS_LIB/NGS RealTimeUserGuidelines.v2.1.pdf

Henning, W., Martin, D., Schrock, G., Thompson, G., & Snay, R. (2013). *National Geodetic Survey Guidelines for Real Time GNSS Networks* (ver. 2.2). National Geodetic Survey (NGS), National Oceanic and Atmospheric Administration (NOAA). Retrieved from https://www.ngs.noaa.gov/PUBS_LIB/NGSGuidelinesForRealTimeGNSSNetworksV2.2.pdf

Hoaglin, D. C., Iglewicz, B., & Tukey, J. W. (1986). Performance of Some Resistant Rules for Outlier Labeling. *Journal of the American Statistical Association*, *81*(396), 991-999.

Hoaglin, D. C., & Iglewicz, B. (1987). Fine-Tuning Some Resistant Rules for Outlier Labeling. *Journal of the American Statistical Association*, *82*(400), 1147-1149.

Hofmann-Wellenhof, B., Lichtenegger, H., & Wasle, E. (2008). *GNSS - Global Navigation Satellite Systems: GPS, GLONASS, Galileo and more*. Wien, Austria: Springer. 518 pp.

Holden, N. M., Martin, A. A., Owende, P. M. O., & Ward, S. M. (2001). A method for relating GPS performance to forest canopy. *International Journal of Forest Engineering*, *12*(2), 51-56.

Hu, G. R., Khoo, H. S., Goh, P. C., & Law, C. L. (2003). Development and assessment of GPS virtual reference stations for RTK positioning. *Journal of Geodesy*, *77*(5), 292-302.

Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*, *54*(2), 187-211.

Klimánek, M. (2010). Analysis of the accuracy of GPS Trimble JUNO ST measurement in the conditions of forest canopy. *Journal of Forest Science*, *56*(2), 84-91.

Koenker, R. (1981). A note on studentizing a test for heteroskedasticity. *Journal of Econometrics*, *17*(1), 107–112.

Krämer, W., & Sonnberger, H. (1986). *The Linear Regression Model under Test*. Heidelberg, Germany: Physica-Verlag, 190 pp.

Krebs, C. J. (1999). *Ecological Methodology* (2nd ed.). Menlo Park, CA: Benjamin Cummings. 624 pp.

Lachapelle, G., Henriksen, J., & Melgara, T. (1994). Seasonal Effect of Tree Foliage on GPS Signal Availability and Accuracy for Vehicular Navigation. In *Proceedings of the 7th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS 1994)*, Salt Lake City, UT, September 1994, (pp. 527-532).

Landau, H., Vollath, U., & Chen, X. (2002). Virtual Reference Station Systems. *Journal of Global Positioning Systems*, *1*(2), 137-143.

Langley, R. B. (1998a). GPS Receivers and the Observables. In P. J. Teunissen & A. Kleusberg (Eds.), *GPS for Geodesy* (2nd ed.) (pp. 151-185). Berlin, Germany: Springer. 650 pp.

Langley, R. B. (1998b). Propagation of the GPS Signals. In P. J. Teunissen & A. Kleusberg (Eds.), *GPS for Geodesy* (2nd ed.) (pp. 111-149). Berlin, Germany: Springer. 650 pp.

Lawson, A. (1983). Rank Analysis of Covariance: Alternative Approaches. *Journal of the Royal Statistical Society. Series D: The Statistician*, *32*(3), 331-337.

Leick, A., Rapoport, L., & Tatarnikov, D. (2015). *GPS satellite surveying* (4th ed.). Hoboken, NJ: John Wiley & Sons. 840 pp.

Levene, H. (1960). Robust tests for equality of variances. In I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Meadow, & H. B. Mann (Eds.), *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (pp. 278-292). Stanford, CA: Stanford University Press. 517 pp.

Liu, C. J., & Brantigan, R. (1995). Using differential GPS for forest traverse surveys. *Canadian Journal of Forest Research*, *25*(11), 1795-1805.

Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance *F* test. *Review of educational research*, *66*(4), 579-619.

Mander, P. (2012, August 4). How to convert relative humidity to absolute humidity. *CarnotCycle*. Retrieved from https://carnotcycle.wordpress.com/2012/08/04/how-to-convert-relative-humidity-to-absolute-humidity/

Martin, W. E., & Bridgmon, K. D. (2012). *Quantitative and Statistical Research Methods: From Hypothesis to Results*. San Francisco, CA: John Wiley & Sons. 496 pp.

Meyer, T. H., Bean, J. E., Ferguson, C. R., & Naismith, J. M. (2002). The effect of broadleaf canopies on survey-grade horizontal GPS/GLONASS measurements. University of Connecticut, *Department of Natural Resources and the Environment Articles*. *11*. Retrieved from http://digitalcommons.uconn.edu/nrme_articles/11

Meyer, T. H., Arifuzzaman, K., & Massalski, D. (2010). Assessing the Accuracy of GEOID03 GEOID09 in Connecticut. *Surveying and Land Information Science*, *70*(2), 89-101.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding Analysis of Covariance. *Journal of abnormal psychology*, *110*(1), 40-48.

Milliken G. A., & Johnson, D. E. (2002). *Analysis of Messy Data, Volume III: Analysis of Covariance*. Boca Raton, FL: Chapman and Hall/CRC. 624 pp.

Morales, Y., & Tsubouchi, T. (2007). DGPS, RTK-GPS and StarFire DGPS Performance under Tree Shading Environments. In *2007 IEEE International Conference on Integration Technology (ICIT '07)*, *Shenzhen*, *China*, (pp. 519-524).

Næsset, E., Bjerke, T., Øvstedal, O., & Ryan, L. H. (2000). Contributions of differential GPS and GLONASS observations to point accuracy under forest canopies. *Photogrammetric Engineering and Remote Sensing*, *66*(4), 403-407.

Næsset, E. (2001). Effects of differential single-and dual-frequency GPS and GLONASS observations on point accuracy under forest canopies. *Photogrammetric engineering and remote sensing*, *67*(9), 1021-1026.

Parent, J. R., & Volin, J. C. (2014). Assessing the potential for leaf-off LiDAR data to model canopy closure in temperate deciduous forests. *ISPRS Journal of Photogrammetry and Remote Sensing*, *95*, 134-145.

Petovello, M. (2011). GNSS Solutions: Virtual Reference Stations. *Inside GNSS*, *6*(4), 28-31.

Pirti, A. (2005). Using GPS near the forest and quality control. *Survey Review*, *38*(298), 286-298.

Pirti, A. (2016). The seasonal effects of deciduous tree foliage in CORS-GNSS measurements (VRS/FKP). *Tehnicki Vjesnik: Technical Gazette*, *23*(3), 769-774.

Pullen, S., & Rife, J. (2009). Differential GNSS: Accuracy and Integrity. In S. Gleason & D. Gegre-Egziabher (Eds.), *GNSS Applications and Methods* (pp. 87-119). Boston, MA: Artech House. 508 pp.

Quade, D. (1967). Rank Analysis of Covariance. *Journal of the American Statistical Association*, *62*(320), 1187-1200.

Rayleigh, J. W. S. (1919). On the problem of random vibrations, and of random flights in one, two, or three dimensions. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *37*(220), 321-347.

Razali, N. M., & Wah, Y. B. (2011). Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. *Journal of Statistical Modeling and Analytics*, *2*(1), 21-33.

Rempel, R. S., Rodgers, A. R., & Abraham, K. F. (1995). Performance of a GPS animal location system under boreal forest canopy. *The Journal of wildlife management*, *59*(3), 543-551.

Retscher, G. (2002). Accuracy Performance of Virtual Reference Station (VRS) Networks. *Journal of Global Positioning Systems*, *1*(1), 40-47.

Rodríguez-Pérez, J. R., Alvarez, M. F., & Sanz-Ablanedo, E. (2007). Assessment of low-cost GPS receiver accuracy and precision in forest environments. *Journal of Surveying Engineering*, *133*(4), 159-167.

Rutherford, A. (2001). *Introducing ANOVA and ANCOVA: a GLM approach*. London, UK: Sage. 182 pp.

Savage, N., Ndzi, D., Seville, A., Vilar, E., & Austin, J. (2003). Radio wave propagation through vegetation: Factors influencing signal attenuation. *Radio Science*, *38*(5), 1088–2002.

Seo, S. (2006). *A review and comparison of methods for detecting outliers in univariate data sets* (Master theses, University of Pittsburgh).

Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, *52*(3-4), 591-611.

Shiffler, R. E. (1988). Maximum Z Scores and Outliers. *The American Statistician*, *42*(1), 79-80.

Siddiqui, M. M. (1962). Some problems connected with Rayleigh distributions. *Journal of Research of the National Bureau of Standards, Section D: Radio Propagation*, *66D*(2), 167-174.

Sigrist, P., Coppin, M., & Hermy. (1999). Impact of forest canopy on quality and accuracy of GPS measurements. *International Journal of Remote Sensing, 20*(18), 3595-3610.

Sincich, T. (2011). *Business Statistics by Example* (5th ed.). Upper Saddle River, NJ: Prentice Hall. 1179 pp.

Snay, R. A., & Soler, T. (2008). Continuously operating reference station (CORS): history, applications, and future enhancements. *Journal of Surveying Engineering*, *134*(4), 95-104.

Spilker, J. J. (1996). Foliage attenuation for land mobile users. In B. W. Parkinson & J. J. Spilker (Eds.), *Global Positioning System: Theory and Applications*, *1* (pp. 569-583). Washington, DC: American Institute of Aeronautics and Astronautics. 632 pp.

Steel, R. G., Torrie, J. H., & Dickey, D. A. (1997). *Principles and procedures of statistics: A biometrical approach* (3rd ed.). Boston, MA: McGraw-Hill. 666 pp.

Strang, G., & Borre, K. (1997). *Linear algebra, geodesy, and GPS*. Wellesley, MA: Wellesley-Cambridge Press. 624 pp.

Tabachnick, B. G., & Fidell, L. S. (2012). *Using Multivariate Statistics* (6th ed.). Boston, MA: Pearson. 1024 pp.

Talbot, N., Lu, G., Allison, T., & Vollath, U. (2002). Broadcast network RTK-transmission standards and results. In *Proceedings of the 15th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS 2002)*, Portland, OR, September 2002, (pp. 2379- 2387).

Teunissen, P. J., & Kleusberg, A. (1998). GPS Observation Equations and Positioning Concepts. In P. J. Teunissen & A. Kleusberg (Eds.), *GPS for Geodesy* (2nd ed.) (pp. 187-229). Berlin, Germany: Springer. 650 pp.

Texas Department of Transportation. (2016). *TxDOT Survey Manual*. Texas Department of Transportation. Retrieved from http://onlinemanuals.txdot.gov/txdotmanuals/ess/ess.pdf

Tuček, J., & Ligoš, J. (2002). Forest canopy influence on the precision of location with GPS receivers. *Journal of Forest science*, *48*(9), 399-407.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, PA: Addison-Wesley. 688 pp.

Van Sickle, J. (2008). *GPS for Land Surveyors* (3rd ed.). New York, NY: Taylor and Francis. 284 pp.

Verzani, J. (2014). *Using R for introductory statistics* (2nd ed.). Boca Raton, FL: CRC Press. 518 pp.

Vogt, W. P. (2005). *Dictionary of Statistics and Methodology: A Nontechnical Guide for the Social Sciences* (3nd ed.). Thousand Oaks, CA: Sage Publications. 355 pp.

Weber, G., Dettmering, D., & Gebhard, H. (2005). Networked transport of RTCM via internet protocol (NTRIP). In F. Sanso (Eds.), *A Window on the Future of Geodesy: Proceedings of the International Association of Geodesy General Assembly, Sapporo, Japan, 2003, Symposium Series, 128*, (pp. 60-64).

Wegener, V., & Wanninger, L. (2005). Communication options for network RTK. *International Association of Geodesy (IAG), Working Group 4.5.1: Network RTK (2003-2007)*. Retrieved from http://www.wasoft.de/e/iagwg451/wegener/communication.html

Weisberg, S. (2014). *Applied Linear Regression* (4th ed.). Hoboken, NJ: John Wiley & Sons. 384 pp.

Wing, M. G., & Eklund, A. (2007). Performance comparison of a low-cost mapping grade global positioning systems (GPS) receiver and consumer grade GPS receiver under dense forest canopy. *Journal of Forestry*, *105*(1), 9-14.

Zilkoski, D. B., D'Onofrio, J. D., & Frakes, S. J. (1997). *Guidelines for Establishing GPS-derived Ellipsoid Heights (standards, 2 Cm and 5 Cm) Version 4.3* (NOS NGS-58). Silver Spring, MD: National Geodetic Survey (NGS), National Oceanic and Atmospheric Administration (NOAA), U.S. Department of Commerce. Retrieved from ftp://www.ngs.noaa.gov/pub/marti/Precise%20Leveling%20Workshop%20CA/Manuals/NGS-58.pdf

Appendix 3A: Outlier detection of each directional error (ΔE, ΔN, and ΔU) by each level (leaf-on/leaf-off) of tree canopy status

Units in centimeter

	Detection by boxplot †	Detection by standard deviation ‡
Easting errors (ΔE)		
Leaf-on	None	One (–7.14)
Leaf-off	Two (9.17, 7.07)	Two (9.17, 7.07)
Northing errors (ΔN)		
Leaf-on	None	One (7.31)
Leaf-off	None	One (5.70)
Ellipsoid-height errors (ΔU)		
Leaf-on	One (16.15)	One (16.15)
Leaf-off	Two (–11.55, –14.06)	Two (–11.55, –14.06)

† Outlier labeling method by boxplot proposed by Tukey. Tukey (1977) and Hoaglin, Iglewics and Turkey (1986) proposed to use 1.5 as a multiplier for interquartile range (IQR). Later, Howglin and Iglewicz (1987) re-proposed to use 2.2 as a multiplier, because they empirically found that about a half of detected outliers with multiplier 1.5 actually were not an outlier. In our research, we used 2.2 as a multiplier, so that a value $Q_1 - 2.2\times IQR$ and $Q_3 + 2.2\times IQR$ for inner fence.

‡ Outlier labeling method by standard deviation (Shiffler, 1988; Seo, 2006; Sincich, 2011). This method begins with standardization of all the values in the data with $Z = x - \mu / \sigma$, where x is a value of the data, μ is the mean, and σ is the standard deviation. A value whose Z-score is under -3σ or over 3σ is considered as an outlier.

Appendix 3B: Normality tests

Shapiro-Wilk test for normality of directional errors: The null hypothesis of this Shapiro-Wilk's test (Shapiro & Wilk, 1965; Verzani, 2014) is that the residuals of the directional error are normally distributed, and the alternative hypothesis is that the residuals are not normally distributed. The results of the tests for different residual sets from fitted ANCOVAs with two covariates respectively (absolute humidity and sky obstruction) are identical for leaf-on and leaf-off. (Shapiro-Wilk test of whole dataset including leaf-on and leaf-off together is an extra information.) Residuals of all directional errors in leaf-on condition are normally distributed, but only northing error in leaf-off condition is normally distributed. Directional errors in leaf-off condition cannot be estimated as normally distributed.

(a) Shapiro-Wilk test with ANCOVA residuals when surface absolute humidity is a covariate.

	Leaf-on	Leaf-off
Easting errors (ΔE)	$W = 0.9910$ $p = 0.14$	$W = 0.9485$ $p \ll 0.00$
Northing errors (ΔN)	$W = 0.9945$ $p = 0.53$	$W = 0.9934$ $p = 0.31$
Ellipsoid-height errors (ΔU)	$W = 0.9908$ $p = 0.13$	$W = 0.9850$ $p < 0.01$

(b) Shapiro-Wilk test with ANCOVA residuals when sky obstruction is a covariate.

	Leaf-on	Leaf-off
Easting errors (ΔE)	$W = 0.9918$ $p = 0.20$	$W = 0.9480$ $p \ll 0.00$
Northing errors (ΔN)	$W = 0.9961$ $p = 0.81$	$W = 0.9941$ $p = 0.40$
Ellipsoid-height errors (ΔU)	$W = 0.9896$ $p = 0.08$	$W = 0.9817$ $p < 0.01$

Appendix 3C: Homogeneity of variances for each level of independent variable

Bartlett's test, Levene's test, and Fligner-Kelleen's test are available for a test for homogeneity of variances. Bartlett's test has a strength for a normally distributed dataset. However, Levene's test is more robust to departures from normality than Bartlett's test, and Fligner-Killeen's test is still more robust from non-normal data. Fligner-Killeen's test is a nonparametric test. Null hypothesis is that variances of leaf-on and leaf-off are identical, and alternative hypothesis is that variances of leaf-on and leaf-off are different. In this study, results for three different tests are identical.

Directional errors	Test statistics and p-values for homogeneity of variance tests
Easting errors (ΔE)	Bartlett test: $K^2 = 9.42$, $p < 0.00$ Levene's test: $F(1, 500) = 15.28$, $p \ll 0.00$ Fligner-Killeen median test: $FK = 17.29$, $p \ll 0.00$
Northing errors (ΔN)	Bartlett test: $K^2 = 2.74$, $p = 0.10$ Levene's test: $F(1, 500) = 2.60$, $p = 0.11$ Fligner-Killeen median test: $FK = 2.54$, $p = 0.11$
Ellipsoid-height errors (ΔU)	Bartlett test: $K^2 = 24.04$, $p \ll 0.00$ Levene's test: $F(1, 500) = 24.96$, $p \ll 0.00$ Fligner-Killeen median test: $FK = 24.89$, $p \ll 0.00$

Appendix 3D: Linearity between covariate and dependent variable using linear regression

(a) Absolute humidity (g/m^3): the linearity between directional errors (dependent variables) and absolute humidity (covariate) is tested by a linear regression analysis with the estimated coefficient for absolute humidity (b), its t-statistic, and p-value. As an information R^2, F-statistic, and p-value for overall model are also reported. The null hypothesis is that the coefficient of absolute humidity is equal to zero, that means there is no effect.

