NETWORK ANALYSIS OF SURVEY DATA FOR CHARACTERIZATION OF YIELD REDUCING FACTORS OF TROPICAL RICE ECOSYSTEMS IN SOUTH AND SOUTHEAST ASIA

SITH JAISONG

**SUBMITTED TO THE FACULTY OF GRADUATE SCHOOL
UNIVERSITY OF THE PHILIPPINES LOS BAÑOS
IN PARTIAL FULFILLMENT OF THE
REQUIREMENT FOR THE
DEGREE OF**

**DOCTOR OF PHILPSOPHY
(Plant Pathology)**

JUNE 2016

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

Differential networks reveal the dynamics of animal

pests and disease co-occurrences

Introduction

Rice (*Oryza sativa*) is a major crop in South and Southeast Asian. Generally, rice farmers cultivate 2 rice crops per year, with the typical seasonal crop cycles or rotations being rice-rice-fallow or rice-rice-secondary crops (corn, soybean, peanut). The Food and Agriculture Organization of the United Nations (FAO) estimates that approximately 70 percent of total lowland rice area produces 2 rice crops each year. The first crop is cultivated in the wet season, while another is in the dry season. The important role of seasonal cropping in the temporal dynamics of animal pests and diseases has been studied under farmers field survey in South and Southeast Asia by the use of multivariate techniques Savary et al. (2000); Willocquet et al. (2008). The previous studies showed that Injuries profiles (the combination of injuries) differ from season to season in term of weather pattern. In the dry season, crop losses were lower than in the wet season. A previous study based on surveys done in farmers' rice fields in the region of lowland rice were shown to be strongly associated with injury profiles.

In the previous chapter, the co-occurrence networks to yield co-occurrence networks, a methodological approach which has already proved fruitful in a variety of dif-

ferent applications. Plant injuries caused by pests maybe affect yield production. There-fore, in this chapter, I attempted to characterize the patterns of rice injuries by studying the changes in the co-occurrence patterns of rice injuries (*e.g* disease incidence, animal pest injury incidence) at different yield levels.

Differential network analysis aims to compare the connectivity of two nodes at 2 different conditions. As demonstrated by several studies, differential networks can identify important nodes implicated in my fields, and also provide critical novel in-sights not obtainable using other approaches. In this work, I explore the the properties of network of a complex association of rice injuries at different yield levels. Eluci-dating the rice injuries association represents a key challenge, not only for achieving a deeper understanding of injury association (injury profiles) but also for identifying the unique association. Given that the injury association is governed by a complex network of injuries association, it seems natural to explore network properties which may help elucidate some of different association presenting in the different seasons.

In this chapter, I employ a differential network topology method to examine the co-occurrence relationships of rice injuries from survey data. I use graph theory methods to examine the topological feature dynamic of a co-occurrence network corresponding to different seasons, and production environments. The co-occurrence networks were built from differentially co-occurring injuries. I extract significantly differential co-occurring injuries from co-occurrence networks, which represent different seasons, to identify which injuries that may be involved specifically curtain season. I postulate that these selected injuries may contribute to the difference in the co-occurrence patterns in different season. Furthermore, I identified the injuries associated with yield from net-works at different yield levels. Finally, I suggest key injuries that may contribute to yield

reduction under a curtain production environment. The goal is to leverage insights to better understand the rice injury co-occurrence that may contribute to pest management development.

Materials and Methods

Differential co-occurrence network construction

The survey data were pre-processed by using methods described in the previous chapter. Subsequently, I applied the method proposed by Fukushima (2013) to identify differentially co-occurrence links. The difference of co-occurrence of injury x and y between two conditions (A and B) was quantified by Fisher's z-test.

For the pair of x injury and y injury, I denoted the correlation coefficient based on Spearman's correlation coefficient by r_{xy}^A and r_{xy}^A in networks of condition A and condition B, respectively. To test whether the 2 correlation coefficients were significantly different, correlation coefficients for each of the 2 conditions, r_{xy}^A and r_{xy}^B, were transformed into Z_{xy}^A and Z_{xy}^B, respectively.

The Fisher's transformation of coefficient r_{xy}^A is defined by

$$Z_{xy} = \frac{1}{2} \log \left[\frac{1 + r_{xy}}{1 - r_{xy}} \right] \tag{I.1}$$

Next, The p-value of the difference in Z_{xy} values was calculated using the standard normal distribution.

$$p(Z \geq \left| \frac{Z_{xy}^A - Z_{xy}^B}{\sqrt{\frac{1}{N_A - 3} + \frac{1}{N_A - 3}}} \right| \tag{I.2}$$

Next, The p-value of the difference in Z values was calculated using the standard

normal distribution

N_A and N_A represent the sample size for each of condition. The Z has an approximately Gaussian distribution under null hypotheses that the population correlations are equal. The pairwise correlation significants are considered at p-value < 0.05.

Differential co-occurrence network in different seasons

Consider any two injuries x and y in the survey data, let r_{xy}^D and r_{xy}^W be the Spearman's correlation coefficient calculated separately over the samples in dry and wet, respectively. I constructed differential co-occurrence networks that are specified by adjacency matrix $A^{diff} = (A_{xy}^{diff})$ where the entry A_{xy}^{diff} quantified by following:

$$A_{xy}^{diff} = \begin{cases} 1 & \text{when } r_{xy}^D > r_{xy}^W \text{ at } P_{z_{xy}}\text{-value} < 0.05 \\ 0 & \text{when } P_{z_{xy}}\text{-value} > 0.05 \\ -1 & \text{when } r_{xy}^W > r_{xy}^D \text{ at } P_{z_{xy}}\text{-value} < 0.05 \end{cases} \qquad \text{(I.3)}$$

For this differential co-occurrence network,A_{xy}^{diff} equals 1 depending on whether any injury pairs show significantly higher co-occurrence level in dry season than wet season, but -1 is vice versa, and if it equal 0, meaning that co-occurrence level of injury pairs were not different in dry and wet season. I-1 illustrated the differential co-occurrence network at different seasons.

Difference of co-occurrence network of rice injuries at different yield levels

Consider any two injuries x and y in the survey data, let r_{xy}^L and r_{xy}^H be the Spearman's correlation coefficient calculated separately over the samples in L and H yield level, respectively. I constructed differential co-occurrence networks that are specified by adjacency matrix $A^{diff} = (A_{xy}^{diff})$ where the entry A_{xy}^{diff} quantified by following:
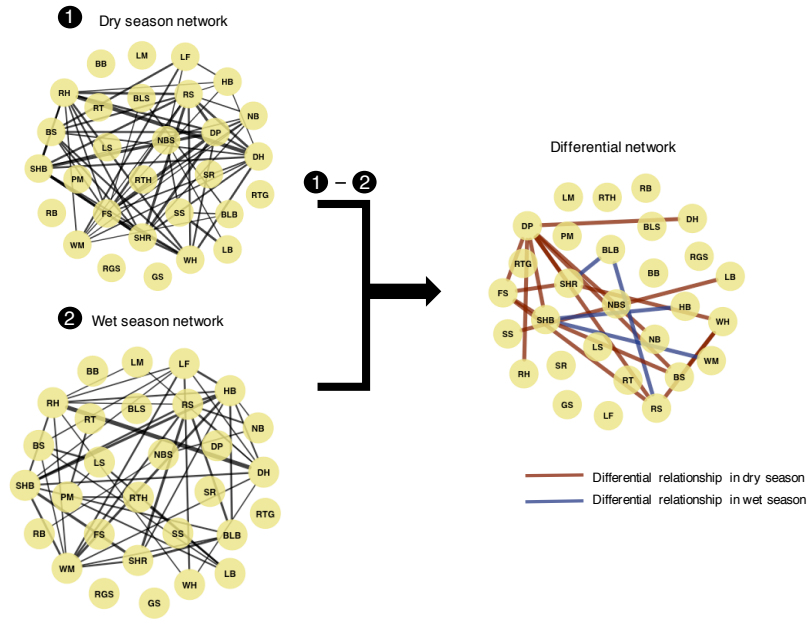
Figure I-1: Schematic showing principle of differential analysis in seasons. Co-occurrence networks are measured in each of two seasons (left) resulting in interactions (black). Dry season is subtracted from wet season to create a differential co-occurrence network (right), in which the significant differential interactions are those that positive (red) or negative (blue) in score after the shift in conditions, which means differential in dry, and wet season, respectively.

