Environmental Modelling and Software 126 (2020) 104667

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: http://www.elsevier.com/locate/envsoft

An open source model for quantifying risks in bulk electric power systems from spatially and temporally correlated hydrometeorological processes

Yufei Su [a], Jordan D. Kern [b,*], Simona Denaro [a], Joy Hill [a], Patrick Reed [c], Yina Sun [d], Jon Cohen [e], Gregory W. Characklis [a,f]

[a] Department of Environmental Science and Engineering, University of North Carolina-Chapel Hill, Chapel Hill, NC, 27516, United States
[b] Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, 27695, United States
[c] Department of Civil and Environmental Engineering, Cornell University, Ithaca, NY, 14853, United States
[d] NextEra Energy, Juno Beach, FL, 33408, United States
[e] Department of Civil and Environmental Engineering, University of California-Davis, Davis, CA, 95616, United States
[f] Center on Financial Risk in Environmental Systems, University of North Carolina-Chapel Hill, Chapel Hill, NC, 27516, United States

ARTICLE INFO

Keywords:
Stochastic hydrology
Weather
Electricity markets
Prices

ABSTRACT

Variability (and extremes) in streamflow, wind speeds, temperatures, and solar irradiance influence supply and demand for electricity. However, previous research falls short in addressing the risks that joint uncertainties in these processes pose in power systems and wholesale electricity markets. Limiting challenges have included the large areal extents of power systems, high temporal resolutions (hourly or sub-hourly), and the data volumes and computational intensities required. This paper introduces an open source modeling framework for evaluating risks from correlated hydrometeorological processes in electricity markets at decision relevant scales. The framework is able to reproduce historical price dynamics in high profile systems, while also offering unique capabilities for stochastic simulation. Synthetic generation of weather and hydrologic variables is coupled with simulation models of relevant infrastructure (dams, power plants). Our model will allow the role of hydrometeorological uncertainty (including compound extreme events) on electricity market outcomes to be explored using publicly available models.

1. Introduction

In recent years, interest has grown in exploring the effects of hydrometeorological variability, and especially extreme events, on the operations of bulk power systems (large, interconnected systems of generation, transmission and load (demand)) (Collins et al., 2018; Forster and Lilliestam, 2011; Franco and Sanstad, 2008; Kern and Characklis, 2017; Staffell and Pfenninger, 2018; Tarroja et al., 2016; Turner et al., 2019; van Vliet et al., 2016, 2012; Voisin et al., 2018). Both droughts and floods compromise the operations of hydroelectric dams (Gleick, 2017; Su et al., 2017; Tarroja et al., 2016), while droughts in particular can also impact thermal power plants that are dependent on cooling water (van Vliet et al., 2016, 2012). Air temperatures influence a range of system components, most notably electricity demand for heating and cooling (Franco and Sanstad, 2008). In addition, as variable energy resources like wind and solar expand their share of the power mix, the grid is becoming more sensitive to fluctuations in wind speeds and solar irradiance (Collins et al., 2018; Staffell and Pfenninger, 2018). By influencing supply and demand for electricity, hydrometeorological processes have direct impacts on pollution (e.g., increased greenhouse gas emissions (Collins et al., 2018; Hardin et al., 2017; Tarroja et al., 2016)), wholesale electricity prices (Boogert and Dupont, 2005; Collins et al., 2018; Seel et al., 2018), and the financial standing of suppliers of electricity (e.g., retail utilities, renewable energy producers) and consumers (Bain and Acker, 2018; Boogert and Dupont, 2005; Foster et al., 2015; Kern and Characklis, 2017; Kern et al., 2015).

However, with few exceptions (Turner et al., 2019), previous investigations fall short in assessing the holistic influence of hydrometeorological variability on bulk power systems. Past research efforts assess operational and financial risks from exposure to variability in a more limited set of hydrometeorological processes (Collins et al., 2018; Kern et al., 2015) (e.g., streamflow and temperatures, or wind speeds and solar irradiance); do not consider these effects within the context of large, interconnected power systems (Kern and Characklis, 2017);

* Corresponding author.
E-mail addresses: yufeisu@live.unc.edu (Y. Su), jkern@ncsu.edu (J.D. Kern).

https://doi.org/10.1016/j.envsoft.2020.104667
Received 9 October 2019; Received in revised form 31 January 2020; Accepted 12 February 2020
Available online 15 February 2020

and/or do not assess impacts probabilistically (Hardin et al., 2017). These shortcomings may be partly attributable to the challenges of modeling bulk electric power systems at sufficient scale and resolution to simulate system operations in a realistic way, and over sufficient time horizons to explore joint uncertainty in multiple, correlated input variables.

Interconnected power systems span areas so large that system operators often have some ability to deal with spatially heterogeneous stressors. For example, a localized power supply shortfall caused by drought in one area might be managed by importing power from other areas where water, and thus electricity from hydropower production and water-cooled generators, is more abundant. From a modeling perspective, this necessitates adopting system topologies that extend beyond a single watershed, state, and region. Hydrometeorological uncertainty and power system risks can also manifest on different time scales. Extreme meteorological and hydrological conditions can have durations on the order of days (floods (Najibi and Devineni, 2017), heat waves), weeks to months (wind "droughts"), and years (hydrological droughts (Andreadis et al., 2005)), whereas power system modeling requires an hourly or sub-hourly time step (Pandžžić et al., 2014). Although stochastic modeling approaches can be used to create large synthetic records of hydrometeorological processes in order to explore risks from extreme events (Brown et al., 2015; Reed et al., 2013), this poses a direct challenge to the use of computationally expensive integer programming within power system models (Pandzžić et al., 2014), making large ensemble Monte Carlo simulations less tractable. Adding to these challenges is the potential presence of significant spatial and temporal covariance among key hydrometeorological processes (Jimenez et al., 2011; Woodhouse et al., 2016). If significant correlations exist, an increased number of model runs may be required to characterize the probability of coincident extremes (e.g., widespread simultaneous hydrological drought, a wind drought, and a heat wave) that may be of particular concern to power system operators (Mazdiyasni and AghaKouchak, 2015; Turner et al., 2019).

