\subsection{Introduction}

Agricultural crop plants are frequently injured, or infected by more than one species of pests and pathogens at the same time. Many of these injuries may affect yields. Because of this co-occurrence in injuries, the idea of ``crop health" has been highlighted and implemented to manage the combination of injuries or so-called injury profiles \citep{Savary\_2006\_Quantification}. Co-occurrence patterns of injuries are beginning to provide important insight into these injury profiles, which possibly present co-occurring or anti-co-occurring (mutually exclusive) relationships between injury-injury. Uncovering these patterns is important to implications in plant disease epidemiology and management. However, there are only a few reports of injury–injury relationships in rice crop systems are currently unknown. This could be a difficult task since complex patterns of injury profiles are related to environmental conditions, cultural practices, and geography \citep{Willocquet\_2008\_Simulating}.

To address this issue, we used in-field surveys as a tool to develop ground-truth databases that allowed us to identify the major yield reducing pests in irrigated lowland rice ecosystems. These sorts of databases provide an overview of the complex relationships between crop, field management, pest injuries, and yields. Several previous studies \cite{Savary\_2000\_Characterization,Savary\_2000\_Quantification,Dong\_2010\_Characterization} and \cite{Reddy\_2011\_Characterizing} involved surveys that were used to characterize injury profiles in an individual production situation (a set of factors including cultural practices, weather condition, socioeconomics, \textit{etc}.) that determine agricultural production, and the injury profiles using nonparametric multivariate analysis such as cluster analysis, correspondence analysis, or multiple correspondence analysis. Their results led to the conclusions that injury profiles (the combination of disease and pest injury that may occur in a given farmer’s field) were found co-occurrence patterns across sites, which are associated at regional scale. For example, stem rot, sheath blight, planthopper, and rice whorl maggot injuries, are high incidence, with low incidence of brown spot, and absence of bacterial leaf blight, leaf blast, and neck blast are a common pattern in tropical Asia from the study of \citet{Savary\_2000\_Characterization}.

Co-occurrence analysis and network theory have recently been used to reveal the patterns of co-occurrence between microorganisms in the complex environments ranging from human gut to ocean and soils \citep{Faust\_2012\_Microbial\_co,Ma\_2016\_Geographic}. Co-occurrence patterns are ubiquitous and particularly important in understanding community structure, offering new insights into potential interaction network. Recent reviews of network based approaches reveal that these tools have demonstrated previously unseen co-occurrence patterns, such as strong non-random association, topology based analysis of large networks has been proven powerful for studying the characteristics of co-occurrence pattern of the communities in ecological community \citep{Williams\_2014\_demonstrating, Barberan\_2012\_Network}, or key actors in social networks \citep{Crowston\_2006\_Hierarchy}. Here, we significantly advance this study by providing a comprehensive understanding of the topological shifts of animal pest injury and disease co-occurrence networks at regional scale.

South and Southeast Asia represent big bowl of rice for the world population. Comparing the topological properties of the node associated with occurrence in the different countries and examining network level topological features can provide us with insight into variation in the co-occurrence patterns of rice injuries in different countries. This approach helps contextualize the animal pests -- disease association by taking to account the complex network of potential association among animal pest and disease occurring in farmers’ fields in these countries. Specifically, we addressed the following questions:(i) How can the co-occurrence relationships of rice injuries be examined from the perspective of network analysis (ii) Which animal injuries and diseases are found commonly close co-occurrence patterns among other variables in order to target to control or monitor. To answer these questions, we performed crop health survey at the farmers’ fields in two different seasons and five countries in South and Southeast Asia and implemented co-occurrence network analysis to examine the topological feature differences across countries. Our main objective was to characterize and better understand co-occurrence networks in the association of rice animal pests and diseases.

