Principal Component Analysis as a tool for analyzing On-farm by Imarhiagbe, O., Osawaru, M. E. and Ogwu, M. C. is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
.

Osawaru *et al.* (2015)
Biological and Environmental Sciences Journal for the Tropics 12(1), June, 2015 ISSN 0794 – 9057

BEST JOURNAL 12(1):514 – 522
Date received: March 05, 2015
Date accepted: May 25, 2015
Printed in Nigeria

PRINCIPAL COMPONENT ANALYSIS (PCA) AS AN IDEAL TOOL FOR ANALYSING ON-FARM RESEARCH DATA

[1, 2]Osawaru, M. E., *[1]Ogwu, M. C and [1]Imarhiagbe, O.
[1]Department of Plant Biology and Biotechnology, University of Benin, Benin City, Nigeria [2]Present address: Department of Biological Sciences, University of Abuja, Gwagwalada Abuja, FCT, Nigeria *Corresponding author: matthew.ogwu@uniben.edu/+234-8077692496

ABSTRACT
On-farm research is a set of procedures for adaptive research aimed at developing recommendations for representative groups of farmers. These groups of farmers' plots are used as the research site, thereby allowing for appropriate site specific technology, valid results and faster adoption. The move towards on-farm research means that researchers, such as plant breeders and agronomists, who have been trained in techniques of on-station research, are now under pressure to move on-farm. As a result of variations between farms and variations between plots within farm, many variables are encountered during data collections; the task of analyzing such enormous variables becomes a problem. One of such key steps in data analysis is finding ways to reduce dimension without sacrificing accuracy. Principal component analysis is a statistical technique in data analysis. It is used to compress higher dimensional dataset to lower dimensional ones. PCA allows the use of variables which are not measured in the same units for example elevation, concentration of nutrient, temperature and pH, making it an ideal tool for on-farm research data analysis.

Keywords: Dataset, On-Farm Research, Principal Component Analysis (PCA)

INTRODUCTION

On-Farm research/trials have gained popularity in the past years with due consideration given to the knowledge, problems and priorities of farming systems. Researchers, such as plant breeders and agronomists, who have been trained in techniques of on-station research, are now under pressure to move On-farm. It is therefore perhaps not surprising that the design of such on-station trials has too often, resulted in miniaturized research. Conditions for on-farm trials are typically less controlled than those for experimental fields. It was assumed that the best technology in research stations is also the best in farmer's fields. However, with recent testing especially in developing countries of humid tropics where there are high variability between farmers' fields, the response to improved crop management is less favorable in on-farm than in research stations. In the case of such inconsistencies, the technology selection process should be done on-farm in comparison to the farmers existing practice under the farmers growing conditions (Gomez and Gomez, 1984).

In on farm research, data collections emanate through large number of variables comprising the dataset. Hence, it is very likely that subsets of variables are highly correlated with each other. The accuracy and reliability of a classification or prediction model gotten from the on-farm experiment will suffer if highly correlated variables or variables that are unrelated to the outcome of interest are included. Superfluous variables can increase the data and data processing costs on a large database. The dimension of a model is the number of independent or input variables used by the model. It has been recommended that one of the key steps in data analysis is finding ways to reduce dimensionality without sacrificing accuracy.

Principal component analysis (PCA) is a classic technique in data analysis. It can be used for compressing higher dimensional data sets to lower dimensional ones. PCA allows the use of variables which are not measured in the same units for example elevation, concentration of nutrients, temperature, pH, and so on which makes it an ideal tool for analyzing on-farm data (Davis, 1986).

THEORETICAL FRAMEWORK

ON-FARM RESEARCH

On-Farm Research is a set of procedures for adaptive research whose purpose is to develop recommendations for representative groups of farmers. In On-farm research, farmers participate in identifying problems and its priorities, managing experiment and evaluating results. On-farm research, therefore, is an adaptive research with farmer's perspectives. The objective of On-farm research is to identify existing inputs or practices that might help to solve major problems of many farmers in a defined study area (Wuest *et al.*, 1991).

SIZE AND SCOPE OF FIELD RESEARCH

Traditional small plot research is conducted on small uniform experimental areas that minimize side effects that often plague field research.

Theoretically, small plot research enhances the researcher's ability to detect true and repeatable differences among the experimental treatments. Small plot research enables researchers to evaluate many treatments in a small area of land which minimizes the land resources required for field plot research. The small plot sizes often require specialized or small-scale research plot equipment.

