A CYBER AGE APPROACH FOR GLOBAL MANAGEMENT OF
BARLEY YELLOW DWARF VIRUS IN WINTER WHEAT

A Thesis in
Entomology
by
Joseph Thaddeus Walls III

© 2014 Joseph Thaddeus Walls III

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

December 2014
The thesis of Joseph Walls III was reviewed and approved* by the following:

Cristina Rosa  
Assistant Professor of Plant Virology  
Thesis Co-Advisor

John Tooker  
Associate Professor of Entomology  
Thesis Co-Advisor

Edwin G. Rajotte  
Professor of Entomology

Joseph Russo  
President and Chief Scientist of ZedX, Inc.

Gary Felton  
Professor and Department Head of Entomology

*Signatures are on file in the Graduate School
ABSTRACT

Precision agriculture often uses computer-based decision-support systems (DSS’s) to disseminate pest and disease information to growers to more efficiently manage agricultural productions. In this thesis, a DSS is developed to be accessed by growers for management of barley yellow dwarf disease caused by *Barley yellow dwarf virus*. The disease devastates grain growing regions around the world in epidemic patterns. It is a well-studied disease, but management can be greatly improved with determination of necessity and optimal timing of insecticide-treated seed, planting date, pest scouting, and foliar insecticide spray treatments, in chronological order. Using published literature and interviews with experts in BYDV epidemiology and agricultural decision-making, dependency networks were used to model field conditions that would logically warrant these management actions. The networks represented nine possible outputs: use insecticide-treated seed, use untreated seed, plant crop immediately, delay planting, scout for aphid vectors of BYDV, do not scout, full foliar insecticide spray, ½ (diluted) insecticide spray, and no insecticide spray. There were a total of 243 total combinations of conditions to reach the seed treatment recommendations, 9,720 to reach the planting date recommendations, 62,208 to reach the scouting recommendations, and 216 to reach the insecticide spray recommendations. In this work, I consider and strive to improve the mechanism for inferring output recommendations even when using only partial data sets.

Inference mechanisms are necessary components of DSS’s to extrapolate outputs from input data to give users recommendations. The dependency networks represent
inference mechanisms that require all input information be present before a management recommendation can be made. This thesis proposes a novel secondary inference mechanism structure to be overlaid onto the dependency networks that uses a numerical, rather than categorical or ordinal, calculation system to handle partial input information. This inference mechanism used the dependency networks as a template to make a prototype numerical representation of importance of field conditions in making management decisions. Secondly it calculates a likely success of these management decisions when executed, and penalizes the grower if he or she executes an incorrect management tactic. The success or penalties are measured in terms of optimum yield. It is also proposed that this secondary inference mechanism can allow a BYD management decision forecast based on pest and disease statuses, as well as real-time recommendations.

The purpose of the DSS developed in this thesis is to show the applicability of implementing DSS’s based on expert knowledge into a platform (iPIPE) that is capable of gathering data from users in a two-way feedback loop. Since future management decisions rely on previous ones the feedback loop allows management practices conducted by the grower to alter future management recommendations given by the system. It also enhances large scale (regional) pest monitoring with input of individual field data. The DSS reported in this thesis will serve as a basis for the evolution of precision management of crop diseases. It will aid in reducing input cost and increasing sustainability and cereal grain yield in BYD management and eventually it will serve as model to better manage other crop pests and diseases.
TABLE OF CONTENTS

List of Figures ........................................................................................................... viii
List of Tables .............................................................................................................. xi
Acknowledgements ................................................................................................... xii

Chapter 1. Introduction ............................................................................................... 1
    Introduction ............................................................................................................. 1
    Literature Review .................................................................................................. 3
        Taxonomy .......................................................................................................... 3
        History of BYD ................................................................................................. 4
        Economic Impacts ............................................................................................... 5
        Aphid Vector Biology ......................................................................................... 5
        Winter Wheat Biology ........................................................................................ 8
        Virus-Host Interaction ....................................................................................... 8
        Vector-Host Interaction ..................................................................................... 10
        Vector-Virus Interaction ..................................................................................... 11
    Management of BYD (Technological and Biological) ............................................ 13
        Management Techniques .................................................................................... 13
        Expert Decision-Support Systems ..................................................................... 16
        Recent Developments in DSS Technology ......................................................... 17

Chapter 2. Modeling the Decision Process for BYD Management ............................. 20
    Introduction ............................................................................................................. 20
    Methods ................................................................................................................. 28
        Dependency Networks ........................................................................................ 28
        Decision Framework .......................................................................................... 29
    Results .................................................................................................................... 33
        Dependency Networks ........................................................................................ 36
        Treated Seeds ..................................................................................................... 37
            Aphid/BYD Problem in Previous Years (History of Field) ............................. 38
            Decision to Buy Treated Seeds ...................................................................... 39
        Planting ............................................................................................................... 40
            Likelihood of Aphid Problem during the Coming Season .............................. 41
            Migration Peak/Susceptible Phase Overlap .................................................... 41
            Delay Planting Decision ............................................................................... 42
        Scouting .............................................................................................................. 44
            Infectivity Risk ............................................................................................... 44
            Current Aphid Risk Status ............................................................................. 44
            Current Plant Susceptibility ......................................................................... 45
            Observational Data of Aphid Population ...................................................... 45
            Decision to Scout ......................................................................................... 46
        Spraying .............................................................................................................. 46
    Discussion .............................................................................................................. 47
Chapter 3. A Proposed Mechanism to Estimate Missing Information for the BYD-DSS
Decision-Support System........................................................................................................52

Introduction...............................................................................................................................52
Decision-Support System (DSS) Components............................................................54
Knowledge Base.........................................................................................................................54
Inference Mechanism ............................................................................................................54
Database and User Interface ...............................................................................................55
BYD-DSS.................................................................................................................................56
Brief Explanation of the BYD-DSS .........................................................................................56
Methods..................................................................................................................................58
Scoring Dependency Networks ...........................................................................................58
Inference Mechanism ............................................................................................................59
Output.....................................................................................................................................59
Threshold.................................................................................................................................59
Score.......................................................................................................................................59
Values....................................................................................................................................60
Weights.................................................................................................................................60
Risks.......................................................................................................................................61
Success of Management Decisions ......................................................................................62
Success.................................................................................................................................62
Results.....................................................................................................................................63
Threshold.................................................................................................................................63
Score.......................................................................................................................................64
Values....................................................................................................................................64
Weights.................................................................................................................................65
Risks.......................................................................................................................................65
Success of Management Decisions ......................................................................................66
Treated Seeds ..........................................................................................................................68
Example of how Success can be Calculated Using Incomplete Information .......69
Planting Date ..........................................................................................................................71
Spraying.................................................................................................................................71
Discussion..............................................................................................................................72

Chapter 4. Discussion ............................................................................................................75

Major Accomplishments .....................................................................................................76
Inference Mechanism .........................................................................................................77
Possible Improvements .....................................................................................................78
Future Work ..........................................................................................................................81

Appendix A. Dependency Networks from Chapter 2 .........................................................83

Appendix B. Novel Inference Mechanism via Scoring System from Chapter 3 .................93

Appendix C. Statistical Analyses of Aphid Migration Data from Italy and Midwest
Unite States ..........................................................................................................................105
LIST OF FIGURES

Figure 2-1. Timeline of BYD decision-making with winter wheat and aphid events overlaid. This timeline is designed for the Pennsylvania region, thus other regions may have slightly different dates, but similar chronology..........................................................29

Figure 2-2. General outline of the BYD-DSS system interconnectedness of conditions listed in List 1. Controllable conditions are in ovals, whereas uncontrollable environmental conditions are in rectangles. Connections are color-coded in chronological order pertaining to the management recommendation they affect. Red connections are calculated to determine necessity of treated seeds first, then green connections to determine timing and necessity of scouting, and finally violet to determine timing and necessity of insecticide spray. Dashed lines indicate where a projective development model will need to be run to output the connected condition. Dotted lines indicate there is an incomplete source of data that may need to be filled in by the user using the system......35

Figure 2-3. Flow chart of entire the BYD-DSS. The BYD-DSS describes the knowledge base of dependency networks reported in this chapter..........................................................36

Figure 2-4. Current plant susceptibility dependency network. Using operators the networks can be read in “IF/THEN” format. For example, the orange path can be read as “IF the plant is less than 2 weeks of age AND a seed treatment/previous spray is in effect on the crop, THEN the crop has a medium susceptibility to infection of BYDV.” Secondly, there can be multiple variables combined in a single path as seen with the red path. This can be read as “IF the plant is less than two weeks of age OR 2-4 weeks of age, AND there is no seed treatment/previous spray in effect, THEN the crop has a high susceptibility to infection of BYDV.”..................37

Figure 3-1. Correlation of Weights, Risks, Values, Scores, Thresholds, and Outputs in the new inference mechanism. Where X refers to any variable of the condition ..............62

Figure 3-2. Scored ‘Decision to Buy Treated Seeds’ dependency network.................................66

Figure 3-3. Flow chart of management decision Success (and failure). Each management decision proposed by the system can be followed or not, and the failure to follow an advised decision carries penalties that can be calculated and seen by the grower on a smart phone app or web based interface. The grower will be able to see a decision tree similar to this one, but the computer will optimize the management program by finding the set of practices leading to the highest Success, profit, and sustainability desired by the user (the latter two are not calculated by this tree). As seen from this figure, each management decision affects the others.................................................................68

Figure A-1. Aphid/BYD Problem in Previous Years Dependency Network...............................83

Figure A-2. Decision to Buy Treated Seeds Dependency Network........................................84
Figure A-3. Likelihood of Aphid Problem During the Coming Season Dependency Network

Figure A-4. Migration Peak/Susceptible Phase Overlap Dependency Network

Figure A-5. Delay Planting Decision Dependency Network

Figure A-6. Infectivity Risk Dependency Network

Figure A-7. Current Aphid Risk Status Dependency Network

Figure A-8. Observational Data of Aphid Population Dependency Network

Figure A-9. Decision to Scout Dependency Network

Figure A-10. Decision to Spray Dependency Network

Figure B-1. Aphid/BYD Problem in Previous Years Scored Dependency Network

Figure B-2. Decision to Buy Treated Seeds Scored Dependency Network

Figure B-3. Likelihood of Aphid Problem During the Coming Season Scored Dependency Network

Figure B-4. Migration Peak/Susceptible Phase Overlap Scored Dependency Network

Figure B-5. Delay Planting Decision Scored Dependency Network

Figure B-6. Current Aphid/BYD Risk Status Dependency Network

Figure B-7. Current Plant Susceptibility Scored Dependency Network

Figure B-8. Decision to Spray Scored Dependency Network

Figure B-8. Decision to Scout Scored Dependency Network

Figure C-1 Illinois R. padi Suction Trap Catches in 2007

Figure C-2. Illinois R. padi Suction Trap Catches in 2009

Figure C-3. Illinois R. padi Suction Trap Catches in 2010

Figure C-4. Indiana R. padi Suction Trap Catches in 2007

Figure C-5. Indiana R. padi Suction Trap Catches in 2009

Figure C-6. Indiana R. padi Suction Trap Catches in 2010

Figure C-7. Iowa R. padi Suction Trap Catches in 2007
Figure C-8. Iowa *R. padi* Suction Trap Catches in 2009 ...................................................... 110
Figure C-9. Iowa *R. padi* Suction Trap Catches in 2010 ...................................................... 110
Figure C-10. Kansas *R. padi* Suction Trap Catches in 2007 ...................................................... 111
Figure C-11 Kansas *R. padi* Suction Trap Catches in 2009 ...................................................... 111

Figure C-12. This figure shows the average catch of *R. padi* in suction traps from Carmagnola, Italy from 1982-1989. This graph omitted Julian date 166 from 1987 because there were 1,692 individuals caught that day, which was skewing that data point .............................................................. 112

Figure C-13. Average *S. avenae* caught in suction traps from 1982-1989. .............................. 112

Figure C-14. Average *S. avenae* caught in suction traps from 1982-1989 without 1985, an outbreak year for *S. avenae* .............................................................. 113

Figure C-15. Average *R. maidis* caught in suction traps from 1982-1989 .............................. 113

Figure C-16. Suction trap catches of *R. padi* from Udine, Italy, 1985. This is an example of a large magnitude migration. The migration starts at about 1100 DD, but is still dispersed around 1500 DD .............................................................. 114

Figure C-17. Suction trap Catches of *R. padi* from Udine, Italy, 1986. This is an example of a small magnitude migration. The migration starts at approximately 1400 DD ....... 114

Figure C-18. *R. padi* suction trap catches plotted against the mean wind speed of the day they were caught. Every data point represents a day from 1983-1992. .............................. 115
LIST OF TABLES

Table 2-1. Conditions used to manage BYD. Controllable management decisions are in bold, whereas uncontrollable environmental conditions used to decide best management practices are in normal font. Each subset describes the conditions used in determining the condition above it. Source of reasoning for influence of each variable in decision process are listed after each condition, if applicable ..................33

Table 3-1. Treated seeds management decision scoring table. This is table 1 of 3 concerning management decision scoring. Others can be seen in Appendix B. Columns 1, 4, and 7 are the conditions and variables present in the Decision to Buy Treated Seeds network. Columns 2, 5, and 8 are the values of each variable. Column 10 is the Score of each combination of variables based on the sum of observed values. Columns 11 and 12 are the Success of treated seeds and untreated seeds, respectively. Columns 3, 6, and 9 are the average Scores given only knowing the variables in columns 1, 4, and 7, respectively (e.g. the average Score given the only data available is that the field has had a history is 48; and for knowing the field has not had a history, the average Success is 32). The equation to calculate the Success of untreated seeds is If Score > 46, then ((66-Score)/20)*100, otherwise 100. 46 is the highest Score that would be suggested to by untreated seeds. 66 is the highest overall Score. 20 is the difference between 66 and 46. Finally, 100 is the highest possible Success.................................................................70

Table B-1. Planting Date management Success and Scores ..................................................102

Table B-2. Insecticide spray management Success and Scores ............................................103

Table D-1. Primers used for BYDV testing .................................................................119

Table D-2. Reagents for each vial of PCR reaction .........................................................119

Table D-3. PCR Cycles .................................................................................................120
ACKNOWLEDGEMENTS

I would like to start off by thanking my co-advisor Cristina Rosa and also my committee member Ed Rajotte. Their patience and guidance helped push me through my research and the writing process to obtain a final product that I can say I am proud of. I would also like to thank my other committee members; Joe Russo, for spending many long evenings working on the project with me, and John Tooker, for helping with many biological questions, helping to push me in the right directions, and initially introducing me to the project team.

Piero Caciagli provided many years of his own data collection to allow statistical analyses of aphid migrations. These data were invaluable in gaining knowledge on aphid phenology. He was also incredibly accommodating on my first trip out of the United States to his home town of Torino, Italy.

To Drs. Elizabeth Borer and Christelle Lacroix of the University of Minnesota for supplying primer sequences for the PCR testing technique and Bryce Falk of the University of California, Davis for mailing positive controls for BYDV species. Thank you all for so generously contributing to the project.

Finally, I want to deeply thank my family and friends, including all the students and faculty of the entomology department. They have been patient and supportive of me despite of all of the missed calls over the past two years. Their support and guidance has led me to where I am today. It is due to my desire to honor them that I have pushed through the difficult times to better myself. Also to my dog, Triumph, who greets me with the same happiness and enthusiasm every day, no matter what. He has provided a consistency and friendship in my life to help me maintain my sanity.

Thank you all so much for your unending support. This thesis would not have been possible without all of you.
Chapter 1

Introduction

Decision-support systems (DSS) are becoming important tools in agriculture. These systems can be used as an integrated pest management (IPM) strategy that helps farmers in planning important crop management decisions to obtain high yield, economic gain, and/or sustainability while following regulations. In a DSS statistical models are generally used to predict the probability of the outcome of any set of inputs or observations. Decision-support systems use these models, in conjunction with heuristic, experience-based algorithms to give the farmer direct suggestions on how to farm and increase crop yield, economic gain, and sustainability.

Technology has advanced enough to allow us to obtain and share information instantaneously via the Internet. Many agricultural processes have become automated using programs that control machines, such as irrigation systems, without human intervention (Auat Cheein & Carelli 2013; Clemmens & Schuermans 2004). These systems are controlled via a feedback loop from monitors detecting environmental conditions, which elicit an action, which in turn changes the environment. The same process can be modeled with disease management for crop diseases; however, instead of only machines and monitors providing information, growers can directly input information from their fields. This allows for an accurate estimation of the best control strategies for any given location.
Barley yellow dwarf (BYD) is a disease caused by *Barley yellow dwarf virus* (BYDV). BYDV has a complex life cycle and is difficult to manage. It is the most economically important viral disease of small grains worldwide. Global statistics are difficult to estimate due to lack of data and misdiagnosis (Miller and Rasochovà 1997); however, individual wheat fields in areas prone to viral infection may experience average yield losses between 11 and 33% and sometimes up to 80% (Miller and Rasochovà 1997, Pike 1990). There are several components in the biology of the disease cycle that must be realized. BYDV is a virus complex, comprising viruses belonging to different genera. These viruses can be transmitted by more than 25 species of aphids worldwide. The aphids feed upon, and transmit BYDV to, various grain crops such as wheat, barley, oats, rye, and corn, as well as non-crop species. The hosts of BYDV include over 150 species of grasses in the family *Poaceae* (Miller and Rasochovà 1997). In Pennsylvania, where this thesis work is based, there are four main aphid vectors of BYDV: *Rhopalosiphum padi* (bird cherry-oat aphid), *R. maidis* (corn leaf aphid), *Sitobion avenae* (English grain aphid), and *Schizaphis graminum* (greenbug).