The modeling scales, resolutions, and ensemble sizes required in exploring the risks to bulk electric systems from hydrometeorological variability present a challenge, and few (if any) models capable of performing this type of analysis are publically available. Given recent increased interest among the research community in modeling interconnected systems (e.g., food-energy-water (Logan, 2015)), a generalizable and open source modeling framework for simulating the influence of correlated hydrometeorological processes on power system dynamics at decision relevant scales would be a valuable addition.

The goal of this paper is to present such a framework: the newly developed California and West Coast Power (CAPOW) systems model. CAPOW was designed by the authors to explore a high profile test-bed–the West Coast of the conterminous United States (U.S.). The bulk electric power systems covering most of the states of California, Oregon and Washington are included, as well as the two major wholesale electricity markets active across these states (current gaps in coverage are the PacifiCorp West, Sacramento Municipal Utility District, Los Angeles Department of Water and Power balancing authorities). CAPOW is comprehensive in its treatment of stochastic weather and streamflow, simulation of relevant infrastructure (reservoir networks, power systems), and evaluation of outcomes (system costs, prices, etc.). While focused on the U.S. West Coast, the steps required in building and executing the CAPOW model (as well as much of the code) are fairly generalizable and can be transferred to other systems and interconnections of interest (Chowdhury et al., 2019). Most grid specific information used in the model is publically available anywhere in the U. S. (generator size, location, fuel type, prime mover type, average heat rate, etc.). Hydrometeorological data used to simulate electricity demand, wind, solar and hydropower production are also available throughout the U.S.; as well as hourly records of renewable energy production in each balancing authority through the EIA. Analogous transmission grid information (bi-directional capacities) is publically available for all WECC areas, and for many (if not all) sub-regions in the eastern interconnection. Note that to transfer the model to other regions, additional capabilities that are not currently in CAPOW may be required (e.g., representing impacts of extreme cold, air temperatures (Henry and Pratson, 2016), and a lack of cooling water availability due to low streamflow and temperatures (Miara et al., 2017; van Vliet et al., 2016, 2012) on thermal power plant functionality). The model is Python-based; all code and data required to run the CAPOW model, as well as some documentation of the model, is available at https://github.com/romulus97/CAPOW_PY36 under the MIT free software license.

2. Methods

Our description of methods parallels the CAPOW model's work flow (Fig. 1), beginning with a discussion of surface water and electric power system topologies, including key physical assets (e.g., power plants, dams/reservoirs) and their connections (i.e., water routing between reservoirs, high voltage transmission pathways). This is followed by a description of CAPOW's unit commitment and economic dispatch (UC/ED) model, which is used to simulate actual power system operations. The methods section ends with a description of our approach for stochastically generating model inputs from historical weather and streamflow data.

2.1. System topology

2.1.1. Electric power

In order to model the West Coast grid (the case study explored here), we first adopt a 21-zone topology of the Western Electricity Coordinating Council (WECC), a regulatory body charged with reducing risks to the Western grid by enforcing standards and assessing reliability (Fig. 2). This topology, which has been used in the past by WECC and other researchers to assist in long term planning exercises (Ho et al., 2016; Mkarov et al., 2010), groups balancing authorities (utility footprints) into multiple zones that are connected via aggregated transmission pathways throughout the region. Each zone-to-zone transmission pathway is associated with bi-directional capacities (i.e., maximum limits on zone-to-zone transfers of electricity) estimated from publically available data (Western Electricity Coordinating Council, 2016).

Each zone in the network consists of: 1) the load (electricity demands) of its member balancing authorities, which fluctuate on hourly, daily, seasonal and annual time scales; and 2) a portfolio of co-located generation resources with which to meet those demands. Comprehensive databases of generators located in each node of the 21-zone WECC topology are publically available from multiple sources (US Environmental Protection Agency, 2018; Western Electricity Coordinating Council System Adequacy Planning Department, 2015). These also contain information on relevant operating characteristics for each generator (e.g., fuel type, capacity, average heat rate) that are used to formulate the UC/ED simulation model.

There are two major trading hubs for wholesale electricity on the U. S. west coast: 1) the Mid-Columbia (Mid-C) market that serves as a hub for much of the Pacific Northwest region; and 2) the California Independent System Operator (CAISO), a competitive wholesale market that manages approximately 80% of California's electricity flow. The 21-zone WECC topology shown in Fig. 2 includes five nodes (red, numbered) that directly correspond to these markets: node 1 (Pacific Northwest) corresponds to the Mid-C market, and nodes 2–5 correspond to the CAISO market. Nodes 2–5 also represent the service areas of three major utilities: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). Currently only these five zones (and power flows among them) are modeled mechanistically using a UC/ED model. No UC/ED models exist outside these five zones. Neighboring zones are considered only in terms of their exchanges of electricity with the core UC/ED zones, and these exchanges

