A Decision Support System for Agriculture Using Natural Language Processing (ADSS)

Prof. Mrs. J.R.Prasad, Prof. R.S.Prasad, Dr. U.V.Kulkarni.

Abstract
The agricultural sector which is core part of the Indian economy, represents 35% of The impact of climate change on agriculture is expected to impact on agricultural productivity and shifting crop patterns
This paper suggests development of a decision support system for agriculture based on the natural language processing. The analytical data about the rainfall pattern, soil structure of the area will be maintained at back end, the system will retrieve the information based on the interaction with the user, which will be a farmer in this case. The authors aim to provide a user friendly decision support system.

Index Terms
Crop forecast, Decision Support System, Natural Language Processing, whether.

I. INTRODUCTION

The choice of any agricultural system is made by considering three main issues: (a) the socioeconomic context (needs, resources, markets, etc.), (b) the natural resource base (current vegetation, soils, water, topography, etc.), and (c) the climate. To some extent, (a) and (b) are themselves governed by (c), so that agricultural production is inextricably linked to climate, as shown in figure 1. More specifically, the success or failure of a given season's crop is highly dependent on that season's weather.
A general seasonal forecast (e.g. will it be wetter or drier than average) may help in deciding what crops to plant, whereas more precise information (e.g. the date of onset of rains), may help to decide on which date to plant.
The spatial and temporal scales of information are also important, and vary from user to user. There is growing awareness and understanding in the end-user community of the economic and social costs and benefits of access to such climate and agricultural sciences.

Mrs. J.R.Prasad is working as an Assistant Professor at the Department of Computer Engineering, VIIT, Pune. She is a member of IAENG: No. 62346
Prof. R.S.Prasad is Assistant Professor and Head of the Department of Computer Engineering, VIIT, Pune. He is a member of IAENG: No. 62347
Dr. U.V.Kulkarni is a Professor at Department of Computer Engineering and Dean, Academics at SGGS, Nanded.

Nevertheless, there is a major need for enhanced research in ways to facilitate the efficient and timely transfer of up to date climate and agricultural information to potential end-users of is information and to identify ways to increase their capacity to use the information to improve their agricultural productivity, resource management practices and to support sustainable development.

A. Goal

To utilize the ability to predict crop variability on the scale of months to a year to improve management and decision-making in respect of crop production at farm and up to national scales.

B. Objectives

- To enhance the interface between the agricultural, meteorological and socio-economic communities and farmers and other stakeholders;
- To establish the forecasting needs for decision making in crop production;
- To develop the capacity for integrated climate and agricultural simulation and prediction for decision making for a range of farming systems; and
- To demonstrate the capability and value of improved climate prediction for improving crop production at farm to national scales as a proof of concept.

II. THE NEED FOR ADSS

The issues of global food security and the environmental consequences of increasing food production to the required levels are a major international concern. The growth in human population over the past century has been closely associated with an increase in food production. Generally, production has increased slightly faster than population so that, for example, 5.8 billion people today have 15% more food than a population of 4 billion had 20 years ago. There is now a reasonable degree of certainty in the projections of population for the next 20-30 years but these become increasingly uncertain with time. Overall the big changes in demography will occur in the next 2 or 3 decades with slower changes occurring after 2050 or thereabouts. Population will increase by about 1 billion per decade for the next 2 to 3 decades with most of this increase occurring in the less developed nations and almost none in the developed countries of Europe and North America. Thus a population of 5.8 billion today will rise to about 8 billion by 2025. There will be large regional differences in the expansion of population with large percentage increases in Africa and West Asia whereas Central Asia will decline in share of population (but not absolute numbers) and South and South East Asia will eventually remain almost constant as a proportion of the total (Fischer and Heilig, 1997).

Given the close association of population and food production, it is possible to estimate the required grain production (wheat, rice and maize together supply about 60% of the total carbohydrate). Overall, the average yield of cereal grains will need to rise from the current level of 2.8 t/ha to about 4.2 t/ha by 2025. This project is a timely initiative that will bring together two major scientific communities with potential benefits to many millions of people.

III. INTELLECTUAL FEASIBILITY OF THE PROJECT

The intellectual justification for the project is based on three main issues:
1. The improved capability for seasonal crop prediction.
2. The opportunity to forward-plan in response to climate forecasts to influence
productivity at field and regional scales; and
3. The need to minimize production and environmental risks and to increase possible
economic and social benefits for the farming community in order to mitigate the economic and social consequences of unforeseen climatic extremes.

