A logistic regression based approach for the prediction of flood warning threshold exceedance

Tommaso Diomede, Luca Trotter, Maria Stefania Tesini, and Chiara Marsigli
ARPAE-SIMC, HydroMeteorological and Climate Service of the Emilia-Romagna Regional Agency for Environmental
Protection and Energy, Bologna, Italy (tdiomede@arpa.emr.it)

A method based on logistic regression is proposed for the prediction of river level threshold exceedance at short (+0-18h) and medium (+18-42h) lead times. The aim of the study is to provide a valuable tool for the issue of warnings by the authority responsible of public safety in case of flood.

The role of different precipitation periods as predictors for the exceedance of a fixed river level has been investigated, in order to derive significant information for flood forecasting. Based on catchment-averaged values, a separation of "antecedent" and "peak-triggering" rainfall amounts as independent variables is attempted. In particular, the following flood-related precipitation periods have been considered: (i) the period from 1 to n days before the forecast issue time, which may be relevant for the soil saturation, (ii) the last 24 hours, which may be relevant for the current water level in the river, and (iii) the period from 0 to x hours in advance with respect to the forecast issue time, when the flood-triggering precipitation generally occurs.

Several combinations and values of these predictors have been tested to optimise the method implementation. In particular, the period for the precursor antecedent precipitation ranges between 5 and 45 days; the state of the river can be represented by the last 24-h precipitation or, as alternative, by the current river level. The flood-triggering precipitation has been cumulated over the next 18 hours (for the short lead time) and 36-42 hours (for the medium lead time).

The proposed approach requires a specific implementation of logistic regression for each river section and warning threshold. The method performance has been evaluated over the Santerno river catchment (about 450 km2) in the Emilia-Romagna Region, northern Italy. A statistical analysis in terms of false alarms, misses and related scores was carried out by using a 8-year long database. The results are quite satisfactory, with slightly better performances for the higher flood warning threshold.

The optimisation of the method has been performed in a "hindcast" mode, that is observed rainfall represents the flood-triggering precipitation. Using rainfall forecasts as predictor of the flood-triggering precipitation causes a degradation of performance; nevertheless, this performance is similar to that provided by a distributed rainfall-runoff model driven by the same rainfall forecast input.

This computationally cheap technique to estimate flood warning exceedance can be applied to each gauged river section, independently of the availability of a rating curve for that section.