WILEY

RESEARCH ARTICLE

Impact of aerosols on reservoir inflow: A case study for Big Creek Hydroelectric System in California

Farzana Kabir[1] | **Nanpeng Yu**[1] | **Weixin Yao**[2] | **Longtao Wu**[3] | **Jonathan H. Jiang**[3] | **Yu Gu**[4] | **Hui Su**[3]

[1]Electrical and Computer Engineering, University of California, Riverside, Riverside, California

[2]Department of Statistics, University of California, Riverside, Riverside, California

[3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

[4]Joint Institute for Regional Earth System Science and Engineering and Department of Atmospheric and Oceanic Science, University of California, Los Angeles, Los Angeles, California

Correspondence
Farzana Kabir, Electrical and Computer Engineering, Suite 343 Winston Chung Hall, University of California, Riverside, Riverside, CA 92521-0429.
Email: farzana.kabir@email.ucr.edu

Abstract

Accurate and reliable reservoir inflow forecast is instrumental to the efficient operation of the hydroelectric power systems. It has been discovered that natural and anthropogenic aerosols have a great influence on meteorological variables such as temperature, snow water equivalent, and precipitation, which in turn impact the reservoir inflow. Therefore, it is imperative for us to quantify the impact of aerosols on reservoir inflow and to incorporate the aerosol models into future reservoir inflow forecasting models. In this paper, a comprehensive framework was developed to quantify the impact of aerosols on reservoir inflow by integrating the Weather Research and Forecasting model with Chemistry (WRF-Chem) and a dynamic regression model. The statistical dynamic regression model produces forecasts for reservoir inflow based on the meteorological output variables from the WRF-Chem model. The case study was performed on the Florence Lake and Lake Thomas Alva Edison of the Big Creek Hydroelectric Project in the San Joaquin Region. The simulation results show that the presence of aerosols results in a significant reduction of annual reservoir inflow by 4–14%. In the summer, aerosols reduce precipitation, snow water equivalent, and snowmelt that leads to a reduction in inflow by 11–26%. In the spring, aerosols increase temperature and snowmelt which leads to an increase in inflow by 0.6–2%. Aerosols significantly reduce the amount of inflow in the summer when the marginal value of water is extremely high and slightly increase the inflow in the spring when the run-off risk is high. In summary, the presence of aerosols is detrimental to the optimal utilization of hydroelectric power systems.

KEYWORDS
aerosol, ARIMAX, dynamic regression, inflow simulation, water-energy nexus

1 | INTRODUCTION

Hydroelectric power plants play a key role in supporting the integration of increasing amounts of wind and solar energy as they have high level of operational flexibility and storage capability. Hydroelectric power plants take on important responsibilities such as flood control, navigation, irrigation, agricultural and urban water supply, and recreation in addition to enhancing the stability of power systems and security of power supply. Hence, it is important to determine the optimal operational schedule of single-stage or multistage hydroelectric power plants. An accurate and reliable reservoir inflow forecast model is in crucial need to enable optimal and efficient scheduling of hydroelectric resources (Gragne, Sharma, Mehrotra, & Alfredsen, 2015; Madsen, Richaud, & Pedersen, 2009; Valipour, Banihabib, & Behbahani, 2013).

Typically, the river run-off in the Sierra Nevada region are highly influenced by meteorological variables such as temperature, precipitation, and snow water equivalent (SWE) (Cayan, Riddle & Aguado, 1993). Because the reservoir inflows of this region are generated by the run-off captured by the reservoirs, therefore, these meteorological variables can be used as explanatory variables in reservoir inflow forecast models. In the past decade, researchers have discovered that the presence of aerosol particles in the atmosphere can exert great influence on the hydrological cycle in a region through the meteorological variables (Barnett, Adam, & Lettenmaier, 2005; Lohmann, 2005; Qian, Gustafson, Leung, & Ghan, 2009; Ramanathan et al., 2001).

Aerosols are a mixture of tiny particles or liquids that are suspended in air and can range from 0.001 to 10 μm in size. A discussion on impact of aerosols on temperature, SWE, and precipitation has been provided

in Section 6. A detailed description of effect of aerosol on precipitation and snow water equivalent in California is provided in Wu et al. (2018). Wu et al. (2018) showed that aerosols reduce precipitation and SWE by 10% over mountain tops in the Sierra Nevada region. This is a result of (both anthropogenic and naturally occurring) aerosols serving as cloud condensation nuclei (CCN), which leads to an increase of nonprecipitating clouds. CCN are aerosol particles that act as the initial sites for condensation of water vapour into cloud droplets. Aerosol deposition on snow increases absorption of solar radiation, leading to warming and further reduction of SWE over mountain tops. As the level of anthropogenic aerosol particles (such as sulfate and carbonaceous aerosols) increases rapidly from preindustrial times to the present-day over urban and industrial regions, their impact is becoming more significant on the hydrological cycle and thereby on reservoir inflow (Charlson, Langner, Rodhe, Leovy, & Warren, 1991; Charlson et al. 1992; Lohmann, 2005; Schwartz, 1996). It is critical to understand and quantify the impact of aerosols on reservoir inflow as it can influence hydropower generation and reservoir operations. It should be mentioned that the atmospheric lifetime of aerosols is very short, typically 2 to 4 days, making their effect on climate and weather more regional and less persistent into the future than those of the long-lived greenhouse gases (Hansson & Bhend, 2015; Verheggen & Weijers, 2010). Aerosols show large spatial and temporal variation in atmospheric aerosols concentrations and properties. Therefore, there are large differences in their effect on climate and weather on a regional basis (Hansson & Bhend, 2015; Penner et al., 2001; Ramachandran & Cherian, 2008; Regayre, et al., 2015; Samset, 2018; Verheggen & Weijers, 2010). Because inflows into the reservoirs are influenced by climatic variables, the impact of aerosols on reservoir inflows should also vary from region to region and should therefore be studied on a regional scale.

The primary objective of this paper is to fit a reservoir inflow forecast model and subsequently quantify the impact of aerosols on inflows into Florence Lake and Lake Thomas Alva Edison in the Big Creek Hydroelectric System. Because Florence Lake and Lake Edison are the higher elevation reservoirs of the system, an accurate forecast of inflow into these reservoirs can also improve the operational efficiency of the system greatly. The Big Creek Hydroelectric System resides in the San Joaquin Valley, which is surrounded by the Sierra Nevada mountain range in the east. San Joaquin Valley has one of the highest pollutant concentrations in the United States due to its unique geographical location. A detailed description of the study area is provided in Appendix B. Autoregressive integrated moving average model (ARIMA) is a well-known univariate time series model frequently used in hydrological forecasting. ARIMA models can predict a time series variable based on its own past values (AR term) and past values of the error term (MA term). Including exogenous variables in ARIMA model improves forecasting accuracy and is commonly known as ARIMAX model or dynamic regression model. In this paper, we first fit a statistical hydrologic model with dynamic regression method where meteorological variables such as temperature, precipitation, and SWE are used as explanatory variables. The best parsimonious dynamic regression model is selected using the Akaike information criterion (AIC), residual diagnostics and goodness of fit. Meteorological variables are then simulated using the Weather Research and Forecasting

model with Chemistry (WRF-Chem) with different aerosol emission levels. These simulated meteorological variables with and without aerosol impacts are fed into the dynamic regression model to quantify the impact of aerosols on reservoir inflow in the Big Creek Hydroelectric System. Detailed analysis of aerosol impacts on temperature, precipitation and SWE in California is not the objective of this study since it has been provided in Wu et al. (2018).

The unique contributions of this paper are listed as follows.

1. We developed an innovative and comprehensive framework for evaluating the impact of aerosols on reservoir inflow. The framework seamlessly integrates the numerical weather forecasting model (WRF-Chem) and the statistical inflow forecasting model (dynamic regression).
2. We fitted a dynamic regression model to forecast daily inflow into the hydroelectric reservoirs. The model coefficients for the meteorological variables provide an intuitive understanding of how temperature, precipitation, and snow water equivalent influence reservoir inflow.
3. We quantified the impact of aerosols on reservoir inflow in the Big Creek Hydroelectric System based on the proposed dynamic regression model and WRF-Chem model. The simulation results show that the presence of aerosols resulted in a reduction of the annual reservoir inflow by 4–14%.

The existing research on the effect of climate change and human activities on streamflow (Gleick & Chalecki, 1999; Knowles & Cayan, 2002; Lettenmaier & Gan, 1990; VanRheenen, Wood, Palmer, & Lettenmaier, 2004) and inflow into reservoirs (Brekke, Miller, Bashford, Quinn, & Dracup, 2004) in the San Joaquin Basin focus on the effect of carbon dioxide and several other greenhouse gases. There are very few literature studying the effect of natural and anthropogenic aerosols on streamflow and reservoir inflow (Givati & Rosenfeld, 2007). Our study focuses on exploring the impact of aerosols on inflow at the Big Creek Hydroelectric System located in the upper San Joaquin River system in the Sierra Nevada Mountains of Central California.

The remainder of the paper is organized as follows. Section 2 summarizes existing studies on statistical inflow forecasting models and discusses the rationality of choosing dynamic regression model. Section 3 presents the overall framework of our study. Section 4 presents the technical methods used in fitting the dynamic regression model to forecast reservoir inflow and the WRF-Chem model. Section 5 describes steps of fitting the dynamic regression model and their goodness of fit. Section 6 shows the evaluation of the WRF-Chem model and the impact of aerosols on inflow into the two hydropower reservoirs. Section 7 concludes the paper by discussing the direction of future research and limitations of the study. A list of acronyms used in this study is provided in Appendix A. Appendix B describes the study area. Lastly, Appendix C provides an overview of the dynamic regression model and a description of the methods used in fitting the model.

2 | LITERATURE REVIEW

This section presents a review of research articles relevant to this paper, which can be grouped into two categories: (a) statistical inflow forecasting models and (b) impact of aerosols on reservoir inflow.

2.1 | Statistical inflow forecasting models

The existing models for hydrological modelling and forecasting can be separated into three groups: time series models (Moeeni, Bonakdari, Fatemi, & Zaji, 2017; Mohammadi, Eslami, & Dardashti, 2005; Papamichail & Georgiou, 2001; Valipour et al., 2013; Valipour, 2015), regression models (Galeati, 1990; Lall & Bosworth, 1994; Mohammadi et al., 2005), and artificial neural network (ANN) models (Coulibaly, Anctil, & Bobée, 2000; Jain, Das, & Srivastava, 1999; Kilinç & Ciğizoğlu, 2005; Mohammadi et al., 2005; Valipour et al., 2013; Xu & Li, 2002). Mohammadi et al. (2005) compared regression, ARIMA, and ANN models to forecast spring inflows into the Amir Kabir reservoir in the Karaj watershed. Valipour et al. (2013) compared ARMA, ARIMA, and the autoregressive ANN models to forecast monthly inflows of the Dez dam reservoir. Both of these studies chose ANN as the best model. Moeeni et al. (2017) compared SARIMA (seasonal ARIMA) and ANN-GA (ANN combined with genetic algorithm) models in making short-term and long-term predictions of monthly inflow into a dam where SARIMA model outperformed the ANN-GA model, especially in forecasting low values. Papamichail and Georgiou (2001) used stochastic SARIMA model to forecast monthly inflow of one or more months ahead into the planned Amopeos Reservoir in Northern Greece which helped evaluate the optimal real time reservoir operation policies. The monthly forecasts were used to generate a synthetic series of monthly inflows that preserves the key statistics of the historical monthly inflows and their persistence Hurst coefficient, providing a probabilistic framework for reservoir design. Monthly means and the monthly standard deviations of the forecasted inflows were close to that of the measured inflows demonstrating the ability of SARIMA models to forecast monthly inflows and generate synthetic series of monthly inflows. Valipour (2015) investigated SARIMA and ARIMA models for long-term run-off forecasting in the United States. They found SARIMA model to be the best model in their study with an error of < 5% for all states. Therefore, ARIMA model can be considered as an effective tool for forecasting reservoir inflow. Including exogenous covariates in ARIMA model helps explain the dynamic relationship between the response time series and the explanatory variable time series and improve forecast accuracy. This is called a dynamic regression model. Dynamic regression model is also referred to as ARIMAX model.

For time series data, using dynamic regression model is preferred over ordinary regression because some of the underlying assumptions of regression model, for example, normal distribution, homoscedasticity and no autocorrelation of error terms, are frequently violated when being applied to time series data (Makridakis, Wheelwright, & Hyndman, 2008). Applying the ARIMA modelling approach to model the information contained in the error term of the regression model can take care of its autocorrelation. The transfer function in the dynamic regression model captures the time-lagged relationship of input variables and the predictor variable. Therefore, the dynamic regression model can also be thought of as a regression model with time-lagged inputs and ARIMA model for disturbances. Lastly, though ANN models might improve forecast accuracy, it is challenging to interpret the impact of aerosols on inflow by examining the weights on the meteorological input variables. In the light of all these considerations, we decided to adopt the dynamic regression model to forecast inflow into the hydropower reservoirs.

2.2 | Impact of aerosols on hydrology

The presence of aerosol articles have impact on the hydrological cycle through its impact on earth's radiative forcing, precipitation and snow water equivalent (Barth et al., 2005; Lohmann, 2005; Ramanathan, Crutzen, Kiehl, & Rosenfeld, 2001). It has been shown that an increase in atmospheric aerosols primarily affects solar radiation entering earth's atmosphere, snow albedo, cloud formation, and precipitation. Aerosol effects can be differentiated in three pathways—aerosol-radiation interaction (ARI) or direct effect, aerosol-snow interaction (ASI), and aerosol-cloud interaction (ACI) or indirect effect. Reflective aerosol particles, such as nitrate and sulfate particles, scatter the solar and thermal radiation and increase planetary albedo cooling both surface and atmosphere (Andreae, Jones, & Cox, 2005; Charlson & Schwartz, 1992; Haywood & Boucher, 2000; Johnson, Shine, & Forster, 2004; Kaufman, Tanré, & Boucher, 2002; Kiehl & Briegleb, 1993; Penner et al., 2006; Quaas, Boucher, Bellouin, & Kinne, 2008). However, light-absorbing aerosols, such as black carbon absorb radiation known as LAA, decrease planetary reflectivity and increase air temperature (Jacobson, 2001; Johnson et al., 2004). Presence of soot particles and dust in snow darkens the surface and reduces the snow albedo through ASI (Chýlek, Ramaswamy, & Srivastava, 1983; Clarke & Noone, 1985; Doherty, Warren, Grenfell, Clarke, & Brandt, 2010; Flanner, Zender, Randerson, & Rasch, 2007; Grenfell, Light, & Sturm, 2002; Hansen & Nazarenko, 2004; Jacobson, 2004; Lee-Taylor & Madronich, 2002; Marks & King, 2013; 2014; Reay, France, & King, 2012; Warren, 1984; Warren & Clarke, 1990; Wiscombe & Warren, 1980; Ye etal., 2012). Snow albedo perturbations increase the surface air temperature and accelerate snowmelt (Barnett, Adam, & Lettenmaier, 2005; Flanner et al. 2007; Hansen & Nazarenko, 2004; Lau, Kim, Kim, & Lee, 2010; Ming et al., 2009; Qian, Gustafson, Leung, & Ghan, 2009; Wiscombe & Warren, 1980; Xu et al., 2009). Further reduction of snow albedo takes place by snow albedo feedback (Brandt, Warren, & Clarke, 2011; Flanner et al. 2007; Hadley & Kirchstetter, 2012; Hansen & Nazarenko, 2004). Snow grain size, shape, and Black carbon-snow mixing type also play important roles in ARI (He et al, 2017; Kokhanovsky, 2013; Liou et al., 2014; Räisänen, Makkonen, Kirkevåg, & Debernard, 2017; Wiscombe & Warren, 1980). Internal mixing of light-absorbing aerosols and snow reduces snow albedo more than external mixing, which enhances the aerosol-induced snow albedo reduction. Nonspherical snow grains tend to show less aerosol-induced snow albedo reductions compared with spherical snow grains. These two opposite effects on snow albedo reductions by light-absorbing aerosols may further influence snowmelt and SWE and have merit for in-depth studies (He et al., 2017; He, Liou, & Takano, 2018; Liou et al., 2014).

ACI or indirect effect of aerosols on climate includes a change in microphysical and optical properties of cloud droplets, which is related to aerosols acting as CCN. Increasing the number concentration of CCN can lead to formation of more cloud droplets, which results in a decrease in cloud droplet radius leading to higher cloud albedo (Jones, Roberts, & Slingo, 1994; Twomey, 1974, 1991). Another effect of

decrease in cloud droplet size is the reduced precipitation through the "second indirect effect" (Ramanathan et al., 2001; Rosenfeld, 2000). This is due to the fact that small water droplets continue to drift in air and are less likely to grow to sufficient size to fall out as precipitation prolonging cloud lifetime (Ackerman, Kirkpatrick, Stevens, & Toon, 2004; Albrecht, 1989; Kaufman, Koren, Remer, Rosenfeld, & Rudich, 2005; Rosenfeld, 2000). Higher cloud reflectivity and increase in cloud lifetime also produce a net cooling effect on earth's surface by shading it from solar radiation. Absorptive aerosols can reduce low-cloud cover through the "semidirect effect" (Hansen, Sato, & Ruedy, 1997; Johnson et al., 2004) leading to positive radiative forcing. Glaciation aerosol effect is a possible counteracting effect where an increase in ice nuclei by anthropogenic aerosols (mineral dusts and a fraction of hydrophilic soot particles) acting as ice nuclei causes supercooled liquids to freeze (Lohmann, 2002; Lohmann & Feichter, 2005). The ice crystals quickly grow at the expense of cloud droplets because the vapour pressure over ice is lower than that over water, leading to more frequent glaciation of supercooled clouds. The precipitation formation via the ice phase is more efficient than in warm clouds, and therefore, the glaciated clouds have a shorter lifetime than supercooled water clouds leading to more precipitation. Chemical nature of the dust determines whether glaciation or warm cloud lifetime effect is larger. Borys, Lowenthal, Cohn, and Brown (2003) showed that the smaller mean droplet size in supercooled cloud caused by anthropogenic aerosols can significantly reduce ice particle riming efficiencies in mid altitude orographic clouds, resulting in lower orographic snowfall rates.