$$A_{xy}^{diff} = \begin{cases} 1 & \text{when } r_{xy}^L > r_{xy}^H \text{ at } P_{z_{xy}}\text{-value} < 0.05 \\ 0 & \text{otherwise} \end{cases} \tag{I.4}$$

For this differential co-occurrence network, A_{xy}^{diff} equals 1 depending on whether any injury pairs show significantly higher co-occurrence level in low yield level than high yield state, and if it equal 0, meaning that co-occurrence level of injury pairs were not or lower different in low yield level state. I-2 illustrated the differential co-occurrence network at different seasons.

Topological properties To investigate the structural properties of differential networks, I calculated topological features for each node in the network with the **igraph** package. This feature set included node degree, clustering coefficient, and betweenness.
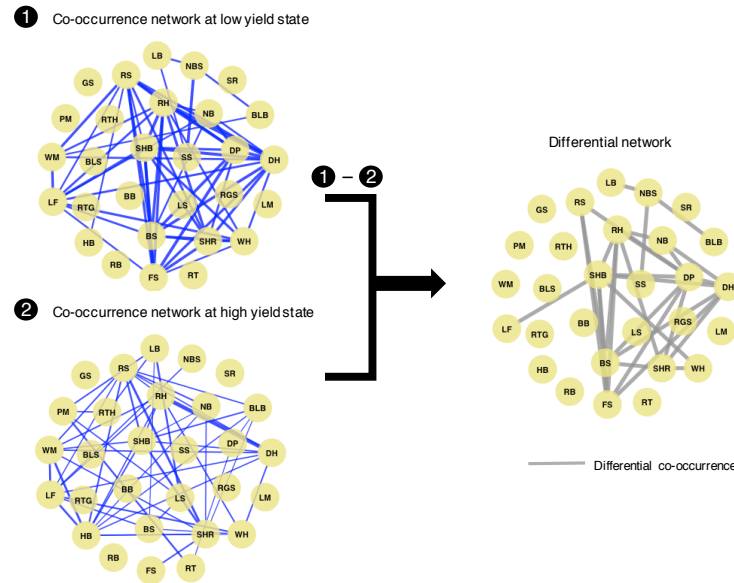
Figure I-2: Schematic showing differential analysis at different yield levels. Co-occurrence networks are measured in each of two different yield levels (left) resulting in interactions (blue). The network at low yield level network is subtracted from high yield level to create a differential co-occurrence network (right), in which the significant differential interactions are those that positive in score after the shift from low to high yield state.

I.1   Results

Construction of differential co-occurrence networks of rice injuries at different seasons

I determined which co-occurrence patterns of rice injuires of survey data were differentially expressed between dry and wet season. The resulted networks are showed in I-3) to Figure.I-7). Differential co-occurrence network (DCON) in seasons at CP (Figure. I-3) reveals SHB, RS, SHB showing significantly different co-occurrence level in both dry and wet season. DCON in season at OD **??** reveals LB different both in dry and wet season. DCON in season at RR I-5 reveals that DP, BS, RTH, LF. DCON in season at WJ (Figure.I-7 revealed that BS, NBS, SHB BLS. I-6

Figure I-3: Differential co-occurrence network of rice injuries in different seasons at Central Plain, Thailand

Figure I-4: Differential co-occurrence network of rice injuries in different seasons at Odisha, India

Construction of differential co-occurrence networks of rice injuries at yield level

In this study, three successive yield classes were defined, in order to enable a better description of actual yield, from low (< 4 ton/ha), medium (4 – 6 ton/ha) , high (> 6 ton/ha) yield levels. Figure I-8 shows the number of farmers' fields surveyed classified in each season, and production environment.

Berry and Widder (2014) recommended that a co-occurrence network will be more reliable, it should be produced using a minimum of 25 samples or observations. From figure I-8, to be able compare the networks at different yield levels, I chose the data set, which are medium and high yield level of CP, low and medium yield level of TM, medium and high yield level at RR, low and medium yield level of TM, and

Legend

Node

BB	=	bug burn
BLB	=	bacterial leaf blight
BLS	=	bacterial leaf streak
BS	=	brown spot
DH	=	deadheart
DP	=	dirty panicle
FS	=	false smut
GS	=	grassy stunt
HB	=	hopper burn
LB	=	leaf blast
LF	=	leaffolder injury
LM	=	leaf miner injury
LS	=	leaf streak
NB	=	neck blast
NBS	=	narrow brown spot
PM	=	panicle mite injury
RB	=	rice bug injury
RGS	=	ragged stunt
RH	=	rice hispa
RT	=	rat injury
RS	=	red stripe
RTG	=	rice tungo
RTH	=	rice thrip injury
SHB	=	sheath blight
SHR	=	sheath rot
SR	=	stem rot
SS	=	silver shoot
WH	=	whitehead
WM	=	whorl maggot injury

