

## ABSTRACT

MEHRA, LUCKY KUMAR. Risk Assessment and Prediction of *Stagonospora nodorum* blotch in Winter Wheat. (Under the direction of Drs. Peter S. Ojiambo and Christina Cowger).

*Stagonospora nodorum* blotch (SNB), caused by the fungus *Parastagonospora nodorum*, is a major disease of wheat (*Triticum aestivum*). Pre-planting factors have been associated with the late-season SNB severity, however, the relative influence of these factors on the risk of the disease has not been determined. The overall goal of this work was to evaluate the combined effect of pre-planting factors and in-season weather on the onset and development of SNB in winter wheat.

Residue from a previously infected wheat crop is one of the pre-planting factors that influences SNB severity and can also be an important source of initial inoculum. Therefore, the effects of infected residue on disease severity and yield were quantified. Experiments were conducted at 10 site-years that included a combination of three locations and three years (2012 to 2014) using a moderately susceptible winter wheat cultivar DG Shirley and a highly susceptible cultivar DG 9012. Across site-years, disease severity ranged from 0 to 50% and increased nonlinearly as residue level increased, with a rapid rise to an upper limit with little change in severity above 20 to 30% soil surface coverage. The effect of residue coverage on yield was only significant for DG Shirley in 2012 and for DG 9012 at Salisbury in 2014. SNB also resulted in significant effects on yield and test weight in the most disease-conducive environments, suggesting that the economic threshold for the disease may be higher than previously assumed and warrants review.

The performance of multiple regression (MR) and three machine learning algorithms (artificial neural networks, categorical and regression trees, and random forests) was

examined in predicting the pre-planting risk of SNB in wheat. Pre-planting factors tested were cultivar resistance, latitude, longitude, previous crop, seeding rate, seed treatment, tillage type, and wheat residue. Disease severity assessed at the end of the growing season was used as the response variable. The models were developed using 431 disease cases (unique combinations of predictors) collected from 2012 to 2014, and these cases were randomly divided into training, validation, and test datasets. A strong relationship was observed between specific pre-planting predictors and late-season severity of SNB whereby latitude, longitude, wheat residue, and cultivar resistance were the most important predictors. Results show that the random forest (RF) algorithm, which explained 79% of the variability within the data, was the most accurate in predicting the risk of SNB. The RF algorithm could allow early assessment of the risk of SNB, facilitating sound disease management decisions prior to planting of wheat.

Management of SNB during the season mainly relies on fungicide sprays after flag leaf emergence. However, the disease can occur much earlier than flag leaf emergence, and the relationship between time of SNB onset and yield in winter wheat is not well understood. SNB incidence was recorded in 435 disease cases collected from 11 counties in North Carolina during the 2012 to 2014 growing seasons. SNB onset was defined as day of the year (doy) when SNB incidence reached 50%. Starting at the end of the tillering growth stage, temperature, relative humidity, and precipitation were combined to calculate the daily infection value (DIV), a weather index ranging from no (= 0) to optimum (= 1) growth of *P. nodorum*. Based on the SNB onset-yield relationship, doy 102 was identified as the onset cutoff point. If onset occurred after doy 102 ('late'), 72% of the cases had above-average yield, while if the onset occurred before doy 102 ('early'), only 13% of the cases had above-

average yield. Binary logistic regression was used to predict onset classes as a function of i) cumulative DIV until the flag leaf was visible and ii) pre-planting factors. Wheat residue and cumulative DIV until stem elongation were found to be significant predictors of SNB onset. The binary logistic model had a correct classification rate of 0.94. The performance of the model in cross-validation techniques was also high. Once validated using independent data, the model could serve as a within-season decision support tool to help growers with decisions on fungicide application for SNB management.

© Copyright 2015 by Lucky Kumar Mehra

All Rights Reserved

Risk Assessment and Prediction of Stagonospora nodorum blotch in Winter Wheat

by  
Lucky Kumar Mehra

A dissertation submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the Degree of  
Doctor of Philosophy

Plant Pathology

Raleigh, North Carolina

2015

APPROVED BY:

---

Peter S. Ojiambo  
Committee Co-Chair

---

Christina Cowger  
Committee Co-Chair

---

Randy Weisz

---

Kevin Gross

## **DEDICATION**

To my parents for inspiring me to work hard

## **BIOGRAPHY**

Lucky Kumar Mehra grew up in a village in district Gurdaspur, Punjab, India. Thanks to the advice of his high school mentor Neeraj Gupta, Lucky applied to Punjab Agricultural University, Ludhiana, India to pursue Bachelor of Science in Agriculture. He thoroughly enjoyed all the agricultural courses, but developed special interest in plant diseases and decided to choose Plant Protection as major during the final year of the degree. He completed his B.Sc. in 2008. Lucky was accepted to University of Georgia, Athens in 2009 to conduct research on a postharvest disease of blueberry under the guidance of Dr. Harald Scherm. After completing his M.S. degree in December 2010, Lucky joined the Department of Plant Pathology at North Carolina State University in the spring of 2011. His doctoral research focused on the epidemiology of *Stagonospora nodorum* blotch of wheat under the guidance of Dr. Peter Ojiambo and Dr. Christina Cowger. He conducted a series of field experiments from 2011 to 2014 across North Carolina to develop models to predict the risk and onset of *Stagonospora nodorum* blotch in winter wheat. The results of those studies are presented in this dissertation.

## ACKNOWLEDGMENTS

The work presented in this dissertation could not have been possible without the help of many incredible people. I would like to take this opportunity to express my appreciation. I would especially like to thank my co-advisors Dr. Peter S. Ojiambo and Dr. Christina Cowger, for their guidance and timely suggestions, both professionally and personally. I thoroughly enjoyed working with both of them. Thanks are also due to my committee members, Dr. Randy Weisz and Dr. Kevin Gross, for providing valuable insights into my dissertation work.

For their support in field work and valuable scientific discussions, I would like to thank my current and previous lab-colleagues from both Ojiambo and Cowger lab: Katie Neufeld, Anna Thomas, Wendy Britton, Thomas Keever, Mary Lewis, Rebecca Echerd, Barbara Stevens, Taylor Howell, Ryan Parks, Matt Bertucci, Tyler Cole, Bradley Van Staaldhuizen, Brandon Clontz, Ian Swaim, Charlie Glover, Joseph Jeffreys, Jennifer Patton-Özkurt, and Anita Panthi. I am also indebted to the staff at research stations and county extension agents for providing tremendous support in managing research plots. Thanks to North Carolina Agricultural Experiment Station and United States Department of Agriculture-Agricultural Research Service for providing the financial support for this project.

Finally, I would like to thank my wife Dr. Sushila Chaudhari for her intellectual input and help in the field work. For their unconditional love and support, I would like to thank my parents, siblings, and cousins. To all my friends, and fellow graduate students in the Department of Plant Pathology, North Carolina State University: Thank you for making my



graduate school experience great with picnics, first Fridays, lunches, dinners, and many special moments. Thank you!

## TABLE OF CONTENTS

<b>LIST OF TABLES .....</b>	<b>viii</b>
<b>LIST OF FIGURES .....</b>	<b>x</b>
 <b>CHAPTER 1 .....</b>	 <b>1</b>
<b>Literature Review .....</b>	<b>1</b>
<b>1.1 Importance of wheat .....</b>	<b>1</b>
<b>1.2 Challenges in wheat production .....</b>	<b>2</b>
<b>1.3 Geographical distribution of SNB and associated yield losses .....</b>	<b>3</b>
<b>1.4 Taxonomy of <i>Parastagonospora nodorum</i> .....</b>	<b>4</b>
<b>1.5 Epidemiology of SNB .....</b>	<b>5</b>
<b>1.6 Management of SNB .....</b>	<b>7</b>
<b>1.7 Necrotrophic effectors or host-selective toxins.....</b>	<b>8</b>
<b>1.8 Prediction models for SNB .....</b>	<b>9</b>
<b>1.9 Justification and objectives of study .....</b>	<b>11</b>
<b>Literature Cited .....</b>	<b>15</b>
 <b>CHAPTER 2 .....</b>	 <b>25</b>
<b>Quantifying the effects of wheat residue on severity of <i>Stagonospora nodorum</i> blotch and yield in winter wheat .....</b>	<b>25</b>
<b>Abstract.....</b>	<b>26</b>
<b>Introduction.....</b>	<b>27</b>
<b>Materials and Methods.....</b>	<b>30</b>
<b>Results .....</b>	<b>36</b>
<b>Discussion.....</b>	<b>41</b>
<b>Acknowledgments .....</b>	<b>47</b>
<b>Literature Cited .....</b>	<b>49</b>

<b>CHAPTER 3</b> .....	68
<b>Predicting pre-planting risk of Stagonospora nodorum blotch in winter wheat using machine learning models</b> .....	68
<b>ABSTRACT</b> .....	69
<b>1. Introduction</b> .....	70
<b>2. Materials and Methods</b> .....	74
<b>3. Results</b> .....	85
<b>4. Discussion</b> .....	91
<b>Acknowledgments</b> .....	98
<b>References</b> .....	99
 <b>CHAPTER 4</b> .....	 116
<b>Predicting the onset of Stagonospora nodorum blotch in winter wheat using pre-planting factors and in-season weather</b> .....	116
<b>Abstract</b> .....	117
<b>Introduction</b> .....	118
<b>Materials and Methods</b> .....	123
<b>Results</b> .....	132
<b>Discussion</b> .....	135
<b>Acknowledgements</b> .....	140
<b>Literature Cited</b> .....	141
 <b>CHAPTER 5</b> .....	 159
<b>Conclusions</b> .....	159

## LIST OF TABLES

### CHAPTER 2

<b>Table 2.1.</b> Number of days of weather conducive to the development of <i>Stagonospora nodorum</i> blotch during the epidemic period in North Carolina.....	53
<b>Table 2.2.</b> Effect of wheat residue levels on the intensity of <i>Stagonospora nodorum</i> blotch and yield characteristics for winter wheat cultivar DG Shirley .....	54
<b>Table 2.3.</b> Impact of levels of wheat residue on the development and severity of <i>Stagonospora nodorum</i> blotch and yield characteristics for winter wheat cultivar DG Shirley .....	56
<b>Table 2.4.</b> Effect of wheat residue levels on intensity of <i>Stagonospora nodorum</i> blotch and yield characteristics for winter wheat cultivar DG 9012 in 2014 .....	58
<b>Table 2.5.</b> Impact of amounts of wheat residue on the development and severity of <i>Stagonospora nodorum</i> blotch and yield characteristics for winter wheat cultivar DG 9012 in 2014.....	59
<b>Table 2.6.</b> Orthogonal polynomial contrasts to test the hypothesis of linear, quadratic and cubic response in <i>Stagonospora nodorum</i> blotch severity and yield with increasing levels of wheat residue cover in plots of cultivar DG Shirley .....	61
<b>Table 2.7.</b> Nonlinear regression analysis of the relationship between amounts of wheat residue cover on the ground and the development of <i>Stagonospora nodorum</i> blotch, using a two-parameter logarithmic model .....	63

### CHAPTER 3

<b>Table 3.1.</b> Description of experimental sites and tillage methods used in a study conducted in North Carolina to identify pre-planting factors that influence the risk of <i>Stagonospora nodorum</i> blotch in winter wheat .....	105
<b>Table 3.2.</b> Independent variables tested for their usefulness in assessing pre-planting risk of <i>Stagonospora nodorum</i> blotch in winter wheat .....	106

**Table 3.3.** Results of multiple regression analysis conducted to explain variation in maximum severity of *Stagonospora nodorum* blotch in winter wheat based on pre-planting variables using data collected in North Carolina from 2012 to 2014 ....107

**Table 3.4.** Classification rates, sensitivity, specificity and prediction accuracy of final models developed using the multiple regression and machine learning techniques to predict the risk of *Stagonospora nodorum* blotch in winter wheat based on pre-planting variables using data collected in North Carolina from 2012 to 2014 ....108

## CHAPTER 4

**Table 4.1.** Experimental sites and tillage methods used in three growing seasons in North Carolina to identify factors that influence the onset of *Stagonospora nodorum* blotch in winter wheat.....145

**Table 4.2.** Description of potential predictors tested to determine their usefulness in predicting the onset of *Stagonospora nodorum* blotch in winter wheat.....146

**Table 4.3.** Results of PROC VARCLUS, variable reduction technique used to eliminate redundant variables .....147

**Table 4.4.** Spearman and Hoeffding coefficients for the correlation between potential predictor variables and time of disease onset .....148

**Table 4.5.** Parameter estimates of the final binary logistic regression model developed to predict whether onset of *Stagonospora nodorum* blotch in winter wheat would occur before mid-April in North Carolina .....149

**Table 4.6.** Performance characteristics of the final binary logistic regression, the 10-fold and leave-one-out cross validation models.....150

## LIST OF FIGURES

### CHAPTER 2

**Figure 2.1.** Effect of different levels of wheat residue on the soil surface on the development of *Stagonospora nodorum* blotch in three years at three experimental sites in North Carolina. The wheat cultivar DG Shirley was evaluated in all site-years (**A-G**), while the cultivar DG 9012 was evaluated only in 2014 at all sites (**H-J**). Data points represent means of disease severity (%) values from five replicates; the bars are standard errors of the mean. ....64

**Figure 2.2.** Relationship between *Stagonospora nodorum* blotch intensity (relative area under the disease progress curve) and levels of wheat residue on soil surface in three growing seasons at three experimental sites in North Carolina. The wheat cultivar DG Shirley was evaluated in all years (**A-C**), while the cultivar DG 9012 was evaluated only in 2014 at all sites (**D**). Symbols (circles and triangles) represent the observed disease data, while the curves represent the predicted disease intensity based on a two-parameter logarithmic model. In each year, the coefficients of correlations ( $r$ ) between observed and predicted values are a range depicting the goodness-of-fit of the model to the data across experimental sites.65

### CHAPTER 3

**Figure 3.1.** Schematic flow of an artificial neural network depicting the input, hidden and output layers. The input layer contains predictors, while the output layer contains the response variable. The hidden layer is composed of the combination ( $\Sigma$ ) and transfer (L) functions and summarizes predictor variables and associated weights (W), applies the transfer function, and sends the result to the output layer. The weights (W1, W2, W3...etc.) link input and hidden layer of the neural network.109

**Figure 3.2.** Performance in predicting *Stagonospora nodorum* blotch of wheat of a multiple regression (MR) model and three machine learning models: artificial neural network (ANN), categorical and regression tree (CART), and random forest (RF). Based on running averages of 15 randomizations of the test dataset which comprised 10% of 431 cases collected in the study. **A**) coefficient of determination; **B**) correct classification rate, i.e., proportion of cases correctly classified as low-severity (<30% disease severity) or high-severity ( $\geq 30\%$  disease severity); **C**) specificity, the proportion of cases correctly classified as low-disease; and **D**) sensitivity, the proportion of cases correctly classified as high-disease. Symbols represent mean values for any given number of randomizations

of the test dataset, while the vertical bars represent standard errors of the mean based on the total number of randomizations evaluated.....110

**Figure 3.3.** Importance of pre-planting variables as identified by the neural network model. Variable importance was determined using the ‘weights’ method of Garson (1991). CR = cultivar resistance rating for *Stagonospora nodorum* blotch, LAT = latitude, LON = longitude, SR = seeding rate, ST = seed treatment, and WR = wheat residue on soil surface. ....111

**Figure 3.4.** Classification and regression tree used to estimate severity of *Stagonospora nodorum* blotch in winter wheat based on pre-planting predictor variables collected in North Carolina from 2012 to 2014. Low (L) disease class = disease severity <30%, and high (H) class = disease severity  $\geq$ 30%. The original tree had 25 nodes but was pruned to seven terminal nodes. Within the tree, predictor variables are shown in rectangles, while response variables are shown in circles. CR = cultivar resistance rating for *Stagonospora nodorum* blotch, LAT = latitude, LON = longitude (degrees W), and WR = wheat residue on soil surface.112

**Figure 3.5.** Performance of the Random forest model depicting: **A)** coefficient of determination as a function of the number of trees in the model and **B)** importance of pre-planting predictor variables. The importance of a predictor variable is determined by the percent increase in the mean square error after permuting a specific predictor variable. CR = cultivar resistance rating for *Stagonospora nodorum* blotch, LAT = latitude, LON = longitude, SR = seeding rate, ST = seed treatment, and WR = wheat residue on soil surface. ....113

## CHAPTER 4

**Figure 4.1.** Location of field experiments conducted in 11 counties in North Carolina in the 2012- 2014 growing seasons to determine factors influencing the onset of *Stagonospora nodorum* blotch in winter wheat. Thick grey lines are approximate boundaries of three geographical regions of the state: Piedmont, Coastal plains, and Tidewater. ....151

**Figure 4.2.** For 390 disease cases (unique combinations of pre-planting factors), **A**, distribution of time of *Stagonospora nodorum* blotch (SNB) onset in wheat, and **B**, a scatter plot showing relationship of SNB onset time with yield. Disease onset occurred before mid-April in 158 cases, and after mid-April in 232 cases. The dashed vertical line represents day of year 102 (mid-April), while the dashed horizontal line represents the mean (10 year average) wheat yield of 4.03 t ha<sup>-1</sup> in North Carolina. ....152

- Figure 4.3.** In field experiments conducted on wheat in three geographical regions of North Carolina, the observed and predicted proportions of cases where onset of *Stagonospora nodorum* blotch occurred before mid-April. ....153
- Figure 4.4.** Distribution of cumulative daily infection values (cDIV) until **A**, three, **B**, two, and **C**, one week before a mid-April threshold (day of year 102) within both early and late onset classes. The lower, middle, and upper horizontal lines of each box are 25<sup>th</sup>, 50<sup>th</sup> (median), and 75<sup>th</sup> percentiles of the data, respectively. The vertical bars extend to 10<sup>th</sup> and 90<sup>th</sup> percentiles, and circles are outliers. The cDIVs measure accumulated favorable weather for *Parastagonospora nodorum* growth. ....154
- Figure 4.5.** Predicted probability for the onset of *Stagonospora nodorum* blotch occurring before mid-April in low (<20%) and high ( $\geq$ 20%) wheat residue cases as a function of cumulative favorable weather for *Parastagonospora nodorum* infection observed from day of year 66 (first week of March) until two weeks before day of year 102 (mid-April). ....155
- Figure 4.6.** Receiver operating characteristics curves for the final binary logistic regression, 10-fold and leave-one-out cross-validated models developed to predict the onset of *Stagonospora nodorum* blotch in winter wheat. ....156



## **CHAPTER 1**

### **Literature Review**

#### **1.1 Importance of wheat**

Wheat (*Triticum aestivum*) is the most important staple food of the world population and occupies more than 220 million hectares of land (Anonymous 2015a; Dixon et al. 2009), which is higher than any other cultivated crop. A majority (>65%) of the world population, especially in the developing countries, is dependent on wheat for nutrition needs (Braun et al. 2010). Wheat production is growing annually but is not keeping up with the world population growth (Dixon et al. 2009). In 2014, the world wheat production was 725 million tons (Anonymous 2015a) and wheat production will need to be about 900 million tons to meet global demand by 2050. The arable land is limited, so the increase in wheat production has to come from higher-yielding varieties and by reduction in yield losses caused by various biotic and abiotic factors. This goal of meeting wheat demand by 2050 is a huge challenge especially for resource-poor countries due to global climate change predictions indicating that wheat production could decrease by 20–30% in developing countries (Braun et al. 2010; Dixon et al. 2009).

The United States ranks third in wheat production worldwide after China and India. In 2014, wheat production in the US was estimated at 61 million tons, which is approximately 8% of world wheat production (Anonymous 2015a). Out of total national production in 2014, the shares of hard red winter, hard red spring, soft red winter, white, and durum wheat market classes were 36%, 27%, 23%, 11%, and 3%, respectively (Anonymous

2015a). In the US, North Carolina (NC) produces mainly soft red winter wheat and contributes about 2% (1.34 million tons) to total national wheat production and about 10% to national soft red winter wheat production (Anonymous 2014). The wheat acreage and production in NC has been steadily increasing, owing to comparatively higher prices of wheat in recent years. The price of soft red winter wheat has been around \$5 per bushel after 2008, as compared to pre-2008 when it was \$3 per bushel or lower (Anonymous 2010, 2014 & 2015a).

## 1.2 Challenges in wheat production

Wheat production is affected by many factors such as diseases, insects, weeds, nematodes, and abiotic disorders. About 10% of yield losses are caused by fungal plant pathogens (Oerke 2006). Worldwide, a majority of losses due to fungal pathogens in wheat production are caused by three rust diseases: stem rust, also known as black rust; leaf rust, also known as brown rust; and stripe rust, also known as yellow rust (Dixon et al. 2009). In NC, the most important wheat diseases are leaf rust (*Puccinia triticina*), powdery mildew (*Blumeria graminis*), Stagonospora nodorum blotch (SNB) (*Parastagonospora nodorum*), and head scab (*Fusarium graminearum*) (Quaedvlieg et al. 2013; Weisz 2013). Due to the impact of plant diseases on yields, better management of these biotic constraints is needed for sustainable wheat production and food security.

Traditional wheat breeding programs have mainly been directed towards obligate pathogens, and thus resistant cultivars are available to manage leaf rust and powdery mildew.

The progress on SNB resistance in many breeding programs has been modest due to the sporadic nature of the disease in screening nurseries. Currently, breeding efforts for SNB resistance are underway (Anonymous 2015b; Francki 2013), and disease management relies primarily on prophylactic fungicide sprays. This dissertation focuses on quantification of factors affecting SNB and development of prediction models, which will lead to better management of this disease.

### **1.3 Geographical distribution of SNB and associated yield losses**

*Stagonospora nodorum* blotch occurs in wheat-growing areas worldwide, but the disease is more prevalent in areas with warm (22–27°C) and moist weather, such as the southeastern United States, parts of Europe, southern Brazil, and Australia (Eyal 1987; Murray and Brennan 2009). The pathogen was first described on nodes of wheat in England (Berkeley 1845). In the United States, first report of losses caused by *P. nodorum* was published in 1898 by Townsend. The disease affects both the quantity and quality of yield (Eyal 1999) and the pathogen is capable of affecting wheat at both seedling and adult stages. Historically, losses up to 50% have been reported, in addition to lower grain quality. The yield losses are highest when flag leaf, F-1 (leaf below flag leaf), and F-2 (leaf below F-1) are infected (Bhathal et al. 2003; Fried et al. 1987; King et al. 1983; McKendry et al. 1995; Shipton et al. 1971). Several review articles have been written over time which underscore the importance of SNB in wheat production (Bergren 1981; Eyal 1999; King et al. 1983;

Lucas et al. 1991; Oliver et al. 2012; Shipton et al. 1971; Solomon 2006; van Ginkel et al. 1999).

#### **1.4 Taxonomy of *Parastagonospora nodorum***

The synonyms of *P. nodorum* are *Depazea nodorum*, *Septoria nodorum*, *Leptosphaeria nodorum*, *Stagonospora nodorum*, and *Phaeosphaeria nodorum* (Quaedvlieg et al. 2013). At the first description, Berkeley (1845) named the fungus *Depazea nodorum*, but later moved it to the *Septoria* genus based on morphological characteristics. Until 1971 (Shipton 1971), the pathogen was known as *Septoria nodorum* (teleomorph *Leptosphaeria nodorum*). In the 1980s, *Septoria nodorum* was classified as *Stagonospora nodorum* based on the length-to-width ratio of conidia, and the teleomorph *L. nodorum* was reclassified as *Phaeosphaeria nodorum* (Cunfer 1997). The length-to-width ratio of *Stagonospora* spp. is less than 10, while it is greater than 10 for *Septoria* spp. In a recent phylogenetic analysis based on the 28S nuclear ribosomal RNA and RPB2 (DNA-directed RNA polymerase II) genes, a new genus name *Parastagonospora* was given to the pathogens causing nodorum blotches of cereals (Quaedvlieg et al. 2013). The pathogen belongs to class Dothideomycetes in phylum Ascomycota, kingdom Fungi (Eyal 1987; Solomon 2006). It is grouped in the order Pleosporales with several devastating plant-pathogenic genera such as *Leptosphaeria*, *Cochliobolus*, *Pyrenophora*, *Alternaria*, *Phoma*, and *Venturia* (Goodwin 2004; Solomon et al. 2006).

## 1.5 Epidemiology of SNB

Primary inoculum in the field can be infected seed (harboring mycelia of the fungus), pycnidiospores (asexual spores) or windborne sexual ascospores from infected wheat debris (Arseniuk et al. 1998; Eyal 1999; Shah et al. 1995). A mark-release-recapture study in Switzerland has shown that sexual recombination happens in the field, and airborne immigrant ascospores can be responsible for primary infection (Sommerhalder et al. 2010). Several other studies also indicate ascospores as a source of primary inoculum (Bathgate and Loughman 2001; Bennett et al. 2007; Sommerhalder et al. 2006). In North Carolina, the presence of ascocarps (sexual structure of ascomycetes) in low frequencies has been reported and could be responsible for epidemic initiation (Cowger and Silva-Rojas 2006). The sexual fruiting body of this fungus is a pseudothecium, which contains asci and plays a major role in the survival of this fungus. Release of ascospores from pseudothecia is highly dependent on weather variables [rainfall >1 mm, temperature above 0 °C, and high (75–95%) relative humidity] (Arseniuk et al. 1998; Bathgate and Loughman 2001).

Studies in New York have shown that initial infection is not solely dependent on immigrant ascospores, and can be caused by seedborne inoculum (Shah et al. 1995; Shah and Bergstrom 2000; Shah et al. 2001). Transmission of the pathogen from seed to coleoptile and the first leaf decreases with increase in temperature (Shah and Bergstrom 2000). It is likely that seed infection plays a relatively bigger role in regions where mean temperature is lower (around 9 °C). Wheat straw infected with *P. nodorum* can also serve as a source of primary

inoculum (Holmes and Colhoun 1975), especially in the case of wheat sown after no-till where wheat debris remains on the soil surface. Secondary spread of *P. nodorum* occurs through rain-splashed pycnidiospores (Shah et al. 2001), which are produced in the asexual fruiting body of the fungus known as the pycnidium (Eyal 1987).

The pathogen can infect all aboveground parts of the plant. Due to its ability to infect nodes of wheat plant, the name “*nodorum*” was assigned as the species epithet. Initial symptoms are small (1–2 mm), dark-brown (chocolate-colored) spots on the leaf. Later, lesions expand and become oval in shape with diffuse yellow halos (Cowger and Weisz 2012). Disease symptoms appear first on the oldest leaves in early spring. Lesions can expand and coalesce, leading to necrosis of the entire leaf. Small (160–210  $\mu\text{m}$  in diameter) fruiting bodies (pycnidia) are formed in the center of mature lesions one to two weeks after infection under high relative humidity (Solomon et al. 2006). Mature pycnidia exude pink-colored cirrhi containing a large number of conidia. With rain-splash, conidia can be dispersed from infected lesions on leaves to spikes, where they cause tan-colored lesions on glumes that are known as glume blotch (Cowger and Weisz 2013). Both conidia and ascospores can germinate and cause infection between 5 and 35 °C (optimum 15 to 25 °C), and penetration can happen directly through the cuticle or opportunistically through stomata (Bergstrom 2010). The optimum temperature for the development of disease symptoms and pycnidial formation is 20 °C (Verreet and Hoffmann 1990).

The rate at which an epidemic spreads is dependent upon latent period, which is defined as the period between inoculation of host tissue and sporulation (Van der Plank

1963). The latent period of *P. nodorum* varies greatly from 6 to 49 days across various studies and is dependent on temperature, moisture, and cultivar (Shearer and Zadoks 1972 & 1974; Stooksbury et al. 1987; Zearfoss et al. 2011). A model has been developed by Zearfoss et al. (2011), quantifying the relationship between *P. nodorum* latent period and temperature under field conditions. The inverse of latent period was linearly related to temperature under outdoor conditions, and was given by this equation (Zearfoss et al. 2011):

$$\frac{1}{\text{Latent Period}} = 0.0072 + 0.0026 \times \text{Mean Temperature}$$

This information can be used in prediction models to incorporate the effect of temperature on latent period and hence on epidemic onset and development. SNB development is also favored by rainstorms, which can cause sudden outbreaks and cause fast vertical spread from lower leaves to upper leaves (Royle et al. 1986).

## 1.6 Management of SNB

*Stagonospora nodorum* blotch is currently managed in wheat using a variety of cultural practices that include crop rotation and tillage ensuring complete burial of residue, as well as the planting of moderately resistant cultivars, use of fungicide-treated seed, and foliar fungicides (Krupinsky et al. 2007; Luke et al. 1983; Milus and Chalkey 1997). While crop rotation and tillage have been shown to reduce the severity of SNB at the end of the season, their effectiveness depends on their widespread adoption because aerial ascospores from adjacent fields may lead to disease development in fields without wheat residue on the soil surface (Bergstrom 2010; Cowger and Silva-Rojas 2006). A few winter wheat cultivars with

partial resistance to SNB are available (Weisz and Cowger 2014) and more breeding efforts are underway to develop resistant varieties for SNB (Cowger and Murphy 2007; Francki 2013). Foliar fungicide sprays are effective in controlling SNB but wheat prices have, historically, been very low and the realized yield response may not be adequate to offset the cost of fungicide treatments (Weisz et al. 2011). The North Carolina Agricultural Chemical Manual recommends multi-site fungicides (e.g. mancozeb); site-specific ones such as strobilurins (e.g. pyraclostrobin, azoxystrobin, and picoxystrobin); and combinations of strobilurin and triazoles (e.g. trifloxystrobin plus prothioconazole) (Koenning 2015).

### **1.7 Necrotrophic effectors or host-selective toxins**

In addition to being an important pathogen of wheat, *P. nodorum* also serves as a model organism for necrotrophic fungal pathogens; its genome was published in 2007 (Hane et al. 2007; Oliver et al. 2012; Solomon et al. 2006). It secretes various host-selective toxins (more recently termed “necrotrophic effectors”) to kill host tissue during colonization (Friesen et al. 2012). Necrotrophic effectors (NE) of *P. nodorum* interact with corresponding sensitivity genes in wheat in an “inverse gene-for-gene” manner (Oliver et al. 2012). Oliver et al. (2012) suggested that, in interactions involving NE and sensitivity genes in wheat, the severity of disease is determined by number and identity of such matches. Thus far, nine such NE-host sensitivity gene interactions have been identified: SnTox1-*Snn1*, SnToxA-*Tsn1*, SnTox2-*Snn2*, SnTox3-*Snn-B1*, SnTox3-*Snn3-D1*, SnTox4-*Snn4*, SnTox5-*Snn5*, SnTox6-*Snn6*, and SnTox7-*Snn7* (Abeysekara et al. 2009; Gao et al. 2015; Friesen et al. 2007; Friesen



et al. 2008; Friesen et al. 2012; Liu et al. 2004a; Liu et al., 2004b; Shi et al. 2015). These interactions occur under field conditions as well (Friesen et al. 2009), and the discoveries have led to increased breeding efforts using marker-assisted selection to find quantitative trait loci associated with SNB resistance in wheat (Oliver et al. 2012). Planting resistant cultivars is one of the most effective ways to manage SNB, and introgression of resistance into commercial wheat lines using these new discoveries is a critical step towards better disease management.

### **1.8 Prediction models for SNB**

The quest to develop prediction models for SNB started in the 1980s, when Tyldesley and Thompson (1980) developed a model that had 71% accuracy in predicting SNB epidemics based on frequency of rainfall in England and Wales. They found that more than 5 days with rainfall greater than 1 mm (“wet days”) in the second half of May and first half of June (stem elongation period) were associated with severe SNB epidemics. Other factors such as sunshine, temperature, and infection level in the previous year’s crop were considered but found to have little effect on SNB. Similar qualitative thresholds were provided by Hansen et al. (1994) in Denmark for both SNB and *Septoria tritici* blotch, where eight days with rainfall  $\geq 1$  mm in a 30-day period starting at stem elongation correlated with disease severity and yield response. Some other studies have also focused on weather variables during the stem elongation period and have not combined weather variables with pre-planting predictors that can influence SNB development (Djurle et al. 1996; Shaw et al.

2008). In 1987, Caristi et al. developed an expert system called EPINFORM to provide estimates of damage caused by SNB and stripe rust in Montana. Their system counted number of infection cycles necessary to cause yield penalty; however, it was assumed that inoculum is present in the field all the time and weather is the only deciding factor in initiating the infection cycle. Quantitative testing of the expert system was not done for SNB due to unavailability of data on yield reduction. A simulation model developed by Djurle and Yuen (1991) found that weather conditions were the most important factors in determining an SNB epidemic, while amount of initial inoculum was found to be less crucial. In general, these modelling efforts have found rainfall to be a significant predictor of end-of-season SNB intensity.

A decision support system developed by Verreet and Hoffmann in 1990 in Germany required monitoring of the wheat crop from the stem-elongation [Zadoks 31 (Zadoks 1974)] to watery-ripe (Zadoks 71) growth stages, but on the positive side, the model was independent of several cultural practices and cultivar resistance level. The model also required the identification and enumeration of pycnidia in order to determine whether thresholds had been met. At different growth stages, specific leaves were designated as test leaves; e.g., top leaves (flag leaf and F-1) were the test leaves during later growth stages (Zadoks 71). At each growth stage, the authors suggested examining 30 to 40 leaves, and on an average if there were >1 pycnidia per test leaf, a fungicide spray was recommended (Verreet and Hoffmann 1990). In 2000, backpropagation models were developed to classify infection periods of *P. nodorum* with correct classification of 84% validation cases (De Wolf

and Franc 2000). However, their study did not answer questions such as when to start weather monitoring in the season, what would be the effect of cultural practices such as crop rotation and tillage, and how many infection periods would be considered as the threshold before a fungicide spray is warranted.

### **1.9 Justification and objectives of study**

Minimum tillage has gained attention because of its benefits, such as reduced soil erosion and runoff, increased water use efficiency, and improved soil physical properties (Shaver et al. 2013). In the case of wheat, this practice has also led to more primary *P. nodorum* inoculum being present in the field. This may lead to early SNB onset and more fungicide sprays for adequate disease control. Given the trend toward reduced tillage, it is critical to determine the relationship between percent wheat residue cover on the ground at planting and in-season SNB severity. Given this relationship, the amount of debris on the ground can act as a predictor variable in a prediction model for this disease.

