Geophysical Research Abstracts,
Vol. 11, EGU2009-4027-1, 2009
EGU General Assembly 2009

Use of normalized radial basis function in hydrology

T. Čotar (1), I. Grabec (2), and M. Brilly (3)

(1) Faculty of Civil Engineering and Geodesy, Jamova 2, 1000 Ljubljana, Slovenia (prevodnik.toni@gmail.com), (2) Faculty for Mechanical Engineering, Askerceva 6, 1000 Ljubljana, Slovenia (igor.grabec@fs.uni-lj.si), (3) Faculty of Civil Engineering and Geodesy, Jamova 2, 1000 Ljubljana, Slovenia (mbrilly@fgg.uni-lj.si)

In hydrology we work a lot with time series of river runoff, temperature and precipitation. Those time series are nonlinear and stohastic. They are ussualy contaminated with noise. Despite their complexity we built models from them. Models can be deterministic or stohastic. We use those models for prediction of future values, for interpolation of missing values and for extrapolation to scenarios that we do not have similar data for.

In our case we used a normalized radial basis function for interpolation of missing Reka river mean monthly runoff. We have mean monthly temperature and monthly precipitation time series from 1851 to 2006 from station in Trieste, Italy. We have Reka river runoff time series from 1952 to 2006 form station Cerkvenikov mlin in Ilirska Bistrica, Slovenia. Those two stations are roughly 40 kilometers apart. From those data alone we tried to estimate Reka river runoff from 1851 to 1951. Data from 1952 to 1990 were used for learning of model, data from 1991 to 2006 were used for verification of model.

For radial basis function w we chose multidimensional normal distribution, where Σ is covariance matrix:

$$w\left(x,\Sigma\right)=\frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}}\exp\left[-\tfrac{1}{2}\left(x-\overline{x}\right)^{T}\Sigma^{-1}\left(x-\overline{x}\right)\right] \quad (1)$$

The prediction or in other words interpolating process is given by:

$$\hat{y}_i\left(x_i\right)=\sum_{n=1}^{N}y_n C_n\left(x_i\right) \quad (2)$$

Weighting coefficient C, which is also a function, is:

$$C_n\left(\mathrm{x_i}\right)=\frac{\mathrm{w}\left(\mathrm{x_i}-x_n,\sigma\right)}{\sum_{\mathrm{k=1}}^{\mathrm{N}}\mathrm{w}\left(\mathrm{x_i}-x_k,\sigma\right)}=\left\{\sum_{\mathrm{k=1}}^{\mathrm{N}}\exp\left[\frac{\|x_i-x_n\|^2-\|x_i-x_k\|^2}{2\sigma^2}\right]\right\}^{-1} \quad (3)$$

The only parameter in this model is standard deviation σ. We find it by optimizing objective function, which is called a coefficient of prediction quality:

All time series were first normalized by their own standard deviations. In that way we make diferrent time series similar to eachother, which simplifies the model. We could even try to optimize the model itself, but we use a simple one. In our case, if q_n means Reka river mean runoff at month n, P_n monthly precipitation in month n and T_n mean temperature in month n and analogly for other indexes, the model reads:

In more complex case C_n would be a function of example also P_{i-1} and T_{i-1} or something like that.

$$Q = 1 - \frac{E\left[(y-\hat{y})^2\right]}{Var(y)+Var(\hat{y})} \quad (4)$$

$$\hat{q}_i\left(P_i, T_i\right) = \sum_{n=1}^{N} q_n C_n\left(P_i, T_i\right) \quad (5)$$

We wrote a model in Matlab. The result for coefficient of prediction quality in verification process was 0,85. That is very good result if we consider bad data correlation, simplicity of the model and simplicity of the method itself. We compared that result with results from using machine learning methods. The results were comparable. But the difference is that, using normalized radial basis function is much more physicaly intuitive and it is based on some physical concepts of entropy and biological concepts of learning. But the problem, that we encountered, lied in the interpolating nature of method used. So, it worked very well in middle ranges of values, but slightly worse at minimal and maximal values of our data.

The method used is very simple and intuitive. It is very simple to use. It is also useful for removing noise from data and for continuous control. For the control purpose it is used so that every new data gets in the model and model adapts itself by slightly changing weighting coefficients. This function of control and noise removal is going to be tested in the future.

LITERATURE

Grabec, I., Sachse, W. 2006. Synergetics of Measurement, Prediction and Control. Springer Verlag.

Grabec, I. 2005. Extraction of physical laws from joint experimental data. European Physical Journal.

Kantz, H., Schreiber, T. 1997. Nonlinear time series analysis. Cambridge University Press.