Characterizing rice pest injury co-occurrence patterns at different irrigated lowland rice growing areas by network analysis

Sith Jaisong

October 22, 2015

Abstract

Knowledge of the network structure of rice pest injury co-occurrence patterns helps to understand the formation and characteristics of injury profiles at different rice growing areas in South and Southeast Asia. Therefore, survey data of 420 rice fields over three years across 5 countries, India, Indonesia, Philippines, Thailand, and Vietnam were investigated. The aim of the study was to use network analysis to characterize the patterns of pest injury co-occurrence in specific locations and to determine the key pests using the properties of the network. The results of the weighted co-occurrence networks indicated that ???. Therefore, network analysis provides insights into the pattern of co-occurrence which could be of particular interest for the identification of key factors with regard to different rice growing locations. ....

Introduction

In nature, growing plants can frequently be infected at the same time by more than one species of pests and pathogens. These plant enemies caused injuries, which possibly affect yield productions when they are severe. Crop health is highlighted and implemented to manage a combination of injuries, so-called injury profiles. The co-occurrence patterns of injury are beginning to provide important insight into the injury profiles, which possibly present co-occurring or anti-co-occurring relationships between injury-injury. Uncovering these patterns might have important to implications in plant disease epidemiology and management. However, there are only a few reports of injuryinjury interactions in crop systems and the mechanisms of interactions are currently unknown and this could be a difficult task since complex patterns of injury profiles are related to environmental conditions, cultural practices, and geography (Willocquet et al., 2004).

To address this issue, we use in-field surveys as a useful tool to develop ground-truth databases that allow one identify actual constraints due to pests in an agricultural productions system. These sorts of databases provide an overview of the complex relationships between the crop, its management, pest injuries, yields. Several previous studies (Savary et al., 2000b,a, 2005; Dong et al., 2010; Reddy et al., 2011) involved surveys that have been used to identify relationships in an individual production situation (a set of factors including cultural practices, weather condition, socioeconomics, *etc* that determine agricultural production) and the injury profiles using nonparametric multivariate analysis such as cluster analysis, correspondence analysis, multiple correspondence analysis. Performing correspondence analysis (Savary et al., 1997), they characterized the relationships between categorized levels of variables: actual yield, production situations, and injuries profiles. For example, stem rot and sheath blight are frequently found together with high (mineral) fertilizer inputs, low pesticide use, and good water management in transplanted rice crops and overall, their results led to the conclusions that observed injuries profiles were strongly associated with production situations and the level of actual yields.

We applied the technique from ecological study the co-occurrence analysis and network theory to reveal the patterns of co-occurrence of injury profiles. While existing statistic approach for analysis the survey data e.g. cluster analysis, correspondence analysis, multiple correspondence analysis, the methods were present in this papers has the advantages of ....

Network theory is the study of relationships between entities ('nodes') and connections between these entities ('edges'). Network theory has previously been used effectively to describe social and biological datasets, and it has been shown to be a useful tool for ???. Here, we consider pest injuries as nodes and create an edge between any two injuries if they are co-existing. We give an edge greater weight if the two injuries have strong co-occurrence at either end. The relationships will be more complex when the number of their components increased. A way to systemically model and intuitively interpret such relationships is the depiction as a graph or network. This approach has been widely used and proven very useful in biological studies (Moslonka-Lefebvre et al., 2011). Networks typically consist of nodes, usually representing components, while links between the nodes depict their interactions (Proulx et al., 2005). A correlation network is a type of network in which two nodes are connected if their respective correlation lies above a certain threshold. The construction of this network is obtained from pairwise correlation methods (Toubiana et al., 2013). By using appropriate correlation measure, correlation networks can capture biologically meaningful relationships, and discover valuable information in crop health surveys.

The aim of this study was to analyze how the structure of correlation network of pest injury co-occurrence are different over 5 locations under investigation (West Java, Indonesia, ??, India, Laguna, Philippines, Central plain, Thailand, and Makhong river delta, Vietnam). By quantifying the important aspects of the position of the specific pest injury, information based on network analysis could help to understand the formation and characteristics of co-occurrence patterns. Furthermore, key factors for could be identified with this additional information.

Materials and Methods

Study sites, sampling and data collection

We conducted the surveys located in the South and South East Asia, Kerala, India(Lat , Long), Indonesia (Lat , Long), Philippines (Lat , Long), Central Plain, Thailand (Lat , Long), and Mekong Delta Vietnam (Lat , Long). Theses are the important rice growing areas, where use irrigated lowland rice ecosystem. intensive condition, which grow twice per year. We sampled in , under same standardized protocol described in the IRRI publication, "A survey portfolio to characterize yield-reducing factors in rice", was used for data collection (Savary and Castilla, 2009).