Besides establishing the relationship between wheat debris ground-cover and SNB severity, another goal for this project is to develop a pre-planting risk assessment for SNB of wheat. For the latter goal, a suite of statistical approaches will be used. They include multiple regression and machine learning methods such as neural networks, classification and regression tree, and random forest. Past studies indicate an association between SNB severity and pre-planting variables such as wheat residue on ground, resistance level of wheat variety, location of the field, the type of tillage, seed treatment, and seed rate (Eyal 1999;

Solomon 2006; Weisz 2012). Wheat residue in the field generates elevated inoculum potential and may result in early development of epidemics. Susceptible cultivars are likely to develop high levels of disease. Field location is also important because the disease tends to occur in the western half (mountains and Piedmont) of NC, owing to the prevalence of no-till in that area. The type of tillage also greatly influences the amount of wheat residue on the soil surface. Seed treatment can affect seedborne inoculum of the fungus, and lower seeding and/or emergence rates can lead to less dense stands, encouraging more effective splash dispersal of the pathogen. A pre-planting risk assessment for SNB will be a valuable decision support tool for efficient and profitable management of the disease, and especially can help with decisions that need to be made prior to planting.

After planting, if the disease appears, the only available management option in addition to cultivar resistance is fungicide application. Historically, wheat prices have been very low (Anonymous 2015) and many growers cannot afford unprofitable fungicide sprays. It has already been reported that application of strobilurin or triazole fungicides on the basis of wheat growth stage, and in the absence of fungal diseases such as SNB, is not profitable (Wegulo et al. 2011; Weisz et al. 2011). Further, applying fungicides only when necessary reduces the rate at which pathogens develop resistance to fungicides. Resistance to strobilurins in *P. nodorum* has already been reported (Morzfeld et al. 2004). In Europe, strobilurin resistance in a similar pathogen, *S. tritici*, has caused serious problems (Torriani et al. 2009). Therefore, slowing down the development of resistance to available effective fungicides is an important reason to apply fungicides only when necessary.

One approach to increasing fungicide efficiency is to develop models that evaluate the effect of the above-described pre-planting factors (tillage, crop rotation, wheat residue, cultivar resistance, seed treatment, and seed rate) and in-season weather in order to predict the time of SNB onset. As reviewed above, most of these factors are known to have an effect on end-of-season disease severity, but their effect on time of epidemic onset is little understood. This type of model will be especially useful if predictions can be made early in the season when there is still time to apply fungicides to keep SNB levels low. Given the polycyclic nature of the disease, time of disease onset also affects the extent and magnitude of epidemic development during the season. For example, if disease onset occurs early in the season, there is a higher likelihood that this will result in a more severe epidemic. Due to differences in resistance level of cultivars, inoculum potential, fungicide seed treatment and plant density, time of disease onset will vary considerably from one season to the next. Thus, it is important to identify and explain the reasons for the variation in disease onset. The ability to explain this variation is crucial for successful disease prediction, either with explicit prediction of epidemic onset time or with a negative prognosis that forecasts a disease-free period. Accurate prediction of epidemic onset will be useful in reducing the scouting and disease observation required to determine if fungicide sprays are warranted.

Therefore, based on the above considerations, the specific objectives of this dissertation are to:

1. Quantify the effects of wheat residue on severity of *Stagonospora nodorum* blotch and yield in winter wheat (Chapter 2).

2. Assess the pre-planting risk of *Stagonospora nodorum* blotch of winter wheat (Chapter 3).
3. Identify the factors affecting the onset of *Stagonospora nodorum* blotch in winter wheat (Chapter 4).

### Literature Cited

- Abeysekara, N. S., Friesen, T. L., Keller, B., and Faris, J. D. 2009. Identification and characterization of a novel host-toxin interaction in the wheat-*Stagonospora nodorum* pathosystem. *Theor. Appl. Genet.* 120:117–126.
- Anonymous. 2010. 2010 State Agriculture Review, North Carolina.
- Anonymous. 2014. 2014 Crop Summary- News Release, USDA North Carolina Field Office. Released 13 January 2015. <http://www.ncagr.gov/stats/release/croprelease01.pdf>. Accessed 21 September 2015.
- Anonymous. 2015a. USDA, Economic Research Service, Wheat Data. Online: <http://www.ers.usda.gov/data-products/wheat-data.aspx>. Accessed 15 September 2015.
- Anonymous, 2015b. USDA, Agricultural Research Service -Eastern United States Septoria Nursery. 2015. <http://www.ars.usda.gov/Main/docs.htm?docid=8419&page=8>. Accessed on April 16, 2015.
- Arseniuk, E., Góral, T., and Scharen, A. L. 1998. Seasonal patterns of spore dispersal of *Phaeosphaeria* spp. and *Stagonospora* spp. *Plant Dis.* 82:187–194.
- Bathgate, J. A., and Loughman, R. 2001. Ascospores are a source of inoculum of *Phaeosphaeria nodorum*, *P. avenaria* f. sp. *avenaria* and *Mycosphaerella graminicola* in Western Australia. *Australas. Plant Pathol.* 30:317–322.
- Bergren, B. 1981. Glume blotch (*Septoria nodorum*) and leaf blotch (*Septoria tritici*) on wheat - a review of the literature. *Plant Protection Reports Agriculture* 19. Institute for plant and forestry protection. Swedish University of Agricultural Sciences.

- Berkeley, M. J. 1845. Disease in the wheat crop. *Gard. Chron.* 35:601.
- Bhathal, J. S., Loughman, R., and Speijers, J. 2003. Yield reduction in wheat in relation to leaf disease from yellow (tan) spot and septoria nodorum blotch. *Eur. J. Plant Pathol.* 109: 435–443.
- Bennett, R. S., Milgroom, M. G., Sainudiin, R., Cunfer, B. M., and Bergstrom, G. C. 2007. Relative contribution of seed-transmitted inoculum to foliar populations of *Phaeosphaeria nodorum*. *Phytopathology* 97:584–591.
- Bergstrom, G. C. 2010. Stagonospora nodorum blotch and Stagonospora avenae blotch. Pages 75–77 in: Compendium of wheat diseases and pests. W. W. Bockus, R. L. Bowden, R. M. Hunger, W. L. Morrill, T. D. Murray, and R. W. Smiley, eds. The American Phytopathological Society, St. Paul, Minnesota.
- Braun, H. J., Atlin, G., and Payne, T. 2010. Multi-location testing as a tool to identify plant response to global climate change. Pages 115-138 in: *Climate Change and Crop Production*. C.R.P. Reynolds, ed. CABI, London, UK.
- Caristi, J., Scharen, A. L., Sharp, E. L., and Sands, D. C. 1987. Development and preliminary testing of EPINFORM, an expert system for predicting wheat disease epidemics. *Plant Dis.* 71:1147–1150.
- Cowger, C., and Murphy, J. P. 2007. Artificial inoculation of wheat for selecting resistance to Stagonospora nodorum blotch. *Plant Dis.* 91:539–545.



- Cowger, C., and Silva-Rojas, H. V. 2006. Frequency of *Phaeosphaeria nodorum*, the sexual stage of *Stagonospora nodorum*, on winter wheat in North Carolina. *Phytopathology* 96:860–866.
- Cowger, C., and Weisz, R. 2013. Small grain disease management. Pages 58–70 in: Small grain production guide. R. Weisz, ed. North Carolina Cooperative Extension Service, Publication No. AG-580 (Revised March 2013).
- Cunfer, B. M. 1997. Taxonomy and nomenclature of *Septoria* and *Stagonospora* species on small grain cereals. *Plant Dis.* 81:427–428.
- De Wolf, E. D., and Franc, L. J. 2000. Neural network classification of tan spot and *Stagonospora* blotch infection periods in a wheat field environment. *Phytopathology* 90:108–113.
- Djurle, A., and Yuen, J. 1991. A simulation model for *Septoria nodorum* in winter wheat. *Agric. Syst.* 37:193–218.
- Djurle, A., Ekbom, B., and Yuen, J. E. 1996. The relationship of leaf wetness duration and disease progress of glume blotch, caused by *Stagonospora nodorum*, in winter wheat to standard weather data. *Eur. J. Plant Pathol.* 102:9–20.
- Dixon, J., Braun, H. -J., Kosina, P., and Crouch, J. (eds.). 2009. *Wheat Facts and Futures 2009*. Mexico, D.F.: CIMMYT.
- Eyal, Z. 1999. The *Septoria tritici* and *Stagonospora nodorum* blotch diseases of wheat. *Eur. J. Plant Pathol.* 105:629–641.

- Eyal, Z., Scharen, A. L., Prescott, J. M., and van Ginkel, M. 1987. The Septoria Diseases of Wheat: Concepts and Methods of Disease Management. Mexico, D. F.: CIMMYT.
- Francki, M. G. 2013. Improving *Stagonospora nodorum* resistance in wheat: a review. Crop Sci. 53:355–365.
- Fried, P. M., and Meister, E. 1987. Inheritance of leaf and head resistance of winter wheat to *Septoria nodorum* in a diallel cross. Phytopathology 77:1371–1375.
- Friesen, T. L., Meinhardt, S. W., and Faris, J. D. 2007. The *Stagonospora nodorum* wheat pathosystem involves multiple proteinaceous host-selective toxins and corresponding host sensitivity genes that interact in an inverse gene-for-gene manner. Plant J. 51:681–692.
- Friesen, T. L., Zhang, Z., Solomon, P. S., Oliver, R. P., and Faris, J. D. 2008. Characterization of the interaction of a novel *Stagonospora nodorum* host-selective toxin with a wheat susceptibility gene. Plant Physiol. 146:682–693.
- Friesen, T. L., Chu, C. -G., Liu, Z. H., Xu, S. S., and Faris, J. D. 2009. Host-selective toxins produced by *Stagonospora nodorum* confer disease susceptibility in adult wheat plants under field conditions. Theor. Appl. Genet. 118:1489–1497.
- Friesen, T. L., Chu, C., Xu, S. S., and Faris, J. D. 2012. SnTox5-*Snn5*: a novel *Stagonospora nodorum* effector-wheat gene interaction and its relationship with the SnToxA-*Tsn1* and SnTox3-*Tsn3-B1* interactions. Mol. Plant Pathol. 13:1101–1109.
- Gao, Y., Faris, J. D., Liu, Z., Kim, Y. M., Syme, R. A., Oliver, R. P., Xu, S. S., and Friesen, T. L. 2015. Identification and characterization of the SnTox6-*Snn6* interaction in the *Parastagonospora nodorum*-wheat pathosystem. MPMI 28:615–625.

- Goodwin, S. B. 2004. Minimum phylogenetic coverage: An additional criterion to guide the selection of microbial pathogens for initial genomic sequencing efforts. *Phytopathology* 94:800–804.
- Hane, J. K., Lowe, R. G., Solomon, P. S., Tan, K. C., Schoch CL, et al. 2007. Dothideomycete-plant interactions illuminated by genome sequencing and EST analysis of the wheat pathogen *Stagonospora nodorum*. *Plant Cell* 19:3347–3368.
- Hansen, J. G., Secher, B. J. M., Jørgensen, L. N. and Welling, B. 1994. Thresholds for control of *Septoria* spp. in winter wheat based on precipitation and growth stage. *Plant Pathol.* 43:183–189.
- Holmes, S. J. I., and Colhoun, J. 1975. Straw-borne inoculum of *Septoria nodorum* and *S. tritici* in relation to incidence of disease on wheat plants. *Plant Pathol.* 24:63–66.
- King, J. E., Cook, R. J., and Melville, S. C. 1983. A review of *Septoria* diseases of wheat and barley. *Ann. Appl. Biol.* 103:345–373.
- Koenning, S. R. 2015. Foliar fungicides for wheat leaf disease control. In: 2015 North Carolina Agricultural Chemical Manual, Chapter X—Disease Control. 414p. Online: <http://content.ces.ncsu.edu/disease-control.pdf> . Accessed 17 September 2015.
- Krupinsky, J. M., Tanaka, D. L., Merrill, S. D., Liebig, M. A., Lares, M. T., and Hanson, J. D. 2007. Crop sequence effects on leaf spot diseases of no-till spring wheat. *Agron. J.* 99:912–920.
- Liu, Z. H., Faris, J. D., Meinhardt, S. W., Ali, S., Rasmussen, J. B., and Friesen, T. L. 2004a. Genetic and physical mapping of a gene conditioning sensitivity in wheat to a partially

- purified host-selective toxin produced by *Stagonospora nodorum*. *Phytopathology* 94:1056–1060.
- Liu, Z. H., Friesen, T. L., Rasmussen, J. B., Ali, S., Meinhardt, S. W., and Faris, J. D. 2004b. Quantitative trait loci analysis and mapping of seedling resistance to *Stagonospora nodorum* leaf blotch in wheat. *Phytopathology* 94:1061–1067.
- Lucas, J. A., Bowyer, P., Anderson, H. M., eds. 1999. *Septoria on Cereals: A Study of Pathosystems*. Wallingford, UK: CABI
- Luke, H. H., Pfahler, P. L., and Barnett, R. D. 1983. Control of *Septoria nodorum* on wheat with crop rotation and seed treatment. *Plant Dis.* 67:949–951.
- McKendry, A. L., Henke, G. E., and Finney, P. L. 1995. Effects of *Septoria* leaf blotch on soft red winter wheat milling and baking quality. *Cereal Chem.* 72:142–146.
- Milus, E. A., and Chalkley, D. B. 1997. Effect of previous crop, seedborne inoculum, and fungicides on development of *Stagonospora* blotch. *Plant Dis.* 81:1279–1283.
- Morzfeld, J., Koopman, B., and Hoppe, H. -H. 2004. Development of fungicide resistance of wheat and barley pathogens against strobilurins: A methodological approach. *J. Plant Dis. Protect.* 111:105–114.
- Oerke, E.-C. 2006. Crop losses to pests. *J. Agric. Sci.* 144:31–43.
- Oliver, R. P., Friesen, T. L., Faris, J. D., and Solomon, P. S. 2012. *Stagonospora nodorum*: From pathology to genomics and host resistance. *Annu. Rev. Phytopathol.* 50:23–43.
- Quaedvlieg, W., Verkley, G. J. M., Shin, H. -D., Barreto, R. W., Alfenas, A. C., Swart, W. J., Groenewald, J. Z. and Crous, P. W. 2013. Sizing up *Septoria*. *Stud. Mycol.* 75:307–390.

- Royle, D. J., Shaw, M. W., and Cook, R. J. 1986. Patterns of development of *Septoria nodorum* and *S. tritici* in some winter wheat crops in Western Europe, 1981–1983. *Plant Pathol.* 35:466–476.
- Shah, D. A., Bergstrom, G. C., and Ueng, P. P. 2001. Foci of *Stagonospora nodorum* blotch in winter wheat before canopy development. *Phytopathology* 91:642–647.
- Shah, D., Bergstrom, G. C., and Ueng, P. P. 1995. Initiation of *Septoria nodorum* blotch epidemics in winter wheat by seedborne *Stagonospora nodorum*. *Phytopathology* 85:452–457.
- Shah, D. A., and Bergstrom, G. C. 2000. Temperature dependent seed transmission of *Stagonospora nodorum* in wheat. *Eur. J. Plant Pathol.* 106:837–842.
- Shaver, T. M., Peterson, G. A., Ahuja, L. R., and Westfall, D. J. 2013. Soil sorptivity enhancement with crop residue accumulation in semi-arid dryland no-till agroecosystems. *Geoderma* 192:254–258.
- Shaw, M. W., Bearchell, S. J., Fitt, B. D. L., Fraaije, B. A. 2008. Long-term relationships between environment and abundance in wheat of *Phaeosphaeria nodorum* and *Mycosphaerella graminicola*. *New Phytol.* 177:229–238.
- Shearer, B. L., and Zadoks, J. C. 1972. The latent period of *Septoria nodorum* in wheat. 1. The effect of temperature and moisture treatments under controlled conditions. *Neth. J. Plant Pathol.* 78:231–241.

- Shearer, B. L., and Zadoks, J. C. 1974. The latent period of *Septoria nodorum* in wheat. 2. The effect of temperature and moisture treatments under field conditions. *Neth. J. Plant Pathol.* 78:48–60.
- Shi, G., Friesen, T. L., Saini, J., Xu, S. S., Rasmussen, J. B., and Faris, J. D. 2015. The wheat *Snn7* gene confers susceptibility on recognition of the *Parastagonospora nodorum* necrotrophic effector SnTox7. *Plant Genome* 8: doi: 10.3835/plantgenome2015.02.0007
- Shipton, W. A., Boyd, W. R. J., Rosielle, A. A., and Shearer, B. I. 1971. The common septoria diseases of wheat. *Bot. Rev.* 37:231–262.
- Solomon, P. S., Lowe, R. G. T., Tan, K., Waters, O. D. C., and Oliver, R. P. 2006. *Stagonospora nodorum*: cause of stagonospora nodorum blotch of wheat. *Mol. Plant Pathol.* 7:147–156.
- Sommerhalder, R. J., McDonald, B. A., and Zhan, J. 2006. The frequencies and spatial distribution of mating types in *Stagonospora nodorum* are consistent with recurring sexual reproduction. *Phytopathology* 96:234–239.
- Sommerhalder, R. J., McDonald, B. A., Mascher, F., and Zhan, J. 2010. Sexual recombinants make a significant contribution to epidemics caused by the wheat pathogen *Phaeosphaeria nodorum*. *Phytopathology* 100:855–862.
- Stooksbury, D. E., Johnson, J. E., and Cunfer, B. M. 1987. Incubation period and latent period of wheat for resistance to *Leptosphaeria nodorum*. *Plant Dis.* 71:1109–1112.

- Torriani, S. F. F., Brunner, P. C., McDonald, B. A., and Sierotzki, H. 2009. QoI resistance emerged independently at least 4 times in European populations of *Mycosphaerella graminicola*. *Pest Manag. Sci.* 65:155–162.
- Townsend, C. O. 1898. Some important wheat diseases. Maryland Agric. Exp. Sta. Bull. 58:123–127.
- Tyldesley, J. B. and Thompson, N. 1980. Forecasting *Septoria nodorum* on winter wheat in England and Wales. *Plant Pathol.* 19:9–20.
- Van der Plank, J. E. 1963. Plant diseases: epidemics and control. Academic Press, New York, 349 pp.
- van Ginkel, M., McNab, A., and Krupinsky, J., eds. 1999. *Septoria and Stagonospora Diseases of Cereals: A Compilation of Global Research*. Mexico, D.F.:CIMMYT.
- Verreet, J. A., and Hoffmann, G. M. 1990. A biologically oriented threshold decision model for control of epidemics of *Septoria nodorum* in wheat. *Plant Dis.* 74:731–738.
- Wegulo, S. N., Zwingman, M. V., Breathnach, J. A., and Baenziger, P. S. 2011. Economic returns from fungicide application to control foliar fungal diseases in winter wheat. *Crop Protect.* 30:685–692.
- Weisz, R. 2013. Small Grain Production Guide. North Carolina Cooperative Extension Service, Publication No. AG-580 (Revised March 2013).
- Weisz, R., and Cowger, C. 2014. 2014 Wheat variety performance & recommendations. Smartgrains- The Small Grains Fact Sheet. No. 37. Online:

<http://www.smallgrains.ncsu.edu/Misc/VarietySelection.pdf> . Accessed 17 September 2015.

Weisz, R., Cowger, C., Ambrose, G., and Gardner, A. 2011. Multiple mid-Atlantic field experiments show no economic benefit to fungicide application when fungal disease is absent in winter wheat. *Phytopathology* 101:323–333.

Zadoks, J. C., Chang, T. T., and Konzak, C. F. 1974. A decimal code for the growth stages of cereals. *Weed Res.* 14:415–421.

Zearfoss, A.D., Cowger, C., and Ojiambo, P.S. 2011. A degree-day model for the latent period of *Stagonospora nodorum* blotch in wheat. *Plant Dis.* 95:561–567.



## CHAPTER 2

### **Quantifying the effects of wheat residue on severity of *Stagonospora nodorum* blotch and yield in winter wheat**

Published in *Phytopathology*:

Reproduced, by permission, from The American Phytopathological Society. Permission  
Granted on 10/14/2015

Citation:

Mehra, L. K., Cowger, C., Weisz, R., and Ojiambo, P. S. 2015. Quantifying the effects of  
wheat residue on severity of *Stagonospora nodorum* blotch and yield in winter wheat.

*Phytopathology* 105:1417-1426. <http://dx.doi.org/10.1094/PHYTO-03-15-0080-R>.

## Abstract

*Stagonospora nodorum* blotch (SNB), caused by the fungus *Parastagonospora nodorum*, is a major disease of wheat (*Triticum aestivum*). Residue from a previously infected wheat crop can be an important source of initial inoculum, but the effects of infected residue on disease severity and yield have not previously been quantified. Experiments were conducted in Raleigh and Salisbury, North Carolina, in 2012, 2013, and 2014 using the moderately susceptible winter wheat cultivar DG Shirley. In 2014, the highly susceptible cultivar DG 9012 was added to the experiment and the study was conducted at an additional site in Tyner, North Carolina. Four (2012) or six (2013 and 2014) wheat residue treatments were applied in the field in a randomized complete block design with five replicates. Treatments in 2012 were 0, 30, 60, and 90% residue coverage of the soil surface, while 10 and 20% residue treatments were added in 2013 and 2014. Across site-years, disease severity ranged from 0 to 50% and increased nonlinearly ( $P < 0.05$ ) as residue level increased, with a rapid rise to an upper limit and showing little change in severity above 20 to 30% soil surface coverage. Residue coverage had a significant ( $P < 0.05$ ) effect on disease severity in all site-years. The effect of residue coverage on yield was only significant ( $P < 0.05$ ) for DG Shirley at Raleigh and Salisbury in 2012 and for DG 9012 at Salisbury in 2014. Similarly, residue coverage significantly ( $P < 0.05$ ) affected thousand-kernel weight only of DG 9012 in 2014 at Raleigh and Salisbury. Our results showed that when wheat residue was sparse, small additions to residue density produced greater increases in SNB than when residue was abundant. SNB only led to effects on yield and test weight in the most disease-conducive

environments, suggesting that the economic threshold for the disease may be higher than previously assumed and warrants review.

## **Introduction**

*Stagonospora nodorum* blotch (SNB), caused by the ascomycete fungus *Parastagonospora nodorum* (Quaedvlieg et al. 2013), occurs in wheat-growing areas worldwide, but is more prevalent in areas with warm and moist conditions such as the southeastern United States, parts of Europe, southern Brazil, and Australia (Eyal et al. 1987; Murray and Brennan 2009). Temperatures of 20 to 25°C with high relative humidity (>75%) and frequent rainfall favor the development of the disease (Verreet and Hoffmann 1990). The pathogen is capable of infecting wheat (*Triticum aestivum* L.) at both the seedling and adult stages. Typical disease symptoms include a necrotic leaf blotch as well as discoloration of the head, which is usually known as glume blotch. After about a week following infection and under high humidity, pycnidia begin to form in the necrotic lesions.

SNB affects both the quantity and quality of wheat grain (Eyal 1999). Reduction in yield can occur due to reduced photosynthetic area of the upper leaves, which leads to a decrease in the level of carbohydrates available for seed fill (Verreet and Hoffmann 1989), and hence, lower thousand-kernel weight (Gilbert and Tekauz 1993). Yield losses up to 50% due to SNB have been reported in Europe and the United States (Bhathal et al. 2003; Fried et al. 1987; King et al. 1983; McKendry et al. 1995; Shipton et al. 1971), while in Australia, the disease has been reported to account for 12% of yield losses in wheat (Murray and Brennan

2009). SNB reduced test weight, which is a measure of grain quality, by about 6% in untreated compared to fungicide treated plots in Arkansas, United States (Milus and Chalkley 1997). The milling quality of grain can also be reduced by infection of the developing seed, causing shriveling, discoloration, and reduced flour yield (Bhathal et al. 2003; Gilbert and Tekauz 1993).

The primary sources of inoculum for initial infection in the field can include infected seed harboring mycelia of the fungus, pycnidiospores or conidia (asexual spores) from infected wheat debris, or windborne sexual ascospores (Arseniuk et al. 1998; Eyal 1999; Shah et al. 1995). Pseudothecia, which are the sexual fruiting bodies of *P. nodorum*, contain asci and play a major role in the survival of the pathogen. A mark-release-recapture study in Switzerland showed that sexual recombination occurs in the field, and airborne immigrant ascospores can be responsible for primary infection (Sommerhalder et al. 2010). Several other studies have also indicated that ascospores are important as a source of primary inoculum (Bathgate and Loughman 2001; Bennett et al. 2007; Sommerhalder et al. 2006). In North Carolina, ascospores could play a role in initiating SNB epidemics (Cowger and Silva-Rojas 2006), while both ascospores and seedborne inoculum are involved in initial plant infections in New York (Shah and Bergstrom 2000; Shah et al. 1995; Shah et al. 2001).

*Stagonospora nodorum* blotch is managed through cultural practices, fungicide application, and planting resistant cultivars (Engle et al. 2006; Luke et al. 1983; Milus and Chalkley 1997). The primary cultural management practices include crop rotation and burying infected wheat residue by tillage to reduce the survival of the pathogen from one

season to the next (Luke et al. 1983; Milus and Chalkley 1997). Wheat straw infected with *P. nodorum* can serve as a major source of primary inoculum (Holmes and Colhoun 1975), especially in the case of wheat planted no-till with wheat debris from the previous crop remaining on the soil surface.

The widespread adoption of conservation tillage to reduce soil erosion and conserve moisture (Shaver et al. 2013) has resulted in an increase of crop residue on the soil surface, hence availability of more primary inoculum, and has been linked to an increase in the severity of SNB in the eastern United States (Shaner and Buechley 1995). The effect of tillage on the epidemiology of a disease is complex, but the impact on disease intensity generally relates to the amount of residue remaining after planting (Boosalis et al. 1986). Although the amount of wheat debris on the ground has been shown to influence severity of SNB (Luke et al. 1983; Milus and Chalkley 1997), a quantitative relationship between the level of soil-surface residue and SNB severity is still lacking. These previous studies (Luke et al. 1983; Milus and Chalkley 1997) have also not related levels of residue to yield in wheat. For example, Luke et al. (1983) reported that residue from previous crops was an important inoculum source that influenced the level of SNB, but they neither specified the amount of residue in their plots nor quantified the relationship between wheat residue and severity of SNB. Similarly, Milus and Chalkley (1997) reported a higher incidence and severity of SNB in plots infested with wheat residue than in clean plots, but they did not specify the amount of residue in the infested plots nor relate residue levels to disease intensity.

Pre-planting management factors that influence the risk of SNB include crop residue, crop rotation, disease history, cultivar selection, seed treatment, and location of field. Quantifying the effects of residue on the severity of SNB and yield is fundamental in developing models that can be used to make SNB management decisions prior to planting of a wheat crop. Therefore, the specific objectives for this study were to quantify the effect of amount of wheat residue on severity of SNB, and determine the impact of wheat residue on yield and yield components in soft red winter wheat. This work was part of a larger study designed to determine how pre-planting disease management factors influence the risk of SNB on winter wheat during the growing season.

## **Materials and Methods**

**Field sites and wheat cultivars.** Field experiments were conducted during the 2011-2012, 2012-2013 and 2013-2014 growing seasons (hereafter referred to as 2012, 2013, and 2014, respectively) at the Lake Wheeler Road Field Laboratory in Raleigh and at the Piedmont Research Station in Salisbury, North Carolina. In 2014, experiments were also conducted at a grower's field in Tyner, North Carolina. All field plots were prepared using conventional tillage and within each location, experiments were conducted in different field plots every year. The previous crops in the field before conducting the experiments were as follows: fallow (no crop) at Raleigh, corn at Salisbury, and soybean at Tyner. The cultivar Dyna-Gro Shirley (hereafter referred to as DG Shirley), which is moderately susceptible to *P. nodorum*, was used at all sites due to its resistance to other wheat diseases such as powdery

mildew and leaf rust that are prevalent in the region. In addition, a highly susceptible cultivar DG 9012 was used in the 2014 experiments to determine whether results obtained on DG Shirley (in 2012 and 2013) were broadly similar to those of a more susceptible cultivar. DG 9012 is moderately resistant to leaf rust but moderately susceptible to powdery mildew.

Certified seed was used in this study and seed for the two cultivars was purchased from commercial seed suppliers in eastern North Carolina. All production practices, such as seeding rate, seed treatment, and planting date, were performed according to the Small Grain Production Guide for North Carolina (Weisz 2013). Prior to planting, seed was treated with Vitavax (Chemtura, Australia), a product that is moderately effective in controlling *P. nodorum* in seed. Additionally, three replicate samples of 20 kernels were arbitrarily selected and plated on SNAW (a selective medium), and as described below were found to be free of *P. nodorum*. Seed was planted using a research plot planter with 0.20-m row spacing. Experimental plots were 1.0 m wide and ranged from 5.5 m to 6.4 m in length, except in Salisbury in 2012, where the plots were 7.3 m long. To minimize interplot interference, each plot was surrounded by buffer plots of the same size as treatment plots. In 2012, buffer plots were planted with wheat cultivar NC-Neuse, which is moderately resistant to *P. nodorum*, while barley, which is a non-host of *P. nodorum*, was planted in buffer plots in 2013 and 2014.

Weather data in Raleigh and Salisbury were acquired using standard automated weather stations located less than 1.5 km away from the field plots. Temperature and relative humidity at Tyner were recorded using portable data loggers (WatchDog, Spectrum

Technologies, Plainfield, IL), while rainfall was monitored from the nearest automated national weather station. Typically, weather data were captured from late February until the last disease assessment date. When necessary, irrigation was conducted at the Raleigh and Salisbury sites during dry periods to promote disease development, and the amount of irrigation water was recorded.

**Wheat residue treatments and experimental design.** In 2012, four wheat residue treatments were tested, namely 0, 30, 60, and 90% of plot area covered with residue. The 0% treatment corresponded to conventional tillage representing complete burial of wheat residue; 30 and 60% residue treatments simulated gradations of minimum tillage; and the 90% residue treatment simulated no-tillage (Stubbs et al. 2004). In 2013 and 2014, residue levels of 10 and 20% soil surface coverage were added to the experiment to better investigate the low end of the range, for a total of six treatments. Wheat residue used in each year was obtained from a single source, with a different commercial wheat grower providing the residue used in each year. The residue was obtained from Rowan (2012), Forsyth (2013) and Stokes (2014) counties in the Piedmont region of North Carolina, where SNB epidemics occur frequently, and in each case the residue was baled immediately after harvest of the crop. The baled residue was stored under cover and then transported to the experimental sites for application to plots in mid-winter. In all cases, SNB was present in the previous crop and the presence of *P. nodorum* pycnidia was confirmed on the straw by microscopic examination prior to application of treatments.



The percentage of residue cover in a plot for each treatment was estimated as follows. First, dry weights of residue needed to provide 95 to 100% soil surface coverage were established, using the line transect method (Wollenhaupt and Pingry 1991). From this relationship, the equivalent dry weight of residue to be applied for each treatment was then determined. To avoid smothering of seedlings, residue was applied to plots between mid-January and mid-February when the wheat was at growth stage Zadoks GS 21 to 23 (main stem and one to three tillers, Zadoks et al. 1974). Residue application in *Septoria* screening nurseries has shown that the application of residue at planting or before GS 21 results in patchy stands (C. Cowger, *personal observation*). At each site, the experiment was laid out in a randomized complete block design, with five replicates of each residue treatment.

**Disease assessment.** To quantify disease severity, the percentage of leaf area showing symptoms of SNB was estimated on a whole-plot basis. Briefly, disease severity in each plot was assessed in five distinct sections (~to 0.1 m<sup>2</sup>) spread across the entire plot. The percent leaf area with SNB symptoms in each section was estimated visually and the mean leaf area across all the five sections was the overall disease severity for the plot. The lowest leaves that senesced due to maturity or abiotic factors were not included in the visual assessment. Typically, the first disease assessment was recorded when SNB lesions at least 1 cm in length were visible on lower leaves. This first assessment was between Zadoks GS 37 (flag leaf visible) and GS 45 (swollen boot) growth stages. Disease was assessed every 2 weeks for a total of three to four assessment records depending on the intensity and

progression of the disease. Final disease severity was recorded around Zadoks GS 85 (soft dough) when the flag leaf and the F-1 leaf started to senesce. The same evaluator conducted all assessments of disease severity at all locations.

**Harvest and grain assay.** In each year, grain from all plots was harvested at plant maturity using a small-plot research combine. Grain moisture and test weight were then determined using a grain analysis computer (GAC 2100, Dickey-John Corp., Auburn, IL). Subsequently, grain weight from each treatment plot was used to determine the corresponding yield after adjusting the grain weight to moisture content of 13.5% (Cowger et al. 2010). In the 2013 and 2014 trials, one thousand kernels were also randomly selected from each treatment plot and weighed to determine the thousand-kernel weight. Thousand-kernel weight is a measure of seed size, and is thus related to both yield and test weight.

To determine the incidence of kernels colonized by *P. nodorum* in grain, a sample of 20 arbitrarily selected kernels from each replicate was plated on SNAW (Manandhar and Cunfer 1991). Two 100 mm × 15 mm Petri dishes, each containing 10 kernels, were used per replicate for each treatment, and there were five replicates per treatment. Kernels colonized by *P. nodorum* were assessed 7 days after plating under a 15-watt near-ultraviolet light (model TL-D BLB, Philips, Somerset, NJ). *P. nodorum* produces metabolites in oxgall (an ingredient of SNAW medium) that fluoresce under near-ultraviolet light (Manandhar and Cunfer 1991). The percentage of infected kernels was calculated as the number of colonized kernels divided by the total number of plated kernels and multiplied by 100.

**Data analysis.** Disease severity over time was used to calculate the area under the disease progress curve (AUDPC) as described by Campbell and Madden (1990). To allow for direct comparison of AUDPC values obtained from different epidemic durations at different site and years, relative AUDPC (RAUDPC) was calculated as  $RAUDPC = AUDPC/\Delta t$ , where  $\Delta t$  = time between the first and last disease assessment.