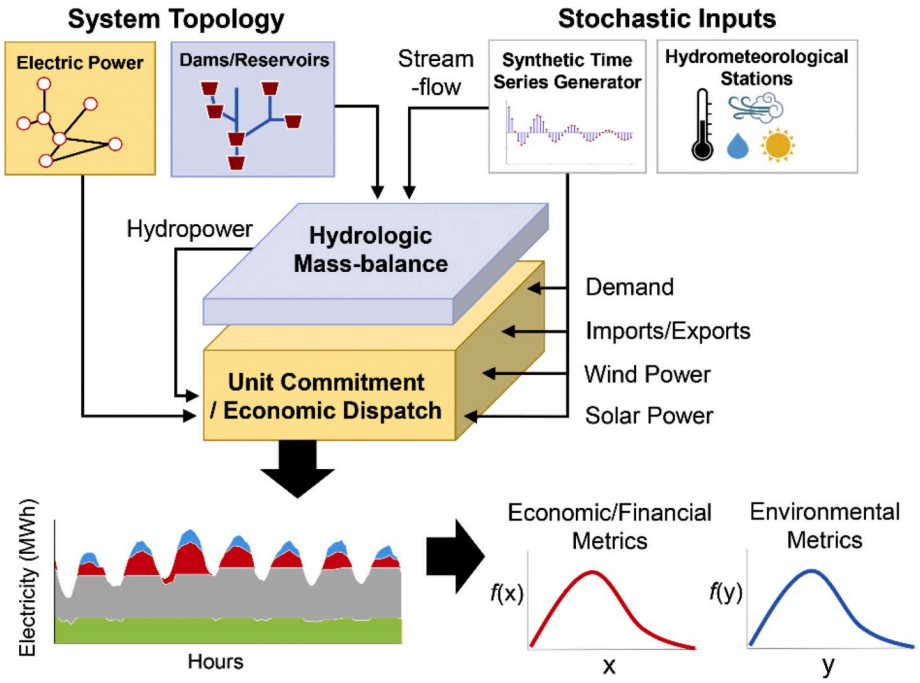

Fig. 1. Model workflow. Topologies of relevant electric power and surface water infrastructure are defined first, and then synthetic time series inputs are used to drive stochastic simulation of a power system (unit commitment/economic dispatch) model. Model outputs include the least cost generation schedule, total system costs, estimated wholesale prices, and emissions.

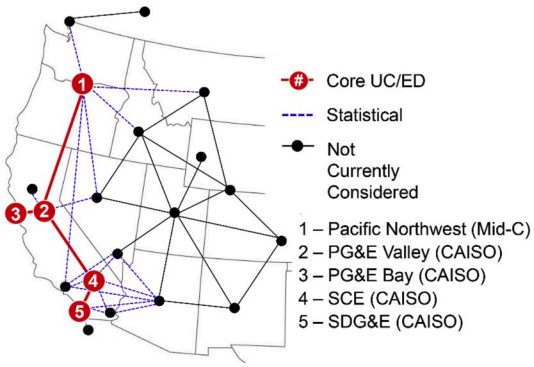

Fig. 2. Power system topology used in the CAPOW modeling framework. The five red zones, comprising collections of balancing authorities (load centers and generation assets), are mechanistically modeled using unit commitment economic dispatch (UC/ED) models. Blue lines represent exchanges (imports/exports) of electricity with adjacent zones that are represented statistically. Black dots represent zones in the WECC system that are not currently represented in CAPOW. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

are modeled statistically (see Supplemental Material).

2.1.2. Dams and reservoirs

Recent analyses of the impacts of drought on power generation in the Western U.S. (Harto et al., 2011) suggest that cooling water issues from low streamflow and high water temperatures pose a minor threat to thermal power plants in the region. Rather, the primary mechanism through which hydrologic extremes can impact power system operations is through variability in hydropower generation. Within the WECC topology shown in Fig. 2, hydropower capacity makes up 58% of installed generating capacity in zone 1 (Pacific Northwest), 18% of generating capacity in zone 2 (PG&E Valley), and 4% of capacity in zone 3 (SCE) (US Environmental Protection Agency, 2018). Fig. S2 in the Supplemental Material section maps major (>5 MW) hydroelectric dams that participate in balancing authorities located within the five numbered zones that make up the UC/ED model. These dams primarily fall within the Columbia River Basin, which spans several Northwestern U.S. states and Canada, as well as the Sacramento River, San Joaquin River, and Tulare Lake basins in California.

Publically available hydrologic mass balance models exist for 85% of the hydropower capacity in the Pacific Northwest (versions of HYSSR, developed by the U.S. Army Corps of Engineers to simulate the Federal Columbia River Power System; and a ResSim model that simulates the operations of Federal dams in the Willamette River Basin). Models exist for only 12% of the hydropower capacity in California (the ORCA model (Herman and Cohen, 2019), which simulates the operations of major storage/flood control dams). In California, much of the state's hydropower capacity is privately owned and located in high altitude areas of the Sierra Nevada Mountains. Little information about the operation of these dams is publically available, so hydropower production at these projects is simulated via an alternative approach in which hydropower production at upstream dams is predicted using observed streamflow downstream. First, for major high altitude hydroelectric dam in the Sierra Nevada Mountains, a corresponding downstream storage reservoir or stream gauge on the same river is identified. In order to predict upstream hydropower generation at a given dam using observed streamflow downstream, the calendar year is broken into four seasons: winter, spring, summer, and fall. Each season is assumed to follow a different set of "operating rules" that translate observed downstream flows into estimates of upstream hydropower production. Rules are fitted using the differential evolution algorithm in the SciPy library of Python, based on root mean squared error (RMSE) between observed and simulated hydropower production for each upstream dam.

About 15% of hydropower capacity in the Pacific Northwest and 20% of hydropower capacity in California are within the five core WECC zones that make up the UC/ED model but fall outside the four river basins mentioned above and are not associated with publically available models. These projects are modeled by scaling hydropower generation from nearby dams. A more detailed description of how hydropower

production is simulated on a daily basis can be found in the Supplemental Material.

2.2. Unit commitment and economic dispatch model

The power system and reservoir network topologies described above form the basis of a unit commitment/economic dispatch (UC/ED) model that we use to simulate the operation of the five numbered WECC zones in Fig. 2, which include the Mid-C and CAISO markets. Simulating the UC/ED model for a single year at an hourly time step takes approximately 6 h using the CPLEX solver on a 16-core machine with 2.5 GHz processors using a Linux operating system. What follows is a general overview of the model's structure and functionality. A mathematical formulation of the UC/ED model can be found in the Supplemental Material.

We coded the UC/ED model in Python using the Pyomo mathematical optimization package, structuring it as an iterative, mixed integer linear program. Over a user-defined operating horizon (e.g., 48 h), deterministic optimization is used to minimize the cost of meeting demand for electricity and operating reserves (including unit start costs, no load costs, fuel costs, and penalties associated with transferring electricity between zones), subject to constraints on individual generators and transmission paths. Costs are minimized by strategically "dispatching" (scheduling) generation from flexible generation resources (natural gas power plants, hydroelectric dams and system imports) on an hourly basis. Variable renewable energy (wind and solar) are not dispatchable (they can be consumed only when available); as such, they are typically treated as "electricity demand reduction" within a zone, but can be also curtailed during periods of oversupply.

