



Predictive data-driven models of the evapotranspiration process

I. El-Baroudy (1), A. Elshorbagy (1), S. K. Carey (2), O. Giustolisi (3), D. Savic (4)

(1) Centre for Advanced Numerical Simulation (CANSIM), Department of Civil & Geological Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9 Canada
(IBE140@mail.usask.ca / Fax: (306) 966-5427 / Phone: (306) 966-5348)

(2) Department of Geography & Environmental Studies, Carleton University, Ottawa, ON K1S 5B6 Canada

(3) Technical University of Bari, Engineering Faculty, Department of Civil & Environmental Engineering, Taranto, via Turismo, 8, 74100, Italy

(4) School of Engineering, Computer Science and Mathematics, University of Exeter, Exeter, Devon, EX4 4QF United Kingdom

Evapotranspiration accounts for more than two thirds of the precipitation losses at the global scale while the remaining third contributes to both soil storage and stream flow. Precipitation and stream flow records exhibit high variation relative to evapotranspiration; i.e. evapotranspiration requirements are achieved from annual precipitation before stream flows. Therefore, reliable estimates of actual evapotranspiration are crucial for effective watershed modeling and water resource management, yet direct measurements of the evapotranspiration losses are difficult and expensive.

This paper investigates the utility of the data-driven techniques in modeling actual evapotranspiration measured by eddy covariance system. The authors use the Evolutionary Polynomial Regression in comparison with Artificial Neural Networks and Genetic Programming, used in a previous research work. In this research, evapotranspiration is modeled using the day-time hourly data (8 am – 8pm) of the environmental variables; Net Radiation (NR), Ground Temperature (GT), Air Temperature (AT), Wind Speed (WS) and Relative Humidity (RH). This paper uses two case studies in

the Mildred Lake mine, the South Bison Hill (SBH) and the South West Sand Storage (SWSS), which are located to the north of Fort McMurray, Alberta, Canada. The performance of the modeling tools, ANN, GP, and EPR, is evaluated based on three error measures; (i) Root Mean Squared Error (RMSE), (ii) Mean absolute relative Error (MARE), and (iii) Correlation Coefficient (R). Furthermore, this research extends the utility of these techniques by exploring the effect of previous states (time lags) of the meteorological input variables on characterizing the actual evapotranspiration. The hysteresis effects due to the variations in the diurnal behaviour, morning and afternoon hours, of the different environmental variables, are also investigated in this study.

The models developed using the Evolutionary Polynomial Regression, based on the two case studies at the Mildred Lake mine, AB, Canada, provided comparable performance to the models of Genetic Programming and Artificial Neural Networks. Moreover, the Evolutionary Polynomial Regression provided simpler models than those developed by the other data-driven techniques, particularly in one of the case studies. The inclusion of the previous states of the input variables and modeling separate time periods, morning and afternoon, enhanced the performance of the developed model, which in turn confirms the dynamic nature of the evapotranspiration process.