%As the size and extent of biological data sets grow, scientists turn to new techniques, such as network analysis, to understand biological complexity over large scales. For network analysis of microbial datasets, topological ‘co-occurrence’ networks are generated from correlative metrics, where nodes represent observed variables, and significant correlations are represented by the edges connecting them. We used an unprecedented

\material and methods

I designed a statistical approach written in R v. 3.0.1 \cite{}. All scripts necessary to replicate this analysis are included in the appendix. The analysis presented in this chapter is designed to contract network models of co-occurrence patterns of rice injuries at different levels across cropping seasons (wet, and dry season), and production environments (summarized in Figure?). I considered co-occurrence to be positive rank correlations (Spearman’s correlation) between pairs of injures within each dataset with the strength of the relationship represented by the correlation coefficient (Figure 1B). We only considered negative and positive co-occurrence relationships based on strength of correlation ($\rho$ from the Spearman’s correlation) at values (p-value < 0.05)

\subsubsection{Network construction}

The co-occurrence network was inferred based on the Spearman correlation matrix constructed with \texttt{R} function \texttt{cor.test} with parameter method `Spearman' (package stats) was used for calculate Spearman's correlation coefficient ($\rho$), which is defined as the Pearson correlation coefficient between the ranked variables \cite{R\_2015}. The nodes in this network represent injuries and the edges that connect these nodes represent correlations between injuries. Based on correlation coefficients and $P$-values for correlation, we constructed co-occurrence networks.The cutoff of $P$-values was 0.05. Network properties were calculated with the \textbf{igraph} package \citep{Csardi\_2010\_igraph}. All farmers' fields were divided into groups by country and season. The impact of each sample group on the Spearman correlation value of each edge in the network was assessed by Spearman correlation value of these fields. The network of each of group was detected the community structures by maximizing the modularity measure over all possible partitions by using \texttt{cluster\\_optimal} function of \textbf{igraph} package \cite{Brandes\_2008\_Modularity}.

\subsubsection{Topological feature analysis}

We calculated the topological features for each network with the \textit{igraph} package. We measured two levels of network topologies. The node features that we focused on are node degree, betweenness, and clustering coefficient. Node degree is measured by the number of the edges (connections) of a node has. Betweenness of a node is defined by the number of of shortest paths going through a node, and the local clustering coefficients of a node is the ratio of existing edges connecting a node's neighbors to each other to the maximum possible number of such edges.

Global features including network clustering coefficient, average path length, and diameter were measured for each network. The network clustering coefficient measures the degree to which nodes of the network tend to cluster together and is a measure of the connectedness of the network and is indicative of the degree of relationships in the network. Average path length is the average number of steps along the shortest paths for all possible pairs of network nodes, and diameter is the greatest distance between any pair of nodes. In our analysis, both diameter and average path length are considered measures of the size of the network. Larger networks are less connected, meaning that the likelihood of a strong connection between any two randomly selected species is low.

The network clustering coefficient are considered measures of the complexity of the network. The networks are more complex, the network has higher clustering coefficient, and shorter average path length.

Node were further classified by ranking all nodes according to three node features, partitioning this ranked list into three equally value of each node property. Nodes with high rank value in top third proportion of node degree, and high rank value in top third proportion of betweenness are recognized as indicator in co-occurrence network of rice injuries.

\newpage

\textbf{Result}

\textbf{Prevalence of injuries across sites and seasons}

In the previous chapter,

Differences among sites and sea- sons for injury prevalence (percent fields affected by a given injury; 37) were sum- marized (Table 3). Among the injuries caused by pathogens, SHB showed the highest prevalence, exceeding 50% in any site–season combination. SHR, BS, and SR occurred in decreasing order of prevalence level. At the other end of this spectrum, RTD was observed in one site and one season only, and thus was not further considered in the analyses. Insect injuries appear to have higher prevalence than those due to pathogens. Most insect injuries were omnipresent, often with prevalence exceeding 80%. Weed infestation, both above and below the rice crop canopy, was omni- present and had the highest prevalence levels of all injuries. This crude summary also suggests that differences as well as similarities occurred among sites and sea- sons with respect to occurrence of injuries.

% discussion

Rice injuries were found commonly in South and South east Asia, but at different levels of incidence.