The effects of residue level on maximum disease severity, RAUDPC, yield, test weight, thousand-kernel weight and incidence of colonized kernels were determined by analysis of variance using the GLIMMIX procedure in SAS 9.3 (SAS Institute, Cary, NC). Due to differences in the number of treatments and cultivars used in different years, the analysis was conducted separately for each site and year. Replicates were treated as blocks and specified as a random effect, while treatments were specified as a fixed effect in the model. Further, to account for possible extraneous variation due to trends in soil moisture or fertility that could be associated with rows and columns in the plots, a row error model (Kehel et al. 2010) was implemented within the GLIMMIX procedure to separate independent random effects associated with rows of the trial and the independent plot error. The row error model was used to increase accuracy and precision in the estimates of the effect of residue treatment on yield. The random row error model and the base model were compared using the Akaike information criterion (Kehel et al. 2010). Treatment means were separated using Fisher's least significant difference (LSD) test for variables where the treatment effect was significant in the analysis of variance.

Orthogonal polynomials were also calculated where residue levels significantly affected disease severity or yield to determine linear, quadratic and cubic changes in these responses as residue level increased. Given that residue levels in the 2013 and 2014 trials were not equally spaced, the ORPOL function in PROC IML was used to obtain the appropriate coefficients for the CONTRAST statement in the GLIMMIX procedure. Following preliminary analyses to determine the relationship between RAUDPC and residue level, data were fitted to a 2-parameter logarithmic model of the form,  $y = a + b \cdot \ln(x + 1)$ , where  $y = \text{RAUDPC}$ ,  $x = \text{residue level}$ . In this model, the parameter  $a$  is the intercept and  $b$  is the rate parameter. The model was fitted to the disease data by linear regression using PROC REG in SAS. Goodness-of-fit of the model was assessed based on correlation of observed with predicted  $y$ -values and statistical significance of the overall model.

## Results

**Weather variables.** In this section, the conduciveness of temperature, relative humidity, and rainfall (Bergstrom 2010; Eyal et al. 1987) is compared among years for the epidemic period, defined here as the interval from two weeks prior to disease onset until the final disease assessment. Temperature, relative humidity and rainfall were more favorable for disease development in 2012 than in 2013 or 2014. During the epidemic at Raleigh, mean daily temperatures ranged from 4 to 26°C, 5 to 24°C and 1 to 26°C, in 2012, 2013 and 2014, respectively (data not shown). The number of days with average daily temperature between 20 and 25°C, the optimum range for disease development, was also higher in 2012 than in

2013 at Raleigh. A similar trend in days with conducive temperatures was also observed at Salisbury, where the number of days with optimum temperature for disease development was higher in 2012 than in 2013. In 2014, the number of days with optimum temperature was higher at Raleigh compared to Salisbury and Tyner (Table 2.1).

The trend in relative humidity was similar to that for temperature between sites and epidemic years. Relative humidity ranged from 37 to 98% during the epidemic in 2012 at Raleigh, while the corresponding range in 2013 and 2014 was 30 to 95% and 37 to 90%, respectively, at the same site (data not shown). Generally, relative humidity increased during the growing season across all sites and epidemic years. However, the number of days with favorable relative humidity varied across sites and epidemic years. For example, there were more days with  $RH \geq 75\%$  in 2012 than in 2013 and 2014. The number of days with precipitation events and the cumulative precipitation were higher in 2012 than in 2013 or 2014 (Table 2.1). For example, there were more days with measurable precipitation in 2012 than in 2013 or 2014. The cumulative precipitation at Raleigh was 292, 135, and 372 mm in 2012, 2013, and 2014, respectively, and the corresponding rainfall amounts at Salisbury were 338, 290, and 260 mm, respectively.

**Disease development.** Severity of SNB increased over the experimental period during the epidemics in 2012, 2013 and 2014 at all sites (Fig. 2.1). Disease onset occurred much earlier in 2012 than in either 2013 or 2014 across all sites. For example, disease onset on DG Shirley occurred 20 to 30 days earlier at Raleigh and Salisbury in 2012 than in 2013

(Fig. 2.1). In 2014, disease onset on DG 9012 occurred 15 days earlier at Raleigh and Tyner than at Salisbury. Mean maximum disease severity on DG Shirley was approximately 15 to 30% higher in 2012 than in 2013 or 2014 at both Raleigh and Salisbury (Fig. 2.1). Disease severity on DG 9012 was higher at Raleigh than at Salisbury or Tyner in 2014 (Fig. 2.1). As expected from individual disease severity values, RAUDPC values were also considerably higher in 2012 than in 2013 or 2014.

The effects of residue level on maximum disease severity and RAUDPC were highly significant ( $P < 0.001$ ) across all sites and epidemic years for DG Shirley (Table 2.2). In 2012, disease severity was significantly ( $P < 0.05$ ) lower in plots with no residue than in plots with residue at both Raleigh and Salisbury (Table 2.3). In addition, disease severity in plots with 60 and 90% residue levels was significantly higher than that in plots with 30% residue (Table 2.3). The effects of residue level on disease severity were slightly less variable at Salisbury than Raleigh as evidenced by lower standard errors at the Salisbury site (Table 2.3). Significant ( $P < 0.05$ ) differences in disease severity in plots with no residue and those with >20% residue were also observed in 2013. However, disease severity in plots with 10 and 20% residue level was not significantly different from that in plots with no residue except at Salisbury where SNB in plots with 20% residue was significantly different from no residue (Table 2.3). Similar effects of residue levels on disease severity were also observed in 2014 but residue effects on SNB were less variable in 2014 than in 2012 (Table 2.3). The effects of residue level on disease severity for DG 9012 in 2014 (Table 2.4) were similar to

those observed on DG Shirley except that mean disease severity was higher on DG 9012 (Table 2.5) than on DG Shirley.

Trend analysis using orthogonal polynomials showed that disease severity on DG Shirley increased linearly with residue level as indicated by a highly significant ( $P < 0.005$ ) linear orthogonal polynomial (Table 2.6). In addition, the increase in disease severity with residue level also displayed a significant ( $P < 0.05$ ) quadratic trend with the exception of Raleigh in 2013 (Table 2.6). Similar results were observed for DG 9012 in 2014 (data not shown). The quadratic contrast is a test of whether a quadratic term can be included given that a linear term is already in the model. It serves as a hierarchical test of a quadratic model (with both linear and quadratic terms) versus a linear model. Thus, the significant results of the general ANOVA hypothesis were primarily explained by both a linear and quadratic relationship between residue level and disease severity.

**Relationship between disease intensity and residue level.** The effects of residue levels on disease intensity were well described ( $P < 0.05$ ) by the 2-parameter logarithmic model (Fig. 2.2). The parameters  $a$  and  $b$  were significantly ( $P < 0.05$ ) different from zero for data collected in all sites and epidemic years (Table 2.7). The correlations between observed vs. predicted disease intensity values were high and significant ( $P < 0.05$ ) and ranged from  $r = 0.76$  to  $0.98$  (Table 2.7). In addition, plots of residuals versus observed or predicted values did not reveal any specific pattern.

**Yield, yield components, and infected kernels.** Effects of residue levels on grain yield varied among experimental sites and seasons of this study (Tables 2.2 and 2.4). For example, the effects of residue level on grain yield were significant ( $P < 0.05$ ) at Raleigh and Salisbury in 2012 on DG Shirley (Table 2.2) and at Salisbury in 2014 on DG 9012 (Table 2.4). Marginal effects of residue on yield were observed at Raleigh in 2013 ( $P = 0.0782$ ) and at Tyner in 2014 ( $P = 0.0857$ ), while no significant ( $P > 0.4622$ ) effects were observed at the other site-years (Tables 2.2 and 2.4). Where the effects of residue level were significant, yield in plots with no residue was significantly higher than yield in plots that had residue. In the latter case, yield generally decreased with increasing residue level, although mean yields were not always significantly different (Tables 2.3 and 2.5). Residue level only had a significant effect on test weight at Raleigh in 2012 (Tables 2.2 and 2.4). Similarly, the effects of residue on thousand-kernel weight were not significant except at Raleigh and Salisbury in 2014 on DG 9012 where thousand-kernel weight generally decreased with increasing residue level (Tables 2.4 and 2.5).

Residue treatments did not significantly ( $P > 0.05$ ) affect incidence of kernel infection for DG Shirley (Table 2.2). However, significant ( $P < 0.05$ ) differences in kernel infection were observed on DG 9012 in 2014 at Raleigh and Salisbury but not at Tyner (Table 2.4). At Raleigh, kernel infection for DG 9012 in 2014 ranged from 40 to 75%, and residue levels  $\geq 30\%$  resulted in kernel infection levels that were significantly higher than those in residue-free control plots. At Salisbury, the 90% residue level resulted in DG 9012 kernel infection percentages significantly higher than in the untreated plots or lower residue



levels. Percent infected kernels at Salisbury was significantly lower than that at Raleigh, ranging from 15 to 45% (Table 2.5).

## **Discussion**

To the best of our knowledge, this is the first study to quantify the effects of the amount of residue on soil surface on the severity of SNB and yield in winter wheat. Our results showed that disease severity was significantly affected by the amount of residue on the soil surface at all experimental sites and growing seasons. In addition, we observed a non-linear relationship between the amount of residue on the soil surface and disease severity, where fairly small increases in residue density produced a rapid increase in disease severity. The amount of residue on the soil surface also affected grain yield, but the effects were only significant when disease onset occurred early and weather conditions during the season were conducive for rapid disease development. The latter has an important implication since in the Mid-Atlantic region of the United States, SNB is one of the foliar diseases most frequently targeted for fungicide applications to wheat.

Infected wheat residue on the soil surface harbors *P. nodorum* between crops and can serve as a source of primary inoculum for initiation of new disease epidemics at the start of the growing season (Duczek et al. 1999; Krupinsky et al. 2007). Several studies have linked infected residue on the soil surface to increased severity of SNB in the field (Holmes and Colhun 1975; Luke et al. 1983; Milus and Chalkley 1997). However, in all these reports, the amount of residue on the soil in field experiments was not specified. In the present study,

disease severity increased with increasing amounts of residue on the soil surface and then leveled off above 30% residue. An increase of residue level up to 30% resulted in greater increase in disease severity than when residue coverage levels were increased to 60 or 90%. It is possible that the amount of inoculum on the surface that is available for splash-dispersal increases as residue levels increase until 30% residue coverage. However, at higher residue levels, a higher proportion of total inoculum is trapped in the lower layers of residue and therefore unavailable for dispersal to cause infection. In some cases, a residue level of 10% resulted in disease severity that was not different from that in plots with 30% residue level. In other words, the 10% residue level provided adequate inoculum for an epidemic equivalent to that initiated from 30% residue. This suggests that relatively small amounts of residue on the soil surface can provide sufficient inoculum for the development of SNB epidemics. Similar observations have also been reported for gray leaf spot in corn (de Nazareno et al. 1993). The 10% residue level on the soil surface is equivalent to conditions of conventional tillage with incomplete burial of residue (Stubbs et al. 2004). We predict that SNB will progress at a fast rate and reach a high maximum severity in such a field and thus, fields would need to be plowed in a manner that ensures complete burial of infected residue if conventional tillage is expected to result in a decrease in SNB inoculum.

The disease intensity (RAUDPC) at all experimental sites was higher in 2012 than in 2013 or 2014 for DG Shirley, irrespective of residue treatment. Similarly, disease onset was earlier in 2012 than in 2013 and 2014 growing seasons, regardless of the residue treatments. Given that SNB is a polycyclic disease, earlier onset was likely responsible for the higher

disease intensities in 2012. Weather factors probably account for differences in timing of disease onset and severities of SNB epidemics observed in this study. More days with optimal temperatures, relative humidity, and rain were observed in 2012 than in 2013 or 2014. Although grain yield in plots with no residue was generally higher than yield in plots with residue, significant effects of residue level on yield were only observed during the 2012 season. The levels of disease severity in 2013 and 2014, which were limited by less favorable weather, were not high enough to significantly reduce grain yield especially for DG Shirley. These findings are similar to those of Milus and Chalkey (1997) who also observed significant differences in yield between plots with residue and plots with no residue only during severe epidemics of SNB. Similarly, Eyal (1981) reported significant reductions in grain yield only during severe epidemics of SNB in Israel.

In our study, residue-associated impacts on yield occurred in environments where disease levels in residue-treated plots were elevated early in the epidemic (Figs. 2.1A, 2.1B, and 2.1I). Since residue was applied in late January or early February, conducive weather in February was likely an important determinant of whether yield would be affected by SNB. It is reasonable that if disease severity is elevated early in the growing season, the risk of disease-related yield impacts may be greater than if the same disease level is not reached until later in the season, due to polycyclic disease development. Infected residue used in the present study was stored under shelter prior to field application and thus did not undergo the weathering to which residue left in the field is exposed after harvest. Thus, the residue used in this study may have had a higher inoculum efficiency than residue left in the field.

However, the increase in SNB with residue application is qualitatively similar when either artificially inoculated residue is introduced in the field or naturally weathered residue left in the field is evaluated as an inoculum source (Holmes and Colhoun 1975; Milus and Chalkley 1997).

Our results support the conclusion that yield and test weight, the variables of interest to wheat growers, are only negatively impacted by the highest levels of disease intensity, despite the disease-elevating effects of unincorporated wheat residue. Although further work on the economic threshold for SNB in the Mid-Atlantic states is needed, our results suggest that the threshold may be higher than is often presumed. In this study, we selected cultivars with moderate or high susceptibility in order to maximize the chances of detecting significant effects of residue level on disease symptoms, yield, and test weight. However, many high-yielding wheat cultivars are moderately resistant to *P. nodorum*, and use of these cultivars is more likely to prevent yield and test weight reductions attributable to SNB.

It is noteworthy that at Raleigh and Salisbury where precipitation was higher than at Tyner in 2014, the susceptible cultivar DG 9012 had a higher incidence of infected grain associated with higher residue levels. While this was not accompanied by a negative test weight impact, it could lead to higher levels of disease in subsequent wheat crops via seed transmission if the harvested grain were saved for planting. The thousand-kernel weight of DG 9012 was negatively affected by residue level at those same two locations, indicating the impact of residue on both seed infection and seed size. This trend was, however, not observed for the moderately susceptible DG Shirley. In their work on the control of SNB

using cultural practices, Luke et al. (1983) reported a significant negative correlation between SNB severity and thousand-kernel weight or seed infection when disease severity was low ( $\leq 20\%$ ). In the present study, we observed no effects of SNB on thousand-kernel weight and seed infection under our epidemic conditions except in 2014 on the susceptible cultivar DG 9012, for which the mean disease severity was  $>20\%$ . Increased cultivar susceptibility to SNB was found to be positively correlated with increased seed infection by *P. nodorum* (Cunfer 1978). However, we are unable to directly compare our results with those of Luke et al. (1983) given that we assessed disease severity on a whole-canopy basis, while Luke et al. (1983) assessed disease severity only on glumes and flag, F-1 and F-2 leaves.

*Stagonospora nodorum* blotch developed in residue-free plots, although the resultant levels of disease severity were much lower than those in plots with residue. Development of SNB in clean plots with no wheat residue was also reported in a previous study (Milus and Chalkey 1997). Dispersal gradients of *P. nodorum* pycnidiospores from a source are usually shallow with a horizontal distance of up to 1 m (Brennan et al. 1985b) in the absence of free vegetation between plots (Holmes and Colhoun 1975), although the spores can also be dispersed to a horizontal distance of 10 m in splash micro-droplets under windy conditions (Brennan et al. 1985a). The experimental plots in our study were surrounded with buffer plots and it is unlikely that disease in residue-free plots was due to pycnidiospores from adjacent plots. Given that pathogen-free seed was used for planting, it is possible that windborne ascospores played a role in initial infection in the residue-free plots. Ascocarps of *P. nodorum* have been reported in North Carolina albeit at low frequency (Cowger and Silva-

Rojas 2006). Thus, both pycnidiospores and ascospores may play a role in initiating SNB epidemics in North Carolina, although ascospores of *P. nodorum* have not been found in other parts of the eastern United States (Cunfer 1998; Shah et al. 1995). In other pathosystems, earlier onset of disease epidemics has been reported to occur in plots with residue than in residue-free plots (Payne et al. 1987), an observation that is attributed to the higher inoculum density in plots with residue. In the present study, however, timing of SNB epidemic onset was similar in both residue-free plots and plots with residue. This is probably because our experiments were conducted in residue-free sites and we introduced residue in plots only after plants were well-tillered to avoid smothering of plants. Our results are similar to those reported for gray leaf spot in corn, where disease epidemics started about the same time regardless of residue level in the experimental plots (de Nazareno et al. 1993). Given that residue remains relatively undisturbed in no-tillage or minimum-tillage fields, it is expected that onset of SNB epidemics will occur much earlier in no-tillage fields than observed in the present study.

Important conclusions that are relevant to the control and management of SNB can be drawn from the present study. First, the increase in disease severity with increasing residue level was nonlinear; residue levels of 10 to 30% produced a rapid increase in disease severity compared to high residue levels under favorable environmental conditions. This finding is relevant where residue is used as an input variable in assessing the likelihood and severity of SNB during the season. Residue levels and other cultural practices can be useful predictors in determining the pre-planting risk for SNB that can be used to guide growers in making

disease management decisions prior to planting (Paul and Munkvold 2004). Second, our results strengthen the previous recommendations to rotate wheat with a non-host (Krupinsky et al. 2007; Luke et al. 1983; Milus and Chalkey 1997) to reduce the amount of initial inoculum in subsequent seasons when wheat is planted. The level of resistance to SNB will also influence the effects of residue on disease severity, with impacts being greater on more susceptible than moderately resistant cultivars. Where a susceptible cultivar is planted in the presence of infected residue, fungicides may be necessary to prevent negative impacts on yield when environmental conditions during the season are favorable for disease development. But the economic threshold for fungicides to manage SNB may be higher than previously believed, as yield and test weight losses appear to be associated with SNB only in more severe epidemics. The acreage under reduced tillage is increasing in wheat-growing areas of the United States (Horowitz et al. 2010), an outcome of which will be increased inoculum availability and the risk of more severe epidemics of SNB, particularly where wheat follows wheat. Thus, cultural practices aimed at reducing infected residue will remain important in the overall SNB management strategy.

### **Acknowledgments**

This study was supported by Hatch Funds from North Carolina Agricultural Experiment Station for Project Number NC02251, and by the USDA-ARS. The authors thank staff at Piedmont Research Station at Salisbury and Lake Wheeler Road Field Laboratory at Raleigh, North Carolina, for managing the research plots; and Ryan Parks,

Bradley Van Staaldhuizen, Tyler Cole, Brandon Clontz, Ian Swaim, Charlie Glover, Barbara Stevens, Joseph Jeffreys, and Jennifer Patton-Özkurt for their assistance in planting and harvesting.



## Literature Cited

- Arseniuk, E., Góral, T., and Scharen, A. L. 1998. Seasonal patterns of spore dispersal of *Phaeosphaeria* spp. and *Stagonospora* spp. *Plant Dis.* 82:187–194.
- Bathgate, J. A., and Loughman, R. 2001. Ascospores are a source of inoculum of *Phaeosphaeria nodorum*, *P. avenaria* f. sp. *avenaria* and *Mycosphaerella graminicola* in Western Australia. *Austral. Plant Pathol.* 30:317–322.
- Bennett, R. S., Milgroom, M. G., Sainudiin, R., Cunfer, B. M, and Bergstrom, G. C. 2007. Relative contribution of seed-transmitted inoculum to foliar populations of *Phaeosphaeria nodorum*. *Phytopathology* 97:584–591.
- Bergstrom, G. C. 2010. *Stagonospora nodorum* blotch and *Stagonospora avenae* blotch. Pages 75–77 in: *Compendium of wheat diseases and pests*. W. W. Bockus, R. L. Bowden, R. M. Hunger, W. L. Morrill, T. D. Murray, and R. W. Smiley, eds. The American Phytopathological Society, St. Paul, Minnesota.
- Bhathal, J. S., Loughman, R., and Speijers, J. 2003. Yield reduction in wheat in relation to leaf disease from yellow (tan) spot and septoria nodorum blotch. *Eur. J. Plant Pathol.* 109:435–443.
- Boosalis, M. G., Doupnik, B. L., and Watkins, J. E. 1986. Effect of reduced tillage on plant diseases. Pages 389–408 in: *No-tillage and Surface Tillage Agriculture*. M. A. Sprague and G. B. Triplett, eds. John Wiley & Sons, New York.
- Brennan, R. M., Fitt, B. D. L., Taylor, G. S., and Colhoun, J. 1985a. Dispersal of *Septoria nodorum* pycnidiospores by simulated raindrops in still air. *J. Phytopathol.* 112:281–290.

- Brennan, R. M., Fitt, B. D. L., Taylor, G. S., and Colhoun, J. 1985b. Dispersal of *Septoria nodorum* pycnidiospores by simulated rain and wind. J. Phytopathol. 112:291–297.
- Campbell, C. E., and Madden, L. V. 1990. Introduction to Plant Disease Epidemiology. John Wiley & Sons, New York.
- Cowger, C., and Silva-Rojas, H. V. 2006. Frequency of *Phaeosphaeria nodorum*, the sexual stage of *Stagonospora nodorum*, on winter wheat in North Carolina. Phytopathology 96:860–866.
- Cowger, C., Weisz, R., Anderson, J. M., and Horton, J. R. 2010. Maize debris increases barley yellow dwarf virus severity in North Carolina winter wheat. Agron. J. 102:688–695.
- Cunfer, B. M. 1978. The incidence of *Septoria nodorum* in wheat seed. Phytopathology 68:832–835.
- de Nazareno, N. R. X., Lipps, P. E., and Madden, L. V. 1993. Effect of levels of corn residue on the epidemiology of gray leaf spot of corn in Ohio. Plant Dis. 77:67–70.
- Duczek, L. J., Sutherland, K. A., Reed, S. L., Bailey, K. L., and Lafond, G. P. 1999. Survival of leaf spot pathogens on crop residue of wheat and barley in Saskatchewan. Can. J. Plant Pathol. 21:165–173.
- Engle, J. S., Madden, L. V., and Lipps, P. E. 2006. Distribution and pathogenic characterization of *Pyrenophora tritici-repentis* and *Stagonospora nodorum* in Ohio. Phytopathology 96:1355–1362.
- Eyal, Z. 1981. Integrated control of Septoria diseases of wheat. Plant Dis. 763–768.

- Eyal, Z. 1999. The septoria tritici and stagonospora nodorum blotch diseases of wheat. Eur. J. Plant Pathol. 105:629–641.
- Eyal, Z., Scharen, A. L., Prescott, J. M., and van Ginkel, M. 1987. The Septoria Diseases of Wheat: Concepts and Methods of Disease Management. Mexico, D. F.: CIMMYT.
- Fried, P. M., and Meister, E. 1987. Inheritance of leaf and head resistance of winter wheat to *Septoria nodorum* in a diallel cross. Phytopathology 77:1371–1375.
- Gilbert, J., and Tekauz, A. 1993. Reaction of Canadian spring wheats to *Septoria nodorum* and the relationship between disease severity and yield components. Plant Dis. 77:398–402.
- Holmes, S. J. I., and Colhoun, J. 1975. Straw-borne inoculum of *Septoria nodorum* and *S. tritici* in relation to incidence of disease on wheat plants. Plant Pathol. 24:63–66.
- Horowitz, J., Ebel, R., and Uda, K. 2010. ‘No-Till’ Farming is a Growing Practice. United States Department of Agriculture-Economic Research Service. Economic Information Bulletin Number 70. Washington, DC.
- Khel, Z., Habash, D. Z., Gezan, A. A., Welham, S. J., and Nachit, M. M. 2010. Estimation of spatial trend and automatic model selection in augmented designs. Agron. J. 102:1542–1552.
- King, J. E., Cook, R. J., and Melville, S. C. 1983. A review of *Septoria* diseases of wheat and barley. Ann. Appl. Biol. 103:345–373.

- Krupinsky, J. M., Tanaka, D. L., Merrill, S. D., Liebig, M. A., Lares, M. T., and Hanson, J. D. 2007. Crop sequence effects on leaf spot diseases of no-till spring wheat. *Agron. J.* 99:912–920.
- Luke, H. H., Pfahler, P. L., and Barnett, R. D. 1983. Control of *Septoria nodorum* on wheat with crop rotation and seed treatment. *Plant Dis.* 67:949–951.
- Manandhar, J. B., and Cunfer, B. M. 1991. An improved selective medium for the assay of *Septoria nodorum* from wheat seed. *Phytopathology* 81:771–773.
- McKendry, A. L., Henke, G. E., and Finney, P. L. 1995. Effects of *Septoria* leaf blotch on soft red winter wheat milling and baking quality. *Cereal Chem.* 72:142–146.
- Milus, E. A., and Chalkley, D. B. 1997. Effect of previous crop, seedborne inoculum, and fungicides on development of *Stagonospora* blotch. *Plant Dis.* 81:1279–1283.
- Murray, G. M. and Brennan, J. P. 2009. The Current and Potential Costs from Wheat Diseases in Australia. Wagga Wagga, Australia: NSW Agriculture.
- Paul, P. A., and Munkvold, G. P. 2004. A model-based approach to preplanting risk assessment for gray leaf spot of maize. *Phytopathology* 94:1350–1357.
- Payne, G. A., Duncan, H. E., and Adkins, C. R. 1987. Influence of tillage on development of gray leaf spot and number of airborne conidia of *Cercospora zeae-maydis*. *Plant Dis.* 71:329–332.
- Quaedvlieg, W., Verkley, G. J. M., Shin, H. -D., Barreto, R. W., Alfenas, A. C., Swart, W. J., Groenewald, J. Z., and Crous, P. W. 2013. Sizing up *Septoria*. *Stud. Mycol.* 75:307–390.

- Shaner, G., and Buechley, G. 1995. Epidemiology of leaf blotch of soft red winter wheat caused by *Septoria tritici* and *Stagonospora nodorum*. Plant Dis. 79:928–938.
- Shah, D. A., and Bergstrom, G. C. 2000. Temperature dependent seed transmission of *Stagonospora nodorum* in wheat. Eur. J. Plant Pathol. 106:837–842.
- Shah, D., Bergstrom, G. C., and Ueng, P. P. 1995. Initiation of *Septoria nodorum* blotch epidemics in winter wheat by seedborne *Stagonospora nodorum*. Phytopathology 85:452–457.
- Shah, D. A., Bergstrom, G. C., and Ueng, P. P. 2001. Foci of *Stagonospora nodorum* blotch in winter wheat before canopy development. Phytopathology 91:642–647.
- Shaver, T. M., Peterson, G. A., Ahuja, L. R., and Westfall, D. J. 2013. Soil sorptivity enhancement with crop residue accumulation in semi-arid dryland no-till agroecosystems. Geoderma 192:254–258.
- Shipton, W. A., Boyd, W. R. J., Rosielle, A. A., and Shearer, B. I. 1971. The common septoria diseases of wheat. Bot. Rev. 37:231–262.
- Sommerhalder, R. J., McDonald, B. A., and Zhan, J. 2006. The frequencies and spatial distribution of mating types in *Stagonospora nodorum* are consistent with recurring sexual reproduction. Phytopathology 96:234–239.
- Sommerhalder, R. J., McDonald, B. A., Mascher, F., and Zhan, J. 2010. Sexual recombinants make a significant contribution to epidemics caused by the wheat pathogen *Phaeosphaeria nodorum*. Phytopathology 100:855–862.

- Stubbs, T. L., Kennedy, A. C., and Schillinger, W. F. 2004. Soil ecosystem changes during the transition to no-till cropping. *J. Crop Improv.* 11:105–135.
- Verreet, J. A., and Hoffmann, G. M. 1989. Physiological reactions of infections by *Septoria nodorum* in different growth stages of wheat. Pages 63–65 in: *Septoria of Cereals*. Proc. Workshop Septoria Dis. Cereals, 3<sup>rd</sup>. P. M. Fried, ed. Swiss Fed. Res. Stn. for Agronomy, Zurich.
- Verreet, J. A., and Hoffmann, G. M. 1990. A biologically oriented threshold decision model for control of epidemics of *Septoria nodorum* in wheat. *Plant Dis.* 74:731–738.
- Weisz, R. 2013. Small Grain Production Guide. North Carolina Cooperative Extension Service, Publication No. AG-580 (Revised March 2013).
- Wollenhaupt, N. C., and Pingry, J. 1991. Estimating Residue using the Line-transect Method. Publication No. A3553. Cooperative Extension, University of Wisconsin, Madison, WI.
- Zadoks, J. C., Chang, T. T., and Konzak, C. F. 1974. A decimal code for the growth stages of cereals. *Weed Res.* 14:415–421.

**Table 2.1.** Number of days of weather conducive to the development of *Stagonospora nodorum* blotch during the epidemic period <sup>a</sup> in North Carolina

Site	Weather variable								
	Temperature <sup>b</sup>			Relative humidity <sup>c</sup>			Rainfall <sup>d</sup>		
	2012	2013	2014	2012	2013	2014	2012	2013	2014
Raleigh	16	14	21	25	19	10	26	17	19
Salisbury	14	10	15	21	16	8	27	23	14
Tyner <sup>e</sup>	-	-	10	-	-	18	-	-	15

<sup>a</sup> The epidemic period was defined as the interval starting two weeks prior to disease onset until final disease assessment.

<sup>b</sup> The number of days when average daily temperature ranged between 20 and 25°C.

<sup>c</sup> The number of days with relative humidity  $\geq 75\%$ .

<sup>d</sup> The number of days with measurable rainfall.

<sup>e</sup> Experiments were conducted at Tyner only in 2014.

**Table 2.2.** Effect of wheat residue levels on the intensity of *Stagonospora nodorum* blotch and yield characteristics for winter wheat cultivar DG Shirley

Site/ Variable <sup>a</sup>	2012 <sup>b</sup>				2013 <sup>b</sup>				2014 <sup>b</sup>			
	ndf	ddf	<i>F</i>	<i>P</i> > <i>F</i>	ndf	ddf	<i>F</i>	<i>P</i> > <i>F</i>	ndf	ddf	<i>F</i>	<i>P</i> > <i>F</i>
Raleigh												
MaxDS	3	12	80.55	0.0001	5	20	8.09	0.0003	5	20	31.37	0.0001
RAUDPC	3	12	181.07	0.0001	5	20	18.62	0.0001	5	20	108.77	0.0001
PKC	3	16	0.43	0.7358	5	20	1.53	0.2259	5	20	1.84	0.1507
Yield	3	16	6.81	0.0036	5	20	2.35	0.0782	5	20	0.83	0.5454
TW	3	12	9.50	0.0017	5	20	1.01	0.4374	5	20	0.54	0.7420
TKW	-	-	-	-	5	24	1.01	0.4337	5	24	1.23	0.3268
Salisbury												
MaxDS	3	12	245.51	0.0001	5	20	8.82	0.0002	5	20	25.26	0.0001
RAUDPC	3	12	392.15	0.0001	5	24	16.82	0.0001	5	20	55.77	0.0001
PKC	3	12	0.30	0.8229	5	24	1.00	0.4406	5	24	0.71	0.6225
Yield	3	9	9.17	0.0041	5	19	0.97	0.4622	5	24	0.77	0.5828
TW	3	12	2.53	0.1065	5	19	0.64	0.6748	5	24	0.56	0.7323
TKW	-	-	-	-	5	20	0.80	0.5623	5	20	1.26	0.3177
Tyner												
MaxDS	-	-	-	-	-	-	-	-	5	20	6.54	0.0009
RAUDPC	-	-	-	-	-	-	-	-	5	20	21.22	0.0001
PKC	-	-	-	-	-	-	-	-	5	20	2.62	0.0557
Yield	-	-	-	-	-	-	-	-	5	20	2.28	0.0857
TW	-	-	-	-	-	-	-	-	5	20	0.99	0.4488
TKW	-	-	-	-	-	-	-	-	5	24	1.03	0.4224

<sup>a</sup> MaxDS is maximum disease severity (%), RAUDPC is relative area under the disease progress curve, and PKC is percent kernels colonized by *Parastagonospora nodorum*. Yield is yield (t ha<sup>-1</sup>), TW is test weight (kg m<sup>-3</sup>), while TKW is thousand-kernel weight (g).



<sup>b</sup> Four residue levels were evaluated in 2012, while 6 residue levels were evaluated in the 2013 and 2014 growing seasons. Thousand-kernel weight was not recorded in the 2012 experiments and the experiments were conducted at Tyner only in 2014.