On-farm research targets fields by virtue of their larger size, are typically more variable than smaller fields used for small-plot research. The greater within-field variability introduces a lot of background "noise" that can mask true differences in the measured responses between treatments. On-farm research allows for the use of commercial-scale field equipment and yield monitoring, but because individual plot size with OFR is larger (often equipment width by length of field), the number of treatments that can be evaluated per acre of land is fewer than with small-plot research.

REASONS FOR ON-FARM RESEARCH

- On-farm research gives high quality results regarding the suitability of the investigated technological innovations under small-, medium-, and large-scale farming conditions.
- On-farm research leads to more appropriate site specific technology and aid faster adoption (Ashby, 1986).
- Since farmers are the adopters and often the innovators of new farming techniques, it would therefore be unwise to undertake farm-based research without involving farmers in the research process as much as possible (Wuest *et al.,* 1991).
- Using accepted methods of on-farm testing, farmers can achieve experimental precision comparable to those of intensive university research trials (Spencer, 1993).
- On-farm research broadens the range of validity of conclusions beyond the narrow confines of a research institute (Statistical Services Centre, 1998).

STAGES OF ON-FARM RESEARCH

CIMMYT (1988) identified five procedures of an on-farm research program, which is still very much valid today and include:

(i) Diagnosis
(ii) Planning
(iii) Experimentation
(iv) Assessment / Evaluation of results
(v) Recommendation and diffusion

i. **DIAGNOSIS:** This involves collecting and analyzing information to design On-farm experiments. It requires an understanding of farmers' circumstances and practices are made in order to;

(a) Understand the farming system, and system interaction;
(b) Identify possible productivity problems and
(c) Begin to develop hypotheses on possible solutions.

On farm research ought to begin with an understanding of farmer's conditions. This requires an examination of farmer's fields and interviews with farmers. The diagnosis is used to help identify major factors that limit farm productivity and to help specify possible improvements. The Information will be useful in planning future experiments.

ii. **PLANNING:** The planning of On-farm research is used to identify experimental factors to be included in on-farm experiments, as well as to suggest other research activities. There are six practical steps for research planning:

- Listing problems
- Ranking problems
- Identifying causes of problems
- Diagram problems and causes
- Listing possible solutions
- Screening possible solutions for feasibility

iii. **EXPERIMENTATION**: On-farm experiments are conducted in the fields of representative farmers and to examine a small number of experimental variables. Those experiments may be described and classified in a number of ways, but regardless of classification most of them progress from exploring production problems, to testing possible solutions, and then to verifying recommendations and demonstrating them with farmers.

iv. **ASSESSMENT/EVALUATION OF RESULTS:** The results of the on-farm experiments should be analyzed carefully. The analysis requires an assessment of farmers' reactions and opinions. A thorough agronomic interpretation and careful statistical and economic analysis will be helpful to this stage. The results of the assessment are then used to plan future research and to make recommendations for farmers.

v. **RECOMMENDATION:** When researchers are confident that they have enough information, then they can formulate recommendations to be demonstrated on a larger area in the farmers' fields (CIMMYT, 1988).

TYPES OF ON-FARM RESEARCH

On-farm research differs according to the chosen objective, the selected location, the type of design used and the degree of farmer involvement into:

Type 1: Researcher-designed and Managed Trial –The objective is to assess the performance of the new technology under various biophysical conditions. When research has shown promising results in on-station trials, researcher would wants to evaluate the new technology in multi location as the on-station trial does not represent the wide range of conditions (e.g. soil fertility, weed flora, altitude, rainfall, farmer's conditions).

Type 2: Researcher-designed and Farmer-managed Trials- The objective is to get farmer feedback on specific prototypes and to conduct economic analysis. Type 1 trials have confirmed that the new technology will work in farmer's conditions, therefore it is planned to implement the trial on a wider scale with active involvement of the farmers. Researchers are interested in getting the information on biophysical, economical and farmers' assessment of the technology.

Type 3: Farmer-designed and Managed Trials- Here, farmers carry out the experiments on their own; the objective of type 3 is to assess farmer innovation and acceptability. Farmers are aware of a given technology, they like what they see and would like to experiment it by themselves or farmers are aware of a problem and would try some methods to solve them. Researchers want to know to which extent and how a technology is adapted by farmers. Constant monitoring and recording of farmer's comments is necessary.