Some injuries of this study is relatively low prevalence of areas that have been reported such as leaf blast, brown spot. It could be implied that these injuries strongly depended on locations or climatic conduction to develop, so they were not observed at all locations or seasons during survey were conducted. Another reason is that there are some factors such as the utility of resistance verities in the farmer’s fields surveyed, so we could observe those injuries at low level of incidence. The similar reasoning could explain to many injuries such as BLB, RB, NB, which widely occur in rice growing areas.

From the survey data of rice injuries observed in farmers’ fields, I analyzed the interaction and build the network based on that data. The methods applied for building a co-occurrence network of rice injuries were adapted from ecological studies. Usually, relationships were assessed using Pearson correlation. However, the use of the Pearson correlation coefficient is problematic because it requires the variables are applied with similar measure, and the variable values are normally distributed. Additionally, Pearson correlation can only capture linear relationships. Due to the fact that the assumptions of Pearson correlation are not fit with the survey data. The alternative is provided by using Spearman’s rank correlation coefficient, which is also widely used in ecological studies.

The exploration of co-occurrence networks is a useful method for determining interactions of co- occurring injuries. Network analysis has also suggested important injuries in networks. The important injuries were selected from the node features such as node degree, clustering coefficient, and betweenness. The betweenness represents the importance of the control potential that an injury exerts over the associations of other injuries in the network. The clustering coefficients are indicative of the potential spreading of the incidence of injuries through the network. As activated injury can activate other injuries, a more densely connected network facilitates injury activation \cite{Williams\_2014\_demonstrating}. In the network of dry season in West Java, Indonesia, BLB and SR can be the targets to be monitored because they have high betweenness, which indicated that they are more likely to present then others injuries.

It is good attempt to detect communities in a network because it can reveal information about the networks that is maybe not easy to detect by simple observation. The communities are groups of nodes that are densely connected among their node members, and slightly connected with the rest of the network. In this study, we detected node community based on the optimization of the modularity of a sub-network, which is an approach that widely applied in many fields \cite{Liu\_2014\_Detecting}. Even though, the groups of injury profiles from this study are different from seasons and countries, but they are reasonable because some rice injuries present in some seasons (seasonal occurrence) such as gall midge injury \cite{Krishnaiah\_2004\_Rice}. The result also showed similar groups of injuries to the patterns of injuries profile from the study of \cite{Savary\_2000\_Characterization}.

%SR is also best to be the target because it also associated with the RS, which has high clustering coefficients also because RS is potentially able to be co-found with many other injuries. BLB also show the association with the injuries, which present high clustering coefficient. Both of BLB and SR are intermediate clustering coefficient., so they are formed concurrence between groups, but less potentially to present complex association. DH and WH are less associated with other injuries, so they are difficult to employ other injuries to.

It is the good to monitored WM and RB and SHB also even the GLH has no betweenness, but it has high clustering coefficients, and connected with two high betweenness node. It would be also good target to monitored also because when it present, it also frequency appear together.

We explored the characteristics of rice injury profiles

เราเองก็ยังไม่รุ้เลยว่า การที่ จัดกลุ่มขึ้นมาได้นั้น มันเป็นเพราะว่าอไร รู้แต่ว่า มัน close realtion

Communities are groups that are densely connected among their members, and sparsely connected with the rest of the network. Community structure can reveal abundant hidden information about complex networks that is not easy to detect by simple observation.

With regard to pest management development, understanding that the relationships of rice injuries could simplify pest management decisions. As applied here, the node degree, clustering coefficients, and betweenness described the interaction of injuries in the networks. Therefore, they may be useful in finding a good indicator for monitoring pest in rice fields. High-betweenness node High- clustering coefficient node

Community structure in complex network can reveal hidden information about complex networks that is maybe not easy to detect by simple observation such as nodes, which are clustered. In this study, we detected node community based on the optimization of the modularity of a sub-network, which is a popular approach \cite{Liu\_2014\_Detecting}. Even though, the groups of injury profiles from this study are different from seasons and countries, but they are reasonable because some rice injuries present in some seasons (seasonal occurrence) such as gall midge injury \cite{Krishnaiah\_2004\_Rice}. from the groups of injury profile from \cite{Savary\_2000\_Characterization}.