**Table 2.3.** Impact of levels of wheat residue on the development and severity of *Stagonospora nodorum* blotch and yield characteristics for winter wheat cultivar DG Shirley

Variable <sup>x</sup>	Residue level (%) <sup>y</sup>	2012 <sup>z</sup>		2013		2014		
		Raleigh	Salisbury	Raleigh	Salisbury	Raleigh	Salisbury	Tyner
MaxDS	0	18.8 c	20.4 c	25.4 d	19.2 e	19.8 d	8.0 d	19.4 d
	10	-	-	26.4 d	20.6 de	27.0 c	20.8 c	20.4 cd
	20	-	-	29.4 cd	24.2 cd	29.2 bc	24.2 bc	24.4 ab
	30	36.6 b	29.8 b	36.8 ab	25.2 bc	30.0 b	23.4 bc	22.6 bc
	60	44.2 a	34.8 a	33.6 bc	28.6 ab	33.4 a	26.4 ab	24.4 ab
	90	47.2 a	36.0 a	40.0 a	29.6 a	34.2 a	28.8 a	26.4 a
	Mean	36.7	30.3	31.9	24.6	28.9	21.9	22.9
	SE	2.71	1.43	1.31	0.90	0.96	1.37	0.69
RAUDPC	0	8.8 c	9.4 d	8.0 d	7.8 d	4.8 d	3.5 d	6.7 d
	10	-	-	8.6 cd	9.7 c	9.5 c	9.3 c	9.6 c
	20	-	-	10.2 c	10.9 c	11.7 b	12.8 b	10.9 bc
	30	23.6 b	17.2 c	13.8 ab	11.3 bc	12.3 b	12.2 b	10.2 bc
	60	25.4 b	21.1 b	13.0 b	13.1 b	14.3 a	14.9 a	11.4 b
	90	27.6 a	23.3 a	15.9 a	15.3 a	14.8 a	16.1 a	13.2 a
	Mean	21.3	17.7	11.6	11.4	11.2	11.5	10.3
	SE	1.73	1.22	0.60	0.50	0.65	0.81	0.41
PKC	0	25.0	17.0	16.0 b	7.00	43.0 b	29.0	32.0 b
	10	-	-	17.0 b	5.00	45.0 b	35.0	24.0 b
	20	-	-	22.0 ab	7.00	55.0 ab	34.0	35.0 ab
	30	31.0	20.0	27.0 a	10.0	53.0 ab	26.0	30.0 b
	60	28.0	17.0	21.0 ab	11.0	53.0 ab	40.0	32.0 b
	90	24.0	21.0	23.0 ab	14.0	63.0 a	32.0	44.0 a
	Mean	27.0	18.8	21.0	9.00	52.0	32.7	32.8
	SE	2.31	1.81	1.50	1.34	2.39	2.31	1.88
Yield	0	8.7 a	8.0 a	11.1 a	9.4	8.6	8.5	6.9 a
	10	-	-	10.4 ab	9.7	8.6	8.1	6.5 ab
	20	-	-	10.1 b	9.5	8.1	8.9	6.1 b
	30	8.1 ab	7.6 b	10.4 ab	9.9	8.0	8.7	6.8 a
	60	7.3 c	7.4 b	9.8 b	9.7	8.0	8.6	6.6 ab

	90	7.8 bc	7.4 b	10.4 ab	9.4	8.4	8.3	6.2 b
	Mean	7.9	7.6	10.3	9.6	8.3	8.5	6.5
	SE	0.15	0.12	0.13	0.11	0.19	0.13	0.11
TW (×10)	0	70.2 a	73.1 a	69.7	68.4	74.9	72.0	72.9
	10	-	-	68.5	69.0	75.2	72.2	73.8
	20	-	-	70.0	68.6	75.2	72.0	73.7
	30	68.5 b	72.9 ab	68.8	69.2	75.4	72.3	74.2
	60	68.0 b	72.7 ab	69.2	68.3	75.2	72.6	74.0
	90	68.2 b	72.6 b	69.9	68.5	75.1	72.3	72.9
	Mean	68.7	72.8	69.4	68.7	75.2	72.2	73.6
	SE	0.26	0.08	0.30	0.17	0.10	0.12	0.25
TKW	0	-	-	37.0	40.9	38.0 ab	38.0 a	37.7
	10	-	-	35.2	42.1	39.4 a	37.9 ab	36.7
	20	-	-	35.8	39.9	38.5 ab	37.7 ab	37.1
	30	-	-	34.2	41.7	37.8 b	37.9 ab	37.5
	60	-	-	35.8	41.8	38.5 ab	38.0 ab	36.9
	90	-	-	36.5	41.2	37.9 ab	37.1 b	36.1
	Mean	-	-	35.8	41.3	38.3	37.7	37.0
	SE	-	-	0.41	0.37	0.23	0.14	0.23

<sup>x</sup> MaxDS is maximum disease severity (%), RAUDPC is relative area under the disease progress curve, and PKC is percent kernels colonized by *Parastagonospora nodorum*. Yield is yield (t ha<sup>-1</sup>), TW is test weight (kg m<sup>-3</sup>), while TKW is thousand-kernel weight (g). Values presented are means and for each variable within a site and year, values followed by the same letter are not significantly different based on Fisher's LSD at  $\alpha = 0.05$ . Mean separation is shown only for variables with significant treatment effects in analysis of variance.

<sup>y</sup> Percentage of residue cover on the soil surface and SE = standard error of the mean.

<sup>z</sup> Residue cover levels of 10 and 20% were not tested in 2012; thousand-kernel weight was not recorded in 2012.

**Table 2.4.** Effect of wheat residue levels on intensity of *Stagonospora nodorum* blotch and yield characteristics for winter wheat cultivar DG 9012 in 2014

Site	Variable <sup>a</sup>	ndf	ddf	<i>F</i>	<i>P</i> > <i>F</i>
Raleigh	MaxDS	5	20	16.60	0.0001
	RAUDPC	5	20	115.26	0.0001
	PKC	5	20	3.52	0.0190
	Yield	5	20	0.94	0.4751
	TW	5	20	0.42	0.8291
	TKW	5	20	2.86	0.0418
Salisbury	MaxDS	5	20	52.86	0.0001
	RAUDPC	5	20	82.54	0.0001
	PKC	5	23	3.30	0.0216
	Yield	5	20	2.93	0.0382
	TW	5	20	0.24	0.9378
	TKW	5	20	2.88	0.0407
Tyner	MaxDS	5	24	2.54	0.0555
	RAUDPC	5	20	24.74	0.0001
	PKC	5	20	0.99	0.4511
	Yield	5	20	0.31	0.9020
	TW	5	24	0.83	0.5422
	TKW	5	20	1.39	0.2694

<sup>a</sup> MaxDS is maximum disease severity (%), RAUDPC is relative area under the disease progress curve, and PKC is percent kernels colonized by *Parastagonospora nodorum*. Yield is yield (t ha<sup>-1</sup>), TW is test weight (kg m<sup>-3</sup>), while TKW is thousand-kernel weight (g).

**Table 2.5.** Impact of amounts of wheat residue on the development and severity of *Stagonospora nodorum* blotch and yield characteristics for winter wheat cultivar DG 9012 in 2014

Variable <sup>y</sup>	Residue level (%) <sup>z</sup>	Site		
		Raleigh	Salisbury	Tyner
MaxDS	0	29.6 d	15.0 d	23.4 c
	10	34.0 cd	27.2 c	24.0 bc
	20	39.0 bc	28.0 bc	26.0 abc
	30	42.4 b	31.0 b	26.8 ab
	60	49.2 a	34.8 a	26.8 ab
	90	49.8 a	36.8 a	27 a
	Mean (SE)	40.7 (1.61)	28.8 (1.38)	25.7 (0.45)
RAUDPC	0	7.5 f	6.4 d	11.1 d
	10	13.3 e	14.3 c	13.2 c
	20	16.6 d	16.1 c	15.3 b
	30	18.5 c	18.3 b	15.4 b
	60	20.4 b	21.8 a	16.5 ab
	90	22.4 a	23.1 a	16.9 a
	Mean (SE)	16.5 (0.85)	16.7 (1.07)	14.7 (0.43)
PKC	0	41.0 b	16.0 b	32.0
	10	62.0 ab	28.0 b	47.0
	20	59.0 ab	26.0 b	45.0
	30	75.0 a	24.0 b	41.0
	60	74.0 a	28.8 ab	35.0
	90	75.0 a	45.0 a	38.0
	Mean (SE)	64.3 (3.56)	27.9 (2.58)	39.7 (2.46)
Yield	0	8.1	8.5 ab	5.6
	10	8.1	8.0 bc	5.6
	20	7.9	7.9 c	5.8
	30	8.2	8.6 a	5.6
	60	7.8	7.9 c	5.8
	90	8.0	8.0 bc	5.6
	Mean (SE)	8.0 (0.11)	8.2 (0.14)	5.7 (0.12)
TW (×10)	0	77.2	74.3	76.2
	10	77.2	74.1	75.9
	20	77.2	74.0	76.2
	30	77.0	74.2	74.3
	60	76.9	73.9	76.2
	90	77.1	74.5	75.2
	Mean (SE)	77.1 (0.10)	74.2 (0.17)	75.6 (0.34)
TKW	0	38.0 a	38.2 a	35.7
	10	37.0 ab	38.2 a	34.7
	20	36.7 abc	38.1 ab	35.6
	30	35.9 bc	37.5 abc	34.9
	60	35.4 c	37.0 bc	34.9
	90	36.5 abc	36.8 c	35.4
	Mean (SE)	36.6 (0.27)	37.6 (0.23)	35.2 (0.16)

<sup>y</sup> MaxDS is maximum disease severity (%), RAUDPC is relative area under the disease progress curve, and PKC is percent kernels colonized by *Parastagonospora nodorum*. Yield is yield (t ha<sup>-1</sup>), TW is test weight (kg m<sup>-3</sup>), while TKW is thousand-kernel weight (g). Values presented are means and for each variable within a site and year, values followed by the same letter are not significantly different based on Fisher's LSD at  $\alpha = 0.05$ . Mean separation is shown only for variables with significant treatment effects in the analysis of variance.

<sup>z</sup> Percentage of residue cover on the soil surface and SE = standard error of the mean.

**Table 2.6.** Orthogonal polynomial contrasts to test the hypothesis of linear, quadratic and cubic response in *Stagonospora nodorum* blotch severity and yield with increasing levels of wheat residue cover in plots of cultivar DG Shirley

Variable <sup>a</sup>	Site	Test	2012 <sup>b</sup>			2013 <sup>b</sup>			2014 <sup>b</sup>		
			ndf,ddf	<i>F</i>	<i>P</i> > <i>F</i>	ndf,ddf	<i>F</i>	<i>P</i> > <i>F</i>	ndf,ddf	<i>F</i>	<i>P</i> > <i>F</i>
Disease severity	Raleigh	General	3,12	80.55	0.0001	5,20	8.09	0.0003	5,20	31.37	0.0001
		Linear	1,12	213.69	0.0001	1,20	30.80	0.0001	1,20	115.82	0.0001
		Quadratic	1,12	27.18	0.0002	1,20	1.40	0.2501	1,20	29.67	0.0001
		Cubic	1,12	0.78	0.3950	1,20	2.18	0.1558	1,20	6.59	0.0184
	Salisbury	General	3,12	245.51	0.0001	5,20	8.82	0.0002	5,20	25.26	0.0001
		Linear	1,12	654.45	0.0001	1,20	39.38	0.0001	1,20	75.66	0.0001
		Quadratic	1,12	82.00	0.0001	1,20	4.11	0.0561	1,20	25.65	0.0001
		Cubic	1,12	0.09	0.7721	1,20	0.01	0.9131	1,20	16.92	0.0005
RAUDPC	Raleigh	General	3,12	181.07	0.0001	5,20	18.62	0.0001	5,20	108.77	0.0001
		Linear	1,12	422.86	0.0001	1,20	76.38	0.0001	1,20	389.58	0.0001
		Quadratic	1,12	97.97	0.0001	1,20	3.85	0.0639	1,20	119.35	0.0001
		Cubic	1,12	22.39	0.0005	1,20	3.07	0.0953	1,20	26.62	0.0001
	Salisbury	General	3,12	392.15	0.0001	5,24	16.82	0.0001	5,20	55.77	0.0001
		Linear	1,12	1093.1	0.0001	1,24	80.09	0.0001	1,20	198.40	0.0001
		Quadratic	1,12	81.05	0.0001	1,24	1.78	0.1948	1,20	51.82	0.0001
		Cubic	1,12	2.31	0.1547	1,24	1.88	0.1830	1,20	19.22	0.0003
Yield	Raleigh	General	3,16	6.81	0.0036	-	-	-	-	-	-
		Linear	1,16	12.63	0.0026	-	-	-	-	-	-
		Quadratic	1,16	5.98	0.0264	-	-	-	-	-	-
		Cubic	1,16	1.81	0.1973	-	-	-	-	-	-
	Salisbury	General	3,9	3.30	0.0041	-	-	-	-	-	-
		Linear	1,9	5.95	0.0007	-	-	-	-	-	-
		Quadratic	1,9	3.21	0.0651	-	-	-	-	-	-
		Cubic	1,9	0.75	0.9963	-	-	-	-	-	-

<sup>a</sup> Disease severity refers to maximum disease severity (%) recorded in the respective treatment plots; RAUDPC is relative area under the disease progress curve; yield is measured as  $\text{t ha}^{-1}$ .

<sup>b</sup> Residue levels significantly affected yield in 2012 but not in 2013 and 2014 and thus orthogonal contrasts were conducted only for yield data in the 2012 growing season.



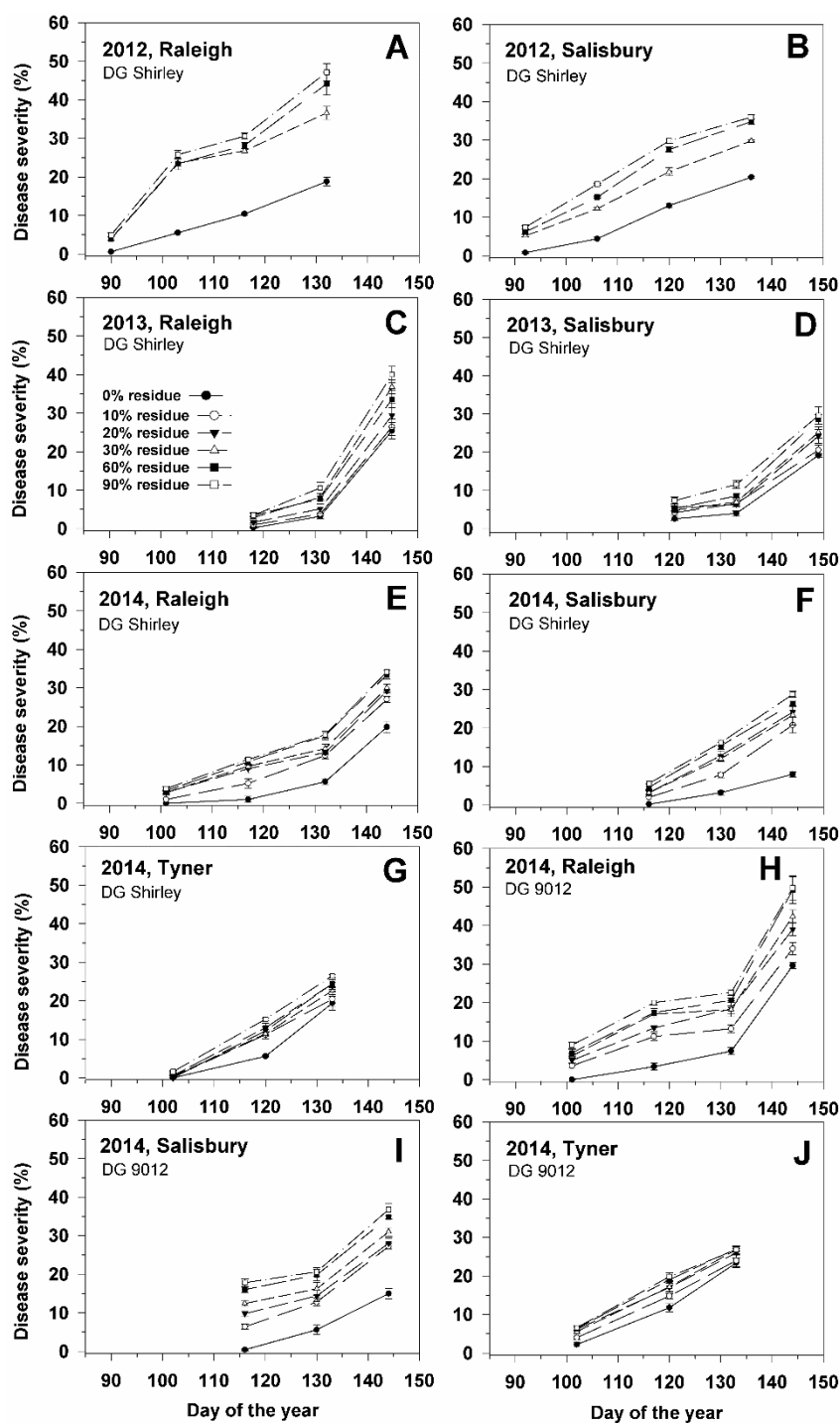
**Table 2.7.** Nonlinear regression analysis of the relationship between amounts of wheat residue cover on the ground and the development of *Stagonospora nodorum* blotch, using a two-parameter logarithmic model<sup>x</sup>

Year	Location	Cultivar	Parameter (SE) <sup>y</sup>		Goodness-of-fit	
			<i>a</i>	<i>b</i>	<i>r</i> <sup>z</sup>	<i>P</i> > <i>F</i>
2012	Raleigh	DG Shirley	8.83 (0.723)	4.15 (0.206)	0.978	0.0001
	Salisbury	DG Shirley	9.04 (0.615)	2.89 (0.176)	0.968	0.0001
2013	Raleigh	DG Shirley	6.75 (0.889)	1.66 (0.272)	0.756	0.0001
	Salisbury	DG Shirley	6.97 (0.653)	1.51 (0.200)	0.819	0.0001
2014	Raleigh	DG Shirley	4.57 (0.413)	2.29 (0.126)	0.960	0.0001
	Salisbury	DG Shirley	3.32 (0.635)	2.79 (0.194)	0.938	0.0001
	Tyner	DG Shirley	6.53 (0.455)	1.30 (0.139)	0.870	0.0001
2014	Raleigh	DG 9012	6.81 (0.675)	3.30 (0.207)	0.949	0.0001
	Salisbury	DG 9012	5.85 (0.770)	3.71 (0.236)	0.948	0.0001
	Tyner	DG 9012	10.8 (0.500)	1.33 (0.153)	0.855	0.0001

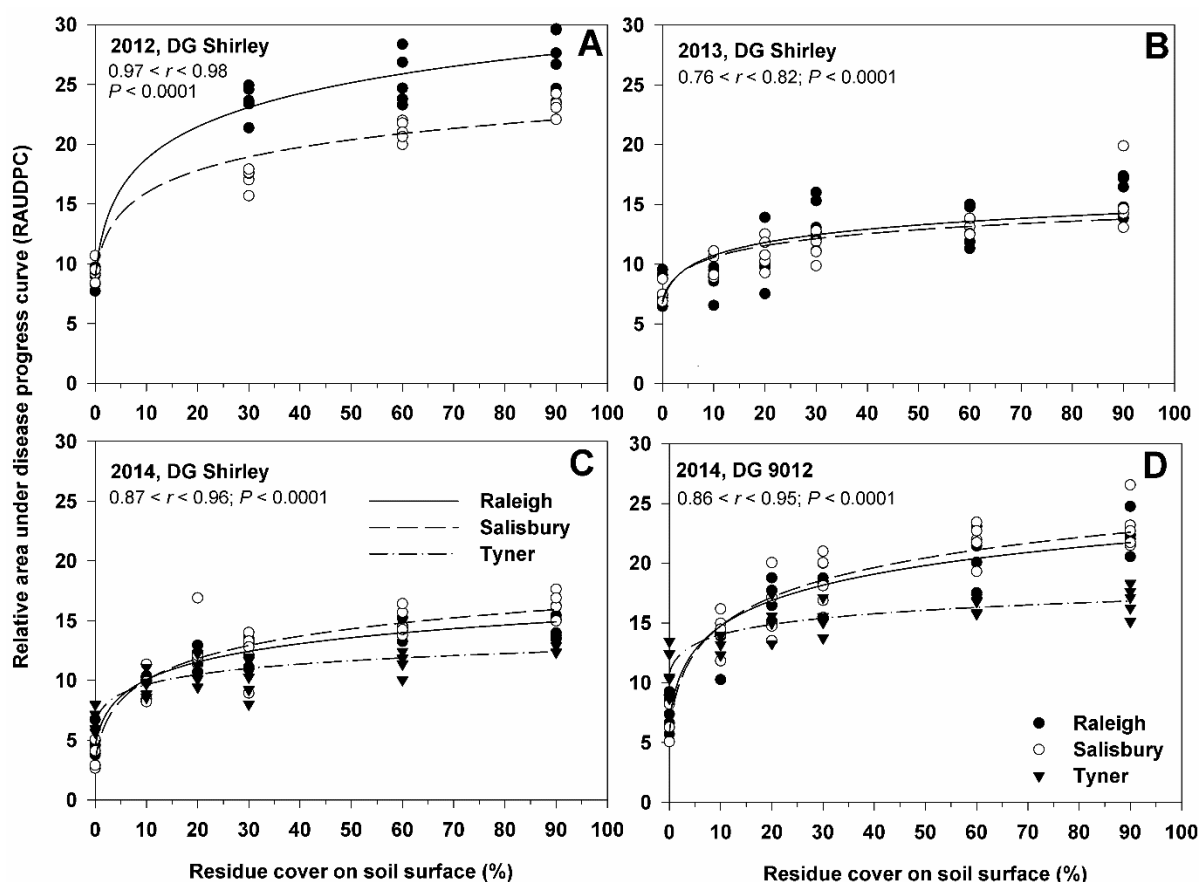
<sup>x</sup> Two-parameter model fitted to the data was of the form:  $y = a + b \cdot \ln(x+1)$ , where  $y$  = relative area under the disease progress curve (RAUDPC),  $x$  = amount of residue,  $a$  = intercept, and  $b$  = rate parameter.

<sup>y</sup> Values in parenthesis are standard errors (SE) of the parameter estimates.

<sup>z</sup> Correlation coefficient between observed and predicted disease intensity (RAUDPC) values.



**Figure 2.1.** Effect of different levels of wheat residue on the soil surface on the development of *Stagonospora nodorum* blotch in three years at three experimental sites in North Carolina. The wheat cultivar DG Shirley was evaluated in all site-years (A-G), while the cultivar DG 9012 was evaluated only in 2014 at all sites (H-J). Data points represent means of disease severity (%) values from five replicates; the bars are standard errors of the mean.



**Figure 2.2.** Relationship between *Stagonospora nodorum* blotch intensity (relative area under the disease progress curve) and levels of wheat residue on soil surface in three growing seasons at three experimental sites in North Carolina. The wheat cultivar DG Shirley was evaluated in all years (A–C), while the cultivar DG 9012 was evaluated only in 2014 at all sites (D). Symbols (circles and triangles) represent the observed disease data, while the curves represent the predicted disease intensity based on a two-parameter logarithmic model. In each year, the coefficients of correlations ( $r$ ) between observed and predicted values are a range depicting the goodness-of-fit of the model to the data across experimental sites.

### CHAPTER 3

## **Predicting pre-planting risk of *Stagonospora nodorum* blotch in winter wheat using machine learning models**

To be submitted to *Agricultural Systems*:

L.K. Mehra <sup>a</sup>, C. Cowger <sup>a,b</sup>, K. Gross <sup>c</sup>, and P.S. Ojiambo <sup>a,\*</sup>

<sup>a</sup> *Department of Plant Pathology, North Carolina State University, Raleigh, NC 27695, USA*

<sup>b</sup> *United States Department of Agriculture-Agricultural Research Service, Raleigh, NC 27695, USA*

<sup>c</sup> *Department of Statistics, North Carolina State University, Raleigh, NC 27695, USA*

## ABSTRACT

Pre-planting factors have been associated with the late-season severity of *Stagonospora nodorum* blotch (SNB), caused by the fungal pathogen *Parastagonospora nodorum*, in winter wheat (*Triticum aestivum*). The relative importance of these factors in the risk of SNB has not been determined and this knowledge can facilitate disease management decisions prior to planting of the wheat crop. In this study, we examined the performance of multiple regression (MR) and three machine learning algorithms (artificial neural networks, categorical and regression trees, and random forests) in predicting the pre-planting risk of SNB in wheat. Pre-planting factors tested as potential predictor variables were cultivar resistance, latitude, longitude, previous crop, seeding rate, seed treatment, tillage type, and wheat residue. Disease severity assessed at the end of the growing season was used as the response variable. The models were developed using 431 disease cases (unique combinations of predictors) collected from 2012 to 2014 and these cases were randomly divided into training, validation, and test datasets. Models were evaluated based on the regression of observed against predicted severity values of SNB, sensitivity-specificity ROC analysis, and the Kappa statistic. A strong relationship was observed between late-season severity of SNB and specific pre-planting factors in which latitude, longitude, wheat residue, and cultivar resistance were the most important predictors. The MR model explained 33% of variability in the data, while machine learning models explained 47 to 79% of the total variability. Similarly, the MR model correctly classified 74% of the disease cases, while machine learning models correctly classified 81 to 83% of these cases. Results show that the random

forest (RF) algorithm, which explained 79% of the variability within the data, was the most accurate in predicting the risk of SNB, with an accuracy rate of 93%. The RF algorithm could allow early assessment of the risk of SNB, facilitating sound disease management decisions prior to planting of wheat.

*Keywords:* Disease risk, Machine learning, Random forest, Variable importance, *Stagonospora nodorum* blotch, Wheat

## 1. Introduction

*Stagonospora nodorum* blotch (SNB) of wheat (*Triticum aestivum*), caused by *Parastagonospora nodorum*, is a major disease of wheat worldwide (Solomon et al., 2006; Quaedvlieg et al., 2013). The disease affects both the quantity and quality of yield, and losses up to 50% have been reported in susceptible cultivars (Eyal, 1987; Bhathal et al., 2003). In the United States, the disease is becoming more prevalent in part due to the increased adoption of minimum tillage (Shaner and Buechley, 1995). Wheat residue infected with *P. nodorum* serves as a source of primary inoculum for infecting the wheat crop in the field (Holmes and Colhoun, 1975). Minimum tillage favors the survival of the SNB pathogen in wheat residue left on the soil surface from the previous cropping season (Milus and Chalkey, 1997), a practice that ensures inoculum availability to initiate SNB epidemics at the start of the growing season. The cropping area under minimum tillage is increasing in wheat-growing regions of the United States (Horowitz et al., 2010), making severe epidemics of

SNB more likely, especially in rotations where wheat was planted the previous season, as is the case with wheat double-cropped with soybeans.

SNB is currently managed in wheat using a variety of methods that include crop rotation, tillage, planting moderately resistant cultivars, fungicide-treated seed and foliar fungicides (Luke et al., 1983; Milus and Chalkey, 1997; Krupinsky et al., 2007). Pre-planting factors such as crop rotation and tillage have been shown to reduce the severity of SNB at the end of the season, but their effectiveness depends on their widespread adoption given that airborne ascospores from adjacent fields may lead to disease development even where there is no wheat residue on the soil surface (Cowger and Silva-Rojas, 2006). Foliar fungicides can be effective in controlling SNB, but in periods when wheat prices are low, the realized yield response may not be adequate to offset the cost of fungicide treatments (Weisz et al., 2011). Complete resistance in wheat to SNB is currently not available and cultivar resistance ranges from moderately resistant to highly susceptible. Cultivar resistance also interacts with wheat residue in the field to influence SNB severity, with residue effects being greater on a highly susceptible than a moderately susceptible cultivar (Mehra et al., 2015).

As for all plant diseases, an SNB epidemic is an outcome of an ecological process that involves the interaction between a population of host wheat plants and *P. nodorum*, a process that is influenced by the environment at different temporal and spatial scales (Madden et al., 2007). Variability in the susceptibility of the wheat plant population and pathogenicity of *P. nodorum* determine the extent of subsequent spread of SNB. Together with host plant resistance, the environment, defined broadly to include pre-planting factors

that influence inoculum availability and weather during the growing season, determines the severity of SNB epidemics. Interaction among various elements of the pathosystem dictates that SNB epidemics will exhibit complicated behavior over different temporal and spatial scales (Madden et al., 2007). A fundamental goal in botanical epidemiology is to predict the risk of disease at various spatio-temporal scales (Madden, 2006). Information on the expected risk of a disease epidemic can aid growers in making better informed disease management decisions when seeking to reduce potential yield losses. Development of models to understand disease dynamics and predict the risk of disease outbreak to facilitate decision-making is an integral component of plant disease management (Jeger and Xu, 2015; Scherm et al., 2006).

Pre-planting factors such as crop rotation, type of tillage, planting of a less susceptible cultivar and amount of wheat residue in the field can influence SNB risk during the growing season. The decision to use any of these factors singly or in combination can reduce the risk of SNB and its impact on yield at the end of the growing season. However, such a decision-making tool is currently not available for SNB. The choice to use a moderately resistant cultivar at a given location should be based on previous history of SNB at the location, whether wheat was planted the previous season and the type of residue management practiced in the field. Clearly, the use of moderately resistant cultivars in SNB management can be improved through a selective combination of host resistance with other pre-planting factors. For example, while the amount of wheat residue is related to disease severity, other factors such as cultivar susceptibility influence the magnitude of that



relationship (Mehra et al., 2015). In addition, field location also appears to be an important pre-planting predictor of SNB. For example, in North Carolina, SNB tends to occur in the western (Piedmont region) and northeastern (Tidewater region) parts of the state (Weisz, 2013). Given that several pre-planting factors can potentially influence the risk of SNB, there is a need to develop a decision-making criterion that takes the effect of these factors, singly or in combination, into consideration. A pre-planting risk assessment model can provide critical information to guide SNB management decisions prior to planting of the wheat crop.

Predicting the risk of SNB by relating pre-planting factors to the severity of the disease can involve working with data that is complex and unbalanced. For example, the relationship between pre-planting factors (e.g., wheat residue) and SNB severity can be strongly nonlinear and could involve high-order interactions (Mehra et al., 2015). When the interest lies in developing a model to predict a disease severity class, often the goal is to produce an accurate classifier for the disease class and to uncover the predictive structure of the problem. Traditionally, regression analysis has been the most popular modelling technique in predicting disease risk (De Wolf et al., 2003; Gent and Ocam, 2009). In recent years, accurate classifiers have been developed using machine learning methods, which are capable of synthesizing regression or classification functions based on available data (Gutierrez, 2015). Unlike traditional methods, machine learning methods can deal with complex and non-linear relationships between predictors and a response and are also able to process multifaceted and noisy data (Recknagel, 2001; Garzón et al., 2006).

Among machine learning methods, categorical and regression tree (CART) and artificial neural networks (ANN) methodologies have been used to predict the risk of plant diseases (De Wolf and Franci, 2000; Paul and Munkvold, 2004; Kim et al., 2014). The random forest (RF) technique, which is an extension of CART, has been shown to have greater accuracy among machine learning methods (Svetnik et al., 2003; Garzón et al., 2006) and also provides a measure of the importance of each candidate predictor. However, the RF algorithm (Breiman, 2001) has not been applied in botanical epidemiology to assess the risk of disease development. This study compares multiple regression, CART, ANN and RF modelling approaches for their accuracy in predicting the risk of SNB using pre-planting factors. The objective of the study was to develop risk assessment models that can be used to guide management decisions for SNB before planting of the wheat crop.

## **2. Materials and Methods**

### *2.1. Field sites and data collection*

Field experiments were conducted at twelve sites in eleven counties in North Carolina during the 2011/2012, 2012/2013, and 2013/2014 (hereafter 2012, 2013, and 2014, respectively) growing seasons (Table 3.1). Experimental sites were chosen to represent areas with different histories of SNB, varying cropping practices and a range of weather conditions. In each year, wheat was planted at each site in a conventionally prepared field, a no-tilled field, or both. Crop production practices at each site followed standard recommendations for North Carolina (Weisz, 2013) but with no fungicide applications.

Planting was earlier in the western than in the eastern part of the state, ranging from 25 September to 8 November in each year.

For the 2012 experiment, five soft red winter wheat cultivars (Branson, Dyna-Gro Dominion, Dyna-Gro Shirley, SS8700 and SY9978) were planted at each site in a randomized complete block design with six replications. Cultivars had levels of resistance to *P. nodorum* ranging from 3 to 6 based on a scale of 1 (most resistant) to 9 (most susceptible). Resistance ratings were determined based on performance in the United States Department of Agriculture-Agricultural Research Service Septoria screening nursery in North Carolina (Anonymous, 2015). These cultivars had similar heading dates and generally possessed resistance to other foliar fungal wheat pathogens prevalent in the state.

In 2013 and 2014, two additional factors, seed treatment and seeding rate, were varied at each site. In these years, the experiment was laid out as a split-split-plot design with six replicates. Seed treatment was the main plot factor, seeding rate the sub-plot factor, and wheat cultivar the sub-sub-plot factor. Two levels of both the seed treatment factor, carboxin + thiram-treated or -untreated seed and the seeding rate, the standard rate (380 seeds m<sup>-2</sup>) versus a reduced rate (20% reduction of standard seeding rate) were evaluated. Seed was treated with imidacloprid insecticide to minimize the incidence of barley yellow dwarf virus. In 2013, four cultivars were used: Dyna-Gro 9012, Dyna-Gro Shirley, P26R20, and SS8641. These four cultivars were also planted in 2014, except for SS8641 which was replaced by USG3438. Cultivars tested in the first year of the study were replaced in subsequent years to generate a wide range of disease resistance ratings across the entire study.

At each site, longitude, latitude, previous crop, and wheat residue cover on the ground were recorded. Latitude and longitude data were determined by locating the position of the field site on Google Maps (Google Inc., CA). The number 1 was assigned to fields where wheat was the previous crop, while 0 was assigned to fields where the previous crop was not wheat. The amount of wheat residue cover on the ground was determined using the line transect method (Wollenhaupt and Pingry, 1991).

Disease severity was assessed visually on a whole-canopy basis by estimating the percentage of SNB severity in the plot (Mehra et al., 2015). Two to four assessments were made at most sites, while in a few cases, only one assessment was made due to either a shorter wheat season or low levels of disease. The response variable was maximum disease severity (MaxDS), i.e., disease severity at the last assessment, recorded around Zadoks growth stage 85, which is the soft dough stage (Zadoks et al. 1974), and averaged over the six replications. Values of MaxDS predicted from various modelling approaches were later categorized to generate a binary predicted response variable of low disease severity ( $<30\%$ ) and high disease severity ( $\geq 30\%$ ). The 30% severity cutoff on a whole-canopy basis corresponds to approximately 20% disease severity on the flag leaf (L.K. Mehra, unpublished results), which has been shown to result in yield loss in wheat (Bhathal et al., 2003). Thus, 30% disease severity is a useful threshold for risk assessment and making management decisions for SNB.

## *2.2. Modelling approach*

Predictive modeling tries to find good rules for predicting the response variable based on the value of predictor variables in the dataset. In this study, MaxDS from a unique combination of predictor variables or disease cases (Table 3.2) was considered as the response variable. A total of 431 disease cases were obtained from the three years of the study across all experimental sites. Each modelling technique described below involved two independent steps. In the first step, the entire dataset was split randomly into training (70%), validation (20%) and test (10%) datasets using the procedure SURVEYSELECT in SAS (version 9.4, SAS Institute, Cary, NC). This splitting was conducted 15 times (i.e., 15 randomizations of the data). Each time, a model was developed using the training dataset and optimized using the validation dataset, and the predictive ability of the model was tested on the test dataset. In the second step in the modelling process, the final model was developed using all the disease cases collected in the study.

Two predictive modelling techniques, multiple regression and machine learning models, were applied in this study to predict the risk of SNB based on pre-planting variables. Within the machine learning paradigm, three predictive models were selected, ranging from the simple classification and regression trees to the more complex Breiman's random forest algorithm. To our knowledge, this is the first use of the RF algorithm to predict disease risk in plant-based agricultural systems.

During exploratory data analysis, previous crop, tillage type, and wheat residue were found to be highly correlated, while the three variables were not correlated with other pre-

planting factors. Thus, the SAS PROC VARCLUS with option MAXEIGEN = 0.9 was used to eliminate two redundant predictors (Nelson 2001). The VARCLUS variable reduction procedure identifies clusters of variables that are highly correlated among themselves but as uncorrelated as possible with variables in other clusters. Subsequently, wheat residue and five other pre-planting factors (Table 3.2) were considered as independent variables in the ensuing modeling exercise.