	Leaf-on	Leaf-off
Easting errors (ΔE)	Estimated coefficients: $b = 1.12\times10^{-3}$ $t = 3.71$ $p \ll 0.00$	Estimated coefficients: $b = 5.07\times10^{-4}$ $t = 1.38$ $p = 0.17$
	For overall model: $R^2 = 0.05$ $F(1, 239) = 13.74$ $p \ll 0.00$	For overall model: $R^2 = 7.34\times10^{-3}$ $F(1, 259) = 1.92$ $p = 0.17$
Northing errors (ΔN)	Estimated coefficients: $b = 1.31\times10^{-3}$ $t = 4.05$ $p \ll 0.00$	Estimated coefficients: $b = 8.58\times10^{-5}$ $t = 0.22$ $p = 0.83$
	For overall model: $R^2 = 0.06$ $F(1, 239) = 16.39$ $p \ll 0.00$	For overall model: $R^2 = 1.81\times10^{-4}$ $F(1, 259) = 0.05$ $p = 0.83$
Ellipsoid-height errors (ΔU)	Estimated coefficients: $b = -1.03\times10^{-3}$ $t = -1.44$ $p = 0.15$	Estimated coefficients: $b = -1.49\times10^{-3}$ $t = -2.18$ $p = 0.03$
	For overall model: $R^2 = 8.64\times10^{-3}$ $F(1, 239) = 2.08$ $p = 0.15$	For overall model: $R^2 = 1.81\times10^{-2}$ $F(1, 259) = 4.76$ $p = 0.03$

(b) Sky obstruction (%): the linearity between directional errors (dependent variables) and sky obstruction in percent (covariate) is tested by a linear regression analysis with the estimated coefficient for sky obstruction (b), its t-statistic, and p-value. As an information R^2, F-statistic, and p-value for overall model are also reported. The null hypothesis is that the coefficient of absolute humidity is equal to zero, that means there is no effect.

	Leaf-on	Leaf-off
Easting errors (ΔE)	Estimated coefficients: $b = -2.15\times10^{-2}$ $t = -3.26$ $p < 0.01$	Estimated coefficients: $b = -1.04\times10^{-2}$ $t = -1.14$ $p = 0.26$
	For overall model: $R^2 = 4.25\times10^{-2}$ $F(1, 239) = 10.60$ $p < 0.01$	For overall model: $R^2 = 4.96\times10^{-3}$ $F(1, 259) = 1.29$ $p = 0.26$
Northing errors (ΔN)	Estimated coefficients: $b = 2.40\times10^{-3}$ $t = 0.36$ $p = 0.72$	Estimated coefficients: $b = 6.89\times10^{-3}$ $t = 0.70$ $p = 0.48$
	For overall model: $R^2 = 5.47\times10^{-4}$ $F(1, 239) = 0.13$ $p = 0.72$	For overall model: $R^2 = 1.89\times10^{-3}$ $F(1, 259) = 0.49$ $p = 0.48$
Ellipsoid-height errors (ΔU)	Estimated coefficients: $b = 2.54\times10^{-3}$ $t = 0.18$ $p = 0.86$	Estimated coefficients: $b = -3.87\times10^{-2}$ $t = -2.28$ $p = 0.02$
	For overall model: $R^2 = 1.34\times10^{-4}$ $F(1, 239) = 3.20\times10^{-2}$ $p = 0.86$	For overall model: $R^2 = 1.97\times10^{-2}$ $F(1, 259) = 5.21$ $p = 0.02$

Appendix 3E: Homogeneity of regression slope between covariate and independent variable (canopy status) using ANCOVA with interaction terms

ANCOVAs with interaction term of independent variable and covariate are pretested. When regression slopes for all the group (leaf-on/leaf-off) are parallel, there are no interactions between a covariate and an independent variable, thus they are independent.

(a) Test between canopy status (independent variable) and absolute humidity (g/m^3, covariate). The estimated coefficient for the interaction term (b), its t-statistic, and p-value are reported. The null hypothesis is that the coefficient of the interaction term is equal to zero, that means there is no interaction effect.

	Estimated coefficient / Interaction (Canopy × Absolute humidity)
Easting errors (ΔE)	$b = -7.17\times10^{-4}$ $t = -1.43$ $p = 0.15$
Northing errors (ΔN)	$b = -1.13\times10^{-3}$ $t = -2.37$ $p = 0.02$
Ellipsoid-height errors (ΔU)	$b = -4.62\times10^{-4}$ $t = -0.45$ $p = 0.65$

(b) Test between canopy status (independent variable) and sky obstruction (%, covariate). The estimated coefficient for the interaction term (b), its t-statistic, and p-value are reported.

	Estimated coefficient / Interaction (Canopy × Sky Obstruction)
Easting errors (ΔE)	$b = 1.11\times10^{-2}$ $t = 0.95$ $p = 0.34$
Northing errors (ΔN)	$b = 4.49\times10^{-3}$ $t = 0.37$ $p = 0.71$
Ellipsoid-height errors (ΔU)	$b = -4.12\times10^{-2}$ $t = -1.74$ $p = 0.08$

Appendix 3F: Homoscedasticity of residuals for each level of independent variables (leaf-on and leaf-off), tested by NCV score test and studentized Breusch-Pagan test

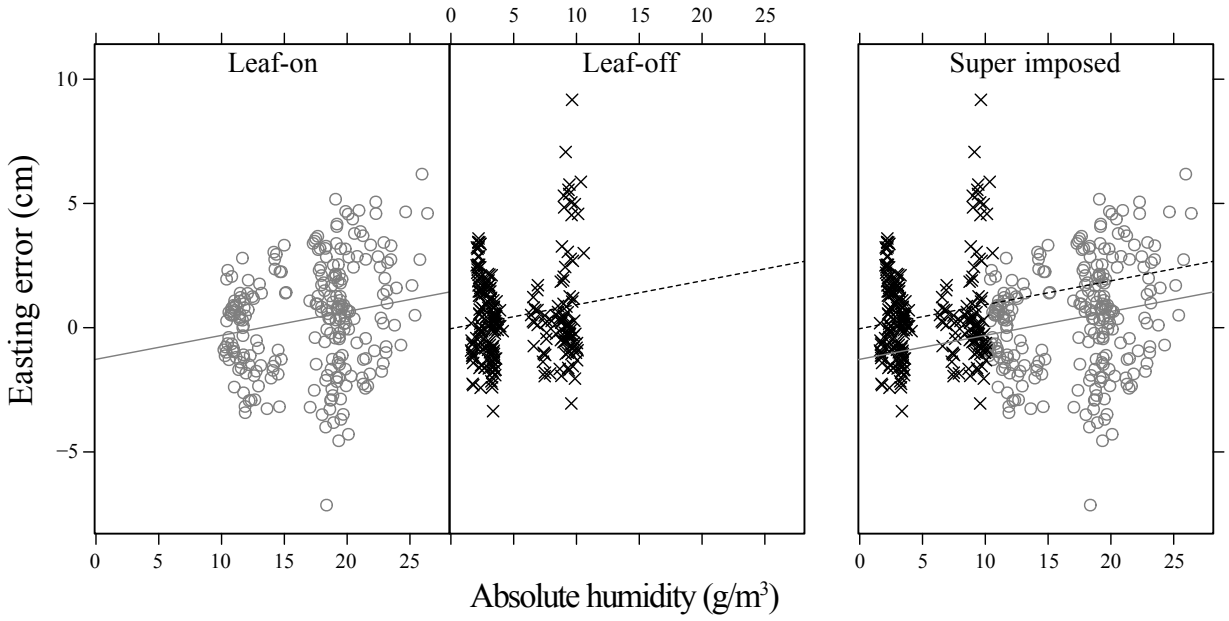
(a) Absolute humidity (g/m^3): the null hypothesis of NCV score test or Breusch-Pagan test is that the variance of the residuals is constant across all values of absolute humidity.

		Non-constant Variance (NCV) Score Test	Studentized Breusch-Pagan Test
Easting Errors (ΔE)			
	Leaf-on	$df = 1$ $\chi^2 = 9.53$ $p < 0.01$	$df = 1$ $BP = 10.61$ $p < 0.01$
	Leaf-off	$df = 1$ $\chi^2 = 22.13$ $p \ll 0.01$	$df = 1$ $BP = 11.06$ $p \ll 0.01$
Northing Errors (ΔN)			
	Leaf-on	$df = 1$ $\chi^2 = 4.95$ $p = 0.03$	$df = 1$ $BP = 4.01$ $p = 0.04$
	Leaf-off	$df = 1$ $\chi^2 = 1.44$ $p = 0.23$	$df = 1$ $BP = 1.37$ $p = 0.24$
Ellipsoid-height Errors (ΔU)			
	Leaf-on	$df = 1$ $\chi^2 = 0.15$ $p = 0.70$	$df = 1$ $BP = 0.14$ $p = 0.71$
	Leaf-off	$df = 1$ $\chi^2 = 14.73$ $p \ll 0.01$	$df = 1$ $BP = 10.18$ $p < 0.01$

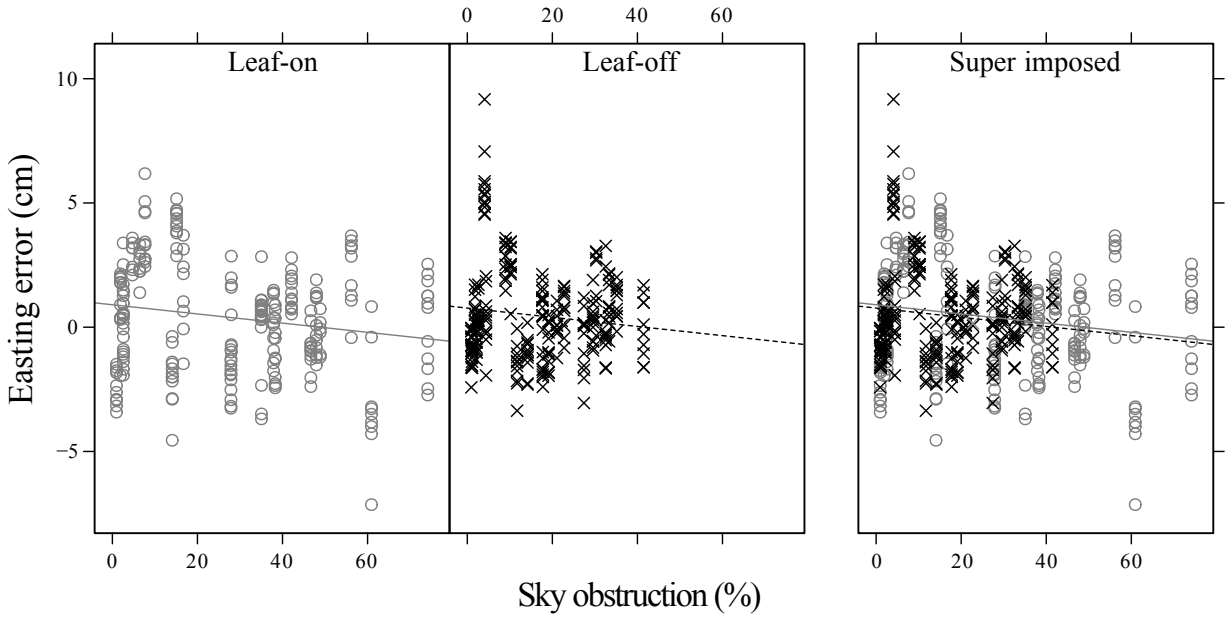
(b) Sky Obstruction (%): the null hypothesis of NCV score test or Breusch-Pagan test is that the variance of the residuals is constant across all values of sky obstruction.

	Non-constant Variance (NCV) Score Test	Studentized Breusch-Pagan Test
Easting Errors (ΔE)		
Leaf-on	$df = 1$ $\chi^2 = 2.11$ $p = 0.15$	$df = 1$ $BP = 2.65$ $p = 0.10$
Leaf-off	$df = 1$ $\chi^2 = 24.87$ $p \ll 0.01$	$df = 1$ $BP = 11.94$ $p \ll 0.01$
Northing Errors (ΔN)		
Leaf-on	$df = 1$ $\chi^2 = 0.02$ $p = 0.89$	$df = 1$ $BP = 0.02$ $p = 0.89$
Leaf-off	$df = 1$ $\chi^2 = 4.16$ $p = 0.04$	$df = 1$ $BP = 4.15$ $p = 0.04$
Ellipsoid-height Errors (ΔU)		
Leaf-on	$df = 1$ $\chi^2 = 5.11$ $p = 0.02$	$df = 1$ $BP = 4.70$ $p = 0.03$
Leaf-off	$df = 1$ $\chi^2 = 0.20$ $p = 0.66$	$df = 1$ $BP = 0.14$ $p = 0.71$

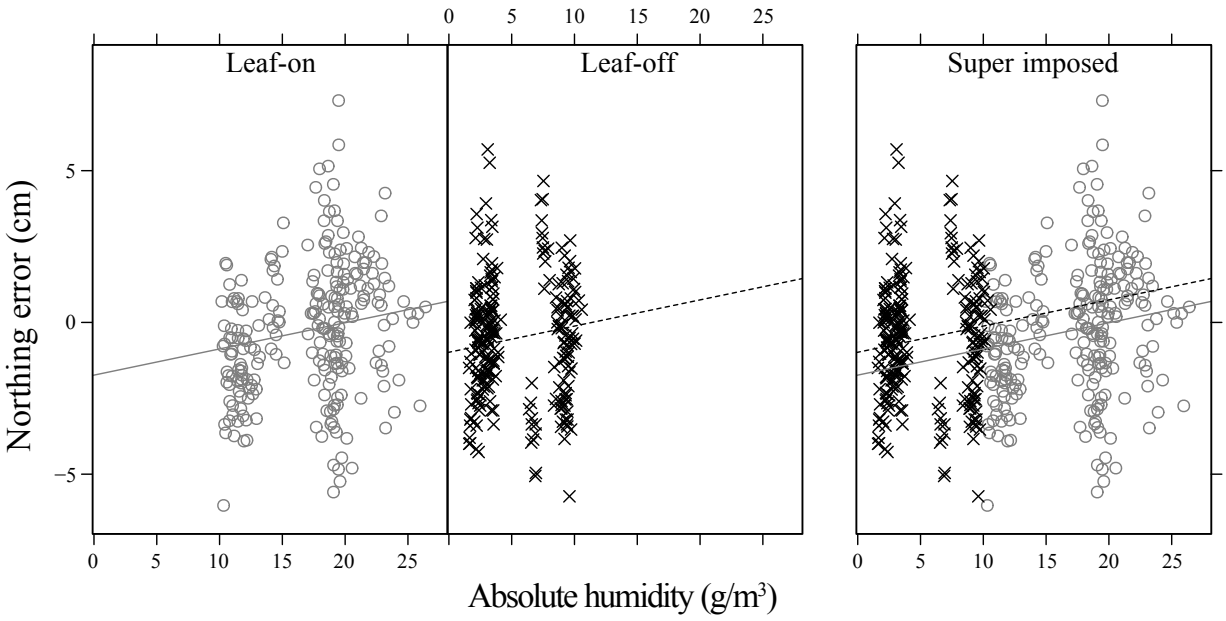
Appendix 3G: Scatterplots comparing leaf-on and leaf-off groups for ANCOVA



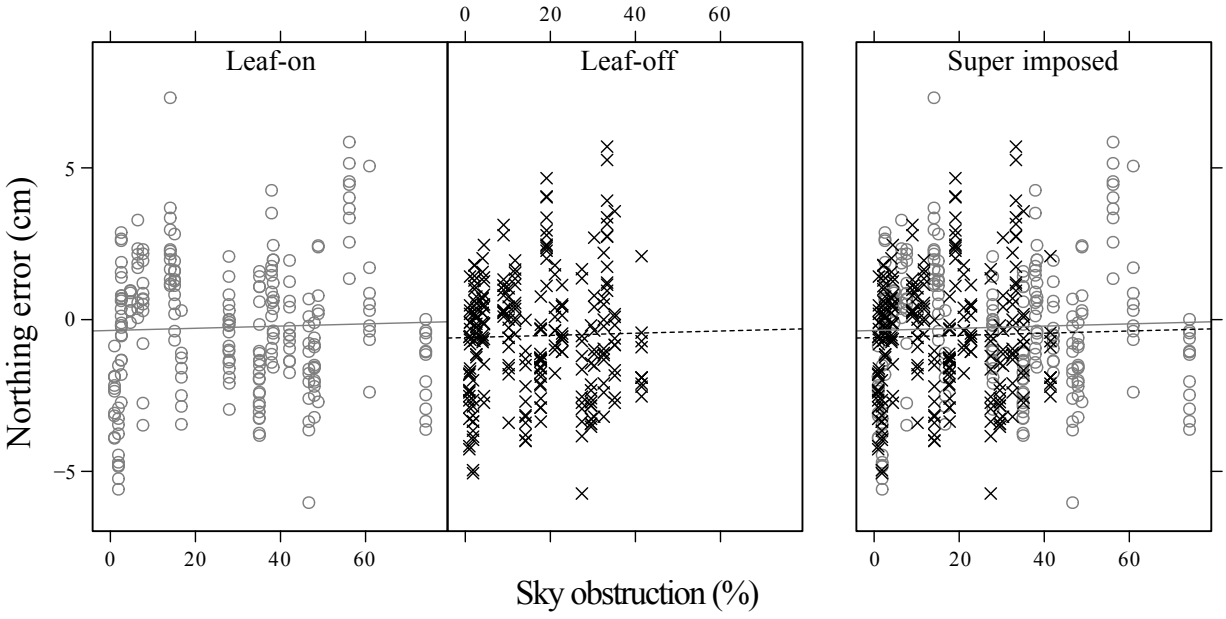
(a) Scatterplots of easting errors (cm) for leaf-on and leaf-off conditions with surface absolute humidity (g/m^3) as a covariate. For all the plots, leaf-on expressed with hollow circles and solid lines, and leaf-off expressed with cross symbols and dashed lines.



