PREDICTING PRODUCTION OF CRUDE PALM OIL BASED ON WEATHER ATTRIBUTES

ONG EI LIN

UNIVERSITI UTARA MALAYSIA

2009

PREDICTING PRODUCTION OF CRUDE PALM OIL BASED ON WEATHER ATTRIBUTES

A project submitted to the Faculty of Information Technology in partial fulfillment of the requirement for the degree
Master of Science (Information Technology)
Universiti Utara Malaysia

By

ONG EI LIN

PERMISSION TO USE

In presenting this project in partial fulfilment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the University Library may make it freely available for inspection. I further agree that permission for copying of this project in any manner, in whole or in part, for scholarly purposes may be granted by my supervisor, in her absence, by the Dean of the Faculty of Information Technology. It is understood that any copying or publication or use of this project or parts thereof for financial gain should not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my project.

Request for permission to copy or to make use of material in this project, in whole or in part should be addressed to:

Dean of the Faculty of Information Technology
Universiti Utara Malaysia
06010 UUM Sintok
Kedah Darul Aman

ABSTRACT

In hydrological cycle, water is the important source for rainfall forecasting. Hence, rainfall forecasting becomes a critical issue in equatorial country like Malaysia. Rainfall can affect environment and plantation activities and agriculture in Malaysia. In Malaysia, Meteorological Department collects weather information for each state in Malaysia. Rainfall prediction is important because it can produces the useful information to the palm oil production and recommending appropriate prevention climate change such as floods warning advise as well as managing water resource operations. For instances, Malaysian Palm Oil Board (MPOB) has given a lot of information about the palm oil production and its effect due to the climate changes. In this study, the analysis on weather data from the year 1996-2005 for five states such as Kedah, Kelantan, Malacca, Penang and Perak was carried out. In the initial study, regression analysis has been conducted to determine the relationship of the weather attributes and palm oil variable such as Fresh Fruit Bunches, Oil Extraction Rate and Crude Palm Oil production. However, the results were not so encouraging, therefore CBR approach has been attempted to solve the current problem then reuse the information and knowledge based that have been stored in the cases. The similarity measurement can be determined effectively between cases. Therefore, similarity measurement between cases in the rainfall and palm oil case base is the important element in CBR. The performance of each similarity measure is evaluated based on attribute’s weight, and classification accuracy, In general, the similarity values achieved at most is 99.33%.

ABSTRAK

Dalam kitaran hidrologi, air merupakan sumber utama bagi taburan hujan. Oleh demikian, ramalan taburan hujan merupakan suatu isu yang penting kepada negara-negara khatulistiwa seperti Malaysia. Taburan hujan banyak mempegaruhi aktiviti sector perladanan and pertanian di Malaysia. Jabatan kaji cuaca telah menyediakan maklumat tentang taburan hujan bagi setiap negeri Malaysia. Maklumat yang didapati daripada ramalan taburan hujan boleh meningkatkan pertumbuhan kelapa sawit, memberi amaran bencana alam seperti banjir dan mengurus saluran air dengan sempurna. Disamping itu, Malaysian Palm Oil Board (MPOB) juga banyak memberi maklumat untuk meningkat pertumbuhan dan productiviti kelapa sawit, disamping itu menunjukkan memusnahkan pertumbuhan kepala sawit kesan dari bencana alam. Data taburan hujan daripada tahun 1996 hingga 2005 dianalisis berdasarkan lima negeri yang terlibat iaitu Kedah. Kelantan, Melaka, Pulau Pinang and Perak. Regression analysis perlu dilaksanakan terlebih dahulu untuk menentukan perhubungan antara parameter taburan hujan dengan parameter kepala sawit yang terdiri daripada fresh fruit bunches dan oil extraction rate. Walaubagaimanapun, hasil yang di dapati tidak memuaskan. oleh itu, pendekatan case based reasoning (CBR) digunakan. CBR merupakan satu pendekatan yang berupaya menyelesaikan masalah semasa, menggunaan maklumat dan pengetahuan sedia ada dalam kes tersebut. Dalam teknik CBR, similariti di antara kes bagi taburan hujan dan kelapa sawit boleh ditentukan melalui pengukuran similariti dan ia merupakan elemen utama dalam kajian CBR. Kebolehupayaan setiap pengukuran similariti dinilai berdasarkan pemberatan parameter dan ketepatan pengkelasan. kesimpulannya, nilai similariti yang tertinggi dalam kajian ini ialah 99.33%.