### 2.2.1 Multiple regression (MR) model

Regression analysis is one of the most popular techniques for predictive modelling. A multiple regression model with more than one predictor can be written as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + \varepsilon, \quad (1)$$

where  $y$  is the response variable (i.e., MaxDS),  $\beta_i$  is the regression coefficient,  $x_i$  is the  $i^{\text{th}}$  pre-planting predictor for  $i = 1, 2, \dots, m$ , and  $\varepsilon$  is the random error term.

In the first step of the modelling process, the MR model was implemented using the SAS procedure GLMSELECT with the BACKWARD variable selection method and sub-options CHOOSE = validate, STOP = validate and MAXSTEP = 26. The type and designation of predictor variables evaluated in the MR model is summarized in Table 3.2. To allow for direct comparison of model coefficients, latitude and longitude were standardized to have a mean of zero and a standard deviation of 1.0 prior to regression analysis (Schielzeth, 2010). Interactions between qualitative and quantitative predictors, and quadratic terms of quantitative predictors, were also included in the base model. The models obtained

from the 15 randomizations of the training dataset were used to make predictions for the test dataset, and the prediction accuracy of the MR model was determined by linear regression of observed against predicted values of MaxDS.

In the second step of the modelling process, the final MR model was developed as described in the first step above using all the 431 disease cases. A significance level of 0.1 was used as the basis for variable retention. If a quadratic effect or an interaction was significant after variable selection, the main effects of variables comprising the quadratic effect or the interaction were included in the model to preserve the hierarchy. In the final step of the modelling approach, predicted values of MaxDS were assigned to the low- or high-disease severity class based on a disease severity threshold of 30%. The proportion of correctly classified cases, specificity (the proportion of true positives), and sensitivity (the proportion of true negatives) were then calculated.

### *2.2.2 Artificial neural network (ANN) model*

Artificial neural network (ANN) models are analytical techniques that were originally developed by researchers attempting to mimic the neurophysiology of the human brain (Ripley, 1996). These models predict new cases after going through a learning process with existing data. An ANN is commonly divided into three or more layers: an input layer, a hidden layer(s), and an output layer (Fig. 3.1). The input layer contains the input nodes (the input variables or predictors for the network), while the output layer contains the desired output of the system, and the hidden layer contains a series of nodes that are associated with

transfer (or activation) functions. Each layer of the ANN is linked by weights that are determined through a learning algorithm.

A three-layer feed-forward network with back-propagation learning algorithm was used to predict the risk of SNB based on pre-planting predictors (Ripley, 1996; Franci, 2004). In the back-propagation algorithm, the weights of the network are trained by minimization of an error function ( $E$ ) of the form:

$$E = \frac{1}{2} \cdot [\sum (t_p - y_p)^2], \quad (2)$$

where  $t$  and  $y$  are the predicted and actual observed outputs, respectively, of the  $p^{\text{th}}$  training pattern (Rojas, 1996). A single-hidden-layer architecture was used, with the number of nodes in the hidden layer serving as a tuning hyper-parameter of the whole model (Sarle, 2002). The logistic function of the form:

$$L(\mathcal{G}) = \frac{1}{1 + e^{-\mathcal{G}}}, \quad (3)$$

was used as the activation function of the hidden layer that transfers the summed inputs to the output layer in which

$$\mathcal{G} = \sum_{i=1}^n w_i x_i + \theta_i, \quad (4)$$

where  $w_i$  is the weight of the input value,  $x_i$  and  $\theta_i$  is the bias term (Venable and Ripley, 1999). The importance of predictors in the ANN model was determined using the Garson algorithm (Garson, 1991), as a recent study comparing methods for quantifying variable importance in ANN found this algorithm to be the most accurate (Fischer, 2015). The



algorithm determines the relative importance of a variable by partitioning absolute values of hidden-output weights into components associated with each variable node and the importance of all variables sums to 100%. A weight decay value of 0.001 was used to avoid overfitting of the ANN by penalizing large weights that could increase the variance of output (Bishop, 1995). The ANN model was implemented in the R statistical computing environment using the *caret* package. The *tuneGrid* function in *caret* was used to determine the number of nodes required in the hidden layer for optimal performance of the model. As recommended by Ripley (1996), the 100 ANN were fitted and the output from all the networks was averaged using the *avNNet* function within the *caret* package in R version 3.2.2 for Windows.

### 2.2.3 Classification and regression tree (CART) model

In the CART modelling technique, an empirical tree represents a segmentation of the data that is created by applying a series of simple rules. CART models generate rules that can be used for prediction through a repetitive process of splitting. Given a training dataset  $\mathbf{L}$  with  $N$  cases, consisting of  $m$  predictors  $\mathbf{X}_i$  ( $i=1,...,m$ ) as the input space  $\mathbf{X}$  and the response variable  $y$ , the CART algorithm recursively partitions the input space to obtain a tree predictor (with  $y'_\eta$  as the predicted response for the sample  $\mathbf{X}_\eta$ ):

$$y'_\eta = T_L(\mathbf{x}_\eta). \quad (5)$$

Starting with the entire input space  $\mathbf{X}$ , CART attempts to find a binary partition to increase the response purity in the subspaces formed by the partition. The partition is defined

as a hyperplane perpendicular to one of the coordinate axes of  $\mathbf{X}$ . The purity of the resulting subspaces depends on the homogeneity of the response classes. Several criteria are available to facilitate selection of the binary splits, depending on whether  $y'_\eta$  is a categorical or continuous response (Breiman et al., 1984). Binary partitioning is repeated in each new subspace until subspace response homogeneity is achieved. The maximal tree is usually overfitted and algorithms are used to constrain the overfitting by pruning the tree to its best generalization size. The prediction for a particular subspace is the majority vote (for classification if  $y'_\eta$  is categorical) or the average (for regression if  $y'_\eta$  is continuous) of training responses in that subspace.

In this study,  $y'_\eta$  was a continuous response variable and thus, a regression tree within the CART modelling approach was implemented using the ‘Decision Tree’ method in the ‘Partition’ modeling option in the JMP Pro statistical package (v11.2, SAS Institute, Cary, NC). Given that the response variable  $y$  was a continuous response, binary partition was based on maximizing the *LogWorth* statistic:

$$LogWorth = -\log_{10}(p\text{-value}), \quad (6)$$

where  $p$ -value is the probability calculated from the sum of squares due to the differences in the means of the two groups formed from the partition (Su et al., 2009). A ‘validation’ column was provided to differentiate between the training, validation, and test datasets. The training and validation datasets were used to avoid overfitting the tree and to stop the splitting of tree nodes automatically when the coefficient of determination ( $R^2$ ) from the

validation subset was better than the next ten splits (Breiman et al., 1984). A minimum split size of eight was specified in this study. The prediction formula was saved in the JMP data table, and used to predict MaxDS for the test subset. The prediction accuracy of the CART model was determined by simple linear regression of observed on predicted MaxDS values for the test dataset. Predicted values of MaxDS were then assigned to the low- or high-disease class, and the proportion of correctly classified cases, specificity, and sensitivity of the model were calculated as described above.

#### 2.2.4 Random forest (RF) model

A random forest is a collection of tree predictors  $T_L(\mathbf{x}_\eta; \theta_k)$ , where  $K$  is the number of trees indexed by  $k = 1, \dots, K$ ;  $\mathbf{x}_\eta$  is defined as above and has a vector length  $p$  associated with input vector  $\mathbf{X}$ ; and  $\theta_k$  are independent and identically distributed random vectors that indicate a training dataset  $\mathbf{L}$ . The dataset  $\mathbf{L}$  is assumed to be independently drawn from the joint distribution of  $(\mathbf{X}, Y)$  and comprises  $\eta$   $(p + 1)$ -tuples  $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_\eta, y_\eta)$ . When the response is a continuous variable as in the present study, the final predictor  $y_\eta$  for a sample  $\mathbf{x}_\eta$  is the average over all trees:

$$y'_\eta = \frac{1}{K} \sum_{k=1}^K T_L(\mathbf{x}_\eta; \theta_k). \quad (7)$$

As  $K \rightarrow \infty$ , the Law of Large Numbers ensures

$$E_{\mathbf{x}, Y}[Y - y'_\eta(\mathbf{X})]^2 \rightarrow E_{\mathbf{x}, Y}[Y - E_\theta y_\eta(\mathbf{X}; \theta)]^2, \quad (8)$$

in which the quantity on the right is the prediction error, and convergence in that equation ensures the lack of overfitting in random forests (Breiman, 2001).

In this study, the RF model was developed using the *randomForest* package in the R environment (Liaw and Wiener, 2002). The RF classifier requires the definition of two parameters for generating a prediction model: the number of classification trees desired ( $K$ ), and the number of prediction variables ( $m$ ) to select randomly at each node to make the tree grow. Here,  $k = \{1, 2, \dots, K\}$  trees were grown in the forest and the final predictor was the average across  $K$  trees. At each node within a tree, a given number of predictors was randomly chosen, and the best predictor was used to split the node. RF uses the Gini index to split a node and selects the split with the lowest impurity at each node (Breiman et al., 1984). The process was repeated across the subsequent nodes to grow the tree.

Each tree was developed using approximately two-thirds of cases as training dataset  $\mathbf{L}$ , which was used to make a prediction for the remaining one-third of cases as the “out-of-bag” dataset. To control variance and overfitting, the number of predictors used at each node ( $m = 1$  to 6) was evaluated using the function *tuneRF* of *randomForest* package in R and optimized using the “out-of-bag” error estimate (Liaw and Wiener, 2002). The  $R^2$  of the prediction on the out-of-bag dataset was taken as the prediction accuracy of the tree. A test dataset was also used in order to compare the RF model with the MR, ANN and CART models. The RF algorithm also provides a measure of variable importance in the modelling, and the importance is derived from the contribution of each variable accumulated along all nodes and all trees where it is used (Breiman, 2001). The predicted values of MaxDS were

assigned to the low- or high-disease class. The proportion of correctly classified cases, the specificity, and the sensitivity of the model were calculated as described above.

#### *2.2.5 Assessment of model performance*

The Receiver Operating Characteristics (ROC) curve served to evaluate the performance of the models. Specifically, we estimated the area under the ROC curve (AUC), a threshold-independent index widely used in ecology. The ROC is based on the concept of class-dependent accuracy, which can be tabulated through a confusion matrix (McPherson et al., 2004). Points on the ROC are defined by the sensitivity and specificity indicators. The AUC ranges from 0.5 (random accuracy) to a maximum value of 1, which represents the most accurate model theoretically achievable. Two additional measures were calculated for each model: the coefficient of determination from regression of observed on predicted disease values, and the Kappa statistic (Monserud and Leemans, 1992). Kappa ( $\kappa$ ) is a measure of agreement of model prediction beyond random chance and has a range of  $\kappa = 0$  to 1. Values of  $\kappa$ :  $<0.4$  = low degree of similarity,  $\kappa$ : 0.40 to 0.55 = acceptable degree of similarity, while  $\kappa$ : 0.55 to 0.70 = good, 0.70 to 0.85 = very good, and  $> 0.85$  = excellent agreement beyond random chance (Monserud and Leemans, 1992).

### **3. Results**

A total of 431 unique disease cases were obtained from the field experiments in North Carolina, with 35, 236, and 160 cases recorded in 2012, 2013, and 2014, respectively. Based

on our defined disease threshold, 297 (69%) were classified as low-disease cases, while 134 (31%) were classified as high-disease cases. The high-disease class contained 11, 42, and 19% of total cases in 2012, 2013, and 2014, respectively. Different modelling approaches were used to determine pre-planting factors that influenced the severity of SNB, and those factors were subsequently used to predict the risk of SNB. The results obtained for each predictive method and the accuracy of models developed are presented below.

### *3.1. Multiple regression (MR) model*

The results of the MR analysis indicated that specific pre-planting predictor variables significantly contributed to the MR model ( $F$ -statistic = 51.55;  $P < 0.0001$ ,  $n = 431$ ). Cultivar resistance, longitude and wheat residue were the most important pre-planting factors identified by MR to influence the severity of SNB (Table 3.3). The quadratic effect of longitude also influenced the risk of SNB. Latitude, seed treatment with fungicide, and seeding rate were not significant predictors of MaxDS and hence were not included in the final model.

Based on 15 randomizations of the test dataset, the proportion of variability ( $R^2$ ) of MaxDS in the test dataset explained by factors identified by MR was low with an average of 0.32 (Fig. 3.2A). When MaxDS in the test dataset was classified as low- or high-disease, the average correct classification rate of the MR model was 0.74 (Fig. 3.2B). The specificity of the model was very high with an average rate of 0.91 (Fig. 3.2C), while average sensitivity was the lowest among four modeling approaches with an average rate of 0.38 (Fig. 3.2D).

The final MR model developed using all the 431 disease cases in the study explained 33% of the variation in MaxDS (Table 3.3). The final MR model had a correct classification rate of 0.74 (Table 3.4). The sensitivity of the final MR model was low with a value of 0.40, while the specificity was very high with a value of 0.90 (Table 3.4).

### *3.2 Artificial neural network (ANN) model*

Analysis of the disease cases using the ANN methodology indicated that latitude, longitude, wheat residue, and cultivar resistance were the most important predictor variables (in decreasing order) that affected disease severity, with relative importance values ranging from >10 to 32% (Fig. 3.3). Seeding rate and seed treatment were found to be of less importance (<10%) in the ANN (Fig. 3.3). Increasing the number of nodes in the hidden layer reduced the root mean squared error (RMSE) and the final architecture of the ANN model was optimized at 12 nodes in the hidden layer based on lowest cross-validated RMSE.

Based on 15 randomized test datasets, the ANN model identified factors that explained a moderate proportion of the variability in MaxDS, with a mean of 0.63 (Fig. 3.2A). MaxDS in the test dataset was correctly classified as low- or high-disease at an average rate of 0.78 (Fig. 3.2B). The specificity of the ANN model was high with a mean rate of 0.91 (Fig. 3.2C), while sensitivity of the model was low with an average rate of 0.49 (Fig. 3.2D). When all the disease cases in the study were used for model development, the final ANN model accounted for 73% of the total variation in MaxDS (Table 3.3). Values of MaxDS predicted by the final ANN model and assigned to the low- or high-disease classes

resulted in a correct classification rate of 0.83 (Table 3.4). The sensitivity of the final ANN model was moderate with a rate of 0.60, while the specificity was very high with a rate of 0.93 (Table 3.4).

### *3.3 Categorical and regression tree (CART) model*

The final tree chosen based on the lowest Akaike's Information Criterion had 25 nodes. The tree was further pruned to 7 terminal nodes without compromising the classification ability of the tree. The correct classification rates for the 7-node and 25-node trees were 0.83 and 0.85, respectively. Longitude, latitude, cultivar resistance, and wheat residue were identified by the CART model as the most important predictor variables influencing the severity of SNB, and were used in the final tree model (Fig. 3.4). Seeding rate and seed treatment were not identified as important factors affecting disease severity and were not used in the final tree model.

The proportion of total variability in MaxDS explained by the CART model based on the test dataset was high with an average rate of 0.72 (Fig. 3.2A). The correct classification rate of MaxDS in the test dataset was also high with a mean rate of 0.80 (Fig. 3.2B). Based on the test dataset, the specificity of the CART model was high with an average rate of 0.88; while average sensitivity of the model was moderate with a mean rate of 0.63 (Fig. 3.2). The proportion of variation in MaxDS explained by the final CART model was 0.47 (Table 3.4). The average rate at which predicted values of MaxDS were correctly assigned to the low- or high-disease class based on the final CART model was high with a rate of 0.83. The final



CART model had a moderate degree of sensitivity with a value of 0.55, while the specificity of the model was the highest with a value of 0.95 (Table 3.4).

### *3.4 Random Forest (RF) model*

As expected, the number of trees in the RF model influenced the proportion of variability in MaxDS explained by the model. Increasing  $K$  from 1 to 30 trees increased  $R^2$  from 0.74 to 0.77 (Fig. 3.5) and an additional increase in  $K$  from 31 to 100 resulted in a marginal increase in  $R^2$  to 0.79. The value of  $R^2$  stabilized with  $K$  ranging from 150 to 300 trees and the final RF model was obtained by aggregating 300 base models. The optimized number of variables used at each split in the final RF model was  $m = 3$ . In decreasing order of importance, the variables were longitude, wheat residue, cultivar resistance, and latitude (Fig. 3.5). Like all other modeling methods, seeding rate and seed treatment were not identified by the RF model as having an important effect on disease severity.

Based on randomized test datasets, the proportion of variability in MaxDS explained by the RF model was significantly ( $P < 0.05$ ) higher than that of all other models tested with an average of rate of 78% (Fig. 3.2A). The average correct classification rate of MaxDS in the test dataset for the RF was 0.81 which was significantly ( $P < 0.05$ ) higher than that of the MR and ANN models, but not significantly different from that of the CART model (Fig. 3.2B). Based on the test dataset, the average specificity of the RF model was 0.87, which was significantly ( $P < 0.05$ ) lower than that of other models tested except the CART model (Fig. 3.2C). The sensitivity of the RF model was the highest among the four modelling techniques

with a mean of 0.67, which was significantly ( $P < 0.05$ ) higher than the sensitivity of other models except the CART model (Fig. 3.2D). The final RF model explained the highest amount of variation in MaxDS with a value of 79% (Table 3.4). The rate at which predicted values of MaxDS were correctly assigned to the low- or high-disease classes in the final RF model was 0.81. The final RF model also had the highest sensitivity among all the models tested with a rate of 0.69 (Table 3.4).

### *3.5 Model performance*

The accuracy of the models was assessed based on coefficients of determination of the final models, sensitivity-specificity ROC analysis and the Kappa statistic. As indicated above, the final RF model had the highest coefficient of determination followed by the ANN and CART models, while the  $R^2$  value for the MR was the lowest (Table 3.4). The RF also had the highest area under the ROC curve ( $AUC = 0.93$ ), followed by the ANN and CART models, while AUC value for the MR model was the lowest with an AUC of 0.81 (Table 3.4). The MR model had the lowest Kappa value ( $\kappa = 0.37$ ), which indicates a low degree of agreement of MR model predictions beyond random chance. However, Kappa values for CART, ANN and RF models were comparably higher with values ranging from  $\kappa = 0.55$  to 0.57 (Table 3.4), which indicates an acceptable degree of agreement of the three machine learning model predictions beyond chance.

#### **4. Discussion**

Pre-planting factors previously correlated with the late-season severity of SNB were used to develop risk assessment models that could be useful in making disease management decisions prior to planting of the wheat crop. Two techniques, multiple regression and machine learning models, were used to find good rules for predicting MaxDS from eight pre-planting predictor variables using data collected across diverse ecological conditions, disease histories and cropping practices in North Carolina. Models developed in this study identified longitude, latitude, cultivar resistance and amount of wheat residue as significant predictors of SNB severity. Assessment of the accuracy of the models using the ROC curve analysis showed that the RF model was the most accurate classifier for assessing the risk of SNB. To our knowledge, this work represents the first use of RF to predict disease risk in plant-based systems and the first study to develop pre-planting risk assessment models for SNB in wheat.

In winter wheat, the role of field location, previous crop, cultivar resistance, amount of wheat residue on the soil surface, seed treatment and tillage in the development of SNB is well documented (King et al., 1983; Luke et al., 1983; Leath et al., 1993; Stover et al., 1996; Milus and Chalkley, 1997; Solomon et al., 2006; Weisz, 2013). However, the relative importance of each of these factors to the severity and risk of SNB has never been determined. In addition, a clear understanding of the most important pre-planting factors that influence the risk of SNB was previously lacking. The three machine learning models developed in this study identified longitude, latitude, cultivar resistance, and wheat residue as significant predictors of the SNB severity. The MR model also identified all these factors,

except latitude, as significant predictors of SNB. These results also validate previous reports on the effect of location, cultivar resistance and wheat residue on the severity SNB (e.g., Luke et al., 1983; Holmes and Colhoun, 1975; Weisz, 2013), and indicate that these predictors are useful in predicting the risk of SNB in winter wheat.

SNB is frequently problematic in the western and northeastern parts of North Carolina (Weisz, 2013). Thus, it is not surprising that longitude was an important predictor of SNB in North Carolina, with all the models identifying longitude as one of the two most important predictors, along with latitude. The effect of longitude on SNB severity can be seen directly in the MR model by the significant quadratic effect of standardized longitude which indicates that predicted disease severity is higher in the east and west and lower in the central parts of the state. This was especially evident under conventional tillage, where high levels of SNB were observed in western parts of the state (data not shown). The importance of longitude as a predictor can be explained by variation in environmental conditions that favor the development of SNB. In this study, rainfall amounts recorded in experimental sites in the eastern or western part of the state were 10 to 80% higher than in sites located in the central part of the state (Mehra et al., unpublished results). The level and frequency of precipitation are both known to favor the development of SNB in wheat (Verreet and Hoffmann, 1990). The median latitude was slightly higher in the high-disease class than in the low-disease class but the effects of latitude on SNB severity were highly dependent on the previous crop.

Previous crop, tillage and wheat residue were highly correlated and wheat residue was identified as the best predictor for the risk of SNB among these three pre-planting factors. The limited importance of previous crop and tillage could be explained by the fact that these two variables are an indirect measure of the survival of *P. nodorum* from one season to the next, a characteristic that is better reflected by wheat residue. Similar observations were also reported in a study that evaluated the importance of pre-planting factors for the risk of gray leaf spot of maize (Paul and Munkvold, 2004). In addition, even under conventional tillage with incomplete burial of residue, 10% of residue remains on the soil surface (Stubbs et al., 2004), which can result in an end-of-season disease severity similar to that in fields with 30% residue that can result from no-till fields (Mehra et al., 2015). The amount of residue in no-till fields also depends on the previous crop, with higher residue levels resulting when wheat is planted after wheat as compared to when another crop is planted between two wheat crops. The tillage effects on disease severity in the latter case would not be easily distinguishable from the effects of tillage with complete burial of residue since both practices would result in little or no residue to influence disease severity.

None of the modeling approaches identified seed treatment or seeding rate as useful predictors of SNB risk. The seed used in the present study was certified and was free of *P. nodorum*, which explains why seed treatment was not an important factor. The reduced seeding rate used in this study may not have been enough lower to generate a significance difference in SNB compared to the normal seeding rate. Further reductions in seeding rates

may result in high SNB compared to the normal rates, but growers are not likely to use such low seeding rates due to potential yield penalties.

Predicting the risk of disease plays an important role in the decision making and planning process for disease management in plant pathosystems. Historically, regression models have been widely used to predict epidemics of plant diseases (Rosso et al., 2003; Uddin et al., 2003; Paul and Munkvold, 2005; Del Ponte et al., 2006; Olatinwo et al., 2008). In this study, while the MR model had a very high specificity and thus, was a very good predictor of low disease severity at the end of the season, the model explained less than 50% of total variability in the dataset. The MR model was also a poor predictor of the high-disease class as evidenced by its low sensitivity. The MR model was also the least accurate in predicting the risk of SNB as indicated by its lowest AUC. Highly accurate decision rules that combine high levels of sensitivity and specificity will be required for any predictive system to be useful in management of plant diseases (Gent et al., 2013). Regression analysis has the advantage of simplicity and produces a model equation with parameter estimates that can be directly related to scientific hypotheses and thus, has been the main choice for modeling disease risk in botanical epidemiology. However, other approaches such as artificial neural networks and decision trees have been useful in developing predictive models in several fields (Gutierrez, 2015). Application of these alternative methods has been limited but is slowly gaining interest in plant disease epidemiology.

Artificial neural networks have been used to model the risk of disease development in plant-pathosystems (Batchelor et al., 1997; De Wolf and Franc, 2000; Chakraborty et al.,

2004). However, applications of ANN in these systems did not establish the relative importance of predictor variables. The ANN algorithm used in the present study allowed for estimation of the relative importance of each predictor variable in assessing the risk of SNB. The ANN model developed here performed much better than the MR model, with a good balance between model sensitivity and specificity. Often, prediction is more important than explanation in standard back-propagation ANN models, and model construction is not easily understood (Frasconi et al., 1993; Hastie et al., 2009), which has created a perception of a 'black box' that has limited the use of ANN models. The inability to easily calculate standardized coefficients for each independent variable and the difficulty in interpreting weights from the network analysis are also other weaknesses of ANN models (Frasconi et al., 1993; Ottenbacher et al., 2004). In this study, transparency was increased and the explanatory power of the ANN model was improved by determining the relative contribution and importance of each predictor variable to the prediction of SNB severity. The number of nodes in the hidden layer required in the optimized model was twelve, which was greater than the number of important predictors in the model. This suggests that the relationship between MaxDS and pre-planting variables is nonlinear and hence will not be fitted well by the MR model without adding higher-order interactions and polynomial terms. In contrast, nodes in the hidden layer of the ANN model intrinsically captured the nonlinearity between MaxDS and predictor variables.

The CART modelling technique has previously been applied to predict the risk of disease in plant- and forest-pathosystems (Rosso and Hansen, 2003; Paul and Munkvold,

2004; Fan et al., 2006; Kelly et al., 2007; Copes and Scherm, 2010; Kim et al., 2014). In this study, the CART model performed better than MR and was as accurate as the ANN model in predicting the risk of SNB. Simplicity of the modelling approach is one notable attribute of CART that allows for determination of variable importance at each node (De'ath and Fabricius, 2000). CART also generates an intuitive tree diagram that illustrates the relationship between the response and the predicted variables. The tree indicated that several combinations of predictor variables could result in the same disease-severity class, and that longitude, latitude, wheat residue and cultivar resistance strongly influenced MaxDS. High levels of SNB occurred west of longitude 80.6, which is the Piedmont and foothills of the Appalachian Mountains, while low and high severity occurred east of longitude 80.6 (the Coastal Plain and Tidewater regions) depending on the latitude, amount of wheat residue and cultivar resistance. For example, east of longitude 76.7 (in the Tidewater), low disease occurred north of latitude 36.2 (roughly, north of the Albemarle Sound), while high disease occurred south of latitude 36.2 when a highly susceptible cultivar was planted. The final CART model, pruned to 7 terminal nodes, predicted SNB classes as well as the fully grown tree with 25 terminal nodes, and thus, it is likely to generalize well on an independent dataset (Breiman et al., 1984).

The RF algorithm, which has previously not been used to predict the risk of disease in plant-based agricultural systems, produced the most accurate model to predict the risk of SNB in winter wheat. Like the other models evaluated, RF identified location, wheat residue and cultivar resistance as the key predictors affecting the risk of SNB. The key advantages of



RF include its non-parametric nature, high classification accuracy, and capability of determining variable importance. However, it can be difficult to understand the rules used to generate the final classification due to the large number of trees generated from resampling the same dataset. The number of trees,  $K$ , and predictor variables used at each node,  $m$ , influence the accuracy of the RF classifier. In this study, different values of  $K$  (1 to 300) and  $m$  (1 to 6) were evaluated to optimize these parameters in the final classifier for a total of 1,800 different RF models to predict the risk of SNB. As  $K$  is increased, the generalization error decreases and converges to a limit (Breiman, 2001). However, the value of  $m$  (which is constant during forest growth) affects both the correlations between the trees and the strength of the individual trees. Reducing  $m$  reduces correlation and strength, while increasing  $m$  increases both. Thus, it is preferable to use a large value for  $K$  and a small value for  $m$  to reduce the generalization error and correlation between trees in the forest.

Historically, wheat prices have been low in the United States, which has reduced profit margins for growers (Weisz et al., 2011). Thus, only the most accurate predictor models that guide pre-planting management decisions to minimize unprofitable spray application are likely to be acceptable to risk-averse producers as decision tools. In this regard, the RF model could be a useful pre-planting decision management tool for SNB, as it performed better during internal validation than the other models developed in this study. The model can be used to guide the selection of a specific combination of pre-planting factors that will result in a reduce risk of SNB. In addition, where cultural management practices such as tillage type and crop rotation are difficult to alter in order to reduce the

amount of wheat residue, the pre-planting model can provide a quantitative assessment of SNB risk in those situations to facilitate informed decision-making. Although the machine learning models and especially the ANN and RF models developed in this study had a high internal validation accuracy, the models need to be validated with independent data before they can be integrated into a management program for SNB. Such an independent validation of these models would focus on using disease cases collected from growing seasons with a wide range of disease severity levels from locations with diverse cultural practices. Comprehensive economic management decisions for SNB in winter wheat can be made by combining prediction models developed in this study with yield-loss models of wheat.

### **Acknowledgments**

This study was supported by funds from North Carolina Agricultural Experiment Station Hatch Project NC02251, USDA-ARS, and NSF Ecology and Evolution of Infectious Diseases Initiative (EEID) grant number 10-15825. The authors would like to thank staff at research stations and county extension agents, for managing the experimental plots; and Ryan Parks, Bradley Van Staaldhuizen, Tyler Cole, Brandon Clontz, Ian Swaim, Charlie Glover, Barbara Stevens, Joseph Jeffreys, and Jennifer Patton-Özkurt for their assistance with field experiments.

## References

- Anonymous, 2015. USDA, Agricultural Research Service -Eastern United States Septoria Nursery. 2015. <http://www.ars.usda.gov/Main/docs.htm?docid=8419&page=8>. Accessed on April 16, 2015.
- Batchelor, W.D., Yang, X.B., Tschanz, A.T., 1997. Development of a neural network for soybean rust epidemics. *Trans. ASAE* 40, 247–252.
- Bhathal, J.S., Loughman, R., Speijers, J., 2003. Yield reduction in wheat in relation to leaf disease from yellow (tan) spot and septoria nodorum blotch. *Eur. J. Plant Pathol.* 109, 435–443.
- Bishop, C., 1995. *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford, 482 pp.
- Breiman, L., 2001. Random forest. *Mach. Learn.* 45, 5–32.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. *Classification and Regression Trees*, Wadsworth, Inc., Monterey, California.
- Chakraborty, S., Ghosh, R., Ghosh, M., Fernandes, C.D., Charchar, M.J., Kelemu, S., 2004. Weather-based prediction of anthracnose severity using artificial neural network models. *Plant Pathol.* 53, 375–386.
- Copes, W.E., Scherm, H., 2010. Rhizoctonia web blight development on container-grown azalea in relation to time and environmental factors. *Plant Dis.* 94, 891–897.
- Cowger, C., Silva-Rojas, H.V., 2006. Frequency of *Phaeosphaeria nodorum*, the sexual stage of *Stagonospora nodorum*, on winter wheat in North Carolina. *Phytopathology* 96, 860–

866.

- De'ath, G., Fabricius, K.E., 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology* 81, 3178–3192.
- Del Ponte, E.M., Godoy, C.V., Li, X., Yang, X.B., 2006. Predicting severity of Asian soybean rust epidemics with empirical rainfall models. *Phytopathology* 96, 797–803.
- De Wolf, E. D., Francl, L. J. 2000. Neural network classification of tan spot and *Stagonospora nodorum* blotch infection periods in a wheat field environment. *Phytopathology* 90, 108–113.
- De Wolf, E.D., Madden, L.V., Lipps P.E., 2003. Risk assessment models for wheat Fusarium head blight epidemics based on within-season weather data. *Phytopathology* 93, 428–435.
- Eyal, Z. 1999. Integrated control of Septoria diseases of wheat. *Plant Dis.* 65, 763–768.
- Fan, Z., Kabrick, J.M., Shifley, S.R., 2006. Classification and regression tree based survival analysis in oak-dominated forests of Missouri's Ozark highlands. *Can. J. For. Res.* 36, 1740–1748.
- Frasconi, P., Gori, M., Tesi, A., 1993. Successes and failures of back-propagation: a theoretical investigation. In: Omidvar, O., ed., *Progress in Neural Networks*. Ablex Publishing Toronto, Canada, pp. 205–242.
- Fischer, A., 2015. How to determine the unique contributions of input-variables to the nonlinear regression function of a multilayer perceptron. *Ecol. Model.* 309–310, 60–63.
- Garson, G.D., 1991. Interpreting neural-network connection weights. *AI Expert* 6, 46–51.