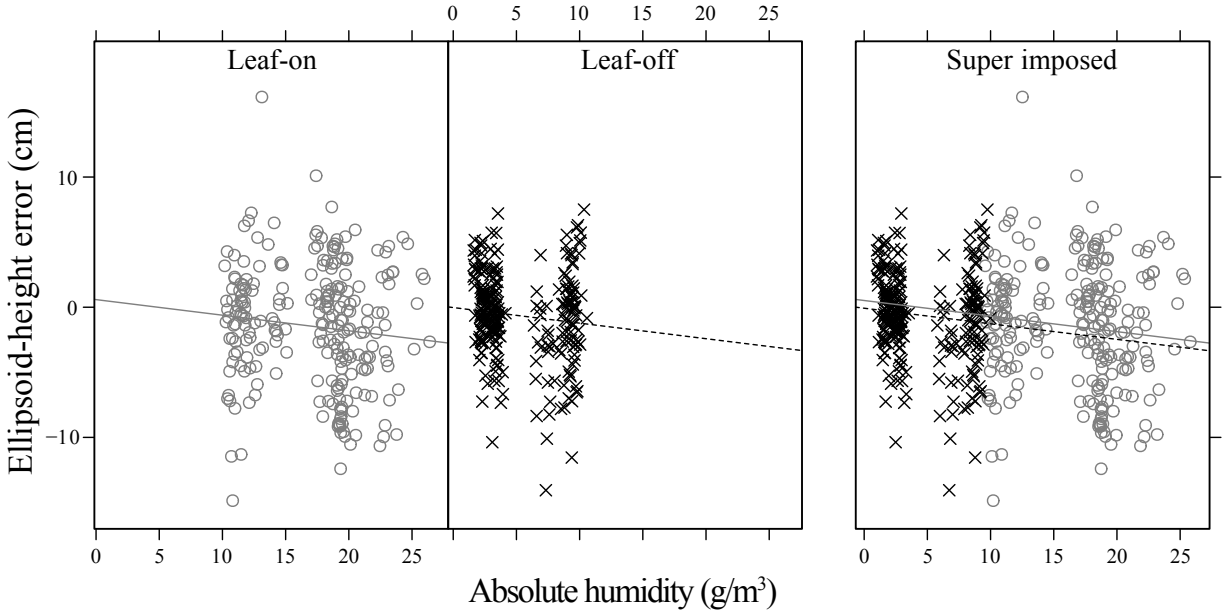
(b) Scatterplots of easting errors (cm) for leaf-on and leaf-off conditions with sky obstruction (%) as a covariate.



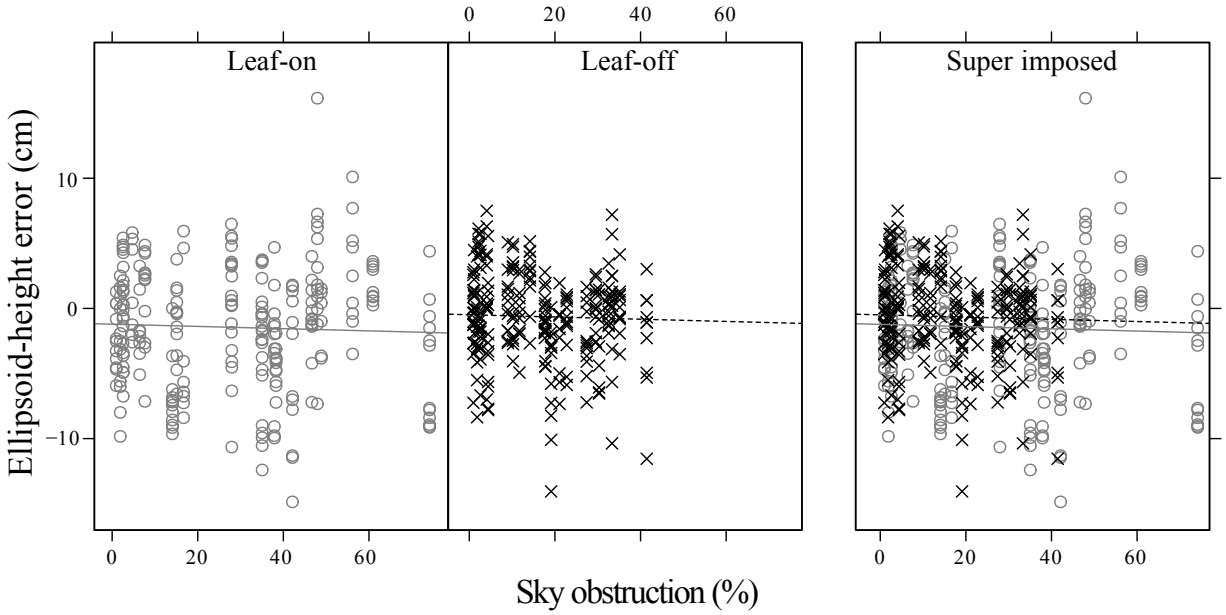
(c) Scatterplots of northing errors (cm) for leaf-on and leaf-off conditions with surface absolute humidity (g/m3) as a covariate.



(d) Scatterplots of northing errors (cm) for leaf-on and leaf-off conditions with sky obstruction (%) as a covariate.

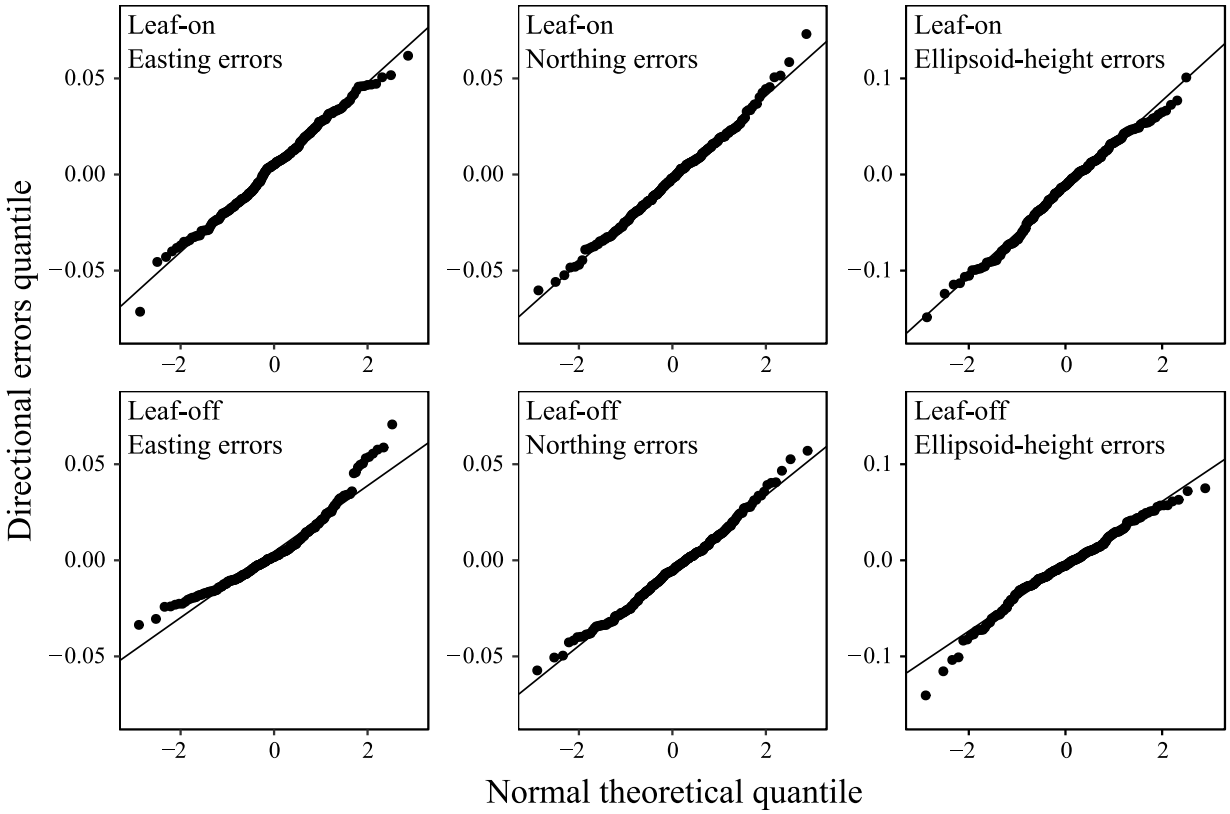


(e) Scatterplots of ellipsoid-height errors (cm) for leaf-on and leaf-off conditions with surface absolute humidity (g/m3) as a covariate.



(f) Scatterplots of ellipsoid-height errors (cm) for leaf-on and leaf-off conditions with sky obstruction (%) as a covariate.

Appendix 3H: Normal Q-Q plots of RTK positioning data



Normal Q-Q plots of RTK directional errors to the theoretical quantile.

Chapter 4

Comparing the Predictive Ability of Four Models of Habitat Suitability and Environmental Favorability

Abstract

Four bentgrass (*Agrostis spp.*) habitat models were compared to determine whether any of the underlying models had better predictive ability than the rest. The models were a global-support habitat suitability (GHS) model, a global-support environmental favorability (GEF) model, a local-support habitat suitability (LHS) model, and a local-support environmental favorability (LEF) model. The models produced raster maps for their predictions, and several statistical tests and evaluations were performed with the samples from each map to find the differences and predictive abilities of each model. The Friedman's rank sum test followed by Conover's *post hoc* test with Bonferroni correction verified each model is significantly different ($p \ll 0.01$) in the distribution of its predictions. Comparisons with Cochran's Q test revealed that predictive abilities at least one map was different from the others ($p \ll 0.01$). A pairwise McNemar's *post hoc* test suggested the GHS model is different from the rest. Each prediction map was validated using ground-truth data from field surveys, and a χ^2 independence test revealed that GHS model was (just barely) independent to actual bentgrass presence/absence ($p = 0.076$) but other three models were not independent ($p < 0.01$) to ground-truth data. (Here, independence is undesirable. It means that the model's predictions are not correlated with ground-truth data.) Their accuracies are similar at around 67% for all four models. However, because there were

different numbers of field plots where bentgrass was present (18/86) and absent (68/86), these accuracies are unbalanced. A calibration to compensate, the balanced accuracy of GHS model dropped to 61%, but balanced accuracy of other three maps raised to around 71%. The GEF, LHS, and LEF models have much higher sensitivity values (around 78% for all) than GHSPM (50%). However, GHS (72%) has slightly higher specificity than other three models (around 65%). Positive predictive values are similar for all four prediction maps (from 32% to 38%), and GEF, LHS, and LEF models have much higher negative predictive values (around 91%) than the GHS model (84%). These results show that GEF model, LHS model, and LEF model have stronger predictive ability than the GHS model. However, the predictive ability of GEF model, LSH model, and LEF model appear to be identical.

4.1 Introduction

4.1.1 Overview

Global navigation satellite system (GNSS) technology makes it easy to collect precise positioning data everywhere on Earth's surface unless there is a barrier to signal propagation. With the precise GNSS data and a geographic information system (GIS), researchers can perform map-based precise ecological modeling. Ahrens, Chung, Meyer, and Auer (2011) used a habitat suitability model and an environmental favorability model to create two maps that predict the presence/absence of bentgrass species based on field data collected by GNSS positioning technology; their research was a part of predictive ecological risk assessment study for creeping bentgrass (*Agrostis stolonifera*) engineered for herbicide resistance. **Habitat suitability** is the probability of finding the creature being studied at a certain place. An **environmental favorability model** is a habitat-suitability model that has accounted for any imbalance in the

presence/absence data. Habitat suitability maps can be used to inform conservation planning, ecological assessment, impact study of a species management, and identification of ecological steps on species habitat (Burgman, Breininger, Duncan, & Ferson, 2001; Gray, Cameron, & Kirkham, 1996; Rand & Newman, 1998; Van Horne & Wiens, 1991). Creeping bentgrass is a perennial routinely planted in golf courses, and a transgenic, glyphosate-resistant creeping bentgrass was created by commercial laboratories because it might offer cost benefits and the simplification of landscape management. These models and maps were built to help understand whether bentgrass dispersion was possible from source (a golf course) to the surrounding feral population, and to see if forests, lakes, and wetlands surrounding the golf course can to be a natural barrier to this dispersion. Both of the predictive maps were created using statistical global-support logistic regression models. (**Global support** means that a single-equation modeled whole study area; *i.e.* the entire data set was used to create the models.)

There are also **local-support** models. Local-support models partition the data set into neighborhoods, and each neighborhood is fitted with its own prediction equation. Local-support models necessarily have finer spatial resolution than global-support models so there is an expectation that they could be more accurate. Conversely, as with all statistical models, each individual per-point local-support model is built from fewer data than a global-support model so there is likewise an expectation that they could be less accurate. These notions interact in complicated, nonlinear ways, and this suggests it would be valuable to explore which of these approaches was best for this data set, which is the subject of the present study.

The data used in this study are identical to Ahrens *et al*. (2011). However, the goals of this study are different than theirs, namely, the comparison of models vs. making an ecological assessment. There are four bentgrass habitat models in this study: global-support habitat

suitability (GHS) model, global-support environmental favorability (GEF) model, local-support habitat suitability (LHS) model, and local-support environmental favorability (LEF) model. Statistical analyses are needed to compare these models, so the presence/absence data were randomly split into two portions: a calibration set (70%) and a validation set (30%) (Wang, Zhang, & Li, 2012; Wang, Zhang, & Li, 2013; Wang, Zhang, Li, Lin, & Zhang, 2014). The validation set was used to quantify the predictive ability of each model. Initially, four probability maps were built from the statistical models based on logistic regression analyses: global-support habitat suitability map (GHSM), global-support environmental favorability map (GEFM), local-support habitat suitability map (LHSM), and local-support environmental favorability map (LEFM). These four probability maps were reclassified as four predictive maps in dichotomous scale (presence/absence of bentgrass) by receiver operating characteristic (ROC) analysis: global-support habitat suitability prediction map (GHSPM), global-support environmental favorability prediction map (GEFPM), local-support habitat suitability prediction map (LHSPM), and local-support environmental favorability prediction map (LEFPM). There were two approaches to compare the four habitat models: (1) compare differences among the four statistical models, and (2) compare how well each model predicted bentgrass presence/absence by comparison with the ground-truth data.

4.1.2 Background

4.1.2.1 Logistic regression

Logistic regression, developed by David Cox in 1958 (Cox, 1958; Walker & Duncan, 1967), is a non-linear regression whose dependent variable is of dichotomous scale (Allison,

2001; Hosmer, Lemeshow & Sturdivant, 2013; Kleinbaum & Klein, 2006). The independent variables can be any of ratio, interval, ordinal, and categorical (nominal) scale (Hosmer *et al*., 2013).

Logistic regression fits a sigmoidal curve, or logistic function, between the values of the dichotomous dependent variable – values of zero and one only – to the values of the independent variables. The logistic function is derived in terms of **odds**, which is the ratio of the probability of presence and the probability of absence:

$$odds = \frac{\text{Probabiblity of presence}}{\text{Probabiblity of absence}} = \frac{Pr}{1 - Pr} \ .$$

The series for $\frac{x}{1-x} = \sum_{i=1}^{\infty} x^i$ for $0 \leq x < 1$, and the series for $e^x = \sum_{i=0}^{\infty}(x^i/i!)$, and these are the same series apart from the constants. Therefore, it is natural that logistic regression supposes that

$$\frac{Pr}{1 - Pr} = e^z, \tag{4.1}$$

where z is a linear polynomial in the independent variables: $z = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k$. Solving (4.1) for z yields

$$z = \ln(\frac{Pr}{1 - Pr}) \ , \tag{4.2}$$

so z is the natural logarithm of odds, the log-odds, called a **logit.** Solving this equation for probability Pr yields the logistic function

$$Pr = \frac{1}{1 + e^{-z}} \ . \tag{4.3}$$

The values of z's parameters, the b_i, are determined by the maximum likelihood method. So the probability of presence is value of (4.3) using z's parameters multiplying the values of the independent variables.

In some cases, an independent variable of a logistic regression model can be a dichotomous, nominal, and ordinal scales. The contrast coding system is a common approach for representing the categorical-scaled variables (Institute for Digital Research and Education, UCLA, 2006). A variable in a nominal or ordinal scale has several discrete values; therefore, it cannot be expressed with just one variable in the regression model. Instead, using multiple (so-called) dummy variables, one for each discrete value, makes it possible to express categorical-scaled variables with set of simple surrogate sub-variables. In this study the dummy coding consists of sub-variables coded with zeros and ones (Allison, 2001).