Injury variables were also simplified. Although a very large number of pathogens, insects, and weeds are harmful to rice, many are seldom considered to cause yield losses. Diseases such as narrow brown spot, bacterial leaf streak, leaf scald, and leaf smut, and insects such as rice bugs, rice hispa, and defoliators in general are not considered to represent major, widespread, yield-reducing factors. The study therefore concentrated on injuries listed in Table 1. A second aspect pertains to the injury mechanisms, and Table 1 includes injuries that were grouped in the field assessment procedure according to their nature: light stealers (BLB, BS, LB: proportion of injured leaves), senescence accelerators (BLB, SHB, LB: proportion of injured leaves, except for SHB), tissue users (leaves: RWM, LF: proportion of injured leaves; tillers: SR, SHB, DH: proportion of injured tillers; panicles: SHR, WH: proportion of injured panicles), assimilate sappers (PH: number of insects sampled), turgor reducers (at the tiller level: SR, SHB: proportion of injured tillers; at the panicle level: NB: proportion of injured panicles), and stand reducers (WA and WB)

Crop health survey data were collected through surveys comprising 420 farmers' fields from 2010 to 2012 for wet and dry seasons in different production en-

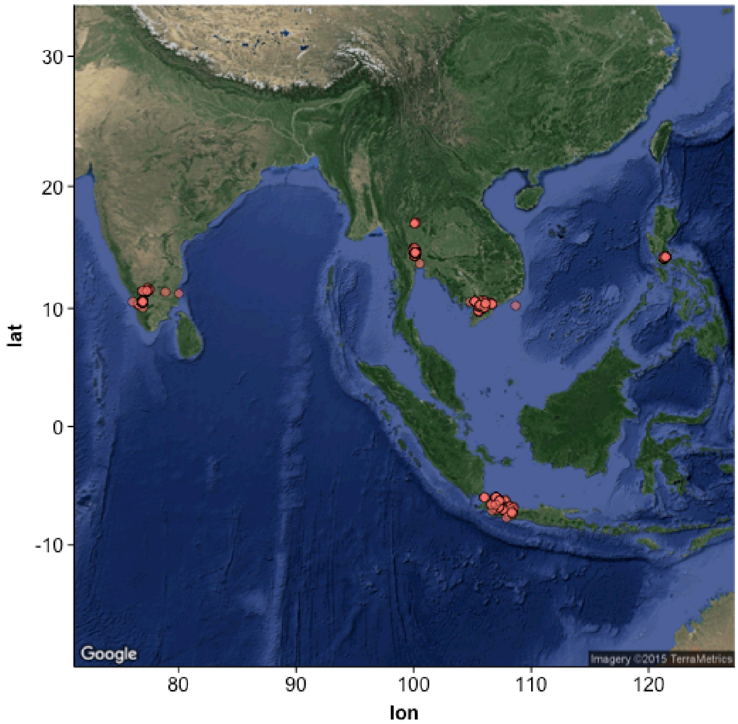

Figure 1. Survey sites

vironments across South and South East Asia. The survey protocol described in the IRRI publication, "A survey portfolio to characterize yield-reducing factors in rice", was used for data collection (Savary and Castilla, 2009). The variables collected included patterns of cropping practices, crop growth measurement and crop management status assessments, measurements of levels of injuries caused by pests, and direct measurements of actual yields from crop cuts. The data collected can be classified into three groups: cropping practices, injuries, and actual yield measurements.

Co-occurence analysis

We considered co-occurrence both of positive and negative correlations based on Spearman's rank based correlation between paris of pest injuries within each dataset with the strength of relationship (ρ from Spearman's correlation) represented by the correlation coefficient. The coefficients with p-values less than p = 0.05 were considered. Negative correlations (indicative of anti-co-occurrence relationship) were also included in analysis. The `R` function `cor.test` with parameter method 'Spearman' (package stats) was used for calculate Spearman's correlation coefficient (ρ), which is defined as the Pearson correlation coefficient between the ranked variables. These correlation relationships were generated for each pair of injury within each location replicate as long as both injury had incidence value greater than 0. We made a network of co-occurrence relationships within each locations based on the strength of the correlation (ρ from the Spearman's correlation), and co-occurrence relationships were only included if they occurred across all locations. Though this method has been illustrated to produce some spurious co-occurrence relationships among data, this rank-based correlation statistic does not require any transformation of variables to fit assumptions of normality and may outperform Pearsons correlations. To increase our level of stringency that may reduce the appearance of spurious co-occurrences within our networks, pairwise relationships had to be consistent across all datasets of a given ecosystem type, greatly reducing the number of co-occurrence pairs.