This chapter provides a brief overview of *Barley yellow dwarf virus*, the biology of its vectors and hosts, and the utility of decision-support engines in managing BYDV and, in general, insect-vectored pathogens. This thesis will report on a novel construction of a single decision-support system by modeling the logic for managing BYD in winter wheat, and using computer databases to accommodate real-time farmer input and meteorological data to calculate the most efficient course of action for any given location in the world. Wheat is the focal crop due to its importance in agriculture and food production and the economic damage from BYD worldwide.
Literature Review

Taxonomy

The vectors of BYDV are commonly known as aphids. Aphids are insects (Class Insecta) in the order Hemiptera, family Aphididae. This family is very diverse, having 500 genera and over 4,400 species, approximately 250 of which are important crop pests, mainly due to their ability to vector pathogens as they feed on plant phloem. The four main vectors of BYDV belong to the following genera and species: Rhopalosiphum padi, R. maidis, Sitobion avenae, and Schizaphis graminum.

There are five viruses in the United States that cause the disease known as barley yellow dwarf. These are Barley yellow dwarf virus (BYDV-PAV, BYDV-MAV, BYDV-RMV, BYDV-SGV), and Cereal Yellow dwarf virus (CYDV-RPV). The transmission efficiency of each virus type depends, in part, on the vector species (Rochow 1969; Johnson and Rochow 1972). BYDV’s belong to the family Luteoviridae; their nucleic acids are made of a single linear molecule of positive-sense, single stranded RNA ranging from 5.7 to 6.0 kb in size (Miller et al. 2002; Smith and Barker 1999; Miller and Rasochovà 1997). The viruses BYDV-PAV and BYDV-MAV belong to the genus Luteovirus, in 1999 the BYDV-RPV species was reclassified as Polerovirus and commonly named Cereal yellow dwarf virus (CYDV) (Smith and Barker 1999). BYDV-RMV and BYDV-SGV have yet to be officially assigned to genera (Wu et al. 2011), so for simplicity they will be referred to as BYDV’s. BYDV-PAV is transmitted most efficiently by Rhopalosiphum padi and Sitobion avenae, BYDV-MAV specifically by Sitobion avenae, BYDV-RPV specifically by R. padi, BYDV-RMV by R. maidis, and BYDV-SGV most efficiently, but not specifically, by Schizaphis graminum (Rochow...
BYDV can infect a wide range of grasses in the family *Poaceae* (formerly *Gramineae*). The host species that will be focused on primarily in this thesis will be *Triticum aestivum* (winter wheat).

**History of BYD**

BYD first became a problem in the United States in 1890. The disease reached epidemic proportions and was widespread throughout the south and Midwest that year (Webster and Phillips 1912). The “greenbug” vector, now known as *Schizaphis graminum*, is thought to have been introduced to the United States from England in 1882 (Webster and Phillips 1912). Another epidemic occurred in 1907 and T.F. Manns of Ohio published the first report of the disease. The disease was thought to affect oats alone, but was later realized to be more of a generalist. Manns believed the disease was caused by a bacterium and transmitted by “plant lice,” which later came to be known as aphids. Although he was correct in assuming transmission by the aphids, the claim of the disease caused by a bacterium was incorrect. In 1951, another epidemic year, the disease spread to barley crops in California. During this outbreak Oswald and Houston were first to describe the causal agent of the disease as a virus transmitted by multiple aphid vectors. It is likely that the virus that caused this outbreak was BYDV-PAV (Gray and Gildow 2003).

The evolution of BYD came from a close association between the virus, aphid vector and host (Martin et al. 1990). *Luteoviruses*, are unique in that they are limited to
plant phloem, their interactions with aphid vectors, and limitation to specific families of plants (Martin et al. 1990). It is likely that *Luteoviruses* evolved due to recombination with other plant viruses, probably sharing a common ancestry with *Carnation mottle virus* or *Southern bean mosaic virus* (Martin et al. 1990).

**Economic Impacts**

BYDV is the most economically important cereal grain virus in the world (Lister & Ranieri 1995). BYD annually causes between 11 and 33% global yield loss in wheat (Lister & Ranieri 1995). This range is so large because BYDV symptoms are easily misdiagnosed by an untrained eye as nutrient deficiencies or physical stress (Comeau 1990). There have been reports of up to 80% yield loss in infected fields (Pike 1990). In 2012 the Food and Agriculture Organization of the United Nations, statistics division (FAOSTAT) estimated a global production of wheat at 671,496,872 metric tonnes. Assuming a conservative global BYD yield loss at 11%, a yield loss of approximately 83 million metric tonnes can be attributed to the disease.

**Aphid Vector Biology**

Aphids are insects of the order *Hemiptera*, which can be distinguished by mouth parts specialized for piercing and sucking. Aphids use these straw-like mouth parts to feed on the phloem of plants. Certain aphids, such as *R. padi*, *S. avenae*, and *R. maidis* specialize in feeding on grains and grasses of the family *Poaceae*. These aphids normally reproduce via parthenogenesis, generally a single female founds an entire colony.
However they can be holocyclic (sexual) or anholocyclic (parthenogenetic). Studies conducted in the United Kingdom show a strong occurrence of an anholocyclic overwintering stage for *R. padi* and *S. avenae*, although they can occasionally be found to be holocyclic (Pons et al. 1995; Zhou et al. 1995; Hand 1989). *Rhopalosiphum maidis* shows anholocyclic overwintering stages as well, but has also been found to be completely holocyclic (Huggett et al. 1999). The United States has a relatively similar climate to the UK, so anholocyclic overwintering of these aphids are likely to occur in the US as well, depending on the severity of the winter. Preliminary data from suction traps also shows male alatae to be quite rare in the Midwest United States (Suction trap data from Doris Lagos at the University of Illinois at Urbana-Champaign), suggesting that most colonies are founded by a single female aphid making all the individuals in a colony clones until a male occurs allowing sexual reproduction.

As most insects, aphids are ectothermic poikilotherms; thus, they rely entirely upon ambient temperature for internal heat. Due to this, aphids are unable to develop and mature when temperatures fall below a certain threshold. The base developmental threshold of many aphid species, below which no development occurs, is 5°C (Turl 1980). The base developmental threshold of *R. padi* has been more specifically found to be 5.8°C with the upper threshold, above which the rate of development does not increase, is 25.1°C (Elliot and Kieckhefer 1989). Within the upper and lower threshold limits degree days can be calculated to predict development and migration of insects.

The phenologies of aphid vectors are cyclic and may have one or more migratory phases. Aphid migrations are preceded by the production of alatae that, under the right conditions, will take flight and migrate to new locations. Migratory patterns are a species-
specific trait that has been studied extensively. *R. padi* and *R. maidis* have two distinct migratory periods in most temperate climates (such as the United States and Northern Italy), one in late spring and another in autumn. *S. avenae* has a single migratory peak in late spring, but only a very small migration in autumn (Coceano et al. 2009). Temperature and rainfall/moisture are significant factors in predicting onset of BYDV aphid vector migrations (Thackray et al. 2009; Kendall et al. 1992; Fabre et al. 2006). Alate aphids also require a minimum temperature before internal body temperature is high enough to allow flight muscles to function. The threshold at which *R. padi* and *S. avenae* take flight on 50% of the days at a given temperature are days with a daily maximum temperature of 15.5°C and 16°C, respectively (Walters and Dixon 1984). These temperatures did vary, however, based on other variables such as substrate and wind speed, so it is difficult to interpret these results in the context of the single temperature variable. Dry and Taylor conducted lab experiments in 1970 on temperature thresholds for different aphid species flight. They found that the temperature threshold for *R. maidis* flight was different based on the substrate. Alatae had higher thresholds when on host plants, but lower thresholds when on glass. Since these studies were not conducted under field conditions, and much variability is observed in flight temperature threshold, it cannot be assumed that temperature would be a constraint in flight take-off in field conditions. Preliminary data from Piero Caciagli (Appendix C) indicates there may be more of an influence in average temperature fluctuation on alate flight rather than a specific temperature threshold. Concerning the fall migration of these aphids in August, September, and October overnight temperatures will be much lower than the maximums, so it is likely that migration flight occurs during specific parts of the day similar to other
aphid species (Lewis and Taylor 1965). Wind speed also affects the take-off of aphid alatae. Aphids are poor fliers and have difficulty flying in strong winds, so when strong winds are present flight take-off is often delayed. Time to take-off is positively correlated with wind speed, thus, the greater the wind speed the longer individual aphids will delay flight (Walters and Dixon 1984). Preliminary data from Piero Caciagli also shows a strong negative correlation between wind speed and aphid flight (Appendix C).

**Winter Wheat Biology**

Wheat production strategies vary due to climate and management goals of the farmer. Winter wheat is used for animal feed and for flour, which is used in most bread and baked goods. Winter wheat differs from spring wheat in that, as the name implies, it is planted in autumn and develops throughout the winter and early spring. Normal planting dates of winter wheat vary based on regional climate. Typical dates of planting generally range from September through October. For example, the United States Department of Agriculture (USDA) suggests winter wheat planting in Pennsylvania occurs between 15 September and 15 October (USDA 2010). Harvest dates usually range from late May through July; the USDA suggests 10-30 July in Pennsylvania (USDA 2010).

**Virus-Host Interaction**

Barley yellow dwarf virus, like most plant viruses, is named for the symptoms it causes in its hosts. Infected plants exhibit symptoms of stunted growth and discoloration
due to chlorosis of the leaves. Severe infections of BYDV in wheat can cause up to a 66% reduction in plant size, decreased heading, and an obvious yellow coloring due to chlorosis (Oswald and Houston 1953). Chlorosis in BYDV infected plants is caused by an accumulation of soluble carbohydrates and decrease in the nitrogen, calcium, and magnesium concentrations in the infected leaf tissues (Goodman et al. 1965; Riedell et al. 2007). BYDV infection causes approximately a 45% reduction in photosynthesis per plant, or a 25% reduction when comparing equal masses of healthy and infected plant tissues (Jensen & Van Sambeek 1972). This 25% reduction in photosynthesis is accompanied by a 65% reduction in chloroplasts (Jensen & Van Sambeek 1972). The plant must increase its chloroplasts' output under these circumstances so it reallocates its nutrient concentration to photosynthetic processes and away from growth, which causes the stunted nature of infected plants. These symptoms progress the most when temperatures hover around 25°C and more extreme temperatures inhibit infection and symptom development (De Wolfe 2002).

Susceptible Poaceae species may become infected with the virus at any time throughout their life cycle, but winter wheat is most susceptible and damaged during its seedling stages, specifically when inoculated before Growth Stage 31 (GS 31), the beginning of stem elongation (Zadoks et al. 1974; Smith and Sward 1982; Oswald and Houston 1953). When infected after tillering, during stem elongation, yield losses are dramatically decreased (Smith and Sward 1982).

Resistance to BYDV does not naturally occur in wheat; however there has been some work to isolate a resistance gene found in other wild relatives (Zhang et al. 2009). The bdv2 gene isolated from plants in the genus Thinopyrum is the most widely used
gene to study resistance (Zhang et al. 2009; Ayala-Navarrete et al. 2013). Other resistance genes have been isolated from Thinopyrum as it is widely used in BYDV resistance studies (Ma & Tomita 2013).

Strains of BYDV often differ in the severity of symptoms. Plumb (1974) listed BYDV-PAV and BYDV-RPV as “severe” and BYDV-MAV as “mild.” This makes R. padi the most dangerous and destructive of the BYDV aphid vectors.

**Vector-Host Interaction**

The primary hosts of R. padi, R. maidis, and S. avenae are the bird cherry-oat (Prunus padus), maize (Zea mays), and in general grasses, respectively. Rhopalosiphum padi, R. maidis, and S. avenae therefore do not feed specifically on Triticum aestivum, though it is one of the preferred secondary hosts (Leather and Dixon 1982; Lushai et al. 1997). Rhopalosiphum maidis, on the other hand, will feed on wheat but generally prefers barley (Foott 1977). The most preferred secondary host of R. padi is Lolium perenne, or perennial ryegrass, because aphids are highly attracted to this grass species and perform well on it (Leather and Dixon 1982).

Different aphid species have different fecundities based on many variables. Sitobion avenae has a much higher reproductive rate in wheat fields than R. padi or R. maidis (Coceano et al. 2009). This observation is especially apparent in seedling wheat plants when R. padi and S. avenae colonize the same plant (Chongrajitanameteekul et al. 1991). On a heading plant, S. avenae prefer the head, whereas R. padi prefer lower parts of the plant such as the stalk or leaves (Chongrajitanameteekul et al. 1991; Vereijken
Although *S. avenae* may be a stronger competitor once a field is colonized, *R. padi* and *R. maidis* have a more competitive migratory phase (Coceano et al. 2009). This is apparent in the two major migrations and number of individuals of *R. padi* and *R. maidis* caught in suction traps compared with the single migration of *S. avenae* (Coceano et al. 2009).

Direct damage from aphids is actually quite rare in the autumn. Yield losses may occur in the spring, generally from *S. avenae* due to the large spring migration, but only under extreme infestation circumstances (Oakley and Walters 1994). However, these yield losses cannot be considered direct because most of the damage comes from a fungal pathogen in the aphid honeydew (Vereijken 1979). In fact, one of the most devastating outcomes of aphid infestation is due to the transmission of BYDV in the autumn. Thus, managing the autumn migration of aphids is just as important in preventing yield loss as managing spring migration of aphids.

**Vector-Virus Interaction**

BYDV are transmitted in a circulative, persistent, nonpropagative manner (i.e. no replication within the vector), by aphids (Miller and Rasochova 1997). The acquisition of viral particles by the aphid vector alters the aphid’s behavior to promote the spread of virus. It has been shown that aphids viruliferous for BYDV prefer healthy plants, whereas non-viruliferous aphids prefer diseased plants (Ingwell et al. 2012). Virions are acquired during feeding from infected phloem cells. Once the virions are ingested by the aphid, they must pass through the midgut and/or hindgut into the hemocoel. Hindgut epithelium
recognition of virus particles is *Luteovirus* specific, but not serotype specific, as any of the serotypes can pass to the hemocoel of any vector, even if the vector is unable to transmit that specific serotype (Miller and Rasochova 1997). The virions are then actively transported to the salivary glands (Miller and Rasochova 1997). Virions contain structural proteins needed for transport through the salivary glands, and these proteins, together with unknown host proteins, are responsible for vector specificity (Chay et al. 1996). The infected aphid remains viruliferous for the rest of its life, hence persistent transmission. Aphids require a minimum acquisition access period for BYDV transmission that ranges between 15 minutes and 3 hours (Gray et al. 1991). The acquisition access period and transmission efficiency is dependent on virus titer and age of the plant and infection stage (Gray et al. 1991).

BYDVs infect monocot grasses, and not all BYDV aphid vectors feed primarily on these grasses (e.g. *R. padi*). Thus, the virus may only be present in a population after migration from the primary host. Proportions of viruliferous migrants in a given migration can range from 0% to over 10% (Coceano et al. 2009; Plumb 1976). In Italy, average viruliferous migrant percentages are 11.22, 0.71, 7.71, and 1.85 for *R. padi*, *S. avenae*, *R. maidis*, and *Sc. graminum*, respectively (Coceano et al., 2009). As low as these percentages may seem, considering a migration may consist of millions of aphids, the actual number of viruliferous aphids is quite high.
Management of BYD (Technological and Biological)

Management Techniques

BYD is a difficult disease to manage. As can be seen in the preceding literature review, there exists an array of complex interactions between virus, vector, and host making disease epidemics and outbreaks extremely difficult to track. In recent years several strategies have been implemented to better manage BYDV vectors, and therefore BYD. Some techniques include, but are not limited to, planting insecticide treated seeds, a delay of planting, and foliar insecticide sprays.

Planting insecticide treated seeds is important if the crop is likely to emerge during times of high aphid vector migration; generally earlier in the year (Stewart 2013; Royer et al. 2005; Gourmet et al. 1996). The seed treatment protects the crop for approximately two weeks after emergence (Paulsrud et al. 2001). Since the first few weeks after emergence are the plant’s most susceptible ages to infection of BYDV, insecticide treated seeds can provide a useful barrier against disease onset. However, it is not always a 100% effective management option since the treatment wears off after a few weeks, and often needs further management tactics later in the season (Kennedy and Connery 2012; Stewart 2013). In addition, treated seeds are more expensive than untreated seeds, as using treated seeds generally entails mixing regular seed with costly insecticides. In addition, using seeds treated with insecticides for several years can leave significant concentrations of insecticide behind in the soil, often disrupting the natural soil and aquatic ecosystems (Krupke 2012; Goulson 2013). Insecticide seed treatments are often systemic and also have non-target effects on pollinators and other beneficial arthropods (Goulson 2013). These non-target effects can be alleviated to increase
agricultural sustainability by only using seed treatment when other measures are likely to be inefficient.

