## Application of logistic regression and GIS in identifying shrimp farming provinces in Thailand

THAN PE † *University of Western Sydney, AU*
PAUL THURSTAN SMITH *University of Western Sydney, AU*

The Network of Aquaculture Centres in Asia-Pacific (NACA) carried out an Aquaculture Sustainability and Environment Survey of 451 shrimp farms in Thailand in 1995. Two hundred and seventy-six variables were measured in the survey and our interest was to find the variables that could identify ten provinces practising prawn farming. The method of logistic regression analysis was applied and it was found that the variables which play an important role in identifying the provinces were dominant soil type, number of farming years, source of farm water, depth of production ponds, number of aquaculture farms within three kilometers of the farm, number of aquaculture farms sharing the water supply, number of farms discharging effluent into the water supply canal, environmental impact assessment, site selection to avoid impacts of other users, design of separate water supply/drainage system, retention of mangrove buffer zone, effluent treatment pond, existence of water or sediment related problem, and other problems. It was found that the logistic regression model could identify a minimum of 89.58% of the farms in Surat Thani province and a maximum of 99.61% of the farms in the provinces of Chanthaburi and Samut Sakhon. The results, using different types of variables, are shown on a map of Thailand using the MAPINFO package of Geographical Information System (GIS).

† POSTAL: *University of Western Sydney (Macarthur), P.O. Box 555,*
*Campbelltown NSW 2560, AUSTRALIA.*
FAX: +61 2 4620 3025
EMAIL: t.pe@uws.edu.au

## Kullback-Leibler distance, probability and football prediction

DAVID L. DOWE † *Monash University, AU*
MATTHEW DORAN *Oakton Computing, AU*
GRAHAM E. FARR *Monash University, AU*
A. JOHN HURST *Monash University, AU*
DAVID POWELL *Monash University, AU*
TORSTEN SEEMANN *Monash University, AU*

Gamblers in particular and others generally are not as interested in a 'right'/'wrong' prediction as they are in a probabilistic prediction of an outcome. Are the bookies' pay-out odds being offered on an event reasonable given our perceived probability (e.g., 52% or 97%), or what are the risks of some medical procedure? A logarithmic scoring function for probabilistic predictions used by Wallace and Patrick (1993, Machine Learning 11, 7-22) and Dowe and Krusel dates back to I.J. Good in the 1950s, is equivalent to fully-invested gambling (Cover and Thomas, 1991, Elements of Information Theory, Wiley) and can also be thought of as a Kullback-Leibler distance minimisation problem. A probabilistic football tipping competition using these ideas was initiated at Monash University in 1995 and a Gaussian competition based on Kullback-Leibler distance was also developed by the first author in 1996 (Dowe, Farr, Hurst and Lentin, 1996, Mathematics and Computers in Sport). We describe the performance of a tipping method developed in 1997 and used in The Australian newspaper in 1998 which combines an Elo system used for rating teams modified to include home-ground advantage with bookies' odds, and the general benefits of Bayesian integration and other ways of combining predictors.

† POSTAL: *School of Computer Science and Software Engineering, Monash University,*
*Clayton VIC 3168, AUSTRALIA.*
FAX: +61 3 9905 5146
EMAIL: dld@cs.monash.edu.au
WWW: http://www.csse.monash.edu.au/~dld/