A SURVEY PORTFOLIO TO CHARACTERIZE YIELD-REDUCING FACTORS IN RICE[1]

S. Savary & N.P. Castilla
Plant Breeding, Genetics, and Biotechnology Division, IRRI

[1]This document is an updated version of :

Savary, S., Elazegui, F.A. and Teng, P.S., 1996. A survey portfolio for the characterization of rice pest constraints. IRRI Discussion Paper No 18. 32 p.

© IRRI, 2009

creative commons

This discussion paper is copyrighted by the International Rice Research Institute (IRRI) and is licensed for use under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License (Unported). Unless otherwise noted, users are free to copy, duplicate, or reproduce, and distribute, display, or transmit any part or all of this publication, and to make translations, adaptations, or other derivative works under the following conditions:

ATTRIBUTION. The work must be attributed, but not in any way that suggests endorsement by IRRI or the author(s).

NON-COMMERCIAL. This work may not be used for commercial purposes.

SHARE ALIKE. If this work is altered, transformed, or built upon, the resulting work must be distributed only under the same or similar license to this one.

NOTES:
- For any reuse or distribution, the license terms of this work must be made clear to others.
- Any of the above conditions can be waived if permission is obtained from the copyright holder.
- Nothing in this license impairs or restricts the author's moral rights.
- Fair dealing and other rights are in no way affected by the above.
- To view the full text of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/

Concepts **3**
A few useful concepts 3
Production situation 3
Pests 3
Injury, damage, and loss 3
Potential, attainable, and actual yields 3
Precision vs. accuracy 3
Development of a survey protocol 4

A Protocol to Assess Pest Constraints to Rice Production: Survey procedure and information for assessing injuries in farmers' fields **5**
Principles 5
Developing a sampling strategy and a feasible survey plan 5
General information (Form 1) 6
Landform 6
Approximate field area 6
Crop establishment method 6
Plant material 6
Crop density 7
Zinc deficiency 7
Overall crop status 7
Crop and pest status (Form 2) 9
First passage in the crop 9
Crop development stage and water status 9
Development stage of the crop 9
Water status 10
Second passage in the crop 10
Crop growth 10
Assessment of weed infestation 11
Insect pests 12
Diseases 12
Final remark 12

Simplified protocol with two visits **13**
Principle 13
General information (Form 1) 13
Crop and pest status (Form 2)—at booting stage and at ripening stage 13

List of variables included in the recording forms **14**

A working example **18**
Nature of the variables 18
Compaction over time: injury variables 18
A hypothetical survey data set 18
Categorization: compaction over range 19
Contingency tables and chi-square tests 20
Correspondence analysis 20
Cluster analysis and other analytical approaches 23

Illustrating the concept of recommendation domain in IPM: a case study in Eastern Uttar Pradesh **27**

References **31**

Concepts

Within any geographic area where crop health management is to be improved or adapted to agricultural changes, there is a need for characterization to define domains for research and technology development and subsequent technology extrapolation (Teng and Savary 1992, Savary et al 1994). Characterization requires that reliable, representative field data be available. A common approach is to collect such data using field surveys.

Surveys are a means to characterize constraints to crop production and represent an essential component of systems research in plant protection. The general methodology was reviewed in, e.g., Teng (1987), Campbell and Madden (1990), and Savary et al (1995; 2006).

The rice crop system is so complex that a large number of descriptors is, on principle, required for its characterization. Each rice field can be seen as a unique realization of one combination of many attributes. These attributes encompass a growing crop, its physical environment, its diseases and pests, and farmers' actions, which influence the whole system. The number of rice pests (diseases, insects, and weeds) to be considered and the levels at which they vary from field to field reflect this diversity (Teng 1990, Litsinger 1992). The prime objective of a survey is to provide an adequate account of this diversity; the prime objective of survey data analysis, however, is to resolve this diversity to a degree that allows interpretations.

This document presents a set of field techniques to address this diversity and describes statistical methods to exploit information that is diverse in nature, precision, and accuracy.

A few useful concepts

Production situation

The concept of production situation was introduced by De Wit and Penning de Vries (1982) to describe the set of factors — physical, biological, and socioeconomic — that determine agricultural production. This concept is used here within the restricted scope of an individual rice field and from the limited view point of plant protection. A quick overview of the variables listed in the following survey procedure indicates how this concept is operationalized in a survey: this list includes a few key components of rice crop management (which reflect the physical and socioeconomic environments of a crop and their interactions) and a series of rice pests (that is, insects, pathogens, and weeds).

Pests

By *pest*, we refer here to any agent that may be harmful to a crop, and potentially reduce its yield. The term encompasses pathogens, animals, and weeds.

Injury, damage, and loss

An injury is the visible, measurable result of the biological activity of a pest. An injury may, or may not, translate into a damage i.e., a yield reduction; this yield reduction may, in turn, translate into a loss, often measured in economic terms (Zadoks and Schein 1979, Zadoks 1985). Our prime focus here is represented by the injuries caused by a series of pests. The difficult issue of estimating damage (yield losses) from survey data alone is addressed in the last section of this document.

Potential, attainable, and actual yields

The definitions from FAO (Chiarappa 1971) are used here. The *potential yield* of a given genotype under given climatic (radiation, temperature) conditions is in practice *limited* by a number of factors in a farmer's field: the supply of water or nutrient may not coincide with the needs of the plants at a particular development stage. This *attainable yield* may further be *reduced* by environmental factors such as storms, floods, or pests. The result is the *actual yield,* that is harvested, and which can be measured in a farmer's field.

Precision vs. accuracy

Ideally, measurements should be both precise (as consistent as possible over successive samples) and accurate (as close as possible to the "true" value [Forbes and Jeger, 1987]). In the context of a survey, the

accumulation of superfluous observations often is counterproductive: as in any tedious field work, *accuracy* has been shown to sharply decline with the number of successive observations within a sampling unit (i.e., within a field), i.e., *precision*. A balance must be found between *accuracy*, which is essential, and *precision*, which can be needed, but in a second stage only. One should also consider the cost of characterization research, which is among the most expensive activities in agricultural research. Accuracy must therefore come first, and is best achieved with comparatively few observations per statistical unit (i.e., fields), which in turn can be distributed over the largest possible number of farmers' fields (statistical units of a survey) so as to increase the survey representativeness. One must also realize that true limits exist to the precision of some measurements: for instance, coefficients of variation in the range of approximately 20% are the norm for farmers' yield estimates in surveys on cereals (Church and Austin 1983).

Development of a survey protocol

The survey protocol described below takes into account the above concepts, and operationalizes them. Operational definitions (Zadoks and Schein 1979) — the translation of concepts into a series of specific steps leading to measurements — are sometimes difficult to establish (Butt and Royle 1978). A clear definition of objectives can partly alleviate such difficulties. The survey protocol described here aims at:

(1) characterizing rice cropping practices,
(2) quantifying injuries, and
(3) measuring links between cropping practices, pest injuries, and actual rice yields.

Our task was considerably simplified by the pioneering work done by Elazegui et al (1990) in the Central Plain of Luzon, Philippines.

The development of a survey portfolio, as here, follows successive steps that are similar to the development of a model:
1. A detailed, preliminary, field survey procedure is established;
2. This procedure is used in one reference site for a series of cropping seasons;
3. A set of analytical techniques is tested to (a) exploit the acquired information, and (b) select statistical approaches that best fulfill the objectives of the characterization work;
4. The field survey procedure is then both revised and simplified in view of (a) its use at new sites, (b) reliance on fewer observers per site, and (c) stability towards differences amongst field observers across sites;
5. The analytical techniques are reviewed so as to best address objectives both on site-specific data sets as well as on combined data sets across different sites.
This document pertains to the end-product of the above steps.

The methodology described below can be used for several purposes in rice research:

- identify the main yield reducing factors in a range of production situations (Du et al., 1997; Elazegui et al., 1990; Sta Cruz et al., 2001; Savary et al., 1997; 2000a; 2000b);
- contribute to quantify yield losses caused by individual or combined harmful organisms to rice (Savary et al., 2006);
- guide crop-health research (Zadoks et al., 1995) and the deployment of crop-health technologies.

It has been extensively used in the Philippines, in India, in Vietnam, and in China. One example, derived from work conducted in India is used at the end of this document as an illustration. The original analytical framework used then is described in some detail. Since then, new statistical tools have been developed that can be very powerful in analyzing data generated from survey work.

A Protocol to Assess Pest Constraints to Rice Production: Survey procedure and information for assessing injuries in farmers' fields

Principles

This survey protocol is meant to address the individual farmer's field. The system being considered here is a farmer's field and it encompasses components such as the crop at its various development stages, the farmer (and field operations such as crop establishment techniques, crop husbandry practices, and inputs), and the pests that may affect the standing rice crop throughout its cycle, from crop establishment to harvest. In this survey procedure, we call pest any harmful agent that may reduce yield: weeds, insects, and pathogens.

While the focus of this protocol is on the individual farmer's field, any survey must address an adequate number of fields, so that analysis is made possible, hypotheses formulated, and interpretations forwarded. While, therefore, the sampling unit of a survey using this protocol consists of a field, this protocol must be applied to a sufficient number of fields to achieve the objectives of a survey. When developing this document, great care has been taken to achieve a reasonable balance between amount of information required at the individual field level (and achieve accuracy at the individual field level) and the amount of work (observations) that each field represents, so as to allow a sufficiently large number of fields to be surveyed (and achieve adequate survey representativeness).

The objectives in using this protocol are:

1. To characterize the patterns of cropping practices and, more generally, achieve a reasonable description of production situations (the set of environmental factors that may determine actual yield of a rice crop). This must include, of course, field operations;
2. To characterize the combinations of pests that may occur in any particular field;
3. To establish links between production situations and pest combinations;
4. To generate a measure of the links between production situations, pest combinations, and variation in actual yield.

Three types of results may thus be expected from a survey conducted using this protocol:

- Determination of the relative importance of rice pests – frequency and intensity of injuries – in the surveyed area;
- Information on patterns of cropping practices that may potentially lead to pest problems;
- And thus, baseline information for the delineation of integrated pest management recommendation domains, based on production situations, and more specifically, on patterns of cropping practices.

Strong emphasis is therefore given to information describing cropping practices as part of production situations as a whole. Only when production situations have adequately been characterized can injuries be addressed and analyzed. This is why, for each field, the survey involves two components which both address cropping practices and injuries: a general information form and a crop and pest status form.