Optimizing planting date can be important in managing aphid vectors of BYDV. Planting later in the year decreases BYDV infections (Kelley 2001; Miller et al. 1991; Irwin & Thresh 1990) because aphid migration coincides with winter-cereal emergence dates (McGrath & Bale 1990; Coceano et al. 2009). However, there is a trade-off between late planting to decrease yield loss due to BYD and winter kill. Wheat that is planted too late in the season may experience much higher rates of winter kill due to immaturity when entering the cold season (Knapp & Knapp 1978; Fowler 1982). Insecticide treated seeds can be used to increase the optimal planting date range for managing this trade-off (Stewart 2013, Gourmet et al. 1996). However, after planting, there is still the option to spray a foliar insecticide to manage the aphid vectors. In fact, aphids require time to acquire and inoculate hosts with BYDV, making insecticide a valid management strategy to decrease infection rate.

BYD generally occurs in epidemic, or outbreak, fashion rather than being consistently present every year in the same location, but predicting outbreak years is difficult, so it is a common practice to spray prophylactically rather than as needed. An economic threshold for BYDV has been estimated at 15 aphids per one foot row of plants (during early post-emergence; Herbert et al., 1999). Necessity of spraying insecticide is also highly dependent on the previous two management decisions (using insecticide treated seeds or delay planting date). A method to track optimal management options based on previous management decisions and pest conditions would be an effective way to decrease unnecessary and expensive management tactics.
There are several available published models that can track aphid populations and disease status to optimize management timing. The statistical model built by Thackray et al., 2009 is very complex model and incorporates many environmental parameters, including soil moisture and drainage, temperature, rainfall, and evaporation among others, to calculate aphid population development. Population development is then used to calculate disease transmission via migration and in-field transmission. Potential yield loss can also be calculated by this statistical model with a 65% calculated $R^2$. It may, however, only be applicable in Mediterranean climates, such as Western Australia.

A model developed in France (Fabre et al., 2006) represents a statistical model available for more temperate climates. This model describes the population dynamics of $R\ padi$, and therefore the potential dynamics of the virus itself. It simply uses an early season aphid count and temperature to model the spatial and temporal spread of the aphids with high accuracy.

These models have the potential to enhance data input to a decision-support system. Available weather engines have the ability to take these models reported in literature and measure the parameters to give a forecast of aphid movement and virus transmission. These are just two examples of statistical models that can be utilized, though several others exist (McElhany et al. 1995; Leclercq-Le Quillec et al. 2000; Kendall et al. 1992).
Expert Decision-Support Systems

Decision-support systems (DSS’s) are computerized methods of taking unstructured data from a user to allow a broader analysis of the impacts of his or her actions (Turban 1993; Cox 1996). Expert DSSs utilize expert knowledge in a field to gauge the impacts of actions. They are an effective method of using expert knowledge to track and provide suggestions on important management decisions a grower makes. Expert DSSs can incorporate statistical model output as well as expert heuristics. There are four components necessary to make an expert DSS functional: databases, knowledge bases, inference mechanisms, and user interfaces (Travis and Latin 1991).

The database of an expert DSS contains the required information for constructing a logical plan of action. It contains factual information to act as somewhat of a “backbone” to the decision process (Zili & Quixin 1989). In agriculture it may contain information such as banned pesticides, re-entry periods, market values, and planting date ranges. For example, if a certain pesticide is banned in an area, the DSS will not be able to suggest it as a management option in the banned region.

The knowledge base provides a model of expert knowledge. It first requires acquisition of knowledge on management practices and may come from many sources such as interviews, scientific literature, simulations, and data analysis (Cullen & Bryman 1988). The knowledge must then be modeled in some manner to give management outputs based on certain conditions.

The inference mechanism uses the knowledge base and available information on conditions to output an optimal management action (Travis & Latin 1991; Zili & Quixin 1989). Not all information on conditions is always available, so it is important to have a
mechanism that can interpret what is available, or be able to estimate missing information. The inference mechanism can, over time, evolve with feedback from the users (Zili & Quixin 1989).

The user interface allows two-way communication between the user and the DSS (Travis & Latin 1991). The user can input his or her information and receive feedback regarding management options. Then, the user can inform the system on the management practices used and the success or failure of each one. A user interface is usually a web-based platform, which have been improving with the evolution of technology.

**Recent Developments in DSS Technology**

Decision-support systems in conjunction with web-based platforms for tracking pest and disease presence are an increasingly used component of precision agriculture (PA). The utility of DSS programs has been shown through a variety of systems. Fabre et al. (2003) synthesized a model predicting BYD outbreaks to aid in planning foliar insecticide sprays. They showed their model reduced BYD control input costs by up to 36%. As described above, Thackray et al. (2009) published a model that could predict yield loss from BYD outbreaks with an $R^2$ of 65%. This may not seem like a high $R^2$, but considering the complexity of the disease cycle, it is one of the better models for BYD prediction.

One of the most successful DSS platforms is the Integrated Pest Management Pest Information Platform for Extension and Education (soybean PIPE). This platform was designed for the soybean rust invasion of 2004 (Isard et al. 2006). It was, and is,
responsible for up to $299 million in annual fungicide savings since 2005, the year it was implemented (Roberts et al. 2006; Hershman et al. 2011). The PIPE arose from a dire need of an early warning system for a disease devastating soybean crops as it moved from its origins in China through Africa and South America. The appearance of the disease in the U.S. in 2004 prompted the development and deployment of the soybean PIPE, and it was quickly adopted and still remains a widely accepted tool for soybean rust, and now other diseases and crops (Bradley et al. 2010). The Integrated Pest Management PIPE (ipmPIPE) is the predecessor of the current Integrated Pest Information Platform for Extension and Education (iPIPE). iPIPE is maintained and publicized by industry, making it a sustainable business plan to disseminate useful information on management decisions (Personal communication with Joe Russo).

This Master’s thesis provides the description of the construction of a DSS for BYD, a complex disease that is difficult to efficiently manage. The DSS will be integrated into the iPIPE platform. The DSS reported is novel in that it is extremely versatile and powerful in projecting management decisions and associating them with potential outcomes. Collaboration between the public and private sectors and the integration into the iPIPE platform will allow for an evolution of the system with the input of a widely untapped data source of individual grower information. Chapter Two describes the knowledge base, all permutations of field observations, statistical model outputs and historical data, organized using dependency networks, into a DSS that is accessible to wheat producers. Chapter Three offers a method to estimate missing input information, so DSS queries can be completed if the user is unable to provide answers to
DSS questions. Chapter Four proposes future work to enhance the DSS and describe how it will fit in the iPIPE.
Chapter 2

Modeling the decision process for BYD management

Introduction

Barley yellow dwarf (BYD), a disease of cereals, has caused major losses in grain yields since the late 1800’s (Webster and Phillips 1912; Manns 1907). Recent advancements in management tactics can decrease virus prevalence in a field. Several such tactics include insecticide treated seeds, altering the planting date, and foliar insecticide spray treatments (Stewart 2013; Kelley 2001; Miller et al. 1991). However, the effectiveness of these tactics can be enhanced if they are linked with interactive and personalized computer-based decision support systems (DSS’s).

Decision-support systems are computerized methods of taking unstructured data from a user to allow a broader analysis of the impacts of his or her actions (Turban 1993; Cox 1996). DSS’s have been used for crop management since the 1980’s (El-Azhary et al. 2000). Expert systems are a type of DSS in which the logic of a human expert is modeled to recommend actions by the user given certain conditions (Travis and Latin 1991). Travis and Latin (1991) list four necessary components of an expert system: a knowledge base, an inference mechanism, a database, and a user interface.

The knowledge base first requires acquisition of knowledge on management practices and may come from many sources such as interviews, scientific literature,
simulations, and data analysis (Cullen and Bryman 1988). This information is compiled in a manner that can be written in a computer language and connects conditions to actions via dependency networks. Dependency networks are pictorial representations of the logical links among observable or predicted situations and management recommendations based on statistical model output and expert heuristic knowledge. They can be read by the computer as IF condition, THEN action statements (Travis and Latin 1991). The action represents the management decision that would be suggested by the experts, the information for which is obtained during the knowledge acquisition phase.

The inference mechanism interprets the knowledge base and searches a database to output a recommended action (Zili and Qiuxin 1989). It can also be improved to use an incomplete set of conditions to give an estimated action, which is useful as most users will not possess knowledge on all conditions (Travis and Latin 1991). The inference algorithm will search the knowledge base for likely outcomes given the limited input. The database will include storage of factual information, including conditions and action restrictions (Zili and Qiuxin 1989). Finally, the user interface, such as a smartphone or website, allows communication between the user and the system (Travis and Latin 1991).

In the field of agriculture and plant pathology, expert systems are not a recent tool (McKinion and Lemmon 1985; Travis and Latin 1991; El-Azhary et al. 2000). There have been many constructed and adopted systems, such as the Penn State Apple Orchard Consultant (Travis et al. 1992). This system was employed and showed considerable interest by growers, who also adopted changes in production practices conducive of integrated pest management (IPM) strategies (Rajotte et al. 1992; Travis et al. 1992). This system and others were developed before the general use of web-based technology in
agriculture (For example: Lemmon 1986; Travis et al. 1992; El-Azhary et al. 2000). In addition, early expert systems required the use of personal computers, which made the expert system inconvenient. With advent of smartphones and the ubiquity of internet access, this inconvenience is disappearing.

Advancing technology such as high-resolution weather forecasting, internet mapping resources and smartphones allow more sophisticated DSS applications. A recent example is the industry sponsored US Department of Agriculture Soybean Rust Information System website. The system was developed for the soybean rust invasion in 2004 and allowed tracking of rust incidence across the country. Certified crop advisors (CCA), extension specialists and other trained personnel could input rust surveillance data from fields. Meteorological/aerobiological models were used to predict the spread of the rust spores from Mexico and the southern U.S. The system helped growers determine necessity and timing of fungicide applications. The availability of this system across the US soybean belt resulted in up to a $299 million benefit during 2005 mainly because the predictive system information gave growers the confidence to eliminate fungicide applications where they were not needed (Roberts et al. 2006). CCA’s and Extension personnel continue to use this program as a useful tool in managing soybean rust (Bradley et al. 2010). The system developed later into the national Integrated Pest Management, Pest Information Platform for Extension and Education (ipmPIPE), which tracked many more crop pests and diseases.

Pest and disease data and forecasts are displayed by the ipmPIPE, however, management strategies are mediated by a human expert. Expert systems can substitute for most expert mediation.
BYD disease is caused by the world’s most economically damaging cereal grain viruses (Lister and Ranieri 1995). These viruses are transmitted by several aphid species and can infect over 150 species in *Poeaceae* (Barnhart 1895; Burnett and D’arcy 1995). BYD has historically been known to cause major problems in epidemic years (Webster and Philips 1912; Oswald and Houston 1953). The disease can cause an average of 11 to 33% yield loss in winter wheat (*Triticum aestivum* (Linnaeus 1758)) in areas prone to infection and potentially over 80% yield loss (Miller and Raschovà 1997; Lister and Ranieri 1995; Pike 1990). Considering wheat is in the top three most economically important food crops in the world, even a small percentage loss in global yield can be substantial (FAOSTAT 2012; Goldschein 2011).

Five viruses causing BYD cause the majority of damage. The viruses are of the family *Luteoviridae*, but are split between the genera *Luteovirus* and *Polerovirus*. The species are *Barley yellow dwarf virus* BYDV-PAV, BYDV-MAV, BYDV-SGV, *Cereal yellow dwarf virus* CYDV-RPV, and CYDV-RMV. These species are transmitted persistently by the aphids (Order: Hemiptera, Family: Aphididae) *Rhopalosiphum padi* (Linnaeus 1758) and *Sitobion avenae* (Fabricius 1775), *R. padi*, *S. avenae*, *R. maidis* (Fitch 1856), and *Schizaphis graminum* (Eastop 1961), respectively (Rochow 1969). *Rhopalosiphum padi*, *S. avenae*, and *R. maidis* are commonly known as the bird-cherry oat, English grain, and corn leaf aphids, respectively. They each have distinctive migratory patterns which often coincide, in part, with wheat growing seasons (Coceano et al. 2009).

There are many environmental conditions, varying regionally, that determine the spread and damage of the disease, which in turn affect management practices. Aphid
vector migrations are highly dependent upon temperature, moisture, wind fields, and size of aphid populations (De Barro 1992; Thackray et al. 2009). These variables are important in predicting migration timing and magnitude. Virus replication and movement within the host plant is also temperature dependent, with an optimal temperature of 25°C and development of symptoms decreasing with variance from this value (De Wolfe 2002). Environmental constraints affecting transmission success of BYDV by aphids include temperature, stochastic weather events, susceptibility of host plant, use of insecticides preventing aphid population growth in crops, and virus titer combined with age of infected plants (Lowles et al. 1996; Power et al. 1991; Jones 1979). These variables can also determine proportion of aphid vector migration that carries the virus, which can be greater than 10% (Coceano et al. 2009; Plumb 1976). Since these variables all likely interact, a model is needed to specify management actions for all combinations of variables.

Barley yellow dwarf is an ideal disease to test the complexity of an insect-vectored pathogen using an expert system based on a PIPE infrastructure. The system presented here will be referred to as the BYD-DSS. BYD is intensely studied, and thus research results as well as expert opinions can be derived from literature, predictive mathematical models, and from human experts. The BYD-DSS can also accommodate information from statistical/mathematical models, weather forecasts, and other databases containing pesticide and crop phenology. Using the BYD-DSS as an inference mechanism of the knowledge base, the suggested management decisions for a given field can be interpreted, which will be given as recommendations to the growers via the user interface.
As mentioned above, some inputs to the BYD-DSS can be outputs of simulation, statistical, and other types of models. There are several simulation models that can be used to predict aphid and virus changes as well as suggestions of management options for aphid vectors of BYDV. Several published models can be integrated into the BYD-DSS and include, but are not limited to, the Home Grown Cereals Authority (HGCA) BYD management guide which uses plant age to determine susceptibility to BYDV (HGCA 2004), an aphid migration/BYD epidemic predictor designed for Australia (Thackray et al. 2009), a BYD spread model designed for Britain (Kendall et al. 1992), and an aphid/BYD spread model designed for France (Fabre et al. 2006). These can all be used to enhance the accuracy of this expert system in all locations around the globe.

Other resources that can be used to enhance the system will include a weather engine similar to Skybit (ZedX, Inc & Meso, Inc. 1998). This weather forecasting system will be useful in tracking development and migrations of aphid vectors and can also utilize the models mentioned above to determine migration severity. This weather engine predicts events at 1 km resolution. The BYD-DSS can also interact with remote pest forecasting systems via the internet and smartphones, not only to deliver recommendations to the user, but also use user-updated local information to modify pest predictions in real time.

The BYD-DSS offers decision support at several times during the season (Fig 1). The decisions addressing prophylactic measures, include insecticide seed treatment (Stewart 2013), altering planting date to avoid aphid migrations (Miller et al. 1991), scouting for aphids, and use a foliar insecticide. Insecticide treated seeds should be planted to increase yield if winter wheat is sown before USDA-recommended dates.
and/or if the risk of BYDV transmission is high (Stewart 2013; Royer et al. 2005). Planting treated seeds during USDA-suggested planting dates or after risk of BYDV transmission shows little to no increase in yield compared to non-treated seeds (Stewart 2013; Royer et al. 2005). Early planting of winter wheat seed may be desirable to minimize winter kill of seedlings in regions that experience harsh winters, in which case planting treated seeds becomes necessary (Stewart 2013; Fowler 1982; Knapp and Knapp 1978). However, concerning BYD management, history of disease in a location is most important in determining benefits of planting treated seeds.

Planting date is arguably the most important decision in BYD management. There are many consequences of early planting, such as more severe aphid infestations, Hessian fly infestations, and army worm infestations (Stewart 2013). On the other hand, a major consequence of planting too late is higher percentage of winter kill (Knapp and Knapp 1978). Therefore, it is necessary to find an optimal planting date to reduce these risks. Fall migration of alate BYDV vectors is driven by aphid crowding, which is in turn driven by optimal meteorological and environmental conditions for development (De Barro 1992). Once each vector species reaches its critical point of overcrowding, alatae are produced (De Barro 1992). These alatae then migrate to a new host. Each aphid species has its own respective migratory phenology, which is relatively predictable using calendar date (Coceano et al. 2009). Migrations of these vector aphids often coincide with winter wheat phonological susceptibility to infection. Winter wheat is most susceptible and damaged by BYDV when inoculated during its seedling stages, specifically before Growth Stage 31 (GS 31), which is the beginning of stem elongation (Zadoks et al. 1974; Smith and Sward 1982; Oswald and Houston 1953). When
inoculated after the tillering phase, during stem elongation, yield losses are dramatically
decreased (Smith and Sward 1982).

Scouting, or walking fields to determine pest presence and population changes, is
another potential practice to manage aphid vectors of BYDV that can be done after
planting. Hiring a CCA to scout fields has associated costs and can be more useful with
optimal timing. Some growers may wish to do the scouting themselves to save on costs.
Even if a grower chooses to scout his/herself this practice is time consuming and to
obtain a more accurate knowledge of the risk of BYDV infection it requires the ability to
identify specific aphid vector species and knowledge of BYDV strains most common in
the area. Pest and disease forecast models can be used as an alternative or sentinel to
scouting. In reality, there is no benefit to scouting if the pest or disease is known not to
affect a certain area. Also, there would be little need to scout for BYDV vectors after
tillering, since wheat susceptibility to the disease decreases after this stage (Smith and
Sward 1982). These conditions and others are important to consider when a grower is
deciding whether or not scouting is necessary.