These trial types are not strictly defined; rather they are points along a continuum. However, you do not need to start with type 1 and proceed to type 3. Farmer participation is furthermore relevant to type 1 trial as well.

Table 1: Types of On-farm Research According to Different Objectives

OBJECTIVES	TRIAL	TRIAL DESIGN	TRIAL MANAGEMENT TYPE
Biophysical feasibility	1	Researcher-led	Researcher-led
Profitability, farmers Assessment of prototype	2	Researcher-led	Farmer-led
Acceptability: Farmers' own innovations	3	Farmer-led	Farmer-led

Source: Rudebjer, 2001

MEASUREMENTS OF ON-FARM EXPERIMENTAL DATA

In on-farm research, we can distinguish between three types of measurement.

(I) Measurement of the type that are taken in on-station trials. These are usually yield components, time to flowering, milk yields, disease scores, and so on.

(II) Measurements of concomitant variables. These can be at a plot level, for example problems of water logging, or at a farm level, for example rainfall or soil type. Some variables, such as dates of sowing and weeding, and other management practices may be at either level.

(III) Measurements of the farmers' opinions. These are gotten from informal discussions or questionnaires (SSC, 1998).

ANALYSIS OF ON-FARM DATA

As with the design, the analysis of the data will use a mixture of methods that are appropriate for the analysis of experimental and survey data. The analysis can be viewed in three stages:

(I) Analysis of Questionnaire-type Data- These set of data results from interviews and other observations. This information is normally at the farmer level, though some questions can relate to particular plots.

(II) Analysis of Yield Type Data- This information is mainly at the plot level, though with some observations at the farm level.

(III) Combination of (I) and (II) Above- using the results from interviews to understand the variation in yield type data.

The type of trial will dictate the proportion of time spent at each stage. One extreme might be farmer designed and managed trial, within which, the main objectives relate to their choices and opinions. Most of the analysis effort would therefore be on (I) above. In some researcher-designed and managed trials the yield data is of particular importance, in which case most of the time is spent on (II).

Experiments with sufficient within-site replication and detailed measurements of yield response can have separate within-site analyses initially, then a combined analysis. This is usually only the case for researcher-designed and managed trials. Others will use the data within a single analysis.

However there are two main differences between on-station and on-farm trials that have a bearing on the analysis. One is that with on-farm trials we expect a farm by treatment interaction, and one of the objectives of the trial is often to explore this interaction. The other difference is that there is now variation at different levels, there is variation between farms because of characteristics such as different agro climatic conditions, management practices, as well as variation between plots within farms. A statistical tool like PCA would be appropriate to analyze as much of such variation as possible.

STATISTICAL ANALYSIS

Statistical analysis refers to a collection of methods used to process large amounts of data. Approaches used may range from some simple analyses on different subgroups of the data to more sophisticated modeling of the whole data set. The analysis is often to evaluate relationships between biophysical responses, environmental, and social variables.

The data are also used to understand reasons for farmer assessments. These may be turned into decision trees for farmers or maps of recommendation domains.

In analyzing on-farm trials data we should be ready to:

(I) split the data up into subsets, e.g. groups of similar farms;

(II) Omit particular plots, e.g. the farmer's own treatment; or particular farms;

(III) Pay close attention to comments made about individual plots, e.g. "crop eaten by animals" may mean that a recorded yield of zero should be treated as a missing value;

(IV) use additional information, both at the farmer level and at the plot level, e.g. farmers may be classified as wealthy or poor, or plots may have information about pest damage;

(V) In the absence of within-farm replication, use treatment contrasts at the farm level to investigate the farm by treatment interaction, or investigate the interaction using the additional farmer information (as in (iv) above);

(VI) report on, and possibly follow up on, particular farmers who show interesting results (SSC, 1998).

PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components (Davis, 1986).

The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to the preceding components (Cattel, 1966).

PCA was invented in 1909 by Karl Pearson. Now it is mostly used as a tool in exploratory data analysis and formatting predictive models.

APPLICATIONS OF PRINCIPAL COMPONENT ANALYSIS

1. PCA as a multivariate technique can be use in analyzing relationships among several quantitative variables.
2. PCA can be use to analyze variables that are measured on different units.
3. PCA provides information about the relative importance of each variable in characterizing the objects.
4. PCA is used to reduce the number of variables of the data set, but retain most of the original variability in the data. A small number of these new variables will usually be sufficient to describe the observational objects (Rencher, 2002).