Edge

— = differential co-occurrence in dry season

— = differential co-occurrence in wet season

Figure I-5: Differential co-occurrence network of rice injuries in different seasons at Red River Delta, Vietnam
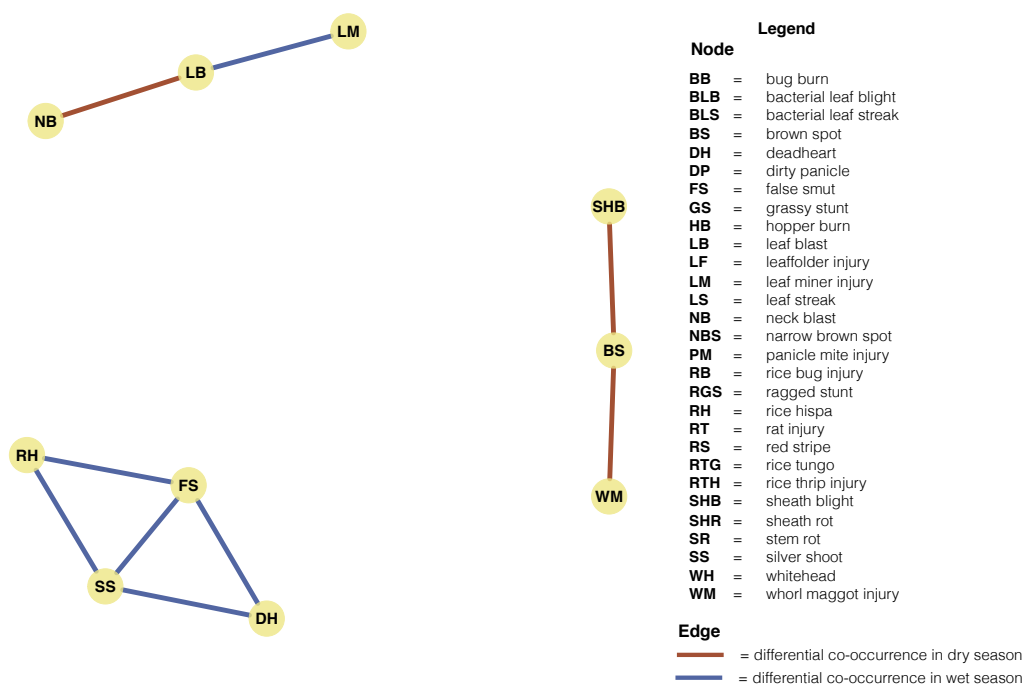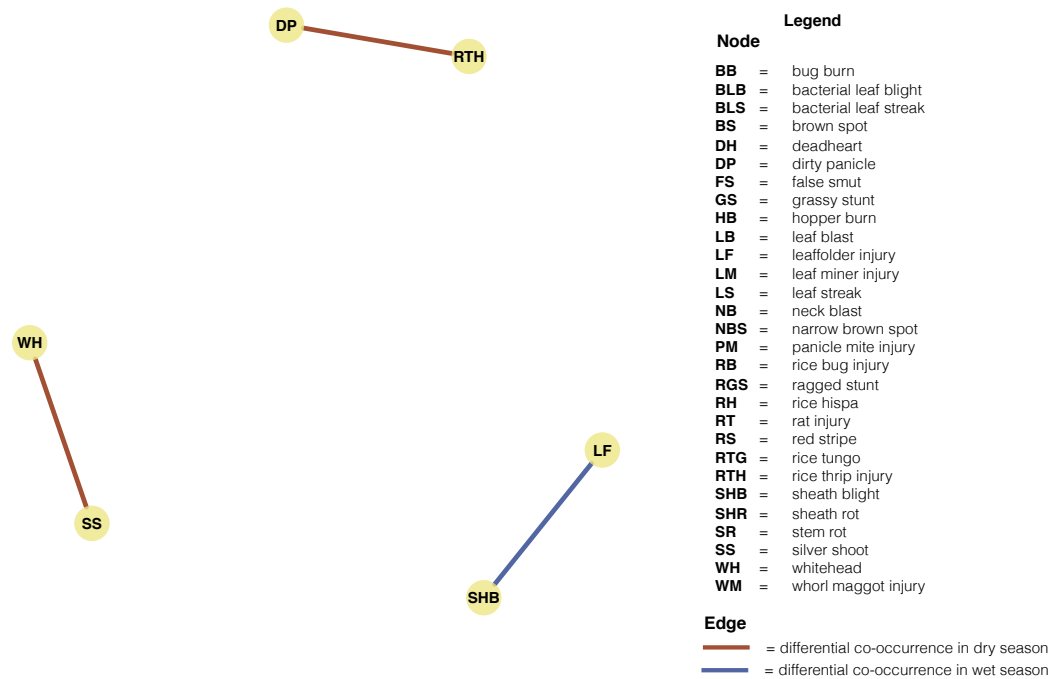
Figure I-6: Differential co-occurrence network of rice injuries in different seasons at Tamil Nadu, India

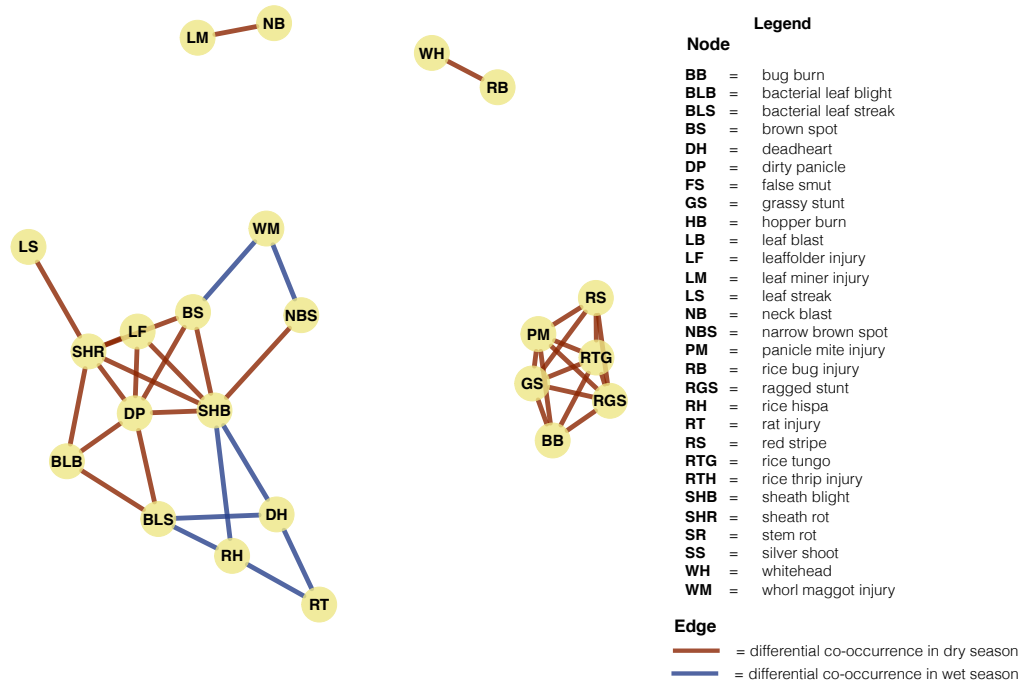Figure I-7: Differential co-occurrence network of rice injuries in different seasons at West Java, Indonesia
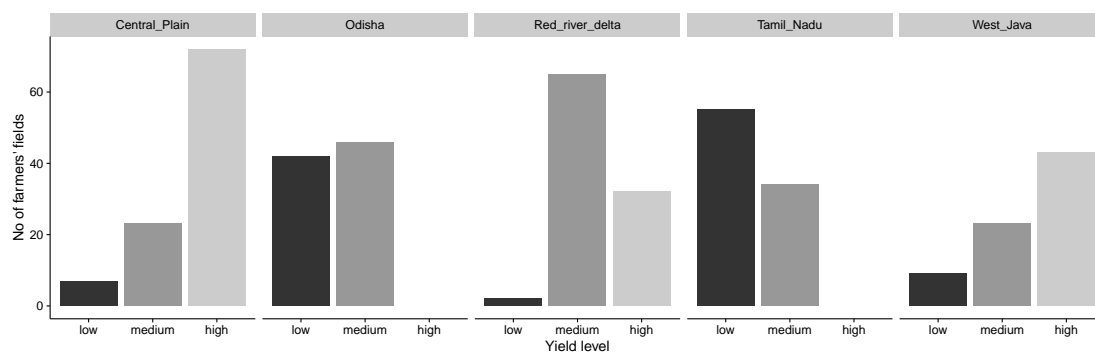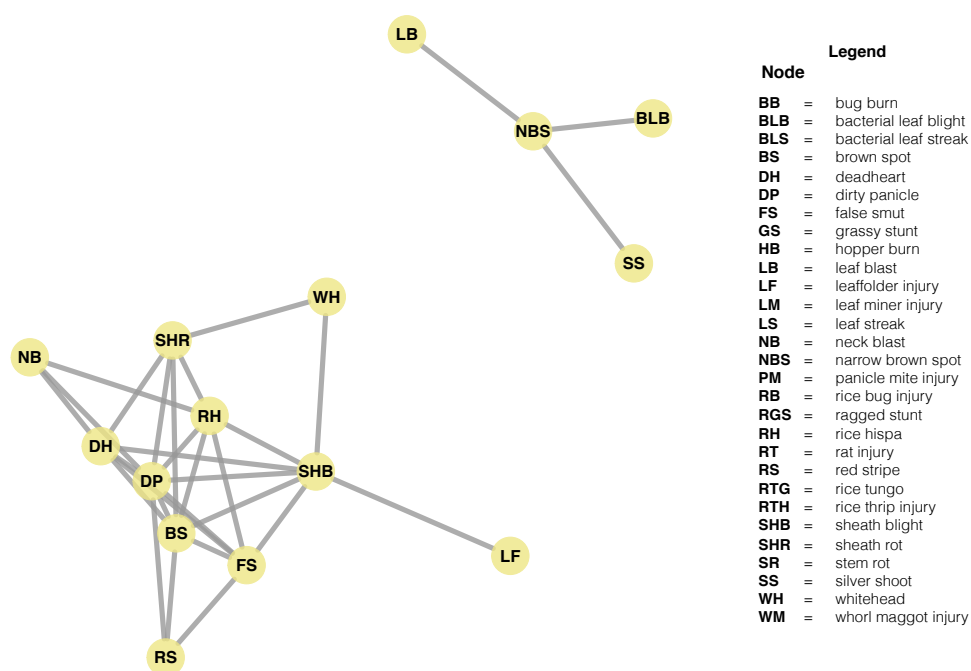
Figure I-8: .