A single iteration of the UC/ED model yields system costs and the least cost generating schedule over the operating horizon (e.g., hours 1–48); however, only the first 24 h of the solution is stored. The remaining solution (hours 25–48) is discarded, and the whole process shifts one day into the future. The next iteration of the model identifies a solution for the hours 25–48, while again looking 48 h into the future (i.e., at hours 25–72). This ensures that the model does not have perfect foresight over unreasonably long time horizons when making decisions with path dependency (e.g., turning on baseload power plants with high "minimum up" times).

Simulation of the UC/ED model creates hourly time series outputs that track provision of electricity and operating reserves by each generator, the flow of electricity among zones, plant specific and system wide emissions of CO_2, total operating costs, and wholesale electricity prices. CO_2 emissions from each power plant are calculated using historical EPA eGrid data that are used to estimate the kg CO_2 per MWh emissions for each plant. Note that total operating costs essentially refers to the value of the objective function in each hour (the cumulative start, no load, and fuel costs across every power plant in every hour). On the other hand, wholesale electricity prices ($/MWh) are dynamic measures of the marginal value of electricity in each market, i.e., how much generators would be paid to sell their electricity in each hour. Within the optimization, wholesale prices are estimated for each zone as the shadow cost of an energy balance constraint at each zone (i.e., the change in objective function value associated with a 1 MWh increase in demand at each zone). Calculating the shadow costs requires the UC/ED model to first be solved in mixed integer form, and then resolved as a linear program (keeping all binary variables fixed from the integer solution) in order to access dual values for relevant constraints in Pyomo. This yields a separate time series of wholesale electricity prices for each of the five WECC zones represented in the core UC/ED model. Prices in the Mid-C market are assumed to be equivalent to prices for the Pacific Northwest zone. To represent the CAISO market, prices for the four relevant zones in California (PG&E Valley, PG&E Bay, SCE, and SDG&E) are weighted to determine an overall price for the market, with the weights fitted via regression ($R^2 = 0.75$, $p < 1e\text{-}3$) on observed values over the period 2012–2016.

2.3. Stochastic inputs

The primary stochastic inputs to the UC/ED model are electricity demand (hourly), wind and solar power production (hourly), and available hydropower production (daily) for each numbered zone in Fig. 2. Several hydrometeorological processes (air temperatures, wind speeds, solar irradiance and streamflow) in turn drive these power system inputs. In the following section, we describe our approach for generating synthetic hydrometeorological time series.

2.3.1. Hydrometeorological variables

2.3.1.1. Air temperatures, wind speeds, and solar irradiance. We collect observed air temperatures, wind speeds, and solar irradiance data within major cities (where electricity demand is highest) and in areas known to have large amounts of installed wind and solar power capacity. Records of daily average temperature and wind speed over the period 1998–2017 come from NOAA's Global Historical Climatological Network (GHCN) for seventeen meteorological stations distributed throughout the Western U.S. (Table 1). Global horizontal irradiance data come from the National Renewable Energy Laboratory's National Solar Radiation Database (NSRDB) (Sengupta et al., 2018); both "clear sky" and observed irradiance data are acquired at a 30-min resolution and then aggregated to daily sums.

Each weather station provides the data necessary to generate 365-day profiles of average temperature and wind speed for their respective locations. We use solar irradiance data to created 365-day profiles of average "clear sky" (cloudless) conditions (Fig. 3).

$$TP_n = \frac{1}{Y}\sum_{y=1}^{Y} T_{n,y} \tag{1}$$

$$WP_n = \frac{1}{Y}\sum_{y=1}^{Y} WS_{n,y} \tag{2}$$

$$SP_n = \frac{1}{Y}\sum_{y=1}^{Y} S_{n,y} \tag{3}$$

Where,

TP_n = average temperature on calendar day n across Y years (°C)
$T_{n,y}$ = observed temperature on calendar day n in year y (°C)
WP_n = average wind speed on day n across Y years (m/s)
$WS_{n,y}$ = observed wind speed on day n in year y (m/s)
SP_n = average clear sky irradiance on day n across Y years (W/m^2)
$S_{n,y}$ = observed clear sky irradiance on day n in year y (W/m^2)

Synthetic values of air temperatures, wind speeds, and solar irradiance are then generated by combining these average profiles (e.g. blue series in panel A of Fig. 3) with stochastic representation of the autocorrelated "residuals" that deviate from these repeating signals (e.g. the gray series in panel A of Fig. 3). Average temperature and wind profiles are subtracted from observed temperature and wind speed values; this yields a daily record of zero-mean residuals (i.e., deviations from average temperature and wind speed for each calendar day over the period 1998–2017). Observed irradiance is subtracted from average clear sky irradiance, yielding a daily record of "losses" due to cloud effects.

$$RT_d = T_d - TP_n \tag{4}$$

$$RW_d = WS_d - TW_n \tag{5}$$

$$IL_d = SP_n - I_d \tag{6}$$

Where,

RT_d = residual temperature on day d (°C)
RW_d = residual wind speed on day d (m/s)

Table 1
Seventeen weather stations in the Global Historical Climatological Network and National Solar Resource Database that provide daily mean air temperature and wind speed data used in development of stochastic inputs.

