Journal of Hydrology 531 (2015) 1003–1019

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Multivariate time series modeling of short-term system scale irrigation demand

Kushan C. Perera [a,*], Andrew W. Western [a,*], Biju George [b], Bandara Nawarathna [c]

[a] Department of Infrastructure Engineering, The University of Melbourne, Australia
[b] International Centre for Agricultural Research in the Dry Areas (ICARDA), Cairo, Egypt
[c] Hazards, Warnings and Forecasts Division, The Bureau of Meteorology, Australia

ARTICLE INFO

Article history:
Received 24 July 2015
Received in revised form 14 October 2015
Accepted 4 November 2015
Available online 10 November 2015
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Sheng Yue, Associate Editor

Keywords:
Irrigation demand
Time series
Multivariate
ARMAX
Exogenous variables
Forecasting

SUMMARY

Travel time limits the ability of irrigation system operators to react to short-term irrigation demand fluctuations that result from variations in weather, including very hot periods and rainfall events, as well as the various other pressures and opportunities that farmers face. Short-term system-wide irrigation demand forecasts can assist in system operation. Here we developed a multivariate time series (ARMAX) model to forecast irrigation demands with respect to aggregated service points flows ($ID_{CGi,ASP}$) and off take regulator flows ($ID_{CGi,OTR}$) based across 5 command areas, which included area covered under four irrigation channels and the study area. These command area specific ARMAX models forecast 1–5 days ahead daily $ID_{CGi,ASP}$ and $ID_{CGi,OTR}$ using the real time flow data recorded at the service points and the uppermost regulators and observed meteorological data collected from automatic weather stations. The model efficiency and the predictive performance were quantified using the root mean squared error (RMSE), Nash–Sutcliffe model efficiency coefficient (NSE), anomaly correlation coefficient (ACC) and mean square skill score (MSSS). During the evaluation period, NSE for $ID_{CGi,ASP}$ and $ID_{CGi,OTR}$ across 5 command areas were ranged 0.98–0.78. These models were capable of generating skillful forecasts (MSSS $\geqslant$ 0.5 and ACC $\geqslant$ 0.6) of $ID_{CGi,ASP}$ and $ID_{CGi,OTR}$ for all 5 lead days and $ID_{CGi,ASP}$ and $ID_{CGi,OTR}$ forecasts were better than using the long term monthly mean irrigation demand. Overall these predictive performance from the ARMAX time series models were higher than almost all the previous studies we are aware. Further, $ID_{CGi,ASP}$ and $ID_{CGi,OTR}$ forecasts have improved the operators' ability to react for near future irrigation demand fluctuations as the developed ARMAX time series models were self-adaptive to reflect the short-term changes in the irrigation demand with respect to various pressures and opportunities that farmers' face, such as changing water policy, continued development of water markets, drought and changing technology.

1. Introduction

Travel time between the source (reservoir) and irrigated fields leads to significant challenges for irrigation system operators in terms of reacting to short-term irrigation demand fluctuations arising from variations in weather including very hot periods and rainfall events, as well as the various other pressures and opportunities that farmers face. Modern irrigation system automation technologies have recently been developed with the aim of improving delivery efficiency and level of service, an example of which has been implemented in northern Victoria, Australia (GMWater, 2010). These systems can deliver water with very short lead times due to their fully automated distribution network (NVIRP, 2010a). They also deliver a step change in the amount of real time data available through monitoring and telemetry of all regulators within the system. A remaining challenge is that while the system automates water control in the distribution canals, bulk delivery of water to the command area still needs to be planned as it is subject to substantial travel time delays. Short-term irrigation demand forecasting has the potential to assist in system operation but has often been constrained by limited information on both current demand levels and likely future weather. This research aims to capitalize on the opportunity presented by the large amount of consistent real-time irrigation system data now available and expanding numerical weather predictions to forecast irrigation demand. In particular, this paper develops and evaluates a time series model to forecast short-term irrigation demand based on recent irrigation demand and weather data.

* Corresponding authors.

http://dx.doi.org/10.1016/j.jhydrol.2015.11.007
0022-1694/

The dynamic and nonlinear nature of short-term irrigation demand results from complex interactions between biophysical (crop–soil–climate interaction), behavioral (farmer and system operator attitudes that influence management decisions) and supply (supply source, seasonal allocation, permanent entitlement) factors (Zaman et al., 2007), and is very challenging to forecast. Various irrigation demand forecast models have been developed to predict irrigation water demand aimed at meeting a variety of farmer and/or operator objectives including: irrigation scheduling (George et al., 2003; George et al., 2000; Rao et al., 1988); yield/profit optimization (Prasad et al., 2006; Smout, 2005; Umamahesh, 2002); and water allocation (Paul, 2000; Rao, 1990). Practical application of these models at system scale is often limited by the lack of required data or the expense of data acquisition (Ticlavilca et al., 2011). Depending on the degree of data availability, two distinct modeling architectures have been used to forecast irrigation demand; process-based (conceptual) and data-driven (statistical) (Alfonso et al., 2011; Pulido-Calvo and Gutierrez-Estrada, 2009). Process-based models use the physical concepts associated with the irrigation demand while data-driven models are trained to map the relationship between influential factors and irrigation demand with no detailed considerations about the internal structure of the physical processes (Alfonso et al., 2011). In an operational context with good real-time dynamic system response data available, data-driven models are often preferred as these measurements encapsulate a wide range of information about the many factors influencing the current system behavior.