While scouting, growers or CCA’s should look for densities of BYDV vectors, and
potentially BYD symptoms to manage the disease. The literature suggests an
economic threshold of 15 aphid vectors per 1 ft. row in fall, but significant damage can
still occur even if aphids do not reach this density (Herbert et al. 1999). These aphids are
generally dispersed in an aggregated pattern (in patches) due to individual “founder”
females, so it is important to scout a large portion of a field to better estimate aphid
populations. The remedial management option if there is a high population of aphids is a
foliar insecticide spray (Herbert et al. 1999) to be considered when seedlings are 0-4 weeks old and if seedlings emerged from non-treated seeds.

All of these tactics are designed to prevent a crop’s exposure to aphid vectors of BYDV during its most susceptible stage. Optimal timing of these treatments can be determined by using environmental conditions, and can be modeled by dependency networks. This paper describes the use of dependency networks in modeling BYD management. Inputs to the model are: environmental variables including temperature, moisture, and wind speed and direction; pest assessments including aphid trap data, pest history, and disease history; crop production practices including planting date, scouting and plant growth. Some of these data are derived from predictive mathematical models, and other factors are heuristic, derived from human experts.

The knowledge base that underpins the BYD-DSS and is depicted by the dependency networks will be integrated into a web/smart phone app. To our knowledge no dynamic, projective, and location-specific DSS’s have been developed using a website and smart phone app created to address a complex insect-vectored viral disease cycle.

**Methods**

**Dependency Networks**

Dependency networks were constructed using design networks LibreOffice Draw (The Document Foundation, open source) on the Ubuntu operating system. The inputs to
the networks were listed at the bottom of the networks (referred to as conditions). Input condition values were classified into ordinal categories (e.g. high, medium and low) called variables. The range of values within a variable is defined. Variable range can come from forecast models, direct user input, or other dependency networks. Combinations of variables from the inputs, denoted by Boolean operators produced an output for a particular dependency. Outputs were placed at the top of the network. An output could be the input to another network or present a recommendation to the user. A path describes a single combination of a variable from each condition, operators, and outputs. All possible combinations of variables for the decisions were connected to satisfy all possible paths, using different colored arrows.

**Decision Framework**

The overall framework for selecting and linking dependencies is shown in Figure 2-1, and each element of the timeline is described in Table 2-1. First, appropriate dependencies were assigned to a portion of the winter wheat growing season (Figure 2-1). Next, input condition information was identified for each decision (Table 2-1). Finally, the relationships among input conditions and the resulting decision options were connected by operators and displayed in dependency networks.

![Figure 2-1](image)

**Figure 2-1.** Timeline of BYD decision-making with winter wheat and aphid events overlaid. This timeline is designed for the Pennsylvania region, thus other regions may have slightly different dates, but similar chronology.
### Treated Seeds: PA July 30th – Sep 15th

- Aphid migration prediction for the coming season (Coceano et al. 2009)
- Crop damage from BYD in the previous season: Immediate history of the field
- Aphid/BYD problem in previous years: History of the field (Personal communication with Piero Caciagli, Ed Rajotte, and Joe Russo)
  - Average aphid counts in previous years
  - Average BYD prevalence in previous years
  - Average yield loss due to BYD in previous years

### Planting Date: PA Sep 15th – planting (USDA 2010)

- Extent of overlap between crop susceptible phase and aphid migration
  - Current weather conditions for aphid development (Personal communication with Piero Caciagli)
  - Prediction of migration peak stats based on suction trap observations (Yearly suction trap data from Piero Caciagli published in Coceano et al. 2009; Appendix C)
  - Expected time to emergence if seeds sown immediately (Plant growth model from ZedX, Inc.)
- Likelihood of having an aphid problem this season
  - Predicted magnitude of migration
- **Aphid/BYD problem in previous years: History of the field**  
  (Personal communication with Piero Caciagli, Ed Rajotte, and Joe Russo)

  - Average aphid counts in previous years
  - Average BYD prevalence in previous years
  - Average yield loss due to BYD in previous years

  - Use of treated seeds, resistant variety, GMO (if developed in the future), etc…(Stewart 2013; Royer et al. 2005; Gourmet et al. 1996)

  - Ideal planting date for winter wheat in the grower’s region compared to the current date (USDA 2010; Stewart 2013; Fowler 1982; Knapp and Knapp 1978)

- **Scouting: Between planting and wheat GS 32/stem extension**

  - Current plant susceptibility to infection of BYDV

    - Current age of plant (Personal communication with Fred Gildow; Smith and Sward 1982; Oswald and Houston 1953)

    - Seed treatment or previous spray in effect (Stewart 2013; personal communication with John Tooker)

  - Current aphid/BYD risk

    - Aphid counts in the region

    - Changes in number of migrant alatae

    - Risk of aphids being virulent
- % of virulent migrants likely to be present (Coceano et al. 2009; Plumb 1976)

- Aphid/BYD problem in previous years: History of the field (Personal communication with Piero Caciagli, Ed Rajotte, and Joe Russo)
  - Average aphid counts in previous years
  - Average BYD prevalence in previous years
  - Average yield loss due to BYD in previous years

- Grower observations of aphid migration (Personal communication with Joe Russo; gives user ability to calibrate models with observations)

- Aphids seen in grower’s field

- Increase in aphids in the field

- Aphids been seen in nearby fields

- Spraying: Between planting and wheat GS 32/stem extension (user may wish to override decision to scout, in which case scouting decision can be used as spraying decision)
  - Population of alate migrants increasing, decreasing, or the same
  - Projected aphid field count (Personal communication with Joe Russo; aphid development model)
Aphid/BYD problem in previous years: History of the field (Personal communication with Piero Caciagli, Ed Rajotte, and Joe Russo)

- Average aphid counts in previous years
- Average BYD prevalence in previous years
- Average yield loss due to BYD in previous years

Table 2-1. Conditions considered when managing BYD. Controllable management decisions are in bold, whereas uncontrollable environmental conditions used to decide best management practices are in normal font (controllable management actions may be considered an environmental condition if the management action has already been implemented at the time of the current recommendation. For example, treated seeds at time of planting are already determined by the grower’s previous actions). Each subset describes the conditions used in determining the condition above it. Source of reasoning for influence of each variable in decision process are listed after each condition, if applicable.

Results

Overall, considering all possible combinations of growing season period, field history, weather and pest predictions, previous grower decisions, and pest observations, there were nine possible recommendations managed by the BYD-DSS (called: Treated seed, no treated seed, plant immediately, delay planting, scout, do not scout, full spray, ½ spray, no spray). There were 243 total combinations of variables to reach the seed treatment recommendation, 9,720 to reach the planting date recommendation, 62,208 to reach the scouting recommendation, and 216 to reach the spray recommendation. However, there are an unlimited number of possibilities when the time component is introduced. All recommendations are shown in Appendix A.
Figure 2-2 shows how all conditions and management decisions interact. It shows how each input condition interacts to alter the management recommendation, and how each management decision alters future recommendations. Figure 2-3 shows that the dependency networks do not function separately, but require several interactions with other web-based databases and mechanisms to give real-time recommendations. They also need input from models and weather engines to determine variables of pest conditions. It is important to have feedback from the user on management practices conducted since that alters later recommendations. Users are also able to input data from their respective fields when other analyses for inputs are unavailable. Once all conditions are entered, either by the user, models, or automated sensors (e.g. weather forecasting stations); an optimal management recommendation can be generated.
Figure 2-2. General outline of the BYD-DSS system interconnectedness of conditions listed in Table 1. Controllable conditions are in ovals, whereas uncontrollable environmental conditions are in rectangles. Connections are color-coded in chronological order pertaining to the management recommendation they affect. Red connections are calculated to determine necessity of treated seeds first, then green connections to determine planting date, blue connections to determine timing and necessity of scouting, and finally violet to determine timing and necessity of insecticide spray. Dashed lines indicate where a projective development model will need to be run to output the connected condition. Dotted lines indicate there is an incomplete source of data that may need to be filled in by the user using the system.
**Figure 2-3.** Flow chart of entire the BYD-DSS. The BYD-DSS describes the knowledge base of dependency networks reported in this chapter.

**Dependency Networks**

Figure 2-4 shows an example of a dependency network. The Current Plant Susceptibility network estimates a crop’s susceptibility to infection at any given time throughout the season. The Current Plant Susceptibility network estimates a crop’s susceptibility to infection at any given time throughout the season. For this network it is assumed that plants that are greater than 4 weeks old are less susceptible than plants 2 to 4 weeks old, which are in turn less susceptible than plant less than 2 weeks old. Plants that have already reached stem elongation are at very little risk of damage from infection. The other condition in this network, having a seed treatment or previous spray, determines whether a crop may be much less vulnerable to infection since a seed treatment or spray will likely kill most potential aphid vectors before inoculation.
Figure 2-4. Current plant susceptibility dependency network. Using operators the networks can be read in “IF/THEN” format. For example, the orange path can be read as “IF the plant is less than 2 weeks of age AND a seed treatment/previous spray is in effect on the crop, THEN the crop has a medium susceptibility to infection of BYDV.” Secondly, there can be multiple variables combined in a single path as seen with the red path. This can be read as “IF the plant is less than two weeks of age OR 2-4 weeks of age, AND there is no seed treatment/previous spray in effect, THEN the crop has a high susceptibility to infection of BYDV.”

Treated Seeds

The first decision that needs to be addressed for the winter wheat growing season is the decision to buy insecticide treated seeds. The conditions of highest importance in this decision will be the aphid/BYD history of the field and the previous season yield loss due to BYD. Considering that a forecast of aphid movement and virulence for a date possibly several months in the future will likely have a large amount of error, it will be more useful to use the history to estimate a field's risk (Personal communication with
Piero Caciagli, Ed Rajotte, and Joe Russo). Thus, a model that describes a forecast of aphid or BYDV spread in a region will carry lower importance in the decision, but may still be used to roughly predict severity of aphid/BYD problem.

*Aphid/BYD Problem in Previous Years (History of field) (Figure A-1)*

Field history is important in every decision a grower makes throughout the wheat growing season, therefore this network is present in every management decision that is mapped. However, as the time of aphid migration approaches, it will be more effective to use real-time data to determine the dynamics of the aphid migration (Personal communication with Joe Russo). Thus, as the season progresses, history of the field will have less influence on the decisions a grower will make about the aphid threat.

Three conditions that can determine whether a given field has had an aphid/BYD problem in previous years are average aphid vector populations (Herbert et al. 1999), average BYD prevalence in previous years (measured in percentage of field infected), and the average yield loss from BYD in previous years.

There is an easily defined correlation between BYD prevalence and yield loss (Perry et al. 2000; Mckirdy et al. 2002), though it is likely that this correlation varies based on the year and location since there is indication of temperature gradients in which BYD symptoms are able to develop (De Wolfe 2002). Thus, it is important to view each field and region as a separate entity. This means, for example, that because a field may have a high prevalence of BYD, it will not necessarily experience a high yield loss due to other conditions, such as temperature constraining symptom development. To account for this potential gradient, the user will input his or her field’s data of BYD prevalence and
yield loss. The network also takes into account that fields with high BYD prevalence are more prone to epidemic years if conditions become optimal for the virus. In other words, a region may experience high prevalence but low yield loss or vice versa.

The logic behind the aphid count condition in this network is simple. A high average aphid count will be considered to be 15-25 or more aphids per 1 ft. row of plants in autumn (Herbert et al. 1999). There is a correlation between aphid count and BYD prevalence due to variance in efficacy of uptake and transmission of the virus by aphids (Lowles et al. 1996; Power et al. 1991; Jones 1979). As with the prevalence/yield loss interaction, there will likely be a regional gradient of these variables, and therefore a gradient of BYD prevalence (Seabloom et al. 2010). To simplify this interaction the number of aphids in the past will be compared to BYD prevalence to determine risk. For example, a high aphid count and low prevalence indicates that although the aphids are present, transmission efficacy in the given area is low. Again, this network takes into account that a high aphid count, under the right conditions is capable of producing outbreaks even though historically BYD may occur at low frequencies.

Decision to Buy Treated Seeds (Figure A-2)

The three conditions used to aid a grower’s decision on buying treated seeds are whether the farm has had an aphid/BYD problem in previous years, a probable aphid migration prediction for the coming season, and yield loss from BYD in the previous season. These conditions will be determined by the Aphid/BYD Problem in Previous Years network, model output, and user-input data, respectively.
The aphid problem in previous years variable addresses the issue as to whether a grower’s field is generally at risk of aphid and BYD problems. The average is the best estimate of a potential problem while knowing nothing about future environmental variables of conditions. Thus, a field’s average aphid/BYD problem is a good basis for projecting the coming year’s problem. Due to this, this variable has the most influence in this decision.

The probable aphid migration count for the following season is a variable predicting severity of aphid migration that will be determined by a statistical model that can be used early enough in the season that it can give the grower enough time to purchase treated seeds if necessary. This means it will use climate and biofixture (a landmark in a species’ phenology) data obtained at least one month before the next season’s planting dates.

The yield loss from BYD the previous season variable is similar to the variables determining aphid problem in previous years, however, instead of looking at the long-term problem of BYD and aphid data it will look at the immediate past. This means that recent trends in the field will affect the outcome of the decision (Personal communication with small grain growers in Italy).

Planting

The planting date decision occurs after seed purchase and will use the range of potential planting dates of winter wheat (USDA 2010) to balance threat of aphid migration from planting too early and winter kill from planting too late. This decision is seminal as it will be designed to eliminate the necessity of buying treated seeds, scouting,
and spraying insecticides; which are costly to the grower and/or the environment. This decision will also take into account that there are other methods of evading the disease in the field (e.g. spraying insecticide), so it will not be unconditionally dominating.

*Likelihood of Aphid Problem during the Coming Season (Figure A-3)*

The variables considered when determining the likelihood of having an aphid problem in the coming season are whether the field has had an aphid/BYD problem in previous years and an aphid count for the approaching season. The data for these variables are obtained by the Aphid/BYD Problem in Previous Years network, and a predictive statistical model, respectively.

The aphid problem in previous years variable allows the history of the field to play a role in this decision. However, the influence of it is decreased by the model predicting the magnitude of the migration. Considering this model will be run at a time much closer to the actual migration there will be many more variables that will account for variation in the migration (Personal communication with Joe Russo).

*Migration Peak/Susceptible Phase Overlap (Figure A-4)*

Aphid migration timing in comparison with wheat growth is important in vector avoidance. Migration timing can be estimated from climate and weather predictions as well as suction trap data. To determine this, the variables of current weather conditions for alate development, time until end of migration, and expected time to seed germination are considered. The data for these variables can come from weather forecasting engines, which can also forecast parameters of aphid development predictive models and a wheat growth models.
Current weather information can give a rough estimate of the magnitude of the migration five weeks in the future since ideal conditions cause development and crowding of aphids, which causes production of alatae (personal communication with Piero Caciagli; Coceano et al. 2009; De Barro 1992; Appendix C).

Using real-time suction trap data that informs about the number of alatae flying over the trap each day and the species of aphids flying, the approximate number of days until large migrations of alatae can be estimated (Observed during analysis of suction trap data from Italy). There is often a period of time before the autumn migration when suction traps catch few aphids sporadically for two to three weeks before a sustained catch is maintained. This information is important in determining when the main portion of the migration will occur.

Wheat plants are most susceptible to BYDV before the tillering phase (Smith and Sward 1982). Thus, it can be generally assumed that younger plants will be the most susceptible to BYD.

**Delay Planting Decision (Figure A-5)**

The decision of whether to plant immediately or postpone planting uses the conditions of overlap of migration and crop susceptible phase, likelihood of aphid/BYD problem during the coming season, whether the grower has resistant crop (i.e. treated seeds, resistant/tolerant wheat varieties, GMO’s [if developed in the future], etc…), and the ideal planting date for the region. These data are obtained from Migration Peak/Susceptible Phase Overlap network, Likelihood of Aphid/BYD Problem during the Coming Season network, user input, and USDA planting dates, respectively.
The migration peak/susceptible phase overlap condition has a very strong correlation with the amount of time from the ideal planting date for the region. The migration peak/susceptible phase condition will influence the decision towards delay planting if the migration will occur before the wheat emerges and the ideal planting date is in the future. However, if the ideal planting date is in the past the decision will be influenced towards planting immediately. On the other hand, if the migration will occur during the crop’s susceptible stages, then the planting date will essentially be pushed later in the season delaying the suggested planting date.

The likelihood of aphid/BYD problem during the coming season variable is most highly correlated with peak/susceptible overlap condition. These two conditions assess the projected risk of the crop to BYDV infection if planted at any given point during the potential planting dates. Thus, a lower likelihood of having a problem will mean peak/susceptible overlap will have less influence on the decision. On the other hand, a higher likelihood of having an aphid problem will mean peak/susceptible overlap will have greater influence.

Having crop resistance to aphids or BYD is a dominating variable in this decision because it will make the crop essentially invulnerable to aphids and BYD for a period of time, thus it would be best to plant before or on the ideal planting date (Stewart 2013; Kennedy and Connery 2012).

If the ideal planting date is greater than three weeks away, then it will always be suggested that the grower delays planting until closer to the date because there may be other consequences from planting too early (McMullen and Ransom 2009; Mississippi State Coordinated Access to the Research and Extension System 2010). Conversely, if
the grower has not planted later than three weeks past the ideal planting date, then it will be suggested that he or she plants immediately to decrease winter kill (Knapp and Knapp 1978).