The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible (Harper, 1999). PCA can be done by eigenvalue decomposition of a data covariance matrix or singular value decomposition of data matrix; usually after mean centering (and normalizing or using Z-score) the data matrix for each attributes.

The results of a PCA are usually discussed in terms of component scores, sometimes called factor scores (the transformed variables values corresponding to a particular data point) and loading (the weight by which each standardized original variable should be multiplied to get the component score (Jolliffe, 2002).

PCA is the simplest of the true eigenvector based multivariate analyses. Often, it's operation can be thought of as revealing the internal structure of the data in a way that best explains the variance in the data, when a multivariate dataset is visualized as a set of coordinate in a high-dimensional data space (1 axis per variable).

PCA can supply the user with a lower-dimensional picture, a "shadow" of this object when viewed from it's (in some sense) most informative view point. This is done by using only the first few principal components so that the dimensionality of the transformed data is reduced (Stathis and Myrinidis, 2009).

USING PCA TO ANALYSE ON FARM RESEARCH DATA (A Hypothetical Example).

Chemical and textural properties were measured on soil from 18 farmers' fields in Yamrat, Bauchi State, Nigeria (Table 1). The table has 18 observational units (Fields), each with measured variables (soil characteristics).

The questions which arise are:

1. which soil properties are correlated (relationship)
2. which soil properties contribute most to the overall variance in soil characteristics
3. How the number of variables can be reduced without losing too much information.

Table 1: Soil Characteristics of 18 Farmers' Fields in Yamrat, Bauchi State, Nigeria.

Field	pH	OC	TN	P	K	Ca	Mg	Mn	SAND	SILT	CLAY
1	5.7	0.49	0.044	5.2	0.17	2.25	0.57	0.07	42	48	10
2	7.1	0.39	0.039	1.1	0.29	4.30	1.12	0.07	54	30	16
3	6.0	0.54	0.045	24.4	0.31	2.66	0.71	0.07	48	44	8
4	5.5	0.34	0.035	2.2	0.21	2.10	0.63	0.10	54	38	8
5	6.2	0.54	0.045	3.1	0.32	4.40	1.01	0.10	52	32	16
6	5.8	0.32	0.037	4.0	0.15	1.88	0.42	0.04	68	26	6
7	6.0	0.29	0.032	10.3	0.38	4.91	0.89	0.08	58	30	12
8	6.1	0.27	0.045	4.4	0.31	4.02	0.94	0.07	58	30	12
9	5.9	0.21	0.039	12.3	0.15	1.81	0.57	0.03	68	26	6
10	6.4	0.10	0.025	7.9	0.18	2.59	0.68	0.05	72	22	6
11	6.3	0.45	0.044	4.9	0.23	2.42	0.70	0.04	68	22	10
12	5.8	0.18	0.039	11.0	0.17	2.66	0.67	0.10	58	34	8
13	6.6	0.25	0.030	2.9	0.26	2.73	0.78	0.07	60	32	8
14	6.2	0.66	0.058	2.5	0.17	4.34	0.92	0.10	52	34	14
15	5.3	0.09	0.038	33.2	0.17	3.38	0.73	0.10	60	30	10
16	6.1	0.52	0.043	34.5	0.32	4.40	1.01	0.10	52	36	12
17	6.0	0.22	0.030	27.5	0.14	2.55	0.63	0.10	68	24	8
18	6.6	0.47	0.042	4.1	0.18	4.14	0.91	0.15	54	36	10

(Mutasaers *et al.,* 1997).

Table 2: Correlation Coefficients of Soil Characteristics of 18 Farmers' Fields

	OC	TN	P	K	Ca	Mg	Mn	SAND	SILT	CLAY
pH	0.25	–0.01	–0.41	0.30	0.39	0.60	–0.01	0.04	–0.22	0.39
OC		0.77	–0.19	0.30	0.36	0.39	0.25	–0.66	0.52	0.55
TN			–0.11	0.08	0.33	0.33	0.19	–0.56	0.42	0.53
P				0.01	0.02	–0.06	0.18	0.03	0.04	–0.17
K					0.64	0.67	0.03	–0.37	0.17	0.56
Ca						0.91	0.52	–0.39	0.08	0.83
Mg							0.44	–0.40	0.08	0.86
Mn								–0.47	0.39	0.36
SAND									–0.92	–0.53
SILT										0.16

(Mutasaers *et al.,* 1997).