ACKNOWLEGMENT

Firstly, I would like to express my deep gratitude to both of my supervisors, Associate Professor Fadzilah Siraj and Puan Nur Azzah Binti Abu Bakar, who guide me patiently in this endeavour, beside their wonderful support and assisting me to complete this project.

Secondary, I would like to extend my thanks and gratitude to Malaysian Meteorological Services Centre and Malaysian Palm Oil Board, have give their cooperation during this research, where they have fully support in all the needed information for this project.

Finally, I owe the warmest gratitude to my beloved family members and friends, who have give the great supporting me throughout this project.

Last and not least, a special appreciate to all those, encouraged and helped me to complete this project. Thanks to all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
BN	Bayesian Network
CBR	Case-Based Reasoning
FFB	Fresh Fruit Bunches
NN	Neural Network
MPOB	Malaysia Palm Oil Board
OER	Oil Extraction Rate
SIM	Similarity

CHAPTER 1

INTRODUCTION

1.1 Background

Oil palm is the most productive oil seed crops in which one hectare of oil palm can produce more than 4.5 metric tons of palm oil. Malaysia and Indonesia are the two countries responsible for over 80% of the world palm oil production; the palm oil sector is one of the most important commodities in Malaysia. Malaysian palm oil has been accepted globally and recognized internationally as one of the major oils and fats in the world. The Malaysian Palm Oil Board (MPOB) is responsible for forecasting Malaysia's crude palm oil (CPO) production which is expected to reach more than 20 million tones. The Malaysian Palm Oil Association (MPOA), on the other hand, has listed out the important of palm oil industry for Malaysia and the world as follows:

1. Third pillar of nation's economy and catalyst for rural development and resultant political stability.
2. Provides direct employment to 400,000 people plus other multiplier effects and spin-offs.
3. Major foreign exchange earner RM27 billion in 2003.
4. Feeding the world Malaysian palm oil is consumed in more then 150 countries world-wide.

The contents of the thesis is for internal user only

REFERENCES

Aamodt, A. (1993). Explanation-Driven Retrieval, Reuse and Learning of Cases. *Proceedings of European Workshop on Case-Based Reasoning*, p. 279-384.

Aamodt, A and Plaza, E. (1994). Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches, p. 39–59.

Avesani, P., Perini, A. & Ricci, F. (2000). Combining CBR and Constraint Reasoning in Planning Forest Fire Fighting.

Bartlmae, K., & Riemenschneider, M. (2005). Case Based Reasoning for Knowledge Management in KDD-Projects Concepts, Organizational Setting, Categorization into KM and Application in the case of Knowledge Discovery in Databases.

Basak, J., Sudarshan,A., Trivedi,D. (2004). Weather Data Mining Using Independent Component Analysis, Journal of Machine Learning Research 5 (2004), p. 239–253.

Basiron, Y. (2007). Palm oil production through sustainable plantations. Malaysian Palm Oil Council, Jalan Perbandaran, Kelana Jaya, Selangor, Malaysia.

Bhaskar, N. (2001). Estimation of Rain Height from Rain Rate using Regression-based Statistical Model: Application to SeaWinds on ADEOS-II. *Master of Science thesis, Department of Electrical Engineering and Computer Science*.

Breese, J., Heckerman, D. and Kadie, C. (1998).Uncertainty in Artificial Intelligence: Proceedings of the Fourteenth Conference.

Bremnes, J. B. (2006). A comparison of a few statistical models for making quartile wind power forecasts. Wind Energy 9, p. 3–11.

Bogner, K., Thielen, J., deRoo, A.(2007). Evaluation of an ensemble based medium range probabilistic flood forecasting system for the Upper Danube catchment. *Geophysical Research Abstracts*. Joint Research Centre, Natural Hazards Action (NAHA).