Station ID	Name	Variables	Latitude	Longitude
USW00024232	SALEM AIRPORT MCNARY FIELD, OR	Wind/ temps	44.90° N	123.00° W
USW00024221	EUGENE MAHLON SWEET FIELD, OR	Wind/ temps	44.12° N	123.21° W
USW00024233	SEATTLE TACOMA INTERNATIONAL AIRPORT, WA	Wind/ temps	47.45° N	122.30° W
USW00024131	BOISE AIR TERMINAL, ID	Wind/ temps	43.56° N	116.22° W
USW00024242	PORTLAND TROUTDALE AIRPORT, OR	Wind/ temps	45.54° N	122.39° W
USW00024157	SPOKANE INTERNATIONAL AIRPORT, WA	Wind/ temps	47.62° N	117.53° W
USW00024163	PASCO TRI CITIES AIRPORT, WA	Wind/ temps	46.26° N	119.11° W
USW00093193	FRESNO YOSEMITE INTERNATIONAL, CA	Wind/ temps	36.77° N	119.71° W
USW00023230	OAKLAND METRO INTERNATIONAL AIRPORT, CA	Wind/ temps	37.71° N	122.21° W
USW00023174	LOS ANGELES INTERNATIONAL AIRPORT, CA	Wind/ temps	33.94° N	118.40° W
USW00023188	SAN DIEGO INTERNATIONAL AIRPORT, CA	Wind/ temps	32.73° N	117.19° W
USW00023232	SACRAMENTO EXECUTIVE AIRPORT, CA	Wind/ temps	38.51° N	121.49° W
USW00023293	SAN JOSE, CA	Wind/ temps	37.33° N	121.88° W
USW00023234	SAN FRANCISCO INTERNATIONAL AIRPORT, CA	Wind/ temps	37.62° N	122.37° W
USW00023160	TUCSON INTERNATIONAL AIRPORT, AZ	Wind/ temps	32.11° N	110.93° W
USW00023183	PHOENIX AIRPORT, AZ	Wind/ temps	33.43° N	112.00° W
USW00053123	LAS VEGAS AIR TERMINAL, NV	Wind/ temps	36.21° N	115.19° W
NSRDB 154166	NATIONAL SOLAR RESOURCE DATABASE #1	Irradiance	40.45° N	121.66° W
NSRDB 13631	NATIONAL SOLAR RESOURCE DATABASE #2	Irradiance	38.57° N	121.7° W
NSRDB 111895	NATIONAL SOLAR RESOURCE DATABASE #3	Irradiance	36.81° N	119.38° W
NSRDB 93873	NATIONAL SOLAR RESOURCE DATABASE #4	Irradiance	35.09° N	117.3° W
NSRDB 83553	NATIONAL SOLAR RESOURCE DATABASE #5	Irradiance	34.05° N	118.38° W
NSRDB 82442	NATIONAL SOLAR RESOURCE DATABASE #6	Irradiance	33.93° N	115.9° W
NSRDB 77068	NATIONAL SOLAR RESOURCE DATABASE #7	Irradiance	33.33° N	114.7° W

IL_d = irradiance "losses" on day d (W/m^2)

Residual temperatures and wind speeds, as well as irradiance losses, are then mean-shifted to eliminate negative values and log-transformed to approximate a Gaussian distribution. The residuals/losses for each calendar day of the year are then divided by their respective standard deviations, in order to control for seasonal heteroscedasticity.

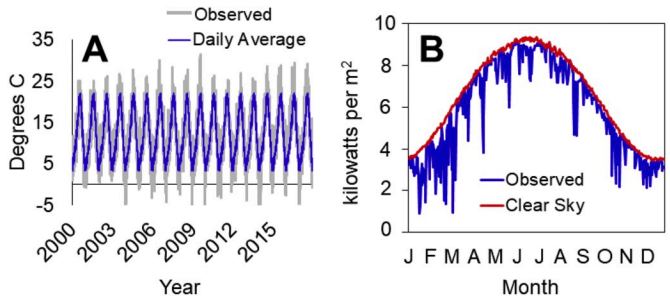

Fig. 3. (A) Daily average and observed temperatures for USW00024232 (Salem, OR). (B) Daily average clear sky conditions and one year of observed irradiance for NSRDB 11895.

$$WRT_d = \widehat{RT_d}/\sigma T_n \quad (7)$$

$$WRW_d = \widehat{RW_d}/\sigma W_n \quad (8)$$

$$WIL_d = \widehat{IL_d}/\sigma IL_n \quad (9)$$

Where,

WRT_d = whitened residual temperature on day d WRW_d = whitened residual wind speed on day d WIL_d = whitened irradiance losses on day d $\widehat{RT_d}$ = mean shifted, log-transformed residual temperature on day d (°C)

$\widehat{RW_d}$ = mean shifted, log-transformed residual wind speed on day d (m/s)

$\widehat{IL_d}$ = mean shifted, log-transformed irradiance losses on day d (W/m^2)

σT_n = standard deviation of transformed temperature residuals on calendar day n σW_n = standard deviation of transformed wind speed residuals on calendar day n σIL_n = standard deviation of transformed irradiance losses on calendar day n

We then model the resultant "whitened" residuals and irradiance losses using a vector autoregressive (VAR) model, in order to capture observed covariance across variables. VAR models describe the behavior of a set of k variables over a given time period as a linear function of their past values and random samples from a multivariate normal distribution. Simulated values of each variable are stored in a $k \times 1$ vector, y_t, which has as its i^{th} element, $y_{i,t}$, the value of the i^{th} variable at time t. The "lag" of the model (i.e., the number of previous time steps that are accounted for when estimating values in y_t) is denoted by the parameter p.

$$y_t = C + A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + \varepsilon_t \quad (10)$$

Where,

$C = k\,x\,1$ vector of constants
$A_i = k\,x\,k$ matrix of coefficients
$\varepsilon_t = k\,x\,1$ vector of error terms
t = time period
p = model lag

Simulation of y_t proceeds through random sampling of noise (ε_t) from a multivariate normal distribution with a covariance matrix estimated from whitened residuals and irradiance losses for the period 1998–2017. The number of lags considered is determined via the Akaike Information Criteria.

A fitted VAR model is used to simulate daily, whitened temperature and wind speed residuals and irradiance losses for each GHCN and NSRDB site considered, for as many years as desired. Simulated values are then "un-whitened" by reversing Equations (7)–(9) (thus restoring heteroscedasticity and non-normality); they are then added back to the 365-day profiles (reversing Equations (4)–(6)), yielding synthetic daily

records of temperature and wind speeds.

2.3.1.2. Streamflow. Streamflow patterns on the west coast of the U.S. are driven by runoff from precipitation as rain and, largely, the melting of snow accumulated during the winter. Both total annual streamflow and the within year distribution of streamflow experienced in this region are known to be influenced by temperatures (Null et al., 2010). At the same time, there are significant correlations among the 85 separate, spatially distributed streamflow gauges that drive CAPOW's simulation of dam operations and hydropower production.