Network analysis

Network models were illustrated the co-occurrence patterns of pest injuries within same locations, where injuries represent nodes and the presence of a co-occurrence relationship based on correlation is represented by an edge. by igraph package where each network was the union of positive or negative correlation coefficients (less than 0.25 or greater than 0.25) that were consistent within each location.

We were also interested in generating statistics that describe the network that may be important for understanding co-occurrence relationships. We produced network statistics that describe the position and connectedness of injuires within each co-occurrence network. Global network properties including the density, heterogeneity, centralization were computed by using `fundamentalNetworkConcepts` function from WGCNA package and for the basic properties such as number of nodes, edges can be computed by using functions from igraph package. Addtionally, we also calculate the smallworldness index of the network by using `smallworldness` function qgraph package. For the node-wise peopterties including node degree, which is the number of co-occurrence relationships that an injury is involved in a network using the `degree` function from igraph package. We also calculated betweenness scores for each node (injury) using the `betweenness` function from igraph package, which is defined by the number of paths through a focal microbial node. Additionally, we calculated clustering coefficients, and eigenvector using the `transitivity` function for comparison to other networks.

Key factor analysis

Identification key factors in a network is very useful and widely used in social science. The way to identify key factors is to compare relative values of centrality such as eigenvector centrality and betweenness. Its apparent that many measures of centrality are correlated **?**. The residue of linear relationship between igenvector centrality and betweeness and regress betweeness on eigenvector centrality can used as an indicators of key factors. A node or factors with higher levels of betweenness and lower eigenvector centrality can be inferred that is central to the functioning of the network. Nodes with lower levels of betweeness and higher eigenvector centrality may be inferred that they are key to the functioning of the network.

Results

	country	Node	Link	diameter	Connect	Geodesic	Density	SW	CENT	HETERO
1	PHL	21	33	1.85	0.28	2.26	0.07	0.82	0.19	0.89
2	VNM	20	47	1.24	0.53	2.05	0.08	1.27	0.08	0.47
3	THA	18	63	1.11	0.47	1.81	0.15	1.07	0.20	0.70
4	IDN	22	48	1.40	0.46	2.25	0.06	1.24	0.07	0.64
5	IND	10	17	1.38	0.55	2.02	0.14	1.71	0.14	0.51

0.1 Indonesia

Discussion

References

Dong, K., Chen, B., Li, Z., Dong, Y., and Wang, H. (2010). A characterization of rice pests and quantification of yield losses in the japonica rice zone of yunnan, china. *Crop Protection*, 29(6):603–611.

Moslonka-Lefebvre, M., Finley, A., Dorigatti, I., Dehnen-Schmutz, K., Harwood, T., Jeger, M. J., Xu, X., Holdenrieder, O., and Pautasso, M. (2011). Networks in plant epidemiology: from genes to landscapes, countries, and continents. *Phytopathology*, 101(4):392–403.

Proulx, S., Promislow, D., and Phillips, P. (2005). Network thinking in ecology and evolution. *Trends in Ecology & Evolution*, 20(6):345–353.

Reddy, C. S., Laha, G. S., Prasad, M. S., and Krishnaveni, D. (2011). Characterizing multiple linkages between individual diseases, crop health syndromes, germplasm deployment, and rice production situations in India. *Field Crops.*

Savary, S. and Castilla, N. (2009). A survey portfolio to characterize yield-reducing factors in rice. *IRRI Discussion Paper No 18*, page 32.

Savary, S., Castilla, N. P., Elazegui, F., and Teng, P. S. (2005). Multiple effects of two drivers of agricultural change, labour shortage and water scarcity, on rice pest profiles in tropical asia. *Field Crops Research*, 91(2):263–271.

Savary, S., Elazegui, F., Pinnschmidt, H., Castilla, N., and Teng, P. (1997). A new approach to quantify crop losses due to rice pests in varying production situations. *IRRI discusion paper series no20. International Rice Research Institute, PO Box*, 933.

Savary, S., Willocquet, L., Elazegui, F. A., Castilla, N. P., and Teng, P. S. (2000a). Rice pest constraints in tropical Asia: quantification of yield losses due to rice pests in a range of production situations. *Plant disease*, 84(3):357–369.

Savary, S., Willocquet, L., Elazegui, F. A., Teng, P. S., Van Du, P., Zhu, D., Tang, Q., Huang, S., Lin, X., and Singh, H. M. (2000b). Rice pest constraints in tropical Asia: characterization of injury profiles in relation to production situations. *Plant Disease*, 84(3):341–356.