Scouting

After the grower has planted, it is important to decide when and if to scout for BYD vectors. This decision will rely heavily on real-time and up-to-date information from other growers using the system or established sentinels, such as suction traps (similar to PIPE). A slight influence of the history of the field will be present, but would not be as applicable to the situation as any deviation from the average history will be apparent in real-time data.

*Infectivity Risk* (Figure A-6)

The Infectivity Risk network uses conditions of aphid problem in previous years and the probable percentage of viruliferous aphids in a migration to determine what the chances are of a migratory population carrying the virus. Aphid/BYD problem in previous years provides information on a field’s probable prevalence and yield loss from BYD come harvest. Combining these two variables gives information on whether a location, in general, has the right conditions to support development of the virus, and therefore assesses the risk.

*Current Aphid Risk Status* (Figure A-7)

The Current Aphid Risk Status combines the likelihood of having a viruliferous migratory population and number of aphids that are likely to be in the field. These conditions can be viewed as multiplicative as having a high infectivity risk and high
population count means a large number of infective aphids are probable, but low risk and
low population means a very low number of infective aphids are probable. This network
uses aphids already established in farms nearby, and therefore a likely count for the given
field, and a change in suction trap catches to determine whether the sheer number of
aphids is going to be high enough in a field to cause problems. The number of aphids per
1 ft. row plants in nearby farms uses 5 instead of 15 as a threshold for the “high” number
of aphids because scouting should be conducted before aphid densities reach economic
threshold levels (Herbert et al. 1999). Changes in suction trap catches determines whether
the migration numbers are rising or falling, which combined with number of aphids
already in the field gives a good estimate of whether aphid densities will soon reach
threshold levels (Observed during analysis of suction trap data from Italy).

Infectivity risk, as stated above, will be a good indication of whether the aphid
migratory population will have a chance of transmitting the disease.

Current Plant Susceptibility (Figure 2-4)

The logic behind this network is explained under Figure 2-4 above.

Observational Data of Aphid Population (Figure A-8)

The Observational Data of Aphid Population network is basically giving the user
a chance to enhance the model’s decisions by telling the program what he or she has
actually seen in the field. The network asks the grower whether he or she has actually
seen aphids in the field, whether there has been an increase in the number of aphids in the
field, and whether any neighbors have observed aphids in their fields. This recalibrates
the networks to be more accurate for a more specific location. The outcome of this
network is the determination of whether the grower’s observations are indicative of a migration that is on the rise (incoming), stationary, or not present.

**Decision to Scout (Figure A-9)**

This network is the final step in aiding the farmer’s decision on whether or not to hire a scout for aphids in the field. All the variables in this network are outputs from other networks listed above in this section. It takes current plant susceptibility, current aphid risk status, and observational data from the grower to determine whether a field is at high enough risk to warrant hiring a scout. Fields with higher crop susceptibility will influence the decision towards scouting immediately, while lower susceptibility influences the decision towards delaying scouting to save money. Higher aphid risk statuses influence the decision towards scouting and lower influence towards delaying scouting. The observational data is an enhancement of the aphid risk variable and works in the same way.

**Spraying (Figure A-10)**

A grower should decide to spray his/her winter wheat crop for aphids when the aphid population density reaches a critical level called economic threshold at approximately 15 aphids per 1 ft. row plants (Herbert et al. 1999), though it is determined by weighing the costs of treatment against the cost of losing yield. This network uses information from scouting and forecasting to determine when a field will likely exceed the economic threshold. The aphid/BYD problem in previous years influences the decision by determining whether the virus will likely cause problems in the location of
the field, since there may be a difference in number of aphids present and ability of virus to develop and cause symptoms in the plant.

Discussion

The BYD-DSS represents a way of capturing what is known about BYD management and making it available to growers in a user-friendly and timely format. The BYD-DSS quickly analyzes which of the thousands of possible BYD situations the grower is facing and gives a recommendation for that situation. Required inputs either come directly from the grower or from remote sources, such as web-based weather forecasts.

The BYD-DSS’s logic is the product of BYD research found in the literature as well as decision-making rules created by human experts after many years of experience. Hence, the BYD-DSS has a ‘personality’ that reflects the knowledge and ideas of the system’s builders. A BYD expert system built by other experts may produce different results.

The dependency networks for this program will show a projected set of best-management practices throughout the winter wheat season for any location from which a user may login. The model will be updated as the season progresses to give more up-to-date information on suggested management practices. This projection will be accomplished by assigning weights and risk values to each controllable and environmental condition and variable for the computer to calculate probabilities. Weights control the importance of an input in a network and values control the importance of
variables. The computer will sum the values to obtain a score that will determine the output of that network. Some data points may not currently be available, such as percentage of virulent aphids in a migration, so a constant, or average, value will be used by the computer in its place. This scoring system is similar to the Peanut Rx system developed by the University of Georgia, University of Florida, and Auburn University (Culbreath et al. 2009). These risk values will be calculated to determine probability of the suggested best course of action providing high, medium, or low yield. Users will be able to see the variables that determine the suggested best course of action and change them if he or she believes it necessary, which will cause the mechanism to recalculate the probabilities of actions and yields. This is important because there are often scenarios in which growers are unable to access his or her fields due to non-environmental factors. This is accounted for by allowing the user to manipulate decisions to go down a different path. The models will then be adjusted to report on best management decisions for factors other than yield, such as economic gain, sustainability, and others.

The BYD-DSS can evolve. As more research is done, or as new environmental monitoring technology is developed, the expert system can be modified to accommodate the changes, becoming more accurate and useful.

The decision framework mapped with these dependency networks will later be adapted to be integrated into a website and smart phone app to disseminate real-time decision-support to growers around the world. This program will be similar to others such as Pennsylvania’s PIPE system (Isard et al. 2006), but will be the first of its kind to dynamically model a complex pathogen, vector, host system.
The BYD-DSS has not been validated by testing its effectiveness in wheat fields in real time. This should be done before general release.

There are many other factors determining a crop’s success other than BYD management decisions. Treated seeds may increase slug damage (personal communication with John Tooker). Planting date may affect both BYDV and Fusarium head blight conversely (Kelley 2001). Scouting for many other pests and diseases is almost always mandatory at some point in the season. To address these issues more expert systems addressing any winter wheat disease should be created to make an overall winter wheat growing recommendation guideline. With more expert systems such as this one for other pests and diseases, a computer can combine and optimize all possible conditions for all pests and diseases and output a set of recommendations. Some management recommendations may not affect other pests or diseases. However, when recommendations do interact between multiple pests and diseases the computer can optimize the recommendation. For example, treated seeds may decrease BYD and other aphid transmitted viruses and increase slug damage. If this were to be the case, then the computer could search multiple networks to determine the most optimal seed treatment decision.

Many fields may have conflicting data with surrounding areas. Thus, it may be useful to look at moving averages of conditions over a region to give the best management recommendations. For example, a field that is rotated with other crops or has never grown winter wheat before may not have a history of aphids or disease. In this case, the history of the surrounding area will provide sufficient information to give an optimal recommendation.
Any new information technology must face the adoption process. Human factors and attitudes must be addressed as the BYD-DSS is introduced to the grower community. Grower attitudes, relationships with the developers, availability of technical training and support will all play a role in successful adoption. These access conditions are described in Rajotte and Bowser (1991). Increasing access to web-based information will also allow quicker adoption of systems similar to this one in the future. Growers’ attitudes and desires towards smartphone technology and this app specifically were evaluated by Babbie (2014). The app will be tailored to growers’ desires to benefit their operations in the field and ultimately increase yield, decrease insecticide use, increase sustainability and overall provide a template for a system that can create a more efficient agricultural environment.

Many areas do not see a high yield loss from BYD in most years. It is quite possible that much of the damage from BYD is misdiagnosed as nutrient deficiencies (Mckirdy and Jones 1993) or that current cultivars exhibit fewer symptoms of the disease, but still experience yield loss (Personal communication with Fred Gildow). Ability to detect virus conditions used in these dependency networks will be greatly enhanced in the near future with the development and improvement of unmanned drones using infrared to detect pathogens in individual crops (Jones 2013). This technology could also be easily adapted to obtain information on aphid vector populations by distinguishing species by cornicle size, which is currently a highly labor-intensive task (Personal communication with Piero Caciagli). These monitors will help eliminate any human error present in detection of variables necessary in managing BYD. Drone-based inputs can easily be used in the BYD-DSS.
The use of dependency networks to model management recommendations for an expert system is a simple method of accomplishing this goal. Using these networks it is easy to break the management process down to see the impact of each condition. It allows the creator to combine a very complex set of conditions into a few easily understood figures. It is also possible to follow a single path of conditions using the dependency networks alone. However, with this task being accomplished by a computer, the process will be made user-friendly. When expert systems are created in this way for other pests and diseases a general and optimal winter wheat management plan can be inferred and will greatly increase the efficiency of winter wheat production.
Chapter 3

A proposed mechanism for estimating missing information in the BYD-DSS knowledge base for the barley yellow dwarf decision-support system

Introduction

‘Precision agriculture’ may be defined as a “method of automating site-specific management using information technology” (Bongiovanni & Lowenberg-Deboer 2004). In recent years precision agriculture has been adopted in agricultural programs to greatly increase food security by increasing yields and reducing labor (Gebbers and Adamchuk 2010). Data integration tools that can be used in precision agriculture are decision-support systems (DSS’s) (Zhang et al. 2002). DSS’s are computerized methods of taking unstructured data from a user to allow a broader analysis of the impacts of his or her actions (Turban 1993; Cox 1996). Expert systems are a type of DSS that use expert knowledge to determine a logical management plan. Travis and Latin (1991) list the following components of an expert system: a knowledge base, an inference mechanism, a database, and a user interface.

In chapter 2 the knowledge base (information sources to build dependency networks) and deterministic inference mechanism (dependency networks) of a barley yellow dwarf (BYD) decision support system was proposed and detailed for the management of the disease in winter wheat. This DSS is hereby referred to as the BYD-
DSS. The BYD-DSS knowledge base and inference mechanism are represented by dependency networks that require that all input information be supplied prior to a recommendation. In reality this may not always be possible, so a way of estimating missing information is necessary. While a user may guess at missing information, it would be preferred to have a more quantified, logical method.

In this chapter a model is proposed to integrate the knowledge base contained in the BYD-DSS using a novel inference mechanism to optimize management decisions for unknown information.

The novel inference mechanism consists of a philosophical quantitative scheme that finds an acceptable cumulative correlation among inputs, operators, and outputs. These cumulative correlations were chosen so as to remain true to the heuristic and expert knowledge modeled in the dependency networks from chapter 2. To validate this mechanism, field trials should be conducted to gain better understanding of the interactions between the input conditions. Currently, the numbers are based on the thesis author’s understanding of the interactions. The BYD-DSS and new inference mechanism will use the Barley yellow dwarf virus/disease complex as a model to provide a template platform that can be adapted to more effectively manage crop pests and diseases that follow a similar dynamic.

This expert system is likely to mature and improve in two ways. First, it may gain access to web-based databases that are accumulations of the experiences of other growers’ BYD experiences. Secondly, the grower’s direct observations can be replaced by information from remote sensing, robotic monitoring systems, and more sophisticated
mathematical models. As the system evolves in these ways elements of this proposed inference mechanism may be changed or eliminated.

**Decision-support system (DSS) components**

**Knowledge base**

The knowledge base is an accumulation of information about management practices that can come from interviews of experts, scientific literature, simulations, and data analysis (Cullen & Bryman, 1988). This information may be interconnected and expressed with the use of dependency networks, which can be seen as pictorial representations of the relationship between the information that is contained in the knowledge base (Travis & Latin 1991). The networks can be read by a computer as IF condition, THEN action statements (Travis & Latin 1991). Each action represents the management decision that would be suggested by experts to address a particular situation given specific conditions.

**Inference mechanism**

The inference mechanism is responsible for making calculations and interpretations of the knowledge base. It can be viewed as a type of artificial intelligence (Travis and Latin 1991). Users often do not possess all information that can be utilized in making a decision, so an inference mechanism capable of filling in the gaps and extrapolating results is desirable (Travis and Latin 1991). This is extremely useful in agriculture because often times some information on conditions necessary to make a
decision may not be readily available. The inference mechanism adjusts for this by assuming the most likely scenario based on other information given.

**Database and User Interface**

The database consists of factual information, such as rules and regulations that cannot be manipulated, directly or indirectly, by the user (Zili and Qiu 1989). This database is queried by the inference mechanism to ensure that management tactics conform to the relationships stored in the database.

The user interface allows two-way communication between the DSS and the user (Travis and Latin 1991). It can be a stand-alone application on a personal computer, tablet or smartphone, or web-based. First, the user provides responses to DSS queries from the knowledge base about conditions specific to the user’s situation. Then, appropriate outputs of models from the database and knowledge base are interpreted by the inference mechanism to provide a recommendation. This interaction can be done multiple times throughout a management process.

The Integrated Pest Information Platform for Extension and Education (iPIPE) is a program designed and maintained by industry. It has a user interface in the form of a smartphone app and a website. iPIPE is an ideal platform for decision-making. It evolved from a combination of the NCSU APHIS Plant Pest Forecasting System (NAPPFAST) (Magarey et al. 2007) and the Integrated Pest Management Pest Information Platform for Extension and Education (ipmPIPE) platforms. NAPPFAST was a platform designed and developed in 2002 by the North Carolina State University in collaboration with ZedX, Inc. as a general pest and disease tracking and forecasting system (Magarey et al. 2007). ipmPIPE was a platform that was evolving from the USDA Soybean Rust Information
System in 2005 to better manage the diseases including soybean rust (Isard et al. 2006). Both platforms have been proven effective in managing diseases, with the ipmPIPE saving growers up to $299 million in fungicide costs alone in 2005 (Roberts et al. 2006). It has also shown considerable adoption and utility by certified crop advisors (Bradley et al. 2010). This platform will be responsible for incorporating the BYD-DSS once it is complete.

**BYD-DSS**

**Brief explanation of the BYD-DSS**

As reported in chapter 2, the BYD-DSS represents the knowledge base of a DSS for management of barley yellow dwarf (BYD) disease, which is caused by the aphid-vectored barley yellow dwarf virus (BYDV). BYDV infects approximately 150 Poaceae species worldwide and is a devastating disease of barley, rye, wheat, and many other cereal grain crops (D’arcy and Burnett 1995). Being a complex and sporadic disease that is difficult to predict with simple mathematical models and causing an estimated 11-33% yield loss in wheat annually (Lister and Ranieri 1995), BYD is a disease that is in dire need of expert-based precision management.

The BYD-DSS is used to determine the necessity of BYD management tactics given certain conditions. If these conditions are not met, then no management action is necessary. The BYD management practices addressed by the BYD-DSS are the use of insecticide treated seeds, planting date alterations, scouting for virus vectors (aphids), and foliar insecticide spray (See chapter 2 for more details). All of these management tactics are aimed at decreasing aphid vector populations, as vector management is the only
prophylactic measure useful to avoid plant infections. Insecticide treated seeds will protect the crop from aphid vectors during its most susceptible stages, the first several weeks of growth (Smith and Sward 1982; Stewart 2013). However, treated seeds are expensive, harmful to the environment, not always effective, and can lead to insecticide resistance (Stewart 2013). Planting date is an extremely important decision in BYD management and can help alleviate need for insecticides (Thackray et al. 2009). This is because the later the grower plants, the less the probability for seedlings to be exposed to viruliferous aphids during fall migration. The result of scouting for aphids in the winter wheat fields is used to consider if there is a need to spray insecticide. Finally, insecticidal spray on seedlings and young plants has environmental costs and is not always necessary in managing BYD, for instance if the vector and virus are not currently present in an area.

In addition to the literature and knowledge of the thesis author, most of the knowledge contained in the dependency networks of the BYD-DSS currently comes from Dr. Piero Caciagli (CNR, Italy), an expert in barley yellow dwarf epidemiology with over three decades of experience in the field. He has confirmed that the outputs of the networks are decisions an expert would make given the input conditions.

On their own, the BYD-DSS dependency networks are not able to output optimal management recommendations unless all input queries are determined. However, once it is integrated into a web-based platform that has the capability of using an inference mechanism to interpret the management strategies even when most of the input conditions are not available, it will be a much more useful management tool.

In this chapter a scoring mechanism to estimate missing data in the dependency networks is proposed. The numbers seen in the scoring system are based on the thesis
author’s knowledge of interactions of input conditions and output recommendations. In addition to the scoring mechanism a separate calculation called Success is proposed. Success calculates the probable benefit of a management decision using the scoring mechanism and penalizes the grower if the incorrect management decision is made. This “Success” will later be directly correlated with probable yield.

**Methods**

**Scoring Dependency Networks**

To incorporate the BYD-DSS in a system able to output management decisions without knowledge of all input conditions, I added a secondary inference mechanism to the dependency networks. This inference mechanism was loosely based on the concept of the Peanut Rx program developed by the University of Georgia (Culbreath et al. 2009). The novel inference mechanism included a scoring system for environmental/pest conditions that determine necessity of a management action.

A benefit of this novel inference mechanism is that measurements of the accuracy of recommendations given certain input conditions can be calculated. Algorithms to determine efficacy of management tactics directly correlate with expected yield (explained in detail below). This scoring system was based on the broad literature available on BYDV, on the thesis author’s and experts’ opinions. The inference mechanism reported here is a prototype and represents a preliminary mechanism which will be built upon as more access to grower-input data is attained. It is not currently based
on any experimental data, as hard data for this does not currently exist in an analyzable format.