medium and high yield level of WJ.

In this study, an differential co-occurrence network (DCON) in yield depicts the associations of injury pairs presenting in lower yield state but absent in higher yield state (Figures. **??**, I-10, I-11a, **??**, **??**). DCON

Discussion

Legend

Node

BB	=	bug burn
BLB	=	bacterial leaf blight
BLS	=	bacterial leaf streak
BS	=	brown spot
DH	=	deadheart
DP	=	dirty panicle
FS	=	false smut
GS	=	grassy stunt
HB	=	hopper burn
LB	=	leaf blast
LF	=	leaffolder injury
LM	=	leaf miner injury
LS	=	leaf streak
NB	=	neck blast
NBS	=	narrow brown spot
PM	=	panicle mite injury
RB	=	rice bug injury
RGS	=	ragged stunt
RH	=	rice hispa
RT	=	rat injury
RS	=	red stripe
RTG	=	rice tungo
RTH	=	rice thrip injury
SHB	=	sheath blight
SHR	=	sheath rot
SR	=	stem rot
SS	=	silver shoot
WH	=	whitehead
WM	=	whorl maggot injury

(a) Differential co-occurrence network of rice injuries in different yield levels at Central Plain, Thailand. The layout of the network graph is based on the Fruchterman-Reingold algorithm, which places nodes with stronger or more connections closer to each other.
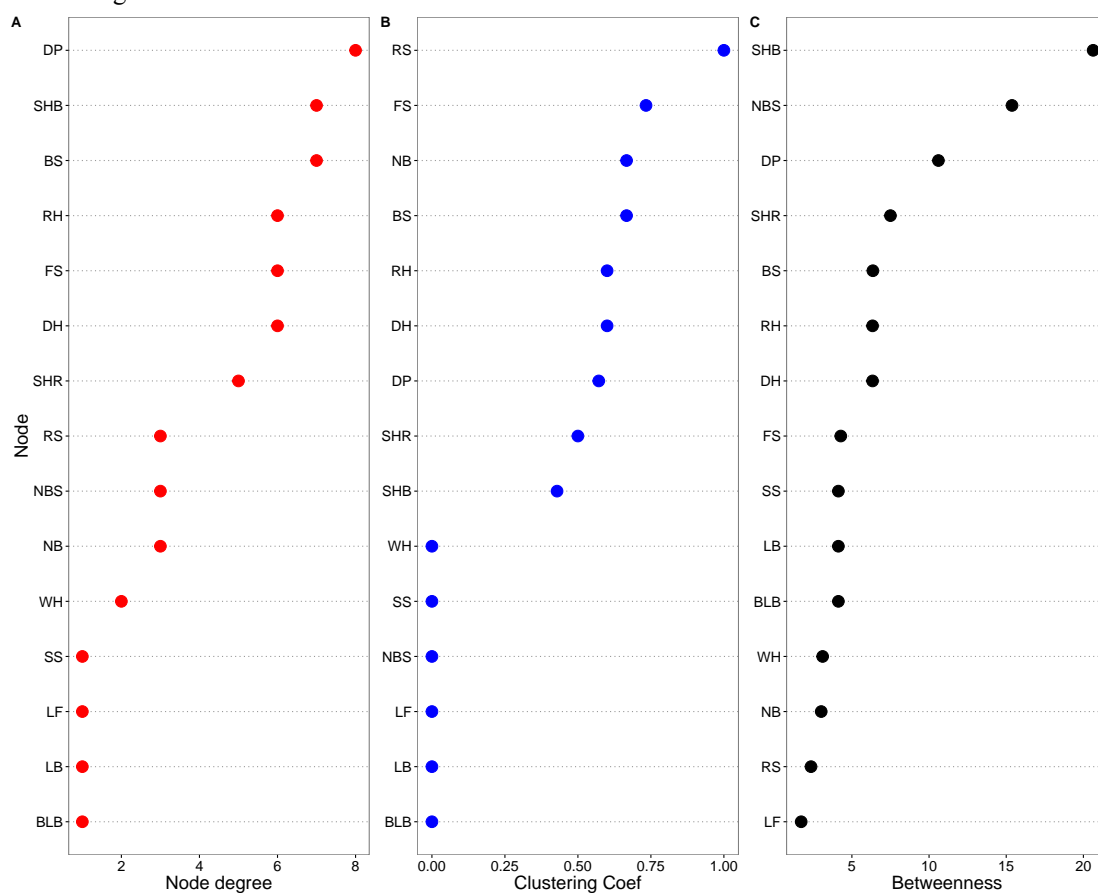
(b) Three centrality measures of the nodes in co-occurrence network of rice injuries in dry season at Central Plain. A: node degree, B:clustering coefficient, and C:Betweenness.

Figure I-9: Rice injuries in dry season in Central Plain, Thailand

Legend

Node

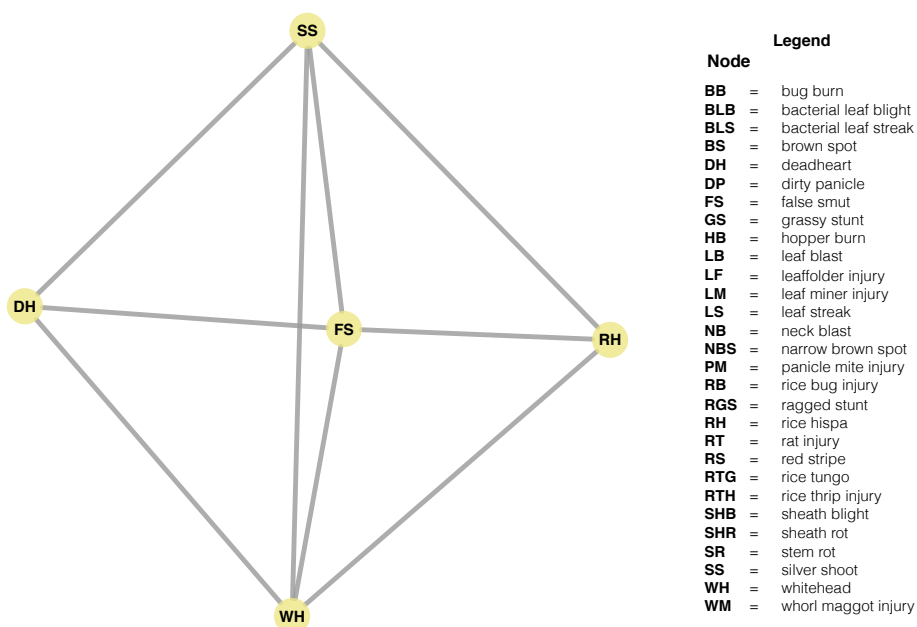|		
BB	=	bug burn
BLB	=	bacterial leaf blight
BLS	=	bacterial leaf streak
BS	=	brown spot
DH	=	deadheart
DP	=	dirty panicle
FS	=	false smut
GS	=	grassy stunt
HB	=	hopper burn
LB	=	leaf blast
LF	=	leaffolder injury
LM	=	leaf miner injury
LS	=	leaf streak
NB	=	neck blast
NBS	=	narrow brown spot
PM	=	panicle mite injury
RB	=	rice bug injury
RGS	=	ragged stunt
RH	=	rice hispa
RT	=	rat injury
RS	=	red stripe
RTG	=	rice tungo
RTH	=	rice thrip injury
SHB	=	sheath blight
SHR	=	sheath rot
SR	=	stem rot
SS	=	silver shoot
WH	=	whitehead
WM	=	whorl maggot injury