Developing a sampling strategy and a feasible survey plan

Every effort has been made to make this survey procedure as easy, flexible, and time-saving as possible, while at the same time, efficient in generating adequate representations of individual fields. The effort needed to use this procedure should not, however, be under-estimated. For a well-trained team of four observers, it takes approximately 30 minutes to accomplish observations in a farmer's field and to fill the survey form. Time is also needed to meet and interview farmers, especially in the beginning of the cropping season, when fields are selected for the survey, and later in the crop cycle, to ascertain which field operations have been conducted and when.

At the beginning of the cropping season, fields are to be selected for the survey. These fields should best represent the diverse cropping practices encountered in the target survey area. It is suggested that a strategy in two phases be used:

In a first phase, a few villages (two may be enough) might be carefully selected as representative of villages of the site to be surveyed. In each village, 7 to 10 fields may then be chosen as representative of the diversity of cropping practices and environments that prevail in each village. This stratified sampling method is a good way to save considerable amount of time. This first phase, lasting a whole cropping season, may be called a 'pilot' phase, where the survey methodology is being tested by observers. Its overall result is a relatively small sample of fields (15 - 20), and experience.

In the second phase, the sample size (number of surveyed fields) has to be increased to 80-100 per cropping season. There are two ways to achieve this: increasing the number of villages where fields are selected for the survey, or increasing the number of fields per village. Experience shows that if the target survey area is relatively homogeneous, there is often relatively little benefit in increasing the number of villages beyond 3 or 4. Often, a good option is to increase the number of fields per village to 20-30, and have 3 - 4 villages surveyed. This results in having the surveyed fields clustered in a few villages. This saves a large amount of time, while achieving adequate survey representativeness.

The survey considers each field as a unique individual (a sampling unit), where yield is considered as the outcome of a series of factors. Some of these factors do not vary during the cropping season (e.g., land type or planting date), whereas others do. More specifically, each individual field may be considered to be exposed to a series of pests, diseases, and weeds, each of them with a specific pattern of variation over time. Therefore, information on the crop and on the biotic constraints are to be gathered throughout the cropping season. Assessments are to be made at four main development stages:

- tillering (20-30 days after transplanting or seeding)
- booting
- early dough
- maturity

At the end of the survey, each surveyed field should therefore be represented by five forms, the first providing general information, and the additional four representing the four successive visits where crop status and weed, disease, and insect injuries have been assessed.

General information (Form 1)

Field number: in each village, a sample of fields is to be taken and each of them assigned a code number. It is very useful to locate and indicate each field position in the village and number it separately.

Landform

Each field should be assigned to one of the possible landform categories.

Approximate field area

The size of the field should be, as much as possible, representative of farmers' field area. A gross estimate is sufficient, provided that more precision is given when fields of smaller area are considered — e.g., 0.1, 0.3, 0.5, 0.8, 1, 1.5, 2, 3 ha.

Crop establishment method

As much details as necessary should be provided. For example: transplanted, dry seeded, or wet seeded rice. Information about specific cropping practices early in the cropping season such as beushaning is also very useful.

Plant material

The rice variety or hybrid needs to be specified. When the rice variety is a 'local' or 'traditional' one, additional information is needed: name, expected crop cycle duration, a brief description of the plant, its habit, and grain type.

Crop density

The basis for estimating rice crop density in both transplanted and direct seeded rice is the number of tillers per square meter. The density of tillers per square meter provides a measure of potentially yield-contributing units. In direct seeded rice, this information can be derived directly from the average number of tillers per 10x10 cm quadrats (this information is to be taken on 10 sampling points at each visit, see below). In transplanted rice, the number of tillers per square meter can be derived from the number of tillers in individual hill (this information taken on 10 hills at each visit, see below), and the number of hills per square meter. Therefore, the number of hills (transplanted) or plants (direct seeded) per square meter only needs to be entered in the first form (general information). These estimates should be made at tillering stage, using a one-square meter quadrat (transplanted), or 10 quadrats (10x10 cm, direct seeded) in a portion of the field that is representative of its average status.

Zinc deficiency

Occurrence of zinc deficiency is to be assessed once in the cropping season, at the tillering stage. The rating system combines four classes of areas (0 %, more than 0 and less than 10 %, more than 10 and less than 25 %, and more than 25 %), and three classes of symptom intensity (light, moderate, and severe). The rating scale is as follows:

0: No zinc deficiency
1: More than 0%, less than 10% field area affected, light symptoms
2: More than 0%, less than 10% field area affected, moderate symptoms
3: More than 0%, less than 10% field area affected, severe symptoms
4: More than 10%, less than 25% field area affected, light symptoms
5: More than 10%, less than 25% field area affected, moderate symptoms
6: More than 10%, less than 25% field area affected, severe symptoms
7: More than 25% field area affected, light symptoms
8: More than 25% field area affected, moderate symptoms
9: More than 25% field area affected, severe symptoms

Overall crop status

This variable is intended to give an overall description of the rice crop stand. It may be helpful to derive a categorized estimate of the attainable yield of the field, that is, of the yield the crop would have produced in the absence of yield reduction by pests, diseases, and weeds.

The status of a rice crop can be assessed at many different stages. One reasonable stage to do so is booting stage. This is because this stage coincides with the time when the foliage has been fully grown, and yet has usually not started to senesce; it also is a stage when most of the field operations have translated into a visible effect on the crop; and it also is a stage when most diseases and pests have not yet achieved their highest possible intensity, except a few of them, and except relatively infrequent circumstances that would have favored early and dramatic injuries to be already visible. The table below provides a framework to rate a rice crop status at booting stage.

The assessment of crop status should, as much as possible, disregard the presence of weeds (which may affect the regularity in height of the stand) and of diseases and pests (which may affect its color). One first element to consider is the overall crop canopy structure, from very regular and homogeneous, to irregular with many off-type plants, to very irregular and even some patches of stunted growth. A second element is the average color of the canopy, from uniformly very green, to pale or yellowish, or even with large fraction of plants discolored. A third element to consider is crop density, from complete soil (or flood water) coverage, to incomplete, with gaps in the canopy. It is hard to determine which of the three elements would determine most the crop performances. In the table indicated below, equal weight is roughly given to each of them. The three elements, also, are often associated, thus the arrangement of the table.

Overall crop canopy structure A		Foliage color B		Crop density C		Overall crop assessment and rating (**A + B + C**)	
Description	Point mark	Description	Point mark	Description	Point mark	Category	Total
Regular height, very homogeneous	**5**	Homogeneous dark green	**5**	Full crop closure	**5**	**Very good**	**15**
A few irregularities in height and/or off-types	**4**					**Good**	**12-14**
Fairly regular, some irregularities in height and/or off-types	**3**	Green, homogeneous	**3**	Some gaps (less than 1% of ground cover)	**3**	**Average**	**8-11**
Not regular, irregular crop height and/or many off-types	**2**	Irregular color, or overall pale / yellowing color	**2**	Gaps (more than 1%, but less than 5% of ground cover)	**2**	**Poor**	**6-7**
Very irregular crop height and/or many off-types, or overall stunted growth	**1**	Pale or yellowish, or large fraction of plants discolored	**1**	Many gaps (over 5% of ground cover)	**1**	**Very poor**	**3-5**

Overall, this table should provide the observer with a tool to guide a decision towards declaring a rice crop stand at booting as "very good", "good", "average", "poor", or "very poor". These categories in turn may be associated with attainable yield levels, that is, levels of yield that would be achieved if no measurable yield-reducing effect would result from pest (diseases, insects, weeds) injuries. Only ranges are indicated, that correspond to the proposed five crop status categories. These ranges will vary, and should not be considered prescriptive: A given category of crop status at booting stage will correspond to attainable yields that vary within a range, reflecting numerous effects influencing

- the un-injured yield: the plant genotype, radiation, and temperature;
- and the actual, harvested yield itself: lodging, and accumulation (or appearance) of injuries beyond the booting stage of the considered crop.

Thus, assessing a crop stand should at best be considered a rough estimate of the attainable (un-injured) yield. The procedure indicated above cannot, however, be considered to produce a reliable estimate of the actual (harvested) yield. The information however is very useful as it does contribute to documenting the efficiency of crop management practices.

• *Direct yield measurement:* in each of the fields, three 10 m^2 areas (2 x 5 meters) are to be separately harvested, dried (to 14 % humidity), and separately weighed. The three harvest areas should be taken at random over the field area as indicated in Figure 2.1.

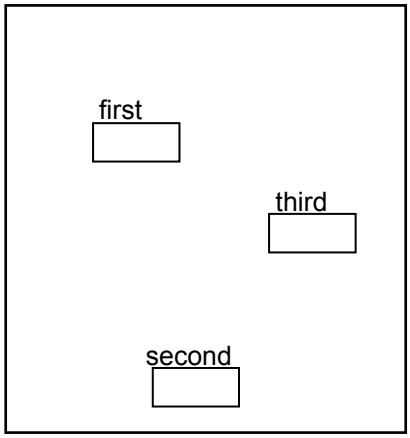


Figure 2.1. Example of random location of three crop cut areas in a field for yield estimation.

Crop and pest status (Form 2)

First passage in the crop

At each visit in the field, the following procedure is recommended: during a first passage in the field, an overall look at the crop is taken, its water status rated, its development stage assessed, and the pests and diseases that are present identified (Figure 2.2). During this first passage, some of the prevailing diseases, insect pests, and weeds are likely to be recognized; nevertheless, as more attention will be paid to individual tillers in the second passage, additional pests and diseases are usually identified. The main purpose of this first passage is, in addition to filling the development stage and water status section of the form, to have an overall impression of the field situation and the changes since the previous visit.

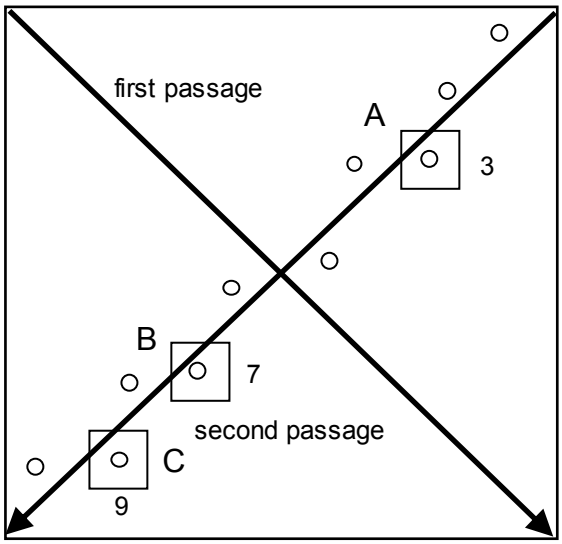


Figure 2.2. Path of sampling in a farmer's field showing 2 passages

Crop development stage and water status

At each visit in the field, two qualitative characteristics of the crop are rated: the development stage of the crop and the water status of the field. These two characteristics can be assessed during the first passage in the field.