4.1.2.2 Habitat suitability model and environmental favorability model

Habitat suitability models date back to early 1980s when the habitat suitability index (HSI) model was proposed by U.S. Fish and Wildlife Service (1980, 1981) for evaluating wildlife habitat and the effects of the management activities, and they have been widely applied such as Elith (2000), Franco, Brito, and Almeida (2000), Larson, Dijak, Thompson, and Millspaugh (2003), Madsen and Prang (2001), Pearce and Ferrier (2000), and Pereira and Itami (1991). The habitat suitability model is a statistical model that estimates the probability to find a species based on the values of the independent variables based on functional relationships between the species and ecological variables suitable for habitat representations, usually indexed with 0–1 scale (Larson *et al*., 2003). The index value 0 represents a place that is not appropriate as a habitat, and the value 1 represents a place that is the most suitable habitat.

There is a drawback of using logistic regression for habitat suitability modeling. Logistic regression pre-supposes the ratios of the presences and absences are identical, and its results are sensitive that ratio (Hosmer *et al*., 2013). Real, Barbosa, and Vargas (2006) proposed a modified habitat suitability model they called an environmental favorability model that accounts for this unbalance as seen in (4.4). Their model is equivalent to (4.3) except the denominator's sum involves the ratio of presences and absences as

$$Fav = \frac{e^z}{\frac{N_1}{N_0} + e^z} \quad , \tag{4.4}$$

where *Fav* is the probability of bentgrass presence, N_0 is the number of study plots without bentgrass, N_1 is the number of study plots with bentgrass, and z is as above.

4.1.2.3 Global- and local-support models, and geographically-weighted logistic regression

A global-support model is a single-equation model that covers the entire study area and that does not incorporate the spatial relationships between neighbor data points. This model is suitable for identifying the pattern of an entire region, and it emphasizes the area's similarities (Fotheringham, Brunsdon, & Charlton, 2003). For example, the habitat suitability map and environmental favorability map created by Ahrens *et al*. (2011) are based on global-support models.

Tobler's law is, "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). By conceptualizing spatial data as realizations of an underlying random process, we can rephrase Tobler's law to be suggesting that spatial data are usually auto-correlated, so the values in nearby locations tend to be similar. If the variance of the

process is constant across the landscape, then the process is **stationary**. Conversely, if the variance of the process varies across the landscape, then the process is **non-stationary**. Global-support models have a better conceptual fit with a stationarity assumption. Local-support models were proposed to deal with spatial-nonstationary processes. Longley and Batty (1996) defined a local-support model to be a statistical model that covers only a portion of a study area. If the non-stationarity varies slowly over space, then it seems reasonable that constructing a model over a small region that is nearly locally stationary should provide a better estimate than fitting a single model over an entire region. Thus, each sample point has its own equation with parameter values, a.k.a. local statistics, that reflect the point's spatial relationships to the neighbor data points (Fotheringham & Wegener, 1999; Fotheringham, Brunsdon, and Charlton, 2003). These equations have the same variables as a global model but different parameter values. Emphasizing differences across space, the local-support model is more proper to investigate the variation throughout the study area, and the model is often used to find exceptions or hot-spots on the map (Fotheringham, Brunsdon & Charlton, 2003). This study employs geographically weighted logistic regression for its local-support models.

The models' parameter values are affected by the values of the neighbors, and it is reasonable that the values of nearby neighbors should be given more weight than the values of more distant neighbors. A **spatial weighting function** is a rule to provide the weightings given the distance between the point-of-interest and a neighbor point. The weighting scheme is called a **kernel**, and the kernel's decay rate over space is called its **bandwidth**. Various kernel functions have been proposed such as Gaussian, exponential, box-car, bi-square, and tri-cube (Gollini, Lu, Charlton, Brunsdon, & Harris, 2013). The Gaussian kernel embodies Tobler's law by providing a

continuously decreasing weighting with separation. Its equation is a re-scaling of the famous Gaussian distribution probability density function (PDF):

$$wgt = e^{-\frac{1}{2}(\frac{dist}{bw})^2}, \tag{4.5}$$

where wgt is the weight factor for the kernel, e is the base of the natural logarithm, $dist$ is the (non-negative, finite) distance between a point-of-interest and a neighbor point, and bw is the bandwidth, which is a surrogate for variance in the Gaussian PDF (Gollini *et al*., 2013). The bandwidth is a constant if and only if the (unknown) spatial process is stationary.

There is a separate interpretation of bandwidth apart from its interpretation as process variance. Bandwidth controls the size – spatial extent – of the kernel in some sense, so it also controls which points form a neighborhood. If the data were sampled on a regular grid, then the neighborhoods can be of fixed size and shape, and such are called **fixed kernels**. However, if the samples are irregular, fixed kernels can over-sample where samples are dense and under-sample where samples are sparse. **Adaptive kernels** vary the bandwidth to account for this. To find optimal bandwidths, computer automated procedures are applied based on leave-one-out cross-validation (CV) scores computed for all possible bandwidths, and the bandwidth with the smallest CV score is best (Harris, Brunsdon, & Charlton, 2011).

4.1.2.4 The receiver operating characteristic (ROC) analysis

The ROC analysis was developed as an object-detection technology for radar during World War II (Swets, 1973). Currently, ROC analysis is applied to radiology, psychology, medicine, biometrics, machine learning, data mining, other automated detection problems (Akobeng, 2007; Swets, 1988; Swets, Dawes, & Monahan, 2000; Zweig & Campbell, 1993). In many cases of

Table 4.1: Confusion matrix (2×2) for dichotomous outcomes.

	Actual Positive (1)	Actual Negative (0)	Column Total
Predicted Positive (1)	TP (n_{11})	FP (n_{10})	$TP + FP$ $(n_{1.})$
Predicted Negative (0)	FN (n_{01})	TN (n_{00})	$FN + TN$ $(n_{0.})$
Row Total	$TP + FN$ $(n_{.1})$	$FP + TN$ $(n_{.0})$	$TP + FP + FN + TN$ (n)

science research, it is necessary to discriminate between two options such as yes or no, good and bad, presence or absence, and male or female. A ROC analysis is for these dichotomous decision-making procedures. The ROC analysis can be defined as a quantitative model to estimate a binary classification from a continuous-scale dataset. The ROC analysis is composed of two steps. First, establish an optimal threshold or cut-off value; and second, classify continuous data into a binary scale with the optimal threshold. Table 4.1 shows a common 2×2 confusion matrix for a class containing positive and negative vs. a predictive class which is classified as positive and negative by a certain test (Fawcett, 2006). The correctly identified positive is called a true positive (TP), and the incorrectly identified positive, i.e., when it actually is negative, is called a false positive (FP), a.k.a. "Type I error". The correctly identified negative is called a true negative (TN), and the incorrectly identified negative is called a false negative (FN), a.k.a. "Type II error". The ratio of TP to actual positives is called sensitivity, and the ratio of TN to actual negatives is called specificity. The equations for $Sensitivity$, $Specificity$, and 1 – $Specificity$ are shown in (4.6) – (4.8).

$$Sensitivity = \frac{TP}{TP + FN} \tag{4.6}$$

$$Specificity = \frac{TN}{TN + FP} \tag{4.7}$$

$$1 - Specificity = \frac{FP}{TN + FP} \tag{4.8}$$

The best cutoff point maximizes both true positive and true negative determinations. However, the relationship of sensitivity and specificity is compensational, so if sensitivity is increased, specificity is decreased, and vice versa. When sensitivity is fixed to a certain value, specificity is also fixed according to the fixed sensitivity value, resulting in a sensitivity-specificity pair. There are many choices determining cutoff points, but maximizing $Sensitivity + Specificity$ has been applied to find best cutoff value. It is identical to maximize Yuden's index in (4.9). This is because the maximize specificity and minimizing 1 – $Specificity$ which is false positive rate are same concept.

$$Yuden's\ Index = Sensitivity\ + Specificity - 1 \tag{4.9}$$

4.1.2.5 Friedman's rank sum test and Cochran's Q test with pairwise *post-hoc* test

Friedman's rank sum test is a non-parametric multiple comparison test for analyzing randomized complete block designs. This test is a non-parametric counterpart of the one-way repeated-measure ANOVA. The purpose of the Friedman's rank sum test is to see differences between multiple treatments. There are several assumptions and precautions for Friedman's rank sum test. First, a group should be randomly sampled from the population. Second, this sampled group should be measured on three or more occasions or by different methods; these are called treatments. Third, a dependent variable should be in ordinal or continuous scale, but the samples do not need to be normally distributed or to follow the assumptions for one-way repeated-

measure ANOVA. As a non-parametric test, Friedman's rank sum test performs rank transformation within a block. A block of observations is the observations from a particular individual in the sampled group who has been observed multiple times. In this study, the "individuals" forming the blocks are pixels in the maps. The null hypothesis of the Friedman's test is: the populations represented by each method have the same distribution, which can be understood to mean there is no difference between each method. In this study a method is one of the ways of constructing a map. When the result of Friedman's rank sum test shows there is a difference, a *post hoc* test should be followed to find which method is different, and either a Conover's Test (Conover & Iman, 1979; Conover, 1999) or a pairwise Wilcoxon's signed rank test is widely applied. When pairwise tests are performed *post hoc*, a general alpha adjustment for multiple comparisons might be applied to adjust the p-values to reduce the chance of increasing incorrectly rejecting null hypothesis (type I error). The Bonferroni correction is one of frequently using methods of alpha adjustment for multiple comparisons (or adjustment for multiplicity, in short), and this correction is used in this study. In this study, Friedman's rank sum test is the method to test the differences of distributions between four probability maps created by each habitat models.

Cochran's Q test is a non-parametric test to verify whether multiple treatments have identical effects. The data structure of Cochran's Q test is very similar to that of Friedman's rank sum test, which is a randomized complete block design. However, the dependent variable of Cochran's Q test is in dichotomous scale (coded as 0 and 1). The null hypothesis of Cochran's Q test is that the proportion of successes (coded as 1) is equal for all treatments. However, the hypothesis can be described more simply as that there are no differences between the treatments (Sheskin, 2011), or the treatments are equally effective. When the test result shows there are

differences in the proportion of successes, a *post hoc* test should be used. A pairwise McNemar's test is usually applied for the *post hoc* test after Cochran's test. A McNemar's test (McNemar, 1947) is identical to Cochran's Q test when there are only two treatments. Because pairwise tests are performed as *post hoc*, the p-values of each pairwise test should be adjusted, and the Bonferroni correction is chosen. In this study, Cochran's Q test is the method to test the differences of predictive ability between four prediction maps in binary scale representing each habitat models.

4.1.2.6 χ^2 test, measure of association, and binary classification analysis

A χ^2 test of independence is a statistical test between two categorical variables from a single population to test whether the variables are independent. The χ^2 test has a null hypothesis that the two categorical variables are independent. When this null hypothesis is rejected, then the two categorical variables are dependent but there is no indication how correlated they might be.

The test statistic is given in (4.10), where O_i is an element of the set of observations and E_i is an element of the set of expectations. When this statistic is close to zero, the differences between sum of observed values and expected values are small. In this case, the chance that the two variables are independent is increased.

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \tag{4.10}$$

Measures of association or association coefficients are indices that measure the strength of a relationship between two categorical variables. Yule's Q and the Φ coefficient are often applied to measure the association strength between two possibly related dichotomous variables. Yule's Q is a measure of association to compare two dichotomous variables, and this index is a

transformation of odds ratio, taking values between –1 and +1. In a 2×2 cross table of counts, the pair of matching observations are called concordant, and the other pair are called discordant. Yule's Q converges to positive one or negative one when the concordant or discordant counts become extremely different. Knoke, Bohrnstedt, and Mee (2002) mentioned a rule-of-thumb to interpret Yule's Q as follows: virtually no relationship (0 – ±0.25), weak relationship (±0.25 – ±0.50), moderate relationship (±0.50 – ±0.75), and strong relationship (±0.75 – ±1.00). Yule's Q can be calculated with (4.11), where n_{11} is the number of TP, n_{00} is the number of TN, n_{10} is the number of FP, and n_{01} is the number of FN (Table 4.1).

$$Q = \frac{n_{11}n_{00} - n_{10}n_{01}}{n_{11}n_{00} + n_{10}n_{01}} \tag{4.11}$$

The Φ coefficient, introduced by Karl Pearson, is also a measure of the association degree between two dichotomous variables. The range of the Φ coefficient is also from –1 to 1, and zero means no association. The interpretation of Φ coefficient is quite similar to the Pearson's correlation coefficient. Similar to Yule's Q, when the concordant or discordant counts become extreme, the Φ coefficient converges to +1 or -1. The interpretation of Φ coefficient suggested by Davis (1971) are as follows: no relationship (0.00), negligible relationship (±0.01 – ±0.09), low relationship (±0.10 – ±0.29), moderate relationship (±0.30 – ±0.49), substantial relationship (±0.50 – ±0.69), and very strong relationship (±0.70 – ±1.00). There is a special need for caution interpreting Yule's Q and Φ when the two values are calculated simultaneously because Yule's Q generally has a higher absolute value than Φ coefficient. The equation to calculate the Φ coefficient is (4.12). where n_{11} is the number of TP, n_{00} is the number of TN, n_{10} is the number of FP, n_{01} is the number of FN, $n_{1.}$ is $n_{10} + n_{11}$, $n_{0.}$ is $n_{00} + n_{01}$, $n_{.0}$ is $n_{00} + n_{10}$, and $n_{.1}$ is $n_{01} + n_{11}$ (Table 4.1).

$$\Phi = \frac{n_{11}n_{00} - n_{10}n_{01}}{\sqrt{n_{1.}\, n_{0.}\, n_{.0}\, n_{.1}}} \tag{4.12}$$

When a predictive model has been built, a predictive-ability evaluation of this model is necessary. A binary classification analysis is an evaluation method for predictive ability of the model with dichotomous results. The purpose of this binary classification analysis is to compare the predictive-model result to the actual data or equivalent and quantify this predictive ability. For the predictive-ability evaluation, a new dataset that is not used for the predictive-model building is required. One of the strategies for this evaluation is to use a data subset before the model building. One data subset is used for predictive model building, and the other data subset is used for evaluation. After the predictive values are calculated from the model, a 2×2 confusion matrix, a special form of contingency table is built. Some of important parameters can be derived from the values of this confusion matrix including prevalence, detection prevalence, accuracy, balanced accuracy, sensitivity (a.k.a. true positive rate (TPR)), specificity (a.k.a. true negative rate (TNR)), positive predictive value (PPV), and negative predictive value (NPV).

The definitions of these parameters are as follows. Prevalence is a proportion of actual positives from total observations. Equation (4.13) shows the prevalence, where TP is true positive, FN is false negative, FP is false positive, and FN is true negative.

$$Prevalence = \frac{TP + FN}{TP + FP + FN + TN} \tag{4.13}$$

Detection Prevalence is a proportion of correctly identified positives from total observations. Equation (4.14) shows the detection prevalence, where TP is true positive, FP is false positive, FN is false negative, and TN is true negative.

$$Detection\ Prevalence = \frac{TP + FP}{TP + FP + FN + TN} \tag{4.14}$$

Accuracy is a ratio of correctly classified values regardless of true or false from total observations. Equation (4.15) represents the accuracy, where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{4.15}$$

Balanced accuracy is mean of sensitivity and specificity. Balanced accuracy is a calibrated accuracy for the data with unbalanced positive/negative ratio. When the number of positive and negative classes are equal, balanced accuracy has same value as accuracy. Equation (4.16) shows how to calculate balanced accuracy.

$$Balanced\ Accuracy = \frac{Sensitivity + Specificity}{2} \tag{4.16}$$

Sensitivity (a.k.a. true positive rate, TPR) is a proportion of correctly identified positives from all the actual positives. Equation (4.6) above shows the calculation of sensitivity. Specificity (a.k.a. true positive rate, TPR) is a proportion of correctly identified negatives from all the actual negatives. Equation (4.7) above shows the calculation of specificity.

Positive Predictive Value (PPV) (a.k.a. precision) is a rate of the actual positives from correctly classified positives. Equation (4.17) shows the PPV, where TP is true positive, and FP is false positive.

$$PPV = \frac{TP}{TP + FP} \tag{4.17}$$

Negative Predictive Value (NPV) is a rate of the actual negatives from correctly classified negatives. Equation (4.18) shows the NPV, where TN is true negative, and FN is false negative.

$$NPV = \frac{TN}{TN + FN} \tag{4.18}$$

4.1.3 Literature review

Ecologists use spatially-explicit maps to help assess the potential ecological impact of genetically modified (GM) crops (Marvier *et al*., 2008). There has also been increased interest in prediction models on the distribution of species and their habitats from species presence/absence data, relevant ecological variables, mathematical methods, and GIS (Hirzel, Le Lay, Helfer, Randin, & Guisan, 2006). This approach has been applied to the distribution models of invasive plants (Evangelista *et al*., 2008; Jacobs & Macisaac, 2009; Nielsen, Hartvig, & Kollmann, 2008; Sánchez-Flores, Rodriguez-Gallegos, & Yool, 2008), and as an identification model of critical habitat for endangered animals (Bellis *et al*., 2008; Glenz, Massolo, Kuonen, & Schlaepfer, 2001; Pereira & Itami, 1991). Ahrens *et al*. (2011) used spatially explicit logistic regression to produce maps of habitat suitability and environmental favorability for creeping bentgrass, which already exists as a GMO: it has been made glyphosate resistant by genetic modification.