IV. APPROACH TO DEVELOP A DECISION SUPPORT SYSTEM FOR AGRICULTURE.

Because there are many approaches to decision-making and because of the wide range of domains in which decisions are made, the concept of *decision support system* (DSS) is very broad. A DSS can take many different forms. We aim to develop a DSS which will be a computerized system for helping farmers make decisions. A user will take a decision which will be a choice between alternatives based on estimates of the values of those alternatives. This system aims to help people working alone or in a group gather intelligence, generate alternatives and make choices. The system will support the choice making process involving the estimation, evaluation and/or the comparison of alternatives. In practice, references to DSS are usually references to computer applications that perform such a supporting role [1].

There is no universally accepted taxonomy of DSS either. Different authors propose different classifications. Using the relationship with the user as the criterion, Haettenschwiler [12] differentiates *passive*, *active*, and *cooperative DSS*. A *passive DSS* is a system that aids the process of decision making, but that cannot bring out explicit decision suggestions or solutions. An *active DSS* can bring out such decision suggestions or solutions. A *cooperative DSS* allows the decision maker (or its advisor) to modify, complete, or refine the decision suggestions provided by the system, before sending them back to the system for validation. The system again improves, completes, and refines the suggestions of the decision maker and sends them back to her for validation. The whole process then starts again, until a consolidated solution is generated.

The important types of decision support systems are listed as follows

- A **model-driven DSS** emphasizes access to and manipulation of a statistical, financial, optimization, or simulation model. Model-driven DSS use data and parameters provided by users to assist decision makers in analyzing a situation; they are not necessarily data intensive. *Dicodess* is an example of an open source model-driven DSS generator [14]
- A **communication-driven DSS** supports more than one person working on a shared task; examples include integrated tools like Microsoft's NetMeeting or Groove [15]
- A **data-driven DSS** or data-oriented DSS emphasizes access to and manipulation of a time series of internal company data and, sometimes, external data.
- A **document-driven DSS** manages, retrieves and manipulates unstructured information in a variety of electronic formats.
- A **knowledge-driven DSS** provides specialized problem solving expertise stored as facts, rules, procedures, or in similar structures. [13]. Our aim is to develop a data-driven DSS for agriculture industry.

V. ROLE OF NATURAL LANGUAGE PROCESSING IN IMPLEMENTING THE ABOVE DSS

In theory, natural language processing is a very attractive method of human-computer interaction. Early systems working with restricted vocabularies, worked extremely well, leading researchers to excessive optimism which was soon lost when the systems were extended to more realistic situations with real-world ambiguity and complexity.

Natural language understanding is sometimes referred to as an AI-complete problem, because natural language recognition seems to require extensive knowledge about the outside world and the ability to manipulate it. The definition of "understanding" is one of the major problems in natural language processing.

The Natural language processing approach to develop the said DSS will address solutions to the following sub-problems.

A. Segmentation

In most spoken languages, the sounds representing successive letters blend into each other, so the conversion of the analog signal to discrete characters can be a very difficult process. Also, in natural speech there are hardly any pauses between successive words; the location of those boundaries usually must take into account grammatical and semantical constraints, as well as the context.

B. Text segmentation

Written languages do not have single word boundaries either, so any significant text parsing usually requires the identification of word boundaries, which is often a non-trivial task.

C. Word sense disambiguation

Many words have more than one meaning; we have to select the meaning which makes the most sense in context.

D. Syntactic ambiguity

The grammar for natural languages is ambiguous, i.e. there are often multiple possible parse trees for a given sentence. Choosing the most appropriate one usually requires semantic and contextual information. Specific problem components of syntactic ambiguity include sentence boundary disambiguation.

E. Imperfect or irregular input

Foreign or regional accents and vocal impediments in speech; typing or grammatical errors, OCR errors in texts.

F. Speech acts and plans

Sentences often don't mean what they literally say; for instance a good answer to "Can you pass the salt" is to pass the salt; in most contexts "Yes" is not a good answer, although "No" is better and "I'm afraid that I can't see it" is better yet. Or again, if a class was not offered last year, "The class was not offered last year" is a better answer to the question "How many students failed the class last year?" than "None" is.