We make use of a Gaussian Copula to preserve the relationship between total annual streamflow and temperatures in stochastically generated samples. First, observed daily average temperatures (1953–2008) at the seventeen meteorological stations are converted to heating and cooling degree days, which measures deviations from 18.33 degrees C (65 degrees F).

$$HDD_{d,s} = \max(18.33 - T_{d,s},\ 0) \tag{11}$$

$$CDD_{d,s} = \max(T_{d,s} - 18.33,\ 0) \tag{12}$$

Where,

$HDD_{d,s}$ = heating degree days on day d at station s

$CDD_{d,s}$ = cooling degree days on day d at station s

$T_{d,s}$ = average near surface air temperature on day d ($^{\circ}C$) at station s

Total annual HDDs and CDDs are calculated, providing coarse measures of the "hotness" of a given year's summer and the "coldness" of a given year's winter. Total annual HDDs and CDDs and total annual streamflow are then transformed into quantile space by calculating the empirical cumulative probability distribution for each variable.

$$P = \mathrm{P}(Q \geq q) \tag{13}$$

Where,

Q = total annual streamflow or degree days at a given site

Empirical probabilities are transformed again into a uniform distribution ranging from −1 to 1 as follows, ensuring a mean of 0 across every variable.

$$Y = 2(P - 0.5) \tag{14}$$

The covariance matrix C across all the variables at every site is estimated, and then synthetic records of total annual streamflow and total annual HDDs and CDDs are generated by taking random samples from a multivariate normal distribution with mean 0 and covariance matrix C, then back-transforming (reversing equations (13) and (14)).

The next step is to match total annual streamflow and total annual HDDs and CDDs simulated via the Copula method with the synthetic daily temperatures generated in the previous section using a vector-autoregressive (VAR) approach. Synthetic daily temperatures simulated using the VAR approach are converted to total annual HDDs and CDDs. For each year of synthetic data desired, we select a single year of total annual HDDs and CDDs generated using the VAR approach, and then calculate the weighted average across every GHCN station. Weights are determined by the fraction of average annual flow across the 85 stream gauges that is contained within each GHCN station's surrounding area:

$$WT_s = \frac{\sum_{g=1}^{G} AVF_g}{AVT} \tag{15}$$

Where,

WT_s = weight assigned to meteorological stationAVF_g = average annual flow at gauge site g closest to stationAVT = average annual flow across all 85 stream gauges

The weighted total annual HDDs and CDDs from the VAR model are compared alongside pairs of weighted total annual HDDs and CDDs generated using the Copula method. The smallest mean squared error difference is identified; then the total annual streamflow values generated via the Copula method are paired with the corresponding daily temperatures (and also wind speeds and solar irradiance) generated via VAR.

Disaggregating total annual streamflow values down to a daily time step must be done in a manner that considers the potential influence of temperatures on the timing of streamflow throughout the year. For example, Fig. 4 shows the relationship between winter and spring temperatures and the timing of streamflow at two major reservoirs in California. The top panel (A) shows 19 years (1997–2015) of weighted average temperatures across the GHCN stations, calculated using weights from Equation (15). Lines are colored according to the mean temperature experienced over the first 24 weeks of the year; the dark red line indicates the year with the hottest temperatures over this period (2015), and the dark blue line indicates the year with the coolest temperatures (2010). In panels B and C, those same line colors are then used to plot contemporaneous "full natural" (unregulated) flows at Folsom Dam (panel B) and Oroville Dam (panel C) in California (two large storage dams for which there are long historical flow records). Flows are

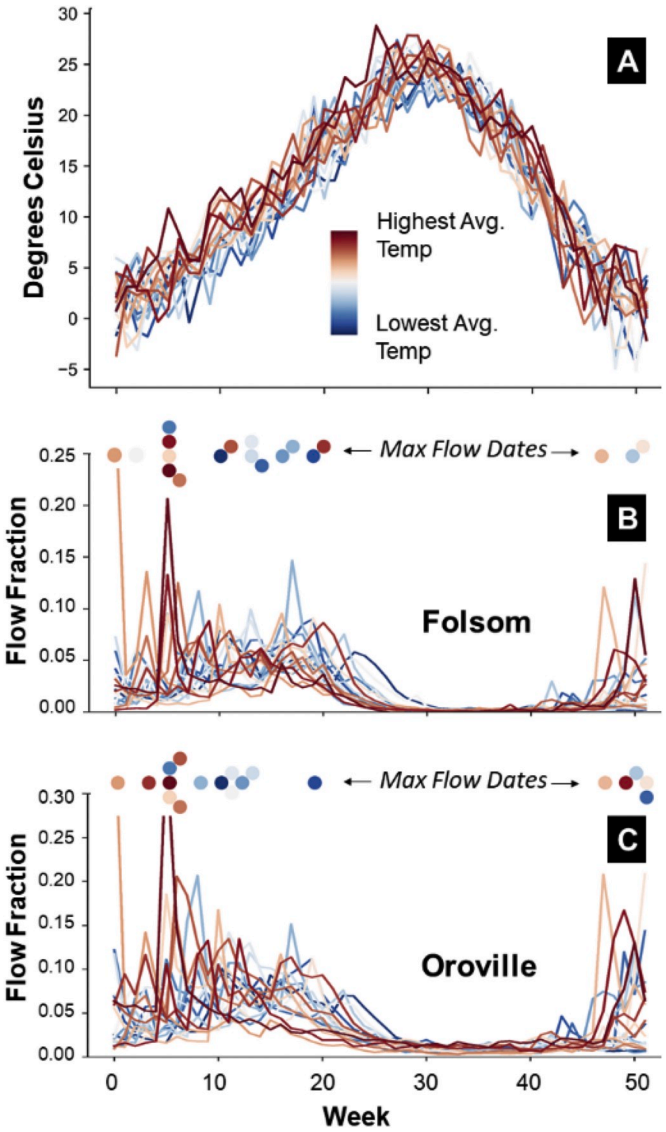

Fig. 4. (A) Weighted average temperatures for the period 1997–2015, colored according to mean temperatures experienced during the first 24 weeks of the year; (B) associated unregulated daily flow fraction profiles for Folsom Dam, with a swarm plot indicating the week of maximum unregulated streamflow; (C) similar data for Oroville Dam.

shown in terms of standardized "fractions" that are created by dividing by total annual flows at each site. At the top of panels B and C, swarm plots identify the week of maximum streamflow. For both dams, years with higher average winter and spring temperatures (red hued circles) tend to be associated with earlier peak streamflow, indicating earlier snowmelt and/or major precipitation events.