- Garzón, M.B., Blazek, R., Neteler, M., de Dios, R.S., Ollero, H.S., Furlanello, C., 2006. Predicting habitat suitability with machine learning models: The potential area of *Pinus sylvestris* L. in the Iberian Peninsula. *Ecol. Model.* 197, 383–393.
- Gent, D.H., Ocamo, C.M., 2009. Predicting infection risk of hop by *Pseudoperonospora humuli*. *Phytopathology* 99, 1190–1198.
- Gent, G.H., Mahaffee, W.F., McRoberts, N., Pfender, W.F., 2013. The use and role of predictive systems in disease management. *Annu. Rev. Phytopathol.* 51, 267–289.
- Gutierrez, D.D., 2015. *Machine Learning and Data Science: An Introduction to Statistical Learning Methods with R*. Technics Publications, Basking Ridge, NJ.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The Elements of Statistical Learning*, Second Edition. Springer, New York, USA.
- Holmes, S.J.I., Colhoun, J., 1975. Straw-borne inoculum of *Septoria nodorum* and *S. tritici* in relation to incidence of disease on wheat plants. *Plant Pathol.* 24, 63–66.
- Horowitz, J., Ebel, R., Uda, K., 2010. ‘No-Till’ Farming is a Growing Practice. United States Department of Agriculture-Economic Research Service. Economic Information Bulletin Number 70. Washington, DC.
- Jeger, M.J., Xu, X.M., 2015. Modelling the dynamics of a plant pathogen and a biological control agent in relation to flowering pattern and populations present on leaves. *Ecol. Model.* 313, 13–28.
- Kelly, M., Guo, Q., Liu, D., Shaari, D., 2007. Modeling the risk for a new invasive forest disease in the United States: An evaluation of five environmental niche models. *Comput.*

- Environ. Urban. Syst. 31, 689–710.
- Kim, K.S., Beresford, R.M., Walter, M., 2014. Development of a disease risk prediction model for downy mildew (*Peronospora sparsa*) in boysenberry. *Phytopathology* 104, 50–56.
- King, J.E., Cook, R.J., Melville, S.C., 1983. A review of *Septoria* diseases of wheat and barley. *Ann. Appl. Biol.* 103, 345–373.
- Krupinsky, J.M., Tanaka, D.L., Merrill, S.D., Liebig, M.A., Lares, M.T., Hanson, J.D., 2007. Crop sequence effects on leaf spot diseases of no-till spring wheat. *Agron. J.* 99, 912–920.
- Leath, S., Scharen, A.L., Lund, R.E., Dietz-Holmes, M.E., 1993. Factors associated with global occurrences of *Septoria nodorum* blotch and *Septoria tritici* blotch of wheat. *Plant Dis.* 77, 1266–1270.
- Liaw, A., Wiener, M. 2002. Classification and Regression by randomForest. *R News* 2, 18–22.
- Luke, H.H., Pfahler, P.L., Barnett, R.D., 1983. Control of *Septoria nodorum* on wheat with crop rotation and seed treatment. *Plant Dis.* 67, 949–951.
- Madden, L.V. 2006. Botanical epidemiology: some key advances and its continuing role in disease management. *Eur. J. Plant Pathol.* 115, 3–23.
- Madden, L.V., G. Hughes, and F. van den Bosch. 2007. The study of plant disease epidemics. APS, St. Paul, MN.
- McPherson, J.M., Jetz, W., Rogers, D.J., 2004. The effects of species' range sizes on the

- accuracy of distribution models: ecological phenomenon or statistical artefact? *J. Appl. Ecol.* 41, 811–823.
- Mehra, L.K., Cowger, C., Weisz, R., Ojiambo, P.S., 2015. Quantifying the effects of wheat residue on severity of *Stagonospora nodorum* blotch and yield in winter wheat. *Phytopathology* 105, 1417–1426.
- Milus, E.A., Chalkley, D.B., 1997. Effect of previous crop, seedborne inoculum, and fungicides on development of *Stagonospora* blotch. *Plant Dis.* 81, 1279–1283.
- Monserud, R.A., Leemans, R., 1992. Comparing global vegetation maps with the Kappa statistics. *Ecol. Model.* 62, 275–293.
- Nelson, B.D. 2001. Variable reduction for modeling using PROC VARCLUS. In: *Proceedings of the Twenty-Sixth Annual SAS® Users Group International Conference*, Cary, NC. Paper No. 261-26. <http://www2.sas.com/proceedings/sugi26/p261-26.pdf>. Accessed on 10 September 2015.
- Olatinwo, R.O., Paz, J.O., Brown, S.L., Kemerait, R.C., Jr., Culbreath, A.K., Beasley, J.P., Jr., Hoogenboom, G., 2008. A predictive model for spotted wilt epidemics in peanut based on local weather conditions and the *Tomato spotted wilt virus* risk. *Phytopathology* 98, 1066–1074.
- Ottenbacher, K., Linn, R.T., Smith PM, Illig, S.B., Mancuso, M., Granger C.V., 2004. Comparison of logistic regression and neural network analysis applied to predicting living setting after hip fracture. *Ann. Epidemiol.* 14, 551–559.
- Paul, P.A., Munkvold, G.P. 2004. A model-based approach to preplanting risk assessment for

- gray leaf spot of maize. *Phytopathology* 94, 1350–1357.
- Paul, P.A., Munkvold, G.P., 2005. Regression and artificial neural network modeling for the prediction of gray leaf spot of maize. *Phytopathology* 95, 388–396.
- Quaedvlieg, W., Verkley, G.J.M., Shin, H.-D., Barreto, R.W., Alfenas, A.C., Swart, W.J., Groenewald, J.Z. Crous, P.W. 2013. Sizing up *Septoria*. *Stud. Mycol.* 75, 307–390.
- Recknagel, F., 2001. Applications of machine learning to ecological modelling. *Ecol. Model.* 146, 303–310.
- Ripley, B.D., 1996. *Pattern Recognition and Neural Networks*. Cambridge University Press, London.
- Rojas, R., 1996. *Neural Networks: A Systematic Introduction*. Springer-Verlag, Berlin, New-York.
- Rosso, P.H., Hansen, E.M., 2003. Predicting Swiss needle cast disease distribution and severity in young Douglas-fir plantations in coastal Oregon. *Phytopathology* 93, 790–798.
- Sarle, W.S., 2002. Neural Network FAQ, part 1 of 7: Introduction. Documented online as FAQ 1473. SAS Institute, Cary, NC. Link: <ftp://ftp.sas.com/pub/neural/FAQ.html>.
- Scherm, H., Ngugi, H.K., Ojiambo, P.S., 2006. Trends in theoretical plant epidemiology. *Eur. J. Plant Pathol.* 115, 61–73.
- Schielzeth, H., 2010. Simple means to improve the interpretability of regression coefficients. *Methods Ecol. Evol.* 1, 103–113.
- Solomon, P.S., Lowe, R.G.T., Tan, K-C., Waters, O.D.C., Oliver, R.P., 2006. *Stagonospora*



- nodorum*: cause of Stagonospora nodorum blotch of wheat. Mol. Plant Pathol. 7, 147–156.
- Stubbs, T.L., Kennedy, A.C., Schillinger, W.F., 2004. Soil ecosystem changes during the transition to no-till cropping. J. Crop Improv. 11, 105–135.
- Su, X., Tsai, C., Wang, H., Nickerson, D.M., Li, B., 2009. Subgroup analysis via recursive partitioning. J. Mach. Learn. Res. 10, 141–158.
- Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P., Feuston, B.P., 2003. Random forest: a classification and regression tool for compound classification and QSAR modeling. J. Chem. Inform. Comput. Sci. 43, 1947–1958.
- Uddin, W., Serlemitsos, K., Viji, G., 2003. A temperature and leaf wetness duration-based model for prediction of gray leaf spot of perennial ryegrass turf. Phytopathology 93, 336–343.
- Venables, W.N., Ripley, B.D., 1999. Modern Applied Statistics with S-Plus (3<sup>rd</sup> edition). Srpinger, New York, 501 pp.
- Verreet, J.A., Hoffmann, G.M., 1990. A biologically oriented threshold decision model for control of epidemics of *Septoria nodorum* in wheat. Plant Dis. 74, 731–738.
- Weisz, R., 2013. Small Grain Production Guide. North Carolina Cooperative Extension Service, Publication No. AG-580 (Revised March 2013).
- Weisz, R., and Cowger, C., 2014. 2014 Wheat variety performance & recommendations. Smartgrains- The Small Grains Fact Sheet. No. 37. Online:

<http://www.smallgrains.ncsu.edu/Misc/VarietySelection.pdf>. Accessed 17 September 2015.

Weisz, R., Cowger, C., Ambrose, G., Gardner, A., 2011. Multiple mid-Atlantic field experiments show no economic benefit to fungicide application when fungal disease is absent in winter wheat. *Phytopathology* 101, 323–333.

Wollenhaupt, N.C., Pingry, J., 1991. Estimating Residue using the Line-transect Method. Publication No. A3553. Cooperative Extension, University of Wisconsin, Madison, WI.

Zadoks, J.C., Chang, T.T., Konzak, C.F. 1974. A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421.

**Table 3.1.** Description of experimental sites and tillage methods used in a study conducted in North Carolina to identify pre-planting factors that influence the risk of *Stagonospora nodorum* blotch in winter wheat

Site	Field type <sup>b</sup>	County	Region	Tillage method <sup>a</sup>		
				2012	2013	2014
Aurora	Grower	Beaufort	Tidewater	CT	- <sup>c</sup>	-
Caswell Farm	Research	Lenoir	South-Central	-	CT, NT	-
Cunningham Station	Research	Lenoir	South-Central	CT	CT, NT	CT, NT
Hertford	Grower	Perquimans	Tidewater	-	CT, NT	-
Lake Wheeler Road	Research	Wake	South-Central	CT	CT	-
Monroe	Grower	Union	Piedmont	-	NT	-
Piedmont Station	Research	Rowan	Piedmont	NT	CT	CT, NT
Tidewater Station	Research	Washington	Tidewater	NT	CT, NT	CT, NT
Tyner	Grower	Chowan	Tidewater	-	-	CT, NT
Rowland	Grower	Robeson	South-Central	CT	-	-
Upper Mt. Station	Research	Ashe	Piedmont	NT	CT, NT	-
Walkertown	Grower	Forsyth	Piedmont	-	CT, NT	CT, NT

<sup>a</sup> CT = conventional tillage with complete burial of residue, and NT = no-tillage.

<sup>b</sup> Research fields were managed by personnel on the research station, while growers' fields were managed by individual growers. In some instances, both conventional tillage and no-tillage experiments were conducted in separate fields at the same site.

<sup>c</sup> No experiment was conducted at the site in that year.

**Table 3.2.** Independent variables tested for their usefulness in assessing pre-planting risk of *Stagonospora nodorum* blotch in winter wheat

Predictor variable	Type	Designation in modelling approach <sup>a</sup>	
		CART, ANN, RF	Multiple regression
Cultivar resistance	Ordinal	1 to 9 <sup>b</sup>	1 to 9 <sup>b</sup>
Latitude	Continuous	Non-standardized	Standardized <sup>c</sup>
Longitude	Continuous	Non-standardized	Standardized <sup>c</sup>
Previous crop	Dichotomous	Wheat, no wheat	1 (wheat) or 0 (no wheat)
Seed rate	Dichotomous	Standard, reduced <sup>d</sup>	1 (standard), 0 (reduced)
Seed treatment	Dichotomous	Yes, no	1 (yes) or 0 (no)
Tillage type	Dichotomous	No-till, conventional	1 (yes), 0 (no)
Wheat residue	Continuous	Non-standardized	Non-standardized

<sup>a</sup> CART = Classification and regression tree model, ANN = Neural networks model and RF = Random forest model.

<sup>b</sup> Cultivar resistance ranges from 1 (= most resistant) to 9 (= most susceptible).

<sup>c</sup> Latitude and longitude were standardized to have mean = 0 and standard deviation = 1, prior to fitting the multiple regression model.

<sup>d</sup> Reduced seed rate is 80% of standard rate of about 380 seeds/m<sup>2</sup>.

**Table 3.3.** Results of multiple regression analysis conducted to explain variation in maximum severity of *Stagonospora nodorum* blotch in winter wheat based on pre-planting variables using data collected in North Carolina from 2012 to 2014

Variable	Estimate	Standard error	<i>t</i> -value	<i>P</i> >   <i>t</i>
Intercept	0.20	2.157	0.09	0.9276
Cultivar resistance	2.24	0.401	5.60	0.0001
Longitude (LON)	1.99	0.569	3.49	0.0005
Wheat residue	0.05	0.017	3.21	0.0014
[LON] <sup>2</sup>	10.90	0.886	12.32	0.0001

**Table 3.4.** Classification rates, sensitivity, specificity and prediction accuracy of final models developed using the multiple regression and machine learning techniques to predict the risk of *Stagonospora nodorum* blotch in winter wheat based on pre-planting variables using data collected in North Carolina from 2012 to 2014

Test statistic	MR <sup>a</sup>	Machine learning model <sup>a</sup>		
		ANN	CART	RF
Coeff. of determination ( $R^2$ ) <sup>b</sup>	0.33	0.73	0.47	0.79
Correct classification <sup>c</sup>	0.74	0.83	0.83	0.81
Sensitivity <sup>d</sup>	0.40	0.60	0.55	0.69
Specificity <sup>d</sup>	0.90	0.93	0.95	0.86
AUC <sup>e</sup>	0.76	0.91	0.89	0.93
Kappa (SE) <sup>f</sup>	0.33 (0.049)	0.57 (0.044)	0.55 (0.044)	0.55 (0.043)

<sup>a</sup> MR = Multiple regression model, ANN = Neural network model, CART = Categorical and regression tree model and RF = Random forest model.

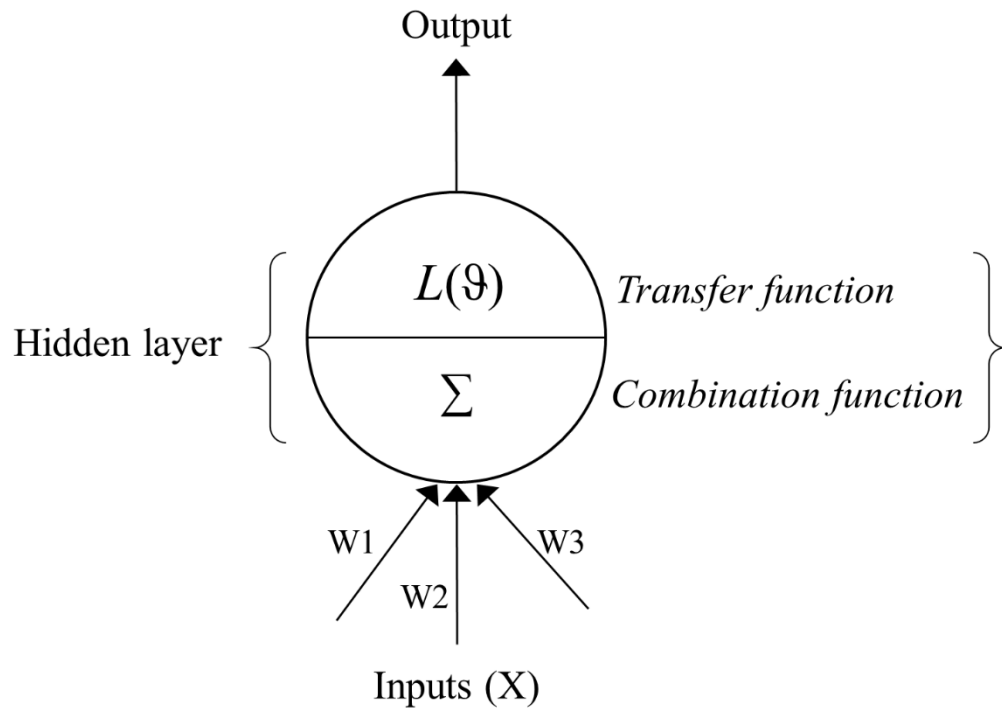
<sup>b</sup> Coefficient of determination that indicates the amount of variation in MaxDS explained by the predictor variables from regression of observed on predicted disease values.

<sup>c</sup> Correct classification rate refers to the proportion of correctly classified cases within the dataset.

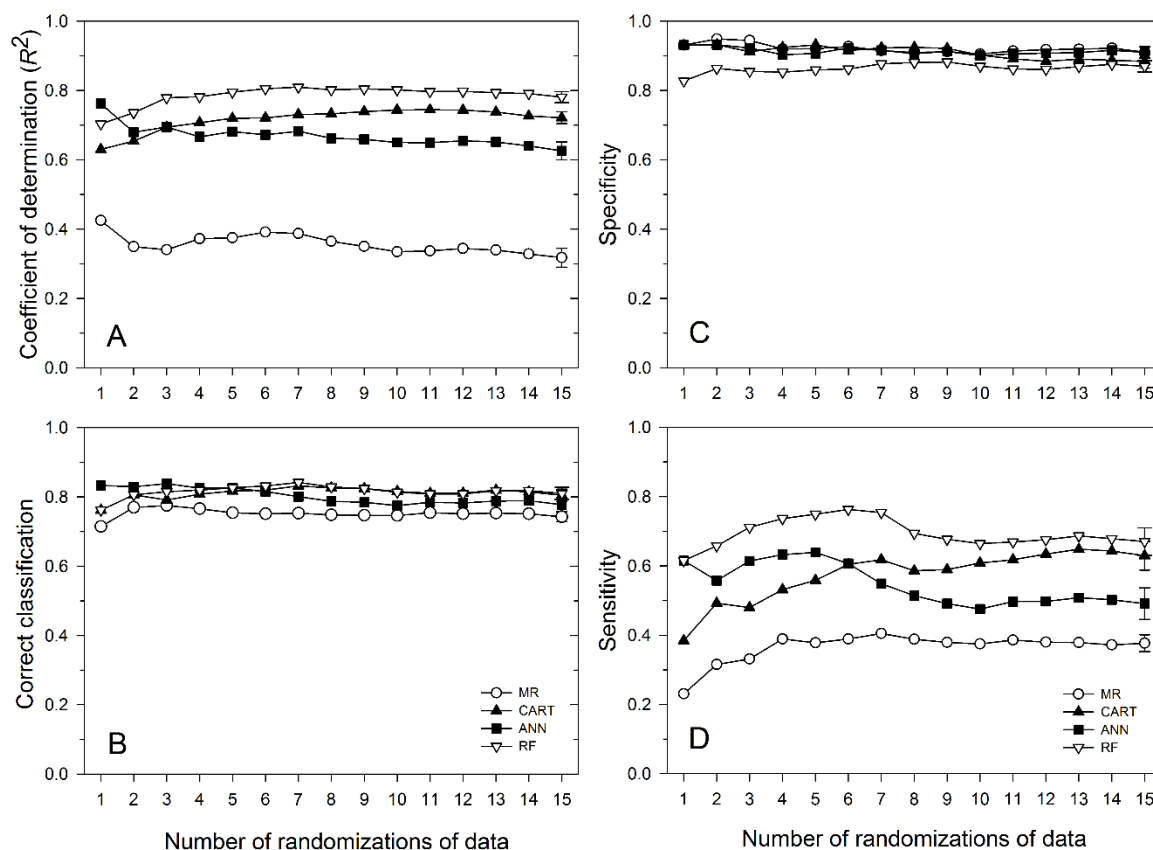
<sup>d</sup> Sensitivity is the proportion of cases correctly classified as high-disease ( $\geq 30\%$  disease severity) and Specificity is the proportion of cases correctly classified as low-disease ( $< 30\%$  disease severity).

<sup>e</sup> Area under the receiver operating curve, an estimator of the prediction accuracy of the model.

<sup>f</sup> Statistic for an estimator of the degree of agreement between observed values and model predictions beyond random chance; SE = standard error.

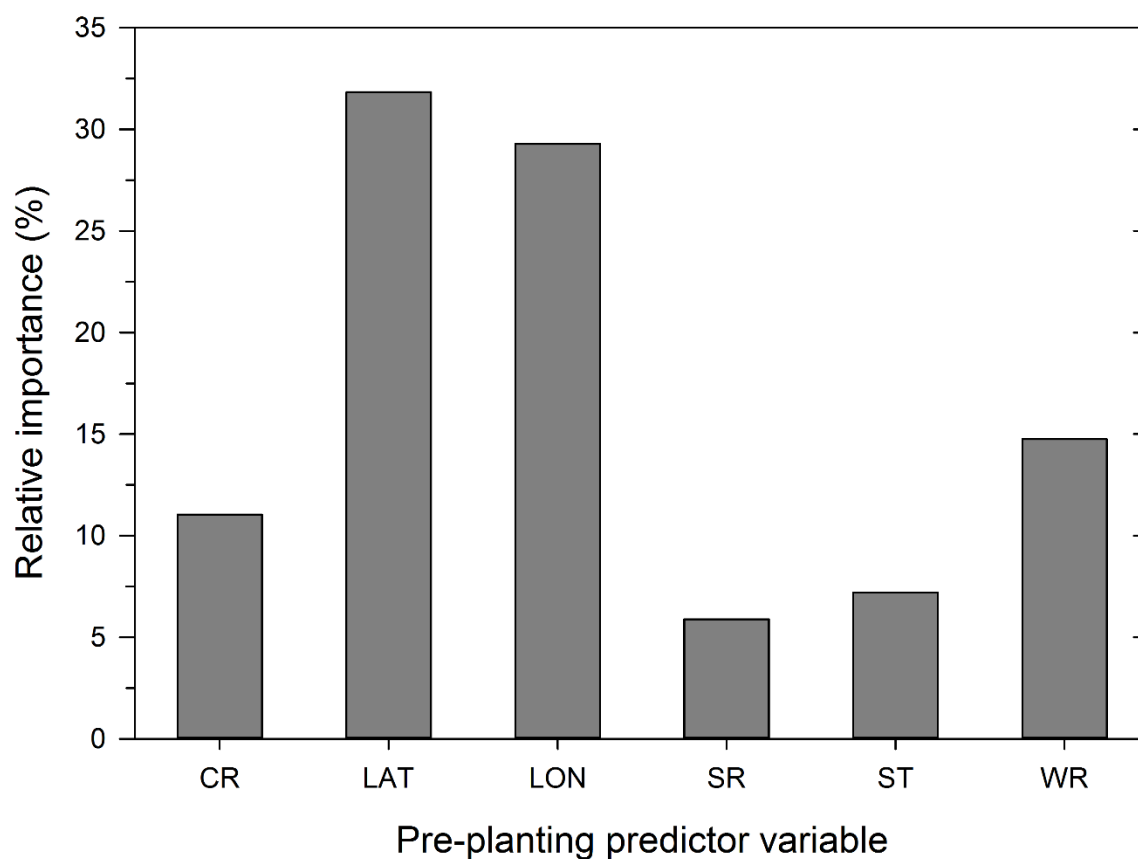


**Figure 3.1.** Schematic flow of an artificial neural network depicting the input, hidden and output layers. The input layer contains predictors, while the output layer contains the response variable. The hidden layer is composed of the combination ( $\Sigma$ ) and transfer ( $L$ ) functions and summarizes predictor variables and associated weights ( $W$ ), applies the transfer function, and sends the result to the output layer. The weights ( $W_1$ ,  $W_2$ ,  $W_3$ ...etc.) link input and hidden layer of the neural network.

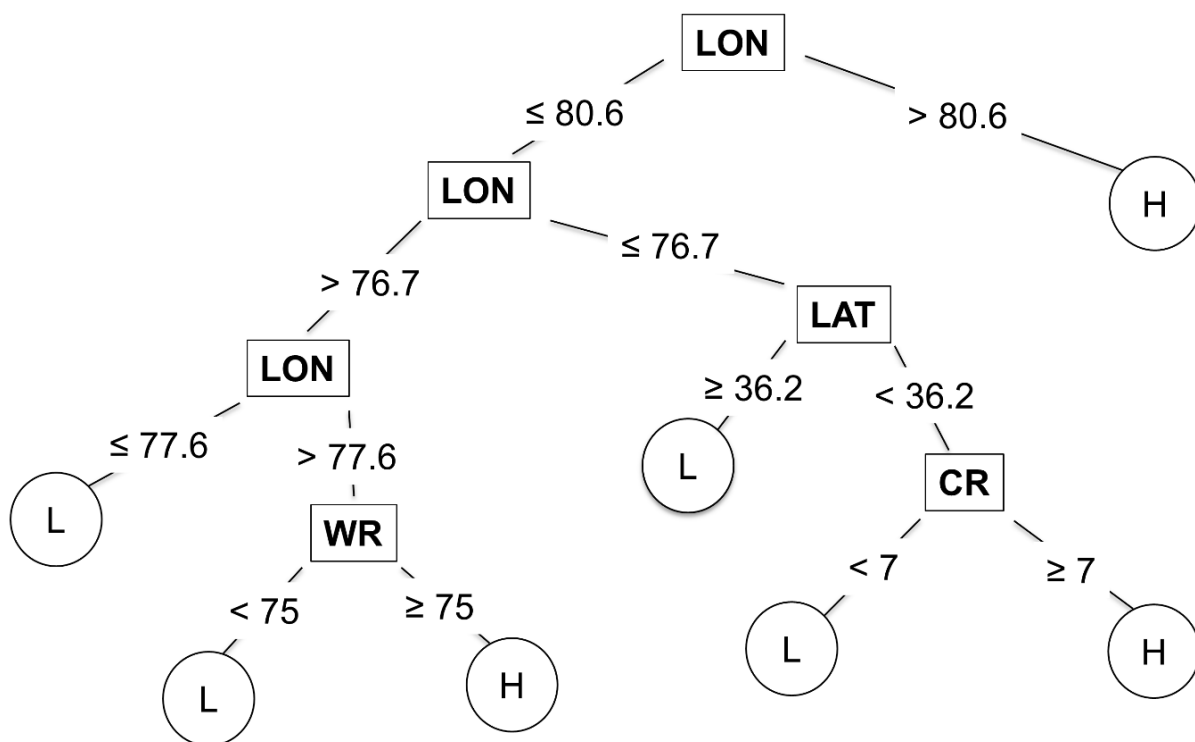


**Figure 3.2.** Performance in predicting *Stagonospora nodorum* blotch of wheat of a multiple regression (MR) model and three machine learning models: artificial neural network (ANN), categorical and regression tree (CART), and random forest (RF). Based on running averages of 15 randomizations of the test dataset which comprised 10% of 431 cases collected in the study. **A)** coefficient of determination; **B)** correct classification rate, i.e., proportion of cases correctly classified as low-severity (<30% disease severity) or high-severity ( $\geq 30\%$  disease severity); **C)** specificity, the proportion of cases correctly classified as low-disease; and **D)** sensitivity, the proportion of cases correctly classified as high-disease. Symbols represent mean values for any given number of randomizations of the test dataset, while the vertical bars represent standard errors of the mean based on the total number of randomizations evaluated.

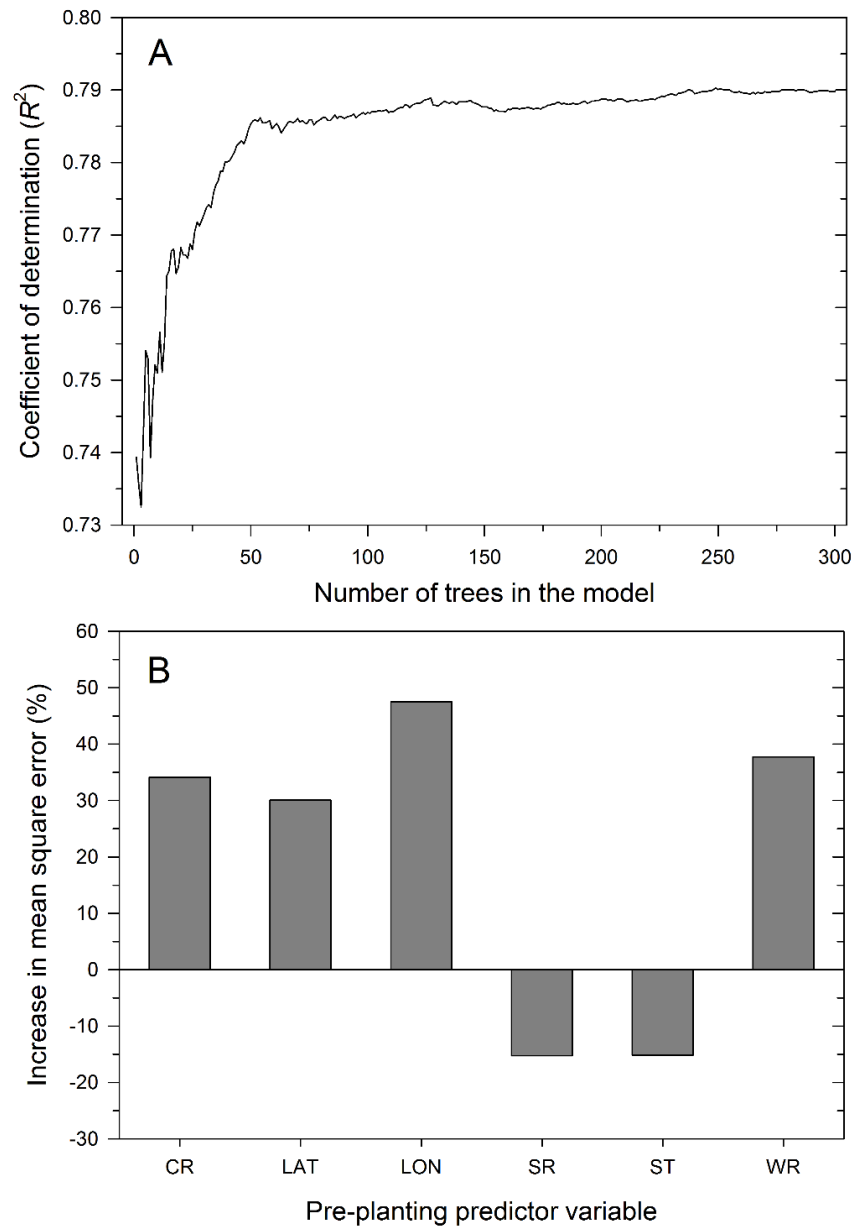




**Figure 3.3.** Importance of pre-planting variables as identified by the neural network model. Variable importance was determined using the ‘weights’ method of Garson (1991). CR = cultivar resistance rating for *Stagonospora nodorum* blotch, LAT = latitude, LON = longitude, SR = seeding rate, ST = seed treatment, and WR = wheat residue on soil surface.



**Figure 3.4.** Classification and regression tree used to estimate severity of *Stagonospora nodorum* blotch in winter wheat based on pre-planting predictor variables collected in North Carolina from 2012 to 2014. Low (L) disease class = disease severity <30%, and high (H) class = disease severity  $\geq$ 30%. The original tree had 25 nodes but was pruned to seven terminal nodes. Within the tree, predictor variables are shown in rectangles, while response variables are shown in circles. CR = cultivar resistance rating for *Stagonospora nodorum* blotch, LAT = latitude, LON = longitude (degrees W), and WR = wheat residue on soil surface.



**Figure 3.5.** Performance of the Random forest model depicting: **A)** coefficient of determination as a function of the number of trees in the model and **B)** importance of pre-planting predictor variables. The importance of a predictor variable is determined by the percent increase in the mean square error after permuting a specific predictor variable. CR = cultivar resistance rating for *Stagonospora nodorum* blotch, LAT = latitude, LON = longitude, SR = seeding rate, ST = seed treatment, and WR = wheat residue on soil surface.

## **CHAPTER 4**

### **Predicting the onset of *Stagonospora nodorum* blotch in winter wheat using pre-planting factors and in-season weather**

To be submitted to *Plant Disease*:

L. K. Mehra, C. Cowger, and P. S. Ojiambo

First and third authors: Department of Plant Pathology, North Carolina State University, Raleigh, NC 27695; second author: United States Department of Agriculture-Agricultural Research Service, Department of Plant Pathology, North Carolina State University, Raleigh, NC 27695

## Abstract

*Stagonospora nodorum* blotch (SNB), caused by *Parastagonospora nodorum*, is a serious disease of wheat worldwide. In the United States, the disease is prevalent in the Mid-Atlantic States, and its management mainly relies on fungicide applications after flag leaf emergence. However, the disease can occur much earlier than flag leaf emergence, and the relationship between time of SNB onset and yield in winter wheat is not understood. The relationship between SNB onset and yield was explored and a binary logistic regression model was developed to predict the time of SNB onset. SNB incidence was recorded in 435 disease cases collected from 11 counties in North Carolina during the 2012 to 2014 growing seasons, where a case was a unique combination of the pre-planting factors of cultivar resistance, latitude, longitude, previous crop, seed treatment, seed rate, tillage type, and wheat residue. SNB onset was defined as day of the year (doy) when SNB incidence reached 50%. Starting at the end of the tillering growth stage, temperature, relative humidity, and precipitation were combined to calculate the daily infection value (DIV), a weather index ranging from 0 (= no *P. nodorum* growth) to 1 (= optimum *P. nodorum* growth). The time of SNB onset explained 32% of the variation in yield, and doy 102 was identified as the threshold for the time of disease onset. If onset occurred after doy 102 ('late' onset), 72% of the cases had above-average yield, while if the onset occurred before doy 102 ('early' onset), only 13% of the cases had above-average yield. To develop a model to predict SNB disease onset, disease onset was categorized into early and late classes and used as the response variable. Binary logistic regression was used to predict onset classes as a function of i)

cumulative DIV until the flag leaf was visible and ii) pre-planting factors. Wheat residue and cumulative DIV until start of stem elongation were found to be significant ( $P < 0.0001$ ) predictors of SNB onset. The binary logistic model had a correct classification rate of 0.94 and specificity and sensitivity rates of 0.91 and 0.99, respectively. Model accuracy based on the area under receiver operating characteristic curve was high with a prediction accuracy of 0.94. The performance of the model in leave-one-out and 10-fold cross-validation techniques was also high. Once validated using independent data, the model could serve as a within-season decision support tool to help growers with decisions on fungicide application for SNB management.

## Introduction

*Stagonospora nodorum* blotch (SNB) of wheat (*Triticum aestivum*), caused by *Parastagonospora nodorum*, is a major disease of wheat globally (Quaedvlieg et al. 2013; Solomon et al. 2006). However, the disease is more prevalent and severe in areas with warm (20–27 °C) and humid weather and frequent rainfall, as is the case in the southeastern United States, parts of Europe and Australia, and southern Brazil (Bergstrom 2010; Eyal 1987; Solomon et al. 2006). The disease affects both the quantity and quality of yield (Eyal 1999) and losses up to 50% have been reported, in addition to lower grain quality (Bhathal et al. 2003; Fried et al. 1987; King et al. 1983; Shipton et al. 1971).

Previous studies have reported an association between the intensity of SNB and pre-planting factors such as wheat residue, resistance level of wheat cultivar, location of the field,

type of tillage, and seed treatment (Eyal 1999; Mehra et al. 2015; Milus and Chalkley 1997; Solomon et al. 2006; Weisz 2013). Wheat residue infected with *P. nodorum* can serve as a source of primary inoculum (Holmes and Colhoun 1975) for infection of wheat, especially where wheat is sown after conservation tillage practices resulting in wheat debris on the soil surface. Wheat residue can harbor both pycnidia (i.e., asexual fruiting bodies) and pseudothecia (i.e., sexual fruiting bodies), and hence gives rise to both rain-splashed pycnidiospores and wind-dispersed ascospores (Bergstrom 2010; Cowger and Silva-Rojas 2006). Secondary spread of the pathogen inoculum occurs through rain-splashed pycnidiospores (Shah et al. 2001). One of the cultural practices to manage SNB is planting wheat in conventionally tilled fields. However, minimum tillage has gained popularity (Horowitz et al. 2010) because of benefits related to reduced soil erosion, runoff and improved soil physical properties (Shaver et al. 2013). In the case of wheat, this practice has led to increased survival and availability of primary inoculum of *P. nodorum* in the field. A crop rotation program with a non-host crop for at least two years can be helpful in reducing the risk to wheat of infection by *P. nodorum*. However, crop rotation does not eliminate the risk of infection by windborne ascospores (Bergstrom 2010) and will be less effective where ascospores are an important source of primary inoculum. Infected seed can also act as a source of primary inoculum (Shah et al. 1995; Shah and Bergstrom 2000; Shah et al. 2001), and hence it is recommended to use certified seed for planting and fungicide seed treatment where farmer-saved seed is to be planted (Cowger and Weisz 2013).

There are no wheat cultivars with complete resistance to *P. nodorum*, but cultivars with moderate resistance are available. Wheat breeding programs have mainly targeted biotrophic pathogens, where R-genes can be introgressed, while the same option is not available against SNB. Thus, in the US cultivars are available with high levels of resistance to leaf rust (caused by *Puccinia triticina*) and powdery mildew (caused by *Blumeria graminis* f. sp. *tritici*), but not to diseases caused by necrotrophs such as SNB (Bergstrom 2010; Cowger and Weisz 2013). The modest progress on SNB resistance in many breeding programs is due to several factors including the sporadic nature of disease epidemics in screening nurseries and the complexity of assembling multiple QTLs in a single genotype. Currently, while breeding efforts for SNB resistance are on-going, management of the disease relies substantially on prophylactic fungicide sprays (Crook et al. 2012; Stover et al. 1996).