2.2.1 | Impact of aerosols on water resources

Surface run-off is a major component of the hydrological cycle. It is defined as water from precipitation, snowmelt, or other resources that flows over the land surface. Few studies were conducted to examine the impact of anthropogenic aerosols on water resources. Painter et al. (2010) studied the effect of dust radiative forcing on snow and run-off from the Upper Colorado River Basin. Disturbance of soil surfaces in the Colorado Plateau and biological crusts occurred in mid-1800s due to dramatic growth in grazing, agriculture and resource exploration. They used the Variable Infiltration Capacity model with postdisturbance and predisturbance impacts of dust on snow albedo and estimated the impact on run-off from the Upper Colorado River Basin from 1996 to 2003 at Lees Ferry, Arizona. Dust loading observed in 2005–2008 was used in the study. They found that the resulting short duration of snow covers leads to a 3-week early peak run-off and a decreased annual run-off (5%) due to increase in evotranspiration from earlier loss of snow cover. The magnitude of difference in run-off increased with the magnitude of the annual run-off. A follow-up of the study was performed by Deems, Painter, Barsugli, Belnap, and Udall (2013) developing a new snow albedo decay parametrization based on observations of levels of dust loading in 2009–2010 as they were unprecedentedly high, being on the order of five times that of 2005–2008. The extreme dust scenario caused the peak snowmelt to occur an additional 3 weeks earlier and further reduced the annual inflow by 1%.

In addition to studying the impact of the deposition of soot aerosol on snow and the resulting impact on snowpack, Qian et al. (2009) also studied its effect on the hydrological cycle in the western United States. They performed a yearlong simulation of WRF-Chem to simulate an annual cycle of soot aerosol deposition on snow and used it to estimate snow albedo perturbations induced by the soot within the western United States. This was followed by three regional climate simulations at Columbia River Basin, the Sacramento-San Joaquin River Basin, the Central Rockies, and the Sierra Nevada mountains. They used WRF in meteorology only mode (WRF-RCM) to capture precipitation, snowpack, and run-off, but with or without the perturbed snow albedo. They found that snow albedo reduction and the snow albedo feedback accelerated snowmelt and altered the streamflow that includes a trend towards earlier melt dates. In western United States, the main contribution to total run-off during winter is surface run-off generated by liquid rain. During spring, both precipitation and snowmelt contribute to run-off. As a result of warming in the soot-perturbed simulation, there are significant reductions in snowpack during the snowy winter period, which are reflected in reduced snow accumulation and more run-off during winter and less snow melt during spring. Run-off increases during late winter because the higher surface temperature in the soot-perturbed simulation causes more precipitation to come in the form of rain rather than snow. By contributing directly to run-off or by causing snowmelt, a higher percentage of rainfall versus snowfall during the cold season increases run-off. As less snow accumulates during winter, run-off as a result of snowmelt decreases during late spring. Qian, Flanner, Leung, and Wang (2011) used a global climate model to simulate the effect of black carbon and dust in snow on the hydrological cycle of the Tibetan Plateau. They found that surface air temperature increased by around 1°C averaged over the Tibetan Plateau and the spring snowpack was reduced due to the presence of black carbon and dust in snow. This had a significant impact on the hydrology, with the discharge increasing during late winter and early spring and decreasing during late spring and early summer showing a trend towards earlier melt dates.

Matt, Burkhart, and Pietikäinen (2018) developed a snow algorithm that allowed for the deposition mass flux of difference species of light-absorbing aerosols as an input variable for application in a rainfall–run-off model allowing determination of the effect of various light-absorbing aerosols at the catchment scale. They demonstrated the effect of black carbon deposition on snow on the hydrologic cycle through the implications for snowmelt and discharge generation on a remote southern Norwegian catchment over a period of 6 years. Their results indicate a significant impact of black carbon in snow at the catchment with run-off increasing in the spring followed by a decrease in discharge because of a trend towards earlier melt date and decrease in the catchment's snow-covered area.

Givati and Rosenfield (2004, 2005) quantified the suppression of orographic precipitation by anthropogenic aerosols over hills downwind of major coastal urban areas in California and Israel and subsequently extended it in Givati and Rosenfeld (2007) to study the impact of anthropogenic aerosols on available water resources in the Sea of Galilee in northern Israel and outflows of the main springs of Jordan River where large portion of water resources result from orographic precipitation. In Givati and Rosenfeld (2004, 2005), they defined the suppression of orographic precipitation as a reduction in the orographic enhancement factor R_o, where R_o is defined as the ratio between the precipitation amounts in the hills to the precipi-

tation in the upwind lowland. Time series of R_o from 1880 to 2000 based on annual precipitation from rain gauges downwind of major urban areas was compared with rain gauges sidewind of the area. A decrease in R_o with time at locations downwind of air pollution sources was explained by the increase in small-particulate air-pollution emissions with the growth of urban areas. The suppression rate was found to be 15–20% in hilly areas in California and Israel. Such decreasing trend was not found in hills downwind of pristine areas. They applied this methodology in Givati and Rosenfeld (2007) to measure trends of the ratio of annual precipitation between hilly to upwind lowland rain gauges and subsequently quantified the trend in orographic precipitation in the catchment areas. Then, they related it to trends in run-off and spring outflows by examining the relation of the trends of the spring outflow and the recharging area of the springs, thereby correlating the loss of precipitation to loss of overall water inflow. They concluded that, air pollution is the main reason behind the suppression of orographic precipitation over the hilly areas and the subsequent decreasing trend in the available water in the Sea of Galilee.

These studies of impact of aerosols on hydrology focuses only on one of the aerosol sources or pathway and few focus on reservoirs. Our study presents a complete account of the aerosol impacts from different sources through three pathways on two hydropower reservoirs in the Sierra Nevada region of California.

3 | FRAMEWORK

This study aims at quantifying the impact of aerosol particles on inflow into Florence Lake Reservoir and Lake Thomas Alva Edison and calculating daily inflow forecasts for these two reservoirs. A dynamic regression model was fitted to forecast the inflow that uses meteorological variables like daily mean temperature, accumulative snow water equivalent, and incremental precipitation as explanatory variables. Observed inflow data and observed meteorological variables data were split into a training set and a test set. Test set was formed by withholding the data for the last water year from the model identification and estimation process, and the rest are used as the training set. The training data set was used to estimate the model parameters. The forecasting accuracy of the model was assessed by performing out-of-sample forecasting on the test set. Forecasts of the predictor variables considering the impact of aerosols on regional climate were calculated in the San Joaquin Valley of California using a version of Weather Research and Forecasting model with Chemistry (Grell et al., 2005; Zhao et al., 2014) with fully coupled aerosol-meteorology-snowpack. Meteorological variable forecasts without the impact of aerosols were also calculated for the same region. Both forecasts were used as respective test sets for calculating inflow forecasts with and without impact of aerosols. Yearly and seasonally aggregated inflow forecasts were then compared with an aim to quantify the impact of aerosols on inflow into Lake Edison and Florence Lake. The procedure is summarized schematically in Figure 1.

4 | TECHNICAL METHODS

4.1 | Dynamic regression model

A dynamic regression model or ARIMAX model (Pankratz, 1991) uses time-lagged explanatory variables to forecast the dependent variable while modelling the error term with an ARIMA model (Box, Jenkins, Reinsel, & Ljung, 2015). Reasons for choosing dynamic regression model were discussed in Section 2. The model can be written as Equation 1.

$$Y_t = \mu + \sum_{i=1}^{M} \frac{\omega_i(B)}{\delta_i(B)} B^{b_i} X_{i,t} + \frac{\theta(B)}{\phi(B)} a_t, \qquad (1)$$

Y_t = dependent variable
$X_{i,t}$ = ith explanatory variable
$\omega_i(B)$ = numerator polynomial of the transfer function
$\delta_i(B)$ = denominator polynomial of the transfer function
b_i = dead time for input $X_{i,t}$
B = backshift operator
$\phi(B)$ = autoregressive operator
$\theta(B)$ = moving-average operator
a_t = white noise.

We conducted model fitting by applying relevant theory to choose the input variables and then following standard methodology for fitting dynamic regression models. The linear transfer function (LTF) method suggested by Pankratz (1991) was applied here to specify the transfer functions, and the methodology described by Box et al. (2015) was applied to determine ARMA order of the error time series. Finally, the coefficients of the entire model were estimated, and the model was checked for adequacy. An overview of the dynamic regression model and linear transfer function method is provided in Appendix C.

4.1.1 | Performance metrics

Out-of-sample forecasting was performed to assess the forecasting accuracy of the model (Makridakis et al., 2008). Some of the sample data at the end of the time series were withheld as the test data set. They were not used in the model identification and estimation process. The fitted model was used to forecast the response variable. Root mean square error (RMSE), mean absolute error (MAE), Nash-Sutcliffe efficiency (NSE), a modified version of NSE (Garrick, Cunnane, & Nash, 1978; Legates & McCabe Jr., 1999), percent bias (PBIAS), and RMSE-observations standard deviation ratio (RSR; Moriasi et al., 2007) were used as accuracy metrics to evaluate the performance of the proposed model and the benchmark models introduced in Section 5.4.2.

Moriasi et al. (2007) used three quantitative statistics, for model evaluation, namely Nash-Sutcliffe efficiency (NSE), percent bias (PBIAS), and RMSE-observations standard deviation ratio (RSR). In addition to the above three statistics, we used a modified version of NSE proposed by Garrick et al. (1978) and Legates and McCabe Jr. (1999) for model evaluation.

MAE and RMSE

RMSE and MAE values signify the goodness of fit of the forecast to the observed inflow and hence can evaluate the performance of the dynamic regression model.

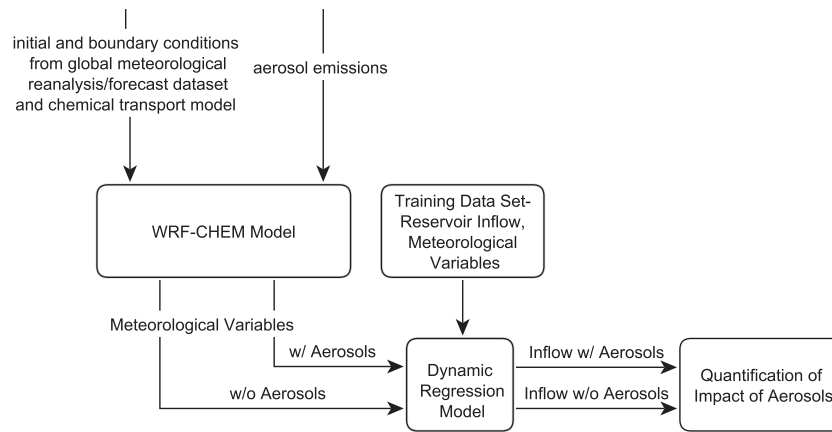$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(Y_{fi} - Y_{oi})^2}{n}}, \qquad (2)$$

FIGURE 1 The overall framework for quantifying the impact of aerosols on reservoir inflow

$$MAE = \sum_{i=1}^{n} \frac{|Y_{fi} - Y_{oi}|}{n}. \quad (3)$$

Here, i denotes the day in a water year, Y_{fi} represents the forecasted inflow on day i, Y_{oi} denotes the observed inflow on day i, and n is the number of days in the water year.

NSE and modified NSE

The Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe 1970) is the most widely used indicator in hydrology because of its flexibility to apply to different types of mathematical models and intuitive interpretability (McCuen, Knight, & Cutter, 2006; Ritter & Muñoz-Carpena, 2013; Schaefli & Gupta, 2007). It has been used widely in streamflow and run-off predictions (Criss & Winston, 2008; Krause, Boyle, & Bäse, 2005; Li, Luo, Jiang, Wan, & Li, 2017; Moriasi et al., 2007; Noh et al., 2016; Schaefli & Gupta, 2007). It is a normalized measure comparing the mean square error generated by a model simulation to the variance of the observed values. NSE effectively compares the performance of a particular model to a simple model that uses mean of the observed values as prediction. NSE ranges from minus infinity to 1, with higher values indicating better forecast.

$$NSE = 1 - \frac{\sum_{i=1}^{N} (Y_{oi} - Y_{fi})^2}{\sum_{i=1}^{N} (Y_{oi} - \bar{Y}_o)^2}. \quad (4)$$

Here, $\bar{Y}_o$ is the mean of the observed inflow of the test period. Several researchers have suggested modifications to the NSE owing to its limitations, such as using mean of the observations as the baseline model and possible effect of outliers on NSE (Garrick et al., 1978; Krause et al., 2005; Legates & McCabe Jr., 1999; Oudin, Andréassian, Mathevet, Perrin, & Michel, 2006). Garrick et al. (1978) termed the use of mean of observed values as primitive, and proposed using the daily mean value of the predictor variable for the calibration period so that the baseline model can indicate seasonal variation of the predictor variable. This modification of NSE is also recommended by the World Meteorological Organization (1986). Schaefli and Gupta (2007) recommends using benchmark models appropriate to the particular case

study. The modified NSE or benchmark efficiency (BE) indicates performance improvement of the hydrologic model over the benchmark model and can be written as

$$BE = 1 - \frac{\sum_{i=1}^{N} (Y_{oi} - Y_{fi})^2}{\sum_{i=1}^{N} (Y_{oi} - Y_{bi})^2}. \quad (5)$$

Here, Y_{bi} represents the forecasted inflow on day i by the benchmark model. Because NSE is calculated by squaring the deviations between observation and model-calculated values, it is sensitive to extreme values leading to misevaluation of model performance (Criss & Winston, 2008; Krause et al., 2005; Legates & McCabe Jr., 1999; Willmott, Robeson, & Matsuura, 2012). Legates and McCabe Jr. (1999) proposed NSE′, a modification of NSE using absolute values instead of squared deviation reducing the effect of squared terms. In Legates and McCabe (2012), they maintain the recommendation of NSE and NSE′ because of their intuitive interpretability and having a fundamental meaning at zero. In general, NSE′ has a lower value than NSE.

$$NSE = 1 - \frac{\sum_{i=1}^{N} (Y_{oi} - Y_{fi})^2}{\sum_{i=1}^{N} (Y_{oi} - \bar{Y}_o)^2}. \quad (6)$$

PBIAS

Percent bias (PBIAS) measures the average tendency of the simulated data to be larger or smaller than their observed counterparts (Gupta, Sorooshian, & Yapo, 1999). Optimal value of PBIAS is 0 with low magnitude values indicating accurate model simulation. Positive values indicate model underestimation bias and negative values indicate model overestimation bias.

$$PBIAS = \frac{\sum_{i=1}^{N} (Y_{oi} - Y_{fi}) \times 100}{\sum_{i=1}^{N} Y_{oi}}. \quad (7)$$

RSR

RSR standardizes RMSE using the observation standard deviation and is calculated as the ratio of the RMSE and standard deviation of measured data with lower RSR indicating better model prediction performance.

$$RSR = \frac{RMSE}{STDEV_{obs}} = 1 - \frac{\sqrt{\sum_{i=1}^{N}\left(Y_{oi} - Y_{fi}\right)^2}}{\sqrt{\sum_{i=1}^{N}\left(Y_{oi} - \bar{Y}_o\right)^2}}. \tag{8}$$

4.2 | WRF-Chem model

The WRF-Chem model (Grell et al., 2005) is a weather research and forecasting system that simulates chemistry and aerosols simultaneously with meteorology. This model has been extensively used to study regional air quality and their interactions with weather and climate (e.g., Barnard, Fast, Paredes-Miranda, Arnott, & Laskin, 2010; Chapman et al., 2009; Fast et al., 2014; Fast et al., 2012; Qian et al., 2009; Wu, Su, & Jiang, 2011a; 2011b; 2013; Wu et al., 2017; Wu et al., 2018; Zhao et al., 2010; Zhao et al., 2013; Zhao et al., 2014). In this study, we used the WRF-Chem version 3.5.1 which includes aerosol interactions with radiation, cloud and snowpack (Zhao et al., 2014). In the WRF-Chem control (CTRL) experiment, the model is run at 4 km horizontal resolution with the model domain covering California and surrounding regions. The major components of aerosols (nitrate, ammonium, elemental carbon, primary organic matter, sulfate, sea salt, dust, water, and other inorganic matter) are simulated in the model along with their physical and chemical processes. Anthropogenic aerosol emissions are obtained from US EPA 2005 National Emissions Inventory (NEI05; US EPA, 2010). Aerosol emissions comprise SO_4, NO_3, EC, organic aerosols, and total $PM_{2.5}$ and PM_{10} masses. Anthropogenic emissions are updated every hour to account for diurnal variability. However, their seasonal variation is not considered in the simulation. Biogenic emissions are calculated online using the Model of Emissions of Gases and Aerosols from Nature (Guenther et al., 2006). Biomass burning emissions are obtained from the Global Fire Emissions Database version 2.1, with 8-day temporal resolution and monthly updates (Environmental Sciences Division, O. R. N. L., 2013). However, year-to-year variability in biomass burning aerosols is not taken into account. Dust emissions are calculated using the DUST TRANsport model scheme (Shaw, Jerry Allwine, Fritz, Rutz, Rishel, & Chapman, 2008) following Wu et al. (2017). Sea salt emissions are derived from the PNNL-updated sea salt emission scheme that includes the correction of particles with a radius less than 0.2 μm (Gong, 2003) and dependence on sea surface temperature (Jaeglé, Quinn, Bates, Alexander, & Lin, 2011).

To derive aerosol optical properties (e.g., extinction, single-scattering albedo, and the asymmetry parameter for scattering) as a function of wavelength, Mie calculations (Ghan et al., 2001) are used. ARI is included in the shortwave and longwave radiation schemes (Fast et al., 2006; Zhao, Liu, Ruby Leung, & Hagos, 2011). ACI is effectively simulated in WRF-Chem (Chapman et al., 2009) by linking simulated cloud droplet number with shortwave radiation and micro-

physics schemes. In this version of WRF-Chem (Zhao et al., 2014), aerosol snow interaction is implemented by considering the deposition of aerosol on snow and the subsequent radiative impacts through the SNICAR (SNow, ICe, and Aerosol Radiative) model (Flanner & Zender, 2005; 2006). The Morrison double-moment microphysics scheme (Morrison, Thompson, & Tatarskii, 2009), rapid radiative transfer model for general circulation models model shortwave and longwave radiation schemes (Iacono et al., 2008), and Community Land Model Version 4 land surface scheme (Lawrence et al., 2011) are the physics parametrizations used in the simulations. The Yonsei University planetary boundary layer scheme (Hong, Noh, & Dudhia, 2006) is used in the simulations. The initial and boundary conditions are provided by the European Center for Medium-Range Weather Forecasts Interim Re-Analysis (Dee et al., 2011) for meteorology and the global Model for Ozone and Related chemical Tracers, version 4 (Emmons et al., 2010) for chemistry. Wu et al. (2018) showed that the model simulations reproduced the spatial and temporal variation of observed precipitation well. More details of the model set-up can be found in Wu et al. (2018).