Toubiana, D., Fernie, A. R., Nikoloski, Z., and Fait, A. (2013). Network analysis: tackling complex data to study plant metabolism. *Trends in Biotechnology*, 31(1):29–36.

Willocquet, L., Elazegui, F. A., Castilla, N., Fernandez, L., Fischer, K. S., Peng, S., Teng, P. S., Srivastava, R., Singh, H., Zhu, D., et al. (2004). Research priorities for rice pest management in tropical asia: a simulation analysis of yield losses and management efficiencies. *Phytopathology*, 94(7):672–682.

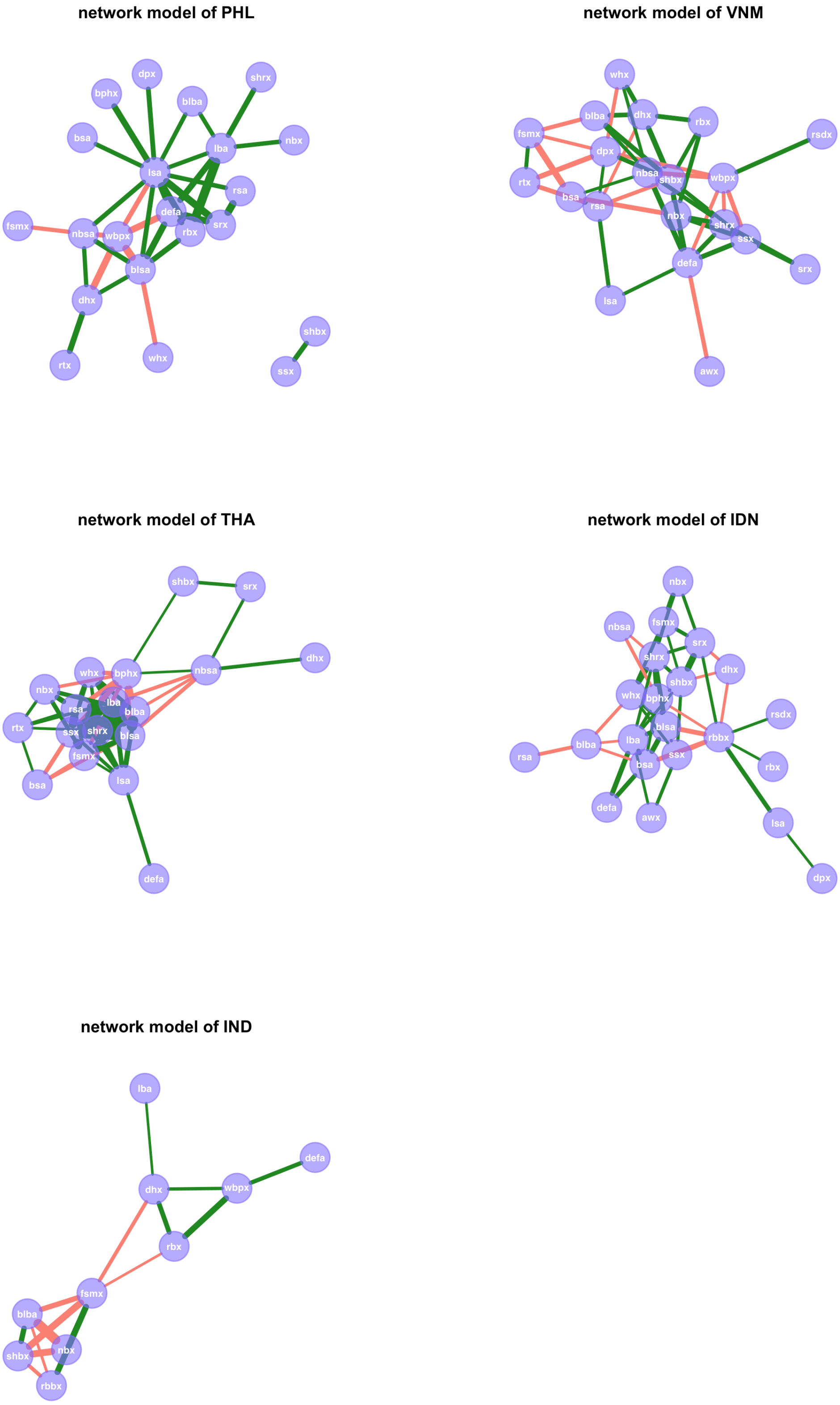

Figure 2. Enter the caption for your figure here. Repeat as necessary for each of your figures