Development stage of the crop

The development stage must be based on observation of rice plants in the crop, which are considered to represent the average development of the stand. It is strongly recommended, in the first passage, to look at several plants, assess their development stage, and enter the development stage of the majority of plants in the form. Development stage is assessed using the following scale:

10 : Seedling
20 : Tillering
30 : Stem elongation/panicle initiation
40 : Booting
50 : Heading
60 : Flowering
70 : Milk
80 : Dough
90 : Ripening
100: Fully mature

Water status

The water status of the field is to be rated using the following key:

Without standing water

1: Soil dry and hard
2: Soil moist and hard
3: Soil moist and soft
4: Soil wet and hard
5: Soil wet and soft

With little standing water (water level lower than 5 cm)

6: Soil hard
7: Soil soft

With adequate water (water level between 5 and 15 cm)

8: Soil hard
9: Soil soft

With too much water (water level higher than 15 cm)

10: Soil hard
11: Soil soft

Second passage in the crop

The second passage involves getting quantitative information on crop growth and intensity of injuries due to diseases, insect pests, and weeds. This information is primarily based on 10 individual hills (or quadrats) taken at random (white circles) along the path across the field.

Crop growth

In addition to crop development stage and water status (first passage), three quantitative variables representing crop growth are recorded:

- the number of tillers per hill (or number of plants per 10x10 cm quadrat in direct seeded rice) — see note below,
- the number of panicles per hill or quadrat, and
- the number of leaves per tiller.

These counts are essential for further estimates, and each of them is to be assessed on each of the 10 individual hills or quadrats, which are taken as the basic sample to represent a given field.

On each hill (or plant for direct seeded rice), the number of tillers is counted, and the average number of leaves per tiller is entered. This average number of leaves per tiller should be the rounded mean for five tillers in each sampled hill (see Figure 2.3), or within the 10 x 10 cm quadrat.

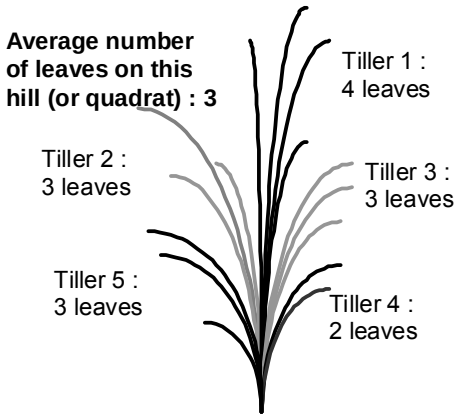


Figure 2.3. Counting leaves and tillers on a rice plant (hill or quadrat).

Note: to assess the number of plants per 10x10 cm quadrat in direct seeded rice, a wire quadrat, 10 x 10 cm can be placed onto the ground by the observer, and the 100 square centimeter area delimited by the quadrat is used for counts and observations.

Assessment of weed infestation. There are many ways to assess weed infestation. It may be based on estimates of

- the height of weeds relative to the height of the rice crop canopy,
- the percentage weed cover above and below the rice crop canopy, or
- the contribution of different weed types to weed infestation.

Three areas within the field, approximately 1 square meter each, are chosen at random along the second path across the field to assess weed infestation. These areas (A, B, and C) are arbitrarily assigned as the areas surrounding hills numbers 3, 7, and 9 of the sampling scheme (Figure 2.2).

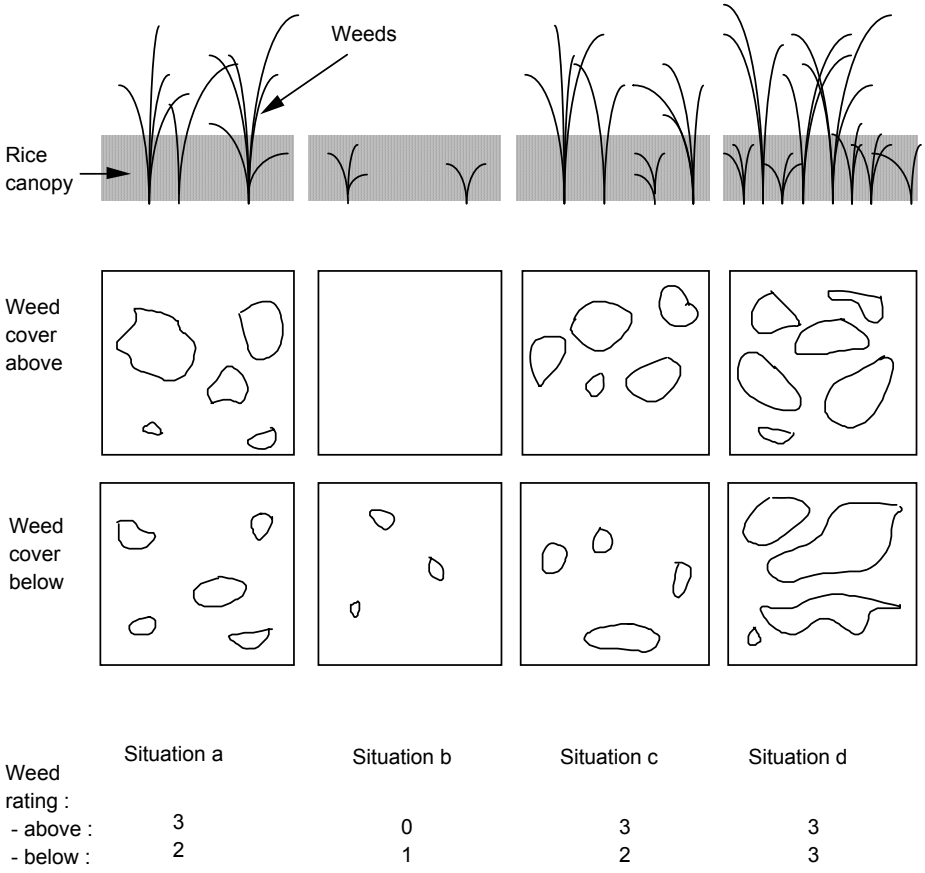


Figure 2.4. A sketched procedure to assess weed infestation.

As mentioned above, the height of the weeds relative to the rice canopy has to be considered. Two ratings have to be made, representing weed infestation above and below the rice canopy (Figure 2.4). These two ratings are based on percent weed cover, the percentage ground area covered by the canopy of weeds, irrespective of rice canopy.

As for the percentage weed cover, the following classes are to be considered:

0 : No weed
1 : Weed cover below 10 % (low)
2 : Weed cover above 10 % and below 30 % (moderate)
3 : Weed cover above 30 % and below 60 % (high)
4 : Weed cover above 60 % (very high)

Using this scale, two ratings are to be given: one pertaining to the fraction of weed canopy growing above the rice crop canopy and the other to the fraction of weed canopy growing below the rice crop (Figure 2.4).

The next information to be entered in the recording form is the nature of the weeds involved in these ratings. Four broad categories of rice weeds can be considered:
S: Sedges; BD: Broad leaf dicots; G: Grassy weeds; SD: Small leaf dicots.

In each of the four areas (A, B, and C) surrounding hill numbers 3, 7, and 9, the frequency of each of the four types of weeds can be entered in the form, by indicating their respective ranks. As an example in area A, sedges (S: 1) might be ranked first, grassy weed (G: 2) second, while there are no representatives of the two other types (BD: 0 and SD: 0). The dominant species (or genera) may be entered in the form to document this information.

Insect pests. Three broad categories of insect pests are considered: insects that reduce the number of (potentially or actually) fertile tillers, insects that reduce the photosynthetically active leaf area, and sucking or grain-damaging insects.

The first category primarily consists of stem borers, represented by the number of deadhearts (DH) per hill or quadrat, and the number of whiteheads (WH) per hill or quadrat. Gall midge also falls under this category. For each hill or quadrat, the numbers of deadhearts and whiteheads are entered in the recording form.

The second category, leaf-feeding insects, is represented by insects such as whorl maggots (WM), leaffolders (LF) and rice hispa (RH). For these pests, the information to be entered is the number of damaged leaves counted in each of the 10 sampled hills or quadrats. Other leaf-feeding damage, such as those caused by armyworms, are to be entered in the "other" category.

The third category is represented by planthoppers (brown planthopper [BPH] and white-back planthoppers [WPH]), rice bugs (RB) and green leafhoppers (GLH); the insects of this group are to be represented by numbers of individuals, rather than resulting damage.

The population levels of the two first pests, BPH and WPH, can be assessed using the standard tapping method, whereas the third, RB, can be estimated from direct counts. The fourth pest, GLH, can be assessed from 5 sweep samplings (represented by sweep no. in the recording form) over the rice canopy. Each of the 5 sweep samplings result from a series of 10 sweep strokes, and the insects caught are collected in small vials containing 70% alcohol. Insect identification and counts are made in the laboratory.

When observed, the presence of beneficial organisms (e.g., predators) should also be indicated in the appropriate section.

Diseases. Three broad categories of rice diseases are considered: foliar diseases, diseases that affect the tiller or the panicle as a whole, and systemic (viral) diseases.

The first category includes red stripe (RS), bacterial blight (BB), leaf blast (LB), *Helminthosporium* brown spot (BS), bacterial leaf streak (BLS), and narrow brown spot (NBS). Each of these diseases can be assessed as the number of diseased leaves counted in each of the 10 sampled hills (or 10x10 cm quadrat).

The second category includes sheath blight (ShB), sheath rot (ShR) false smut (FSM), and neck blast (NB). For these diseases, the information to be entered is the number of infected tillers (ShB, ShR) or panicles (FSM, NB) per hill (or 10x10 cm quadrat).

The third category is represented by diseases which intensities cannot be assessed accurately enough on the basis of a 10-hill sample. When presence of such diseases is recognized during the first passage across the field, a third, rapid passage across the field is necessary, where the number of infected hills out of a total of 50 hills taken at random is counted. Only one disease is specifically considered under this category: rice tungro disease (RTD).