Wu et al. (2017, 2018) evaluated the model performance on simulating aerosols and meteorological variables in California. It has been shown that the model reasonably captures the distribution and seasonal variability of aerosols from October to June but underestimates aerosols from July to September. Because precipitation, snowpack, and inflow are mainly within October–June, the underestimation of aerosols in July–September has limited impacts on our results. The model reproduced the seasonal variations of temperature, precipitation, and SWE in California with some overestimation of temperature and SWE. In a CLEAN simulation, we turned off local aerosol emissions and set aerosols from boundary conditions as zero, but kept chemical components from boundary conditions with aerosol chemistry on. The CCN in the CLEAN experiment was on the order of 10 cm^{-3}, representing a clean environment. The simulations of clouds, precipitation, and radiation are reasonable in the CLEAN run. Thus, meteorological variables from the WRF-Chem CTRL and WRF-Chem CLEAN simulations represent conditions with and without considering impact of aerosols, respectively. Aerosol impacts on temperature, precipitation, and SWE were investigated in Wu et al. (2018) and discussed in Sections 1 and 6.

5 | FITTING INFLOW FORECASTING MODEL

In this section, we explain how to fit the dynamic regression model to forecast reservoir inflow of Florence Lake and Lake Edison, which are part of the Big Creek Hydroelectric Project in California.

5.1 | Data description

The data set contains the daily average reservoir inflow in cu ft/s for five consecutive water years 2010–2014. A water year or a hydrological year is a 12-month period between October 1 of one year and September 30 of the next year. To predict the reservoir inflow, we collected the meteorological data such as the daily air temperature, SWE, and incremental precipitation data from the website of California
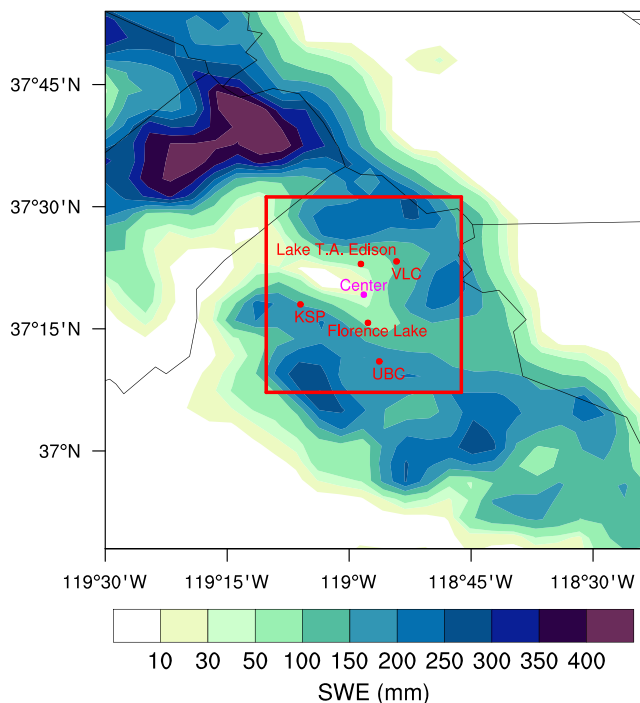
FIGURE 2 Study area with grid box and weather stations identified. Snow depth distribution averaged over water year 2013 is overlaid on the map

data exchange centre. Data for meteorological variables, inflow, and WRF-Chem simulations used in this study can be found in Department of Water Resources, C. (2017). The meteorological data were collected and averaged over three weather stations of Kaiser Point (KSP), Volcanic Knob (VLC), and Upper Burnt Corral (UBC) located within the $0.4 \times 0.4°$ grid box with centre at $(37.32°N, -118.97°E)$. The study area with the grid box is shown in Figure 2 with the snow depth distribution map averaged over water year 2013 overlaid on it. The observations of these meteorological variables are plotted in Figure 3.

5.2 | Predictor/variable selection

Selection of appropriate predictors or explanatory variables is essential for accurate forecast and simple model interpretation. The inflows are generated by the run-off captured by the reservoirs from the San Joaquin River. Streamflow in the Sierra Nevada region has high correlation with temperature, SWE, and precipitation (Cayan et al., 1993). Therefore, these three variables were included in the model to forecast the inflow.

Being a mountainous region, run-off in the Sierra Nevada region is dominated by snowmelt. Maximum run-off in the San Joaquin watershed occurs during the snowmelt run-off period (April–July; Serreze, Clark, Armstrong, McGinnis, & Pulwarty, 1999; Stewart, Cayan, & Dettinger, 2004). Accordingly, most of the reservoir inflows occur in the late spring and early summer between April and July in both Florence Lake and Lake Edison (Figure 3). Therefore, snowmelt during this period is a useful predictor for reservoir inflow. Snowmelt can be calculated by $\max(SWE_{t-1} - SWE_t, 0)$.

To handle the seasonality, four dummy variables were introduced in Table 1 to represent four periods in a year. These periods are

early spring, late spring, early summer, and late summer. We also added interaction terms between the four meteorological variables, temperature, SWE, precipitation, and snowmelt with seasonal dummy variables to model different effects of meteorological variables in different seasons. Because there is a lag of several months between the peak snow accumulation and peak inflow in our study area as seen from Figure 3, lagged snow water equivalent in the late summer was included in the dynamic regression model to capture this effect. To choose the appropriate lag, a stepwise regression is performed with up to 3-month lag of SWE as regressors along with current and lagged temperature, precipitation, and snowmelt as independent variables. Statistically significant lags of SWE are chosen as suitable candidates for inclusion in the final model. Eighty days lagged SWE minimized the AIC of the model during training period and is therefore chosen to be included in the final model. The complete list of variables used in fitting the statistical dynamic regression model is tabulated in Table 2.

5.3 | Model fitting

We explored the model performance with and without natural log transformation of the response and explanatory variables and chose untransformed variables for further model fitting as it offered better prediction results and model interpretation. Steps of fitting the dynamic regression model for Florence Lake inflow forecast is described here. Similar procedure can be followed for Lake Edison. Fitting dynamic regression model has three stages: (a) model identification, (b) model estimation, and (c) model diagnostic checking.

5.3.1 | Model identification

As the first step to identify the appropriate dynamic regression model, a free-form distributed lag for the transfer function of the explanatory variables was estimated with a low order regular AR term as proxy for the disturbance series autocorrelation pattern. A multiple regression model was formed, and stepwise regression was performed to preliminarily select candidate variables and their time lags for fitting the free-form distributed lag model. The orders of $v(B)$ for the explanatory variables in the free-form distributed lag model were determined to be 15 based on their t-test statistics. It can be argued that the inflow is zero when the explanatory variables are zero i.e. when there is no snowmelt or precipitation and the temperature is 0°F. Therefore, no constant term was included in the model.

The disturbance series N_t was then checked for stationarity by augmented Dickey-Fuller test and found to be stationary. A parsimonious rational distributed lag transfer function model of order (b, r, h) was identified by comparing the estimated impulse response weights with theoretical impulse response weight patterns. To demonstrate the process, the estimated impulse response weights of the variable X_4D_3 that corresponds to snowmelt in early summer are shown in Table 3 and plotted in Figure 4.

There are six significant v weights at Lags 1, 2, 4, 7, 9, and 15 having t value more than 2.0. This suggests that the dead time, $b_1 = 1$. Because the six significant impulse response weights follow an exponential decay pattern, the order of the denominator operator was determined to be $r = 1$. The number of unpatterned
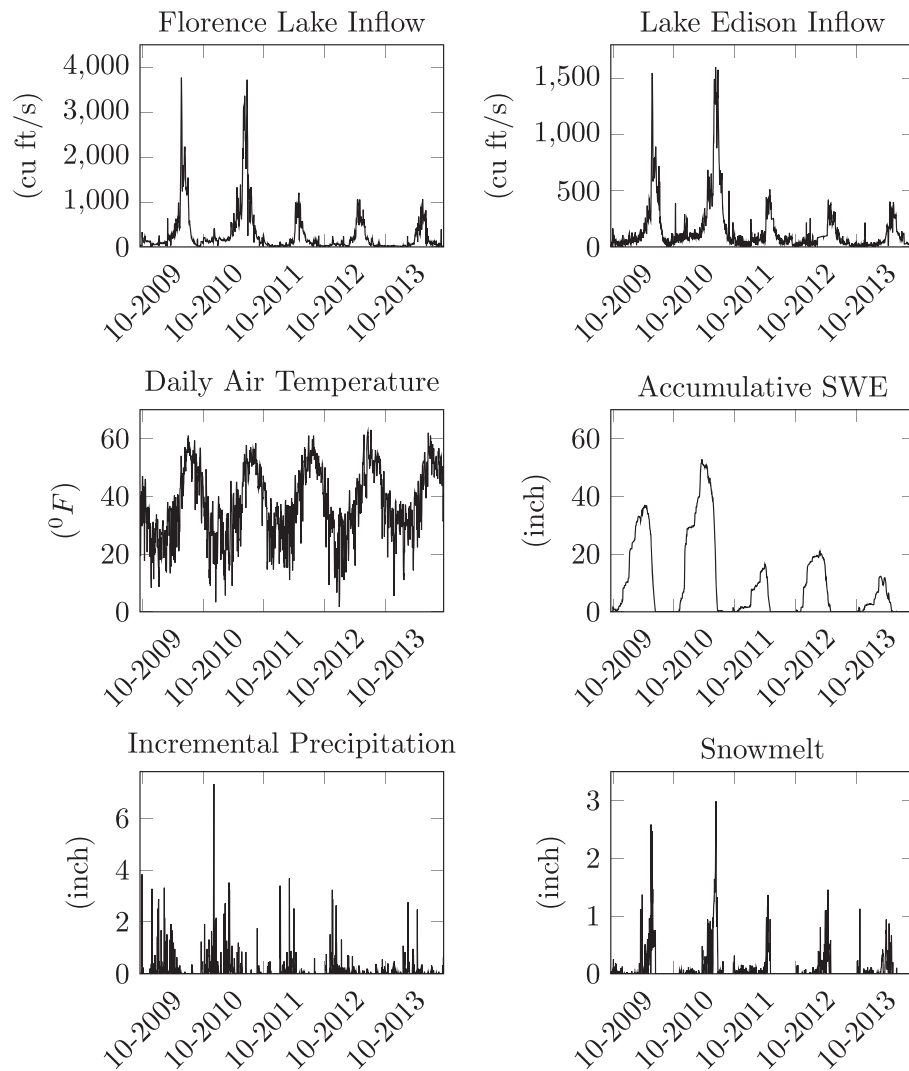
Florence Lake Inflow

Lake Edison Inflow

Daily Air Temperature

Accumulative SWE

Incremental Precipitation

Snowmelt

FIGURE 3 Response variables (inflow at Florence Lake and Lake Edison) and the explanatory variables (daily average temperature, SWE, incremental precipitation, and snowmelt averaged over three weather stations—UBC, KSP, and VLC) for water year 2010–2014

TABLE 1 Description of dummy variables used in the dynamic regression model

Dummy variables	Description	Season
D_1	1 if Date 03/21–04/30, 0 otherwise	Early Spring
D_2	1 if Date 05/01–05/31, 0 otherwise	Late Spring
D_3	1 if Date 06/01–06/21, 0 otherwise	Early Summer
D_4	1 if Date 06/22–09/22, 0 otherwise	Late Summer

TABLE 2 List of variables used in the dynamic regression model

Variables	Symbols
Reservoir inflow	Y
Temperature	X_1
SWE	X_2
Precipitation	X_3
Snowmelt	X_4
Dummy variables	D_1, D_2, D_3, D_4
Interaction terms	$X_1D_1, X_1D_2, X_1D_3, X_1D_4, X_2D_1, X_2D_2,$ $X_2D_3, X_2D_4, X_3D_1, X_3D_2, X_3D_3,$ $X_3D_4, X_4D_1, X_4D_2, X_4D_3, X_4D_4$

terms is $u = 0$. Finally, the order of the numerator operator is $h = u + r - 1 = 0 + 1 - 1 = 0$. The order of the rational distributed lag transfer function for input variable snowmelt in early summer (X_4D_3) was thus determined to be $(b, r, h) = (1, 1, 0)$, and the transfer function could be written as $\frac{\omega_j}{1-\delta_j B}B$. Similar procedure was followed for other input variables, and the dynamic regression model with parsimonious rational distributed lag transfer function was determined.

5.3.2 | Model estimation

An estimate of the parameters of the dynamic regression model was obtained at this stage. An appropriate ARMA model was identified for the error series N_t, and the entire model was refit using the ARMA model for error and the transfer function for the input variables. The parameter estimates of all candidate models were estimated by maximum likelihood estimation.

TABLE 3 Impulse response weights of input variable X_4D_3

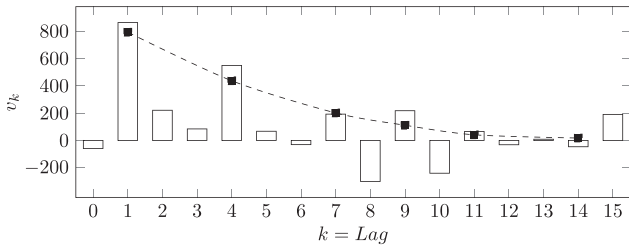| Lag	Estimate	t-value	Lag	Estimate	t-value	p-value
0	−59.55	−0.67	8	−300.71	−4.03	< 0.0001
1	862.75	10.83	9	217.35	3.33	0.0009
2	220.73	2.88	10	−240.18	−3.94	< 0.0001
3	84.11	1.01	11	65.52	1.09	0.2769
4	548.99	6.67	12	−31.77	−0.57	0.5714
5	67.09	0.84	13	8.60	0.15	0.8773
6	−30.91	−0.42	14	−46.42	−0.81	0.4191
7	192.40	2.56	15	190.53	3.39	0.0007

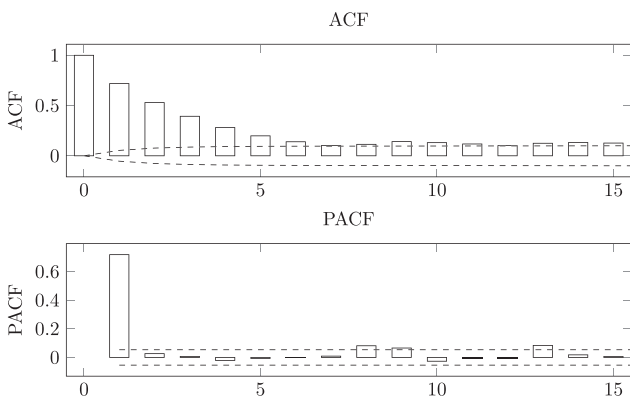**FIGURE 4** Impulse response weights of input variable X_4D_3

FIGURE 5 Autocorrelation function (ACF) and partial autocorrelation function (PACF) of noise series N_t

First, the model was fit using only the transfer function of the input variables. The orders of AR and MA component of the model were identified by matching empirical autocorrelation patterns, for example, autocorrelation function (ACF) plot and partial autocorrelation function (PACF) plots of the residual series with the AR and MA signature patterns. The ACF and PACF plots of the residual series are plotted in Figure 5. Both ACF and the PACF exhibit large spikes that gradually die out indicating that they have both autoregressive and moving averages properties. Though the ACF decays rather slowly and cuts off at Lag 6, an AR order of $p = 6$ is not realistic. An AR order of $p = 1$ was selected based on the AIC and ACF of the residuals. The PACF cuts off at Lag 1. Therefore, the final ARMA model for the error series was determined to be $(p, q) = (1, 1)$. At this stage, the dynamic regression model for Florence Lake can be written as Equation 9.

5.3.3 | Diagnostic checking

The Ljung-Box test for white noise was used to statistically evaluate the degree to which the residuals are free from serial correlation.

For seasonal time series, the lag for Ljung-Box test is recommended to be $h = \min\left(2m, \frac{T}{5}\right)$ where m is the period of seasonality and T is the sample size. In our study, the lag was calculated to be $h = 365 \times 3/5 = 219$. Though the residuals are not perfect white noise after $lag = 25$; for a long time series, this is acceptable. Moreover, Durbin Watson Statistic was calculated to detect presence of autocorrelation in the residuals and found to be 2, which shows that the residuals are not autocorrelated. Normality check of the model residuals was performed by checking a histogram of the residuals and the Q-Q normal plot of the residuals. The residuals were found to be approximately normally distributed, and the Q-Q normal plot is approximately a straight line. The explanatory variables in the final model were checked for multicollinearity. For all explanatory variables, variance inflation factor (VIF) was calculated. A VIF close to 1 for an explanatory variable indicates no correlation of that predictor and the remaining explanatory variables. For all explanatory variables in this model, VIF was found to be < 1.60. Hence, there is no multicollinearity. Because there is no significant residual cross correlation and autocorrelation left, the model is adequate. Similar procedure was followed for Lake Edison. The dynamic regression model for Florence Lake and Lake Edison can be written as Equations 9 and 10.

$$
\begin{aligned}
Y_t = &\left(\omega_{0,0} + B\omega_{0,1} + B^2\omega_{0,2} + B^3\omega_{0,3} + B^4\omega_{0,4}\right) X_{3,t} \\
&+ \left(\omega_{1,0} + B\omega_{1,1} + B^2\omega_{1,2} + B^3\omega_{1,3}\right) X_{1,t}D_{1,t} \\
&+ \left(\omega_{2,0} + B\omega_{2,1} + B^2\omega_{2,2} + B^3\omega_{2,3}\right) X_{1,t}D_{2,t} + \frac{\omega_3}{(1-\delta_3 B)}X_{4,t}D_{2,t} \\
&+ \left(\omega_{4,0} + B\omega_{4,1} + B^2\omega_{4,2} + B^3\omega_{4,3}\right) X_{1,t}D_{3,t} + \frac{\omega_5}{(1-\delta_5 B)}X_{4,t}D_{3,t} \\
&+ \omega_6 B^{80}\left(X_{2,t}\right) D_{4,t} + \frac{\omega_7}{(1-\delta_7 B)}BX_{4,t}D_{4,t} + \frac{(1-\theta_1 B)}{(1-\phi_1 B)}a_t.
\end{aligned}
$$

(9)

$$
\begin{aligned}
Y_t = &\omega_0 X_{1,t} + \omega_1 X_{3,t} + \left(\omega_{2,0} + B\omega_{2,1} + B^2\omega_{2,2} + B^3\omega_{2,3}\right) X_{1,t}D_{1,t} \\
&+ \left(\omega_{3,0} + B\omega_{3,1} + B^2\omega_{3,2} + B^3\omega_{3,3} + B^4\omega_{3,4}\right) X_{1,t}D_{2,t} \\
&+ \frac{\omega_4}{(1-\delta_4 B)}X_{4,t}D_{2,t} + \left(\omega_{5,0} + B\omega_{5,1} + B^2\omega_{5,2} + B^3\omega_{5,3}\right) X_{1,t}D_{3,t} \\
&+ \frac{\omega_6}{(1-\delta_6 B)}X_{4,t}D_{3,t} + \omega_7 B^{80}\left(X_{2,t}\right) D_{4,t} \\
&+ \frac{\omega_8}{(1-\delta_8 B)}BX_{4,t}D_{4,t} + \frac{(1-\theta_1 B)}{(1-\phi_1 B)}a_t.
\end{aligned}
$$

(10)

The estimated parameter values of the dynamic regression models are shown in Tables 4 and 5. The estimated parameter values of the dynamic regression models with their t-values, p-values, and standard error are shown in Tables 4 and 5.