Ahrens *et al*. (2011) produced two maps but this begs the question whether these are the best maps possible given their data set. Visser and de Nijs (2006) suggested that analysts comparing maps can detect temporal and spatial changes (so-called hot-spots), compare different models, methods or scenarios, calibrate or validate land-use models, analyze uncertainty and sensitivity of a model, and assess map accuracy. Dorren and Seijmonsbergen (2003) compared models and maps predicting rockfall runout zones at a regional scale. They constructed a

contingency table to compare the prediction maps created with the patterns of rockfall deposits derived from pre-existing geomorphological field maps. Ayalew, Yamagishi, Marui, and Kanno (2005) created two GIS-based raster maps for landslide susceptibility in Sado Island, Japan; one using an analytical hierarchy process (AHP) and one with logistic regression. They reclassified the rasters into five classes and overlapped the maps with a landslide activity map, which revealed that 70% of the landslide zones fell into the high and very high susceptibility zones in the AHP map, and 63% for logistic regression map, so the AHP map was deemed the better predictor. Wang *et al*. (2012) compared two raster maps of the spatial distribution of soil organic matter created with geographically weighted regression and with kriging. Wang *et al*. (2013) compared two soil total nitrogen predictive maps created by geographically-weighted regression and with cokriging. They divided point-field data in into a training set and a validation set. With the training dataset, they created raster prediction maps with geographically weighted regression, regression kriging, and cokriging. The fieldwork-derived values in the validation point dataset and pixel values on prediction maps in the same locations were used to calculate comparison statistics such as mean error (ME) and root mean square error (RMSE) to determine which map has less average error. Searching the literature suggested there is a gap when the prediction maps are in categorical scale because the statistical analyses for continuous-scale values are not applicable for the maps with categorical property.

4.1.4 Research questions and hypotheses

Research question 1: Are the underlying probability distributions of the predicted habitat suitability values from the four habitat suitability and environmental favorability models identical? To answer this question, Friedman's rank sum test, which is the nonparametric

counterpart of repeated-measures analysis of variance (ANOVA), was applied with set of 200 random sampled from each probability maps in continuous scale.

Research question 2: Are the predictive abilities for bentgrass presence and absence of the four habitat suitability and environmental favorability models identical? To answer this question, Cochran's Q test is applied with set of 200 random sampled from each prediction maps in dichotomous scale.

Research question 3: Are the bentgrass predictions (present/absent) derived by each habitat suitability and environmental favorability model independent or dependent to the actual bentgrass present/absent distributions? If they are dependent, how closely associated are they? To answer this question, a χ^2 test is applied between ground-truth data of bentgrass presence/absence in the validation set (n = 86) and each prediction map in dichotomous scale. The association strength of actual bentgrass presence/absence and predicted values of each model is quantified using two measures of association, namely Yule's Q and the Φ coefficient.

Research question 4: Which of the bentgrass habitat models presents the best prediction result? To answer this question, binary classification analyses are applied between ground-truth data of bentgrass presence/absence in the validation set (n = 86) and each prediction maps in dichotomous scale. The tests are various ratios of the elements of a standard confusion matrix. These ratios are named prevalence, detection prevalence, accuracy, balanced accuracy, sensitivity, specificity, positive predictive value, and negative predictive value. The differences among these tests can reveal differences among the maps.

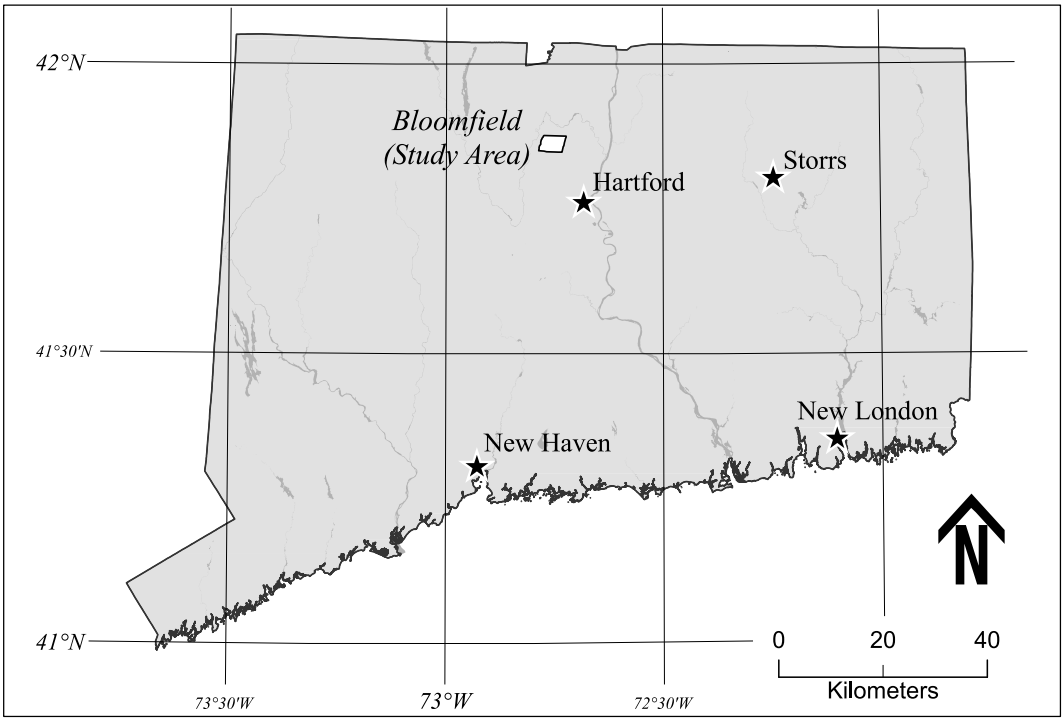
4.2 Material and Methods

4.2.1 Study location

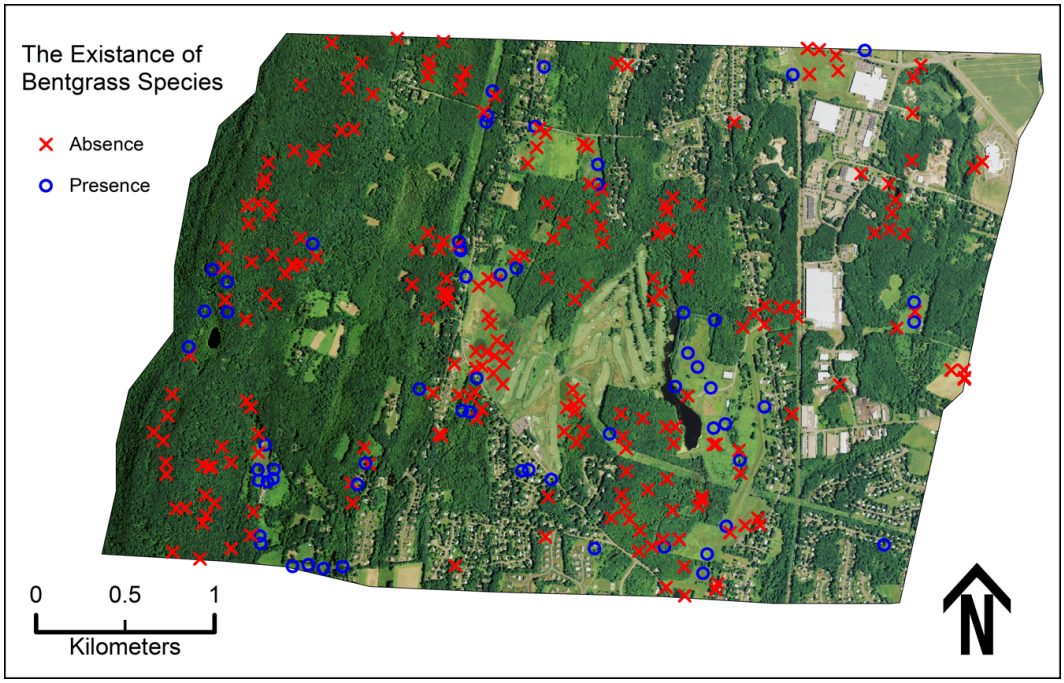
The study area (Figure 4.1) is in Bloomfield, Connecticut, centered at a golf course where bentgrass is planted for putting greens and fairways (golf course manager, personal communication, 2008). The study area covers 8.5 km^2 and is located at 72°45'19"W, 41°51'46"N (Ahrens *et al*., 2011). This area is a part of the Eastern Broadleaf Forest Province subsection of the Lower Connecticut River Valley (Ahrens *et al*., 2011; Metzler & Barrett, 2006). A lake is adjacent to the eastern side of the golf course, and a wetland lines the western side. The golf course is surrounded by a broadleaf forest. A state park covers the western edge of the study area and a power line right-of-way cuts through the study area generally north-to-south. There are recreation areas with managed grasslands, residential areas with lawns, light industry with turf grass, agricultural fields, and a railroad (Ahrens *et al*., 2011).

4.2.2 Field surveying and variable collection

Ahrens *et al*. (2011) studied the possibility of bentgrass escaping from the golf course and migrating elsewhere by creating maps showing bentgrass habitat suitability and environmental favorability. These maps were generated using statistical models created from presence/absence data for plants, standing water, and other variables (Table 4.2). The presence/absence bentgrass data were collected by *in situ* sampling of 289 survey plots from July to October, 2008. The survey plots' geographic coordinates were generated randomly, and the field team found the plots using a TOPCON HiperLite+ GNSS receiver with FC-100 data collector. Using RTK to find the plots that were located in open areas was straightforward; however, as described in



(a) Location of study area in statewide scale.



(b) National Agricultural Imagery Program (NAIP) 2008 aerial orthophoto in study area boundary. The locations for presence/absence of bentgrass are overlapped.

Figure 4.1: Location of the study area and field plots with bentgrass presence/absence information.

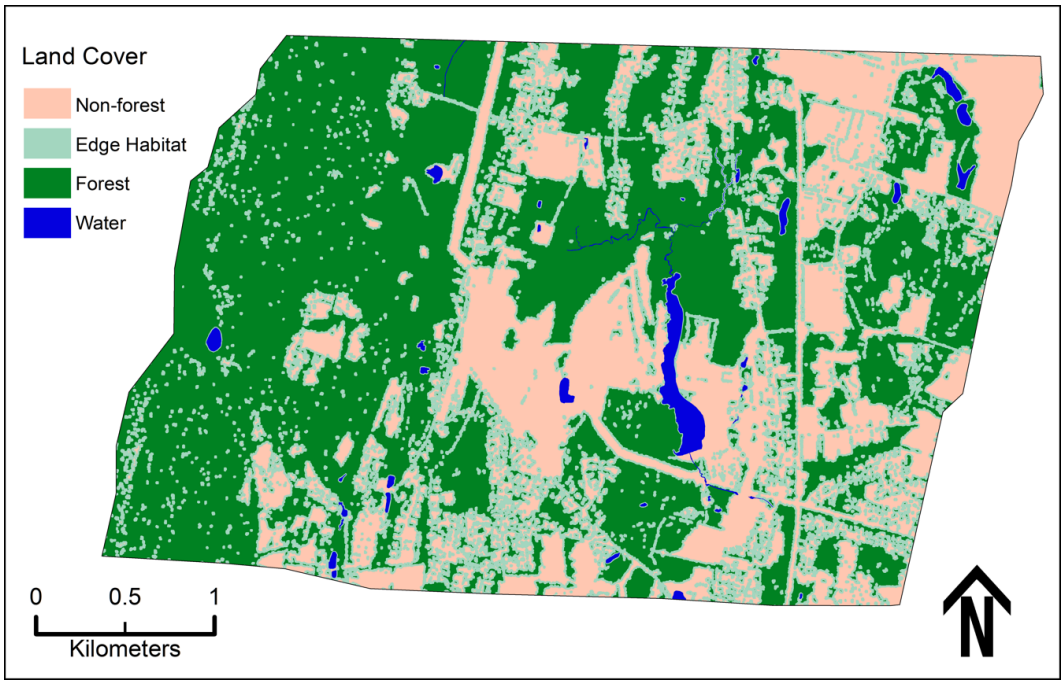
Chapter 3, RTK generally yields a fixed solution in canopy coverage up to around 74%. When a survey plot was beneath the tree canopy such that RTK became inoperable, the plot was found using an autonomous position. Then, the GNSS receiver was erected atop a range pole and set to collect static observations for at least 15 minutes, and the position of the survey plot was determined using post-processing. The survey was conducted inside a 5.64-m (100 m^2) radius circular plot centered at the GNSS receiver. See Ahrens *et al*. (2011) for full details.

The independent variables' values came from using GIS and remote sensing. We produced a land-cover map representing forest, non-forest, and edge habitat between forest and non-forest

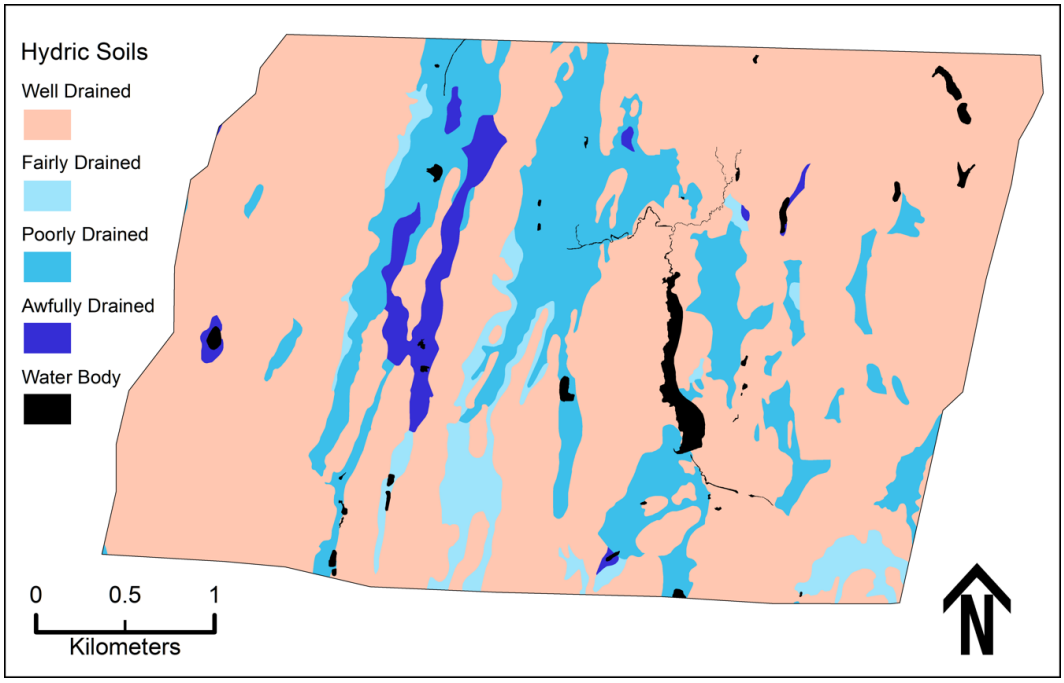
Table 4.2: Description of variables used in the multivariate logistic regression models and GIS mapping. The dependent variable is existence of bentgrass species in a plot, and all the others are independent variables. (Ahrens *et al*., 2011)

Name of the variable	Abbr.	Initial coding	Comments
Dependent Variable			
Existence of bentgrass species	*Bentgrass*	Absent, presence	Recoded as 0 and 1 respectively. Classified as categorical (nominal) factor.
Independent Variables:			
*Percent plot covered by tree canopy	*Tree*	0–25%, 25–50%, 50–75%, and 75–100%	Reclassified as non-forest (0–25%), edge habitat (25–50%), and forest (50–100%). Classified as a categorical (nominal) factor.
Percent plot covered by shrub canopy	*Shrub*	0–25%, 25–50%, 50–75%, and 75–100%	Recoded as I, II, III, and IV respectively. Classified as a nominal variable.
Percent plot covered by herbaceous plant species	*Herb*	0–25%, 25–50%, 50–75%, and 75–100%	Recoded as I, II, III, and IV respectively. Classified as a nominal variable.
Land management evidence by mowing	*Mow*	No evidence of mowing, mowed once per year, mowed more than once per year.	Recoded as I, II, and III respectively. Classified as a nominal variable.
*Land and soil water conditions	*Water*	Well drained and no water found, fairly drained with seasonal flooding, poorly drained with standing water, and awfully drained containing a water body.	Recoded as I, II, III, and IV respectively. Classified as a nominal variable.

*Dummy coding (treatment coding) was applied to categorical variables (*Tree* and *Water*) for GIS map creation with raster calculation.



(a) Forest-cover map digitized from NAIP 2008 aerial orthophoto.



(b) Hydric-soil map created from Natural Resources Conservation Service (NRCS) soil-type polygons joined to the Soil Survey Geographic (SSURGO) database for Connecticut.

Figure 4.2: The forest-cover map and the hydric-soil map.

zone (Figure 4.2 (a)) by digitizing a 2008 National Agricultural Imagery Program (NAIP) aerial orthophoto, shown as the background of Figure 4.3. The NAIP aerial orthophoto was comprised of four bands (red, green, blue, and near-infrared) with 3×3 ft^2 spatial resolution. A normalized difference vegetation index (NDVI) map was calculated from the NAIP orthophoto to make the forest zones more distinctive, and this NDVI map was segmented by *eCognition*, which is an object-based land-cover mapping tool. The forest-zone segments produced by *eCognition* were relatively rough, so the non-forest segments were removed manually using a graphic digitizing tablet. The calculation of NDVI is seen in (4.19).