Final remark. It is important to note that the form provides space in each category where information on additional pests may be entered. Changes or additions in the list of pests should be made without any hesitation, so as to best reflect the pests present in a given site. These changes should, however, be made according to the above guidelines, using the nature of the injuries caused by the corresponding pests.

Simplified protocol with two visits

Principle

The principles of a simplified survey protocol are the same as above, and are meant to enable producing reduced, yet useful, information. The simplified survey protocol involves only two field visits.

It must be recognized that the accuracy of the results generated through using the simplified protocol will be much less robust: this simplified procedure will not provide a history of events occurring in the surveyed fields (including, e.g., water, crop growth, and yield reducing factors). As a result, some important events in the course of a crop cycle in a given field may be missed by observers. Thus, much greater caution will be required in analyzing and interpreting the resulting data.

The procedure for selecting fields is the same as in the complete protocol (see the above procedure). Since the protocol used is simplified, it is advisable to consider increasing the number of fields considered in the survey.

The two visits should be made at, or about, two stages of the rice crop:

- **booting**, and
- **ripening**.

The choice of booting as a key development stage for assessing the crop status is discussed above. These two stages, also, should allow measuring injuries caused by most harmful agents of rice.

The same forms are to be used in the simplified protocol as compared with the complete one. In practice:

- while the complete protocol leads to filling 5 documents (1 copy of form 1, and 4 copies of form 2 – one for each visit),
- the simplified protocol leads to filling 3 documents (1 copy of form 1, and 2 copies of form 2 – one for each visit).

General information (Form 1)

The same information is to be collected (see above), including:

- Field number and code; field location;
- Landform;
- Field area;
- Crop establishment method;
- Plant material;
- Crop density;
- Overall crop status
- Yield estimates from three crop cuts.

Crop and pest status (Form 2)—at booting stage and at ripening stage

- First passage in the crop:
    - Crop development stage;
    - Water status.
- Second passage in the crop:
    - Crop growth;
    - Weed infestation;
    - Insect pests and pest injuries; beneficials;
    - Diseases.

List of variables included in the recording forms

General information: Information gathered per cropping season specific to each field

FA	Approximate field area	[ha]
LF	Landform	-
PC	Crop preceding rice in the field	-
FP	Duration in weeks of fallow period	[number of weeks]
ST	Soil texture	[%]
CEM	Crop establishment method	-
AST	Age of seedlings at transplanting	[days]
NtS	Number of tillers per square meter	[number]
CED	Crop establishment date	[date]
HD	Harvest date	[date]
CVR	Rice variety or hybrid name	-
LDG	Lodging at ripening stage	[% area affected]
CS	Crop status	[a]
Y	Estimated attainable yield	[kg ha^{-1}]

Information gathered at each visit

1. Crop management

WS	Water status	[a]
MAN	Manure input	[b, kg]
FERT	Mineral fertilizer input	[b, kg]
FUN	Fungicide use	[b, kg]
INS	Insecticide use	[b, g or ml]
HERB	Herbicide use	[b, g or ml]

2. Crop growth

Nt	No. of tillers per hill or quadrat	[number of tillers]
Np	No. of panicles per hill or quadrat	[number of panicles]
Nl	No. of leaves per tiller	[number of leaves]

3. Injuries due to animal pests

DH	Deadheart	[number of tillers]
WH	Whitehead	[number of panicles]
GM	Gall midge	[number of tillers]
WM	Whorl maggot	[number of leaves]
LF	Leaffolder	[number of leaves]
RH	Rice hispa	[number of leaves]
OTH	Other rice defoliators	[number of leaves]
BPH	Brown planthopper	[number of insects]
WPH	Whitebacked planthopper	[number of insects]
AW	Armyworm	[number of insects]
RB	Rice bug	[number of insects]
GLH	Green Leafhopper	[number of insects]
SNL	Snail	[number of plants or hills]
RT	Rat	[number of tillers]

4. Injuries due to pathogens

BlB	Bacterial blight	[number of leaves]
LB	Leaf blast	[number of leaves]
BS	Brown spot	[number of leaves]
BLS	Bacterial leaf streak	[number of leaves]

NBS	Narrow brown spot	[number of leaves]
LSm	Leaf smut	[number of leaves]
LSc	Leaf scald	[number of leaves]
RS	Red stripe	[number of leaves]
ShB	Sheath blight	[number of tillers]
ShR	Sheath rot	[number of tillers]
SR	Stem rot	[number of tillers]
FSm	False smut	[number of panicles]
NB	Neck blast	[number of panicles]
RTD	Rice tungro disease	[number of hills or plants]
GSD	Grassy stunt disease	[number of hills or plants]
RSD	Ragged stunt disease	[number of hills or plants]
YS	Yellowing syndrome	[number of hills or plants]
GLD	Glume discoloration	[number of panicles]

5. Injury due to weeds

WA	Weed above crop canopy	[a]
WB	Weed below crop canopy	[a]

[a] based on specific assessment scale
[b] number and dose/quantity applied

Form 1 - General Information date (dd-mm-yy)_____

Field number: _____ Village name: _____

Latitude [1] : _____ Longitude [1]: _____

Location [2]:

Approximate field area (ha): _____

Farmer's name: _____

Landform [3]:

Irrigated: ☐ Rainfed lowland: Upland: ☐

Upper field ☐ Medium ☐ Shallow water ☐

Previous crop: _____ Duration of fallow period (in weeks): _____

Known soil problems [4]: salt ☐ zinc ☐ aluminium ☐ others (specify): ☐ _____

Crop establishment method [4] :

Transplanted ☐ If transplanted, age of seedlings at transplanting (days) _____

Numbers of hills per sqm _____

Direct seeded ☐ Number of plants per sqm _____

Crop establishment date in field (dd-month-yy) :_____ Harvest date (dd-month-yy):_____

Rice variety [4]:_____ Traditional variety ☐ Modern variety ☐ Hybrid ☐

Fertilizer input [4]: ☐ none or:

Manure (kg per ha): _____

N/P/K (kg per ha): _____

Date of application (dd-month-yy): _____

Pesticide use [4]: ☐ none or :

Name of chemical: _____

Rate (dose and volume): _____

Date of application (dd-month-yy): _____

Crop status rating at booting stage [3]:

Very Good ☐ Good ☐ Average ☐ Poor ☐ Very poor ☐

Lodging at last visit (% of field affected): _____

Yield estimates (g dry grain per sqm): 1st area: _____ 2nd area: _____ 3rd area: _____

[1] In degrees (°), minutes (') and seconds ("); e.g., for Los Baños, Philippines: Latitude: 14°10'44"N and Longitude: 121°13'32"E
[2] Sketch a simplified map with field position
[3] Check as appropriate
[4] Check and/or fill as appropriate

Form 2 – Crop and Injuries | **Date (dd-month-yy):** | **Visit No.:**

Crop Information

[1]: according to assessment scale
[2]: specify
[3]: for other systemic diseases, specify

Crop development stage [1] _____ Water status [1]: _____

Hill or quadrat no.	1	2	3	4	5	6	7	8	9	10
No. of tillers per hill or quadrat										
No. of panicles per hill or quadrat										
No. of leaves per tiller										

Weeds

Weed rating [1] :

area	A	B	C
above the rice canopy			
below the rice canopy			

area	A	B	C
Main weed types: S			
BD			
G			
SD			

Major weeds [2]

_____
_____
_____
_____

Animal Pests

	Hill or quadrat no.	1	2	3	4	5	6	7	8	9	10
No. of damaged **plants** or hills	SNL										
	Others [2]										
	Hill or quadrat no.	1	2	3	4	5	6	7	8	9	10
No. of damaged **tillers** per hill or quadrat	DH										
	WH										
	GM										
	RT										
	Others [2]										
No. of damaged **leaves** per hill or quadrat	WM										
	LF										
	Others [2]										
No. of insects per **hill** or **quadrat** (tapping or counting	BPH										
	WPH										
	AW										
	RB										
	Others [2]										

No. of insects (5 sweeps)	Sweep no.	1	2	3	4	5
	GLH					
	Others [2]					

Remarks: beneficial insects [2]

_____
_____

Diseases

	Hill or quadrat no.	1	2	3	4	5	6	7	8	9	10
No. of infected **leaves** per hill or quadrat	BlB										
	LB										
	BS										
	BLS										
	NBS										
	RS										
	Others [2]										
No. of infected **tillers** or **panicles** per hill or quadrat	ShB										
	ShR										
	SR										
	FSM										
	NB										

No. of infected hills or quadrats out of a sample of 50 hills or quadrats:

RTD	
Others [3]	

A working example

Nature of the variables

The survey protocol described above generates a number of variables. Some are quantitative (e.g., yield estimates, injury measurements), while others are, by essence, qualitative (e.g., crop establishment method, variety). Some variables vary over time during a cropping season (e.g., injury measurements), while others do not (e.g., yield). A strategy is needed that allows to compact this information so that it can be analyzed.

Compaction over time: injury variables

Assessments of injuries over the successive development stages represent a large fraction of the data to be analyzed. Considering the objectives of the survey, this information should be translated into new variables, termed here injury variables, which capture the harmful effect of pests. In other words, pest dynamics is not a prime concern (four observations would prove far too poor an information to analyze most of them); rather, the possibly yield-reducing effects of these injuries must be considered.

These injury variables should be made specific to each injury type and should reflect hypotheses on mechanisms that may lead to damage. In analyzing surveys of rice pests, the following choices, for instance, have been made and the corresponding injury variables used:

- bacterial leaf blight: area under BLB intensity progress curve;
- sheath blight: maximum ShB intensity observed throughout the crop cycle;
- deadhearts: maximum deadheart incidence (% tillers damaged) in a crop cycle.

The choice of an injury variable very much depends on the injury itself and on available information on injury mechanisms. Additional examples and details on the rationale for such choices can be found in Savary et al. (1994, 1995).

The outcome of this phase is a considerable reduction in number of variables to address, since four measurements of injuries are summarized into one for each injury considered in the survey.

The variables generated by the survey procedure also vary in range. Some may vary widely, others narrowly. It is worth noting that the distribution frequencies are also likely to be very different (more specifically, while actual yield often exhibits a near-normal distribution, distribution frequencies of injury variables are seldom normal). This issue and several others are addressed in a working example.

A hypothetical survey data set

A working example is shown at the end of this section. In this example, a hypothetical data set from a survey in farmers' field is used. Only five variables are considered:

- Production situation (PS), represented by ordinal, classes;
- Three injury variables: Injury A (IA); Injury B (IB); Injury C (IC); and
- Actual yield (Y).