5.4 | Results and analysis

5.4.1 | Parameter estimates and interpretation

The estimated parameter values of the dynamic regression models with their t-values, p-values, and standard error are shown in Tables 4 and 5. A large value of absolute t-statistic and low p-value (< 0.05) imply that the true parameter value is not 0. It can be observed that temperature, SWE, and precipitation play important roles in forecasting reservoir inflow. For both lakes, snowmelt during spring and summer has a strong and positive correlation with inflow. With

TABLE 4 Parameter estimates for inflow forecast model of Florence Lake

Coefficient	Value	Standard error	t-value	p-value
θ_1	0.07	0.04	1.91	0.06
ϕ_1	0.78	0.02	33.60	< 0.0001
$\omega_{0,0}$	14.06	8.81	1.60	0.1
$\omega_{0,1}$	6.37	9.08	0.70	0.5
$\omega_{0,2}$	9.24	9.10	1.01	0.3
$\omega_{0,3}$	7.74	9.079	0.85	0.4
$\omega_{0,4}$	12.79	8.83	1.45	0.1
$\omega_{1,0}$	1.16	1.27	0.92	0.3
$\omega_{1,1}$	6.17	1.38	4.47	< 0.0001
$\omega_{1,2}$	1.06	1.38	0.77	0.4
$\omega_{1,3}$	2.10	1.26	1.66	0.09
$\omega_{2,0}$	2.13	1.23	1.73	0.08
$\omega_{2,1}$	8.40	1.29	6.51	< 0.0001
$\omega_{2,2}$	−0.50	1.31	−0.38	0.7
$\omega_{2,3}$	0.06	1.23	0.05	0.9
ω_3	93.18	23.23	4.01	< 0.0001
δ_3	0.95	0.02	50.05	< 0.0001
$\omega_{4,0}$	0.86	1.29	0.66	0.5
$\omega_{4,1}$	3.05	1.12	2.72	0.006
$\omega_{4,2}$	1.97	1.12	1.76	0.08
$\omega_{4,3}$	2.49	1.10	2.25	0.02
ω_5	324.03	38.22	8.48	< 0.0001
δ_5	0.79	0.03	26.68	< 0.0001
ω_6	12.92	1.94	6.66	< 0.0001
ω_7	149.15	19.66	7.59	< 0.0001
δ_7	0.92	0.01	62.27	< 0.0001

TABLE 5 Parameter estimates for inflow forecast model of Lake Edison

Coefficient	Value	Standard error	t-value	p-value
θ_1	−0.03	0.03	−0.89	0.4
ϕ_1	0.86	0.02	49.14	< 0.0001
ω_0	0.52	0.24	2.25	0.02
ω_1	21.55	2.75	7.85	< 0.0001
$\omega_{2,0}$	0.31	0.46	0.69	0.5
$\omega_{2,1}$	2.39	0.45	5.36	< 0.0001
$\omega_{2,2}$	0.97	0.44	2.17	0.03
$\omega_{2,3}$	0.56	0.43	1.30	0.2
$\omega_{3,0}$	1.16	0.45	2.58	0.0099
$\omega_{3,1}$	2.18	0.42	5.15	< 0.0001
$\omega_{3,2}$	1.75	0.43	4.06	< 0.0001
$\omega_{3,3}$	0.64	0.42	1.51	0.1
$\omega_{3,4}$	0.47	0.37	1.27	0.2
ω_4	56.90	12.65	4.50	< 0.0001
δ_4	0.31	0.23	1.35	0.2
$\omega_{5,0}$	0.58	0.46	1.26	0.2
$\omega_{5,1}$	0.68	0.37	1.80	0.07
$\omega_{5,2}$	1.38	0.37	3.69	0.0002
$\omega_{5,3}$	2.12	0.37	5.71	< 0.0001
ω_6	126.80	14.58	8.70	< 0.0001
δ_6	0.84	0.02	34.52	< 0.0001
ω_7	2.48	0.88	2.82	0.0048
ω_8	85.75	8.24	10.40	< 0.0001
δ_8	0.94	0.008	107.91	< 0.0001

high t-values, snowmelt is the most important variable in explaining the variability of inflow. This result is consistent with the fact that the run-off in the Sierra Nevada region is dominated by snowmelt. Prior season's SWE is also found to be a useful predictor for inflow during late summer. This can be explained by the fact that the snowpack during cold seasons plays a crucial role in run-off and subsequent reservoir inflow during warmer seasons in the Sierra Nevada region. Current season's temperature has a positive correlation with reservoir inflow in early/late spring and early summer. This is because, in higher elevation rivers, warmer temperature produces faster run-off and less snow (Cayan et al., 1993). Apart from early/late spring and early summer, temperature does not have a significant impact at Florence Lake but has moderate impact at Lake Edison. As shown in the model fitting results, the same season precipitation has significant impact on reservoir inflow at both Florence Lake and Lake Edison. As expected, precipitation is positively correlated with inflow since a higher level of precipitation generally results in more inflow.

5.4.2 | Inflow forecast using dynamic regression model

The performance of the dynamic regression model during the calibration period is evaluated using NSE, NSE′, PBIAS, RSR, RMSE, and MAE. The statistical indexes of model performance are shown in Table 6.

Limits of acceptability of the performance metrics depends on model applications and is therefore subjective (Beven, 2006). Follow-ing Motovilov, Gottschalk, Engeland, and Rodhe (1999), the model performance at daily time step is generally considered to be good when NSE ≥ 0.75 and satisfactory when 0.36 ≤ NSE < 0.75. On the basis of examples of various existing models and research data, Moriasi et al. (2007) proposed general performance ratings for these statistics. Those are provided in Table 7. However, because typically model simulations are poorer for shorter time steps than longer time steps (e.g., daily vs. monthly) (Bernard, Dan, Mike, Jeff, & Mazdak, 2007), a less strict performance rating is required for daily time steps used in our study (Moriasi et al., 2007). In general NSE′ has a lower value than NSE. For NSE′, the model can be considered satisfactory if NSE′ ranges from 0.51 to 0.71 (Licciardello, Zema, Zimbone, & Bingner, 2007).

Value of NSE is very high and PBIAS and RSR values are very low during the calibration period indicating an excellent agreement between the observed and simulated inflows into both Florence Lake and Lake Thomas A. Edison. After calibrating the dynamic regression model with the help of historical data, the next step of the study is to simulate inflow corresponding to the future meteorological variables. The forecast period is set to 365 days. The out-of-sample forecasting ability of the model was assessed by forecasting the reservoir inflow for both lakes in water year 2014 using a test set containing average meteorological data for the grid box region. Actual and forecasted inflow for Lake Edison and Florence Lake are plotted in Figure 6. Figure 6 shows good agreement between observed and simulated inflows.

TABLE 6 Error statistics of the dynamic regression model during calibration period 2010–2013

Lake	NSE	NSE′	PBIAS (%)	RSR	RMSE (cu ft/s)	MAE (cu ft/s)
Florence	0.97	0.82	3.79	0.18	115.91	60.09
Edison	0.97	0.82	3.54	0.17	52.49	30.88

TABLE 7 General performance rating for recommended statistics for a monthly time step (Motovilov et al., 1999)

Performance rating	RSR	NSE	|PBIAS|
Very good	$0.00 \leq RSE \leq 0.50$	$0.75 < NSE \leq 1.00$	$\|PBIAS\| \leq 10$
Good	$0.50 < RSE \leq 0.60$	$0.65 < NSE \leq 0.75$	$10 \leq \|PBIAS\| < 15$
Satisfactory	$0.60 < RSE \leq 0.70$	$0.50 < NSE \leq 0.65$	$15 \leq \|PBIAS\| < 25$
Unsatisfactory	$RSR > 0.70$	$NSE \leq 0.50$	$\|PBIAS\| \geq \pm 25$
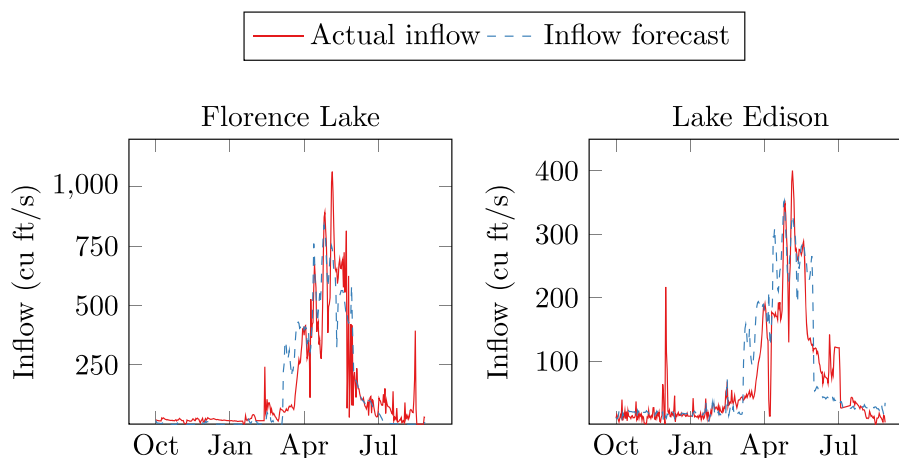
FIGURE 6 Comparison of actual inflow and inflow forecast with observed meteorological variables in water year 2014

Statistical indexes of the model performance during the test period are shown in Table 8.

The NSE of Florence Lake and Lake Thomas A. Edison are 0.78 and 0.72, respectively, which is considered to be good in Motovilov et al. (1999). Comparing NSE, PBIAS, and RSR value of the model with the general performance ratings recommendation in Moriasi et al. (2007), it can be concluded that, for daily step, the fit of the dynamic regression model during the test period is very good for both Florence Lake and Lake Thomas A. Edison. The modified NSE proposed by Garrick et al. (1978) is benchmark efficiency (BE) calculated using the daily mean value of the predictor variable for the calibration period as the baseline model. It is found to be 0.90–0.91 showing very high performance of the dynamic regression model compared with the baseline model. The NSE′ value also indicates good performance of the model. The small root mean square error in Table 8 also indicates that the dynamic regression model is capable of producing a reasonable forecast of inflows into the reservoirs. Moreover, the error in annual inflow forecast for Florence Lake and Lake Edison are only 0.15% and 10%, respectively.

For further comparison, we included a benchmark model, which is a multiple regression model containing the same explanatory variables as the dynamic regression model. The time lagged relationships were incorporated by including the time lagged variables as separate explanatory variables. Inflow of the previous day was also included as an explanatory variable. Statistics indexes of the multiple regression model performance are shown in Table 8. Another benchmark estimator would be a ratio estimator based on the same day inflow from last year, which can be written as Equation 11.

$$Y_t = Y_{t-365} \times \frac{\text{Peak SWE}_{\text{current year}}}{\text{Peak SWE}_{\text{previous year}}}. \tag{11}$$

The values of all the performance metrics indicate that the dynamic regression model consistently outperforms the benchmark multiple regression model and the ratio estimator. PBIAS values of regression model and the ratio estimator are not satisfactory for both Florence Lake and Lake Thomas A. Edison. Moreover, the dynamic regression model is more parsimonious compared with the multiple regression model. From the visual comparison of the observed and simulated inflow in Figure 6, low error in annual inflow forecast, value of the statistical indexes, and their comparison with regression and the ratio estimator, it can be concluded that the dynamic regression model is very good at forecasting reservoir inflows.

5.4.3 | Robustness analysis of the dynamic regression model

Global sensitivity analyis

A global sensitivity analysis (GSA) of the dynamic regression model was performed with the purpose of assessing robustness of the model and

TABLE 8 Forecast error statistics for water year 2014

Lake	Statistics	Dynamic regression	Regression	Ratio estimator
Florence	RMSE (cu ft/s)	100.75	114.33	136.62
	MAE (cu ft/s)	58.47	73.46	79.02
	NSE	0.78	0.57	0.72
	NSE$'$	0.62	0.53	0.53
	PBIAS (%)	0.16	31.47	−26.20
	RSR	0.46	0.65	0.53
Edison	RMSE (cu ft/s)	44.52	50.97	67.85
	MAE (cu ft/s)	27.84	38.12	38.31
	NSE	0.72	0.54	0.67
	NSE$'$	0.58	0.40	0.51
	PBIAS (%)	−8.85	44.32	−31.28
	RSR	0.53	0.68	0.57

simulation results. Input values of the model were perturbed within their reasonable range and subsequent changes in model output were studied (Baroni & Tarantola, 2014; Chu-Agor, Muñoz-Carpena, Kiker, Emanuelsson, & Linkov, 2011; Tomassini, Reichert, Knutti, Stocker, & Borsuk, 2007; Uusitalo, Lehikoinen, Helle, & Myrberg, 2015). Changes in input values automatically cause perturbations in parameter estimates of the dynamic regression model. Little change in resultant output values indicates robustness of the model to perturbations of inputs and parameter estimates and shows the uncertainty of the output variables to be relatively small (Uusitalo et al., 2015).

A qualitative GSA was performed in this study by visual inspection of model predictions. All input values were varied simultaneously within the entire allowable ranges of the input space and the effect on the output was studied (Baroni & Tarantola, 2014; Pianosi et al., 2016). This allowed GSA to evaluate the relative contributions of each input factor to the model output variable and account for effects of non-linear interactions between different inputs (Ciannelli, Chan, Bailey, & Stenseth, 2004; Baroni & Tarantola, 2014; Harper, Stella, & Fremier, 2011; Saltelli et al., 1999). Though local sensitivity analysis where inputs are varied one at a time is more common, it assumes linear relationship between inputs and outputs, making it only a perfunctory sensitivity analysis for most models (Saltelli et al., 2010). GSA does not assume any such specific relationship between inputs and model predictions and, therefore, is recommended for any kind of model (Makler-Pick, Gal, Gorfine, Hipsey, & Carmel, 2011; Rosolem, Gupta, Shuttleworth, Zeng, & de Gonçalves, 2012; Saltelli et al., 2010).

We followed the general probabilistic framework (GPF) based on Monte Carlo simulation for the global sensitivity analysis of deterministic models proposed by Baroni and Tarantola (2014). The flowchart for the GPF can be found in Baroni and Tarantola (2014). As is the norm, output in the sensitivity analysis does not refer to the entire range of temporal inflow variable produced by the model (Pianosi et al., 2016). Rather, it is measured as the variability induced in the model performance measure, RMSE of the test set. In the first step, all sources of uncertainty in the input meteorological variables, U_{xi}, were characterized. Because meteorological data were taken as the average of three weather stations—namely, KSP, UBC, and VLC, uncertainty may arise due to variability of meteorological variables between the calculated average and actual value at the location of Florence Lake and Lake Edison. Errors and approximations in input data measurement are other sources of uncertainty. Because meteorological variable measurements are not available at Florence Lake and Lake Edison, to account for the uncertainty, a grid box of 0.5 × 0.5° with centre at (37.32°N, −118.97°E) was considered. Two more weather stations, Huntington Lake (HNT) and Tamarack Summit (TMR), are located within the grid box along with the three existing weather stations. Average of the meteorological variables in at these five weather stations were calculated. In accordance with the comparison between data from the average of three weather stations and five weather stations, a random error was introduced to the daily nominal value of each variable. Measurement difference in the meteorological variables depend on type of water year and season. As such, unnaturally big variability can be introduced if random values are taken from the probability distribution of the difference time series. To preserve the temporal correlation of the meteorological variables, a random variable following uniform distribution in the interval [0, 1] was multiplied with $X_{5\,stations} − X_{3\,stations}$ where X denotes meteorological variables. The resultant random error was added to the meteorological variable data to produce the perturbed inputs. Each variable was physically constrained to avoid unrealistic values (e.g., negative value for precipitation and SWE).

The realization of each uncertainty was then associated with a scalar input factor $F_i = 1...128$ for $i = 1..3$. The three input factors were assumed to be independent. To minimize the number of model runs, a quasi-randomized, low-discrepancy sampling design called Sobol sequence was used to sample the three discrete uniform distribution according to the method present in Baroni and Tarantola (2014) and Saltelli et al., (2010). No correlations among the three input factors were considered in the sampling design. The simulations were run using a number of sampling points $N = 128$, which corresponds to a number of total number of model runs, $N_R = N(2 \times 3 + 2) = 1,024$. A combination of MATLAB and SAS codes were run to perform the sensitivity analysis.