**Inference mechanism**

A method of inferring an output of networks with missing information, built in this chapter with the use of a scoring system. Refer to Figure 3-1 for a visual representation of the layout of this inference mechanism.

**Output**

End product of a network (can be a management recommendation or an input for another network). Seen in Figure 3-1 at the top of the figure. It is determined by the Score of values in relation to the threshold.

**Where:**

**Threshold**

This threshold is a number that is the minimum (and maximum) Score necessary to solidify a path’s output.

\[
Output_1 > Threshold_1 > Output_2 (> Threshold_2 > Output_3)
\]

**Score**

A Score refers to the sum of the values for any condition in a given dependency network. The Score is used in determining the output of a path using thresholds. **Path:** Connection of values of conditions to outputs using operators represented by colored lines in dependency networks.
\[
Score_{path} = \sum_{1}^{n} Value_{(condition)n}
\]

**Values**

Weights and risks were multiplied to obtain a value. The value represents the combined importance of the weight and risk.

\[
Value = Weight \times Risk
\]

**Weights**

Weights refer to an overall importance of each environmental condition in making a management decision based on the thesis author’s opinion and knowledge previously modeled in Chapter 2. The weights for each condition used as a final dependency network output had to sum to 100%. **Condition:** Environmental or pest status information that is considered necessary in determining the output of a network

\[
\sum_{1}^{n} Weight_{(condition)n} = 100\%
\]
**Risks**

Risks refer to relative disease probability of any variable of a condition compared to the other variables of a condition. These risks must also sum to 100. **Variable**: A separation of conditions into categorical or ordinal ranges.

\[
\sum_{i=1}^{n} Risk_{(variable)_{n}, (condition)_x} = 100
\]

To finalize the inference mechanism of the networks, a threshold between outputs had to be reached. Weights and risks were manipulated, within reason, until Scores achieved thresholds determining outputs similar to paths determined by the BYD-DSS networks built in Chapter 2. Eventually risks, weights, and thresholds will be replaced by analysis of accumulated grower data.
The added benefit of this inference mechanism uses the calculated Scores to determine if a management decision suggested by the system was followed by the user (Success) or to determine the extent of a penalty if the suggested decision was not followed (Failure). The Success of a single management option is set to a maximum value of 100. For each management recommendation in the BYD-DSS the Success of each decision executed by the grower can be added to the Success of each previous and
following decision causing a chain of interacting decisions that will add up to a final Success at the end of the fall (as the management recommendations are ones to be conducted during the fall).

The algorithms determining Success vary for each of the three management decisions addressed and are found in Appendix B. Again, these numbers are, based on magnitude of the Scores from the dependency networks, in a range of 0 to 100 and for now they were assigned based on the author’s understanding of the BYD disease management options.

Results

To illustrate the results of the BYD-DSS dependency networks based on the scoring system outlined above, I will use the dependency network called ‘Decision to Buy Treated Seeds’ (Figure 3-2). All other scored networks can be found in Appendix B. The Infectivity Risk network was omitted from this mechanism and replaced by aphid/BYD problem in previous years. This was done because the percentage of virulent migrants during any given year is generally not available.

Threshold

Thresholds determine outputs of networks as follows:

\[
\begin{align*}
\text{Score}_\text{Red} & \geq 50, \text{Buy treated} \\
\text{Score}_\text{Orange} & \geq 50, \text{Buy treated} \\
\text{Score}_\text{Yellow} & \geq 50, \text{Buy treated}
\end{align*}
\]
\[ \text{Score}_{Dark\ blue} < 50, \text{Buy untreated} \]
\[ \text{Score}_{Light\ blue} < 50, \text{Buy untreated} \]
\[ \text{Score}_{Purple} < 50, \text{Buy untreated} \]
\[ \text{Score}_{Black} < 50, \text{Buy untreated} \]
\[ \text{Score}_{Green} < 50, \text{Buy untreated} \]

Scores

The Scores determined for each path are as follows:

- \textit{Red path} = 28 + 14 + (24 or 12) = 66 or 54
- \textit{Orange path} = 12 + 14 + 24 = 50
- \textit{Yellow path} = 28 + (6 or 0) + 24 = 58 or 52
- \textit{Dark blue path} = 28 + 14 + 4 = 46
- \textit{Light blue path} = 12 + (6 or 0) + 24 = 42 or 36
- \textit{Purple path} = 28 + (6 or 0) + (12 or 4) = 46, 38, 40, or 32
- \textit{Black path} = 12 + (6 or 0) + (12 or 4) = 30, 22, 24, or 16
- \textit{Green path} = 12 + 14 + (12 or 4) = 38 or 30

Values

The values determined for each variable are as follows:

- \( Yes = 40\% \times 70 = 28 \)
- \( No = 40\% \times 30 = 12 \)
- \( High = 20\% \times 70 = 14 \)
Medium = 20% * 30 = 6
Low = 20% * 0 = 0
> 30% = 40% * 60 = 24
10 – 29.9% = 40% * 30 = 12
0 – 9.9% = 40% * 10 = 4

Weights

In determining the necessity of treated seeds to reduce risk to infection of BYDV, the long term history and short term history of the field is most important because long-term forecasts of aphid migration and BYDV transmission will be relatively inaccurate. Thus, these are the most important conditions, they carry the most weight (‘Consistent Aphid/BYD Problem in Previous Years’ = 40% & ‘Crop Damage from BYD Previous Season’ = 40%). A model that forecasts magnitude of aphid migration several (potentially) several months in the future is likely to have a large error and should not carry quite as much weight in the decision (‘Aphid Migration Prediction for Following Season’ = 20%).

Risks

The risk of BYDV inoculation and subsequent BYD damage in fields that have had a long-term history of aphid/BYD is much higher than fields that have not. In Figure 3-1, for the ‘Consistent Aphid/BYD Problem in Previous Years’ condition the risk of 70 is assigned to ‘Yes,’ and 30 is assigned to ‘No.’ For the ‘Aphid Migration Prediction for
the Following Season’ condition the risk of 70 is assigned to ‘High,’ 30 is assigned to ‘Medium,’ and 0 is assigned to ‘Low.’ Finally, in the ‘Crop Damage from BYD Previous Season’ condition the risk of 60 is assigned to ‘>30%,’ 30 is assigned to ’10-29.9%,’ and 10 is assigned to ‘0-9.9%.’

Figure 3-2. Scored ‘Decision to Buy Treated Seeds’ dependency network.

Success of Management Decisions

Figure 3-3 is an example of a flow chart the user may be able to access in the DSS for BYD management. Each decision made (Yes) comes with a Success and a Failure. Failure is calculated simply by 100-Success (as described in Table 3-1). Thus, failure represents a penalty for not conducting the suggested management decision. There will
be many potential combinations of management tactics leading to a final, cumulative Success. This final Success will be translatable into a probable yield value. With this Success/probable yield, the input cost and sustainability of management decisions that are used to reach this Success/yield can be calculated and optimized for profit and sustainability (and others later, if deemed important). In terms of yield, the Success of planting treated seeds and conducting full foliar insecticide spray will always be the greatest. However, these tactics are generally not optimal based on input costs and sustainability.

For example, the highest Success may be 300 (Success of treated seeds + planting date + spray). However, another set of management tactics may lead to a Success of 275 (example Success), but the input cost will be lower and thus the profit will be greater, so this set of tactics will be suggested rather than the tactics leading to the Success of 300. Since the calculations of input cost and sustainability are separate calculations, they are not addressed by the scoring system in this chapter, but will be incorporated into the final inference mechanism.

These values will be based on whether or not the correct decision was made at the correct time, which is in turn based on the first level of scoring described above. If a mistake is made executing a management option, then the yield will be lowered.
Figure 3-3. Flow chart of management decision Success (and failure). Each management decision proposed by the system can be followed or not, and the failure to follow an advised decision carries penalties that can be calculated and seen by the grower on a smart phone app or web based interface. The grower will be able to see a decision tree similar to this one, but the computer will optimize the management program by finding the set of practices leading to the highest Success, profit, and sustainability desired by the user (the latter two are not calculated by this tree). As seen from this figure, each management decision affects the others.

_Treated Seeds_ (Table 3-1)

Success of planting untreated seeds varies based on the following equations:

**Equation 1:**

If(Score > Threshold, then Success = (((66 – Score)/20)*100))

**Equation 2:**
If(Score ≤ Threshold, then Success = 100)

Where 66 represents largest Score in the Treated Seeds dependency network, 20 represents the difference between the largest Score (66) and the largest Score in which it would be suggested to buy untreated seeds (46), and 100 represents the maximum possible Success.

If the Score is less than the threshold, then, according to the dependency networks, untreated seeds should be just as successful as treated seeds. Thus, Success of both planting treated and untreated seed would equal 100.

**Example of how success can be calculated using incomplete information**

As an example, assume the only information on conditions for the treated seeds decision is that he or she had over 30% yield loss attributable to BYD in the previous season. Referring to Table 3-1, notice that the average Score knowing only this piece of information is 50.667 (calculated by the equation:

\[
\text{Value}_{\text{Over 30\%}} + \frac{\text{Value}_{\text{High}} + \text{Value}_{\text{Medium}} + \text{Value}_{\text{Low}}}{3} + \frac{\text{Value}_{\text{Yes}} + \text{Value}_{\text{No}}}{2} = 24 + 6.667 + 20 = 50.667
\]

From this information it is assumed that the average Score knowing only over 30% yield loss is 50.667. Using this average Score the average Success can be calculated using the equations above. To begin, the threshold is 50 for this decision, 50.667 > 50, thus equation 1 is used. **Success = ((66-50.667)/20)*100 = 76.667.** This Success indicates that the average Score will lead to a Success of 76.667 for planting untreated seeds (and a Success of 100 for planting treated seeds). There is a range associated with this, but the average is the most likely scenario. Once more information is
obtained and as time gets closer to the management tactic execution the range will decrease due to more accurate forecasts. These calculations can be accomplished even if the grower and forecasts possess absolutely no information on the conditions by assuming average values for all variables of all conditions.

<table>
<thead>
<tr>
<th>Condition (History)</th>
<th>Value</th>
<th>Average Score (History)</th>
<th>Condition (Prediction)</th>
<th>Value</th>
<th>Average Score (Prediction)</th>
<th>Condition (Dam Prev Season)</th>
<th>Value</th>
<th>Average Score (Dam Prev Season)</th>
<th>Score</th>
<th>Success (Treated seeds)</th>
<th>Success (Untreated seeds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>28</td>
<td>48 High</td>
<td>14 47.333333333 Over 30%</td>
<td>24</td>
<td>50.666666667</td>
<td>66.00</td>
<td>100.00</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>28</td>
<td>48 High</td>
<td>14 47.333333333 10-29.9%</td>
<td>12</td>
<td>38.666666667</td>
<td>54.00</td>
<td>100.00</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>28</td>
<td>48 High</td>
<td>14 47.333333333 0-9.9%</td>
<td>4</td>
<td>30.666666667</td>
<td>46.00</td>
<td>100.00</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>28</td>
<td>48 Medium</td>
<td>6 39.333333333 Over 30%</td>
<td>24</td>
<td>50.666666667</td>
<td>58.00</td>
<td>100.00</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>28</td>
<td>48 Medium</td>
<td>6 39.333333333 10-29.9%</td>
<td>12</td>
<td>38.666666667</td>
<td>46.00</td>
<td>100.00</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>28</td>
<td>48 Medium</td>
<td>6 39.333333333 0-9.9%</td>
<td>4</td>
<td>30.666666667</td>
<td>38.00</td>
<td>100.00</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>28</td>
<td>48 Low</td>
<td>0 33.333333333 Over 30%</td>
<td>24</td>
<td>50.666666667</td>
<td>52.00</td>
<td>100.00</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>28</td>
<td>48 Low</td>
<td>0 33.333333333 10-29.9%</td>
<td>12</td>
<td>38.666666667</td>
<td>40.00</td>
<td>100.00</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>28</td>
<td>48 Low</td>
<td>0 33.333333333 0-9.9%</td>
<td>4</td>
<td>30.666666667</td>
<td>32.00</td>
<td>100.00</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>12</td>
<td>32 High</td>
<td>14 47.333333333 Over 30%</td>
<td>24</td>
<td>50.666666667</td>
<td>50.00</td>
<td>100.00</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>12</td>
<td>32 High</td>
<td>14 47.333333333 10-29.9%</td>
<td>12</td>
<td>38.666666667</td>
<td>38.00</td>
<td>100.00</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>12</td>
<td>32 High</td>
<td>14 47.333333333 0-9.9%</td>
<td>4</td>
<td>30.666666667</td>
<td>30.00</td>
<td>100.00</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>12</td>
<td>32 Medium</td>
<td>6 39.333333333 Over 30%</td>
<td>24</td>
<td>50.666666667</td>
<td>42.00</td>
<td>100.00</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>12</td>
<td>32 Medium</td>
<td>6 39.333333333 10-29.9%</td>
<td>12</td>
<td>38.666666667</td>
<td>30.00</td>
<td>100.00</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>12</td>
<td>32 Medium</td>
<td>6 39.333333333 0-9.9%</td>
<td>4</td>
<td>30.666666667</td>
<td>22.00</td>
<td>100.00</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>12</td>
<td>32 Low</td>
<td>0 33.333333333 Over 30%</td>
<td>24</td>
<td>50.666666667</td>
<td>36.00</td>
<td>100.00</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>12</td>
<td>32 Low</td>
<td>0 33.333333333 10-29.9%</td>
<td>12</td>
<td>38.666666667</td>
<td>24.00</td>
<td>100.00</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>12</td>
<td>32 Low</td>
<td>0 33.333333333 0-9.9%</td>
<td>4</td>
<td>30.666666667</td>
<td>16.00</td>
<td>100.00</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3-1. Treated seeds management decision scoring table. This is table 1 of 3 concerning management decision scoring. Others can be seen in Appendix B. Columns 1, 4, and 7 are the conditions and variables present in the Decision to Buy Treated Seeds network. Columns 2, 5, and 8 are the values of each variable. Column 10 is the Score of each combination of variables based on the sum of observed values. Columns 11 and 12 are the Success of treated seeds and untreated seeds, respectively. Columns 3, 6, and 9 are the average Scores given only knowing the variables in columns 1, 4, and 7, respectively (e.g. the average Score given the only data available is that the field has had a history is 48; and for knowing the field has not had a history, the average Success is 32). The equation to calculate the Success of untreated seeds is If Score > 46, then ((66-Score)/20)*100, otherwise 100. 46 is the highest Score that would be suggested to by
untreated seeds. 66 is the highest overall Score. 20 is the difference between 66 and 46. Finally, 100 is the highest possible Success.

**Planting Date (Table B-1)**

Optimal planting date for a region can be calculated by climate data and plant growth models alone, which weather engines can run. However, aphid BYDV vector migration data is necessary to alter this date to avoid BYD damage. In the table and subsequent equations displayed in Appendix B, 5 weeks are addressed as potential planting dates. They may suggest week -2, which corresponds to 2 weeks before the ideal date calculated only by climate and growth model data, to week 3, which corresponds to 3 weeks after the ideal date. Week 0 corresponds to the same week suggested by the plant growth model and climate data.

**Spraying (Table B-2)**

Optimal foliar insecticide spray timing using aphid population data was determined. Success of conducting a full insecticide spray is always 100, because it will always be more successful in terms of yield, than conducting a half (or diluted) spray or no spray (assuming no cost or sustainability barriers since these will be calculated separately).
**Discussion**

This chapter reports a prototype of an inference mechanism that will be integrated together with the BYD-DSS knowledge base in the BYD disease decision support system. The inference mechanism is used to provide the most probable successful decision taking into account only the available information. A projected set of best management decisions can be calculated from the Scores in this inference mechanism to keep the grower within the scope of a general management goal. This scoring system is a very powerful mechanism because it is not necessary to possess information on all environmental/pest conditions to obtain a management decision. If there is missing information on conditions, the inference mechanism can calculate the average ‘Score’ of a network. With this average ‘Score’ the probable benefit of the management recommendation will be calculated as a ‘Success,’ which is directly correlated with yield.

This is a novel use of big data analysis that will be optimized for the inclusion of the BYD-DSS into the Integrated Pest Information Platform for Extension and Education (iPIPE). iPIPE is an already established platform that contains databases on crop pests and diseases management strategies as well as a user interface available via smartphone app and website. It can receive pest and disease information from sensors and give work orders to task controllers. It also allows a feedback between sensors and task controllers; in this case via a user interface available to growers.

iPIPE incorporates environmental, pest, statistical models, and expert advice data into its platform to report accurate and precise conditions for individual fields. The models are constantly updated to give the grower real-time information. The BYD-DSS reported in this thesis will be integrated into the iPIPE platform using the inference
mechanism proposed in this chapter and then made available to any grower in any location at any time around the world. Since this platform is maintained by industry there is a monetary incentive to sustain the program, giving it a major potential to make a difference in winter cereal grain production.

Before its integration into iPIPE the BYD-DSS will be adjusted to output final wheat yield (which will be directly correlated to Success), profitability and sustainability forecasts. Since this inference mechanism is not based on factual data, at this point it is likely to have many illogical assumptions. However, after many years of accumulated grower input data, the mechanism can be rebuilt to more accurately represent optimal management practices. For now, however, the best inference mechanism that can be built is based on expert and the thesis author’s knowledge.