A 10-meter buffer was created from the border of the forest-zone layer created from the NAIP aerial orthophoto, and the research team named this layer the edge-habitat zone. Any pixels that were neither forest nor edge-habitat were classified as non-forest. These three zones were merged into one raster map.

$$NDVI = \frac{NearInfrared - Red}{NearInfrared + Red} \tag{4.19}$$

The hydric-soil map (Figure 4.2 (b)) was created from a Natural Resources Conservation Service (NRCS) shapefile comprised of soil-type polygons joined to the Soil Survey Geographic (SSURGO) database for Connecticut, which provided drainage class for each soil type. The water bodies on the hydro-soil map were somewhat inaccurate, so they were corrected using the NAIP aerial orthophoto.

4.2.3 Model selection for habitat suitability model

The habitat suitability and environmental favorability models were created using spatially explicit logistic regression. Ahrens *et al*. (2011) used the Akaike Information Criterion (AIC)

Table 4.3: The ranking of the best multivariate logistic regression model to predict the presence of bentgrass species is presented with top 15 models without any interaction (Ahrens *et al*., 2011).

Rank	Synopsis of Models	*K*	*AIC*	Δ*AIC*	*w*	–2log(likelihood)
1	*Tree + Water + Herb + Herb × Tree*	7	237.6	0	5.21×10^{-2}	2.24×10^{2}
2	*Tree + Water + Herb*	6	237.7	< 0.1	5.10×10^{-2}	2.26×10^{2}
3	*Tree + Water + Herb + Shrub + Herb × Tree*	8	238.8	1.1	2.93×10^{-2}	2.23×10^{2}
4	*Tree + Water + Herb + Shrub + Herb × Shrub + Herb × Tree*	9	239.0	1.3	2.69×10^{-2}	2.21×10^{2}
5	*Tree + Water + Herb + Shrub + Herb × Shrub*	8	239.1	1.5	2.50×10^{-2}	2.23×10^{2}
6	***Tree + Water***	**5**	**239.2**	**1.5**	$\mathbf{2.42 \times 10^{-2}}$	$\mathbf{2.29 \times 10^{2}}$
7	*Tree + Water + Herb + Shrub*	7	239.2	1.6	2.35×10^{-2}	2.25×10^{2}
8	*Tree + Water + Mow + Tree × Mow*	9	239.5	1.8	2.09×10^{-2}	2.21×10^{2}
9	*Tree + Water + Mow + Herb + Tree × Mow*	10	239.6	2.0	1.95×10^{-2}	2.20×10^{2}
10	*Tree + Water + Herb + Shrub + Herb × Tree + Shrub × Tree*	9	239.7	2.1	1.85×10^{-2}	2.22×10^{2}

Abbreviations: *K*, number of estimable parameters in a model; *AIC*, Akaike's Information Criteria; *w*, Akaike weights.

(Akaike, 1974) to select the best model. The models were sorted by *AIC* score, and the difference between smallest AIC and each AIC (ΔAIC) was calculated. The model with lower ΔAIC has more power, and models with ΔAIC less than 2 are usually considered as the best choice according to the rule of thumb (Burnham & Anderson, 2002). Ten of tested models had ΔAIC under 2 (Table 4.3), and the 6th-ranked model (*Bentgrass = Tree + Water*) was the only model possible for GIS digital mapping, so that the 6th-ranked model was used for digital mapping of this study. See Ahrens *et al*. (2011) for full details about the data and the modeling.

4.2.4 Data preparation for statistical modeling and mapping

Comparing the predictive ability of the models requires that the presence/absence dataset be randomly split into a calibration set (70% present and 70% absent = 203 plots) and a validation set (30% present and 30% absent = 86 plots) (Table 4.4 and Figure 4.3). The calibration set was used to create the models and maps used here. The validation set was used for statistical analyses for prediction-ability assessments by comparing the map values derived from each model to actual presence and absence of bentgrass collected from the fieldwork.

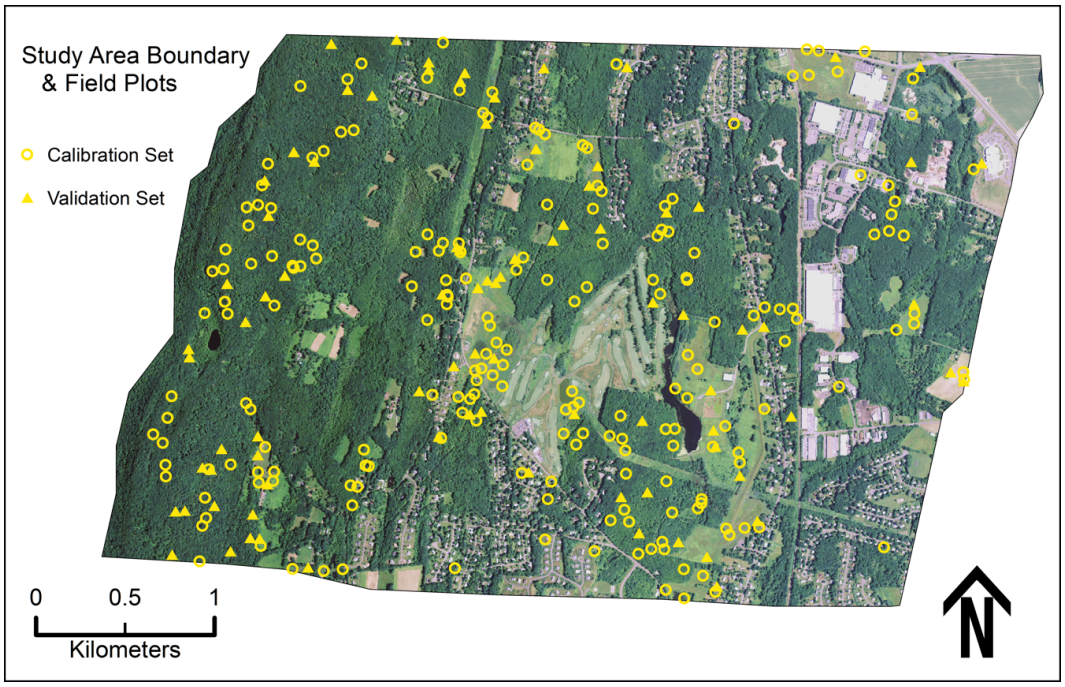


Figure 4.3: Locations of calibration and validation sets.

Table 4.4: Splitting field data into calibration and validation sets.

	Present	Absent	Total
Calibration Set (70%)	43	160	203
Validation Set (30%)	18	68	86
Total	61	228	289

Comparing the four models to one another does not involve the presence/absence data because the question is only whether the models produce the same results or not. Nonparametric tests generally have weaker statistical power than a parametric test, so a power analysis using *G*Power* (Faul, Erdfelder, Buchner, & Lang, 2009) recommended that 200 locations should randomly selected for the comparisons among four maps (Lehmann & D'Abrera, 2006).

4.2.5 Statistical Modeling

4.2.5.1 Habitat suitability model and environmental favorability model with global support

The habitat suitability map and environmental favorability map created by Ahrens *et al*. (2011) are global-support models. Those models were re-created here using identical methodology but with different data: only 70% of the data (calibration set) were used to create the models. These data produced the following habitat suitability model of bentgrass, (4.20) – (4.21).

$$Z_G = 0.4304 + 0.2018 \times Tree_1 - 2.4614 \times Tree_2 \\ -1.6756 \times Water_1 - 1.3508 \times Water_2 - 15.1477 \times Water_3 \quad (4.20)$$

$$HS = \frac{e^{Z_G}}{1 + e^{Z_G}} \quad (4.21)$$

where e is the base of the natural logarithm, Z_G is a logit (or log-odds) of the global habitat suitability model, and HS is habitat suitability.

The environmental favorability model with global support is (4.22), where EF is the environmental favorability, $N_0 = 160$ is the number of study plots without bentgrass, $N_1 = 43$ is the number of study plots with bentgrass species.

$$EF = \frac{e^{Z_G}}{\frac{N_1}{N_0} + e^{Z_G}} = \frac{e^{Z_G}}{\frac{43}{160} + e^{Z_G}} \quad (4.22)$$

4.2.5.2 Habitat suitability model and environmental favorability model with local-support, geographically weighted logistic regression

The local-support models for this study were geographically weighted logistic regressions with adaptive bandwidth and a Gaussian kernel. The local-support models were computed using

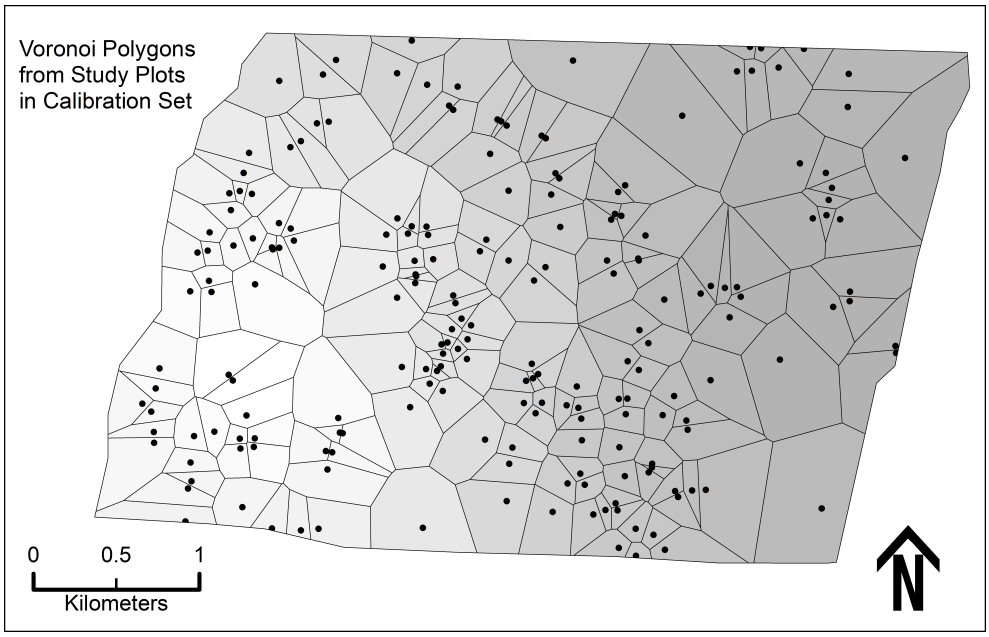


Figure 4.4: Voronoi polygons used to discretize the study area to define the domains of the local-support models. The total number of the polygons is 203.

the *R*-statistical software system with external packages *sp* and *GWmodel* (Gollini *et al*. 2013; Pebesma & Bivand, 2005; Bivand, Pebesma, & Gómez-Rubio, 2013). There are *n* local-support models. The local-support logit $Z_{L,i}$ is shown in (4.23) in which each term includes a subscript $1 \leq i \leq n$ that indicates which equation each term belongs to. Each local-support model has a set of estimated parameters: a constant ($Const_i$), two for trees variables ($CoefTr_{1,i}$, $CoefTr_{2,i}$), and three for water variables ($CoefWtr_{1,i}$, $CoefWtr_{2,i}$, $CoefWtr_{3,i}$). The independent-variable values, $Tree_1$ and $Tree_2$, $Water_1$, $Water_2$, and $Water_3$, came from the forest-cover map and hydric-soil map described above.

$$Z_{L,i} = Const_i + CoefTr_{1,i} \times Tree_{1,i} + CoefTr_{2,i} \times Tree_{2,i} + CoefWtr_{1,i} \times Water_{1,i} + CoefWtr_{2,i} \times Water_{2,i} + CoefWtr_{3,i} \times Water_{3,i} \quad (4.23)$$

Each local model $Z_{L,i}$ is applicable to some zone near a sampling point. We chose to use Voronoi polygons for the zones (Figure 4.4).

4.2.6 Probability maps for bentgrass habitat suitability

The four models produced probability raster maps (GHSM, GEFM, LHSM, LEFM) using the *ArcGIS* raster calculator. Each raster pixel's value is the probability to find bentgrass species at that pixel. The GHSM was calculated with the raster of (4.20) as the input to (4.21); GEFM was calculated with the raster of (4.20) as the input to (4.22); LHSM was calculated with the raster of (4.23) as the input to (4.21); and LEFM was calculated with the raster of (4.23) as the input to (4.22).

4.2.7 Prediction maps for bentgrass in dichotomous scale

For statistical tests to verify the predictive ability of each statistical model, the four probability maps (GHSM, GEFM, LHSM, and LEFM) had to be reclassified as dichotomous maps representing presence and absence of bentgrass. For this process cutoff values of habitat suitability were required. ROC analysis is a method to find cutoff values when a variable is dichotomous, and ROC analyses for each of the maps was performed with the calibration set. The cutoff values of GHSM, GEFM, LHSM, and LEFM were 0.1226, 0.3350, 0.1275, and 0.3420 respectively. When a habitat suitability value of a raster pixel was greater/less than the cutoff value, this pixel was reclassified as 1/0 for presence/absence of bentgrass. The results of this reclassification are four prediction maps (GHSPM, GEFPM, LHSPM, and LEFPM) representing four habitat models respectively. The results are shown in Figure 4.5 and Figure 4.6.

4.2.8 Map comparison among the maps from four modeling methods

The over-arching goal of this study is to determine whether any of these models, or groups of models, produce better predictions than the others. We explored this using two approaches. The first approach is to find differences among the four maps without ground-truth data – i.e., how similar are the predictions without considering their accuracy. For this approach, the four maps were compared at 200 random sample points in two ways. First, were the four maps coming from the same population distribution? Second, were the prediction performances of the four maps identical? Two nonparametric tests for multiple comparisons with a complete-block-design were appropriate. The first question can be answered by Friedman's rank sum test, and the second question can be answered by Cochran's Q test.

The null hypothesis of Friedman's rank sum test is that the tested distributions among the probability maps (GHSM, GEFM, LHSM, and LEFM) are the same, and the alternate hypothesis is that they are different. If the null hypothesis is rejected, a *post hoc* Conover's test with Bonferroni corrections is applied to determine which sample set is different (Conover & Iman, 1979; Conover, 1999). These results were checked using a pairwise Wilcoxon signed rank test.

Cochran's Q tests whether the percentage-of-presence prediction maps (GHSPM, GEFPM, LHSPM, and LEFPM) are the same. If the percentages are not different, then it is reasonable to conclude that the groups' predictive abilities are the same. Cochran's Q test requires that input variables be in the binary scale, so the bentgrass probability maps were reclassified to be dichotomous (present/absent). The pairwise McNemar's test and the McNemar's exact test (a.k.a. binomial test) were chosen for *post hoc* tests after Cochran's Q test. These tests were performed by *R*-statistical package and checked using *SPSS*.

4.2.9 Prediction-ability assessment using ground-truth data

The second approach is to find differences among the four maps against ground-truth data – i.e., how similar are the predictions regarding their accuracy. This approach was considered three ways. First, build four 2×2 contingency tables (GHSPM × validation, GEFPM × validation, LHSPM × validation, and LEFPM × validation) and apply four χ^2 independence tests to see whether the predictions of each maps are independent to actual presence/absence of bentgrass in the field. Here, "validation" means the validation data set comprised of the ground-truth data. In this case, an "independent" result is undesirable: it means the predictions bare no strong resemblance to the ground-truth data.

A χ^2 test only provides whether the predictions and the ground-truth data are independent, but not the degree and direction of the relationship. The relationships can be quantified with Yule's Q and with a Φ coefficent, which are two of measures of association for dichotomous variables. Yule's Q and Φ range from –1 to 1, and the meaning of 0 is no association between two variables, and ±1 is the maximum strength of association between two variables. The interpretation of Yule's Q (Warrens, 2008; Yule, 1912) and a Φ coefficient (Cramer, 1946) are provided by Knoke, Bohrnstedt, and Mee (2002), Davis (1971) and Davenport and El-Sanhurry (1991).

We concluded with a binary classification analysis for each of the prediction map versus actual bentgrass existence. The test statistics for binary classification analysis were prevalence, detection prevalence, accuracy, balanced accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).

4.3 Results

4.3.1 Created maps and their statistics

The four maps in Figure 4.5 are probability maps of bentgrass habitat suitability created by (a) GHSM, (b) GEFM, (c) LHSM, and (d) LEFM. The maps are in continuous scale from 0 to 1 because the habitat suitability on each of the pixel value is the probability of finding bentgrass species at that location. Although the cell values are continuous scale, the number of unique values of each map is restricted because the independent variables used in the mapping models ($Tree$ and $Water$) are in a categorical scale.

The prediction maps (GHSPM, GEFPM, LHSPM, and LEFPM) in Figure 4.6 are the reclassified maps from the probability maps (GHSM, GEFM, LHSM, and LEFM) in Figure 4.5 using cutoff values derived by ROC analysis. The cutoff value for GHSM, GEFM, LHSM, and LEFM are 0.1226, 0.3350, 0.1275, and 0.3420 respectively.

4.3.2 Comparisons between the maps

4.3.2.1 Friedman's rank sum test

From Friedman's rank sum test (n = 200, $\chi^2(3)$ = 479.21, and $p \ll 0.01$), the null hypothesis can be rejected to conclude that the maps stem from different distributions (Table 4.5). Conover's test with a Bonferroni correction indicated that the maps are all different from one another ($p \ll 0.01$). These results were confirmed with a Pairwise Wilcoxon signed-rank test with Bonferroni correction.

Table 4.5: The results of Friedman's rank sum test and *post hoc* test between GHSM, GEFM, LHSM, and LEFM.