Result and analysis

Figure 7 shows the probability distribution of the RMSE of forecasted inflow with perturbed inputs at Florence Lake and Lake Edison in water year 2014. Out of the 1024 model runs performed in the sensitivity analysis, the model that produces median RMSE is selected for analysis of results. Forecasts of the selected perturbed input model are plotted in Figure 8 along with actual inflows and forecasts of reference model
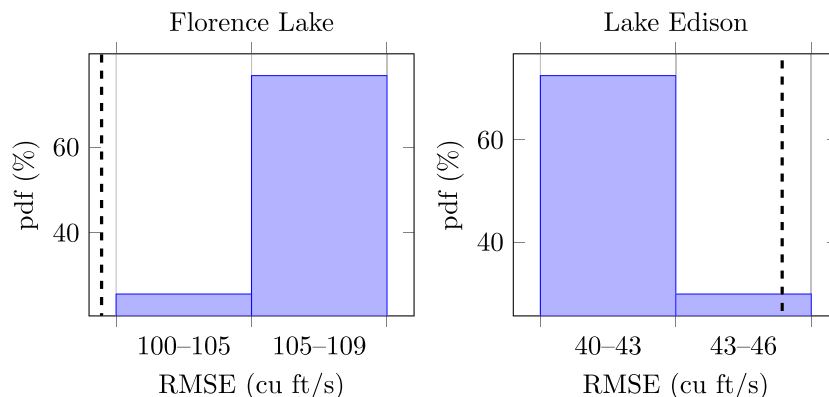
FIGURE 7 Probability distribution function (%) of the RMSE of forecasted inflow with perturbed inputs in water year 2014. RMSE of the reference model is indicated with the dashed line
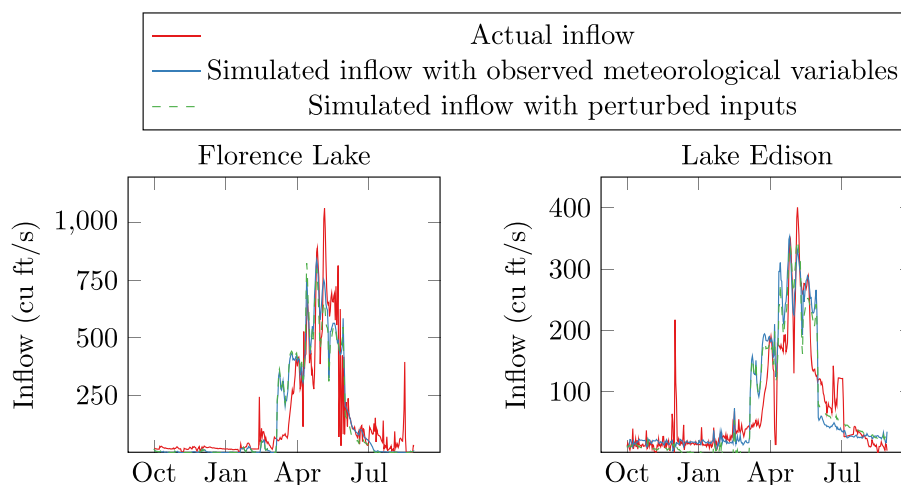

FIGURE 8 Comparison of actual inflow, simulated inflow from observed meteorological variables (reference), and simulated inflow from the selected perturbed input model with median RMSE in water year 2014

for both lakes in water year 2014. The annual inflows of the selected perturbed input model are compared with the observed and reference model inflow forecasts for water year 2014 in Table 9. Here, reference model is the inflow forecast model with unperturbed inputs. Annual inflow for the perturbed model has less than 5% error for both Florence Lake and Lake Edison. The RMSE results and forecasts show a general good performance of the model under perturbed inputs that shows the robustness of the model to perturbed inputs and parameter estimates.

Sensitivity of the dynamic regression model was also performed with respect to number of water years used. Two models were estimated using three and four water years in the training set, respectively, for both lakes. Water years 2013 and 2014 work as the corresponding test sets. Parameter estimates of both models are compared for changes in sign. No parameter estimate changes sign between these two simulations for both Florence Lake and Lake Edison. The annual inflow error is 15% and 5%, respectively, at Florence Lake and Lake Edison when three water years of data are used in the training set to forecast inflow of water year 2013. Therefore, it can be argued that the model is robust against increase in the amount of training data.

6 | QUANTIFYING THE IMPACT OF AEROSOLS ON RESERVOIR INFLOW

6.1 | Evaluation of the WRF-Chem simulation

We investigate the WRF-Chem model performance in our region of interest (the small box in Figure 2). WRF-Chem CTRL and WRF-Chem CLEAN simulation are WRF-Chem model with and without considering impact of aerosols, respectively. As shown in Figure 9, the meteorological variables simulated from the WRF-Chem model are reasonably close to the observed variables. The WRF-Chem model results are highly correlated with the observed meteorological variables for both water years. The correlation coefficients between the observed and WRF-Chem CTRL simulations for temperature and SWE range from 0.88 to 0.97 for 2013. The correlation coefficient for precipitation ranges from 0.66 to 0.73, which is adequate for our model because it has lesser impact on inflow as shown in Tables 4 and 5. Both visual inspection of Figure 9 and one-way ANOVA show that the WRF-Chem model underestimates temperature and precipitation

TABLE 9 Annual reservoir inflow for the selected perturbed input model with median RMSE and reference model in water year 2014

Lake	Actual (acre ft)	Reference model (acre ft)	Perturbed model (acre ft)
Florence	99,979	99,818	94,872
Edison	49,339	54,062	49,270
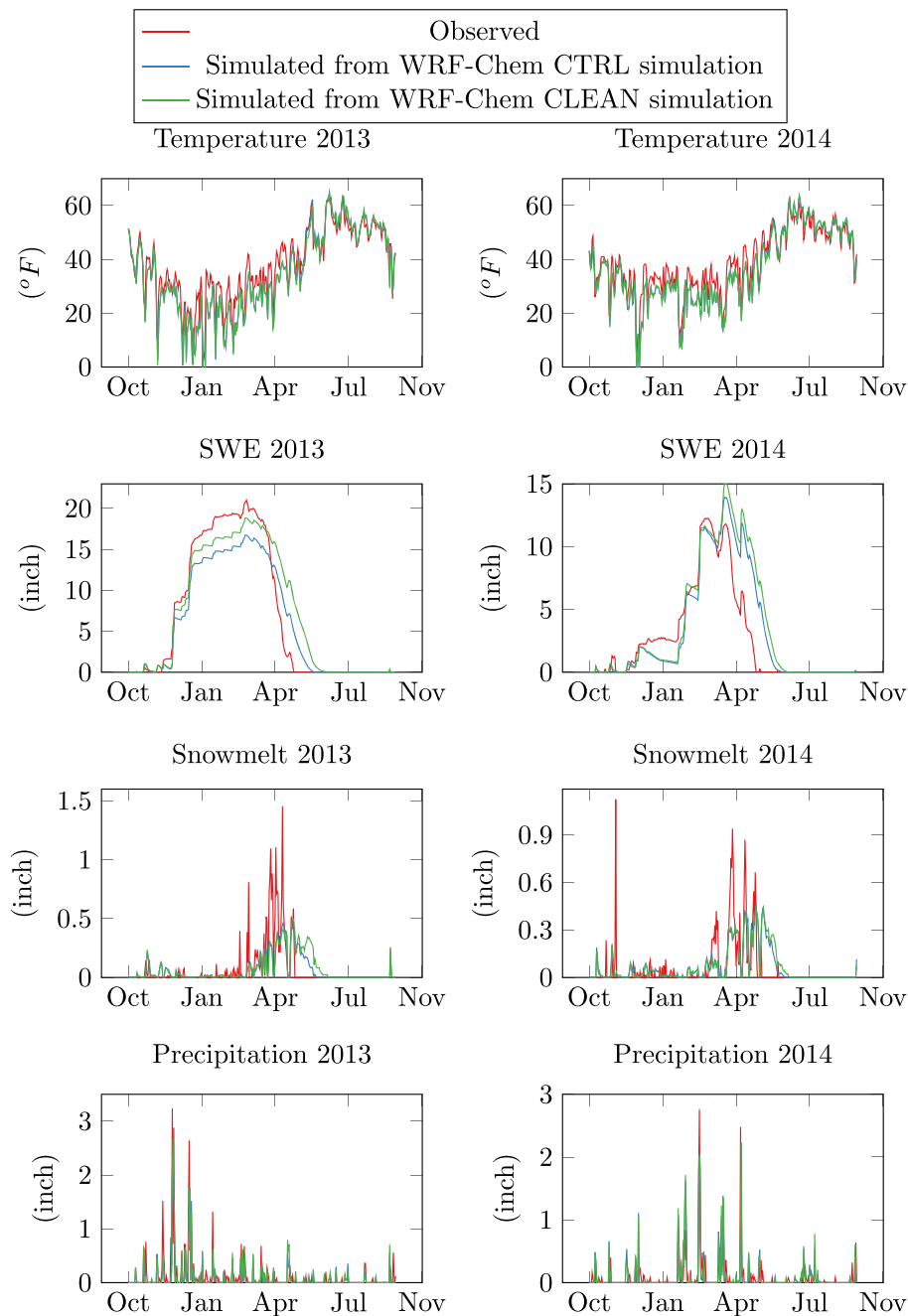
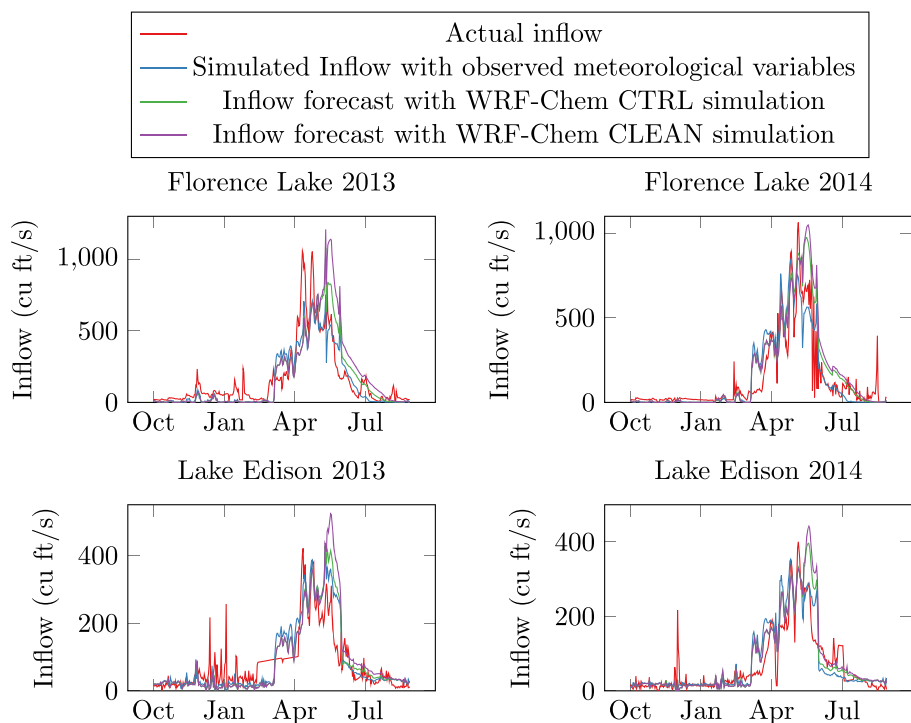**FIGURE 9** Comparison of observed and WRF-Chem CTRL and CLEAN simulated meteorological variables

in our region of interest. The model underestimates the SWE from December to March and overestimates the SWE from April to June.

The meteorological variables simulated from the WRF-Chem CTRL and WRF-Chem CLEAN models are highly correlated. The RMSE of the simulated meteorological variables with the observed variables are shown in Table 10. It can be observed from Table 10 that the meteorological variables from the WRF-Chem CTRL simulations are closer to the observed meteorological variables than those of the WRF-Chem

CLEAN simulations. In general, temperature from WRF-Chem CTRL simulations is higher than those of the CLEAN simulations because aerosol deposition increases impurity of snow (Wu et al., 2018). Precipitation and SWE from WRF-Chem CTRL simulations are lower than the CLEAN simulations. In order to understand how the aerosols affect these two variables, Wu et al. (2018)s examine the effects of ARI, ACI, and ASI separately and found that ACI plays a dominant role in increasing cloud water but reducing precipitation, leading to

TABLE 10 RMSE of WRF-Chem simulated meteorological variables with respect to observed variables

Year	Simulation	Temperature (°F)	SWE (inch)	Precipitation (inch)
2013	CTRL	4.86	2.10	0.20
	CLEAN	4.91	2.61	2.61
2014	CTRL	5.78	2.48	0.23
	CLEAN	5.93	2.85	0.25

FIGURE 10 Comparison of actual inflow, simulated inflow from observed meteorological variables, and WRF-Chem CTRL and CLEAN simulated meteorological variables

reduced SWE. Increase of nonprecipitating clouds can be explained by the fact that more CCN are available for the formation of clouds with smaller cloud droplets when more aerosols are present in the atmosphere. More detailed analysis on aerosol impacts on precipitation and snowpack in our region of interest can be found in Wu et al. (2018). Higher temperature, snow albedo effect, and feedback lead to higher snowmelt with aerosols in the late spring. However, snowmelt is lower with aerosols during early/late summer because of lower prior season's SWE and higher snowmelt in the late spring. From the correlation coefficient of WRF-Chem CTRL and CLEAN simulations, and ANOVA, the difference between the simulated SWE and precipitation in the CTRL and CLEAN simulations is larger in 2013 compared with 2014. In other words, the impact of aerosols on these meteorological variables is stronger in 2013 within the grid box.

6.2 | Quantification of the impact of aerosols on reservoir inflow

The impact of aerosols on reservoir inflow was quantified for two water years 2013 and 2014. In order to quantify the impact of aerosols on inflow, we ran dynamic regression model using the meteorological variables simulated from both the WRF-Chem CTRL and CLEAN models as inputs of the test data set. The actual inflow (red) is com-

pared with simulated inflow from observed meteorological variables (blue) and WRF-Chem simulated meteorological variables (green and purple) in Figure 10. The inflow simulated by the meteorological variables from the CTRL simulations match well with the actual inflow. The difference in inflow between the CTRL and CLEAN simulations ($Inflow_{CTRL}$-$Inflow_{CLEAN}$) represents the impact of aerosols, which is plotted in Figure 11. The observed inflow and meteorological data for 2010–2012 are used as the training set while simulating inflow for water year 2013.

It can be observed from Figure 11 that for the same year, the impact of aerosols on inflow is consistent in direction and similar in magnitude in both lakes. After simulating inflow for both water years under CTRL and CLEAN conditions, annual and seasonal inflows with and without considering the impact of aerosols were calculated. The impact of aerosols on inflow into reservoirs was then calculated by Equation 12.

$$\frac{Inflow_{w/\ Aerosols} - Inflow_{w/o\ Aerosols}}{Inflow_{w/o\ Aerosols}} \times 100\%. \quad (12)$$

The impact of aerosols on weekly inflow in percentages is plotted in Figure 12. Compared with Figure 11, the percentage change in inflows due to aerosols peaks in July–August whereas the difference in inflow magnitude between CTRL and CLEAN simulations peaks in June. This
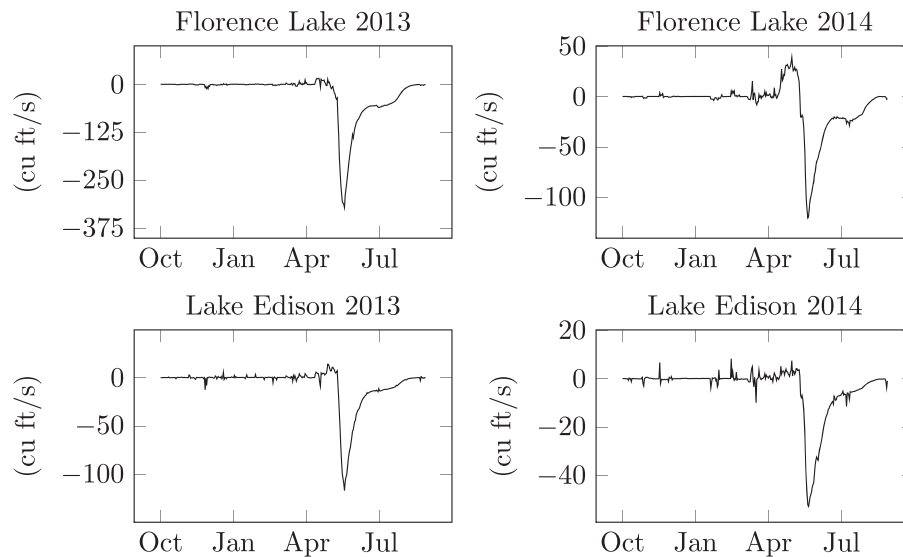
Florence Lake 2013

Florence Lake 2014

Lake Edison 2013

Lake Edison 2014

FIGURE 11 The impact of aerosols on reservoir inflow ($Inflow_{CTRL} - Inflow_{CLEAN}$)

Florence Lake 2013

Florence Lake 2014

Lake Edison 2013
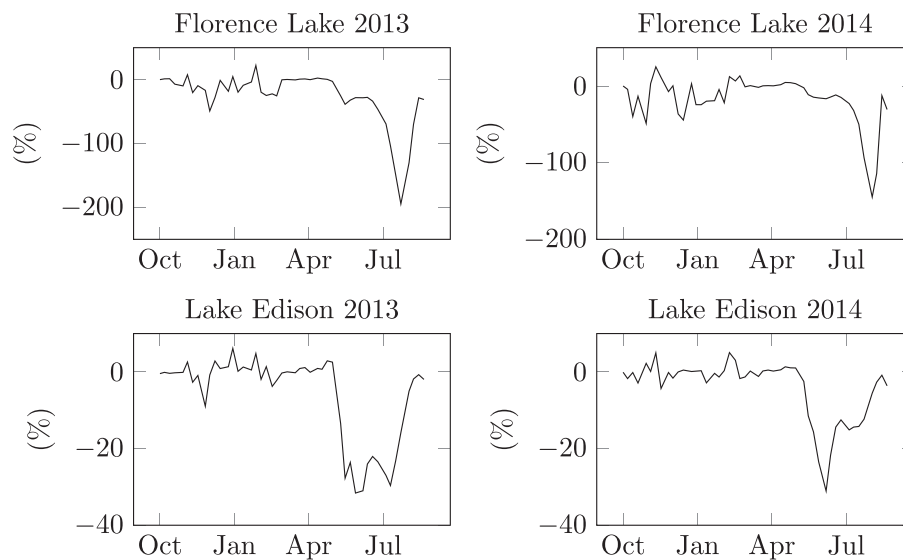
Lake Edison 2014

FIGURE 12 The impact of aerosols on weekly reservoir inflow (%)

is because the volume of inflow is considerably lower in July–August compared with June. The impact of aerosols on annual and seasonal inflow are tabulated in Table 11. For seasonal analysis, we first define the four seasons—fall is defined as the period of 10/01–12/21, winter is defined as 12/22–03/20, spring is defined as 03/21–05/31, and summer is defined as 06/01–09/30.