In order to capture these dependencies between the timing of streamflow and temperatures, we follow a nearest neighbor clustering approach, similar to Nowak et al. (2010). The weights generated in Equation (15) are used to create composite time series of temperatures across the 17 GHCN stations, for both historical and simulated temperature data. For each simulated year, the historical record is searched for a past year that exhibited the most similar winter/spring temperature profile, in terms of mean squared error. The identified historical year is then selected as the basis for determining daily flow fractions at each streamflow gauge site. For the historical year selected, daily flow fractions are calculated as follows:

$$FF_{d,g} = \frac{DF_{d,g}}{AF_g} \tag{16}$$

Where,

$FF_{d,s}$ = flow fraction for day d at streamflow gauge site $gDF_{d,s}$ = observed flow on day d at streamflow gauge site gAF_g = total annual flow observed at gauge site g

Flow fractions for each gauge site are then multiplied by simulated total annual flows to yield a synthetic record of daily flows across the study area.

2.3.2. Power system inputs

The stochastic scenario generation framework permits the exploration of large ensembles of time series for temperatures, wind speeds, solar irradiance, and streamflow. These data are then converted to associated power system inputs for the UC/ED model (time series for each zone of hourly electricity demand, wind and solar availability, daily hydropower production and imports of electricity from other areas in the Western U.S.). Table 2 provides an overview of the different approaches taken to translate raw hydrometeorological variables into power system inputs, as well as their accuracies. Multi-variate regression is used to simulate daily electricity demand, solar and wind power production, and system imports (power flows along WECC Paths listed in Table 2). Daily values are disaggregated down to an hourly time step by sampling from historical profiles. Daily values of available hydropower production are created by passing synthetic streamflow records through mass-balance hydrologic models of dams in the Columbia River basin and major storage reservoirs in California, as well as through a machine learning representation of high altitude hydropower production in California. Detailed descriptions of all models used to translate raw hydrometeorological variables into power system inputs can be found in the Supplemental Material.

Table 2
Model results for power system inputs. R^2 values are based on daily fit for all inputs except hydropower production (weekly). In all cases, regression p-values are less than .01.

Power System Input	R^2 Value	Predictive/Independent Variables	Years
CAISO Solar Power	0.92	Irradiance	2011–2016
Pacific Northwest Wind Power	0.71	Wind speed	2011–2016
CAISO Wind Power	0.71	Wind speed	2011–2016
Pacific Northwest Electricity Demand	0.89	Temperature, wind speed, day-of-week	2010–2016
PG&E Valley Electricity Demand	0.90	Temperature, wind speed, day-of-week	2010–2016
PGE&E Bay Electricity Demand	0.79	Temperature, wind speed, day-of-week	2010–2016
SCE Electricity Demand	0.89	Temperature, wind speed, day-of-week	2010–2016
SDG&E Electricity Demand	0.80	Temperature, wind speed, day-of-week	2010–2016
WECC Path 8	0.83	Temperature, wind speed, day-of-week, Pacific Northwest hydropower	2010–2012
WECC Path 14	0.79	Temperature, wind speed, day-of-week, Pacific Northwest hydropower	2010–2012
WECC Path 3	0.63	Temperature, wind speed, day-of-week, Pacific Northwest hydropower	2010–2012
WECC Path 65	0.85	Temperature, wind speed, day-of-week, Pacific Northwest hydropower, Path 8, Path 14, Path 3	2010–2012
WECC Path 66	0.89	Temperature, wind speed, day-of-week, Pacific Northwest hydropower, Path 8, Path 14, Path 3	2010–2012
WECC Path 46	0.76	Temperature, wind speed, day-of-week, Path 65, Path 66	2010–2012
WECC Path 45	0.88	Temperature, wind speed, day-of-week, Path 46, Path 65, Path 66	2010–2012
WECC Path 24	0.84	Temperature, wind speed, day-of-week, Path 46, Path 65, Path 66	2010–2012
WECC Path 61	0.85	Temperature, wind speed, day-of-week, Path 46, Path 65, Path 66	2010–2012
WECC Path 42	0.90	Temperature, wind speed, day-of-week, Path 46, Path 65, Path 66	2010–2012
Pacific Northwest Hydropower	0.61	Streamflow	2003–2006
CAISO Hydropower	0.85	Streamflow	2001, 2005, 2010, 2011

3. Results & discussion

3.1. Validation of UC/ED formulation

This paper proceeds with a validation of the UC/ED model's ability to reproduce observed power system dynamics (in particular, wholesale electricity prices). Wholesale prices, which are driven by changes in supply and demand, can be viewed as aggregate measures of system performance (high prices can indicate scarcity, and low prices point to abundance). We focus on an extended period of drought that occurred in California over the years 2012–2016. During this period, in-state hydropower generation decreased by an average of 40% (Gleick, 2017), forcing the state to rely significantly more on electricity from natural gas power plants. There has been considerable interest in exploring the impacts of this recent drought on pollutant emissions (Hardin et al., 2017), as well as system costs and prices for retail electricity consumers (Gleick, 2017). Particularly when determining the latter, an understanding of impacts on wholesale electricity prices is necessary. Retail distribution companies in California (PG&E, SCE, and SDGE) all purchase electricity from the CAISO market. If the CAPOW model is able simulate observed wholesale electricity prices over 2012–2016 with accuracy, then the model could also be used to conduct controlled experiments designed to isolate the role of drought (and/or other hydrometorological extremes) on wholesale prices, revenues/costs for utilities, and, ultimately, retail prices for consumers. Natural gas price data used to validate the model (i.e. compare historical CAISO prices across the years 2012–2016) were obtained from EIA's natural gas hub dataset; although these data do not represent the exact price paid by

power plants, they do represent dynamic prices at major gas trading hubs. These day-to-day fluctuations in gas prices are extremely important to capture. EIA's data on the delivered price of natural gas for power plants is typically listed on a monthly/annual time step, which would not allow us to capture more short term, severe price spikes.