The working example includes a set of 90 records (lines), each of them representing a farmer's field.

For the sake of illustration, we can further assume that this hypothetical data set was generated from two visits in the farmers' fields: the first one at maximum tillering stage, where injuries were measured, and the second one, at harvest stage, when yield estimates were made. The injury variables used in this example could further be defined as follows:

- IA is the percent weed cover (% soil area covered by weeds below the rice crop canopy);
- IB is the incidence of rice tungro disease (% plants showing symptoms);
- IC is the proportion of leaf area damaged by leaf-feeding insects (% leaf area damaged).

Using this example, the development of categories will be illustrated, tests of independence between variables will be addressed, and an approach to analyze the complex relationships among variables will be outlined. In the latter, the following strategy (and the underlying hypotheses) will be used: (1) a framework of relationships exists between production situations and injury variables; (2) some injuries may cause yield reductions; and (3) actual yield, Y, is a reflection of complex relationships between injury variables and production situations.

Categorization: compaction over range

Of the five variables considered in the working example, only four exhibit variations in range:
- Injury A (IA) may vary from 1 to 100;
- Injury B (IB) may vary from 0 to 15;
- Injury C (IC) may vary from 0 to about 25; and
- Yield (Y) may vary from 0.5 to 7.5.

Production situation (PS) is an ordinal variable with three classes (PS1, PS2, and PS3), and will not be addressed here.

Categorization is one powerful means to further compact the information. Classes for each of the four quantitative variables can be defined using specific numeric boundaries for each variable:

- IA :	- class 1 (IA1):	$0 \le IA \le 10$	(number of fields n = 35)
	- class 2 (IA2):	$10 < IA \le 30$	(n = 35)
	- class 3 (IA3):	$30 < IA \le 100$	(n = 20)
- IB :	- class 1 (IB1):	$0 \le IB \le 1$	(n = 32)
	- class 2 (IB2):	$1 < IB \le 5$	(n = 38)
	- class 3 (IB3):	$5 < IB \le 100$	(n = 20)
- IC :	- class 1 (IC1):	$0 \le IC \le 5$	(n = 32)
	- class 2 (IC2):	$5 < IC \le 15$	(n = 40)
	- class 3 (IC3):	$15 < IC \le 100$	(n = 18)
- Y :	- class 1 (Y1):	$0 \le Y \le 2$ t ha^{-1}	(n = 20)
	- class 2 (Y2):	$2 < Y \le 4$	(n = 27)
	- class 3 (Y3):	$4 < Y \le 5$	(n = 27)
	- class 4 (Y4):	$4 < Y$	(n = 16)

One reason for establishing these numerical boundaries is to achieve a class-filling that would be more or less balanced among all classes representing a given variable. The number of classes to define for each variable is another important issue. This choice depends on the size of the sample (the number of fields included in the survey) and determines whether chi-square tests are valid or not.

Contingency tables and chi-square tests

The result of categorization of quantitative data is that all the information is brought to the same format: the variables are each represented by a (limited) number of classes. This allows further analysis, combining simultaneously variables which are qualitative in nature and variables which are not.

Building contingency tables is an easy way to explore the relations between paired qualitative or coded quantitative variables. The contingency tables (see below) show bivariate frequency distributions. A chi-square test can be applied to confirm the suggested pattern of relationships in a contingency table. The null hypothesis is the independence of the distribution frequencies of the two variables. The validity of the test depends on the expected sizes of groups, assuming independence of the two distribution frequencies: it should generally be applied to groups with a minimum size. This minimum size was set at 5 by Dagnélie (1973). Gibbons gave a more flexible rule: no more than 20 % of the expected values (assuming independence) should be smaller than 5.

One may want, for example, to test the null hypothesis of independence between injury variable IA and production situations. The following table is formed:

	IA1	IA2	IA3	Totals
PS1	15	13	2	30
PS2	4	12	10	26
PS3	16	10	8	34
Totals	35	35	20	90

with a chi-square value of 12.90.
The expected values (assuming independence between the two distribution frequencies) are:

	IA1	IA2	IA3
PS1	11.67	11.67	6.67
PS2	10.11	10.11	5.77
PS3	13.22	13.22	7.56

and the null hypothesis of independence of the two distribution frequencies is rejected at $P = 0.012$, with 4 degrees of freedom. This result suggests that injury A occurs at higher levels in some production situations (PS2) than in others (PS1).

In practice, major disequilibrium among classes should be avoided; all classes must be represented by a commensurate numbers of individuals (here, fields). This is why only 3 to 4 classes were defined in the previous stage of the analysis. This guideline is essential to the next step, the analysis of a series of contingency tables by means of correspondence analysis.

Correspondence analysis

A number of contingency tables may be built to test bivariate relationships (1) among injuries, (2) between injuries and production situations, (3) between injuries and yield, and (4) between yield and production situations. Correspondence analysis is a very flexible multivariate, non parametric

technique that addresses these sets of relationships in a synthetic way. In this sense, it can be seen as a third means to compact information. As any multivariate technique, correspondence analysis may be designed in such a way that it addresses some issues more specifically than others. In this working example, one strategy may be based on the fact that strong links exist between (1) production situations and injuries, and (2) production situation and yield. This strategy then consists of:

- building the: [production situation x injury] tables: [PS x IA], [PS x IB], and [PS x IC];
- using them as a framework of analysis;
- describing yield variations using this framework and the additional contingency table: [PS x Y].

The procedure used in correspondence analysis is similar to principal component analysis, and involves the computation of eigenvectors and eigenvalues (Benzécri 1973, Greenacre 1984). The sum of the eigenvalues is called the inertia and, with correspondence analysis, equals the chi-square statistic divided by the total number of observations. Each class contributes a fraction to the total inertia; summation of inertia over classes yields the total inertia. Coordinates for axes are defined, based on the eigenvalues. Unlike principal component analysis where the entries of the data matrix are quantitative, correspondence analysis is based on computations on a data matrix of frequencies. Another difference is that classes in the columns and the rows are involved in the same way: each eigenvector is computed from weighted combinations of all the classes that have been selected as 'active' in the analysis. The reason for this is that correspondence analysis is based on a chi-square distance (Benzécri 1973) between classes (Dervin 1988).

Besides the coordinates along the axes, classes are represented by their relative weights, contribution to each axis, and reciprocal contribution to axes. The relative weight (or mass) of each class accounts for the frequency of individuals in the corresponding row (or column). The contribution to an axis is the percentage of inertia of that axis which is derived from a specific class. The reciprocal contribution (or class correlation) represents the proportion of inertia by the specified axis. The sign of the coordinate of each class along each axis indicates the direction that the class takes as it deviates from the origin.

The option chosen to illustrate correspondence analysis in the case of this hypothetical survey data set is therefore to run a correspondence analysis on the [production situation x injury] contingency tables. The coordinates of the yield classes and their reciprocal contributions are computed in a second stage. The latter represents a measure of the representativeness of the [production situation x injury] contingency tables to describe yield variation. The set of contingency tables is shown in Table 4.1.

Outputs from a correspondence analysis using this table are appended to this section. As in principal component analysis, correspondence analysis usually generates a number of axes. Here, because of the few classes involved, the simple analytical approach used, and, mostly, the simplified nature of the hypothetical data used, only two axes are defined. The first axis involves contributions of PS1 and PS3 and injuries such as IA3, IB1, IB3, IC1, and IC3. This first axis may be seen as an axis of decreasing levels of injury B (see co-ordinate values and change of signs of classes of this variable along axis 1). The second axis involves contributions of PS2 and PS3 and of injuries: IA1, IA3, IB2, and especially, IB3. Examination of the reciprocal contributions indicates very high values for all four classes of yield, indicating a excellent overall representativeness of the analysis with respect to yield variation. The coordinates of the Y-classes along the first axis indicate that axis 1 primarily represents a gradient of increasing yields.

Table 4.1. A hypothetical survey in farmers' fields: set of contingency tables used for correspondence analysis.

Variable	Production situation (PS)		
	PS1	PS2	PS3
IA1	15	4	16
IA2	13	12	10
IA3	2	10	8
IB1	16	10	6
IB2	10	16	12
IB3	4	0	16
IC1	4	12	16
IC2	12	12	16
IC3	14	2	2
Y1	0	10	10
Y2	5	12	10
Y3	17	2	8
Y4	8	2	6

Examination of the correspondence analysis graph can be done in several ways. First, one may look at the production situation variables and attempt to detect linkages with other variables. While PS2 seems primarily associated with medium-high injury A (IA2-IA3), with medium level of injury B (IB2) and low-medium levels of injury C (IC1-IC2), production situation PS1 seems predominantly associated with IC3, IA1, and IB1. These two production situations also differ in linkages with yield: PS1 is predominantly associated with medium-high yields (Y3-Y4), while PS2 is linked with medium-low yields (Y1-Y2). Production situation PS3 shares some of the traits of both PS1 and PS2. It does not appear to be particularly linked with any yield level in particular; but it shows a strong linkage with one injury class: IB3.

The graph may also be considered from the perspective of the coded variables which incorporate a sequence: injuries and yield. A path of increasing yield, broadly from left to right, can be outlined. This path is opposed to that of injury A and to that of injury B. This suggests a strong, negative effect of these injuries on actual yields, and this hypothesis can be tested using the corresponding chi-square tests. Injury class IB3 seems particularly linked to production situation PS3. When projected on axis 1, the coordinate of IB3 is intermediate to very low (Y1) or low (Y2) yields. However, as a whole, production situation PS3 is associated to any level of yield, from very low to very high. This would suggest that (1) injury B is particularly frequent in PS3, where (2) it may cause considerable damage. These hypotheses can also be tested using specific chi-square tests

and documented using another analytical approach involving, for instance, [yield x injury] contingency tables.

Cluster analysis and other analytical approaches

This document does not intend to list the many approaches available to analyze survey data on the characterization of crop pests. One group of techniques and an alternative methodology that have been used to characterize rice pests are worth mentioning, however.

Another very broad and extremely powerful set of techniques that share commonalties with correspondence analysis is represented by clustering techniques. These techniques have a strong potential to condense heterogeneous data. Cluster analysis has been used to define production situations and injury profiles (Savary et al., 1997). This type of technique is particularly appealing when a chi-square distance is used, which makes it conceptually compatible with correspondence analysis. An example is briefly outlined at the end of this document.