(a) Differential co-occurrence network of rice injuries in different yield levels at Central Plain, Thailand. The layout of the network graph is based on the Fruchterman-Reingold algorithm, which places nodes with stronger or more connections closer to each other.
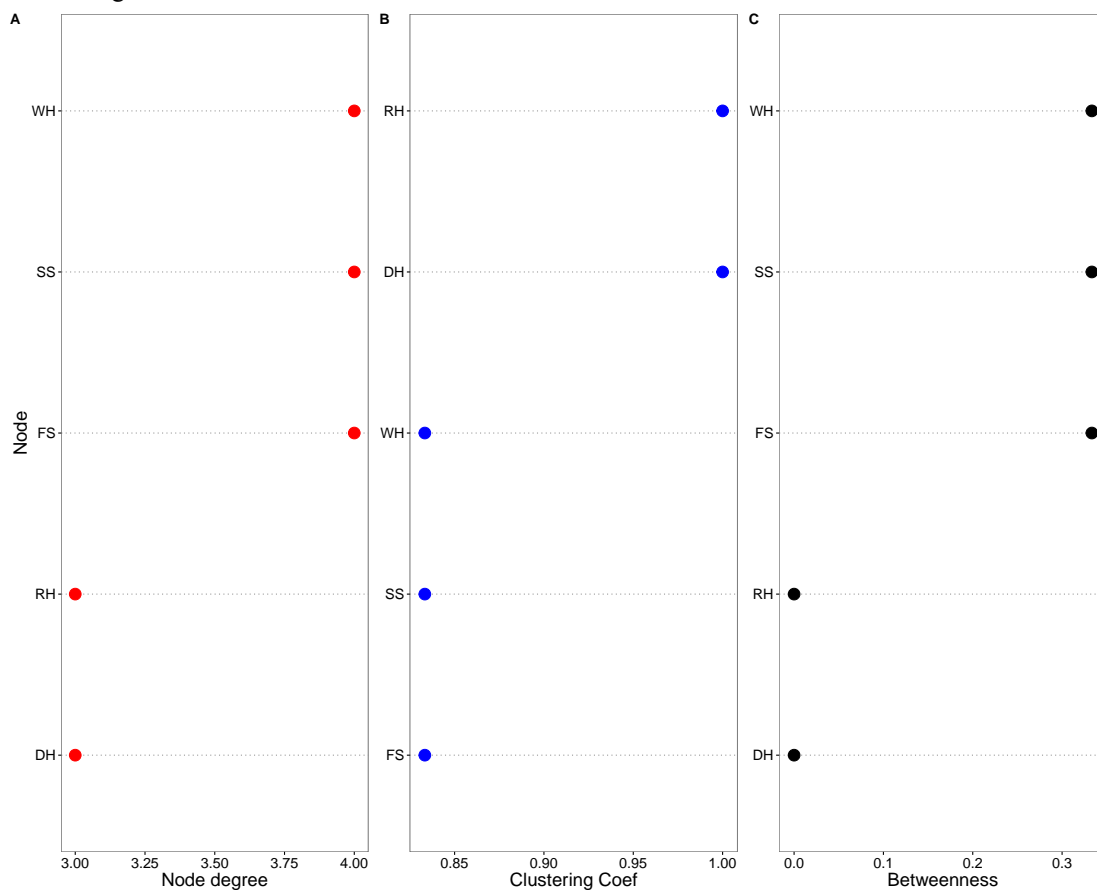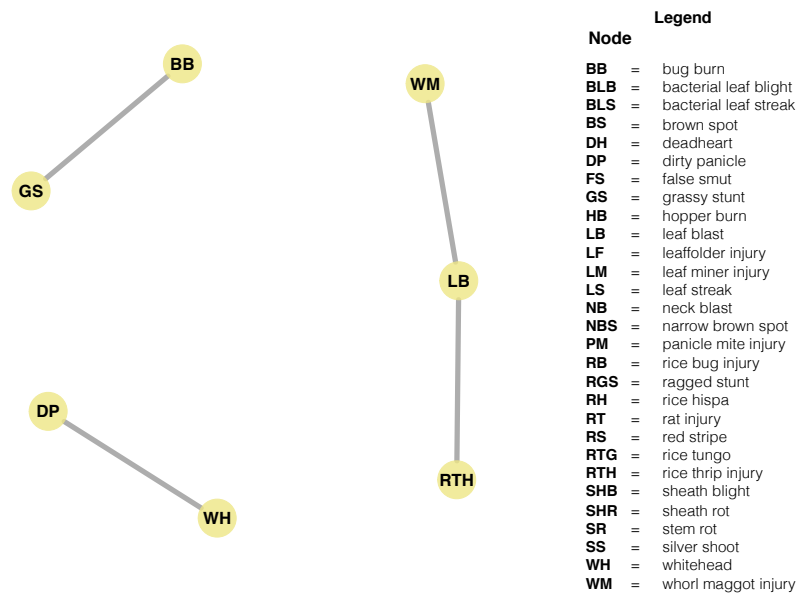

(b) Three centrality measures of the nodes in co-occurrence network of rice injuries in dry season at Central Plain. A: node degree, B:clustering coefficient, and C:Betweenness.

Figure I-10: Rice injuries in dry season in Central Plain, Thailand

Legend

Node

BB	=	bug burn
BLB	=	bacterial leaf blight
BLS	=	bacterial leaf streak
BS	=	brown spot
DH	=	deadheart
DP	=	dirty panicle
FS	=	false smut
GS	=	grassy stunt
HB	=	hopper burn
LB	=	leaf blast
LF	=	leaffolder injury
LM	=	leaf miner injury
LS	=	leaf streak
NB	=	neck blast
NBS	=	narrow brown spot
PM	=	panicle mite injury
RB	=	rice bug injury
RGS	=	ragged stunt
RH	=	rice hispa
RT	=	rat injury
RS	=	red stripe
RTG	=	rice tungo
RTH	=	rice thrip injury
SHB	=	sheath blight
SHR	=	sheath rot
SR	=	stem rot
SS	=	silver shoot
WH	=	whitehead
WM	=	whorl maggot injury

