Quantifying Uncertanties Associated With Climate Variaiblity and Climate Change Studies: the Cox´S Semiparametric Approach

Aline de Holanda Nunes Maia

e-mail:ahmaia@cnpma.embrapa.br
Embrapa Meio Ambiente
Holger Meinke
e-mail:holger.meinke@wur.nl
Wageningen University, Department of Plant Sciences

A wide range of probabilistic approaches have been used for assessing climatic risks and associated uncertainties: multiple linear regression (Martis et al, 2002), logistic regression (Nichols, 1984; Lo et al, 2007), nonparametric approaches based on empirical cumulative distribution functions (Maia & Meinke, 1998; Maia et al, 2004), among others.

In this paper, we quantify risks by using probability of exeedance curves (PECs) and propose the use of the proportional hazards model (also referred to as Cox model - Cox, 1972; CPH model) to investigate the influence of continuous (e.g. oceanic/atmospheric indexes) or categorical predictors (e.g. classes derived from El Niño/Southern Oscillation) on these risks.

CPH model is widely used in medical analyses of survival data that examine the effect of explanatory variables on survival times. This allows the ranking of risk factors and quantitative assessments of their impacts. These methods are well established in medical research and used routinely for risk assessment in medical studies (e.g. Finkelstein et al, 1993; De Lorgeril, 1999; Smith et al, 2001; Gibbons, 2002) where data are rarely normally distributed and often incomplete. To our knowledge, this study is the first using the CPH framework to analyse the linkage among oceanic/atmospheric indexes and climate risks. This constitutes a major step towards better quantification of climate related uncertainties in climate variability and climate change studies.

Some of the advantages of using the CPH model include:

a) the method was developed to accommodate censored data (incomplete information) without the need to balance the data set thereby discarding potentially valuable information;b) in contrast to ordinary least squares multiple linear regression or logistic regression, CPH model does not require assumptions regarding the type of underlying probability distributions; proportionality of hazards is the only assumption necessary for CPH model, and even this can be relaxed via an appropriate generalization;c) it overcomes the problem of having to estimate probabilities of exceedance for each threshold in order to compose a PEC, a limitation of the logistic regression approach as used in Lo et al (2007);d) when compared with the nonparametric approach used by Maia et al (2007), CPH framework is superior because it allows investigating simultaneously influences of many predictors on risks; the contribution of each 'candidate' can objectively be evaluated via likelihood tests;e) methods for assessing uncertainties of PEC (and therefore uncertainties of risk estimates) are readily available. These methods have well established theoretical basis (Kalbfleisch and Prentice, 1980);

We demonstrate the adequacy and usefulness of the proposed approach by analysing the influence of two oceanic/atmospheric indexes on the onset of monsoonal wet season at Darwin, Australia, as suggested by Lo et al (2007): the Southem Oscillation Index (mean of the July and August monthly SOI values) and the first rotated principal component (SST1) of large scale Sea Surface Temperatures anomalies (Drosdowsky and Chambers, 2001).

When applied to grided data, the CPH approach allows objective spatial assessment of either individual or joint influences of such predictors on risks. Mapping the coefficients of the Cox model (which express the magnitude of the predictor's influence) and p-values resulting from likelihood tests provides a complete descriptive and inferential assessment of predictor influence on the risks under investigation. Once the strongest predictors are selected at a location, probabilities of exceeding any particular threshold (preferably within the range of observed data) can be estimated for any combination of predictor values. The resulting risk estimates and their respective uncertainties provide valuable information for decision making in climate sensitive industries.

References

Drosdowsky, W. and L. E. Chambers, 2001: Near-global sea surface temperature anomalies as predictors of Australian seasonal rainfall; Joumal of Climate, 14, 1677-1687.

Cox, D.R, 1972: "Regression Models and Life-Tables (with Discussion)," Joumal of the Royal Statistical Society, Series B, 34, 187–220.

De Lorgeril M.; Salen P.; Martin J-L.; Monjaud I.; Delaye J.; N. Mamelle, 1999: Mediterranean diet, traditional risk factors, and the rate of cardiovascular complications after myocardial infarction. Final report of the Lyon Diet Heart Study. Circulation, 99, 779-85.

Finkelstein, D. M.; Moore, D. F.; D. A. Schoenfeld, 1993: A Proportional Hazards Model for Truncated AIDS Data. Biometrics, 49 (3), 731-740.

Gibbons, L.E.; Teri L.; Logsdon R.; McCurry S.M.; Kukull W.2; Bowen J.; McCormick W.; E. Larson, 2002: Anxiety Symptoms as Predictors of Nursing Home Placement in Patients with Alzheimer's Disease. Journal of Clinical Geropsychology, 8 (4), 335-342.

Kalbfleisch, J.D. and Prentice,R.L.,1980:The Statistical Analysis of Failure Time Data, New York: John Wiley & Sons, Inc.

Lo F., M. Wheeler, H. Meinke and A. Donald,2007: Probabilistic forecasts of the onset of the north Australian wet season. Monthly Weather Review (in press).

Maia, A. de H.N. and H. Meinke, 1999: Non-parametrical survival analysis as a statistical tool in agricultural decision making. International Symposium, Modeling Cropping Systems, Proceedings European Soc. For Agronomy, Div. Agroclimatology and Agronomic Modeling, Lleida, Spain, 21-23 June 1999, 103-104.

Maia, A. H. N.; Meinke, H.; S. Lennox, 2004: Assessment of probabilistic forecast 'skill' using p-values, Proc. 4th International Crop Science Congress, Brisbane, Australia, available at: http://www.cropscience.org.au/icsc2004/poster/2/6/1360 maiaa.htm.

Maia, A. H. N.; H. Meinke, S. Lennox; R. C. Stone, 2007: Inferential, non-parametric statistics to assess quality of probabilistic forecast systems. Monthly Weather Review, 135(2), 351-362.

Martis, A.; G. J. van Oldenborgh; G. Burgers, 2002: Predicting rainfall in the Dutch Caribbean – more than El Niño? Int. J. Climatology, 22, 1219-1234.

Nicholls, N., 1984: A system for predicting the onset of the north Australian wet-season, Journal of Climatology, 4, 425-435.

Nicholls, N.; McBride, J. L.; R.J. Ormerod, 1982: On predicting the onset of the Australian wet-season at Darwin, Monthly Weather Review, 110, 14-17.

Smith, G. E.; O'Brien, P. C.; Ivnik, R. J.; Kokmen, E. G. Tangalos, 2001: Prospective analysis of risk factors for nursing home placement of dementia patients. Neurology, 57:1467-1473.