Most process-based models are based on the soil–water balance equation including biophysical demand and supply factors. These models forecast the irrigation demand by incorporating historical, current (real-time) or future (often from short-term weather forecasts) information on those factors at various spatial scales from paddock (Ejieji and Gowing, 2000) to system (Cai et al., 2011; Wilks and Wolfe, 1998), as well as from lead times of 1–2 days (Cai et al., 2011; Wilks and Wolfe, 1998) to weeks (Ejieji and Gowing, 2000; Wang and Cai, 2009). The reliability of these forecasts is mainly constrained by the prediction uncertainties of the influential factors and various model structural and parameter errors (Ejieji and Gowing, 2000; Wilks and Wolfe, 1998). The prediction uncertainties for inputs increase with the lead time, especially precipitation (Azhar and Perera, 2011). Furthermore, process-based models are also usually structured to capture correlations between short-term weather forecasts and farmer behavior (Austen et al., 2002; Bergez and Garcia, 2010; Ingram et al., 2002). These behavioral factors cannot be directly measured, unlike biophysical or supply factors. This, combined with a lack of consistent flow data often makes quantification of farmer behavior difficult. As a consequence of the above issues, process-based models have most often been used to derive optimal irrigation decisions (Cai et al., 2011; Wang and Cai, 2009) rather than to make volumetric irrigation demand forecasts with a few exceptions (e.g. (Tian and Martinez, 2014)).

Existing data-driven irrigation demand forecast models have been based on a variety of statistical (time series) models as well as artificial intelligence techniques such as artificial or computational neural networks (ANNs or CNNs) or other machine learning techniques. Although, these computational models have been successful in forecasting flood discharges (Nayak et al., 2005) and stream flow (Abrahart and See, 2002; Kasiviswanathan and Sudheer, 2013; Nayak et al., 2012), in which precipitation acts as the forcing variable, the inverse relationship between precipitation and irrigation demand combined with the variety of other influences has limited the irrigation demand forecast performances of these models. Nevertheless, a few CNN models have been developed to forecast one day lead time irrigation demands for several irrigation districts locate in the southern Spain (Pulido-Calvo and Gutierrez-Estrada, 2009; Pulido-Calvo et al., 2007; Pulido-Calvo et al., 2003). Those models didn't combine the precipitation and reference evapotranspiration (ET_0) forecasts which were available from computational models (Kuligowski and Barros, 1998; Tian and Martinez, 2012) or the Numerical Weather Prediction (NWP) models. These studies showed CNN calibration forecast performances that were higher than for univariate time series approaches, but the CNN models were systematically over fitted and the forecast performances declined significantly during validation (Pulido-Calvo and Gutierrez-Estrada, 2009). In an alternative approach, irrigation demand has been expressed as a multivariate output using the Bayesian machine learning algorithm called multivariate relevance vector machine (MVRVM). Using MVRMV, demand volumes for irrigation channels have been forecasted in the arid Sevier river basin, Utah, USA for lead times of one hour up to two days (Ticlavilca et al., 2011) and for 4 days lead times (Alfonso et al., 2011). These models did not consider precipitation as an input (due to the lack of precipitation in the arid climate) and instead used forcing variables such as temperature or ET_0. Overall, these data-driven models are "black-boxes" where the complex interactions between irrigation demand and causal factors remain ill defined. The heuristic nature of selecting the artificial intelligent technique, the network design, the input data set and the calibration period have led to difficulties in model architecture, network parameters, and frequent recalibrations.

Data-driven statistical time series models have also been used and are mostly limited to univariate (Pulido-Calvo and Gutierrez-Estrada, 2009; Pulido-Calvo et al., 2007; Pulido-Calvo et al., 2003), as they were mainly developed for arid-zone agriculture where precipitation does not significantly influence irrigation decisions and the inter-seasonal climate variability is low. These univariate cannot be successfully applied in supplemental irrigated agriculture due to structural constraints (inability to include precipitation effects). Multivariate time series model are an alternative that can be applied to model daily irrigation demand as a multivariate output resulting from bio-physical processes, farmer and operator behaviors and supply factors. These models can capture interactions between time series variables in terms of auto- and cross-correlations as well as linear (trend) and nonlinear (diurnal and seasonal) patterns, depending on the time step and model structure. Multivariate time series models have been widely used in engineering, science, medicine and finance, among other areas. ARMAX (autoregressive moving average with exogenous variable) models have been found to be very successful in forecasting short-term electricity demand (Chao-Ming et al., 2005; Hong-Tzer et al., 1996) and power output for photovoltaic systems (Li et al., 2014), where temperature acts as the exogenous variable. In the field of water resources engineering, ARX(autoregressive with exogenous variable) and ARMAX model structures have also been used extensively to forecast short term runoff (Haltiner and Salas, 1988), stream flow (Chang et al., 2001; Dutta et al., 2012; George et al., 2011; Sun et al., 2014), and urban water demand (Zhou et al., 2002; Zhou et al., 2000)) in which climate variables such as precipitation, ET_0 and temperature act as the exogenous variables. However, we are unaware of any multivariate time series models for forecasting short-term irrigation demand.

Most of the prevailing irrigation demand forecast models have been based on 2–3 years of irrigation. This means that inter-annual variations for a given season (e.g. summer) have not been well captured. The consistent fine scale irrigation demand data available from modernized irrigation systems along with improving short-term weather forecasts is significantly improving data availability and opening new opportunities to employ data-based demand forecasts. In this paper, we develop a multivariate ARMAX model to forecast daily irrigation demand for lead times of up to 5 days. It is structured to capitalize on the opportunity of using