Cattle, H. (1999). Experimental forecast of east African rainfall for October-December 1999. Retrieved September, 2007, from: www.metoffice.gov.uk/research/seasonal/regional/east_africa/pdf/east_africa_1999.pdf.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Thomas, R., Shearer, C., & Wirth, R., (2000). CRISP-DM 1.0 Step-by-step data mining guide.

Clokea, H. L., & Pappenbergerb, F. (2008). Evaluating forecasts of extreme events for hydrological applications: an approach for screening unfamiliar performance measures. *Meteorological Applications*. Wiley InterScience, Einfalt T., Jensen T. S., Klinting A. (2004). Combining NWP data, radar data, raingauge data and a hydrological-hydrodynamical model for flood forecasting in Turkey. *Proceedings of ERAD, Copernicus GmbH*, 6-10 September, Sweden, p. 522–524.

Eric, W. (1997). U.S. Global Change Research Information Office, Suite 250, 1717 Pennsylvania Ave, NW, Washington, DC 20006.

Fairhust, T. H. (1999). Twenty Two Tips for Practical Oil Palm Planters. Better Crops International, vol.13

Fatih, K. (2007). Rainfall forecasting system of Tefer Project, State Hydraulic Works, *Investigation and Planning Department, Flood Forecasting Center*, p. 461-471.

Glastra, R.,Wakker, E. & Richert, W. (2002). Oil Palm Plantation & Deforestation in Indonesia.

Gu, M. & Aamodt, A. (2005). A knowledge-intensive method for conversational CBR. *Proceedings of the Fourth International Conference on Case-Based Reasoning* p.296-311. Chicago, IL: Springer.

Harmon, P. (1992). Case-based reasoning III. *Intelligent Software Strategies,* 8(i).

Jensen, F. V. (1996) An Introduction to Bayesian Networks. NewYork: Springer.

Kim, J., & Miller, N. L. (1996). Simulating Winds and Floods: Regional weather-river prediction and regional climate research. *IEEE Potentials*. 15(4). p. 17-20.

Kourtz, P. (1987). Expert System Dispatch of Forest Fire Control Resources, *AI Applications*, vol.19, p. 1-8.

Li, W. Y. and Chandrasekar, V. (2002). Rainfall Estimation from Vertical Profiles of Reflectivity Using Neural Networks, *Colorado State University Fort Collins,* CO 80523-1373.

Lorenz, E. N. and Bull, A. (1969). Meteorol. Soc., p. 50, 345–349.

Malaysian Palm Oil Association. (2005, Jan 6). RSPO Public Forum on Sustainable Palm Oil, Malaysian Palm Oil Association (MPOA). Retrieved March, 2009, from: http://www.cs.nott.ac.uk/~rxq/files/PhDThesis.pdf

Mcqueen, R. J.(2000) Applying Machine Learning to Agricultural Data. University of Waikato, Hamilton, New Zealand.

Montaner, M., L´opez, B., and Rosa, J. L. D. L. (2003). A taxonomy of recommender agents on the internet. *Articial Intelligence Review,* 19(4): p. 285.330.

Nielsen, H. A., Madsen, H. & Nielsen, T. S. (2006). Using quartile regression to extend an existing wind power forecasting system with probabilistic forecasts. Wind Energy. p. 95–108.

Nunamaker, Jr., Jay F., Minder Chen , Titus D. M. Purdin, (1990). Systems development in information systems research, Journal of Management Information Systems, v.7 n.3, p.89-106

Perrin, O., Rootz´en, H. & Taesler, R. (2006). A discussion of statistical methods used to estimate extreme wind speeds. Theoretical and Applied Climatology. p. 85, 203–215.

Polumetla, A. (2006). Machine Learning Methods for The Detection Of RWIS Sensor Malfunctions. University of Minnesota.

Porter, B.W. & Bareiss, E.R. (1986). PROTOS: An experiment in knowledge acquisition for heuristic classification tasks. *In Proceedings of the First International Meeting on Advnces in Learning (IMAL)*, Les Arcs, France, p. 159-74.