Historically, growers have sought to minimize unprofitable fungicide treatments because wheat prices have been low (Anonymous 2015b). Application of strobilurin or triazole fungicides on the basis of wheat growth stage and in the absence of fungal diseases was found to be unprofitable (Weisz et al. 2011). Confining fungicide application to crops with observable fungal disease improves fungicide efficiency and reduces the rate at which pathogens develop fungicide resistance. Resistance to strobilurins in *P. nodorum* has been reported in Europe (Morzfeld et al. 2004). Strobilurin resistance in a related pathogen, *Septoria tritici*, has caused serious disease-related losses in Europe (Torriani et al., 2009). Slowing the development of resistance to available effective fungicides dictates that the



products be applied only when necessary. One approach to increasing fungicide efficiency is the early application of fungicides based on the onset of disease. An early application may require a lower fungicide dose to keep the pathogen population low, which in turn reduces the selection pressure on the pathogen (van den Bosch et al. 2011). For SNB, this requires determining how the pre-planting factors described above interact with in-season weather factors to influence the time of SNB onset. Most of these pre-planting factors are known to have an effect on the end-of-season disease severity, but their effect on time of epidemic onset has not been determined. Such information would be most useful if SNB onset can be predicted early enough for fungicide applications to significantly reduce the rate of disease development.

Previous studies relating weather variables to the development of SNB have focused on weather during stem elongation (Cook 1977; Djurle et al. 1996; Hansen et al. 1994; Shaw et al. 2008; Tyldesley and Thompson 1980). In general, these authors found rainfall to be a significant predictor of SNB. However, none of the studies examined how weather in combination with pre-planting factors influenced the development of SNB. In addition, temperature is an important factor in the *P. nodorum* infection process (Bergstrom 2010) and affects the latent period of *P. nodorum* (Zearfoss et al. 2011), which in turn influences disease onset.

Given its polycyclic nature, the time of SNB onset will affect the extent and magnitude of epidemic development during the season. In our previous work, we observed that earlier onset of SNB led to higher disease severity at the end of the season and

significant effects of disease on yield occurred when SNB onset occurred early in the season (Mehra et al. 2015). An accurate prediction of epidemic onset will also be useful in reducing unnecessary scouting for the disease needed to determine if fungicide sprays are required.

Given the role of pre-planting factors in disease management, we hypothesized that onset of an SNB epidemic depends on: 1) type of tillage (no-till vs. conventional till), 2) level of cultivar resistance to *P. nodorum*, 3) seeding rate and stand density (with sparser canopy allowing more rain-splash), 4) fungicide seed treatment (with stands from untreated seed developing more disease and earlier in the season), 5) weather variables (temperature, rainfall, and relative humidity [RH]), and 6) location (longitude and latitude). Type of tillage is important because minimum tillage maintains wheat residue on the ground from previous seasons that can increase inoculum availability for infection. Weather variables vary with field location and influence the time and extent of epidemic development, as *P. nodorum* requires frequent rainfall for dispersal and specific conditions for ascospore discharge. Lower seeding rates can lead to sparser crop stands that encourage more efficient splash dispersal of the pathogen, which can also influence epidemic onset. Based on the above considerations, the objective of this study was to explore the relationship between time of SNB onset and yield in winter wheat, and develop a model to predict the onset of SNB based on pre-planting factors and in-season weather.

## Materials and Methods

**Field sites and pre-planting factors.** Field experiments were conducted in 11 North Carolina counties (Fig. 4.1) representing three geographical regions of the state (Weisz 2013), during 2011-12, 2012-13, and 2013-14 (hereafter 2012, 2013, and 2014, respectively). The types of tillage evaluated for a given site-year combination are summarized in Table 4.1. For each site-year, wheat was planted in either a conventionally prepared field (Conv.) or a no-tilled field, or both. Experimental sites were chosen to represent different wheat-growing areas in the state and a range of weather conditions and epidemic history. All production practices at each site followed recommendations from the North Carolina State University small grains production guide (Weisz 2013). In line with those recommendations, planting dates for all sites ranged from the last week in September (in the western mountains) to the first week in November (in the eastern Tidewater zone).

In 2012, five soft red winter wheat cultivars (Branson, Dyna-Gro Dominion, Dyna-Gro Shirley, SS8700, and SY9978) were planted at each site in a randomized complete block design with six replicates (blocks). The cultivars had varying levels of resistance to *P. nodorum* ranging from 3 to 6 on a scale of 1 (most resistant) to 9 (most susceptible). Cultivar resistance ratings were determined from their performance in the Eastern United States Septoria Nursery maintained by the United States Department of Agriculture- Agricultural Research Service (Anonymous 2015c) at Raleigh and Kinston, NC. These cultivars are grown commercially in the state and were chosen for their similar heading dates and resistance to all or at least some of the other prevalent wheat pathogens.

In 2013 and 2014, two more pre-planting factors, seed treatment and seeding rate, were added at all sites. The experimental design at each site was a split-split-plot with six replicates (blocks), with seed treatment as the main plot, seeding rate as the sub-plot, and wheat cultivar as the sub-sub-plot. There were two levels of the seed treatment factor, carboxin+thiram- (Vitavax, Chemtura, Australia) treated and untreated seed. Carboxin+thiram is moderately effective in controlling *P. nodorum* in wheat seed. There were also two levels of seeding rate, standard seeding rate and a reduced rate (i.e., 80% of standard rate). Certified seed was purchased from commercial seed distributors in North Carolina. All planted seed was treated with imidacloprid insecticide to minimize the aphid-transmitted barley yellow dwarf virus. Only four cultivars (Dyna-Gro 9012, Dyna-Gro Shirley, P26R20, and SS8641 [replaced by USG3438 in 2014]) were used in 2013 and 2014 with an SNB resistance level ranging from 4 to 7.

At each site, pre-planting factors such as location (longitude and latitude), previous crop, and amount of wheat residue on the ground were recorded. Latitude and longitude data for each site were determined by locating the position of the field on Google Maps (Google Inc., CA, <https://maps.google.com/>). In each field, the previous crop was recorded on a binary scale with 0 = no wheat crop in the previous year, and 1 = a wheat crop in the previous year. The amount of wheat residue on the ground (%) was determined using the line-transect method (Wollenhaupt and Pingry 1991). The amount of residue in the plots was categorized as low- or high-residue based on a 20% residue cover threshold (Mehra et al. 2015).

**Weather variables.** At field sites on North Carolina Department of Agriculture research stations (Table 4.1), temperature, RH, and precipitation were obtained from existing weather stations. At these weather stations, sensors for temperature, RH, and precipitation were located 2 m above the ground. The sensors recorded data at 1-min intervals; temperature and RH were averaged and precipitation was summed over 30-min intervals. Weather stations were located within 3 km of experiment fields, except for at Caswell Research Farm where the weather station was 8 km from the experimental field.

Weather data for experiments in growers' fields were obtained from portable data loggers (Watchdog model 450, Spectrum Technologies, Plainfield, IL) installed in the fields. Precipitation was measured using a tipping-bucket rain collector placed 2 m above the ground and connected to the data logger placed at the same height. The portable data loggers were protected by radiation shields (Watchdog, Spectrum Technologies, Plainfield, IL) from rain and solar radiation. The portable loggers were programmed to record data at 30-min intervals. At Tyner in 2014, the tipping-bucket was clogged and precipitation data were taken from the nearest State Climate Office of North Carolina weather station, which was 30 km from the experimental site.

Daily infection values (DIV) were calculated from temperature and moisture indices by adapting a method used by Wolf and Verreet (2005) to the wheat-*P. nodorum* pathosystem. Daily infection values were calculated by using weather data collected every 30 min; therefore, a DIV for a specific day at a specific location was an average of 48 values

(one value for each 30-min interval) resulting from the product of temperature and moisture indices as shown below.

$$\text{DIV} = 1 / 48 \sum_{i=1}^{48} (T_i \times M_i) \quad (1)$$

where  $T_i$  and  $M_i$  are temperature and moisture indices for the  $i$ th half-hour or “30-minute interval” of a given day. Temperature index was calculated as a ratio of  $\text{LP}_{\text{opt}}$  and  $\text{LP}_{x_i}$ :

$$T_i = \text{LP}_{\text{opt}} / \text{LP}_{x_i} \quad (2)$$

where  $\text{LP}_{\text{opt}}$  is the latent period of *P. nodorum* at optimum temperature (22.5 °C) (Eyal et al. 1987) and  $\text{LP}_{x_i}$  is the latent period at temperature  $x_i$  °C for the  $i$ th half-hour. The LP was calculated based on the equation developed by Zearfoss et al. (2011):

$$\frac{1}{\text{LP}} = 0.0072 + 0.0026 \times x_i \quad (3)$$

The moisture index was defined on a binary scale: 1 for >75% relative humidity (Bergstrom 2010) or precipitation >0.1 cm (Hansen et al. 1994; Tyldesley and Thompson 1980), and 0 otherwise. The DIV summarizes the influence of temperature and moisture on growth of *P. nodorum*, with 0 indicating no growth and 1 indicating conditions for optimal growth of the pathogen. DIV were accumulated starting from day of the year (DOY) 66, which generally corresponded to the end of tillering with no disease in the plots. Calculation of DIV was stopped at the time of disease onset. The cumulative DIV (cDIV) was calculated for the  $m$ th case by adding all DIV until the day of disease onset:

$$cDIV_m = \sum_{k=66}^{D_m} DIV \quad (4)$$

where  $D_m$  is DOY when onset occurred in the  $m$ th case.

**Disease assessment and definition of disease onset.** Disease incidence and disease severity were both recorded in each plot except in 2012, when only disease severity data was recorded. To assess disease incidence, 10 pairs of wheat tillers (each pair containing two neighboring tillers) were arbitrarily selected in each plot and tagged with a marking flag and cable ties. Tagging of tillers was done in the first two weeks of March, prior to observing SNB symptoms. Starting in late March, plots were visited every two weeks to monitor disease development. Disease assessments were done from mid-spring (starting 25 March) to late spring (ended 30 May). The exception was at Laurel Springs in 2013, where the final disease assessment was made on 11 June 2013; Laurel Springs is cooler than the other sites and has a longer wheat season. A tagged pair of tillers was considered diseased if it had at least one SNB lesion. The disease was identified visually; for confirmation, leaf samples from untagged plants showing similar lesions were later surface-sterilized, placed on water agar to induce pycnidial formation to confirm the presence of *P. nodorum*. Disease incidence was calculated by dividing number of diseased tiller pairs by total number of tagged tiller pairs in the plot and then multiplying the result with 100. The mean disease incidence from six replicates were calculated for each unique combination of pre-planting factors (i.e., case). Disease onset was defined when the average disease incidence for a particular case was approximately 50%.

Additionally, disease severity was assessed visually in each plot on a whole-canopy basis, and was expressed as the percentage of SNB-symptomatic leaf area. Briefly, disease severity in each plot was assessed in five sections of  $\sim 0.1 \text{ m}^2$  each, randomly selected and widely spread within the plot. The percent leaf area with SNB symptoms in each section was estimated visually, and the mean symptomatic leaf area across the five sections represented the overall disease severity for the plot. The lowest leaves that senesced due to maturity or abiotic factors were not included in the visual assessment of disease severity. Two to four assessments were made for most location-year combinations (27 out of 32); only one assessment was made for the remaining five location-years, either due to a shorter growing season or low disease levels.

From the data collected in 2013 and 2014, a relationship between disease incidence and disease severity was determined. Disease incidence of 50% corresponded with 4% disease severity on a whole-canopy basis (data not shown). Hence, in 2012, when disease incidence data were not collected, disease onset was assumed to have occurred when the average disease severity for a case was 4%. In support of that assumption, a 5% disease severity threshold was previously used to indicate disease onset for a related disease on wheat, *Septoria tritici* blotch caused by *Septoria tritici* (Pietravalle et al. 2003). In 98% of cases, disease incidence on a given assessment date was not exactly 50%, while two weeks later it exceeded 50%. In those cases, the date of disease onset was estimated based on the rate of increase of disease incidence from one assessment to the next.



**Disease onset cutoff.** Grain was harvested from each plot using a research combine. The grain was weighed, and yield was calculated after adjusting weight to a 13.5% moisture level. In a preliminary analysis, time of disease onset was plotted against yield to determine the threshold for disease onset. Date of disease onset ranged from DOY 73 to 129. There were 50 cases where defined disease onset never occurred, and these cases were excluded from further analysis (Fig. 4.2A).

A significant linear correlation ( $r = 0.55$ ;  $P < 0.0001$ ) was found between time of disease onset and yield. A linear regression equation of the form:

$$y = -5.93 + 0.098 \times t \quad (5)$$

best described the relationship between yield and time of disease onset ( $R^2 = 0.26$ ,  $P < 0.0001$ ), where  $y$  = yield ( $\text{t ha}^{-1}$ ) and  $t$  = time (DOY) of SNB onset. In North Carolina, the 10-year mean wheat yield is  $4.03 \text{ t ha}^{-1}$  (Anonymous 2015a) which corresponded to DOY 102 (about mid-April). If disease onset occurred before DOY 102, only 13% of the disease cases had higher yields than the North Carolina 10-year average, while if disease onset occurred after DOY 102, 72% of the disease cases had above-average yield (Fig. 4.2B).

Based on this relationship between mean yield and DOY 102, a binary response variable for the time of disease onset (TDO) was defined for each case. If disease onset occurred before DOY 102, TDO was classified as 1 while if disease onset occurred on or after DOY 102, TDO was classified as 0. This threshold for time of disease onset has practical relevance because mid-April is the time when fungicide applications are often made to the winter wheat crop in North Carolina (Weisz 2013). A model that predicts whether

disease onset will occur before mid-April based on pre-planting factors and cumulative weather variables could be useful in guiding fungicide application decisions.

**Variable reduction.** In addition to the eight pre-planting predictors, three potential in-season weather-based predictors (cDIV.1 through cDIV.3) were defined as described in Table 4.2. Briefly, cDIV.x corresponds to accumulated daily infection values until “x” week prior to DOY 102 (the TDO-cutoff). PROC VARCLUS with option MAXEIGEN = 0.9 (SAS 9.3; SAS Institute, Cary, NC) was used to reduce the number of variables by eliminating redundant predictors (Nelson 2001). PROC VARCLUS assigns variables to different clusters on the basis of high intra-cluster correlation and low inter-cluster correlation. The algorithm starts with one cluster, and the cluster is split into two smaller clusters if the second eigenvalue for the cluster is greater than the assigned threshold of 0.9. From each cluster, only one variable is retained for inclusion in subsequent analysis based on its correlation coefficient with the response variable TDO. Spearman correlation coefficients between potential predictors and TDO were calculated using PROC CORR in SAS. In addition, Hoeffding's dependence coefficients between selected predictors and TDO were also calculated to determine if any predictor had a nonlinear relationship with the response. The predictor is suspected to be non-linearly correlated with response if its Spearman correlation coefficient is relatively low and the Hoeffding dependence coefficient is relatively high.

**Model fitting.** A binary logistic model was developed to predict TDO based on selected in-season weather and pre-planting factors using PROC LOGISTIC in SAS. A logistic regression model is usually stated in terms of the probability of the occurrence of an event of interest (e.g., TDO) under certain conditions (e.g., predictor variables) (Harrell, 2001). Here, the binary logistic regression defined in terms of the probability model of  $Y$  (TDO) = 1 given  $m$  predictor variables ( $X$ ) can be stated as

$$P(Y = 1 | X) = \frac{1}{1 + e^{-X\beta}} \quad (6)$$

where  $X\beta = \beta_0 + \beta_1X_1 + \beta_2X_2 + \dots + \beta_mX_m$ . The “all subset selection” (SELECTION = score) method was used to select variables for the model. In the MODEL statement options, BEST = 1 was specified to output only the best one-predictor model, best two-predictor model, and so on.

The model with the lowest Bayesian Information Criteria (BIC) was chosen as the final model from all subsets. The probability cutoff value to classify the predictions into early- or late-onset of SNB was decided by maximizing the sensitivity (proportion of correctly classified events) and specificity (proportion of correctly classified nonevents) based on the receiver operating characteristics (ROC) curve. If the predicted probability was greater than or equal to the probability cutoff, cases were classified as events (onset before mid-April); if it was less, cases were classified as nonevents (onset after mid-April). Model performance was evaluated based on proportion of correctly classified cases, specificity, sensitivity, and area under the ROC curve (AUC). In addition, standard goodness-of-fit

criteria such as concordant and discordant pairs, Somers' *D*, *Gamma*, and Kendall's *Tau-a* were also reported.

**Model validation.** Predictions of the binary logistic model were evaluated using two different cross-validation techniques. First, leave-one-out cross-validation was used by generating a prediction for a case using a model developed without using that case. The cross-validated predictions were obtained by setting `PREDPROBS = crossvalidate` in the `OUTPUT` statement of `PROC LOGISTIC`. Second, a 10-fold cross-validation was conducted by splitting the total number of cases into 10 groups. Each group had same proportion of events as in the full dataset. A model was developed using nine groups and predictions were made on the excluded group and the process was repeated for each group. Model development followed the same steps (including variable reduction using `PROC VARCLUS`) as in the original binary logistic regression model described above. Performance of the cross-validated models was assessed using correct classification, specificity, sensitivity, and AUC. Receiver operating characteristics curves for the binary logistic model and two cross-validation models were generated and used to compare model performance.

## Results

**Disease cases and selected predictor variables.** A total of 435 disease cases were obtained from all field experiments conducted during the three growing seasons (Fig. 4.1). Out of these 435 cases, the defined disease onset occurred on 390 cases. Disease onset before

mid-April accounted for 41% of the 390 cases, while disease onset after mid-April accounted for the remaining 59% (Fig. 4.2). The proportions of cases where SNB onset occurred before mid-April were 0.23, 0.50, and 0.56 in the Piedmont, Coastal Plains, and Tidewater regions of North Carolina, respectively (Fig. 4.3). As expected, of the three weather indices tested (cDIV.1 through cDIV.3), cDIV.1 and cDIV.2 showed the maximum separation between early- and late-onset cases (Fig. 4.4). Generally, the spread of values as shown by the length of vertical lines above and below the box, was greater for early-onset cases than late-onset ones (Fig. 4.4).

Based on the variable reduction technique, the eleven predictor variables examined were clustered into 5 clusters. Seeding rate, seed treatment, and cultivar resistance were all grouped in separate clusters. Latitude, longitude and cDIV.1 through cDIV.3 were grouped in cluster 1, while wheat residue, previous crop, and tillage type were grouped in cluster 2 (Table 4.3). The predictor variables cDIV.2 and wheat residue were selected from cluster 1 and cluster 2, respectively, based on their Spearman correlation coefficient with the response variable (Table 4.4). Based on these results, cultivar resistance, cDIV.2, seeding rate, seed treatment, and wheat residue were selected for use in subsequent analysis. The correlation ranks of these selected predictor variables were similar for the Spearman and Hoeffding's correlation tests (Table 4.4).

**Binary logistic regression model.** The final binary logistic regression model selected using the Bayesian information criterion had two significant predictors, cDIV.2 and wheat

residue (Table 4.5). Cultivar resistance, seed treatment, and seeding rate were not identified as significant predictors of disease onset. As a pre-planting factor, wheat residue can be estimated before planting, while cDIV.2 is SNB-favorable weather up until two weeks before mid-April. Based on the model results, the equation for the final binary logistic regression model can be expressed as:

$$\ln(p / 1 - p) = -5.45 + 2.15 \cdot WR + 1.38 \cdot cDIV.2 \quad (7)$$

where  $p$  is the probability that disease onset will occur before mid-April and WR is the amount of wheat residue at planting. A probability cutoff of 0.34 maximized both sensitivity and specificity, and the model had a sensitivity of 0.99 and specificity of 0.91 (Table 4.6). The percentage of concordant pairs was 94.3%, while that of discordant and tied pairs was 5.4%, and 0.3%, respectively. Other goodness-of-fit criteria for the model such as Somers'  $D$ ,  $\Gamma$ , and Kendall's  $\tau_a$  that are calculated from values of concordant and discordant pairs were 0.89, 0.89, and 0.43 respectively, indicating an excellent fit of the model to the data.

The binary logistic regression model yielded similar predicted proportions of early-onset cases as compared to observed proportions in all three geographical regions of North Carolina (Fig. 4.3). However, the model over-estimated the number of cases where SNB onset would occur before mid-April in the Tidewater region. The relationship of probability of SNB onset before mid-April with cDIV.2 was sigmoidal for both low and high amounts of wheat residue (Fig. 4.5). For a given value of cDIV.2, the probability of SNB onset early in the season was higher in fields with  $\geq 20\%$  wheat residue cover than in fields with  $< 20\%$

residue cover. The correct classification rate was high with a value of 0.94 (Table 4.6). The accuracy of the model as assessed by the ROC analysis was high with an AUC = 0.94 (Table 4.6).

**Model validation.** The results of the 10-fold cross-validation and leave-one-out cross-validation techniques were very similar (Fig. 4.6). The area under the receiver operating characteristics curve (AUC) was exactly same for both validation models and the final model (Table 4.6). However, the correct classification, sensitivity, and specificity rates for the 10-fold cross-validated model were slightly lower than leave-one-out model and the final model (Table 4.6).

## Discussion

The current study illuminated two important aspects of *P. nodorum*-wheat pathosystem. First, wheat yield was found to be associated with time of SNB onset, and the time of disease onset explained 32% of variation in yield. A disease-onset “watershed” period of mid-April was identified, such that SNB onset before that date was associated with significantly higher probability of reduced yield.

Second, the binary logistic regression model identified amount of wheat residue, and favorable weather for infection of wheat by *P. nodorum* until two weeks prior to mid-April (represented by cDIV.2) as important predictors for disease onset. Favorable weather for the infection of wheat by *P. nodorum* was summarized using cumulative daily infection values

that represent the effect of temperature and moisture on infection. The use of DIV has been found to be useful in assessing the most likely periods of absence of *Cercospora* leaf spot in sugar beet fields (Wolf and Verreet, 2005). The model can help growers determine if early SNB onset will occur in their field, with the implication that a fungicide application aimed at preventing SNB-associated yield losses may be profitable. To the best of our knowledge, this is the first study that utilizes a combination of in-season weather and pre-planting cultural factors to predict the onset of SNB in winter wheat.

For a given accumulation of favorable weather, probability of early SNB onset was predicted to be higher in fields with  $\geq 20\%$  wheat residue cover than fields with less residue. This is in agreement with our previous study, where we reported an early SNB onset in residue treated plots (Mehra et al. 2015). The success of the predictive model will depend on the resolution of weather data, which are usually spatially heterogeneous. Especially precipitation may not be identical throughout the area around a weather station or a field. As indicated by Djurle and Yuen (1991), it is also possible that even if disease onset is correctly predicted and does occur, the epidemic may not become severe if weather in the later part of the season is less favorable for disease development.

The quest to develop prediction models for SNB started in the 1980s, when Tyldesley and Thompson (1980) developed a model to predict SNB severity based on frequency of rainfall in England and Wales and the model had an accuracy of 71%. They observed that more than 5 days with rainfall greater than 1 mm in the second half of May and first half of June were associated with severe SNB epidemics. Similar qualitative weather-based



thresholds for epidemic development were provided by Hansen et al. (1994) in Denmark, jointly for both SNB and *Septoria tritici* blotch, where eight days with rainfall  $\geq 1$  mm in a 30-day period starting at stem elongation correlated with disease severity and yield response. However, they did not incorporate effect of pre-planting factors such as wheat residue and other weather variables such as temperature that are well known to play a role in the infection of wheat by *P. nodorum* (Jeger et al. 1981). In the work of Hansen et al. (1994), the time of prediction was 30 days after stem elongation which is late in the season. The epidemic could already be severe by that time and may require more than one fungicide spray to contain it. The binary logistic model developed in our study can be used to predict disease onset at the start of stem elongation period. This early prediction of disease onset should result in more efficient fungicide use since fewer applications would be required to prevent high levels of disease.

Caristi et al. (1987) developed an expert system known as EPINFORM to provide estimates of damage to wheat by SNB and stripe rust in Montana. EPINFORM counts the number of infection cycles necessary to cause a yield penalty, and assumes that inoculum is present in the field all the time and weather is the only deciding factor in initiating the infection cycle. A quantitative validation of the EPINFORM system was not conducted due to the unavailability of data on yield reduction. In another study, Djurle and Yuen (1991) created a simulation model for SNB and found that weather conditions were the most important factors in the development of disease epidemics, while amount of initial inoculum was less important. Our results on the importance of weather for SNB onset largely agree

with those of Djurle and Yuen. However, unlike them, we found initial inoculum, represented by amount of wheat residue on the ground, to be an important predictor of SNB onset. Djurle and Yuen inferred the quantity of initial inoculum from latent invisible and visible lesions on leaves 1-4, which is considerably different from the infected wheat debris in the present study. In addition, their study was designed to simulate development of an SNB epidemic rather than onset of SNB.

In the present study, several agronomically important pre-planting factors were found to have no significant influence on SNB onset. One such factor was cultivar resistance. In wheat, resistance to *P. nodorum* is quantitative and partial (Solomon et al. 2006), and it is known to slow the rate of epidemic spread rather than delay the time of disease onset (Vanderplank 1984). This may explain why cultivar resistance did not influence the time of SNB onset even though it was an important predictor of end-of-season SNB severity (Cowger and Murphy 2007; Weisz and Cowger 2014).

A second factor that did not influence SNB onset was seeding rate. The 20% reduction in standard seeding rate may not have created a sufficiently sparser canopy that better splash dispersal of pycnidiospores from wheat debris could initiate earlier onset of SNB. However, a sparser canopy can result in a microclimate that is less conducive for infection by *P. nodorum* (Bergstrom 2010), so it is possible that the two effects occurred (enhanced splash dispersal and less favorable microclimate) and offset each other. A further reduction in the standard seeding rate might influence SNB onset but regardless of the

direction of that influence, lower seeding rates would not be relevant to commercial wheat production.

A third factor that was non-significant in SNB onset was seed treatment. This can be explained by the use of certified seed in all the experiments and the absence of *P. nodorum* in assays of seed both treated and untreated with fungicide. Seedborne inoculum has been reported to be important in initiating SNB epidemics (Shah et al., 1995) and it reasonable to assume this could have been observed in our study if seed lots were infected. Understandably, the actual effects of infected seed on SNB onset will depend on the levels of seedborne inoculum.

Longitude and latitude were correlated with weather variables and were not included in model development. This could be mainly due to a decrease in temperature as one moves from the southeast to northwest region of the state. As our final model is independent of location, it can be easily tested in other wheat growing areas and adapted for implementation based on the prevailing weather factors and levels of wheat residue.

Once externally validated with independent data, the model developed in this study has good potential to serve as a within-season decision support tool for growers making decisions on fungicide application to manage SNB. Growers can monitor their wheat crop when the model predicts a high probability of early onset to determine if a fungicide spray is needed. Each year, during the first week of April, which on average corresponds to start of stem elongation growth stage (Zadoks 31) in North Carolina, risk maps for the state can be prepared by downloading weather data from weather stations maintained by the North

Carolina State Climate Office or the appropriate state climate office depending on the location of the growers' fields. Alternatively, a web-based application can be developed where users provide their physical address and an estimate of wheat residue on the ground ( $<20\%$  vs.  $\geq 20\%$  residue cover) and the application will forecast an early or late SNB onset. Weather data can be gathered from weather stations nearest to the physical address of the farm to calculate and accumulate daily infection values to input into the model. While this model provides a simple tool to determine whether the onset of SNB will be early or late in a given field, growers will need to monitor their fields to determine when to apply the fungicide spray based on established thresholds to reduce the subsequent rate of development of SNB epidemics. Even if the model predicts a late onset of SNB, it is advisable to monitor the crop for sudden outbreaks of disease due to favorable environmental conditions such as rainstorms that enhance the dispersal of the pathogen.

### **Acknowledgements**

This study was supported by Hatch Funds from North Carolina Agricultural Experiment Station for Project Number NC02251, and by the USDA-ARS. The authors thank staff at research stations and county extension agents, for managing the research plots; and Ryan Parks, Bradley Van Staaldin, Tyler Cole, Brandon Clontz, Ian Swaim, Charlie Glover, Barbara Stevens, Joseph Jeffreys, Sushila Chaudhari, and Jennifer Patton-Özkurt for their assistance in planting, tagging wheat tillers, and harvesting.

## Literature Cited

- Anonymous. 2015a. USDA- National Agricultural Statistics Service. <http://quickstats.nass.usda.gov/>. Accessed on 10 October 2015.
- Anonymous. 2015b. USDA, Economic Research Service, Wheat Data. Online: <http://www.ers.usda.gov/data-products/wheat-data.aspx>. Accessed 15 September 2015.
- Anonymous. 2015c. USDA, Agricultural Research Service -Eastern United States Septoria Nursery Reports 2014-15. <http://www.ars.usda.gov/Main/docs.htm?docid=8419&page=8>. Accessed on 27 July 2015.
- Bergstrom, G. C. 2010. *Stagonospora nodorum* blotch and *Stagonospora avenae* blotch. Pages 75–77 in: Compendium of wheat diseases and pests. W. W. Bockus, R. L. Bowden, R. M. Hunger, W. L. Morrill, T. D. Murray, and R. W. Smiley, eds. The American Phytopathological Society, St. Paul, Minnesota.
- Bhathal, J. S., Loughman, R., and Speijers, J. 2003. Yield reduction in wheat in relation to leaf disease from yellow (tan) spot and septoria nodorum blotch. *Eur. J. Plant Pathol.* 109:435–443.
- Caristi, J., Scharen, A.L., Sharp, E.L., and Sands, D.C. 1987. Development and preliminary testing of EPINFORM, an expert system for predicting wheat disease epidemics. *Plant Dis.* 71:1147–1150.
- Cowger, C., and Murphy, J. P. 2007. Artificial inoculation of wheat for selecting resistance to *Stagonospora nodorum* blotch. *Plant Dis.* 91:539–545.

- Cowger, C., and Silva-Rojas, H. V. 2006. Frequency of *Phaeosphaeria nodorum*, the sexual stage of *Stagonospora nodorum*, on winter wheat in North Carolina. *Phytopathology* 96:860–866.
- Cowger, C., and Weisz, R. 2013. Small grain disease management. Pages 58-70 in: Small grain production guide 2011-12. R. Weisz, ed. North Carolina Cooperative Extension Service, Publication No. AG-580 (Revised).
- Cook, R. J. 1977. Effect of timed fungicide sprays on yield of winter wheat in relation to *Septoria* infection periods. *Plant Pathol.* 26:30–34.
- Crook, A. D., Friesen, T. L., Liu, Z. H., Ojiambo, P. S., and Cowger, C. 2012. Novel necrotrophic effectors from *Stagonospora nodorum* and corresponding host genes in winter wheat germplasm in the southeastern United States. *Phytopathology* 102: 498–505.
- Djurle, A., and Yuen, J. 1991. A simulation model for *Septoria nodorum* in winter wheat. *Agric. Syst.* 37:193–218.
- Djurle, A., Ekbom, B., and Yuen, J. E. 1996. The relationship of leaf wetness duration and disease progress of glume blotch, caused by *Stagonospora nodorum*, in winter wheat to standard weather data. *Eur. J. Plant Pathol.* 102:9–20.
- Eyal, Z. 1999. The septoria tritici and stagonospora nodorum blotch diseases of wheat. *Eur. J. Plant Pathol.* 105:629–641.
- Eyal, Z., Scharen, A. L., Prescott, J. M., and van Ginkel, M. 1987. The Septoria Diseases of Wheat: Concepts and Methods of Disease Management. Mexico, D. F.: CIMMYT.

- Fried, P. M., and Meister, E. 1987. Inheritance of leaf and head resistance of winter wheat to *Septoria nodorum* in a diallel cross. *Phytopathology* 77:1371–1375.
- Hansen, J. G., Secher, B. J. M., Jørgensen, L. N. and Welling, B. 1994. Thresholds for control of *Septoria* spp. in winter wheat based on precipitation and growth stage. *Plant Pathol.* 43:183–189.
- Holmes, S. J. I., and Colhoun, J. 1975. Straw-borne inoculum of *Septoria nodorum* and *S. tritici* in relation to incidence of disease on wheat plants. *Plant Pathol.* 24:63–66.
- Horowitz, J., Ebel, R., and Uda, K. 2010. ‘No-Till’ Farming is a Growing Practice. United States Department of Agriculture-Economic Research Service. Economic Information Bulletin Number 70. Washington, DC.
- Jeger, M. J., Griffiths, E. and Jones, D. G. 1981. Influence of environmental conditions on spore dispersal and infection by *Septoria nodorum*. *Ann. Appl. Biol.* 99:29–34.
- King, J. E., Cook, R. J., and Melville, S. C. 1983. A review of *Septoria* diseases of wheat and barley. *Ann. Appl. Biol.* 103:345–373.
- Mehra, L. K., Cowger, C., Weisz, R., and Ojiambo, P. S. 2015. Quantifying the effects of wheat residue on severity of *Stagonospora nodorum* Blotch and yield in winter wheat. *Phytopathology* 105:1417–1426.
- Milus, E. A., and Chalkley, D. B. 1997. Effect of previous crop, seedborne inoculum, and fungicides on development of *Stagonospora* blotch. *Plant Dis.* 81:1279–1283.