The next step is to compute correlation among the soil characteristics in order to reveal relationships between variables (Table 2). This provides an insight into the relationship between the variables. In this case, a positive correlation is seen between soil Clay, Ca and Mg contents for the soil in these fields.

Table 3: Eigen-values of the Correlation Matrix and the Proportion and Total of Variance Explained by the Five Largest Principal Components

Principal component	Eigenvalue	Difference	Proportion	Cumulative
PRIN1	4.78151	2.55972	0.434683	0.434683
PRIN2	2.22179	0.79831	0.201981	0.636664
PRIN3	1.42348	0.56145	0.129408	0.766071
PRIN4	0.86204	0.07773	0.078367	0.844439
PRIN5	0.78431	—	0.071301	0.915740

(Mutasaers *et. al.,* 1997).

The correlation matrix can now be converted into principal components. The coefficients of the principal components are the eigenvectors of the correlation matrix. Thus, each principal component is a linear combination of the original variable. There are as many principal components that can be computed as there are original variables. However, only the most important ones are of relevance for further analysis. The importance of the principal components calculated from their eigenvalues and their contribution in explaining the overall variance (Table 3).

Table 4: Principal Components with their Percentage Variability.

Principal Component	Percentage Variability (%)
PRIN 1	43.4
PRIN 2	20.2
PRIN 3	12.9
PRIN 4	7.8
PRIN 5	7.1

In the above example, Principal component 1 (PRIN 1) explains 43.4 % of the overall variance, PRIN 2, PRIN 3,PRIN 4, PRIN 5 contribute an additional 20.2 %, 12.9 %, 7.8 % and 7.1 % respectively. All five principal components together in this case explain 91.4 % of the overall variance, and the first three already explain 76.5 %. The remaining six principal components (PRIN 6 to PRIN 11) explain only the residual 8.6 %.

Utmost care is required in the interpretation. Each principal component is a mathematical number without a defined unit or a definite biological meaning. It is a combination of variables measured on different scales. The relative contribution of one or other variables to each principal component, gives an indication of their meaning.

Table 5: Eigenvectors of principal components representing a linear combination of the original variables

	PRIN1	PRIN2	PRIN3	PRIN4	PRIN5
pH	0.175	0.451	-0.290	0.125	-0.265
OC	0.337	-0.210	-0.336	-0.006	0.221
TN	0.288	-0.265	-0.284	0.151	0.583
P	-0.054	-0.151	0.670	-0.166	0.408
K	0.285	0.225	0.142	-0.689	-0.094
Ca	0.379	0.247	0.257	0.093	0.078
Mg	0.391	0.302	0.145	0.049	0.001
Mn	0.237	-0.153	0.408	0.619	-0.305
SAND	-0.349	0.378	0.011	0.152	0.268
SILT	0.223	-0.513	-0.010	-0.200	-0.398
CLAY	0.405	0.165	-0.006	0.051	0.191

(Mutasaers *et al.,* 1997).

The biological meaning of principal components can tentatively be assessed from the relative contribution of different soil characteristics of each principal component according to the eigenvector (Table 5). PRIN 1 is most strongly affected by the soil clay content and the Ca and Mg content, which were seen to be correlated earlier on. PRIN 2 is most strongly associated with soil pH and soil silt content. PRIN 3 is closely related to soil P.