It can be observed from Figure 11 that the difference between WRF-Chem CTRL and CLEAN inflow is negligible during low inflow period. Inflow during fall and winter season is extremely low in both Florence Lake and Lake Thomas Alva Edison. During this period, the main contribution to reservoir inflows comes from surface run-off generated by rainfall. As winter is the wet season in this region, most of the rainfall occur during fall and winter. We also modelled the inflows during this period by temperature because temperature influences precipitation and snow accumulation. From Figure 9, it can be observed that precipitation forecast with and without considering impact of aerosols (WRF-Chem CTRL and WRF-Chem CLEAN) have up to 6% difference. Temperature forecast

from WRF-Chem CTRL and WRF-Chem CLEAN have up to 4% difference. Extremely low inflow combined with small difference in key meteorological variables, temperature, and precipitation leads to small difference between inflows with and without considering the impact of aerosols.

Impact of aerosols is pronounced during the high inflow period from May to June. In general, aerosols lead to slightly higher inflow in the late spring and significantly lower inflow during summer (11–26% reduction) as seen from Figure 11 and Table 11. These results can be explained by the seasonal variation of the impact of aerosols on the meteorological variables. During spring, the presence of aerosols leads to enhanced solar absorption by dust aerosol leading to higher temperature and snowmelt that translate into a higher inflow. On the other hand, aerosols lead to lower precipitation that results in a small reduction in the inflow. The aggregated effect of aerosols on inflow through temperature, snowmelt, and precipitation is slightly higher inflow in the spring. Lower prior season's SWE and lower current season's snowmelt together with lower precipitation result in lower

TABLE 11 Impact of aerosols (%) on annual and seasonal reservoir inflow

Lake	Year	Annual (%)	Fall (%)	Winter (%)	Spring (%)	Summer (%)
Florence	2013	−14	−11	−5	−.01	−26
	2014	−4	−2	−1	2	−11
Edison	2013	−8	−1	0.2	1	−18
	2014	−5	−0.4	0.3	0.6	−11

TABLE 12 Annual reservoir inflow under different aerosol conditions

Lake	Year	Actual (acre ft)	CTRL (acre ft)	CLEAN (acre ft)
Florence	2013	117,390	113,610	13,163
	2014	99,980	119,740	125,410
Edison	2013	58,572	61,004	66,240
	2014	49,339	55,446	58,247

inflow in summer. Low inflow in summer due to impact of aerosol creates the sudden dip in Figure 11 from May to June. This is consistent with the observation in Wu et al. (2018) that over mountaintops in the Sierra Nevada region, surface run-off slightly increases in spring and decreases after April. It is helpful to mention again that Florence Lake and Lake Thomas A. Edison are higher elevation lakes which generate inflow by capturing run-off from the San Joaquin River. The presence of aerosols suppresses precipitation which leads to lower inflow for the Florence Lake during fall and winter. In the Lake Edison, inflow in fall and winter is simulated using precipitation and temperature. Aerosols lead to lower precipitation and higher temperature that translate into lower inflow in fall and slightly higher inflow in winter. The overall effect of aerosols is a reduction in annual inflow by 4–14% for both lakes as shown in Tables 11 and 12. Wu et al. (2018) observed a 10% decrease in surface run-off from October to June in the mountaintops of the Sierra Nevada region due to the impact of aerosols that agrees with our calculated annual impact on reservoir inflow in the region.

The impact of aerosols is more significant in water year 2013 than in water year 2014 for both lakes. This is because the impact of aerosols on the meteorological variables is more pronounced in water year 2013 as seen from the mean of the meteorological variables from CTRL and CLEAN simulations. The annual impact of aerosols is stronger in Florence Lake. This is because the fall and winter inflow are simulated using only precipitation for Florence Lake. For Lake Edison, the fall and winter inflow is simulated using both precipitation and temperature. The higher temperature effect from aerosols offsets some of the reduction in inflow in Lake Edison. Therefore, the annual impact of aerosols on inflow is lower in Lake Edison.

6.3 | Robustness analysis of the estimation of impact of aerosols

It can be observed that the difference between CTRL and CLEAN inflows is between 4% and 15%. On the other hand, the difference between annual observed and simulated inflow varies between 0.1–17% at Lake Florence and 7–9% at Lake Thomas A. Edison, which is in the same range as the impact of aerosols on inflows. However, in Subsubsection 5.4.3, it was shown that the dynamic regression model built is robust against perturbations of input variables and number of water years used in the training set.

Moreover, the most significant impact of aerosols occurs in the late spring and summer when the WRF-Chem simulations of SWE have a large error. Therefore, forecasts of inflow in water years 2013 and 2014 and the impact of aerosol on inflow were calculated with perturbed test set inputs to gain more confidence in the impact of aerosol results. To account for the discrepancy between the observed meteorological variables and WRF-Chem CTRL simulation outputs, the difference between these two time series was calculated for all of the meteorological variables. A random error was introduced to the daily values of the meteorological variables of the WRF-Chem simulated test set. To preserve the temporal correlation of the meteorological variables, a random variable following uniform distribution in the interval $[0, 1]$ was multiplied with $X_{obs} - X_{WRF-Chem\ CTRL}$ where X denotes meteorological variables. The resultant random error was added to the meteorological variables of the test set to generate the perturbed inputs. Both CTRL and CLEAN simulations were perturbed by the same error. Input meteorological variables of the training set are not perturbed. The dynamic regression model was then simulated with the perturbed WRF-Chem CTRL and CLEAN test sets for Florence Lake for both water years; 1,024 model runs were performed in the same fashion as Section 5.4.3. Out of the 1,024 models, the model that produced the median RMSE was selected for analysis of results. The difference in inflow between the CTRL and CLEAN simulations was then calculated and plotted in Figure 13 for the selected model.

It can be observed that, for all of the simulations, the sign and magnitude of the difference in inflow with the selected perturbed input model are similar to the reference model. Here, the reference model denotes the unperturbed input model. The annual and seasonal impacts of aerosol are compared with the reference model in Table 13, which shows that they are similar. Therefore, it can be safely argued that the difference between the inflows arises from the difference between meteorological variables with and without impact of aerosol.

7 | SUMMARY AND CONCLUSION

A comprehensive framework to quantify the impact of aerosols on reservoir inflow was developed by synergistically combining the
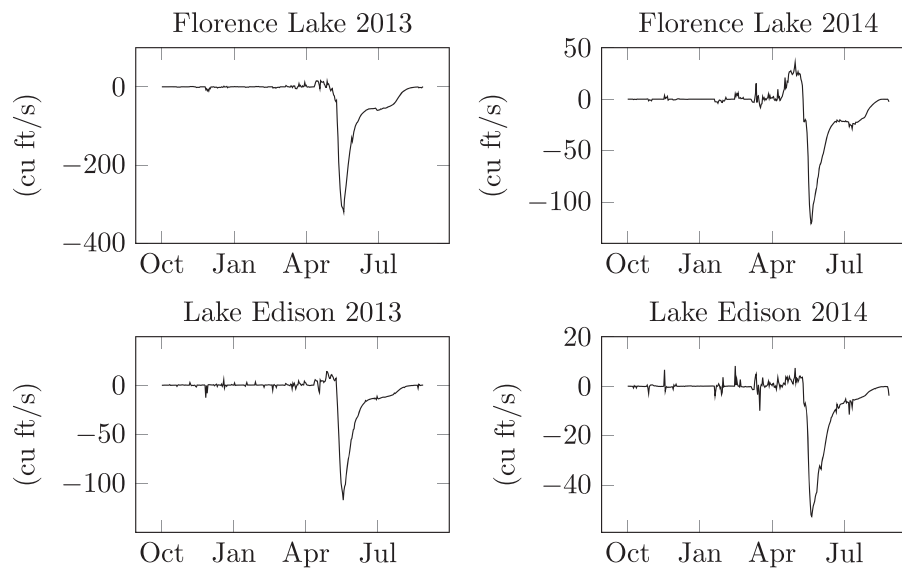
FIGURE 13 The impact of aerosols on reservoir inflow ($Inflow_{CTRL} - Inflow_{CLEAN}$) for the selected perturbed input model with median RMSE

TABLE 13 Comparison of impact of aerosols on annual and seasonal reservoir inflow for the reference and selected perturbed input model with median RMSE

Lake	Year	Model	Annual (%)	Fall (%)	Winter (%)	Spring (%)	Summer (%)
Florence	2013	Reference	−14	−11	−6	−.07	−26
		Perturbed	−13	−10	−5	0.05	−23
	2014	Reference	−4	−2	−1	2	−11
		Perturbed	−4	−0.10	−0.09	2	−10
Edison	2013	Reference	−8	−1	0.2	1	−18
		Perturbed	−8	−1	0.1	1	−17
	2014	Reference	−5	−0.4	0.3	0.6	−11
		Perturbed	−5	−0.2	0.4	0.7	−104

WRF-Chem model and a dynamic regression model. The dynamic regression model can also be leveraged to perform 1-year ahead daily inflow forecast. A case study was conducted using Florence Lake and Lake Thomas Alva Edison of the Big Creek Hydroelectric Project. The dynamic regression model was found to be adequate and performed well compared with the benchmark models.

We investigated the impact of aerosols on the inflow into these hydropower reservoirs over two water years. Aerosols exert influence on reservoir inflows through their influence on meteorological variables. The simulation results show that the presence of aerosols significantly reduces the annual inflow into the hydropower reservoirs of the Big Creek Hydroelectric Project. The impact of aerosols on inflows is pronounced during the high inflow season (spring and summer), whereas it is negligible during the low inflow seasons (fall and winter). Extremely low inflow combined with small difference in key meteorological variables (temperature and precipitation) leads to small difference between inflows with and without considering the impact of aerosols during the low inflow period. Seasonal variation of the aerosol's impact on meteorological variables leads to the seasonal variation of the aerosol's impact on reservoir inflows. Aerosols significantly reduce the amount of inflow in the summer (11–26% reduction) as a result of a reduction in precipitation, snow water equivalent, and snowmelt due to presence of aerosols. Aerosols slightly

increase the inflow in the spring (up to 2%) as a result of an increase in temperature and snowmelt due to presence of aerosols.

The marginal value of water is high in summer, and the run-off risk is high during spring. Therefore, it can be concluded that the presence of aerosol is detrimental to the optimal utilization of hydroelectric power systems. The change in inflow due to impact of aerosol in different seasons with different water and electricity demands can assist the reservoir operators in determining the optimal operation policy for the reservoirs. Further scarcity of reservoir inflow during dry seasons can motivate the San Joaquin River region water resources planners to focus their efforts on mitigation strategies. The findings from this research can provide another justification for stricter environmental regulations to reduce anthropogenic aerosol emissions.

There are a few limitations to our study. The limited historical reservoir inflow data prevented us from capturing the long-term trend in reservoir inflow and evaluating the long-term impact of aerosols on the reservoir inflow. The daily temperature, snow water equivalent, and precipitation records for 2010–2014 were not available at the weather stations located at Florence Lake and Lake Thomas Alva Edison. These additional observations could have improved the quality of inflow forecast. The quantification of the impact of aerosols on inflows into reservoirs strongly depends on the accurate estimation of the difference in meteorological variables between the WRF-Chem

CTRL and CLEAN simulations. WRF-Chem model run in a higher resolution in the grid box could deliver more accurate results. At last, a rigorous outlier detection algorithm can be applied on the inflow data, which can further improve the forecast performance of the dynamic regression model.

Forecast of the inflow into the hydropower reservoirs obtained from this study can assist in optimizing the cascaded hydropower system. The framework for evaluating the impact of aerosols on reservoir inflow is easily extendable to reservoirs located in other regions. The time series of the meteorological variables and reservoir inflows need for dynamic regression model fitting are usually available from the nearby weather stations. WRF-Chem simulations of the relevant meteorological variables with and without aerosol can be performed in other regions with similar experimental set-up. With these data available, the impact of aerosols on inflow into the reservoirs located in other regions can be evaluated in a similar fashion. In the future, we plan to integrate the year ahead inflow forecast of Florence Lake and Lake Edison into the long-term scheduling of the Big Creek Hydroelectric Project. The impact of aerosols on hydroelectric generation and economic value will be assessed. Future studies will also address the drawbacks of the study. Reservoir inflow data from 2015 onwards will be available, and they will be used to forecast reservoir inflow for water year 2017 onwards and quantify the impact of aerosols on reservoir inflow.

ORCID

Farzana Kabir iD http://orcid.org/0000-0003-0518-0672

REFERENCES

Ackerman, A. S., Kirkpatrick, M. P., Stevens, D. E., & Toon, O. B. (2004). The impact of humidity above stratiform clouds on indirect aerosol climate forcing. *Nature*, 432(7020), 1014–1017.

Albrecht, B. A. (1989). Aerosols, cloud microphysics, and fractional cloudiness. *Science*, 245(4923), 1227–1231.

Andreae, M. O., Jones, C. D., & Cox, P. M. (2005). Strong present-day aerosol cooling implies a hot future. *Nature*, 435(7046), 1187–1190.

Barnard, J. C., Fast, J. D., Paredes-Miranda, G., Arnott, W. P., & Laskin, A. (2010). Technical note: Evaluation of the WRF-Chem "Aerosol chemical to aerosol optical properties" Module using data from the MILA-GRO campaign. *Atmospheric Chemistry and Physics*, 10(15), 7325–7340. https://doi.org/10.5194/acp-10-7325-2010

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438(7066), 303–309. https://doi.org/10.1038/nature04141

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438(7066), 303–309.

Baroni, G., & Tarantola, S. (2014). A general probabilistic framework for uncertainty and global sensitivity analysis of deterministic models: A hydrological case study. *Environmental Modelling & Software*, 51(Supplement C), 26–34. https://doi.org/10.1016/j.envsoft2013.09.022

Barth, M., McFadden, J. P., Sun, J., Wiedinmyer, C., Chuang, P., Collins, D., … Stroud, Craig (2005). Coupling between land ecosystems and the atmospheric hydrologic cycle through biogenic aerosol pathways. *Bulletin of the American Meteorological Society*, 86(12), 1738–1742.

Bernard, E., Dan, S., Mike, W., Jeff, A., & Mazdak, A. (2007). A hydrologic/water quality model application. *JAWRA Journal of the AmericanWater Resources Association*, 43(5), 1223–1236. https://doi.org/10.111/j.1752-1688.2007.00105.x

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, 320(1), 18–36. Retrieved from https://www.sciencedirect.com/science/article/pii/S002216940500332X https://doi.org/10.1016/j.jhydrol.2005.07.007

Borys, R. D., Lowenthal, D. H., Cohn, S. A., & Brown, W. O. J. (2003). Mountaintop and radar measurements of anthropogenic aerosol effects on snow growth and snowfall rate. *Geophysical Research Letters*, 30(10), 1538.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control*. Hoboken, New Jersey: John Wiley. (Google-Books-ID: lCy9BgAAQBAJ).

Brandt, R. E., Warren, S. G., & Clarke, A. D. (2011). A controlled snowmaking experiment testing the relation between black carbon content and reduction of snow albedo. *Journal of Geophysical Research: Atmospheres*, 116(D8), D08109.

Brekke, L. D., Miller, N. L., Bashford, K. E., Quinn, N. W. T., & Dracup, J. A. (2004). Climate change impacts uncertainty for water resources in the San Joaquin River Basin, California. *Journal of the American Water Resources Association*, 40(1), 149–164. https://doi.org/10.1111/j.1752-1688.2004.tb01016.x

Cayan, D. R., Riddle, L. G., & Aguado, E. (1993). The influence of precipitation and temperature on seasonal streamflow in California. *Water Resources Research*, 29(4), 1127–1140. https://doi.org/10.1029/92WR028002

Chapman, E. G., Gustafson Jr., W. I., Easter, R. C., Barnard, J. C., Ghan, S. J., Pekour, M. S., & Fast, J. D. (2009). Coupling aerosol-cloud-radiative processes in the WRF-Chem model: Investigating the radiative impact of elevated point sources. *Atmospheric Chemistry and Physics*, 9(3), 945–964. https://doi.org/10.5194/acp-9-945-2009

Charlson, R. J., Langner, J., Rodhe, H., Leovy, C. B., & Warren, S. G. (1991). Perturbation of the northern hemisphere radiative balance by backscattering from anthropogenic sulfate aerosols*. *Tellus A*, 43(4), 152–163. https://doi.org/10.1034/j.1600-0870.1991.00013.x

Charlson, R. J., & Schwartz, S. E. (1992). Climate forcing by anthropogenic aerosols. *Science*, 255(5043), 423–430.

Charlson, R. J., Schwartz, S. E., Hales, J. M., Cess, R. D., Coakley, J. A., Hansen, J. E., & Hofmann, D. J. (1992). Climate forcing by anthropogenic aerosols. *Science*, 255(5043), 423–430.

Chu-Agor, M. L., Muñoz-Carpena, R., Kiker, G., Emanuelsson, A., & Linkov, I. (2011). Exploring vulnerability of coastal habitats to sea level rise through global sensitivity and uncertainty analyses. *Environmental Modelling & Software*, 26(5), 593–604. https://doi.org/10.1016/jenvsoft.2010.12.003

Chýlek, P., Ramaswamy, V., & Srivastava, V. (1983). Albedo of soot-contaminated snow. *Journal of Geophysical Research: Oceans*, 88(C15), 10837–10843.

Ciannelli, L., Chan, K.-S., Bailey, K. M., & Stenseth, N. C. (2004). Non-additive effects of the environment on the survival of a large marine fish population. *Ecology*, 85(12), 3418–3427. https://doi.org/10.1890/03-0755

Clarke, A. D., & Noone, K. J. (1985). Soot in the Arctic snowpack: A cause for perturbations in radiative transfer. *Atmospheric Environment (1967)*, 19(12), 2045–2053.