Once this inference mechanism is finalized, it will be integrated into the iPIPE user interface. This user interface will be altered to better address growers’ needs. A study on growers’ adoption of smartphone precision agriculture programs is being investigated by a sociological component of this project (Gruber 2014). In the past, DSS’s have shown high adoption rates along with management improvements (Roberts et al. 2006; Bradley et al. 2010). Sociological information will make it easier to market the program to growers by showing them that it can benefit farming efficiency.

Precision agriculture and DSS development is the direction that agricultural systems are heading (Jones 2013; Magarey et al. 2010; McBratney et al. 2005; Gebbers & Adamchuk 2010). Many potential inventions that are being developed currently will aid this system’s accuracy. Such developments include drone detection of pathogens in
plants, detection of insect populations using drones, and communication between sensors in the field and databases to name a few (Jones 2013; Primicerio et al. 2012).

With the industry backing this program, it will be able to remain active in agricultural system for an extended time and will also be adapted to manage many other crop pests and diseases. High adoption of this platform will save growers money on pesticide costs and increase sustainability and efficiency of agriculture worldwide.
Chapter 4: Discussion

*Barley yellow dwarf virus* (BYDV) is the causal agent of barley yellow dwarf (BYD) disease, which causes 11-33% global yield loss in winter wheat (Lister & Ranieri 1995). Its occurrence is difficult to predict due to the complexity of interactions between virus, vector, and host, and the disease devastates grain growing regions in epidemic years. Many researchers have developed models to predict the occurrence and spread of BYDV and its vectors. Web-based platforms for precision agriculture crop management are generally available for many crop pests and diseases, but they lack a mechanism for analysis of unstructured user data.

In this thesis I integrated the available BYDV literature together with the knowledge of experts in BYDV epidemiology in an expert decision-support system (DSS). The DSS was built to determine BYD management recommendations for winter wheat based on input conditions, such as history of virus and vector presence in a geographical area, farming decisions and weather forecasting data. The DSS developed in this thesis is to be implemented into the Integrated Pest Information Platform for Extension and Education (iPIPE) maintained by industry. This DSS serves as a model of how expert knowledge can be used to efficiently manage complex pest and pathogen systems in a precise field by field manner. Precision management of BYD will help to increase productivity of fields at risk of BYD while decreasing management costs by applying management practices with more logical timing.
**Major accomplishments**

Aside from using existing mathematical models and expert knowledge on BYDV and its vectors, suction trap data of aphid migration was analyzed to better understand the phenology of the disease. Daily data collected for over 20 years in Northern Italy from two locations having a climate similar to Pennsylvania were provided by Dr. Piero Caciagli (CNR Italy). These data were analyzed for correlations between the environment and aphid vector migration, which can potentially be used to track the aphid vectors. As expected from the literature, *R. padi* and *R. maidis* populations migrate twice per year, while *S. avenae* migrates primarily in the spring with a small migration in the fall. Since winter wheat is most susceptible to BYDV early in the growing season, I looked at what environmental factors could be used to predict the beginning of the fall migration for aphids that serves as main vectors of BYDV in temperate regions. As seen in appendix C, migration of *R. padi* usually starts at various degree day (DD) accumulations, but typically centers regional migration around 1500 DD. The earlier the migration starts, in terms of DD, the more severe the migration tends to be. Also, aphids tend not to fly when the wind speeds exceed 15 km/hour. While suction trap data on aphid migrations are also collected in the US, these data were not collected consistently which prevented in depth analysis of the aphid migration, but can be used to roughly estimate aphid migration for the dependency networks created in chapter 2. Although the analyses in Appendix C and other previously published models may give a good indication of aphid vector migration, often wheat fields are located far from suction traps and at the moment we are lacking field data to demonstrate that aphid migrations in Pennsylvania occur exactly as in Italy. Suction trap data from the Midwest United States was analyzed, but it was not as
complete as the data from Italy, so it was only used to observe calendar date phenology of the migrations. Thus, in my work logical assumptions were made to reduce complexity of conditions leading to management recommendations (e.g. aphid/BYD history of a field can convey information on whether environmental conditions in a location generally allow transmission of BYDV and the progression of symptoms in host plants).

In chapter 2, dependency networks were used to connect field conditions to management recommendations. These recommendations included the use of insecticide treated seed, optimal planting date, timing and necessity of scouting for aphid vectors, and timing and necessity of foliar insecticide sprays. These conditions were then separated into ordinal ranges called variables. Dependency networks connected these variables to a management recommendation, with a total of 72,387 combinations of variables to reach management recommendations. This DSS will be hereby referred to as the BYD-DSS.

**Inference Mechanism**

The dependency networks are a type of deterministic inference mechanism that requires knowledge of every condition to give a management recommendation. The development of a novel inference mechanism capable of interpreting missing data, as reported in chapter 3, will make the DSS more useful to more growers. This will be useful because growers, generally, do not possess all information used as inputs for the dependency networks. Currently, most DSS’s give exact management recommendations that do not change over the course of the season. However, the inference mechanism
proposed in chapter 3 allows for a flexible and dynamic management plan that increases in accuracy as the growing season progresses and real-time pest data comes in. Future management recommendations are also altered by the previous management actions determined by grower feedback. Any amount of field condition information can be put into the BYD-DSS networks (even none) for the numerical inference mechanism to give an optimal management recommendation given the available information. The inference mechanism will then calculate the maximum yield a grower can expect after following all the recommendations, or a penalized yield if the grower fails to meet the recommendation given by the system.

Possible Improvements

Eventually, as grower input data accumulate in the databases, heuristic components of the dependency networks will begin to be replaced by statistical models generated using real data. Models will increase the local accuracy of the DSS. The more access to grower-input information there is, the more accurate the system will eventually be. After many years of data accumulation, the models and dependency networks can be reevaluated to give optimal BYD management recommendations. The knowledge base and inference mechanisms in chapters 2 and 3 are the most accurate methods of making management recommendations given the current statistical model availability.

The BYD-DSS addresses management options for only a single disease in winter wheat (and possibly useful for other winter cereals). However, this is not the only disease
or pest that affects winter wheat. Winter wheat is also not the only crop that encounters problems from disease and pests. Management options addressed by the BYD-DSS may not solely affect BYD management. A more encompassing method would be to accumulate multiple DSSs (similar to the one here) for managing all winter wheat pests and diseases. Since the DSS reported here only regards BYD management it ignores impacts one management option may have on another problem in the fields. For example, treated seeds may increase slug damage and late planting may increase *Fusarium* prevalence (Personal communication with John Tooker, Kelley 2001). To deal with these potential conflicts the numerical mechanism and yield calculations from chapter 3 can be used. A computer could take these yield (or eventually sustainability and profit) calculations from multiple pest and/or disease DSSs and optimize the management regimen. From this a winter wheat DSS could be obtained.

The dependency networks for the BYD-DSS are laborious and often difficult to design. It would take a very long time and many man-hours for similar DSSs to be designed for all the pests and disease of winter wheat. It would have positive effects in the long run; however, there may be simpler ways to accumulate a knowledge base and design inference mechanisms. For example, if the iPIPE platform could have extension agents and consultants answer survey questions on how to best manage certain pests and diseases and what environmental factors are important in making these decisions a computer program could design dependency networks theoretically. To accomplish this, the program could recognize key words in the responses and after analyzing many responses connect most commonly mentioned conditions with management options recommended by the experts. This is a rough idea, but if it were to be studied and
expanded on, it would streamline the process of determining best management practices for other pests and diseases in winter wheat (and other crops eventually).

A series of surveys could also be used to improve the BYD-DSS. There are likely many more management options for BYD not addressed by the dependency networks, e.g. fertilizer or crop rotation. The impact of these management practices on BYD could be obtained by more experts submitting their opinions on optimal management of BYD. The dependency networks are simply the start to a more complex accumulation of knowledge of how to better manage the disease.

As indicated above, this system requires feedback in order to improve. Thus, its success will depend on the adoption rate of the final product. Leaving the management of this product to the industry could create a sustainable business model that will perpetuate the project in the absence of federal government funding. The promise of integration of the DSS into the already established Integrated Pest Information Platform for Extension and Education (iPIPE) removes the labor-intensive task of creating a publically available platform as there is already an established one to adopt it. Also, sociological studies are being conducted by a collaborating sociologist to gauge growers’ acceptance and adoption of web-based platforms, such as the iPIPE (Gruber 2014). These studies will help to increase dissemination of the BYD-DSS.
Future Work

Before the BYD-DSS is integrated into iPIPE it must account for the cost and sustainability of management decisions. Since the Success scores reported in chapter 3 are only correlated to yield, the cost and sustainability of the management practices recommended by the system must be calculated and optimized along with yield.

After the DSS has been integrated into iPIPE grower input data will start accumulating. Instruments to better measure input data for the BYD-DSS will be an important improvement to the system. Drones are already being developed for virus detection in crops (Jones, 2013), and could be easily altered to enhance this system. They could be altered to detect aphid populations at low levels in fields using infrared, or possibly plant stress volatiles. Other automated sensors, can be developed as well. If a suction trap were to be designed to automatically count aphid migrants of certain species (possibly by measuring cornicle length), labor for inputs of this system would be decreased significantly. There are many ways of automating the farming process and making it a more efficient business.

BYDV surveys should be done to better understand the species of viruses present in certain areas. This information can help reduce the number of variables necessary in predicting BYD damage in certain areas. For example, knowledge that a location has only one species of virus that is vectored specifically by a species of aphid indicates that other potential BYDV vectors can be ignored in aphid population calculations. Appendix D gives the protocol of a species-specific BYDV test to be used to see if wheat is infected
by BYDV and to identify the BYDV species. Results from these tests can be used in the scouting and spraying recommendations, and also in aphid migration modeling.

One of the main goals of this project is to educate individuals in the field of agriculture about the possibility and utility of sharing field information to better manage crops. This includes not only growers who can directly use systems similar to this one, but also students. Classes should be implemented to educate students studying agriculture on how to use and create DSS’s similar to this one. Education of this technology will aid in the dissemination of information from the DSS and hopefully lead to greater acceptance of it, which will in turn propagate development of DSS’s similar to the BYD-DSS for many other crop diseases.

This project is given the potential to be very successful due to the expansion of public access to online information via database-searching platforms, such as iPIPE, and availability of portable internet-capable devices. Platforms capable of using multiple databases of user input information to gain knowledge on larger scale pest and disease information allow for advance knowledge of crop pest and disease forecasts. Better forecasts of crop pest and disease information allows for a more optimized management plan. Web-based platforms, such as iPIPE, can combine these data from various databases to give accurate and up-to-date information and forecasts on crop pest and disease statuses. Precision agriculture is the link between this increase in information availability and farm management. Real-time disease management platforms delivered by precision agriculture methods are the future tools to help growers make more informed management decisions to decrease input cost while increasing yield and sustainability.
Appendix A

Dependency Networks from Chapter 2

In this appendix are the deterministic dependency networks as described in chapter 2, before the addition of the novel inference mechanism.

Figure A-1
Figure A-2
Figure A.3

Likelihood of Aphid Problem During the Coming Season (For Delay Planting Decision)

- High Likelihood of Having Aphid/BYDV Problem This Season
- Medium Likelihood of Having Aphid/BYDV Problem This Season
- Low Likelihood of Having Aphid/BYDV Problem This Season

Aphid Problem in Previous Years (From Aphid Problem in Previous Years Network)

High
Low

Probable Aphid Count for Coming Migration Measured in Aphids Per Plant (From Migration Model)
Figure A-5
Figure A-7
Figure A-9
Appendix B

Novel Inference Mechanism via Scoring System from Chapter 3

This appendix displays the dependency networks and management Success calculations as described in chapter 3.

Figure B-1
Figure B.2

Decision to Buy Treated Seeds

Buy Treated Seeds

Don't Buy Treated Seeds

If sum ≥ 50, buy treated seeds
If sum < 50, don't buy treated seeds

Consistent Aphid/BYO Problem in Previous Years (From Aphid/BYO Problem in Previous Years Network)
Yes: 38
No: 12
Weight: 30%/28/12

Aphid Migration Prediction for Following Season (From Model)
High: 14
Medium: 8
Low: 0
Weight: 20%/14/50

Crop Damage from BYO Previous Season (Grower Input)
>30%: 24
10-19%: 12
0-9%: 4
Weight: 40%/24/12/0
Figure B-3
Figure B.5

Delay Planting Decision

Delay Planting

When WH Migration Peaks Likely Gourp Compared to Drop Susceptible Phase If Seeds are Planted Immediately (From Peak/Susceptibility Phase Network)

Expected: 20/11/10

Weight: 40% = 0.16

Likelihood of Asexual Problem During the Coming Season (From the Likelihood of Asexual YDV Problem During the Coming Season Network)

Expected: 10/12/10

Weight: 40% = 0.12

Trained Seeds, Artificially Modified or Any Other Form of Resistant Crop (From the Trained Seeds Decision or Direct Input from Grower)

Expected: 3/3/19

Weight: 20% = 0.06
Figure B.6

Current Aphid/BYDV Risk Status

<table>
<thead>
<tr>
<th>High</th>
<th>Medium</th>
<th>Low</th>
</tr>
</thead>
</table>

If sum $\geq 50$, high  
If sum is $31-49$, medium  
If sum $\leq 30$, low

- Average Number Aphid/yr
- Presence or not
- No Cattle
- Change in Aphid
- BYDV Problem in Previous Years

Details:
- Weight: 4/08/12
- Weight: 7/08/10
- Weight: 12/08/10
- Weight: 11/02/10
Figure B-7
Figure B.8

Decision to Spray

- Full Spray
- Half Spray
- No Spray

If sum>48, full spray
If sum is 32-47, ½ spray
If sum<32, no spray

Current Plant Susceptibility
(From Current Plant Susceptibility Network)
7/09/10
Weight: 80% 28/12.0

Current Aphid/PYDV Risk Status
(From Current Aphid/PYDV Risk Status Network)
5/09/10
Weight: 40% 2/1/24

Observational Data of Aphid Population
(From Observational Data of Aphid Population Network)
5/16/10
Weight: 80% 12/8.0
Figure B-9
The equations determining the Success of management scores of planting on certain dates in comparison to knowing no pest data are as follows:

If(Score < 14, then Success of planting week -2 = 100, week -1 = 40, week 0 = 10, week 1 = 0, week 2 = 0, week 3 = 0)

If(14 ≤ Score ≤ 34, then Success of planting week -2 = 40, week -1 = 100, week 0 = 40, week 1 = 10, week 2 = 0, week 3 = 0)

If(35 ≤ Score ≤ 47, then Success of planting week -2 = 10, week -1 = 40, week 0 = 100, week 1 = 40, week 2 = 10, week 3 = 0)

If(48 ≤ Score ≤ 54, then Success of planting week -2 = 0, week -1 = 40, week 0 = 100, week 2 = 40, week 3 = 10)

If(55 ≤ Score ≤ 61, then Success of planting week -2 = 0, week -1 = 0, week 0 = 10, week 1 = 40, week 2 = 100, week 3 = 40)

If (Score > 61, then Success of planting week -2 = 0, week -1 = 0, week 0 = 0, week 1 = 10, week 2 = 40, week 3 = 100)

Table B-1. Planting Date management Success and Scores

<table>
<thead>
<tr>
<th>Condition (Overlap)</th>
<th>Average Score (Overlap)</th>
<th>Condition (Aphid Prob)</th>
<th>Average Score (Aphid Prob)</th>
<th>Condition (Treated Seeds)</th>
<th>Average Score (Treated Seeds)</th>
<th>Score</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before 8 31.33333 High 28 51.3333333 Yes 0 26.66666667 36 10 40 100 40 10 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before 8 31.33333 High 28 51.3333333 No 20 46.66666667 56 0 0 10 40 100 40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before 8 31.33333 Medium 12 35.3333333 Yes 0 26.66666667 20 40 100 40 10 0 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before 8 31.33333 Medium 12 35.3333333 No 20 46.66666667 40 10 40 100 40 10 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before 8 31.33333 Low 0 23.3333333 Yes 0 26.66666667 8 100 40 10 0 0 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before 8 31.33333 Low 0 23.3333333 No 20 46.66666667 28 40 100 40 10 0 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During 28 51.33333 High 28 51.3333333 Yes 0 26.66666667 56 0 0 10 40 100 40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During 28 51.33333 High 28 51.3333333 No 20 46.66666667 76 0 0 0 10 40 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During 28 51.33333 Medium 12 35.3333333 Yes 0 26.66666667 40 10 40 100 40 10 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During 28 51.33333 Medium 12 35.3333333 No 20 46.66666667 60 0 0 10 40 100 40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During 28 51.33333 Low 0 23.3333333 Yes 0 26.66666667 28 40 100 40 10 0 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During 28 51.33333 Low 0 23.3333333 No 20 46.66666667 48 0 10 40 100 40 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After 4 27.33333 High 28 51.3333333 Yes 0 26.66666667 32 40 100 40 10 0 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After 4 27.33333 High 28 51.3333333 No 20 46.66666667 52 0 10 40 100 40 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After 4 27.33333 Medium 12 35.3333333 Yes 0 26.66666667 16 40 100 40 10 0 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After 4 27.33333 Medium 12 35.3333333 No 20 46.66666667 36 10 40 100 40 10 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After 4 27.33333 Low 0 23.3333333 Yes 0 26.66666667 4 100 40 10 0 0 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After 4 27.33333 Low 0 23.3333333 No 20 46.66666667 24 40 100 40 10 0 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The way in which these ranges were calculated was by subtracting 1 point from the Score for each day after the “ideal planting date”, and adding 3 points to the Score for each day before the “ideal planting date.” The reasoning behind this is that planting too early gives a higher risk for disease than planting later, thus points are subtracted from the Score for planting later, and points added for planting before. A threshold of 44 was used to separate the suggestion to delay planting and the suggestion to plant. In other words, the network would not suggest to plant until the Score reached 44, which determined the Success of planting any given week, with a gradient of Success around the best week for planting taking into account pest conditions.