Qu, R. (2002). Case-Based Reasoning for Course Timetabling Problems. Retrieved October, 2006, from: http://www.cs.nott.ac.uk/~rxq/files/PhDThesis.pdf

Rafiei, D. (1999) Fourier-Transform Based Techniques in Efficient Retrieval of Similar Time Sequences, *Ph.D. thesis, Department of Computer Science, University of Toronto*, Ontario, Canada.

Raftery, A. E., Gneiting, T., Balabdaoui, F. & Polakowski, M. (2005). Using Bayesian model averaging to calibrate forecast ensembles : Monthly Weather Review p. 133, 1155–1174.

Raible, C. C., Bischof, G., Fraedrich, K. and Kirk. E. (1999). Statistical Single-Station Short-Term Forecasting of Temperature and Probability of Precipitation: Area Interpolation and NWP Combination. *Proceedings of American Meteorological Society*, vol 15, p. 203-214.

Ramli, A (2003) Short-Term and Long-Term Projection of Malaysian Palm oil Production. Malaysian Palm Oil Board, P. O. Box 10620, 50720 Kuala Lumpur, Malaysia.

Ridge, G. (2007) Strategic Business Development, Aerospace Sciences and Engineering Division.

Riordan, A. and Sorensen, H. (1995). An intelligent agent for highprecision information filtering. In Proceedings of the CIKM-95 Conference.

Rougegrez, S. (1993) Similarity evaluation between observed behaviors and the prediction of processes. *Proceedings First European Workshop on Case-Based Reasoning*, 1993, Springer-Verlag, Berlin, p.155–166.

Roulston, M. S., D. T. Kaplan, J. Hardenberg, and L. A. Smith (2003). Using medium-range weather forcasts to improve the value of wind energy production. Renewable Energy. 28, p. 585–602.

Siew, W.L. (2000) Analysis of palm and palm kernel oils. Advances In Oil Palm Research. Malaysian Palm Oil Board, Kuala Lumpur, Malaysia. p. 968-1035.

Singhrattna, N., Rajagopalan, B., Clark, M. and Kumar, K. K. (2005). Seasonal Forecasting of Thailand Summer Monsoon Rainfall. *International Journal of Climatology,* November, 2005.

Sloughter, J. M., Gneiting, T., & Raftery A. E. (2008), Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging ,Department of Statistics, University of Washington.

Sundram, K. (2002) Palm Oil: Chemistry and Nutrition Updates. Malaysian Palm Oil Board (MPOB), Kuala Lumpur, Malaysia.

Sycara, K., & Navichandra, D (1992). Retrieval strategies in a case-based design system. In, *Artificial Intelligence in engineering design, Vol.2*, C.Tong and D.Sriram ed. Academic Press.

Toomarian, N.B., Barhen, J., & Aminzadeh, F. (2000). Oil Reservoir Properties Estimation Using Neural Networks.

Vaishnavi, V., & Kuechler, B. (2006). *Design Research in Information Systems*. Retrieved October, 2007, from http://www.isworld.org/Researchdesign/drisISworld.htm.

Watson, I.D., & Abdullah, S. (1994). Developing Case-Based Reasoning Systems: A Case Study in Diagnosing Building Defects. In, Proc. *IEE Colloquium on Case-Based Reasoning: Prospects for Applications*, Digest No: 1994/057, p. 1/1-1/3.

Wilks, D.S. (1999) Multisite generalization of a daily stochastic precipitation generation model. *Journal of Hydrology* 210. p. 178-191.

Xiao, R. R. and Chandrasekar, V., (1997). Development of a Neural Network Based Algorithm for Rainfall Estimation from Radar Observations, *IEEE Transactions on geoscience and remote sensing*, vol.35, no.1, p. 160-170.

Xu, G. and Chandrasekar, V. (2005). Operational Feasibility of Neural-Network-Based Radar Rainfall Estimation, *IEEE Transactions on geoscience and remote sensing*, vol.2, no.1, p. 13-17.