VI. IMPLEMENTATION OF THE PROJECT

Given the conceptual framework for the project, it is appropriate to demonstrate that the
concepts can be applied to practical situations through the execution of this project. This project would involve the application crop predictions to the suitable region of India. The
Selection of a region will depend upon the results of a sequence of tests for suitability.
The first point in identifying a region is to ensure that it represents a homogeneous
Agricultural-climatic zone. Then it must be clear that there is existing capacity in the region to support the research and management activities associated with the project. In particular, a local lead scientist would need to be identified, and “model” farms would need to be found for participation in the development and application of management strategies. The scientific suitability of the region would need to be demonstrated through a sequence of tests, which are described below. There must be sufficient historical data available to ensure that a statistically significant outcome could be expected from an experiment over a few years. This would be demonstrated through analyses and simulations of the use of climate predictions to support agricultural decision making.

A. Historical data

In order for a region to be considered for this project, there must be adequate historical records to support basic analyses of the relationship between climate and agricultural production and general social and economic situation. Thus some initial work may be necessary to quality control each data set and to ensure that the climate, agricultural and socio-economic data are compatible and collocated.

B. Statistical analyses

Using compatible climate and agricultural data, it is first necessary to demonstrate that the crop yield in the selected region has interannual variability correlated with variations in climate. This would ensure that there is some basic influence of climate that may be modified by suitable farm management practices. These statistical analyses need to allow for trends in crop production due to improvements in technology, and they need to focus on interannual variations alone.

C. Predictability

Having demonstrated a statistical relationship between crop production and climate, it is then necessary to show that there is at least some significant predictability of both climate
variations and regional crop yield. Such predictability is required to support any proposed
changes in farm management practices.

D. Simulation studies

Once the basic prediction models have been validated, it is then appropriate to use the crop model to demonstrate that the regional crop productivity is sensitive to some management process at the farm level. For example, it might be shown that varying the amount of fertilizer used on a crop according to seasonal climate variations can lead to increased productivity and profitability of farms in the region. The simulation studies would be carried out using historical climate data. Different management strategies could be tested to determine an optimal farm policy to account for climate variations and hanging socio-economic conditions. These studies could estimate the maximum benefit to cost ratio that could be expected in the demonstration project. The studies should also be repeated using hindcast seasonal climate predictions to estimate the probable benefit to cost ratio from the project. The difference between the two studies accounts for the imperfect accuracy of climate predictions in the region. These studies would provide an estimate of the required duration of a project to ensure that a statistically significant result was obtained; i.e. to ensure that the signal associated with varying farm management practices could be detected above the natural interannual variability of crop production in the region.

E. Operational prediction system

Having demonstrated the theoretical viability of the project, it is then necessary to establish an operational prediction system that will function robustly and efficiently throughout the duration of the project. An objective and repeatable climate prediction system is needed, with appropriate downscaling to link the large scale predictions to the input requirements of the crop models. Similarly, the crop prediction system needs to be stable and robust, so that it can be used consistently throughout the project to support farm-level decision making.

F. Refine management strategy with farm-level partners

The work described in sections 3 and 4 needs to be carried out in close co-operation between the modelers and the farm-level partners in the project. It is clearly necessary to Ensure that any farm-level strategy is practical, affordable and acceptable to farmers. Particular attention should be given to analyzing the economic, social, institutional and cultural factors related to the use of such information by farmers to improve management practices. Thus there would be iteration through activities 3 to 6, as the farm trategies are refined and the models adjusted to optimize the overall outcome.

G. Conduct a "Proof of Concept" prediction trial

It is clear that a substantial amount of work would need to be carried out before an actual prediction trial is conducted in a selected region. The trial would be carried out using consistent procedures and with full documentation, so that valid and robust results can be established. In particular, it would be important that the procedures for prediction and decision-making are objective, leading to an experiment that is repeatable.

VII. ISSUES IN DESIGNING THE PROJECT

1. Develop approaches to predict crop prediction and resource use (e.g., water) at field and regional scale Understand yield reducing factors, their impact on crop yields, and their relation to climate variability.
2. Need to tailor climate predictions for use by crop modelers and need crop modelers to modify their models to meet the constraints of what can be provided by climate prediction community
3. Need to consider scaling issues in working from sub-national to national scales
4. Need for sites where there are good records and a strong climate variability signal
5. Initially a minimum network of sites involving groundnuts and/or rice/wheat in India, rice in Northeast Thailand and China and wheat in Queensland Build on existing projects, regional capabilities and regional/international institutions
6. Essential for climate, agriculture and social scientists to collaborate at all stages and to involve client community at appropriate times Quality data availability and accessibility a necessary condition for site selection
7. Systems analysis of key decisions; factors influencing decision-making and attitudes to climate risks
8. Seek evidence of links between climates variability and production from initial data collection
9. Undertake analysis to demonstrate implications for decision-making at various scales; Include socio-economic analysis as appropriate
10. Client-community expectation and understanding
11. Need to maximize local participation by national teams of scientists and others at all levels to enhance capacity building
12. How current knowledge, perceptions and practices about how climate variability influence crop production
13. What factors (social and cultural systems)