One important objective of rice pest surveys is to quantify the losses caused by pests. When this objective assumes first priority in the analysis, a number of multivariate, parametric techniques are available. One approach may involve the following sequence (e.g., Savary et al 1996):

- A series of factors (F_i) are generated from principal component analysis on injury variables and (coded) variables representing cropping practices;
- The F_i are independent, linear combinations of the original injury variables;
- The F_i are then used as new, synthetic variables in a stepwise multiple regression analysis;
- A multiple regression equation of actual yield (Y) is developed, with the shape

$$Y = a - \Sigma\, b_i\, F_i$$

- This equation is then used to (1) estimate attainable yields, Y_a (setting injury values to 0 in the F_i); (2) estimate actual yields, Y_j, corresponding to specific injuries j (setting each injury value to their means in the F_i); and (3) estimate mean damage due to specific injury j : $YL = Y_a - Y_j$.

The latter approach is appealing as it allows generation of injury-specific estimates of losses. It does not, however, incorporate interactions among injuries, nor does it make justice to changes in production situations, which are 'scaled-down' by regression parameters. It also raises arguments among statisticians, especially with respect to the risk of bias of parameter estimates (this much depends on the properties of the regression model obtained). And, from the plant protection perspective, it does not make use of actual measurements of attainable yield, but of statistical estimates of attainable yield instead. It does, however, have value in generating indices which, with caution, may allow ranking of pest impacts.

CORRESPONDENCE ANALYSIS

FILE NAME: AFCDUMMY.DAT

This file contains 13 classes in rows, among which the 9 first classes are active in this analysis.

ACTIVE CLASSES (COLUMNS):

PS1 PS2 PS3

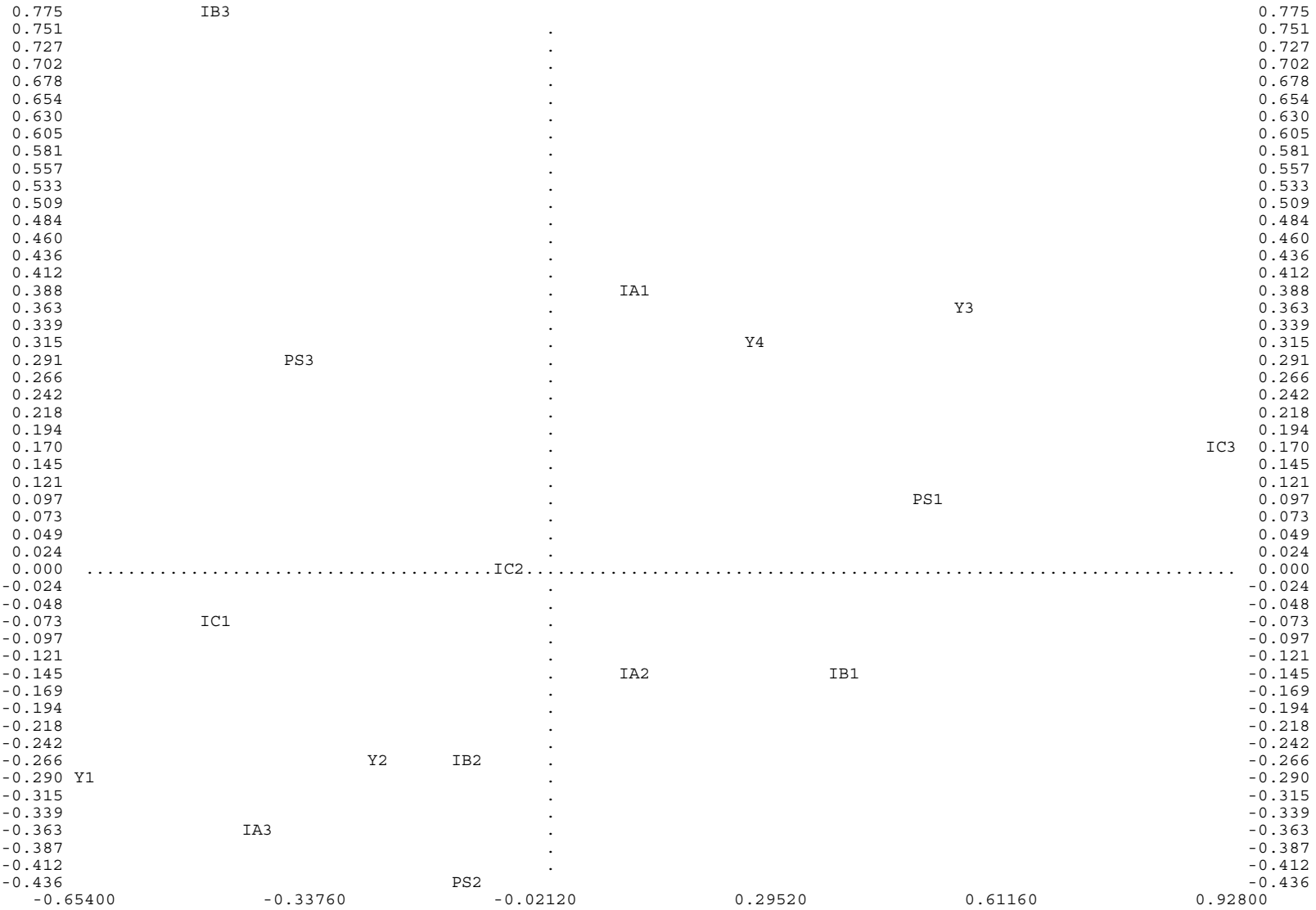
EIGENVECTORS

Factor	Lambda	Inertia	Accum. Inert.	Histogram of eigenvalues
1		0.13323	0.593	0.593
				**********
2	0.09159	0.407	1.000	**********

CO-ORDINATES AND CONTRIBUTIONS ET DES CONTRIBUTIONS (A =CONTRIBUTIONS TO AXES; R = RECIPROCAL CONTRIBUTIONS)

Class	Mass	Axis 1	A	C	Axis 2	A	C
Active classes (columns)							
PS1	0.333	0.504	63.571	96.761	0.092	3.096	3.239
PS2	0.289	-0.142	4.365	8.685	-0.460	66.746	91.315
PS3	0.378	-0.336	32.064	60.730	0.270	30.158	39.270
Active classes (rows)							
IA1	0.130	0.126	1.551	10.672	0.365	18.889	89.328
IA2	0.130	0.116	1.319	36.777	-0.153	3.299	63.223
IA3	0.074	-0.425	10.032	56.574	-0.372	11.200	43.426
IB1	0.119	0.396	13.971	86.716	-0.155	3.113	13.284
IB2	0.141	-0.091	0.878	9.733	-0.278	11.846	90.267
IB3	0.074	-0.461	11.807	26.085	0.776	48.665	73.915
IC1	0.119	-0.434	16.739	96.287	-0.085	0.939	3.713
IC2	0.148	-0.071	0.558	98.979	-0.007	0.008	1.021
IC3	0.067	0.929	43.145	96.852	0.167	2.040	3.148
Additional classes (rows)							
Y1	0.000	-0.655	0.000	81.383	-0.313	0.000	18.617
Y2	0.000	-0.258	0.000	44.530	-0.288	0.000	55.470
Y3	0.000	0.568	0.000	73.145	0.344	0.000	26.855
Y4	0.000	0.296	0.000	49.832	0.297	0.000	50.168

PROJECTION OF CLASSES ON AXES 1 (horizontal) AND 2 (vertical)



A hypothetical data set on rice injuries and yields in farmers fields.

ProdSit	Iac	Ibc	Icc	Yc	Ia	Ib	Ic	Y
PS1	Ia1	Ib1	Ic2	Yc4	5	0	6	5.8
PS2	Ia3	Ib2	Ic2	Yc2	35	2	11	3.2
PS1	Ia2	Ib2	Ic1	Yc3	12	3	3	4.1
PS1	Ia2	Ib1	Ic3	Yc3	25	0	17	4.8
PS2	Ia2	Ib2	Ic3	Yc1	20	3	21	1.5
PS3	Ia1	Ib2	Ic2	Yc3	8	3	7	4.9
PS1	Ia1	Ib1	Ic3	Yc3	4	0	16	4.2
PS2	Ia3	Ib2	Ic2	Yc2	40	5	9	2.5
PS1	Ia2	Ib1	Ic2	Yc4	17	0	6	7.3
PS3	Ia3	Ib3	Ic2	Yc1	35	6	8	1.7
PS3	Ia2	Ib3	Ic3	Yc2	15	5.5	19	3.1
PS1	Ia2	Ib2	Ic3	Yc3	20	2	21	4.5
PS3	Ia1	Ib1	Ic1	Yc4	3	0.5	0	5.3
PS2	Ia2	Ib2	Ic2	Yc2	25	2.5	7	3.9
PS1	Ia1	Ib2	Ic2	Yc3	7	1.5	12	4.3
PS3	Ia3	Ib3	Ic2	Yc1	35	7	11	1.8
PS3	Ia1	Ib1	Ic2	Yc4	10	0.5	8	5.9
PS2	Ia1	Ib1	Ic1	Yc3	2	0	2	5.0
PS1	Ia3	Ib3	Ic2	Yc2	32	5.5	7	2.7
PS1	Ia1	Ib1	Ic3	Yc3	3	0	17	4.9
PS3	Ia3	Ib3	Ic1	Yc2	35	10.5	3	4.0
PS3	Ia2	Ib2	Ic2	Yc3	13	2.5	8	4.1
PS1	Ia1	Ib1	Ic3	Yc4	7	0	21	5.3
PS3	Ia2	Ib2	Ic1	Yc3	20	2	2	4.7
PS1	Ia1	Ib2	Ic2	Yc3	6	3	12	4.6
PS2	Ia2	Ib1	Ic1	Yc2	15	0	5	2.9
PS3	Ia1	Ib3	Ic1	Yc2	10	7.5	4	3.5
PS3	Ia2	Ib3	Ic2	Yc1	25	8.5	9	2.0
PS1	Ia1	Ib3	Ic1	Yc2	5	6.5	1	3.6
PS1	Ia3	Ib1	Ic3	Yc2	32	0.5	18	3.9
PS2	Ia3	Ib2	Ic1	Yc1	35	3.5	2	1.7
PS2	Ia2	Ib1	Ic2	Yc2	15	0	7	2.9
PS3	Ia1	Ib3	Ic1	Yc1	9	7	1	1.8
PS3	Ia1	Ib2	Ic2	Yc2	4	3.5	13	2.8
PS2	Ia3	Ib2	Ic1	Yc1	40	3	5	0.5
PS1	Ia2	Ib1	Ic2	Yc4	12	0	12	6.4
PS3	Ia2	Ib1	Ic1	Yc4	20	0.5	3	6.2
PS2	Ia3	Ib2	Ic2	Yc1	35	4	8	1.3
PS1	Ia2	Ib2	Ic3	Yc3	15	5	23	4.1
PS3	Ia1	Ib2	Ic1	Yc3	3	1.5	2	4.5
PS3	Ia1	Ib3	Ic1	Yc1	7	6.5	4	1.6
PS2	Ia1	Ib1	Ic1	Yc4	5	0.5	3	5.9
PS3	Ia3	Ib2	Ic2	Yc2	32	2	2	2.9
PS2	Ia2	Ib1	Ic1	Yc1	25	0	2	1.9
PS2	Ia2	Ib2	Ic2	Yc2	15	2.5	9	3.8
PS3	Ia1	Ib3	Ic1	Yc1	7	6.5	4	1.6
PS2	Ia1	Ib1	Ic1	Yc4	5	0.5	3	5.9
PS3	Ia3	Ib2	Ic2	Yc2	32	2	7	2.9
PS2	Ia2	Ib1	Ic1	Yc1	25	0	2	1.9