(a) Differential co-occurrence network of rice injuries in different yield levels at Central Plain, Thailand. The layout of the network graph is based on the Fruchterman-Reingold algorithm, which places nodes with stronger or more connections closer to each other.
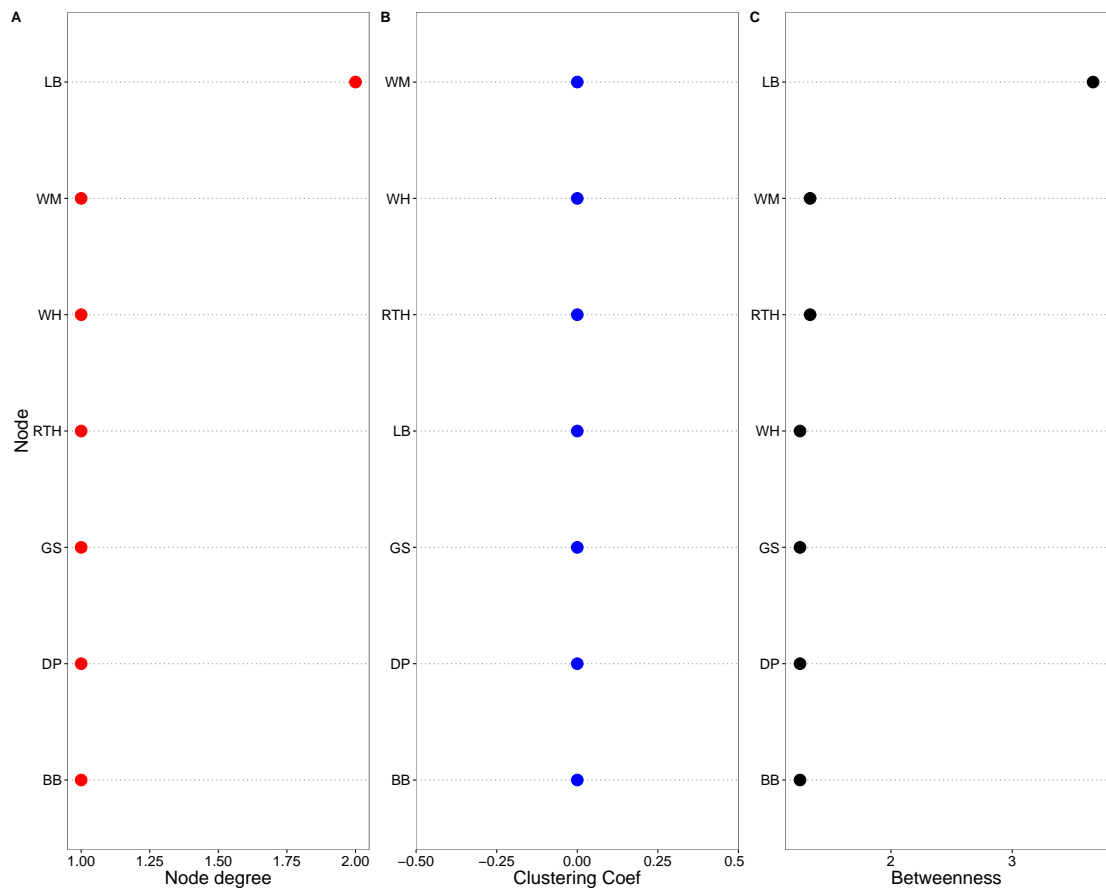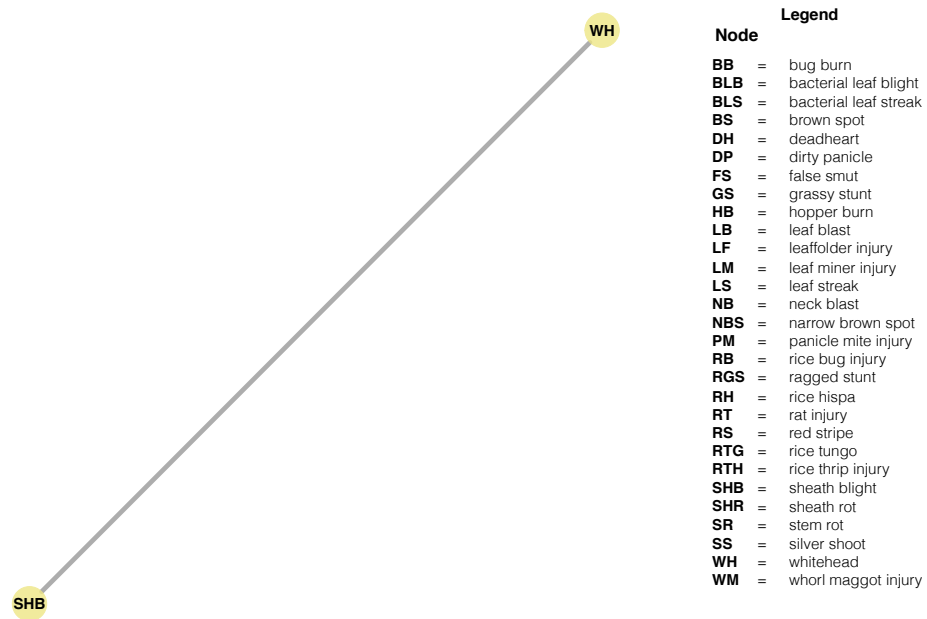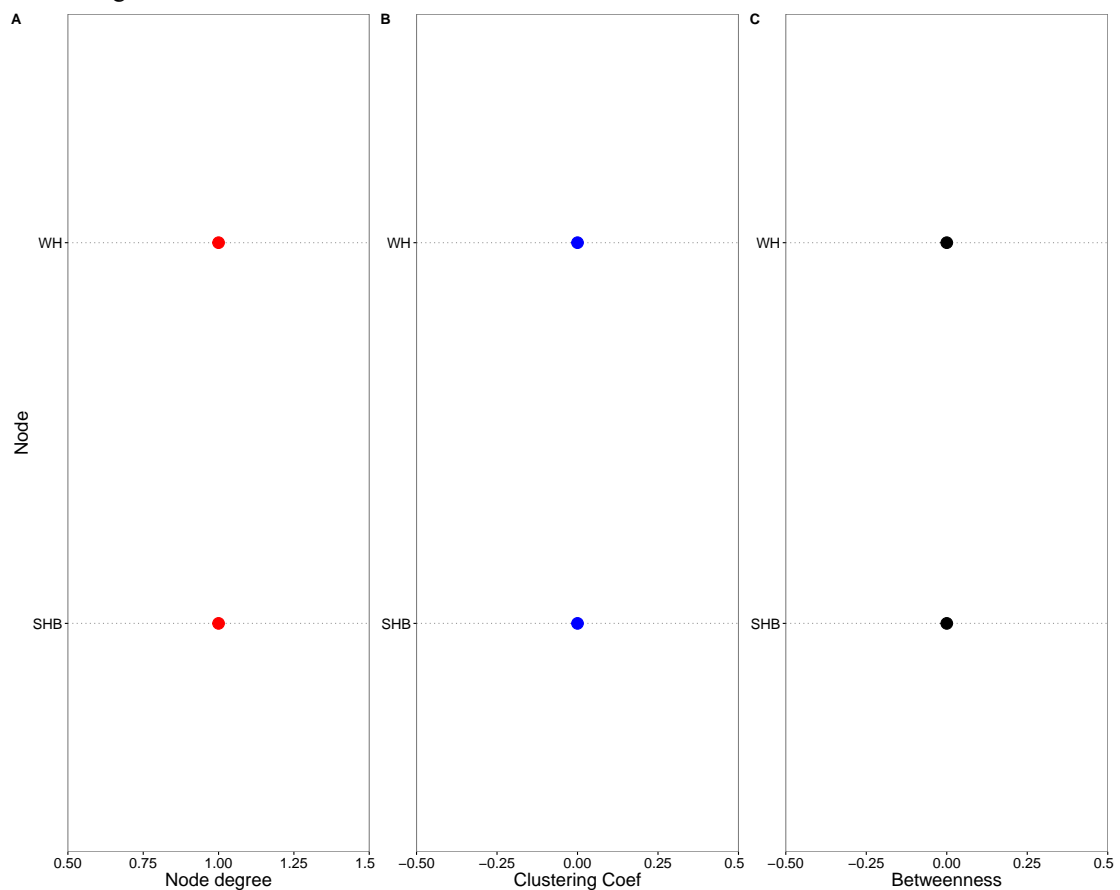

(b) Three centrality measures of the nodes in co-occurrence network of rice injuries in dry season at Central Plain. A: node degree, B:clustering coefficient, and C:Betweenness.