Data descriptions

Sample size (n)		200
Degree of freedom ($d.f.$)		3
Mean ranks	GHSM	1.65
	GEFM	3.68
	LHSM	1.38
	LEFM	3.29

Test statistics

Friedman's χ^2	479.21
asymptotic p-value	$\ll 0.01$

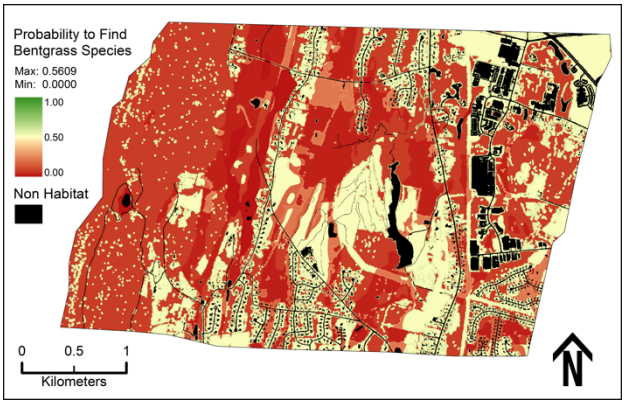
Post hoc test

Conover's test

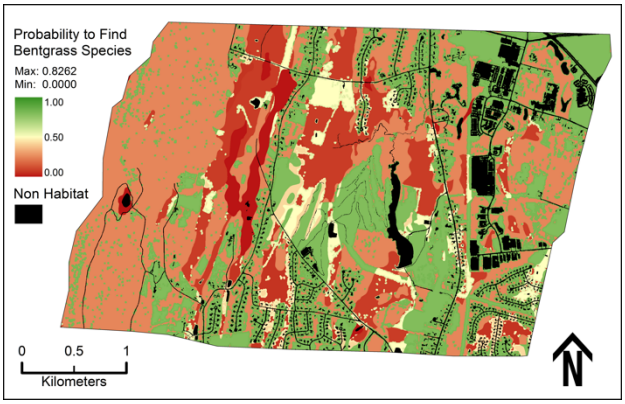
(p-values with Bonferroni correction)

	GHSM	GEFM	LHSM
GEFM	$\ll 0.01$	—	—
LHSM	$\ll 0.01$	$\ll 0.01$	—
LEFM	$\ll 0.01$	$\ll 0.01$	$\ll 0.01$

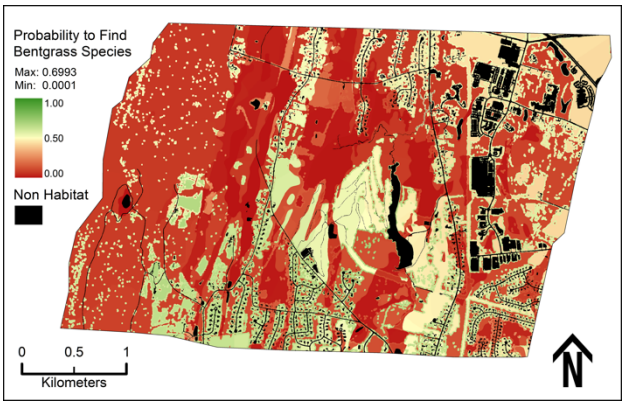
Grouping	Map types
A	GHSM
B	GEFM
C	LHSPM
D	LEFPM



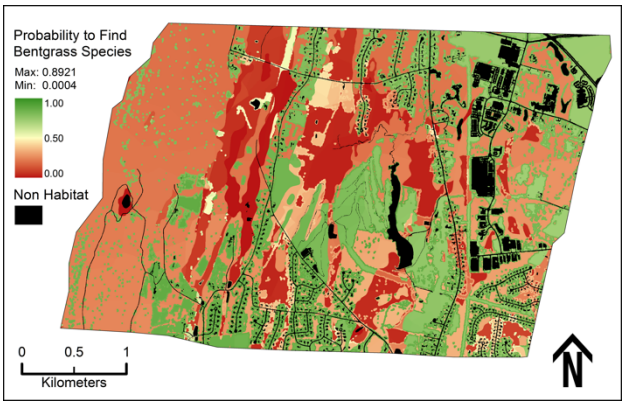
(a) GHSM, 12 unique pixel values



(b) GEFM, 12 unique pixel values



(c) LHSM, 1081 unique pixel values



(d) LEFM, 1081 unique pixel values

Figure 4.5: Maps with habitat suitability values in probability scaled from zero to one.
(a) GHSM, (b) GEFM, (c) LHSM, and (d) LEFM.

4.3.2.2 Cochran's Q test

There was a significant difference in proportion of bentgrass present among GHSPM, GEFPM, LHSPM, and LEFPM according to the result of Cochran's Q test (n = 200, $Q(3)$ = 54.82, and $p \ll 0.01$). This result shows that at least one prediction map has different predictive ability. Accordingly, a pairwise McNemar's test with a Bonferroni correction was performed *post hoc* to show the difference in detail. The *post hoc* grouping shows that GHSPM is significantly different to other three maps ($p \leq 0.001$), but there were no significant differences among GEFPM, LHSPM, and LEFPM. Especially the LHSPM and LEFPM are exactly identical (p = 1.00). Pixel-by-pixel subtraction between LHSPM and LEFPM produced a raster layer of all zeros, confirming that LHSPM and LEFPM are identical (Table 4.6) in spite of the underlying probability maps *not* being identical.

Table 4.6: Cochran's Q test for predictive ability comparison between four habitat maps.

Data descriptions

Sample size (n)		200
Degree of freedom (*d.f.*)		3
Frequencies	Present (1)	Absent (0)
GHSPM	51	149
GEFPM	82	118
LHSPM	75	125
LEFPM	75	125

Test statistics

Cochran's Q	54.82
asymptotic p-value	$\ll 0.01$

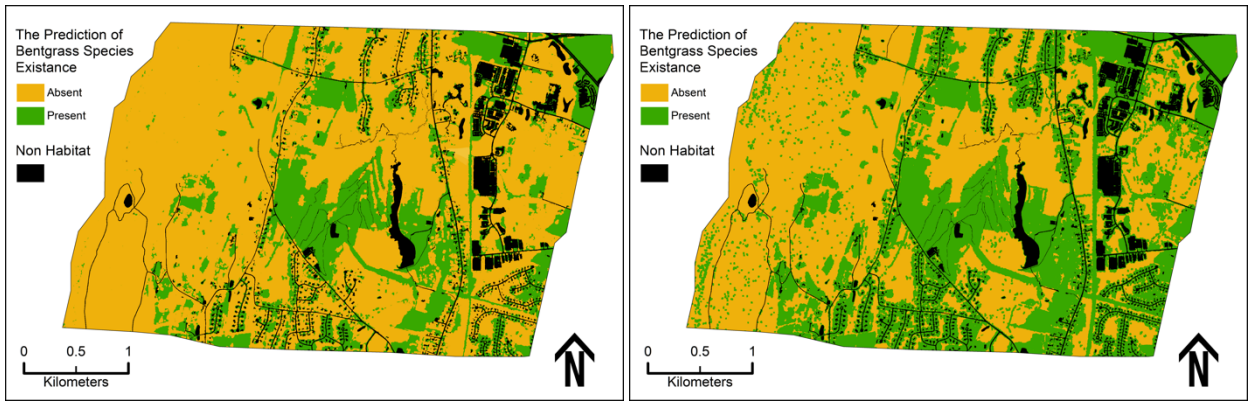
Post hoc test

Pairwise McNemar's test
(p-values with Bonferroni correction)

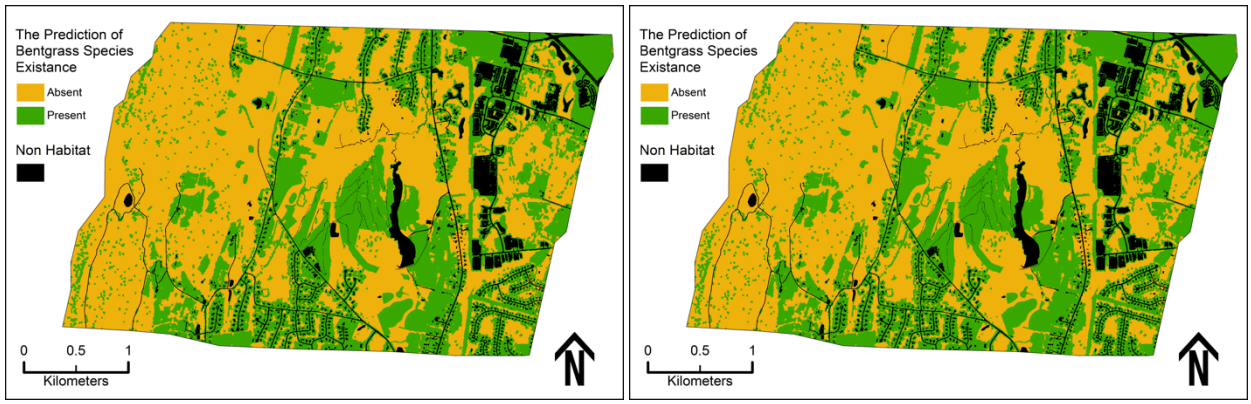
	GHSPM	GEFPM	LHSPM
GEFPM	$\ll 0.01$[*]	—	—
LHSPM	0.01[*]	0.094[†]	—
LEFPM	0.01[*]	0.094[†]	1.000[†]

Grouping	Map types
A	GHSPM
B	GEFPM
B	LHSPM
B	LEFPM

[*] McNemar's Test applied
[†] Binomial Test/Exact McNemar's Test applied because of the small number of misclassified.



(a) GHSPM, cutoff value: 0.1226 (b) GEFPM, cutoff value: 0.3350



(c) LHSPM, cutoff value: 0.1275 (d) LEFPM, cutoff value: 0.3420

Figure 4.6: Prediction maps of bentgrass in dichotomous scale (present/absent), which are reclassified from probability maps in Figure 4.5. ROC analysis was applied to find cutoff values from each habitat models. When habitat suitability was higher than the cutoff, the cell was reclassified as presence of bentgrass, and when habitat suitability value was lower than the cutoff, the cell was reclassified as absence of bentgrass.

4.3.3 Validation of bentgrass predictions

4.3.3.1 Cross tabulation, χ^2 independence test, and strength of association

The confusion matrices for the maps are tabulated in Table 4.7. Pearson's χ^2 test was applied to check the independence between predicted presence/absence of bentgrass by each map and the actual bentgrass presence/absence of the field data. The prediction by GHSPM was independent of the field data (a negative result) (n = 86, $\chi^2(1)$ = 3.15, and p = 0.076). However, bentgrass predictions from the other maps (GEFPM, LHSPM, and LEFPM) are not independent

of the field data (a positive result) (GEFPM / n = 86, $\chi^2(1)$ = 9.66, and p < 0.01; LHSPM / n = 86, $\chi^2(1)$ = 11.22, and p < 0.01; and LEFM / n = 86, $\chi^2(1)$ = 11.22, and p < 0.01) (Table 4.7).

Yule's Q and the Φ coefficient were calculated to quantify the measure-of-association between predicted bentgrass presence/absence to the field data. For GHSPM, there is a moderate strength-of-association (Yule's Q = 0.44 and Φ = 0.19), but other maps (GEFPM, LHSPM, and LEFPM) have strong associations: GEFM has Yule's Q = 0.72 and Φ = 0.36, and both LHSM and LEFM have Yule's Q = 0.75 and Φ = 0.36.

Table 4.7: Contingency tables between actual bentgrass found in the field and predicted values from each map (GHSPM, GEFPM, LHSPM, and LEFPM), and Pearson's χ^2 test of independence results. The size of entire validation set is 86.

(a) GHSPM × actual bentgrass in the field

		Bentgrass in the field		
		Present	Absent	Row total
Predicted Values from GHSPM	Present	9 32.1%	19 67.9%	28 100.0%
	Absent	9 15.5%	49 84.5%	58 100.0%
	Column total	18 20.9%	68 79.1%	86 100.0%

Pearson's $\chi^2(1)$ = 3.15, p-value = 0.076,
Yule's Q = 0.44, Φ = 0.19

(b) GEFPM × actual bentgrass in the field

		Bentgrass in the field		
		Present	Absent	Row total
Predicted Values from GEFPM	Present	14 35.9%	25 64.1%	39 100.0%
	Absent	4 8.5%	43 91.5%	47 100.0%
	Column total	18 20.9%	68 79.1%	86 100.0%

Pearson's $\chi^2(1)$ = 9.66, p-value < 0.01,
Yule's Q = 0.72, Φ = 0.36

(c) LHSPM × actual bentgrass in the field

		Bentgrass in the field		
		Present	Absent	Row total
Predicted Values from LHSPM	Present	14 37.8%	23 62.2%	37 100.0%
	Absent	4 8.2%	45 91.8%	49 100.0%
	Column total	18 20.9%	68 79.1%	86 100.0%

Pearson's $\chi^2(1)$ = 11.22, p-value < 0.01,
Yule's Q = 0.75, Φ = 0.36

(d) LEFPM × actual bentgrass in the field

		Bentgrass in the field		
		Present	Absent	Row total
Predicted Values from LEFPM	Present	14 37.8%	23 62.2%	37 100.0%
	Absent	4 8.2%	45 91.8%	49 100.0%
	Column total	18 20.9%	68 79.1%	86 100.0%

Pearson's $\chi^2(1)$ = 11.22, p-value < 0.01,
Yule's Q = 0.75, Φ = 0.36

Table 4.8: Estimation by binary classifiers between each habitat map and real bentgrass presence/absence field data (unit: %)

Test statistics	Habitat map types			
	GHSPM	GEFPM	LHSPM	LEFPM
Prevalence	20.93	20.93	20.93	20.93
Detection Prevalence	32.56	45.35	43.02	43.02
Accuracy	67.44	66.28	68.60	68.60
Balanced Accuracy	61.03	70.51	71.98	71.98
Sensitivity / True Positive Rate	50.00	77.77	77.77	77.77
Specificity / True Negative Rate	72.06	63.24	66.18	66.18
Positive Predictive Value (PPV)	32.14	35.89	37.84	37.84
Negative Predictive Value (NPV)	84.48	91.49	91.84	91.84

4.3.3.2 Binary classification analyses

Binary classification analyses determined which map has the best predictive ability (Table 4.8). In the table, "Prevalence" means the percentage of ground-truth pixels where bentgrass was found, and the prevalence was around 21%. "Detection Prevalence" means the percentage of ground-truth pixels where bentgrass was correctly predicted to be found. GEFPM, LHSPM, and LEFPM have higher detection prevalence percentages (around 44%) than GHSPM (33%). "Sensitivity" means the percentage of ground-truth pixels correctly classified as present compared to the actual number of present pixels. GEFPM, LHSPM, and LEFPM have higher values (all around 78%) than GHSPM (50%). "Specificity" means the percentage of ground-truth pixels correctly classified as absent compared to the actual number of absent pixels. GHSPM (72%) has higher specificity than other maps (around 65%). "Positive predictive value" means the percentage of the actual number of present pixels compared to the number of pixels predicted present. The maps have similar values (32% to 38%). "Negative predictive value" means the percentage of actual absent compared to the number of pixels predicted absent. GEFPM,

LHSPM, and LEFPM have higher values (around 91%) than GHSPM (84%). "Accuracy" means the proportion of correctly classified pixels, and the maps have similar values around 67%. However, because the prevalence is (only) 21%, the validation data are unbalanced, and accuracy is obscured by this somewhat. "Balanced accuracy" is the arithmetic average of sensitivity and specificity, and the balanced accuracies of GEFPM, LHSPM, and LEFPM are raised to around 71%, but the balanced accuracy of GHSPM drops to 61%.

4.4 Discussion and Conclusions

There were four main questions comparing the four bentgrass habitat models (GHS, GEF, LHS, and LEF models): (1) Do these habitat models have identical distributions? (2) Do these habitat models have identical predictive abilities? (3) How well do the predictions of each model reflect the ground-truth data? and (4) Which of the model has the best predictive ability?

The models produce different results. (1) A Friedman's rank sum test showed the models stem from different underlying probability distributions. (2) A Cochran's *Q post hoc* test revealed that GHSPM is different from the others, which are not different from each other. (3) Indeed, a χ^2 test suggested that GHSPM was independent of the distribution of the ground-truth data. Here, the expectation is that the models should be strongly dependent on the ground-truth data, but that was not the case for GHSPM; however, the others were. (4) The predictive abilities of GEFPM, LHSPM, and LEFPM are better than GHSPM; however, the predictive abilities among GEFPM, LHSPM, and LEFPM are not statistically different. The PPVs are below 40%, so the models as a group don't show strong predictive ability for bentgrass presence. Generally, all the NPVs are high – the lowest NPV is 84% for the global-support habitat suitability model – and all the other models have a NPV over 91%. So, all the models have high

ability to predict bentgrass absence, which was confirmed by the χ^2 tests and the other binary classifiers.

Logistic regression models require input values for their independent variables (obviously), which, in this case, were the *Tree* and *Water* variables. The *Tree* values came from the layer produced by classifying an aerial photograph (Figure 4.2 (a)) and the *Water* values came from a hydro-soils map (Figure 4.2 (b)). Applying a global model is straight forward because there is only one model and many peer scientists, such as Glenz *et al*. (2001), Pereira and Itami (1991), and Store and Jokimäki (2003), applied a global modelling to build habitat suitability maps. However, for local models, there is a question regarding which model to apply: there is one model for each ground-truth sampling point (Brunsdon, Fotheringham, & Charlton, 1998; Fotheringham, Brunsdon, & Charlton, 2003) and it's not obvious which model to use for pixels in between several sampling points. This problem's resolution was to use computational geometry representing proximity (Okabe, Boots, & Sugihara, 1994; Preparata & Shamos, 1985), which is originally proposed by a Russian mathematician Voronoi (as cited in Gold, Remmele, & Roos, 1997). Voronoi polygons tile the plane in a manner dictated by a set of so-called "generating points" that, in this case, were the sampling points. There is one Voronoi polygon per generating point, and a Voronoi polygon is the set of points in the plane nearest its generating point (Aurenhammer, 1991). This naturally divides the study area into convex polygons such that all pixels inside a Voronoi polygon are closer to its sample point that any other sample point (Burrough, McDonnell, & Lloyd, 2015; Gold, 1989). Figure 4.4 shows the set of Voronoi polygons that define the domains of the local models, one per polygon.