- Morzfeld, J., Koopman, B., and Hoppe, H. -H. 2004. Development of fungicide resistance of wheat and barley pathogens against strobilurins: A methodological approach. *J. Plant Dis. Protect.* 111:105–114.
- Nelson, B. D. 2001. Variable reduction for modeling using PROC VARCLUS. In: *Proceedings of the Twenty-sixth Annual SAS® Users Group International conference*, Cary, NC: SAS Institute Inc., 2001. Paper No. 261-26. <http://www2.sas.com/proceedings/sugi26/p261-26.pdf> . Accessed on 10 September 2015.
- Quaedvlieg, W., Verkley, G. J. M., Shin, H. -D., Barreto, R. W., Alfenas, A. C., Swart, W. J., Groenewald, J. Z. and Crous, P. W. 2013. Sizing up *Septoria*. *Stud. Mycol.* 75:307–390.
- Pietravalle, S., Shaw, M. W., Parker, S. R., and Van Den Bosch, F. 2003. Modeling of relationships between weather and *Septoria tritici* epidemics on winter wheat diseases: a critical approach. *Phytopathology* 93:1329–1339.
- Shah, D. A., and Bergstrom, G. C. 2000. Temperature dependent seed transmission of *Stagonospora nodorum* in wheat. *Eur. J. Plant Pathol.* 106:837–842.
- Shah, D. A., Bergstrom, G. C., and Ueng, P. P. 2001. Foci of *Stagonospora nodorum* blotch in winter wheat before canopy development. *Phytopathology* 91:642–647.
- Shah, D., Bergstrom, G. C., and Ueng, P. P. 1995. Initiation of *Septoria nodorum* blotch epidemics in winter wheat by seedborne *Stagonospora nodorum*. *Phytopathology* 85:452–457.



- Shaw, M.W., Bearchell, S.J., Fitt, B.D.L., Fraaije, B. A. 2008. Long-term relationships between environment and abundance in wheat of *Phaeosphaeria nodorum* and *Mycosphaerella graminicola*. *New Phytol.* 177:229–238.
- Shaver, T. M., Peterson, G. A., Ahuja, L. R., and Westfall, D. J. 2013. Soil sorptivity enhancement with crop residue accumulation in semi-arid dryland no-till agroecosystems. *Geoderma* 192:254–258.
- Shipton, W. A., Boyd, W. R. J., Rosielle, A. A., and Shearer, B. I. 1971. The common septoria diseases of wheat. *Bot. Rev.* 37:231–262.
- Solomon, P. S., Lowe, R. G. T., Tan, K-C., Waters, O. D. C., and Oliver, R. P. 2006. *Stagonospora nodorum*: cause of stagonospora nodorum blotch of wheat. *Mol. Plant Pathol.* 7:147–156.
- Stover, R. W., Francl, L. J., and Jordahl, J. G. 1996. Tillage and fungicide management of foliar diseases in a spring wheat monoculture. *J. Prod. Agric.* 9: 261–265.
- Tyldesley, J. B. and Thompson, N. 1980. Forecasting *Septoria nodorum* on winter wheat in England and Wales. *Plant Pathol.* 19:9–20.
- Torriani, S. F., Brunner, P. C., McDonald, B. A., and Sierotzki, H. 2009. QoI resistance emerged independently at least 4 times in European populations of *Mycosphaerella graminicola*. *Pest. Manag. Sci.* 65:155–162.
- Vanderplank, J. E. 1984. *Disease Resistance in Plants*, 2<sup>nd</sup> ed. Academic Press, New York.

- van den Bosch, F., Paveley, N., Shaw, M., Hobbelen, P., and Oliver, R. 2011. The dose rate debate: does the risk of fungicide resistance increase or decrease with dose?. *Plant Pathol.* 60:597–606.
- Weisz, R. 2013. Small Grain Production Guide. North Carolina Cooperative Extension Service, Publication No. AG-580 (Revised March 2013).
- Weisz, R., and Cowger, C. 2014. 2014 Wheat variety performance & recommendations. Smartgrains- The Small Grains Fact Sheet. No. 37. Online: [http://www.smallgrains.ncsu.edu/\\_Misc/\\_VarietySelection.pdf](http://www.smallgrains.ncsu.edu/_Misc/_VarietySelection.pdf) Accessed 17 September 2015.
- Weisz, R., Cowger, C., Ambrose, G., and Gardner, A. 2011. Multiple mid-Atlantic field experiments show no economic benefit to fungicide application when fungal disease is absent in winter wheat. *Phytopathology* 101:323–333.
- Wolf, P. F. J., and Verreet, J. A. 2005. Factors affecting the onset of *Cercospora* leaf spot epidemics in sugar beet and establishment of disease-monitoring thresholds. *Phytopathology* 95:269–274.
- Wollenhaupt, N. C., and Pingry, J. 1991. Estimating Residue using the Line-transect Method. Publication No. A3553. Cooperative Extension, University of Wisconsin, Madison, WI.
- Zadoks, J. C., Chang, T. T., and Konzak, C. F. 1974. A decimal code for the growth stages of cereals. *Weed Res.* 14:415–421.
- Zearfoss, A. D., Cowger, C., and Ojiambo, P. S. 2011. A degree-day model for latent period of *Stagonospora nodorum* blotch in winter wheat. *Plant Dis.* 95:561–567.

**Table 4.1.** Experimental sites and tillage methods used in three growing seasons in North Carolina to identify factors that influence the onset of *Stagonospora nodorum* blotch in winter wheat

Site	County	Region	Field type <sup>b</sup>	Tillage method <sup>a</sup>		
				2012	2013	2014
Aurora	Beaufort	Tidewater	Grower	CT	-	-
Caswell Farm	Lenoir	South-Central	Research	-	CT, NT	-
Cunningham Station	Lenoir	South-Central	Research	CT	CT, NT	CT, NT
Hertford	Perquimans	Tidewater	Grower	-	CT, NT	-
Lake Wheeler Road	Wake	South-Central	Research	CT	CT	-
Monroe	Union	Piedmont	Grower	-	NT	-
Piedmont Station	Rowan	Piedmont	Research	NT	CT	CT, NT
Tidewater Station	Washington	Tidewater	Research	NT	CT, NT	CT, NT
Tyner	Chowan	Tidewater	Grower	-	-	CT, NT
Rowland	Robeson	South-Central	Grower	CT	-	-
Upper Mountain Station	Ashe	Piedmont	Research	NT	CT, NT	-
Walkertown	Forsyth	Piedmont	Grower	-	CT, NT	CT, NT

<sup>a</sup> CT = conventional tillage representing complete burial of residue; NT = no tillage; and “-” indicates no experiment was conducted at the site in that year.

<sup>b</sup> Research fields were managed by personnel on the research station, while grower fields were managed by individual growers. In some cases, both conventional- and no-tillage experiments were conducted in separate fields at the same site.

**Table 4.2.** Description of potential predictors tested to determine their usefulness in predicting the onset of *Stagonospora nodorum* blotch in winter wheat

Variable	Type	Description
Cultivar resistance	Ordinal	1 to 9 <sup>a</sup>
Latitude	Continuous	-
Longitude	Continuous	-
Previous crop	Dichotomous	1 (wheat) or 0 (no wheat)
Seeding rate	Dichotomous	1 (standard) or 0 (reduced) <sup>b</sup>
Seed treatment	Dichotomous	1 (yes) or 0 (no)
Tillage type	Dichotomous	1 (no-till) or 0 (conventional)
Wheat residue	Dichotomous	1 ( $\geq 20\%$ coverage) or 0 ( $< 20\%$ coverage)
cDIV.1 <sup>c</sup>	Continuous	cDIVs until 1 week before day of the year 102
cDIV.2	Continuous	cDIVs until 2 week before day of the year 102
cDIV.3	Continuous	cDIVs until 3 week before day of the year 102

<sup>a</sup> Cultivar resistance ranged from 1 = most resistant to 9 = most susceptible to *Stagonospora nodorum* blotch.

<sup>b</sup> Reduced seeding rate was 80% of standard seeding rate of about 380 seeds/m<sup>2</sup>.

<sup>c</sup> cDIV = cumulative daily infection values, summarizing favorable temperature and moisture until a given day. Calculations of cDIV at all sites commenced at day of year 66 (first week of March).

**Table 4.3.** Results of PROC VARCLUS, variable reduction technique used to eliminate redundant variables

Cluster	Variable	R <sup>2</sup> with		1-R <sup>2</sup> ratio
		Own cluster	Next closest	
Cluster 1	Latitude	0.1860	0.0168	0.8279
	Longitude	0.3058	0.0017	0.6954
	cDIV.1 <sup>a</sup>	0.9549	0.0480	0.0474
	cDIV.2	0.9281	0.0425	0.0751
	cDIV.3	0.9437	0.0485	0.0591
Cluster 2	Wheat residue	0.7129	0.0047	0.2885
	Previous crop	0.7478	0.0069	0.2540
	Tillage type	0.7234	0.0011	0.2769
Cluster 3	Cultivar resistance	1	0.0376	0
Cluster 4	Seeding rate	1	0.0026	0
Cluster 5	Seed treatment	1	0.0376	0

<sup>a</sup> cDIV = cumulative daily infection values summarizing favorable temperature and moisture until a given day. cDIV calculation at all sites commenced at day of year 66 (first week of March).

**Table 4.4.** Spearman and Hoeffding coefficients for the correlation between potential predictor variables and time of disease onset

Variable	Spearman correlation		Hoeffding dependence	
	Coefficient	<i>P</i> -value	Coefficient	<i>P</i> -value
Latitude	0.0044	0.9306	−0.00002	0.3696
Longitude	0.2548	0.0001	0.0125	0.0001
cDIV.1 <sup>a</sup>	0.6674	0.0001	0.0854	0.0001
cDIV.2	0.7488	0.0001	0.1105	0.0001
cDIV.3	0.6016	0.0001	0.0689	0.0001
Wheat residue	0.4316	0.0001	0.0180	0.0001
Previous crop	0.3881	0.0001	0.0115	0.0001
Tillage type	0.1734	0.0001	0.0006	0.2055
Cultivar resistance	−0.0148	0.7716	−0.0029	1
Seeding rate	0.0536	0.2909	−0.0024	1
Seed treatment	0.0536	0.2909	−0.0024	1

<sup>a</sup> cDIV = cumulative daily infection values that summarize the favorable temperature and moisture that has accumulated until a given day. cDIV calculation at all sites commenced at day of year 66 (first week of March).

**Table 4.5.** Parameter estimates of the final binary logistic regression model developed to predict whether onset of *Stagonospora nodorum* blotch in winter wheat would occur before mid-April in North Carolina

Parameter	df	Estimate	Standard error	Wald $\chi^2$	$P > \chi^2$
Intercept	1	-5.45	0.515	111.9	0.0001
Wheat residue ( $\geq 20\%$ )	1	2.15	0.324	44.1	0.0001
cDIV.2 <sup>a</sup>	1	1.38	0.158	77.0	0.0001

<sup>a</sup> Cumulative daily infection values until two week before mid-April. cDIV summarizes favorable temperature and moisture for SNB infection that has accumulated until a given day. cDIV calculation at all sites commenced at day of year 66 (first week of March).

**Table 4.6.** Performance characteristics of the final binary logistic regression, the 10-fold and leave-one-out cross validation models

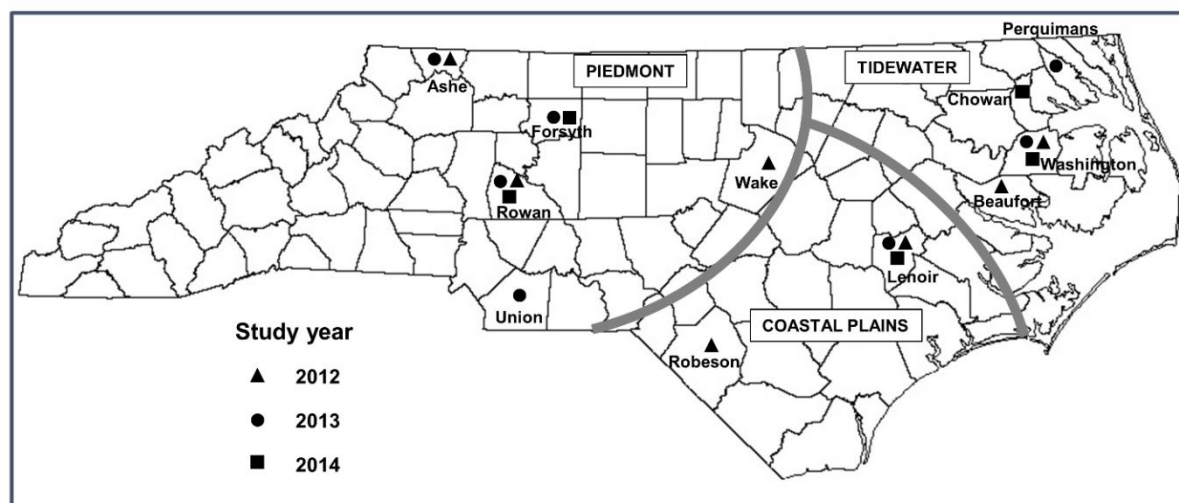
Model	AUC <sup>a</sup>	Correct classification <sup>b</sup>	Sensitivity <sup>c</sup>	Specificity <sup>c</sup>
10-fold cross validation	0.94	0.91	0.97	0.86
Leave-one-out cross validation	0.94	0.94	0.99	0.91
Final binary logistic regression	0.94	0.94	0.99	0.91

<sup>a</sup> Area under receiver operating curve, an estimate of the prediction accuracy of a model.

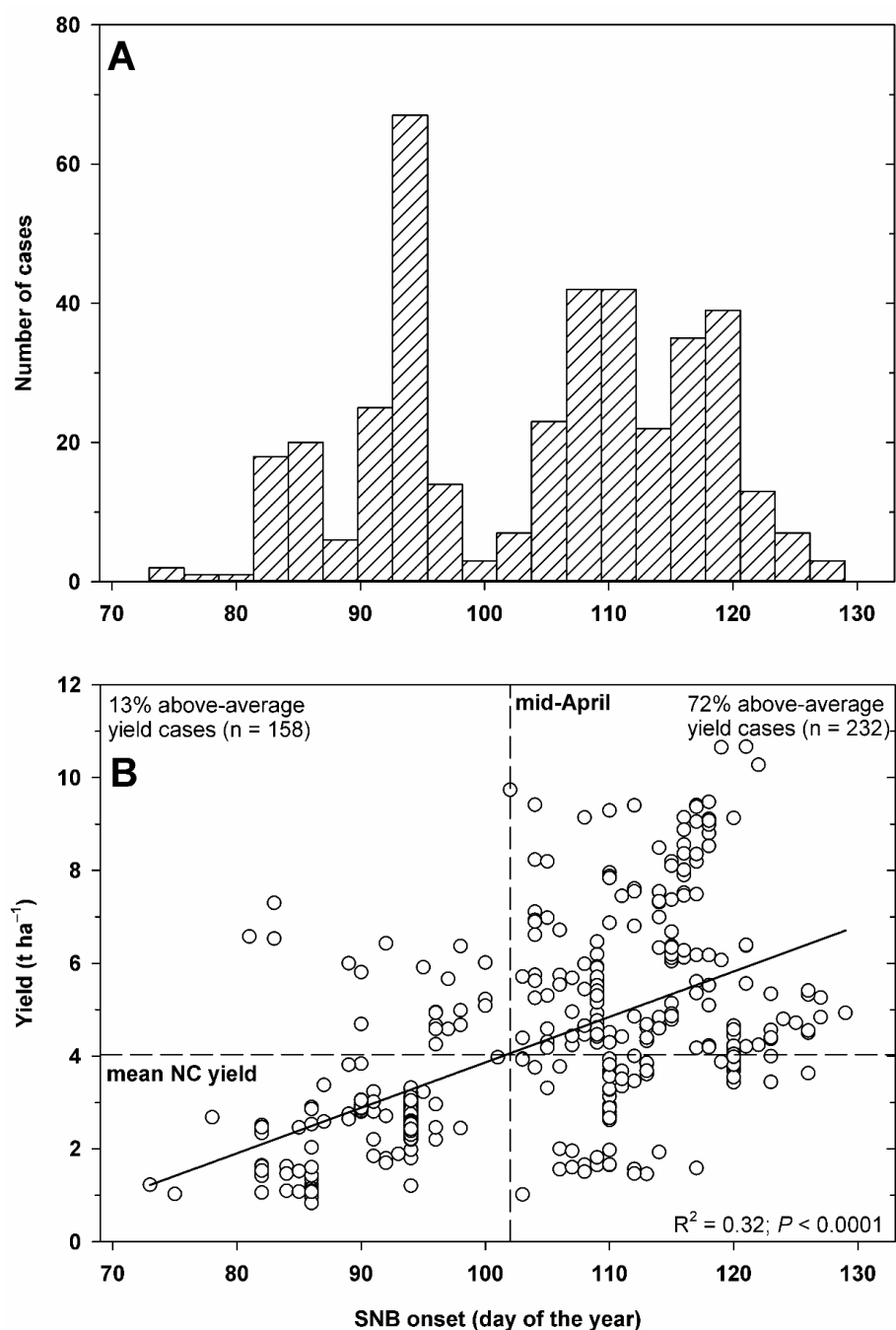
<sup>b</sup> Correct classification is the proportion of correctly classified cases within the dataset.

<sup>c</sup> Sensitivity is the proportion of correctly classified cases in class = 1 (i.e., where disease onset occurred before mid-April) and Specificity is the proportion of correctly classified cases in class = 0 (i.e., where disease onset occurred after mid-April).

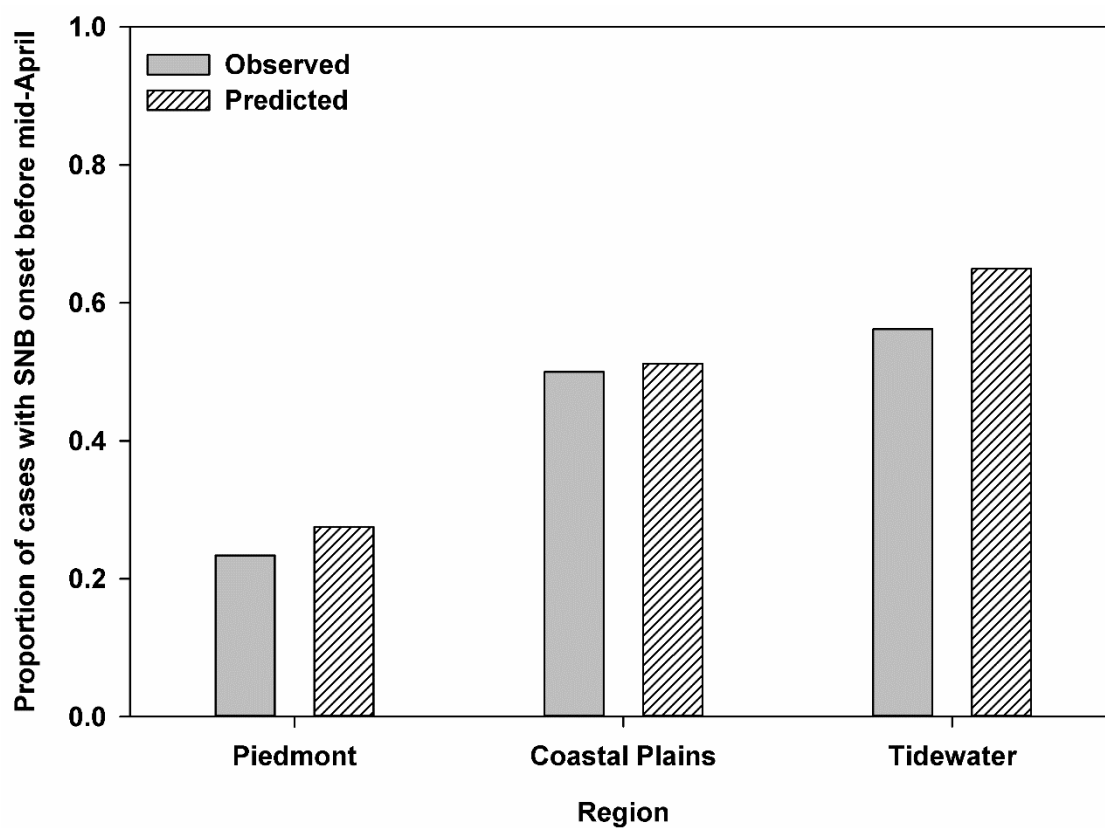




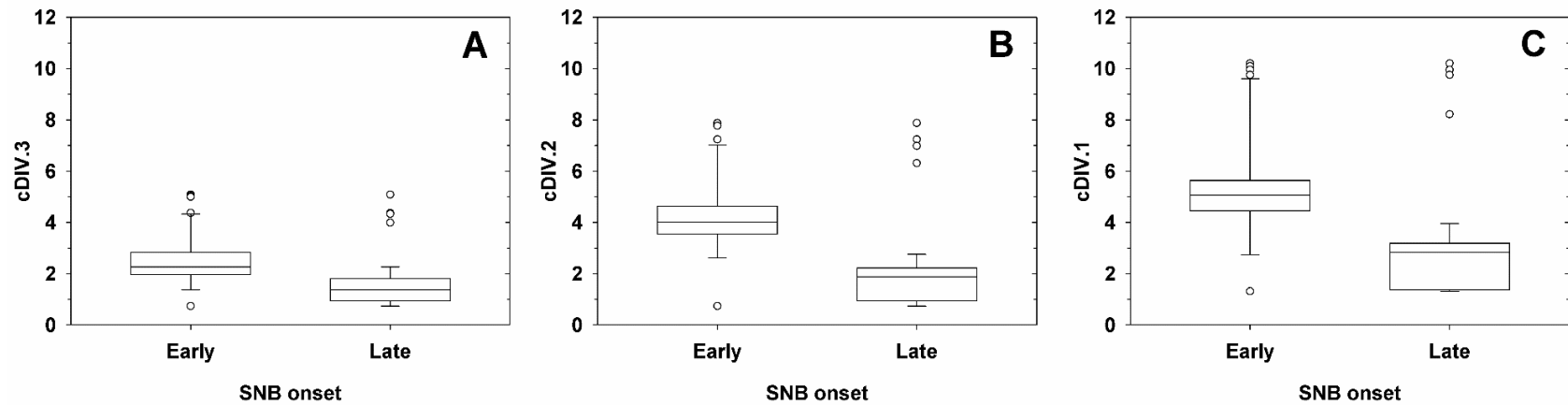
**Figure 4.1.** Location of field experiments conducted in 11 counties in North Carolina in the 2012- 2014 growing seasons to determine factors influencing the onset of *Stagonospora nodorum* blotch in winter wheat. Thick grey lines are approximate boundaries of three geographical regions of the state: Piedmont, Coastal plains, and Tidewater.



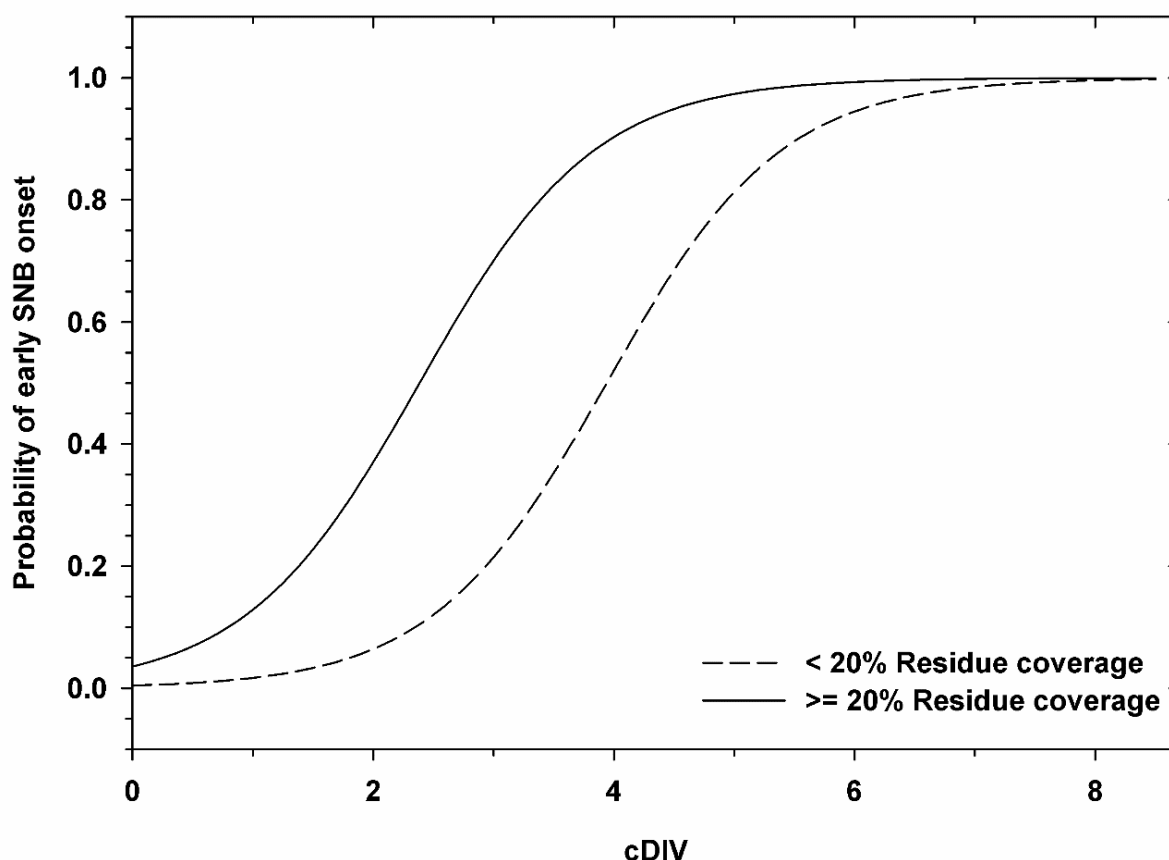
**Figure 4.2.** For 390 disease cases (unique combinations of pre-planting factors), **A**, distribution of time of *Stagonospora nodorum* blotch (SNB) onset in wheat, and **B**, a scatter plot showing relationship of SNB onset time with yield. Disease onset occurred before mid-April in 158 cases, and after mid-April in 232 cases. The dashed vertical line represents day of year 102 (mid-April), while the dashed horizontal line represents the mean (10 year average) wheat yield of  $4.03 \text{ t ha}^{-1}$  in North Carolina.



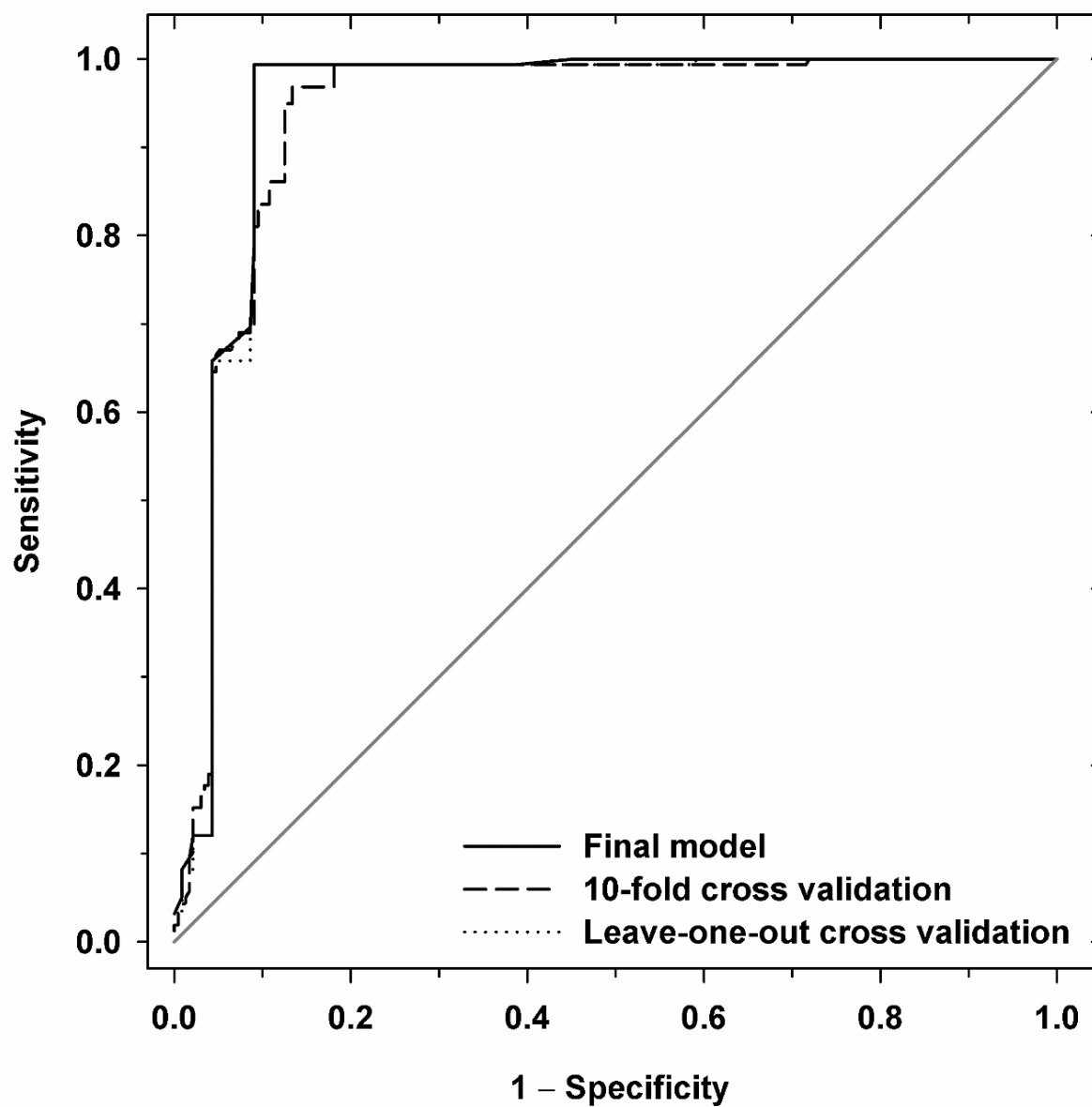
**Figure 4.3.** In field experiments conducted on wheat in three geographical regions of North Carolina, the observed and predicted proportions of cases where onset of *Stagonospora nodorum* blotch occurred before mid-April.



**Figure 4.4.** Distribution of cumulative daily infection values (cDIV) until **A**, three, **B**, two, and **C**, one week before a mid-April threshold (day of year 102) within both early and late onset classes. The lower, middle, and upper horizontal lines of each box are 25<sup>th</sup>, 50<sup>th</sup> (median), and 75<sup>th</sup> percentiles of the data, respectively. The vertical bars extend to 10<sup>th</sup> and 90<sup>th</sup> percentiles, and circles are outliers. The cDIVs measure accumulated favorable weather for *Parastagonospora nodorum* growth.



**Figure 4.5.** Predicted probability for the onset of *Stagonospora nodorum* blotch occurring before mid-April in low (<20%) and high ( $\geq$ 20%) wheat residue cases as a function of cumulative favorable weather for *Parastagonospora nodorum* infection observed from day of year 66 (first week of March) until two weeks before day of year 102 (mid-April).



**Figure 4.6.** Receiver operating characteristics curves for the final binary logistic regression, 10-fold and leave-one-out cross-validated models developed to predict the onset of *Stagonospora nodorum* blotch in winter wheat.

## CHAPTER 5

### Conclusions

The research reported in this dissertation provides a comprehensive assessment of the risk of *Stagonospora nodorum* blotch (SNB) in winter wheat caused by the fungus *Parastagonospora nodorum*. Weather variables and several pre-planting factors such as cultivar resistance, location, previous crop, seeding rate, seed treatment, tillage type, and wheat residue were evaluated as potential predictor variables in the quantitative analysis to evaluate their combined effect on the onset and development of SNB.

Wheat residue infected by *P. nodorum* harbors overwintering fruiting bodies of the pathogen and can therefore serve as one of the primary sources of inoculum. However the effects of different levels of infected wheat residue on the severity of SNB had previously not been quantified. A non-linear relationship between the amount of residue on the soil surface and disease severity was observed in this study. Residue levels up to 10 to 30% soil cover resulted in a rapid increase in disease severity but subsequent increase in residue levels resulted in little to no increase in the disease. The amount of residue on the soil surface also affected grain yield, but the effects were significant only when disease onset occurred early in the season and weather conditions during the season were conducive for rapid disease development. This finding has an important implication in the management of SNB since in the Mid-Atlantic region of the United States, the disease is one of the foliar diseases most frequently targeted for fungicide applications to wheat. The observation that yield and test weight losses appear to be associated with SNB only in more severe epidemics indicates that the economic threshold for the disease may be higher than previously assumed.

Pre-planting factors previously correlated with the late-season severity of SNB were used to develop risk assessment models that could be useful in making disease management decisions prior to planting of the wheat crop. Two analytical techniques, multiple regression and machine learning algorithms, were used to find good rules for predicting the SNB severity at end-of-season using pre-planting predictors. In general, models developed in this study identified longitude, latitude, cultivar resistance and amount of wheat residue as important predictors of SNB severity. These models can provide growers with a quantitative assessment of the risk of SNB at the beginning of season, and a choice can be made regarding which cultivars should be grown to minimize the risk depending on the location and cultural disease management programs.

Additional analysis showed that wheat yield was associated with time of SNB onset, and the onset time explained 32% variation in yield. If the disease onset occurred after mid-April, 72% of disease cases had above-average yield, while only 13% of the cases had above-average yield when disease onset occurred before mid-April. Subsequently, pre-planting factors and weather variables were investigated as potential predictors to determine whether SNB onset would occur before mid-April, and this was accomplished using logistic regression analysis. Out of the 11 potential predictor variables tested, wheat residue, and favorable weather for *P. nodorum* infection until two weeks prior to mid-April were identified as key predictors in the binary logistic regression model. This model can help growers assess if there will be an early or late onset of SNB in their field and possibly the need for fungicide application. If an early onset of SNB (before mid-April) is predicted,



growers can plan fungicide application to prevent potential yield losses associated with SNB. However, if the predicted onset is late (after mid-April), then there is less likelihood of incurring SNB associated yield losses and fungicide applications may not be required.

Based on the risk assessment model of SNB, risk maps for the disease can be prepared and distributed to wheat growers within the state. Each year, during the first week of April, which corresponds to start of stem elongation growth stage in North Carolina, weather data from weather station maintained by North Carolina Climate Office can be downloaded and used to calculate appropriate weather variables. These weather variables can be used as inputs in the binary logistic model and the risk of early onset can be determined where wheat residue levels in the field are known. Alternatively, a web based application can be developed where the user can enter their physical address and an estimate of wheat residue on the ground and the application will output whether the SNB onset will be early or late. The first week of April is time in the growing season when growers have adequate time to make decisions on fungicide applications and the model can provide an early warning for SNB outbreak. Since, the in-season model is not dependent on any given location, it can be easily tested and adapted to other wheat growing areas of the United States. Once validated externally, the models developed in this study have good potential to serve as pre-planting and within-season decision support tools to help growers make informed decisions to manage SNB in their wheat crop.