Figure I-11: Rice injuries in dry season in Central Plain, Thailand

Legend

Node

BB	=	bug burn
BLB	=	bacterial leaf blight
BLS	=	bacterial leaf streak
BS	=	brown spot
DH	=	deadheart
DP	=	dirty panicle
FS	=	false smut
GS	=	grassy stunt
HB	=	hopper burn
LB	=	leaf blast
LF	=	leaffolder injury
LM	=	leaf miner injury
LS	=	leaf streak
NB	=	neck blast
NBS	=	narrow brown spot
PM	=	panicle mite injury
RB	=	rice bug injury
RGS	=	ragged stunt
RH	=	rice hispa
RT	=	rat injury
RS	=	red stripe
RTG	=	rice tungo
RTH	=	rice thrip injury
SHB	=	sheath blight
SHR	=	sheath rot
SR	=	stem rot
SS	=	silver shoot
WH	=	whitehead
WM	=	whorl maggot injury

(a) Differential co-occurrence network of rice injuries in different yield levels at Central Plain, Thailand. The layout of the network graph is based on the Fruchterman-Reingold algorithm, which places nodes with stronger or more connections closer to each other.

(b) Three centrality measures of the nodes in co-occurrence network of rice injuries in dry season at Central Plain. A: node degree, B:clustering coefficient, and C:Betweenness.
