Geophysical Research Abstracts, Vol. 9, 11543, 2007
SRef-ID: 1607-7962/gra/EGU2007-A-11543

Reconciling Hydrological Physically Based Models and Data Driven Models in Terms of Predictive Probability.

E. Todini (1), G. Coccia (1) and C. Mazzeti (2)

(1) Dipartimento di Scienze della Terra e Geologico Ambientali – Università di Bologna. Bologna, Italy, (2) Protezione e Gestione Ambientale Srl (ProGeA) – Bologna, Italy

For many years, hydrologists have debated the appropriateness of using data driven models as opposed to physically based models for flood forecasting and, in particular, for real time flood forecasting. More recently, several tentative have been made for combining the forecasts given by different types of models, by means of Bayesian weighting.

In this paper a new approach is presented that combines the different models in terms of the conditional predictive probability.

Following the work of Krzysztofowicz (1999), by taking advantage of the Normal Quantile Transform, one can transform the observations, namely the predictand y, as well as each of the n model predictions $\hat{y}_i \quad i = 1, ...n$ into a Normal space where their marginal distribution is a Standard Normal distribution. In this space it is possible to build a Multi-variate Normal distribution $f\left(y, \hat{y}_1, \hat{y}_2, ..., \hat{y}_n\right)$ only by estimating the correlation coefficients ρ among the different vectors, namely $\rho_{y,\hat{y}_1}, \ \rho_{y,\hat{y}_2}, \ \rho_{y,\hat{y}_3}, \ldots\ldots., \ \rho_{\hat{y}_1,\hat{y}_2}, \ \rho_{\hat{y}_1,\hat{y}_3}, \ldots\ldots., \rho_{\hat{y}_2,\hat{y}_3}, \ldots\ldots..$

This approach differs from the Bayesian processor proposed by Krzysztofowicz (1999), which implies the independence of the data driven model forecast (in that case a lag-one auto-regressive model) from the physically based model forecast. As a matter of fact, although each model forecast are issued independently, they cannot be statistically independent, given that all of them aim at representing the same quantity, the predictand, with which they are supposed to be correlated.

Given the joint probability density, it is then immediate to obtain the predictive probability as the probability of observing the predictand given the different model outputs,

namely $f\left(y\left|\hat{y}_1, \hat{y}_2, ..., \hat{y}_n\right.\right)$.

Results of this approach are shown for two cases. The first one is a flood forecast problem on the Po river in Italy, based on the combination of (i) an auto-regressive model and a physically based flood routing model, (ii) a Nearest Neighbour and a physically based flood routing model. The second one is the flood forecast in the Parma river based on the combination of a physically based hydrological model (TOPKAPI) and a data driven model (an Artificial Neural Network based model). In particular, in the latter case, the results show the advantage for using the combination approach, which enables to improve the robustness of the forecasts on the one hand by improving the physical model forecasts by means of the ANN ones and on the other hands by reducing the inherent instabilities of the ANN approach on the basis of the hydrological model physical properties.