

AN ALTERED SEMI-PARAMETRIC REGRESSION ENSEMBLE IN SHORT-TERM FLOOD FORECASTING

Le Hoang Tuan¹, To Anh Dung²

¹Department of Mathematics and Physics, University of Information Technology, VNU-HCM

²Faculty of Mathematics and Computer Science, University of Science, VNU-HCM
tuanlh@uit.edu.vn, tadung@hcmus.edu.vn

Abstract

In this paper, the potential and effect of both parametric and non-parametric estimations in a semi-parametric regression model are explored to modelize flood water-level and to forecast the inundation of Mekong Delta, Vietnam. The semi-parametric regression technique is a combination of a parametric regression approach and a non-parametric regression concept. In the process of model building, several amended linear regression models are applied for the parametric component. They are stepwise multiple linear regression, partial least squares solution and multi-recursive regression method. They are used to capture flood's linear characteristics. The non-parametric estimations are solved by two justified algorithms of a smooth function. Furthermore, some modified non-linear regression models based on artificial neural network are also able to obtain flood's non-linear characteristics. Besides, the model's error and convergence are investigated. Then the principle component analysis technique is utilized to set up the final model. Obtained empirical results indicate that the prediction by using the modified semi-parametric regression ensemble model is generally better than those obtained by using the other models presented in this study in terms of the same evaluation measurements. The proposed ensemble model can be used as a reliable and promising alternative flood forecasting tool to upgrade/optimize prediction quality and accuracy further.

This research is funded by Vietnam National University Ho Chi Minh City (VNU-HCM) under grant number C2018-26-05.

Key words: parametric, non-parametric, semi-parametric regression ensemble model, stepwise multiple linear, partial least squares, multi-recursive regression, artificial neural network, estimation, smooth, kernel function, splines, flood, water-level, prediction, forecasting.

MÔ HÌNH TÍCH HỢP HỒI QUY BÁN THAM SỐ CÓ HIỆU CHỈNH TRONG DỰ BÁO LŨ LỤT NGẮN HẠN

Lê Hoàng Tuấn¹, Tô Anh Dũng²

¹Bộ môn Toán – Lý, Trường Đại học Công nghệ Thông tin, ĐHQG-HCM

²Khoa Toán – Tin học, Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM

tuanlh@uit.edu.vn, tadung@hcmus.edu.vn

Tóm tắt

Trong bài báo này, tiềm năng và hiệu quả của phép ước lượng tham số lẫn phi tham số, trong một mô hình hồi quy bán tham số, đều được nghiên cứu nhằm mô hình hóa bề mặt mực nước và dự báo lũ lụt ở vùng Đồng bằng Sông Cửu Long, Việt Nam. Kỹ thuật hồi quy bán tham số là một sự kết hợp giữa cách tiếp cận hồi quy tham số và quan điểm hồi quy phi tham số. Trong quá trình xây dựng mô hình, một vài mô hình hồi quy tuyến tính có cải biên đã được vận dụng cho thành phần tham số. Chúng là hồi quy tuyến tính bội bậc thang, giải pháp bình phương bé nhất riêng phần, và phương pháp hồi quy đa đệ quy. Chúng được dùng để lưu giữ tính chất tuyến tính của lũ lụt. Các phép ước lượng phi tham số thì được giải quyết bằng hai thuật toán có hiệu chỉnh của hàm làm trơn. Ngoài ra, những mô hình hồi quy phi tuyến dựa trên nền tảng mạng nơ ron (có cải tiến) cũng được vận dụng, nhằm giúp cho ta nắm bắt thông tin phi tuyến của lũ lụt. Bên cạnh đó, tính hội tụ và sai số của mô hình cũng được xem xét. Tiếp theo, phương pháp phân tích thành phần chính được áp dụng để cho ra mô hình sau cùng. Những kết quả thực nghiệm đạt được cho thấy rằng việc dự đoán bằng cách dùng mô hình tích hợp hồi quy bán tham số có hiệu chỉnh nhìn chung là tốt hơn nhiều so với những gì đạt được bằng các mô hình còn lại trong phạm vi bài báo này, xét trên cùng đơn vị đo lường và đánh giá. Mô hình tích hợp mà chúng tôi đề xuất có thể được dùng như là một công cụ dự báo lũ lụt thay thế đầy triển vọng và đáng tin cậy, nhất là cho việc cải tiến/tối ưu hóa chất lượng và độ chính xác dự báo.

Nghiên cứu được tài trợ bởi Đại học Quốc gia Thành phố Hồ Chí Minh (ĐHQG-HCM) trong khuôn khổ Đề tài mã số C2018-26-05.

Từ khóa: tham số, phi tham số, mô hình tích hợp hồi quy bán tham số, quy tuyến tính bội bậc thang, bình phương bé nhất riêng phần, hồi quy đa đệ quy, mạng nơ ron nhân tạo, phép ước lượng, tron, hàm hạt nhân, s-plines, lũ lụt, bề mặt mực nước, dự đoán, dự báo.