ISBN: 978-988-98671-8-8

constrain ability of farmers to cope with climate variability and perception of factors
14. Undertake evaluation of the project exercise with clients

Outputs and outcomes

1. Case studies showing how improved capability in climate forecasting can be used in facilitating improved decision-making about crop production
2. Data bases in standard form and available to others - climate, crop production
3. Interim workshops with scientists and end-user communities
4. Final top-level conference to present science and engage policy-makers and other End-users
5. Scientific papers, improved networking and enhance
6. Building scientific capacity to conduct inter-disciplinary studies that involve the participation of clients
7. Evaluation of the value of improved climate variability prediction
8. Establishment of new trans-disciplinary networks

CONCLUSION

Thus the authors foresee improved capability for seasonal crop prediction, Opportunity to forward-plan in response to climate forecasts to influence productivity at field and regional scales. This is high time to minimize production and environmental risks and need to increase possible economic benefits for the farming community to alleviate social consequences of unforeseen climate extremes

REFERENCES

1. Alter, S. L. (1980). Decision support systems: current practice and continuing challenges. Reading, Mass., Addison-Wesley Pub.
2. Druzdzel, M. J. and R. R. Flynn (1999). Decision Support Systems. Encyclopedia of Library and Information Science. A. Kent, Marcel Dekker, Inc.
3. Finlay, P. N. (1994). Introducing decision support systems. Oxford, UK Cambridge, Mass., NCC Blackwell; Blackwell Publishers.
4. Turban, E. (1995). Decision support and expert systems: management support systems. Englewood Cliffs, N.J., Prentice Hall. ISBN 0-024-21702-6
5. Little, J.D.C.(1970, April). "Models and Managers: The Concept of a Decision Calculus." *Management Science*, Vol.16,NO.8
6. Keen, P. G. W. (1978). Decision support systems: an organizational perspective. Reading, Mass., Addison-Wesley Pub. Co. ISBN 0-201-03667-3
7. Moore, J.H.,and M.G.Chang.(1980,Fall)."Design of Decision Support Systems." *Data Base*,Vol.12, Nos.1 and 2.
8. Sprague, R. H. and E. D. Carlson (1982). Building effective decision support systems. Englewood Cliffs, N.J., Prentice-Hall. ISBN 0-130-86215-0
9. Keen, P. G. W. (1980). Decision support systems: a research perspective. Decision support systems : issues and challenges. G. Fick and R. H. Sprague. Oxford ; New York, Pergamon Press.
10. Power, D. J. (1997). What is a DSS? The On-Line Executive Journal for Data-Intensive Decision Support 1(3).
11. Power, D.J. A Brief History of Decision Support Systems DSSResources.COM, World Wide Web, version 2.8, May 31, 2003.
12. Haettenschwiler, P. (1999). Neues anwenderfreundliches Konzept der Entscheidungsunterstützung. Gutes Entscheiden in Wirtschaft, Politik und Gesellschaft. Zurich, vdf Hochschulverlag AG: 189-208.
13. Power, D. J. (2002). Decision support systems: concepts and resources for managers. Westport, Conn., Quorum Books.
14. Gachet, A. (2004). Building Model-Driven Decision Support Systems with Dicodess. Zurich, VDF.
15. Stanhope, P. (2002). Get in the Groove: building tools and peer-to-peer solutions with the Groove platform. New York, Hungry Minds
16. William R. Hersh, "Information Retrieval: A Health and Biomedical Perspective (2nd edition)", *Springer Verlag, 2003*, ISBN 0-387-95522-4
17. W. Bruce Croft and John Lafferty (editors), "Language Modeling for Information Retrieval", *Kluwer Academic Publishers, 2003*, ISBN 1-4020-1216-0, The Kluwer International Series on Information Retrieval, Volume 13