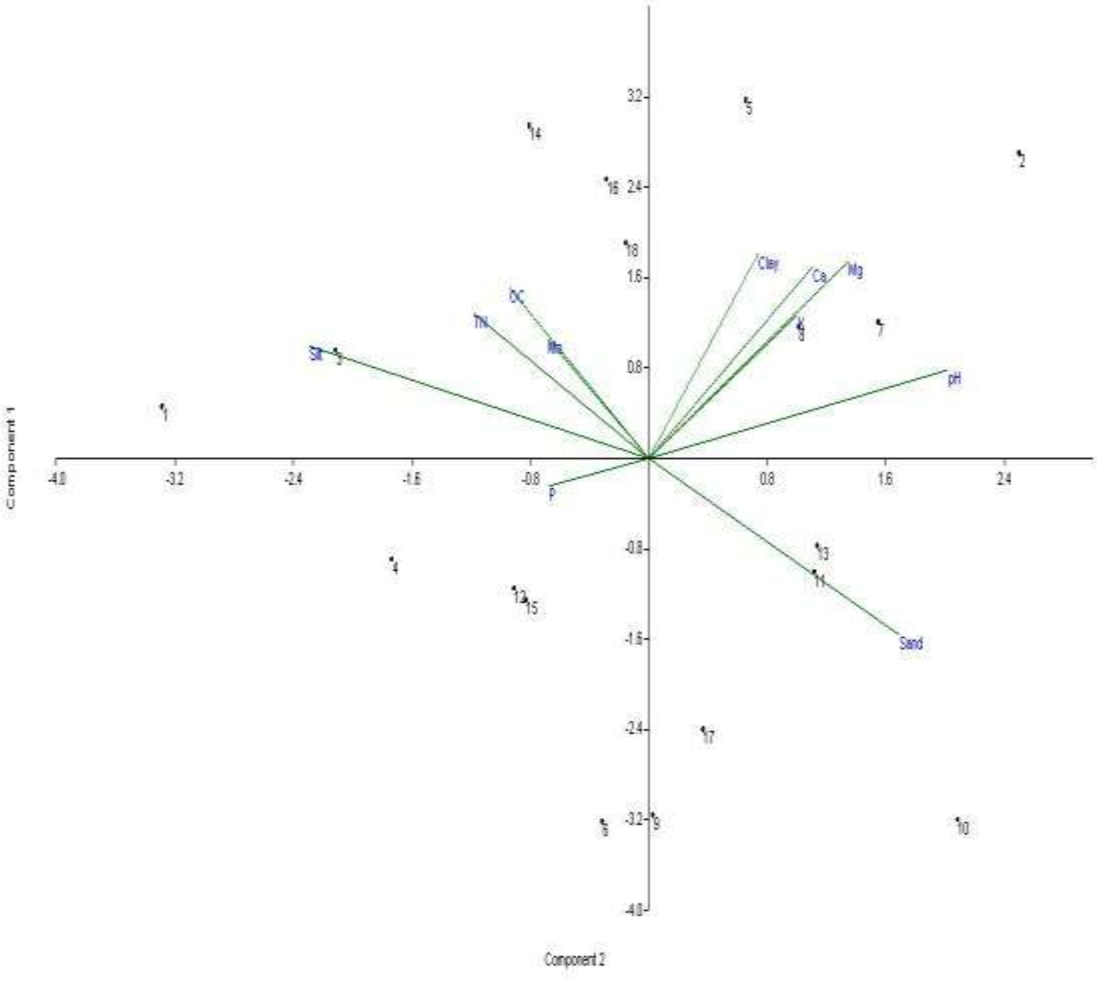

Figure 1: Showing the PCA scatter diagram (bi-plot) for the variables.

Principal Component Analysis chooses the first PCA axis as that line that goes through the centroid, but also minimizes the square of the distance of each point to that line. Thus, in some sense, the line is as close to all of the data as possible. Equivalently, the line goes through the maximum variation in the data.

The second PCA axis also must go through the centroid, and also goes through the maximum variation in the data, but with a certain constraint: it must be completely uncorrelated (i.e. at right angles, or 'orthogonal' to PCA axis 1). For each field, values for the different components, called "principal component scores" are then calculated. These scores are obtained by multiplying the original data matrix with the principal component matrix. The new variables (scores) have Zero mean and a variance equal to the corresponding eigenvalues.

Table 6: Standardized Principal Component scores used as three new variables representing 76.6 % of the variance from the original eleven soil characteristics.

Field	PRIN1	PRIN2	PRIN3
1	0.211	-2.205	-1.021
2	1.238	1.678	-0.698
3	0.439	-1.419	0.075
4	-0.407	-1.164	-0.070
5	1.451	0.438	-0.185
6	-1.470	-0.211	-1.116
7	0.554	1.039	1.039
8	0.534	0.680	-0.073
9	-1.442	0.018	-0.660
10	-1.460	1.397	-0.016
11	-0.458	0.749	-1.380
12	-0.524	-0.611	0.442
13	-0.352	0.762	-0.377
14	1.348	-0.542	-1.140
15	-0.570	-0.557	2.232
16	0.131	-0.193	1.557
17	-1.094	0.245	1.304
18	0.874	-0.104	0.092

A standardization of the scores to unit variance is often recommended. The standard scores are new variables which may be used for further analysis for example, we could used the three new variables (PRIN 1, PRIN 2 and PRIN 3) in a multi-regression analysis instead of the 11 original soil parameters to relate yield of a particular crop to soil characteristic. The three new variables comprise 76.6 % of the total variance from the original 11 soil characteristic; this has made us reduce the data complexity without losing much information.