ProdSit	Iac	Ibc	Icc	Yc	Ia	Ib	Ic	Y
PS3	Ia3	Ib3	Ic1	Yc2	35	10.5	3	4.0
PS3	Ia2	Ib2	Ic2	Yc3	13	2.5	8	4.1
PS1	Ia1	Ib1	Ic3	Yc4	7	0	21	5.3
PS3	Ia2	Ib2	Ic1	Yc3	20	2	2	4.7
PS1	Ia1	Ib2	Ic2	Yc3	6	3	12	4.6
PS2	Ia2	Ib1	Ic1	Yc2	15	0	5	2.9
PS3	Ia1	Ib3	Ic1	Yc2	10	7.5	4	3.5
PS3	Ia2	Ib3	Ic2	Yc3	13	2.5	8	4.1
PS1	Ia1	Ib1	Ic3	Yc4	5	0	6	5.8
PS2	Ia3	Ib2	Ic2	Yc2	35	2	11	3.2
PS1	Ia2	Ib2	Ic1	Yc3	12	3	3	4.1
PS1	Ia2	Ib1	Ic3	Yc3	25	0	17	4.8
PS2	Ia2	Ib2	Ic3	Yc1	20	3	21	1.5
PS3	Ia1	Ib2	Ic2	Yc3	8	3	7	4.9
PS1	Ia1	Ib1	Ic3	Yc3	4	0	16	4.2
PS2	Ia3	Ib2	Ic2	Yc2	40	5	9	2.5
PS1	Ia2	Ib1	Ic2	Yc4	17	0	6	7.3
PS3	Ia3	Ib3	Ic2	Yc1	35	6	8	1.7
PS3	Ia2	Ib3	Ic3	Yc2	15	5.5	19	3.1
PS1	Ia2	Ib2	Ic2	Yc4	20	2	21	4.5
PS3	Ia1	Ib1	Ic1	Yc4	3	0.5	0	5.3
PS2	Ia2	Ib2	Ic2	Yc2	25	2.5	7	3.9
PS1	Ia1	Ib2	Ic2	Yc3	7	1.5	12	4.3
PS3	Ia3	Ib3	Ic2	Yc1	35	7	11	1.8
PS3	Ia1	Ib1	Ic2	Yc4	10	0.5	8	5.9
PS2	Ia1	Ib1	Ic1	Yc3	2	0	2	5.0
PS1	Ia2	Ib3	Ic2	Yc2	32	5.5	7	2.7
PS1	Ia1	Ib1	Ic3	Yc3	3	0	17	4.9

PS1	Ia1	Ib2	Ic2	Yc2	15	2.5	9	3.8

Illustrating the concept of recommendation domain in IPM: a case study in Eastern Uttar Pradesh

IPM strategies are field-dependent. Among the many reasons for this is the fact that the particular combination of pests and diseases (the *injury* profile) that affects a given crop is, to a significant extent, driven by location-specific crop management practices. Another reason is that the amount of yield reduction (*damage*) that can be attributed to a given injury profile depends upon the (*attainable*) yield the crop would have achieved, had these pests been absent. The attainable yield reflects a given production situation and is field-dependent. It is therefore necessary to identify the domains where rice pests and diseases may become constraints to productivity and where different IPM strategies are to be considered.

Surveys have been conducted in 1991-1994 to characterize the injury profile associated with differing production situations in Uttar Pradesh, India. The surveys address monsoon (kharif) season crops in three typical rainfed landforms of two villages. The procedure to analyze this information is exemplified here on the 1992 data, representing a population of 80 fields. The data collected cover a number of categorical (qualitative) or categorized (quantitative) attributes of the agroecosystem as:

• components of the cropping practices such as
  - previous crop in a given field (fallow, wheat, legume, or potato);
  - level of fertilizer inputs (from absent, NPK0, to high, NPK3);
  - application of manure (yes, man0, or no, man1);
  - crop establishment method (transplanted, tr or directly seeded, ds);
  - water stress (low, WS1 to high, WS3);
  - duration of rice crop cycle (short, CD1 to long, CD3);

• components of the injury profiles, represented by the intensities of diseases and pest injuries assessed at three crop development stages, such as:
  - the maximum proportion of tillers with dead hearts or sheath blight;
  - the maximum proportion of panicles with whiteheads, sheath rot, or neck blast;
  - the maximum observed number of armyworms per hill;
  - the area under the number of observed rice bugs per hill;
  - the area under brown spot intensity or leaf folder injury progress curves;

• rice yield, estimated from three sampling areas (1 m^2) each in each field, from very low, Y1 (below 2.4 $t.ha^{-1}$) to high, Y4 (above 3.4 and below 5.2 $t.ha^{-1}$).

The methodology to analyze these data strongly relies on non parametric, multivariate methods that allow handling of qualitative and quantitative information simultaneously. The analysis proceeds in two steps: (i) patterns of cropping practices are characterized using cluster analysis (Figure 4.1), and (ii) these patterns are linked to predominant pests using chi-square tests and correspondence analysis (Figure 4.2).

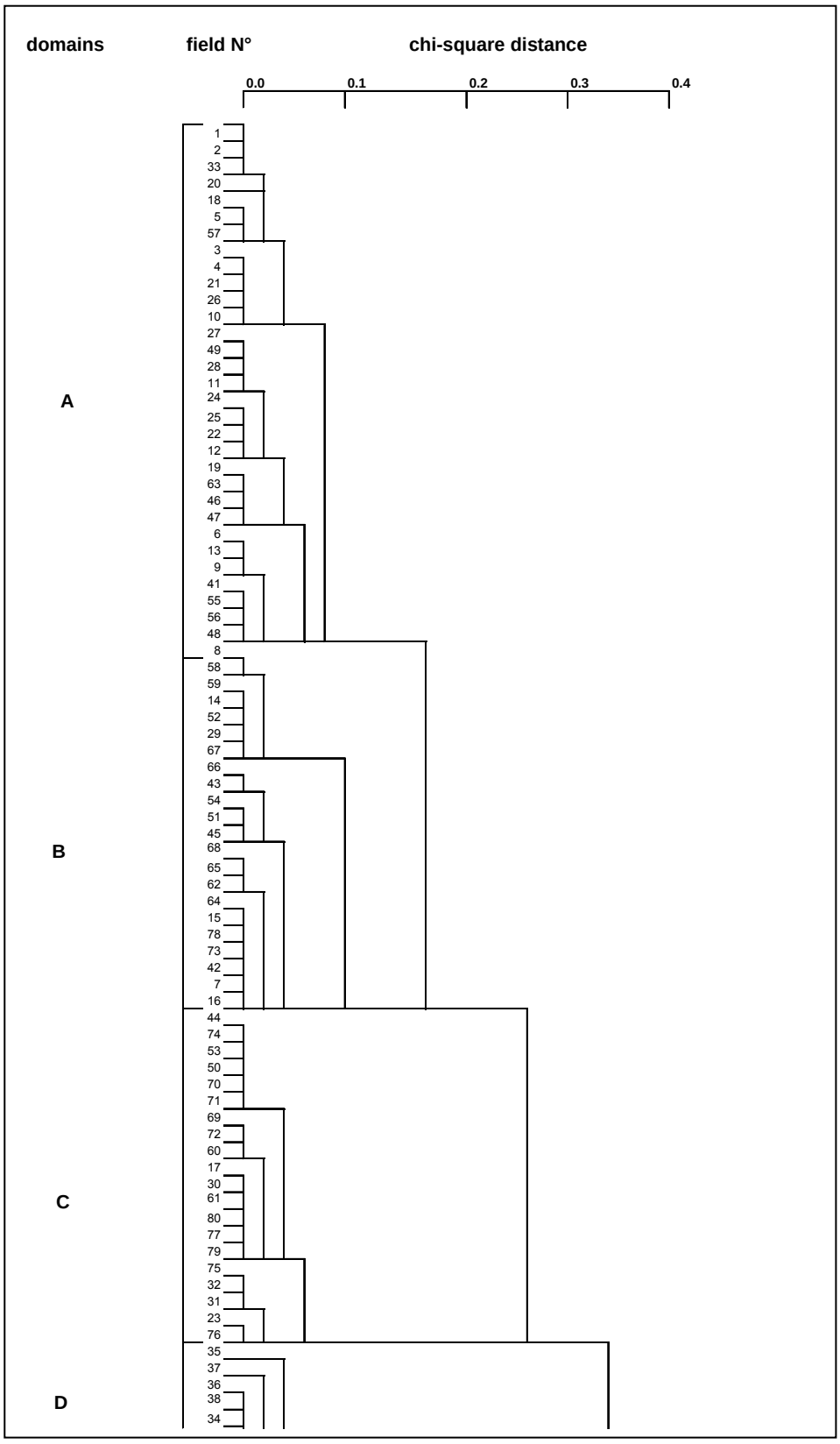
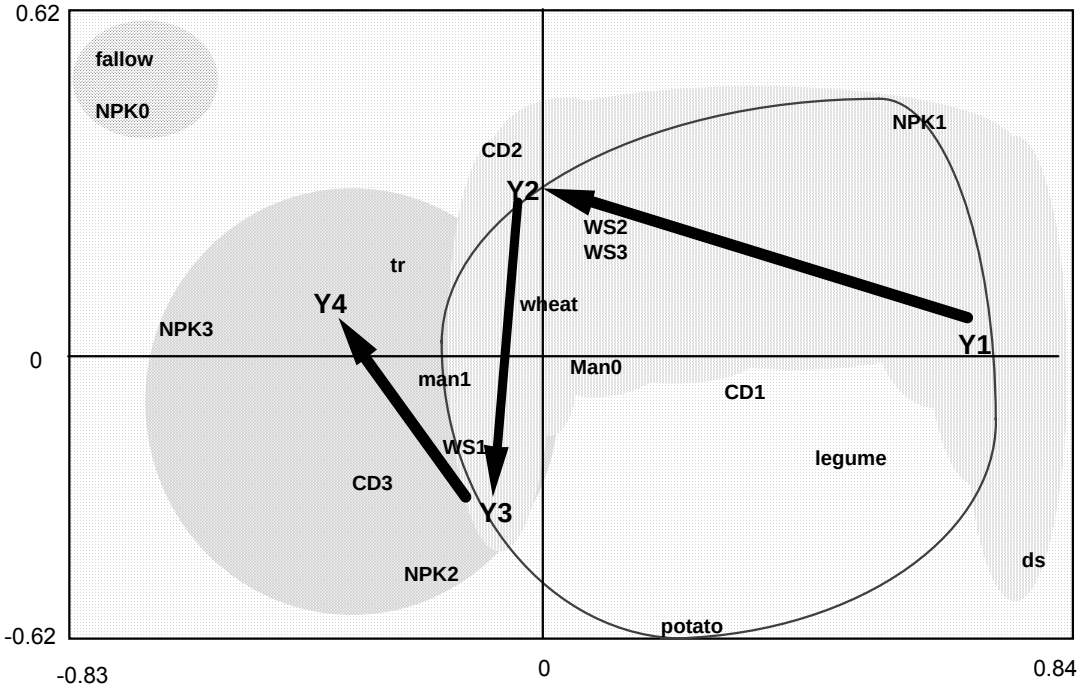