Each Voronoi polygon has its own model, and the union of them constitutes a prediction surface across the study area. This prediction surface is not zero-order continuous: there are

probability discontinuities along the edges of the Voronoi polygons although they are not visible in Figure 4.5 (c) and (d). Meyer (2004) showed this for kriging estimations with local support, and the root cause was the principle that piece-wise estimators are not continuous across edges unless they are constructed to be continuous, and the local-support models in this study were not so constructed.

An interesting feature of the probability maps (GHSM, GEFM, LHSM, and LEFM) is that the values produced by the models are effectively categorical even though the probability maps are technically in continuous scale. The global-support models have only 12 unique values of predicted probability, and the local-support models have 1081 unique values of predicted probability. This categorization is due to the categorical nature of independent variables, $Tree$ and $Water$. $Tree$ has three values and $Water$ has four as explained in the comments of Table 4.2, so the global-support models cannot produce more than $3\times4 = 12$ values. The situation is actually the same for the local-support models despite them producing 1081 values. The difference is partly due to the fact that there are 203 local-support models. If the domain of each model had the 12 combinations of the *Tree* and *Water* variables, then the local-support models would have had 2436 unique values instead of 1081. The difference is also due to each model having different weights based on the spatial separations from its neighbors (4.5). The locations of the sampling points were random, so the weights, which depend on separation, ought not be the same in general. The global-support models produce much coarser results than the local-support models, but this fact *per se*, does not imply the global-support models are inferior. Global models emphasize similarities across the whole area, so this model has frequently applied to search for regularities or general laws of the area, but local models emphasize differences across the study area, so this model usually has applied to search for exceptions and hot-spots

(Fotheringham, Brunsdon, & Charlton, 2003; Unwin & Unwin, 1998). For this reason, global models usually produce simpler or coarse maps compared to the maps derived from local models. Statistical tests are necessary to judge which model is best, if any, or even if the models are different.

A Cochran's Q test can compare the predictive ability between the four models represented in binary scale (Cochran, 1950; Conover, 1999). Accordingly, GHSPM has a different predictive ability than GEFPM, LHSPM, and LEFPM but there is no evidence that GEFPM, LHSPM, and LEFPM have different predictive abilities (Table 4.6). The habitat-suitability models did not correct for the unbalanced data, and the global models are a one-size-fits-all approach. Furthermore, the cutoff value between the *present* and *absent* predictions for the global models came from a ROC analysis (Fawcett, 2006; Hosmer, Lemeshow, & Sturdivant, 2013; Metz, 2000) whose inputs were limited to only the 12 possible values, which seems rather coarse. The global habitat suitability model was afflicted by all three shortcomings. It appears that no handicap alone was sufficient to produce inferior predictions, but all of them together were. Surprisingly, the predictive maps from the local-support models (LHSPM and LEFPM) are identical despite LHSM and LEFM not being identical themselves and LHSPM and LEFPM having different ranges. LHSPM and LEFPM have different cutoff values, which are chosen through a numerical maximization process somewhat akin to a maximum likelihood estimation. The exact details of that *R*-package algorithm (ROCR) are unknown to this author, but having the cutoff-value selection processes ultimate produce identical prediction maps seems notable, and a confirmation of the approach.

The present research offers two scenarios of map comparison that can be applied to other studies. The first scenario is the between-map comparison when there are more than two maps. If

the maps are in continuous scale, the comparison can be applied through a Friedman's rank sum test regardless of the distributions of the sampled data from the maps. If the distribution of each map follows normal distribution, it might be possible to apply ANOVA instead of the nonparametric approach. When the maps are in the binary scale, it is possible to compare whether the predictive ability is identical between maps through a Cochran's Q test.

The second scenario is the single-map comparison to ground-truth data when the map is in the binary scale. The prediction from a map can be tested to elucidate whether this result is dependent or independent to the ground-truth data. As a follow-up, various numerical coefficients can be offered to show how closely this prediction and the actual ground-truth data correlated. Also, various binary classifiers such as accuracy, balanced accuracy, sensitivity, specificity, PPV, and NPV could be other follow-ups and offers various criteria to evaluate the predictive ability of a map. Comparing between these various classifiers of each map, the best predictive model can be distinguished between the models.

The present research focused on the comparisons of continuous scaled maps and binary-scaled maps. However, it would be challenging to extend the predictive ability evaluation to the categorical scaled map with more than three categories for the future research.

4.5 Bibliography

Ahrens, C., Chung, J., Meyer, T., & Auer, C. (2011). Bentgrass distribution surveys and habitat suitability maps support ecological risk assessment in cultural landscapes. *Weed science*, *59*(2), 145-154.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716-723.

Akobeng, A. K. (2007). Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Paediatrica*, *96*(5), 644-647.

Allison, P. D. (2001). *Logistic Regression Using SAS System: Theory and Application*. Cary, NC: SAS Institute. 308 pp.

Aurenhammer, F. (1991). Voronoi diagrams–a survey of a fundamental geometric data structure. *ACM Computing Surveys (CSUR)*, *23*(3), 345-405.

Ayalew, L., Yamagishi, H., Marui, H., & Kanno, T. (2005). Landslides in Sado Island of Japan: Part II. GIS-based susceptibility mapping with comparisons of results from two methods and verifications. *Engineering Geology*, *81*(4), 432-445.

Bellis, L. M., Pidgeon, A. M., Radeloff, V. C., St-Louis, V., Navarro, J. L., & Martella, M. B. (2008). Modeling habitat suitability for greater rheas based on satellite image texture. *Ecological Applications*, *18*(8), 1956-1966.

Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). *Applied Spatial Data Analysis with R* (2nd ed.). New York, NY: Springer. 405 pp.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1998). Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *47*(3), 431-443.

Burgman, M. A., Breininger, D. R., Duncan, B. W., & Ferson, S. (2001). Setting reliability bounds on habitat suitability indices. *Ecological Applications*, *11*(1), 70-78.

Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2nd ed.). New York, NY: Springer. 488 pp.

Burrough, P. A., McDonnell, R. A., Lloyd, C. D. (2015). *Principles of Geographical Information Systems*. Oxford, UK: Oxford University Press. 432 pp.

Cochran, W. G. (1950). The comparison of percentages in matched samples. *Biometrika*, *37*(3/4), 256-266.

Conover, W. J., & Iman, R. L. (1976). *On Multiple-Comparisons Procedures* (Informal Report, No. LA-7677-MS). Los Alamos, NM: Los Alamos Scientific Laboratory. 14 pp.

Conover, W. J. (1999). *Practical Nonparametric Statistics* (3rd ed.). Hoboken, NJ: Wiley. 584 pp.

Cox, D. R. (1958). The regression analysis of binary sequences (with discussion). *Journal of the Royal Statistical Society. Series B*. *20*(2). 215-242.

Cramer, H. (1946). *Mathematical Methods of Statistics*. Princeton, NJ: Princeton University Press. 575 pp.

Davenport Jr, E. C., & El-Sanhurry, N. A. (1991). Phi/phimax: Review and Synthesis. *Educational and Psychological Measurement*, *51*(4), 821-828.

Davis, J. A. (1971). *Elementary Survey Analysis*. Englewood Cliffs, NJ: Prentice-Hall. 195 pp.

Dorren, L. K., & Seijmonsbergen, A. C. (2003). Comparison of three GIS-based models for predicting rockfall runout zones at a regional scale. *Geomorphology*, *56*(1), 49-64.

Elith, J. (2000). Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. In S. Ferson & M. Burgman (Eds.), *Quantitative Methods for Conservation Biology* (pp. 39-58). New York, NY: Springer. 322 pp.

Evangelista, P. H., Kumar, S., Stohlgren, T. J., Jarnevich, C. S., Crall, A. W., Norman, J. B., III., & Barnett, D. T. (2008). Modelling invasion for a habitat generalist and a specialist plant species. *Diversity and Distributions*, *14*(5), 808-817.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861-874.

Fotheringham, A. S., & Wegener, M. (1999). *Spatial models and GIS: New and potential models* (Vol. 7). London, UK: CRC press. 312 pp.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2003). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. West Sussex, UK: John Wiley & Sons. 282 pp.

Franco, A., Brito, J. C., & Almeida, J. (2000). Modelling habitat selection of Common Cranes Grus grus wintering in Portugal using multiple logistic regression. *Ibis*, *142*(3), 351-358.

Glenz, C., Massolo, A., Kuonen, D., & Schlaepfer, R. (2001). A wolf habitat suitability prediction study in Valais (Switzerland). *Landscape and Urban Planning*, *55*(1), 55-65.

Gold, C. M. (1989). Surface Interpolation, Spatial Adjacency and GIS. In J. Raper (Ed.), *Three Dimensional Applications in Geographic Information Systems* (pp. 21-35). London, UK: Taylor & Francis. 280 pp.

Gold, C. M., Remmele, P. R., & Roos, T. (1997). Voronoi methods in GIS. In M. van Kreveld, J. Nievergelt, T. Roos, & P. Widmayer (Eds.), *Algorithmic Foundations of Geographic Information Systems* (pp. 21-35). Berlin, Germany: Springer. 287 pp.

Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. (2013). GWmodel: An R package for exploring spatial heterogeneity using geographically weighted models. *Journal of Statistical Software*, *63*(17), 1-50. Retrieved from https://www.jstatsoft.org/index.php/jss/article/view/v063i17/v63i17.pdf

Gray, P. A., Cameron, D., & Kirkham, I. (1996). Wildlife habitat evaluation in forested ecosystems: some examples from Canada and the United States. In R. M. DeGraaf & R. I. Miller (Eds.), *Conservation of faunal diversity in forested landscapes* (pp. 407-536). New York, NY: Chapman and Hall. 634 pp.

Harris, P., Brunsdon, C., & Charlton, M. (2011). Geographically Weighted Principal Components Analysis. *International Journal of Geographical Information Science*, *25*(10), 1717-1736.

Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C., & Guisan, A. (2006). Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, *199*(2), 142-152.

Hosmer, D. W., Jr., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Hoboken, NJ: John Wiley & Sons. 528 pp.

Jacobs, M. J., & Macisaac, H. J. (2009). Modelling spread of the invasive macrophyte Cabomba caroliniana. *Freshwater Biology*, *54*(2), 296-305.

Kleinbaum, D. G., & Klein, M. (2006). *Logistic Regression: A Self-Learning Text* (2nd ed.). New York, NY: Springer. 514 pp.

Knoke, D., Bohrnstedt, G. W., & Mee, A. P. (2002). *Statistics for Social Data Analysis* (4th ed.). Wadsworth Publishing. 530 pp.

Larson, M. A., Dijak, W. D., Thompson, F. R., III., & Millspaugh, J. J. (2003). *Landscape-level habitat suitability models for twelve wildlife species in southern Missouri* (General Technical Report NC-233). St. Paul, MN: North Central Research Station, Forest Service, U.S. Department of Agriculture. Retrieved from https://www.ncrs.fs.fed.us/pubs/gtr/gtr_nc233.pdf

Lehmann, E. L., & D'Abrera, H. J. (2006). *Nonparametrics: Statistical methods based on ranks*. New York: Springer. 464 pp.

Longley, P. & Batty, M. (ed.) (1996). *Spatial Analysis: Modelling in a GIS Environment*. New York, NY: John Wiley & Sons. 392 pp.

Madsen, A. B., & Prang, A. (2001). Habitat factors and the presence or absence of otters Lutra lutra in Denmark. *Acta theriologica*, *46*(2), 171-179.

Marvier, M., Carrière, Y., Ellstrand, N., Gepts, P., Kareiva, P., Rosi-Marshall, E., Tabashnik, B. E., & Wolfenbarger, L. L. (2008). Harvesting data from genetically engineered crops. *Science, 320*(5875), 452-453.

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, *12*(2), 153-157.

Metz, C. E. (2000). Fundamental ROC analysis. *Handbook of medical imaging*, *1*, 751-769.

Metzler, K. J., & Barrett, J. P. (2006). *The Vegetation of Connecticut, a Preliminary Classification* (Report of Investigations No. 12). Hartford, CT: Department of Environmental Protection, State Geological and Natural History Survey of Connecticut. 109 pp.

Meyer, T. H. (2004). The discontinuous nature of kriging interpolation for digital terrain modeling. *Cartography and Geographic Information Science*, *31*(4), 209-216.

Nielsen, C., Hartvig, P., & Kollmann, J. (2008). Predicting the distribution of the invasive alien Heracleum mantegazzianum at two different spatial scales. *Diversity and Distributions*, *14*(2), 307-317.

Okabe, A., Boots, B., & Sugihara, K. (1994). Nearest neighbourhood operations with generalized Voronoi diagrams: a review. *International Journal of Geographical Information Systems*, *8*(1), 43-71.

Pearce, J., & Ferrier, S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, *133*(3), 225-245.

Pebesma, E. J., & Bivand, R. S. (2005). Classes and methods for spatial data in R. *R news*, *5*(2), 9-13.

Pereira, J., & Itami, R. (1991). GIS-based habitat modeling using logistic multiple regression – A study of the Mt. Graham red squirrel. *Photogrammetric Engineering and Remote Sensing*, *57*(11), 1475-1486.

Preparata, F. P., & Shamos, M. I. (1985). *Computational Geometry: An introduction*. New York, NY: Springer. 398 pp.

Rand, G. M., & Newman, J. R. (1998). The applicability of habitat evaluation methodologies in ecological risk assessment. *Human and Ecological Risk Assessment: An International Journal*, *4*(4), 905-929.

Real, R., Barbosa, A. M., & Vargas, J. M. (2006). Obtaining environmental favourability functions from logistic regression. *Environmental and Ecological Statistics*, *13*(2), 237-245.

Sánchez-Flores, E., Rodriguez-Gallegos, H., & Yool, S. R. (2008). Plant invasions in dynamic desert landscapes. A field and remote sensing assessment of predictive and change modeling. *Journal of Arid Environments*, *72*(3), 189-206.

Sheskin, D. J. (2011). *Handbook of Parametric and Nonparametric Statistical Procedures* (5th ed.). Boca Raton, FL: CRC Press. 1926 pp.

Store, R., & Jokimäki, J. (2003). A GIS-based multi-scale approach to habitat suitability modeling. *Ecological Modelling*, *169*(1), 1-15.

Swets, J. A. (1973). The relative operating characteristic in psychology. *Science*, *182*(4116), 990-1000.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, *240*(4857), 1285-1293.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Better DECISIONS through SCIENCE. *Scientific American*, *283*(4), 82-87.

Tobler, W. R., (1970). A Computer Movie Simulating Urban Growth in the Detroit region. *Economic Geography*, *46*(2), 234-240.

UCLA Institute for Digital Research and Education (IDRE). (2006). *R Library Contrast Coding Systems for categorical variables*. Los Angeles, CA: UCLA Institute for Digital Research and Education, University of California Los Angeles. Retrieved by https://stats.idre.ucla.edu/r/library/r-library-contrast-coding-systems-for-categorical-variables/

Unwin, A., & Unwin, D. (1998). Spatial data analysis with local statistics. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *47*(3), 415-421.

U.S. Fish and Wildlife Service. (1980). *Habitat Evaluation Procedures (HEP)* (ESM 102). Washington, DC: Division of Ecological Services, Fish and Wildlife Service, U.S. Department of Interior. Retrieved from https://www.fws.gov/policy/esm102.pdf

U.S. Fish and Wildlife Service. (1981). *Standards for the Development of Habitat Suitability Index Models for Use in the Habitat Evaluation Procedure* (ESM 103). Washington, DC: Division of Ecological Services, Fish and Wildlife Service, U.S. Department of Interior. Retrieved from https://www.fws.gov/policy/ESM103.pdf

Van Horne, B., & Wiens, J. A. (1991). *Forest bird habitat suitability models and the development of general habitat models* (Fish and Wildlife Research, 8). Washington, DC: Fish And Wildlife Service, U.S. Department of the Interior. 31 pp. Retrieved from http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA322800

Visser, H., & De Nijs, T. (2006). The map comparison kit. *Environmental Modelling & Software*, *21*(3), 346-358.

Walker, S. H., & Duncan, D. B. (1967). Estimation of the probability of an event as a function of several independent variables. *Biometrika*, *54*(1-2), 167-179.

Wang, K., Zhang, C., & Li, W. (2012). Comparison of geographically weighted regression and regression kriging for estimating the spatial distribution of soil organic matter. *GIScience & Remote Sensing*, *49*(6), 915-932.

Wang, K., Zhang, C., & Li, W. (2013). Predictive mapping of soil total nitrogen at a regional scale: a comparison between geographically weighted regression and cokriging. *Applied Geography*, *42*, 73-85.

Wang, K., Zhang, C. R., Li, W. D., Lin, J., & Zhang, D. X. (2014). Mapping soil organic matter with limited sample data using geographically weighted regression. *Journal of Spatial Science*, *59*(1), 91-106.

Warrens, M. J. (2008). On association coefficients for 2×2 tables and properties that do not depend on the marginal distributions. *Psychometrika*, *73*(4), 777-789.

Yule, G. U. (1912). On the methods of measuring association between two attributes. *Journal of the Royal Statistical Society*, *75*(6), 579-652.

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, *39*(4), 561-577.