Coulibaly, P., Anctil, F., & Bobée, B. (2000). Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *Journal of Hydrology*, 230, 244–257. https//doi.org/10.1016/S0022-1694(00)00214-6

Criss, R. E., & Winston, W. E. (2008). Do Nash values have value? Discussion and alternate proposals. *Hydrological Processes*, 22(14), 2723–2725. https//doi.org/10.1002/hyp.7072

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., ... Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. https//doi.org/10.1002/qj.828

Deems, J. S., Painter, T. H., Barsugli, J. J., Belnap, J., & Udall, B. (2013). Combined impacts of current and future dust deposition and regional warming on Colorado River Basin snow dynamics and hydrology. *Hydrology and Earth System Sciences*, 17(11), 4401–4413. https//doi.org/10.5194/hess-17-4401-2013

Department of Water Resources, C. (2017). Reservoir inflow at Florence Lake and Lake Thomas A. Edison in California, meteorological variables at weather stations KSP, VLC, UBC, TMR and HTT for water year 2010-14 and WRF-Chem Simulations of meteorological variables at Big Creek Hydroelectric Project for water year 2013 and 2014. https://doi.org/10.6084/m9.figshare.5705638.v1

Doherty, S. J., Warren, S. G., Grenfell, T. C., Clarke, A. D., & Brandt, R. E. (2010). Light-absorbing impurities in Arctic snow. *Atmospheric Chemistry and Physics*, 10(23), 11647–11680.

Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J.-F., Pfister, G. G., Fillmore, D., ... Kloster, S. (2010). Description and evaluation of the model for ozone and related chemical tracers, version 4 (MOZART-4). *Geoscientific Model Development*, 3(1), 43–67. https//doi.org/10.5194/gmd-3-43-2010

Environmental Sciences Division, O. R. N. L. (2013). Global fire emissions database, version 2 (gfedv2.1). ORNL Distributed Active Archive Center https//doi.org/10.3334/ornldaac/849

Fast, J. D., Allan, J., Bahreini, R., Craven, J., Emmons, L., Ferrare, R., ... Zhang, Q. (2014). Modeling regional aerosol and aerosol precursor variability over California and its sensitivity to emissions and long-range transport during the 2010 CalNex and CARES campaigns. *Atmospheric Chemistry and Physics*, 14(18), 10013–10060. https//doi.org/10.5194/acp-14-10013-2014

Fast, J. D., Gustafson, W. I., Easter, R. C., Zaveri, R. A., Barnard, J. C., Chapman, E. G., ... Peckham, S. E. (2006). Evolution of ozone, particulates, and aerosol direct radiative forcing in the vicinity of Houston using a fully coupled meteorology-chemistry-aerosol model. *Journal of Geophysical Research: Atmospheres*, 111(D21), D21305. https//doi.org/10.1029/2005JD006721

Fast, J. D., Gustafson Jr., W. I., Berg, L. K., Shaw, W. J., Pekour, M., Shrivastava, M., ... Zaveri, R. A. (2012). Transport and mixing patterns over central California during the carbonaceous aerosol and radiative effects study (CARES). *Atmospheric Chemistry and Physics*, 12(4), 1759–1783. https//doi.org/10.5194/acp-12-1759-2012

Flanner, M. G., & Zender, C. S. (2005). Snowpack radiative heating: Influence on Tibetan Plateau climate. *Geophysical Research Letters*, 32(6), L06501. https//doi.org/10.1029/2004GL022076

Flanner, M. G., & Zender, C. S. (2006). Linking snowpack microphysics and albedo evolution. *Journal of Geophysical Research: Atmospheres*, 111(D12). https//doi.org/10.1029/2005JD006834

Flanner, M. G., Zender, C. S., Randerson, J. T, & Rasch, P. J (2007). Present-day climate forcing and response from black carbon in snow. *Journal of Geophysical Research: Atmospheres*, 112(D11), D11202.

Galeati, G. (1990). A comparison of parametric and non-parametric methods for runoff forecasting. *Hydrological Sciences Journal*, 35(1), 79–94.

Garrick, M., Cunnane, C., & Nash, J. E. (1978). A criterion of efficiency for rainfall-runoff models. *Journal of Hydrology*, 36(3-4), 375–381. https//doi.org/10.1016/0022-1694(78)90155-5

Ghan, S., Laulainen, N., Easter, R., Wagener, R., Nemesure, S., Chapman, E., ... Leung, R. (2001). Evaluation of aerosol direct radiative forcing in MIRAGE. *Journal of Geophysical Research: Atmospheres*, 106(D6), 5295–5316. https//doi.org/10.1029/2000JD900502

Givati, A., & Rosenfeld, D. (2004). Quantifying precipitation suppression due to air pollution. *Journal of Applied Meteorology*, 43(7), 1038–1056. https//doi.org/10.1175/1520-0450(2004)043<1038:QPSDTA>2.0.CO;2

Givati, A., & Rosenfeld, D. (2005). Separation between cloud-seeding and air-pollution effects. *Journal of Applied Meteorology*, 44(9), 1298–1314. https//doi.org/10.1175/JAM2276.1

Givati, A., & Rosenfeld, D. (2007). Possible impacts of anthropogenic aerosols on water resources of the Jordan River and the Sea of Galilee. *Water Resources Research*, 43(10), W10419. https//doi.org/10.1029/2006WR005771

Gleick, P. H., & Chalecki, E. L. (1999). The impacts of climatic changes for water resources of the Colorado and Sacramento–San Joaquin River Basins 1. *Journal of the American Water Resources Association*, 35(6), 1429–1441 en. https//doi.org/10.1111/j.1752-1688.1999.tb04227.x

Gong, S. L. (2003). A parameterization of sea-salt aerosol source function for sub- and super-micron particles. *Global Biogeochemical Cycles*, 17(4), 1097. https//doi.org/10.1029/2003GB002079

Gragne, A. S., Sharma, A., Mehrotra, R., & Alfredsen, K. (2015). Improving real-time inflow forecasting into hydropower reservoirs through a complementary modelling framework. *Hydrology and Earth System Sciences*, 19(8), 3695–3714. https//doi.org/10.5194/hess-19-3695-2015

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., & Eder, B. (2005). Fully coupled "online" chemistry within the WRF model. *Atmospheric Environment*, 39(37), 6957–6975. https//doi.org/10.1016/j.atmosenv.2005.04.027

Grenfell, T. C., Light, B., & Sturm, M. (2002). Spatial distribution and radiative effects of soot in the snow and sea ice during the sheba experiment. *Journal of Geophysical Research: Oceans*, 107(C10).

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., & Geron, C. (2006). Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmospheric Chemistry and Physics*, 6(11), 3181–3210. https//doi.org/10.5194/acp-6-3181-2006

Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1999). Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *Journal of Hydrologic Engineering*, 4(2), 135–143. https//doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135)

Hadley, O. L., & Kirchstetter, T. W. (2012). Black-carbon reduction of snow albedo. *Nature Climate Change*, 2(6), 437–440.

Hansen, J., & Nazarenko, L. (2004). Soot climate forcing via snow and ice albedos. *Proceedings of the National Academy of Sciences of the United States of America*, 101(2), 423–428.

Hansen, J., Sato, M., & Ruedy, R. (1997). Radiative forcing and climate response. *Journal of Geophysical Research: Atmospheres*, 102(D6), 6831–6864.

Hansson, H.-C., & Bhend, J. (2015). Causes of regional change-aerosols, *Second assessment of climate change for the Baltic Sea Basin*. Cham: Springer, pp. 441–452. https://link.springer.com/chapter/10.1007/978-3-319-16006-1_24 https//doi.org/10.1007/978-3-319-16006-1_24

Harper, E. B., Stella, J. C., & Fremier, A. K. (2011). Global sensitivity analysis for complex ecological models: A case study of riparian cottonwood population dynamics. *Ecological Applications: A Publication of the Ecological Society of America*, 21(4), 1225–1240.

Haywood, J., & Boucher, O. (2000). Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: A review. *Reviews of Geophysics*, 38(4), 513–543.

He, C., Liou, K.-N., & Takano, Y. (2018). Resolving size distribution of black carbon internally mixed with snow: Impact on snow optical properties and albedo. *Geophysical Research Letters*, 45(6), 2697–2705. https//doi.org/10.1002/2018GL077062

He, C., Takano, Y., Liou, K.-N., Yang, P., Li, Q., & Chen, F. (2017). Impact of snow grain shape and black carbon-snow internal mixing on snow optical properties: Parameterizations for climate models. *Journal of Climate*, *30*(24), 10019–10036. https//doi.org/10.1175/JCLI-D-17-0300.1

Hong, S.-Y., Noh, Y., & Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, *134*(9), 2318–2341. https//doi.org/10.1175/MWR3199.1

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research: Atmospheres*, *113*(D13), D13103. https//doi.org/10.1029/2008JD009944

Jacobson, M. Z. (2001). Strong radiative heating due to the mixing state of black carbon in atmospheric aerosols. *Nature*, *409*(6821), 695–697.

Jacobson, M. Z. (2004). Climate response of fossil fuel and biofuel soot, accounting for soot's feedback to snow and sea ice albedo and emissivity. *Journal of Geophysical Research: Atmospheres*, *109*(D21), D21201.

Jaeglé, L., Quinn, P. K., Bates, T. S., Alexander, B., & Lin, J.-T. (2011). Global distribution of sea salt aerosols: New constraints from in situ and remote sensing observations. *Atmospheric Chemistry and Physics*, *11*(7), 3137–3157. https//doi.org/10.5194/acp-11-3137-2011

Jain, S. K., Das, A., & Srivastava, D. K. (1999). Application of ANN for reservoir inflow prediction and operation. *Journal of Water Resources Planning and Management*, *125*(5), 263–271.

Johnson, B. T., Shine, K. P., & Forster, P. M. (2004). The semi-direct aerosol effect: Impact of absorbing aerosols on marine stratocumulus. *Quarterly Journal of the Royal Meteorological Society*, *130*(599), 1407–1422.

Jones, A., Roberts, D. L., & Slingo, A. (1994). A climate model study of indirect radiative forcing. *Nature*, *370*, 450–453.

Kaufman, Y. J., Koren, I., Remer, L. A., Rosenfeld, D., & Rudich, Y. (2005). The effect of smoke, dust, and pollution aerosol on shallow cloud development over the atlantic ocean. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(32), 11207–11212.

Kaufman, Y. J., Tanré, D., & Boucher, O. (2002). A satellite view of aerosols in the climate system. *Nature*, *419*(6903), 215–223.

Kiehl, J. T., & Briegleb, B. P. (1993). The relative roles of sulfate aerosols and greenhouse gases in climate forcing. *Science*, *260*(5106), 311–314.

Kılınç, I., & Cigizoglu, K. (2005). Reservoir management using artificial neural networks. Proc. 14th. Reg. Director. DSI (State Hydraulic Works), Istanbul, Turkey.

Knowles, N., & Cayan, D. R. (2002). Potential effects of global warming on the Sacramento/San Joaquin watershed and the San Francisco estuary. *Geophysical Research Letters*, *29*(18), 1891. https//doi.org/10.1029/2001GL014339

Kokhanovsky, A. (2013). Spectral reflectance of solar light from dirty snow: A simple theoretical model and its validation. *The Cryosphere*, *7*(4), 1325–1331. https//doi.org/10.5194/tc-7-1325-2013

Krause, P., Boyle, D. P., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, *5*, 89–97. https//doi.org/10.5194/adgeo-5-89-2005

Lall, U., & Bosworth, K. (1994). Multivariate kernel estimation of functions of space and time hydrologic data. In Hipel, K. W., McLeod, A. I., Panu, U. S., & Singh, V. P. (Eds.), *Stochastic and Statistical Methods in Hydrology and Environmental Engineering*. Netherlands: Springer, pp. 301–315. https//doi.org/10.1007/978-94-017-3083-9_22

Lau, W. K., Kim, M.-K., Kim, K.-M., & Lee, W.-S. (2010). Enhanced surface warming and accelerated snow melt in the Himalayas and Tibetan plateau induced by absorbing aerosols. *Environmental Research Letters*, *5*(2), 025204.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., … Slater, A. G (2011). Parameterization improvements and functional and structural advances in version 4 of the community land model. *Journal of Advances in Modeling Earth Systems*, *3*(1), M03001–27. https//doi.org/10.1029/2011MS00045

Lee-Taylor, J., & Madronich, S. (2002). Calculation of actinic fluxes with a coupled atmosphere–snow radiative transfer model. *Journal of Geophysical Research: Atmospheres*, *107*(D24), 4796.

Legates, D. R., & McCabe Jr., G. J. (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, *35*(1), 233–241. https//doi.org/10.1029/1998WR900018

Legates, D. R., & McCabe, G. J. (2012). A refined index of model performance: A rejoinder. *International Journal of Climatology*, *33*(4), 1053–1056. https//doi.org/10.1002/joc.2419

Lettenmaier, D. P., & Gan, T. Y. (1990). Hydrologic sensitivities of the Sacramento-San Joaquin River Basin, California, to global warming. *Water Resources Research*, *26*(1), 69–86. https//doi.org/10.1029/WR026i001p00069

Li, Z., Luo, C., Jiang, K., Wan, R., & Li, H. (2017). Comprehensive performance evaluation for hydrological and nutrients simulation using the hydrological simulation program-Fortran in a mesoscale monsoon watershed, China. *International Journal of Environmental Research and Public Health*, *14*(12), E1599. https//doi.org/10.3390/ijerph14121599

Licciardello, F., Zema, D. A., Zimbone, S. M., & Bingner, R. L. (2007). Runoff and soil erosion evaluation by the AnnAGNPS model in a small mediterranean watershed. *Transactions of the ASABE*, *50*(5), 1585–1593. https//doi.org/10.13031/2013.23972

Liou, K. N., Takano, Y., He, C., Yang, P., Leung, L. R., Gu, Y., & Lee, W. L. (2014). Stochastic parameterization for light absorption by internally mixed BC/dust in snow grains for application to climate models. *Journal of Geophysical Research: Atmospheres*, *119*(12), 7616–7632. https://doi.org/10.1002/2014JD021665

Lohmann, U., & Feichter, J. (2005). Global indirect aerosol effects: A review. *Atmospheric Chemistry and Physics*, *5*(3), 715–737. https//doi.org/10.5194/acp-5-715-2005

Lohmann, U. (2005). Impact of aerosols on the hydrological cycle. *Water and the Environment*, *12*, 104–219.

Lohmann, U. (2002). A glaciation indirect aerosol effect caused by soot aerosols. *Geophysical Research Letters*, *29*(4), 11–1. https//doi.org/10.1029/2001GL014357

Madsen, H., Richaud, B., & Pedersen, C. B. (2009). A real-time inflow forecasting and reservoir optimization system for optimizing hydropower production. *Waterpower XVI*.

Makler-Pick, V., Gal, G., Gorfine, M., Hipsey, M. R., & Carmel, Y. (2011). Sensitivity analysis for complex ecological models—A new approach. *Environmental Modelling & Software*, *26*(2), 124–134. https//doi.org/10.1016/j.envsoft.2010.06.010

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (2008). *Forecasting methods and applications*. India: John Wiley.

Marks, A. A., & King, M. D. (2013). The effects of additional black carbon on the albedo of Arctic sea ice: Variation with sea ice type and snow cover. *The Cryosphere*, *7*(4), 1193.

Marks, A. A., & King, M. D. (2014). The effect of snow/sea ice type on the response of albedo and light penetration depth (e-folding depth) to increasing black carbon. *The Cryosphere*, *8*(5), 1625.

Matt, F. N., Burkhart, J. F., & Pietikäinen, J.-P. (2018). Modelling hydrologic impacts of light absorbing aerosol deposition on snow at the catchment scale. *Hydrology and Earth System Sciences*, *22*(1), 179–201. https://doi.org/10.5194/hess-22-179-2018

McCuen, R. H., Knight, Z., & Cutter, A. Gillian (2006). Evaluation of the Nash-Sutcliffe efficiency index. *Journal of Hydrologic Engineering*, *11*(6), 597–602. https//doi.org/10.1061/(ASCE)1084-0699(2006)11:6(597)

Ming, J., Xiao, C., Cachier, H., Qin, D., Qin, X., Li, Z., & Pu, J. (2009). Black Carbon (BC) in the snow of glaciers in west China and its potential effects on albedos. *Atmospheric Research*, *92*(1), 114–123.

Moeeni, H., Bonakdari, H., Fatemi, S. E., & Zaji, A. H. (2017). Assessment of stochastic models and a hybrid artificial neural network-genetic

algorithm method in forecasting monthly reservoir inflow. *INAE Letters*, *2*(1), 13–23. https//doi.org/10.1007/s41403-017-0017-9

Mohammadi, K., Eslami, H. R., & Dardashti, S. D. (2005). Comparison of regression, ARIMA and ANN models for reservoir inflow forecasting using snowmelt equivalent (a case study of Karaj). *Journal of Agricultural Science and Technology*, *7*(12), 17–30.

Moriasi, D. N., Arnold, J. G., Liew, M. W. Van, Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885–900.

Morrison, H., Thompson, G., & Tatarskii, V. (2009). Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Monthly Weather Review*, *137*(3), 991–1007. https//doi.org/10.1175/2008MWR2556.1

Motovilov, Y. G., Gottschalk, L., Engeland, K., & Rodhe, A. (1999). Validation of a distributed hydrological model against spatial observations. *Agricultural and Forest Meteorology*, *98-99*, 257–277. https//doi.org/10.1016/S0168-1923(99)00102-1

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https//doi.org/10.1016/0022-1694(70)90255-6

Noh, H., Lee, J, Kang, N, Lee, D, Kim, H. S., & Kim, S. (2016). Long-term simulation of daily streamflow using radar rainfall and the SWAT model: A case study of the Gamcheon Basin of the Nakdong River, Korea. https//doi.org/10.1155/2016/2485251

Oudin, L., Andréassian, V., Mathevet, T., Perrin, C., & Michel, C. (2006). Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations. *Water Resources Research*, *42*(7). https://doi.org/10.1029/2005WR004636

Painter, T. H., Deems, J. S., Belnap, J., Hamlet, A. F., Landry, C. C., & Udall, B. (2010). Response of Colorado River runoff to dust radiative forcing in snow. *Proceedings of the National Academy of Sciences*, *107*(40), 17125–17130. https//doi.org/10.1073/pnas.0913139107

Pankratz, A. (1991). *Forecasting with dynamic regression models*. New York: John Wiley.

Papamichail, D. M., & Georgiou, P. E. (2001). Seasonal ARIMA inflow models for reservoir sizing. *Journal of the American Water Resources Association*, *37*(4), 877–885. https//doi.org/10.1111/j.1752-1688.2001.tb05519.x

Penner, J. E., Andreae, M. O., Annegarn, H., Barrie, L., Feichter, J., Hegg, D., ... Pitari, G. (2001). Aerosols, their direct and indirect effects. In Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (pp. 289–348). Cambridge University Press.