Figure 4.1. Patterns of cropping practices in eastern Uttar Pradesh: hierarchical cluster analysis using a Chi-square metric.



domain		cropping practices	pest constraints
A		- High fertilizer input - Some manure - previous crop: wheat predominent - transplanted rice - medium-high yield	- deadhearts - whiteheads - army worms - rice bugs - sheath blight - sheath rot - neck blast
B		- low fertilizer input - some manure - diverse previous crops - water stress often high - transplanted rice predominent - short cycle - low yield	- glume discoloration - (leaf blast)
C		- low fertilizer input - no manure - previous crop: wheat only - medium-low water stress - transplanted or direct-seeded rice - long cycle varieties	- weeds above the rice canopy - glume discoloration
D		- no fertilizer - fallow	- deadhearts - whiteheads - sheath rot

Figure 4.2. Patterns of cropping practices, injuries and actual yields in eastern Uttar Pradesh.

As the crop management practices are essentially described by qualitative attributes (e.g. variety name, previous crop, crop establishment method), a qualitative metric such as chi-square distance is appropriate to aggregate fields into clusters (Figure 4.1). Correspondence analysis is a non parametric multivariate procedure that allows visualization of relationships among variables. Graphs that can be read as maps are generated where various attributes of the agroecosystem are located. Figure 4.2 shows a path of increasing yield levels corresponding to the change in crop management practices. It indicates that recommendation domains associated with the corresponding injury profiles can actually be delineated.

Figure 4.2 shows that a recommendation domain should first be seen in terms of a combination of crop management practices rather than in terms of spatial proximity, and that it does not need to be continuous at the farm level. In other words, two neighboring fields may belong to different recommendation domains. This approach provided a framework for yield loss experiments, where actual damage due to pest combinations was measured in farmers' field experiments in Uttar Pradesh, and in the crop loss data base experiments conducted at IRRI.

References

Benzécri, J.P. 1973. L'Analyse des Données. Vol. 2. L'Analyse des Correspondances. Dunod, Paris.

Butt, D.J. and Royle, D.J. 1980. The importance of terms and definitions for a conceptually unified epidemiology. In: Comparative Epidemiology. A Tool for Better Disease Management. Palti, J. & Kranz, J., Eds. Pudoc, Wageningen. pp. 29-45.

Campbell, C.L. & Madden, L.V. 1990. *Introduction to Plant Disease Epidemiology.* John Wiley and Sons, New York. 532 p.

Chiarappa, L., Ed. 1971. Crop Loss Assessment Methods. FAO Manual on Evaluation and Prevention of Losses by Pests, Diseases, and Weeds. Comm. Agric. Bureaux, Farnham, England. Loose-leafed.

Church, B.M. and Austin, R.B. 1983. Variability of wheat yields in England and Wales. J. Agric. Sci. Camb. 100:201-204.

Dagnélie, P. 1973. "Théorie et Méthodes Statistiques. Applications Agronomiques". Vol. 2. Presses Agronomiques de Gembloux.

Dervin, C. 1988. "Comment Interpréter les Résultats d'une Analyse Factorielle des Correspondances?" Institut Technique des Céréales et Fourrages, Paris.

De Wit, C.T. & Penning de Vries, F.W.T. 1982. L'analyse des systèmes de production primaires. In: *La Productivité des Pâturages Sahéliens.* Penning de Vries, F.W.T. & Djiteye, M.A., Eds. Agr. Res. Rep. 918, Pudoc, Wageningen. pp. 275-283.

Du, P.V. and Savary, S., 1997. A survey of rice constraints in the Mekong delta. International Rice Research Notes. 22:43-44.

Elazegui, F.A., Soriano, J., Bandong, J., Estorninos, L., Jonson, I., Teng, P.S., Shepard, B.M., Litsinger, J.A., Moody, K. & Hibino, H. 1990. Methodology used in the IRRI Integrated Pest Survey. In: Crop Loss Assessment in Rice. IRRI, Los Baños. pp.243-271.

Forbes, G.A. and Jeger, M.J. 1987. Factors affecting the estimation of disease intensity in simulated plant structures. Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz 94:113-120.

Gibbons, J.D. 1976. Non Parametric Methods for Quantitative Analysis. Holt, Rinehart and Winston, New York.

Greenacre, M.J. 1984. Teory and Applications of Correspondence Analysis. Academic Press, London.

Litsinger, J.A. 1992. Crop loss assessment in rice. In: Rice Insects: Management Strategies. Heinrichs, E.A. & Miller, T.A., Eds. Springer-Verlag, New York. pp. 1-65.

Pinnschmidt, H.O., Batchelor, W.D. and Teng, P.S. 1995. Simulation of multiple species damage in rice using CERES-rice. Agricultural Systems 48:193-222.

Savary, S., Elazegui, F.A., Moody, K. and Teng, P.S. 1994. Characterization of rice cropping practices and multiple pest systems in the Philippines. Agricultural Systems 46. 385-408.

Savary, S., Madden, L.V., Zadoks, J.C. & Klein-Gebbinck, H.W. 1995. Use of categorical information and correspondence analysis in plant disease epidemiology. Advances in Botanical Research incorporating Advances in Plant Pathology. Vol 21: 213-240. Academic Press Ltd. London.

Savary, S., Elazegui, F.A., Pinnschmidt, H.O., and Teng, P.S. 1996. Characterization of Rice Pest Constraints in Asia: an Empirical Approach. In: Kropff, M.J., Teng, P.S., Aggarwal, P.K.,

Bouma, J., Bouman, B.A.M., Jones, J.W. & Van Laar, H.H. (Eds) *Application of Systems Approaches at the Field Level.* Kluwer Scientific Publishers, Dordrecht, The Netherlands.

Savary, S., Srivastava, R.K., Singh, H.M. and F.A. Elazegui, 1997. A characterization of rice pests and quantification of yield losses in the rice-wheat system of India. Crop Protection. 16: 387-398.

Savary, S., Willocquet, L., Elazegui, F.A., Teng, P.S., Du, P.V., Zhu, D., Tang, Q., Huang, S., Lin, X. Singh, H.M. & Srivastava, R.K., 2000. Rice pest constraints in tropical Asia: Characterization of injury profiles in relation to production situations. Plant Disease 84:341-356.

Savary, S., Willocquet, L., Elazegui, F.A., Castilla, N. & Teng, P.S., 2000. Rice pest constraints in tropical Asia: Quantification of yield losses due to rice pests in a range of production situations. Plant Disease 84:357-369.

Savary, S., Teng, P.S., Willocquet, L. & Nutter, F.W., Jr., 2006. Quantification and modeling of crop losses: a review of purposes. Annual Review of Phytopathology 44: 89-112.

Savary, S., Teng, P.S., Willocquet, L. & Nutter, F.W., Jr., 2006. Quantification and modeling of crop losses: a review of purposes. Annual Review of Phytopathology 44: 89-112.

Sta Cruz, P.C., Simbahan, G.C., Hill, J.E., Dobermann, A., Zeigler, R.S., Du, P.V., Dela Pena, F.A., Samiayyan, K., Suparyono, Tuat, N.V., Zheng, Z., 2001. Pest profiles at varying nutrient input levels. In: S. Peng, B. Hardy (Eds.) Rice Research for Food Security and Poverty Alleviation. Proceedings of the International Rice Research Conference, 31 March – 3 April 2000. IRRI Los Banos, Philippines, pp. 431-440.

Teng, P.S. 1990. Integrated Pest Management in Rice. An Analysis of the Status Quo with Recommendation for Action. Report to USAID Pesticide and Pest Management Project. 63 p.

Teng, P.S., Ed. 1987. Crop Loss Assessment and Pest Management. APS Press, St Paul. 270 p.

Teng, P.S. & Savary, S. 1992. Implementing the systems approach in pest management. Agricultural Systems 40, 237-264.

Zadoks, J.C. & Schein, R.D. 1979. Epidemiology and Plant Disease Management. Oxford University Press, New York. 427 p.

Zadoks, J.C. 1985. On the conceptual basis of crop loss assessment: the threshold theory. Annu. Rev. Phytopathol. 23, 455-473.

Zadoks, J.C., Anderson, P.K. & Savary, S., 1995. An eco-regional perspective of crop protection problems. In: Eco-regional Approaches for Sustainable Land Use and Food Production. Bouma et al. (Eds.) Kluwer Academic Publishers, Amsterdam. pp. 437-452.