The special feature of PCA is its potential to reduce a large number of variables to a few new values, which compare most of the original overall variance. Its major weakness is that the new variable are purely mathematical concepts (Principal components or Scores), which have no units and are often difficult to interpret biologically (Jackson, 1993).

SOME STATISTICAL SOFTWARE USED FOR PCA ANALYSIS

- GENSTAT - General Statistics
- AGSTATS - Agricultural Statistics
- PAST - Pale Ontological Statistics

RECOMMENDATIONS AND CONCLUSION

Principal Component analysis, which is used to reduce the number of observed variables into a smaller number of variables that account for most of the variance in the data set, will find useful application in on-farm experimentation since a lot of variables are encountered.

PCA also reduces the cost of data processing. It is recommended that plant scientists especially those involved in characterization and evaluation of plant genetic resources would find its application useful in reducing the complexity of the variables; establishing only the significant ones which could be used as a distinguishing feature for the crop plant. Such important features can be used by plant conservationist as a guide for active and base collections of plant genetic resources materials, thereby reducing the "work load" and financial cost of incorporating all original variables in research involving such test plants.

The correct design of experimental studies, the selection of the appropriate statistical analysis of data and the efficient presentation of results are key to the good conduct and communication of science. On-farm research has shown to be site specific, broader and faster adoption as compared to on-station research.

In on farm research, during data collections, one encounter situations where there are large number of variables. A good statistical analysis would be needed to make valid conclusion about such research. Principal component analysis is a powerful tool for reducing a number of observed variables into a smaller number of variables that account for most of the variance in the data set. It is particularly useful when you need a data reduction procedure that makes no assumptions concerning an underlying causal structure that is responsible for co-variation in the data.

Lastly PCA allows the use of variables which are not measured in the same units for example elevation, concentration of nutrients, temperature, pH, and so on, thereby making it a good tool for On-Farm research data analyzes.

REFERENCES

Ashby, J. A. (1986). Methodology for the participation of small farmers in the design of on-farm trials. *Agricultural Administration,* **22**: 1-19.

Cattell, R. B. (1966). The screen test for the number of factors. *Multivariate Behavioral Research,* **1**: 245-276.

CIMMYT. (1988). From Agronomic Data to Farmer Recommendation: An Economic Training Manual. CIMMYT Economics Program, Mexico. 184p.

Davis, J. C. (1986). Statistics and Analysis in Geology. John Wiley and Sons, New York. 832p.

Gomez, K. A. and Gomez, A. A. (1984). Statistical Procedures for Agricultural Research. John Wiley and Sons, New York. 355p.

Harper, D. A. T. (1999). Numerical palaeobiology. John Wiley and Sons, New York. 48p

Jackson, D. A. (1993). Stopping rules in Principal Component Analysis: a comparison of heuristical and statistical approaches. *Ecology*, 74(8): 2204 - 2214

Jolliffe, T. T. (2002). Principal Component Analysis. Springer, New York.487p.

Mutasaers, H. J. W., Weber, G. K., Walker, P. and Fisher, N. M. (1997). A Field Guide for On-Farm Experimentation. IITA Publishing unit, Ibadan. 222p.

Rencher, A. C. (2002). Methods of Multivariate Analysis. John Wiley and Sons, New York. 675p.

Rudebjer, P. G. (2001). Training course on participatory on-farm experimentation and integrated approaches to land management. **In**: Lampung B. and Kutabumi H. (eds). Proceedings of the international centre for research in agro forestry. Southeast Asian Regional Programme, Indonesia. pp 201-218.

Spencer, D. S. C. (1993). Collecting meaningful data on labour use in on-farm trials in sub-Saharan Africa. *Experimental Agriculture,* **29**: 39-46.

Stathis, D. and Myronidis, D. (2009). Principal component analysis of precipitation on Thes-saly region (Central Greece). *Global NEST Journal,* **11**(4): 467-476.

Statistical Services Centre, (1998). On-Farm Trials-Some Biometric Guidelines. Statistical Services centre, University of Reading, Reading. 15p.

Wuest, S., Miller, B., Guy, S., Karow, R., Veserth and Wysochi, D. (1991). Use of On-Farm Research by Farmers for Technology Development. Washington State University Press. Washington D.C. 234p.