Penner, J. E., Quaas, J., Storelvmo, T., Takemura, T., Boucher, O., Guo, H., ... Seland, Ø. (2006). Model intercomparison of indirect aerosol effects. *Atmospheric Chemistry and Physics*, *6*(11), 3391–3405.

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., & Wagener, T. (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, *79*(Supplement C), 214–232. https://doi.org/10.1016/j.envsoft.2016.02.008

Qian, Y., Flanner, M. G., Leung, L. R., & Wang, W. (2011). Sensitivity studies on the impacts of Tibetan Plateau snowpack pollution on the Asian hydrological cycle and monsoon climate. *Atmospheric Chemistry and Physics*, *11*(5), 1929–1948. https//doi.org/10.5194/acp-11-1929-2011

Qian, Y., Gustafson, W. I., Leung, L. R., & Ghan, S. J. (2009). Effects of soot-induced snow albedo change on snowpack and hydrological cycle in Western United States based on Weather Research and Forecasting chemistry and regional climate simulations. *Journal of Geophysical Research-Atmospheres*, *114*(D3), D03108. https//doi.org/10.1029/2008JD011039

Quaas, J., Boucher, O., Bellouin, N., & Kinne, S. (2008). Satellite-based estimate of the direct and indirect aerosol climate forcing. *Journal of Geophysical Research: Atmospheres*, *113*(D5), D05204.

Räisänen, P., Makkonen, R., Kirkevåg, A., & Debernard, J. B. (2017). Effects of snow grain shape on climate simulations: Sensitivity tests with the Norwegian earth system model. *The Cryosphere; Katlenburg-Lindau*, *11*(6), 2919–2942. https://dx.doi.org/10.5194/tc-11-2919-2017

Ramachandran, S., & Cherian, R. (2008). Regional and seasonal variations in aerosol optical characteristics and their frequency distributions over India during 2001-2005. *Journal of Geophysical Research*, *113*(D8), D08207. http://doi.wiley.com/10.1029/2007JD008560

Ramanathan, V., Crutzen, P. J., Kiehl, J. T., & Rosenfeld, D. (2001). Aerosols, climate, and the hydrological cycle. *Science*, *294*(5549), 2119–2124. https://doi.org/10.1126/science.1064034

Ramanathan, V., Crutzen, P. J., Lelieveld, J., Mitra, A. P., Althausen, D., Anderson, J., ... Valero, F. P. J. (2001). Indian Ocean Experiment: An integrated analysis of the climate forcing and effects of the great Indo-Asian haze. *Journal of Geophysical Research-Atmospheres*, *106*(D22), 28371–28398. https//doi.org/10.1029/2001JD900133

Reay, H. J., France, J. L., & King, M. D. (2012). Decreased albedo, e-folding depth and photolytic oh radical and no2 production with increasing black carbon content in Arctic snow. *Journal of Geophysical Research: Atmospheres*, *117*(D14), D00R20.

Regayre, L. A., Pringle, K. J., Lee, L. A., Rap, A., Browse, J., Mann, G. W., ... Woodhouse, M. T. (2015). The climatic importance of uncertainties in regional aerosol cloud radiative forcings over recent decades. *Journal of Climate*, *28*(17), 6589–6607. https://doi.org/10.1175/JCLI-D-15-0127.1

Ritter, A., & Muñoz-Carpena, R. (2013). Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. *Journal of Hydrology*, *480*, 33–45. https://doi.org/10.1016/j.jhydrol.2012.12.004

Rosenfeld, D. (2000). Suppression of rain and snow by urban and industrial air pollution. *Science*, *287*(5459), 1793–1796.

Rosolem, R., Gupta, H. V., Shuttleworth, W. J., Zeng, X., & de Gonçalves, L. G. G. (2012). A fully multiple-criteria implementation of the Sobol method for parameter sensitivity analysis. *Journal of Geophysical Research: Atmospheres*, *117*(D7), D07103. https://doi.org/10.1029/2011JD016355

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, *181*(2), 259–270. https://doi.org/10.1016/j.cpc.2009.09.018

Saltelli, A., Annoni, P., & D'Hombres, B. (2010). How to avoid a perfunctory sensitivity analysis. *Procedia -Social and Behavioral Sciences*, *2*(6), 7592–7594. https//doi.org/10.1016/j.sbspro.2010.05.133

Saltelli, A., Tarantola, S., & Chan, K. P.-S. (1999). A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, *41*(1), 39–56. https://doi.org/10.1080/0040176.1999.10485594

Samset, B. H. (2018). Aerosols and Climate—Oxford Research Encyclopedia of Climate Science. 25.

Schaefli, B., & Gupta, H. V. (2007). Do Nash values have value? *Hydrological Processes*, *21*(15), 2075–2080 en. http://doi.wiley.com/10.1002/hyp.6825

Schwartz, S. E. (1996). The whitehouse effect—shortwave radiative forcing of climate by anthropogenic aerosols: An overview. *Journal of Aerosol Science*, *27*(3), 359–382. https://doi.org/10.1016/0021-8502(95)00533-1

Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A., & Pulwarty, R. S. (1999). Characteristics of the Western United States snowpack from snowpack telemetry(SNOTEL) data. *Water Resources Research*, *35*(7), 2145–2160.

Shaw, W. J., Jerry Allwine, K., Fritz, B. G., Rutz, F. C., Rishel, J. P., & Chapman, E. G. (2008). An evaluation of the wind erosion module in DUSTRAN. *Atmospheric Environment*, *42*(8), 1907–1921. https://doi.org/10.1016/j.atmosenv.2007.11.022

Stewart, I. T., Cayan, D. R., & Dettinger, M. D. (2004). Changes in snowmelt runoff timing in Western North America under a 'Business as Usual' climate change scenario. *Climatic Change*, *62*(1-3), 217–232. https//doi.org/10.1023/B:CLIM.0000013702.22656.e8

Tomassini, L., Reichert, P., Knutti, R., Stocker, T. F., & Borsuk, M. E. (2007). Robust bayesian uncertainty analysis of climate system properties using Markov chain Monte Carlo methods. *Journal of Climate*, *20*(7), 1239–1254. https//doi.org/10.1175/JCLI4064.1

Twomey, S. (1974). Pollution and the planetary albedo. *Atmospheric Environment (1967)*, *8*(12), 1251–1256.

Twomey, S. (1991). Aerosols, clouds and radiation. *Atmospheric Environment. Part A. General Topics*, *25*(11), 2435–2442.

Uusitalo, L., Lehikoinen, A., Helle, I., & Myrberg, K. (2015). An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling & Software*, *63*(Supplement C), 24–31. https//doi.org/10.1016/j.envsoft.2014.09.017

Valipour, M. (2015). Long-term runoff study using SARIMA and ARIMA models in the United States. *Meteorological Applications*, *22*(3), 592–598.

Valipour, M., Banihabib, M. E., & Behbahani, S. M. R. (2013). Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *Journal of Hydrology*, *476*, 433–441. https//doi.org/10.1016/j.jhydrol.2012.11.017

VanRheenen, N. T., Wood, A. W., Palmer, R. N., & Lettenmaier, D. P. (2004). Potential implications of PCM climate change scenarios for Sacramento–San Joaquin River Basin hydrology and water resources. *Climatic Change*, *62*(1-3), 257–281. https//doi.org/10.1023/B:CLIM.0000013686.97342.55

Verheggen, B., & Weijers, E. P. (2010). Climate change and the impact of aerosol. 32.

Warren, S. G. (1984). Impurities in snow: Effects on albedo and snowmelt. *Annals of Glaciology*, *5*(1), 177–179.

Warren, S. G., & Clarke, A. D. (1990). Soot in the atmosphere and snow surface of Antarctica. *Journal of Geophysical Research*, *95*(181), I–1816.

Willmott, C. J., Robeson, S. M., & Matsuura, K. (2012). A refined index of model performance. *International Journal of Climatology*, *32*(13), 2088–2094. https//doi.org/10.1002/joc.2419

Wiscombe, W. J., & Warren, S. G. (1980). A model for the spectral albedo of snow. I: Pure snow. *Journal of the Atmospheric Sciences*, *37*(12), 2712–2733.

World Meteorological Organisation (1986). Intercomparison of models of snowmelt runoff, Operational Hydrology Report No. 23. Secretariat of the World Meteorological Organization. Geneva, Switzerland.

Wu, L., Gu, Y., Jiang, J. H., Su, H., Yu, N., Zhao, C., ... Choi, Y.-S. (2018). Impacts of aerosols on seasonal precipitation and snowpack in California based on convection-permitting WRF-Chem simulations. *Atmospheric Chemistry and Physics*, *18*(8), 5529–5547. https//doi.org/10.5194/acp-18-5529-2018

Wu, L., Su, H., Kalashnikova, O. V., Jiang, J. H., Zhao, C., Garay, M. J., ... Yu, N. (2017). WRF-Chem simulation of aerosol seasonal variability in the San Joaquin Valley. *Atmospheric Chemistry and Physics Discussions*, *17*(12), 7291–7309. https//doi.org/10.5194/acp-2016-981

Wu, L., Su, H., & Jiang, J. H. (2011a). Regional simulations of deep convection and biomass burning over South America: 1. model evaluations using multiple satellite data sets. *J. Geophys. Res.-Atmos.*, *116*(D17), D17208. https//doi.org/10.1029/2011JD016105

Wu, L., Su, H., & Jiang, J. H. (2011b). Regional simulations of deep convection and biomass burning over South America: 2. Biomass burning aerosol effects on clouds and precipitation. *Journal of Geophysical Research-Atmospheres*, *116*(D17), D17209. https//doi.org/10.1029/2011JD016106

Wu, L., Su, H., & Jiang, J. H. (2013). Regional simulation of aerosol impacts on precipitation during the East Asian summer monsoon. *Journal of Geophysical Research-Atmospheres*, *118*(12), 6454–6467. https//doi.org/10.1002/jgrd.50527

Xu, B., Cao, J., Hansen, J., Yao, T., Joswia, D. R., Wang, N., ... Yang, W. (2009). Black soot and the survival of Tibetan glaciers. *Proceedings of the National Academy of Sciences*, *106*(52), 22114–22118.

Xu, Z. X., & Li, J. Y. (2002). Short-term inflow forecasting using an artificial neural network model. *Hydrological Processes*, *16*(12), 2423–2439. https//doi.org/10.1002/hyp.1013

Ye, H., Zhang, R., Shi, J., Huang, J., Warren, S. G., & Fu, Q. (2012). Black carbon in seasonal snow across northern Xinjiang in northwestern China. *Environmental Research Letters*, *7*(4), 44002.

Zhao, C., Chen, S., Leung, L. R., Qian, Y., Kok, J. F., Zaveri, R. A., & Huang, J. (2013). Uncertainty in modeling dust mass balance and radiative forcing from size parameterization. *Atmospheric Chemistry and Physics*, *13*(21), 10733–10753. https//doi.org/10.5194/acp-13-10733-2013

Zhao, C., Hu, Z., Qian, Y., Ruby Leung, L., Huang, J., Huang, M., ... Streets, D. G. (2014). Simulating black carbon and dust and their radiative forcing in seasonal snow: A case study over North China with field campaign measurements. *Atmospheric Chemistry and Physics*, *14*(20), 11,475–11,491. https//doi.org/10.5194/acp-14-11475-2014

Zhao, C., Liu, X., Leung, L. R., Johnson, B., McFarlane, S. A., Gustafson Jr., W. I., ... Easter, R. (2010). The spatial distribution of mineral dust and its shortwave radiative forcing over North Africa: modeling sensitivities to dust emissions and aerosol size treatments. *Atmospheric Chemistry and Physics*, *10*(18), 8821–8838. https//doi.org/10.5194/acp-10-8821-2010

Zhao, C., Liu, X., Ruby Leung, L., & Hagos, S. (2011). Radiative impact of mineral dust on monsoon precipitation variability over West Africa. *Atmospheric Chemistry and Physics*, *11*, 1879–1893. https//doi.org/10.5194/acp-11-1879-2011

APPENDIX A: LIST OF ACRONYMS

ACF	Autocorrelation function
ACI	Aerosol-cloud interaction
AR	Autoregressive
ARI	Aerosol-radiation interaction
ARIMA	Autoregressive integrated moving average
ASI	Aerosol-snow interaction
ANN	Artificial neural network
CCN	Cloud condensation nuclei
GPF	General probabilistic framework
GSA	Global sensitivity analysis
MA	Moving average
MAE	Mean absolute error
NSE	Nash-Sutcliffe efficiency
PACF	Partial autocorrelation function
PBIAS	Percentage bias
RMSE	Root mean square error
SWE	Snow water equivalent
VIF	Variable inflation factor
WRF-Chem	Weather Research and Forecasting model with Chemistry

APPENDIX B: SAN JOAQUIN REGION AND BIG CREEK HYDROELECTRIC PROJECT

The Big Creek Hydroelectric Project is an extensive hydroelectric system that accounts for 12% of California's total hydroelectric generation. The project is located on the upper San Joaquin River system in the Sierra Nevada Mountains of Central California. Sierra Nevada is a mountainous region where most precipitation are retained as snow until temperatures are sufficient for melt (Cayan et al., 1993).

The hydroelectric project is owned and operated by Southern California Edison (SCE), which has a total installed capacity of 1,000 MW accounting for approximately 20% of SCE's total generation capacity. The hydroelectric system includes 27 dams, 23 generating units in 9 power houses, miles of underground tunnels, and 6 major reservoirs with a combined storage capacity of 560,000 acre ft. Water from lakes in higher elevation are routed through the nine powerhouses and discharged to lakes in lower elevations that are connected through tunnels and penstocks. The water travels a combined vertical distance of 6,655 feet before being discharged through the last powerhouse into the San Joaquin River. Florence Lake and Lake Thomas Alva Edison are the higher elevation reservoirs of the system having surface elevation of 7,300 and 7,648 feet, respectively. The dam at Florence Lake captures run-off from the South Fork of the San Joaquin River, diverting it through the Ward Tunnel towards the Portal Powerhouse, which is the first powerhouse in the system to receive water. Lake Thomas Alva Edison is formed by the Vermillion Valley Dam constructed across the Mono Creek, a tributary of the South Fork of the San Joaquin River. It discharges some of its water to the Ward Tunnel and thereby further regulates the water supply to the Portal Powerhouse. Water running through Portal Powerhouse gets discharged into the Huntington Lake where it is in turn diverted to lakes of lower elevation through other power houses. Thus, an accurate forecast of inflow into Lake Thomas Alva Edison and Florence Lake can greatly improve the operational efficiency of the Big Creek Hydroelectric Project.

APPENDIX C: A BRIEF OVERVIEW OF DYNAMIC REGRESSION MODEL

A brief overview of the dynamic regression model is presented in this section. A dynamic regression model is shown in Equation C1. A crucial assumption in dynamic regression is that the explanatory variables are not affected by the dependent variable, that is, there is no feedback between the variables.

$$Y_t = v(B)X_t + N_t, \tag{C1}$$

where

Y_t = dependent variable
X_t = the vector of explanatory variables
$v(B)$ = transfer function
N_t = noise time series.

A free-form distributed lag transfer function model like Equation C2 for M explanatory variables can be estimated where the noise series is approximated by a low order regular AR term proxy. The order of the

transfer function, k_i, is chosen based on the empirical understanding of the model.

$$Y_t = \sum_{i=1}^{M} \sum_{j=0}^{k_i} v_{i,j} X_{i,t-j} + \frac{1}{\phi(B)} a_t, \tag{C2}$$

where

$X_{i,t}$ = ith explanatory variable
$\phi(B)$ = low-order autoregressive proxy
a_t = white noise.

The individual weights $v_{i,j}$ are called impulse response weights. The transfer function can be written in a parsimonious form known as a rational distributed lag transfer function model as shown in Equation (C3).

$$Y_t = \mu + \sum_{i=1}^{M} \frac{\omega_i(B)}{\delta_i(B)} B^{b_i} X_{i,t} + N_t, \tag{C3}$$

where

$\omega_i(B)$ = $\omega_{i,0} + \omega_{i,1}B + \omega_{i,2}B^2 \dots + \omega_{i,h_i}B^{h_i}$
$\delta_i(B)$ = $1 - \delta_{i,1}B - \delta_{i,2}B^2 - \dots - \delta_{i,r_i}B^{r_i}$
b_i = dead time for input $X_{i,t}$.

It should be noted that r_i, h_i, and b_i are constants for the ith explanatory variable. (b, r, h) are the orders of the rational distributed lag transfer function. The numerator of the transfer function model captures the lagged effect of the covariates, and the denominator represents the decaying effects of the covariates. The noise series N_t may have an autocorrelated time structure that can be described by an ARIMA model.

The autoregressive (AR) component in the ARIMA model refers to the lagged values of the dependent variable time series; the moving average (MA) component refers to the lagged error terms, that is, residuals; and the integrated component represents the number of times a time series must be differenced to achieve stationarity. A general notation for ARIMA models is ARIMA (p, d, q), where p denotes the number of autoregressive terms, q denotes the number of moving average terms, and d denotes the number of times a series must be differenced to induce stationarity. Using the general notations of an ARIMA model, the noise series can be written as

$$N_t = \frac{\theta(B)}{\phi(B)} a_t, \tag{C4}$$

where

$\phi(B)$ = $1 - \phi_1 B - \dots - \phi_p B^p$
$\theta(B)$ = $1 - \theta_1 B - \dots - \theta_q B^q$.

Here, a_t is assumed to be white noise. The white noise assumption implies zero mean, normal distribution, and constant variance.

C.1 | Linear transfer function method

The LTF method suggested by Pankratz (1991) was applied in this study to handle multiple inputs. The order of the rational form transfer

function (b_i, r_i, h_i) for each variable i needs to be determined together with the order of ARIMA (p, d, q) model for the noise time series N_t. LTF method uses a free-form distributed lag model to estimate the impulse response weights in Equation (C2) together with an initial autoregressive proxy for the autocorrelation term of the disturbance time series N_t. If N_t is not stationary with time varying mean, then both the input and output time series should be differenced

accordingly. A parsimonious rational form transfer function similar to Equation (C3) is then identified by comparing the estimated impulse response weight pattern with theoretical impulse response weight patterns. The methodology described by Box et al. (2015) is then applied to determine ARMA order of the error time series N_t. Finally, the coefficients of the entire model are estimated and the model is checked for adequacy.