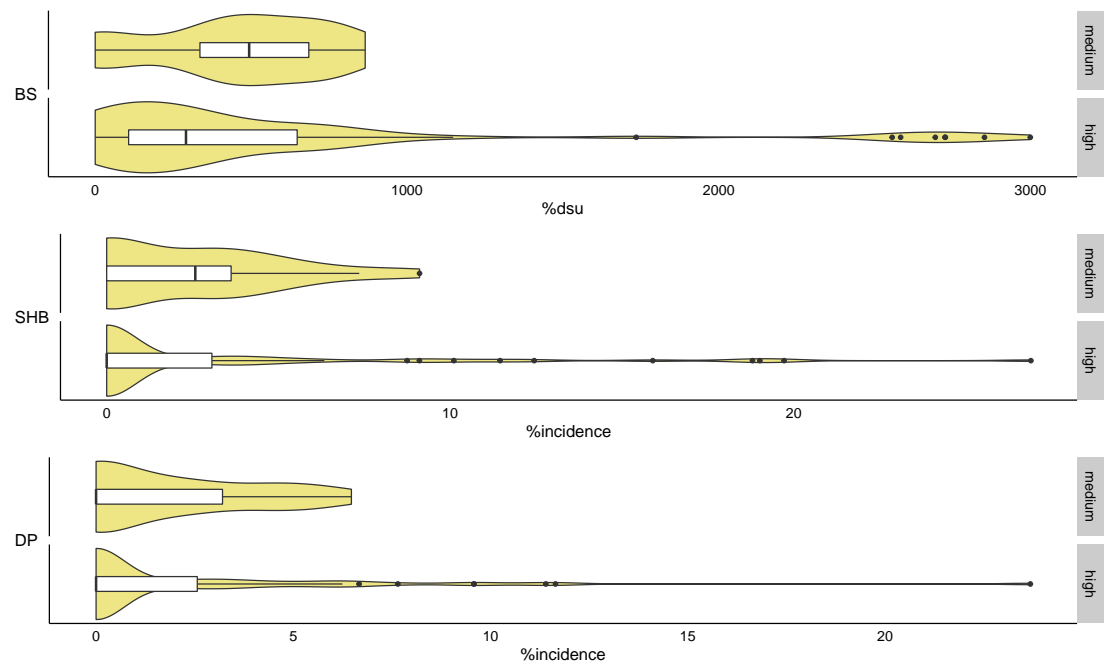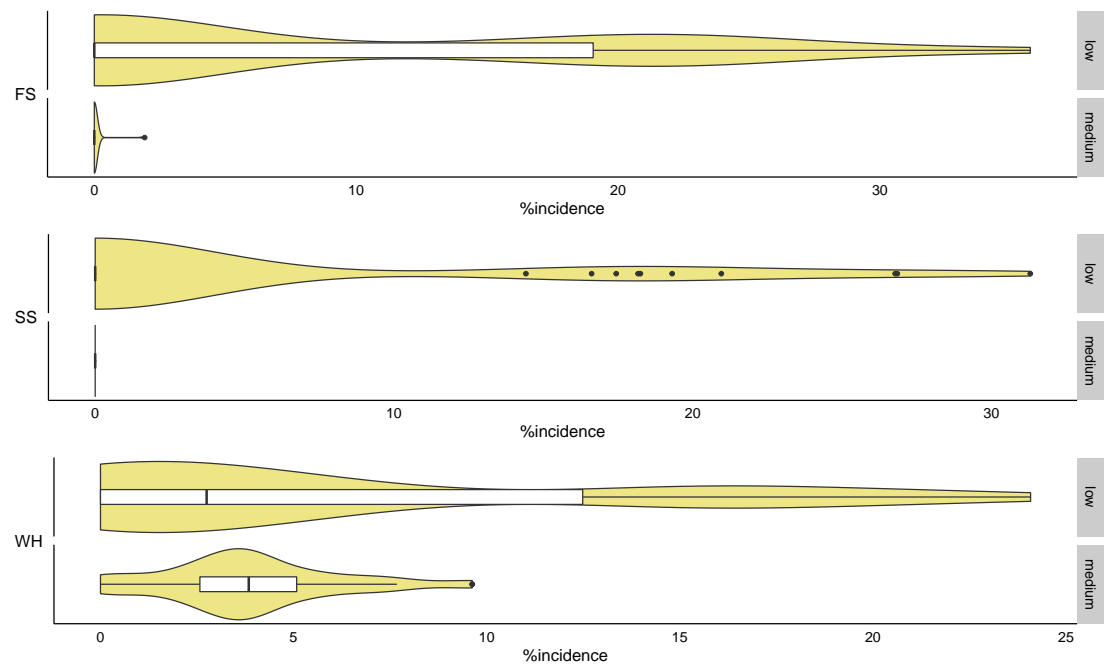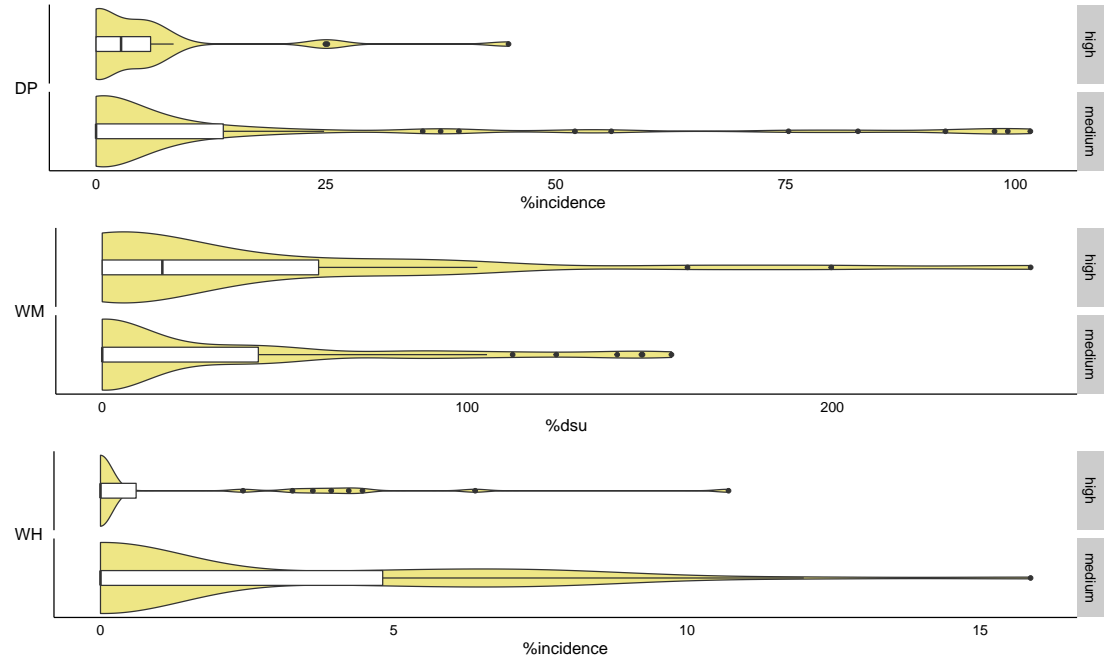
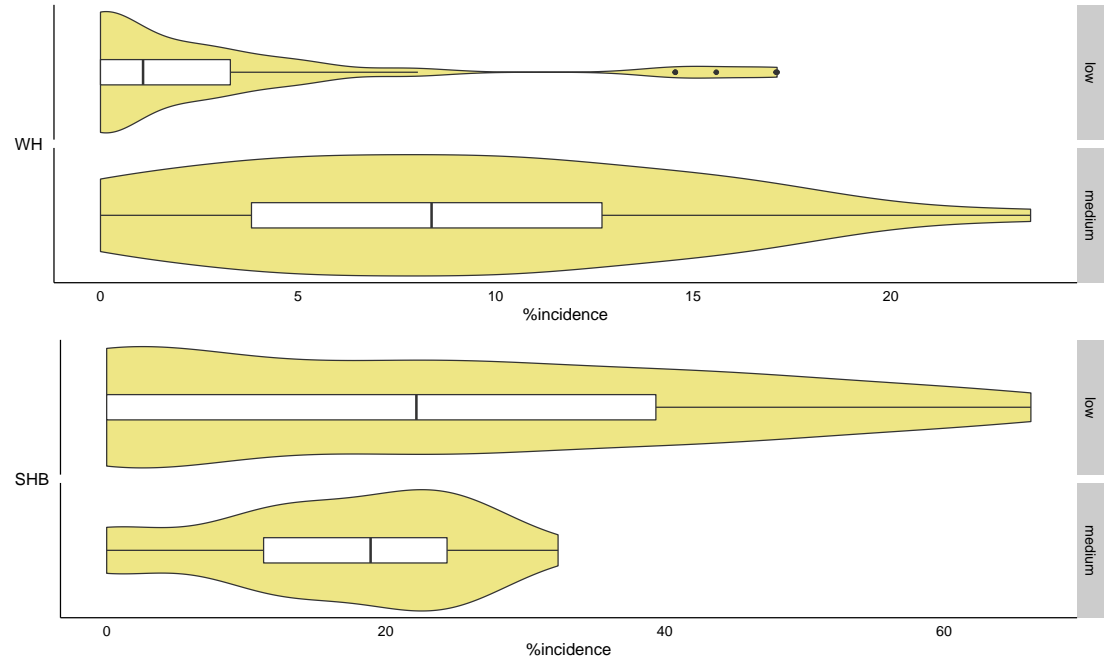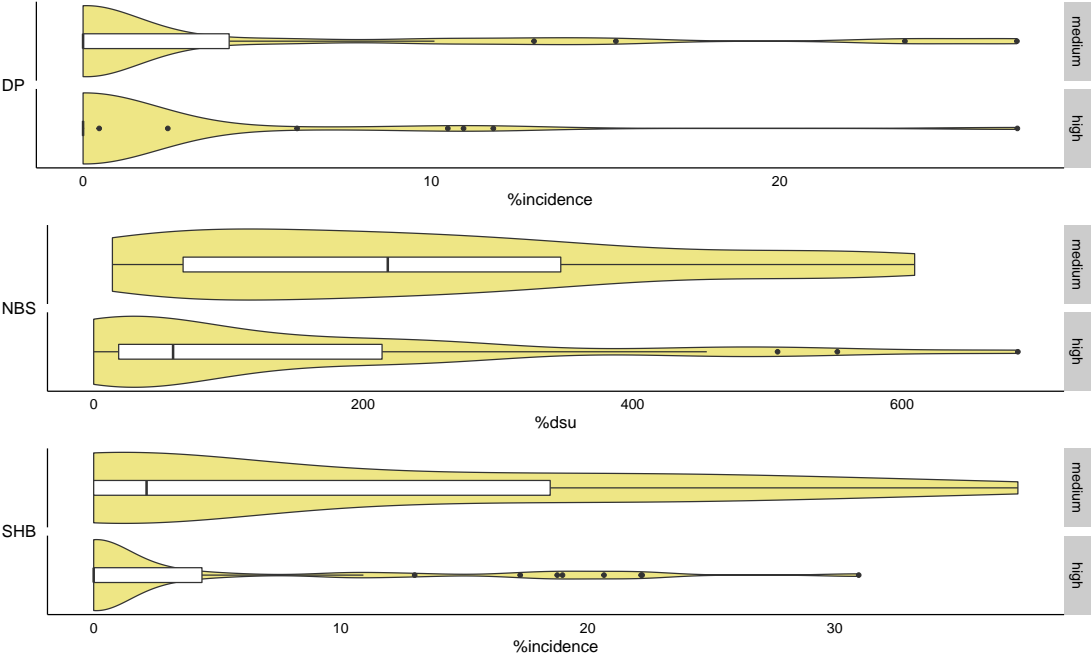Figure I-12: .

Figure I-13: .

Figure I-14: .

Figure I-15: .

Figure I-16: .

Figure I-17: .

LITERATURE CITED

BERRY, D. and WIDDER, S. 2014. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. Front. Microbiol, 5(219):10–3389.

FUKUSHIMA, A. 2013. Diffcorr: an r package to analyze and visualize differential correlations in biological networks. Gene, 518(1):209–214.

SAVARY, S., WILLOCQUET, L., ELAZEGUI, F. A., TENG, P. S., DU, P. V., ZHU, D., TANG, Q., HUANG, S., LIN, X., SINGH, H. M., and SRIVASTAVA, R. K. 2000. Rice Pest Constraints in Tropical Asia: Characterization of Injury Profiles in Relation to Production Situations. Plant Disease, 84(3):341–356.

WILLOCQUET, L., AUBERTOT, J., LEBARD, S., ROBERT, C., LANNOU, C., and SAVARY, S. 2008. Simulating multiple pest damage in varying winter wheat production situations. Field Crops Research, 107(1):12–28.