<table>
<thead>
<tr>
<th>Condition (Plant Suscept)</th>
<th>Value</th>
<th>Average Score (Plant Suscept)</th>
<th>Condition (Aphid Risk)</th>
<th>Value</th>
<th>Average Score (Aphid Risk)</th>
<th>Condition (Observational)</th>
<th>Value</th>
<th>Average Score (Observational)</th>
<th>Score</th>
<th>Full</th>
<th>1/2</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>28</td>
<td>48.00</td>
<td>High</td>
<td>24</td>
<td>44.00</td>
<td>Incoming</td>
<td>12</td>
<td>38.67</td>
<td>64</td>
<td>Full</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>High</td>
<td>28</td>
<td>48.00</td>
<td>High</td>
<td>24</td>
<td>44.00</td>
<td>Stationary</td>
<td>8</td>
<td>34.67</td>
<td>60</td>
<td>Full</td>
<td>100</td>
<td>20</td>
</tr>
<tr>
<td>High</td>
<td>28</td>
<td>48.00</td>
<td>High</td>
<td>24</td>
<td>44.00</td>
<td>Not present</td>
<td>0</td>
<td>26.67</td>
<td>52</td>
<td>Full</td>
<td>100</td>
<td>60</td>
</tr>
<tr>
<td>High</td>
<td>28</td>
<td>48.00 Medium</td>
<td>12</td>
<td>32.00</td>
<td>Incoming</td>
<td>12</td>
<td>38.67</td>
<td>52 Full</td>
<td>100</td>
<td>60</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>28</td>
<td>48.00 Medium</td>
<td>12</td>
<td>32.00</td>
<td>Stationary</td>
<td>8</td>
<td>34.67</td>
<td>48 Full</td>
<td>100</td>
<td>80</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>28</td>
<td>48.00 Medium</td>
<td>12</td>
<td>32.00</td>
<td>Not present</td>
<td>0</td>
<td>26.67</td>
<td>40 1/2</td>
<td>100</td>
<td>100</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>28</td>
<td>48.00 Low</td>
<td>4</td>
<td>24.00</td>
<td>Incoming</td>
<td>12</td>
<td>38.67</td>
<td>44 1/2</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>28</td>
<td>48.00 Low</td>
<td>4</td>
<td>24.00</td>
<td>Stationary</td>
<td>8</td>
<td>34.67</td>
<td>40 1/2</td>
<td>100</td>
<td>100</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>28</td>
<td>48.00 Low</td>
<td>4</td>
<td>24.00</td>
<td>Not present</td>
<td>0</td>
<td>26.67</td>
<td>32 1/2</td>
<td>100</td>
<td>100</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>12</td>
<td>32.00 High</td>
<td>24</td>
<td>44.00</td>
<td>Incoming</td>
<td>12</td>
<td>38.67</td>
<td>48 Full</td>
<td>100</td>
<td>80</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>12</td>
<td>32.00 High</td>
<td>24</td>
<td>44.00</td>
<td>Stationary</td>
<td>8</td>
<td>34.67</td>
<td>44 1/2</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>12</td>
<td>32.00 High</td>
<td>24</td>
<td>44.00</td>
<td>Not present</td>
<td>0</td>
<td>26.67</td>
<td>36 1/2</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>12</td>
<td>32.00 Medium</td>
<td>12</td>
<td>32.00</td>
<td>Incoming</td>
<td>12</td>
<td>38.67</td>
<td>36 1/2</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>12</td>
<td>32.00 Medium</td>
<td>12</td>
<td>32.00</td>
<td>Stationary</td>
<td>8</td>
<td>34.67</td>
<td>32 1/2</td>
<td>100</td>
<td>100</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>12</td>
<td>32.00 Medium</td>
<td>12</td>
<td>32.00</td>
<td>Not present</td>
<td>0</td>
<td>26.67</td>
<td>24 No</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>12</td>
<td>32.00 Low</td>
<td>4</td>
<td>24.00</td>
<td>Incoming</td>
<td>12</td>
<td>38.67</td>
<td>28 No</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>12</td>
<td>32.00 Low</td>
<td>4</td>
<td>24.00</td>
<td>Stationary</td>
<td>8</td>
<td>34.67</td>
<td>24 No</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>12</td>
<td>32.00 Low</td>
<td>4</td>
<td>24.00</td>
<td>Not present</td>
<td>0</td>
<td>26.67</td>
<td>16 No</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0</td>
<td>20.00 High</td>
<td>24</td>
<td>44.00</td>
<td>Incoming</td>
<td>12</td>
<td>38.67</td>
<td>36 1/2</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0</td>
<td>20.00 High</td>
<td>24</td>
<td>44.00</td>
<td>Stationary</td>
<td>8</td>
<td>34.67</td>
<td>32 1/2</td>
<td>100</td>
<td>100</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0</td>
<td>20.00 High</td>
<td>24</td>
<td>44.00</td>
<td>Not present</td>
<td>0</td>
<td>26.67</td>
<td>24 No</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0</td>
<td>20.00 Medium</td>
<td>12</td>
<td>32.00</td>
<td>Incoming</td>
<td>12</td>
<td>38.67</td>
<td>24 No</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0</td>
<td>20.00 Medium</td>
<td>12</td>
<td>32.00</td>
<td>Stationary</td>
<td>8</td>
<td>34.67</td>
<td>20 No</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0</td>
<td>20.00 Medium</td>
<td>12</td>
<td>32.00</td>
<td>Not present</td>
<td>0</td>
<td>26.67</td>
<td>12 No</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0</td>
<td>20.00 Low</td>
<td>4</td>
<td>24.00</td>
<td>Incoming</td>
<td>12</td>
<td>38.67</td>
<td>16 No</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0</td>
<td>20.00 Low</td>
<td>4</td>
<td>24.00</td>
<td>Stationary</td>
<td>8</td>
<td>34.67</td>
<td>12 No</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0</td>
<td>20.00 Low</td>
<td>4</td>
<td>24.00</td>
<td>Not present</td>
<td>0</td>
<td>26.67</td>
<td>4 No</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

Table B-2. Insecticide spray management Success and Scores
Success of conducting a half or no insecticide spray varies based on the following equations:

If(Score > 47, then Success of ½ spray = (((64 - Score)/20)*100) and Success of no spray = 0

If(47 ≤ Score ≤ 32, the Success of ½ spray = 100 and Success of no spray = (((44 - Score)/16)*100))

If(Score<32, then Success of ½ spray = 100 and Success of no spray = 100)

Where 64 represents the largest Score calculated for the Decision to Spray dependency network, 44 represents the largest Score that would suggest a ½ spray, 20 represents the difference between the largest Score (64) and the largest Score that would suggest a ½ spray (44), and 100 represents the highest possible Success.
Appendix C

Statistical Analyses of Aphid Migration Data from Italy and Midwest United States

The following figures were synthesized using data accumulated by Piero Caciagli of the Istituto di Virologia Vegetale in Torino, Italy, and by Doris Lagos of the University of Illinois at Urbana-Champaign. The data represents suction trap catches of aphid vectors of BYDV in several states in the Midwest United States and in Udine and Carmagnola, Italy. The data from the Midwest is from three years: 2007, 2009, and 2010. The figures representing the data from Italy in this appendix are from Carmagnola from the years of 1982-1989. The data was analyzed in these ways to better understand the phenology of aphid vectors of BYDV and to use this information to model better management decisions in the dependency networks from chapter 2.

There are different migration peaks for different species, however, only the fall migration is concerned for this thesis as it carries the greatest threat of BYD yield loss. To attempt to predict the fall migration of aphid vectors, temperature, wind speed, rain, atmospheric pressure, and aphid catch correlations were analyzed with Italian suction trap data. Temperature analyses were conducted using degree days (DD) with a lower developmental threshold of 5.78°C (Elliot and Kieckhefer 1989). Most years showed a dispersion of migration around 1500 DD from the beginning of July. If the migration started showing double digit and sustained catches well before this DD accumulation
(around 1250 DD), then it was usually a very large magnitude migration. However, if it started at approximately 1500 DD or after, the migration was generally weak.

For the wind speed analysis it was found that aphids will generally not fly after the wind speed reaches a certain point. This speed looks to be approximately 15 km/h. Aphids will still fly after wind speeds reach this point, but the number caught drastically decreased, with all of the large catches happening during days with relatively calm wind speeds. Little to no correlation was found between rain and atmospheric pressure for aphid suction trap catches. That is not to say that there is no correlation, just that none were found here.

Midwest Catches

Illinois

2007

![Total Catches in Illinois 2007](image)

Figure C-1
2009

Figure C-2

2010

Figure C-3
Indiana

2007

Figure C-4

2009

Figure C-5
2010

Figure C-6

Iowa

2007

Figure C-7
2009

Figure C-8

2010

Figure C-9
Kansas

2007

Figure C-10

Total Catches in Kansas 2007

2009

Figure C-11

Total Catches in Kansas 2009
Carmagnola, Italy

**Figure C-12.** This figure shows the average catch of *R. padi* in suction traps from Carmagnola, Italy from 1982-1989. This graph omitted Julian date 166 from 1987 because there were 1,692 individuals caught that day, which was skewing that data point.

**Figure C-13.** Average *S. avenae* caught in suction traps from 1982-1989.
Figure C-14. Average *S. avenae* caught in suction traps from 1982-1989 without 1985, an outbreak year for *S. avenae*

Figure C-15. Average *R. maidis* caught in suction traps from 1982-1989
Figure C-16. Suction trap catches of *R. padi* from Udine, Italy, 1985. This is an example of a large magnitude migration. The migration starts at about 1100 DD, but is still dispersed around 1500 DD.

Figure C-17. Suction trap Catches of *R. padi* from Udine, Italy, 1986. This is an example of a small magnitude migration. The migration starts at approximately 1400 DD.
Figure C-18. *R. padi* suction trap catches plotted against the mean wind speed of the day they were caught.

Every data point represents a day from 1983-1992.
Appendix D

BYDV Reverse Transcriptase PCR

The purpose of creating an RT-PCR technique to detect BYDV in host plants was to conduct a survey of the viruses present in Pennsylvania. Christelle Lacroix, a post-doc in Elizabeth Borer’s lab supplied some of the primer sequences, though many were altered slightly after using BLAST from the NCBI website to check sequences. There was not enough time to conduct an exhaustive BYDV survey, but many plants were tested with a single emmer (or einkorn) plant testing positive for BYDV-PAV and CYDV-RPV. The following protocol can be used for future detection of virus in host plants, but needs to be tested and confirmed on positive controls.

**RNA Extraction and Purification Technique**

Several methods were tested to find the best RNA extraction technique. The most successful was using the RNeasy Plant Mini Kit using the following protocol.

1. Place 100 mg plant tissue, 1000 μL RLT buffer, 10 μL β-mercaptoethanol, a metal bead, and sand (until bead is covered) in a 1.5 mL tube.
2. Repeat step 1 so there are 2 tubes with 100 mg plant tissue each.
3. Homogenize samples in a grinder at 300 rpm for 3:00 minutes.
4. Transfer lysate to a QIA shredder spin column (lilac) placed in a 2 mL collection tube.

5. Repeat step 4 for the second tube with 100 mg tissue

6. Centrifuge for 2 min. at full speed

7. Transfer flow-through to new microcentrifuge tube

8. Repeat step 7 for the second tube

9. Add 0.5X volume of 100% ethanol and mix immediately by pipetting

10. Repeat step 9 for the second tube.

11. Quickly transfer both samples in each tube to a single RNeasy Mini spin column (pink) in a 2 mL collection tube

12. Centrifuge at ≥8000g for 15 seconds.


14. Repeat steps 11-13 until all lysate/ethanol has been centrifuged through a single RNeasy Mini spin column.

15. Add 700 μL buffer RW1 to the pink spin column.

16. Centrifuge for 15 sec. at ≥8000g

17. Discard flow-through

18. Add 500 μL buffer RPE to RNeasy spin column

19. Centrifuge for 15 sec. at ≥8000g

20. Discard flow-through

21. Add 500 μL buffer RPE to RNeasy spin column

22. Centrifuge for 2 min. at ≥8000g

23. Discard flow-through
24. Place spin column in a new 1.5 mL collection tube

25. Add 50 μL RNase-free water directly to the spin column membrane

26. Centrifuge for 1 min. at ≥8000g

27. Resulting RNA can then be stored at -20°C or run through RT-PCR

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Product Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPV-F</td>
<td>ATGTTGTACCGCTTGATCCAC</td>
<td>425</td>
</tr>
<tr>
<td>RPV-R</td>
<td>CTGCGTTCTTGCAGCAGG</td>
<td>425</td>
</tr>
<tr>
<td>RMV-F</td>
<td>GACGAGGACGACCAAGTGGA</td>
<td>346</td>
</tr>
<tr>
<td>RMV-R</td>
<td>GCCATACTCCACCTCCGATT</td>
<td>346</td>
</tr>
<tr>
<td>SGV-F</td>
<td>ACCAGATCTTAGCCGGGTTT</td>
<td>238</td>
</tr>
<tr>
<td>SGV-R</td>
<td>CTGGACGTCGACCATTCTCTT</td>
<td>238</td>
</tr>
<tr>
<td>PAV-F</td>
<td>AGAGGAGGGGCAAAATYYTGT</td>
<td>273</td>
</tr>
<tr>
<td>PAV-R</td>
<td>ATKGTGAAGGAATTAATGTA</td>
<td>273</td>
</tr>
<tr>
<td>MAV-F</td>
<td>AGGAGGAGGGCAATCTTTTA</td>
<td>251</td>
</tr>
<tr>
<td>MAV-R</td>
<td>ACCCAAGGTTGATTGTGAGC</td>
<td>251</td>
</tr>
<tr>
<td>Generic-F</td>
<td>CRGMCAGTGGTTRTGG</td>
<td>Variable</td>
</tr>
</tbody>
</table>
### Table D-1. Primers used for BYDV testing

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic-R</td>
<td>TGGTAGGKCTTYAGWARTCC</td>
<td>Variable</td>
</tr>
<tr>
<td>Wheat-F</td>
<td>AAGACCATCACCCCTGGAGGT</td>
<td>169</td>
</tr>
<tr>
<td>Wheat-R</td>
<td>GAGACGGGAGCACAAGTGAA</td>
<td>169</td>
</tr>
</tbody>
</table>

### Table D-2. Reagents for each vial of PCR reaction

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume per Reaction (μL)</th>
<th>Initial Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>GoTaq Buffer</td>
<td>5</td>
<td>5X</td>
</tr>
<tr>
<td>DTT</td>
<td>1.1</td>
<td>0.1 M</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>1.1</td>
<td>25 mM</td>
</tr>
<tr>
<td>DnTP</td>
<td>0.5</td>
<td>10 mM</td>
</tr>
<tr>
<td>GoTaq DNA Polymerase</td>
<td>0.125</td>
<td>5 u/μL</td>
</tr>
<tr>
<td>Promega</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSII</td>
<td>0.035</td>
<td>200 u/μL</td>
</tr>
<tr>
<td>H₂O</td>
<td>12.65</td>
<td>50.6%</td>
</tr>
<tr>
<td>RNA (or H₂O for control)</td>
<td>2</td>
<td>Ng</td>
</tr>
<tr>
<td>Forward Primer</td>
<td>1.25</td>
<td>10 μM</td>
</tr>
<tr>
<td>Reverse Primer</td>
<td>1.25</td>
<td>10 μM</td>
</tr>
<tr>
<td>Cycle Repetitions</td>
<td>Cycle</td>
<td>Temperature (°C)</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------</td>
<td>------------------</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>50.0</td>
</tr>
<tr>
<td>35</td>
<td>2</td>
<td>95.0</td>
</tr>
<tr>
<td>35</td>
<td>2</td>
<td>50.0</td>
</tr>
<tr>
<td>35</td>
<td>2</td>
<td>72.0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>72.0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Table D-3. PCR Cycles.

**RT-PCR Protocol**

1. Dilute primers (Table 1) to 1 mMol
2. Vortex and centrifuge for several seconds
3. Let stand on ice for 10 min.
4. Dilute primers further to 10 μMol
5. Vortex and centrifuge
6. Place reagents into each PCR reaction vial as per Table 2.
7. Place tubes in PCR machine with protocol described by Table 3.

**Making and Running Gel**
1. Combine 1.5% agarose with 0.5 M TAE buffer
2. Heat and mix until no crystals are visible
3. Allow liquid to cool (but not to RT) and add 8 μL SYBR safe
4. Pour liquid into well with comb and allow to solidify
5. Remove comb and place gel in electrophoresis machine with wells at negative end
6. Fill with 0.5 M TAE buffer until thin film covers gel
7. Add 2 μL SYBR safe to positive end of machine
8. Load gel wells with 8 μL samples and a ladder
9. Run electrophoresis until yellow band almost reaches edge of gel
Literature Cited


123


dwarf virus by Rhopalosiphum padi (L.) and Sitobion avenae (F.). Annals of applied biology. 128(1), 45-53.