Fig. 5 compares observed daily average electricity prices in the CAISO market alongside prices simulated by the UC/ED model, showing strong agreement ($R^2 = 0.75$). For the purposes of validating the UC/ED model, we used historical records of temperatures, wind speeds, solar irradiance and streamflow at the sites listed in Table 2. Thus, discrepancies between observed and simulated prices are entirely due to the UC/ED formulation itself and/or discrepancies in fuel prices experienced. In general, the model accurately captures variation in electricity prices on daily time scales and above; although model outputs include hourly prices, hourly price dynamics (e.g., "peak" and "off-peak" patterns) are not as well represented. This is expected for a model reliant on a somewhat abstracted representation of the transmission network.

3.2. Validation of stochastic inputs

The UC/ED model's ability to capture more than 70% of daily variability in CAISO electricity prices suggests that coupling it with stochastic simulations of weather and hydrology would enable probabilistic assessment of a broad set of hydrometeorological risks in wholesale electricity markets. Before using CAPOW in this manner, however, the model's underlying "stochastic engine" (i.e., the suite of approaches used to simulate weather and hydrological variables and relevant power system inputs) must be validated.

3.2.1. Hydrometeorological variables

Given the large geographical extent considered, as well as the highly interconnected nature of the U.S. West Coast grid, it is important that stochastically generated meteorological and hydrological inputs exhibit the same statistical dependencies as the historical record. Fig. 6 shows correlation matrices calculated using historical data from the 17 GHCN stations and 7 NSRDB sites (top left), as well as historical data from the 85 stream gauges (bottom left). These are compared alongside correlation matrices calculated using 1000 years of corresponding stochastic data generated using the approaches described in section 2.3.

Lighter areas show positive correlation (two locations/variables that are more likely to both experience high/low values simultaneously); dark areas show negative correlations. In general, results show a high degree of fidelity between historical and simulated covariance across variables and space. For example, historical and simulated streamflow

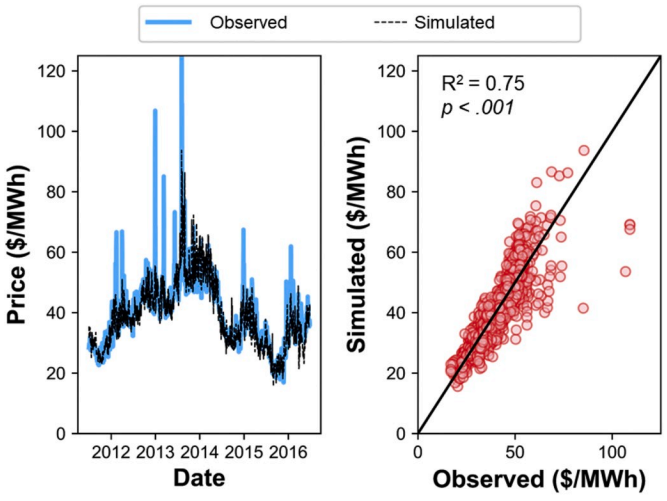

Fig. 5. Daily observed vs. simulated wholesale electricity prices in the CAISO market over the period 2012–2016.

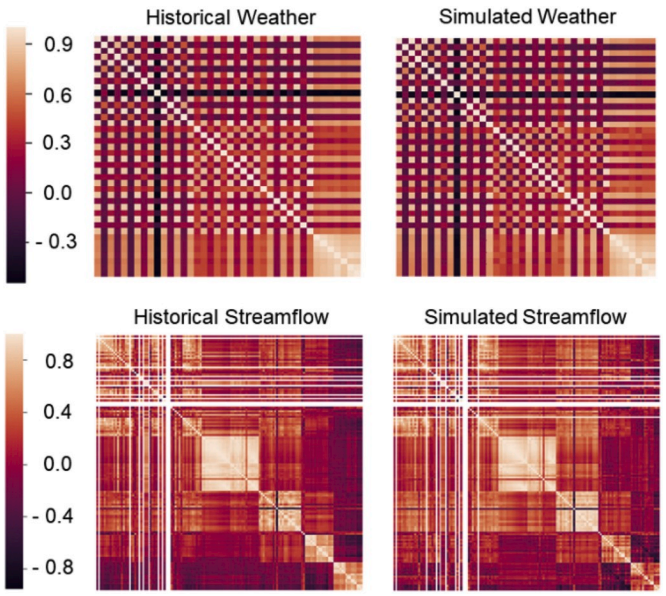

Fig. 6. Historical and simulated covariance matrices for weather variables (top) across the 17 GHCN stations and streamflow (bottom) across the 85 stream gauges considered. Pockets of high values in the bottom figures indicate stream gauges within the same watershed.

correlation matrices both show the same pockets of light values, which are associated with highly correlated stream gauges located within the same watershed. Overall, these results suggest that CAPOW, when run in stochastic mode, is able to capture spatial heterogeneities in weather and hydrological processes (e.g., the likelihood of experiencing high/low temperatures/wind speeds/irradiance/streamflow simultaneously at sites distributed across the entire region).

Equally important, the underlying stochastic engine of CAPOW is able to reproduce observed statistical moments (e.g., mean, standard deviation) in hydrometeorological conditions. Fig. 7 shows close agreement between historical and simulated temperatures and wind speeds across the 17 GHCN stations, in terms of percentile (1st, 50th, and 99th), while also demonstrating the stochastic model's ability to occasionally generate more extreme min/max values than the historical record.

In Fig. 8, a similar comparison is shown using streamflow data. Each panel includes historical (blue/red circles) and simulated (black line) values for each of the 85 stream gauges considered. Red circles represent gauges in California (mostly the Sierra Nevada Mountains) and blue circles represent gauges in the Pacific Northwest (mostly the Columbia River Basin). Each panel represents a different percentile (1st/50th/99th) as well as min/max values. Note that in some cases, negative values are shown. This is an artifact of our use of BPA's modified flow dataset, which consists of historical flows at gauge sites in the Columbia River Basin with modern human withdrawals applied. At certain gauge sites, this results in negative flow values (water is subtracted from reservoir storage). In general, results suggest close agreement between the distributions of historical and stochastically generated streamflow values, while also demonstrating the stochastic model's ability to occasionally generate more extreme min/max values than the historical record.

3.2.2. Power system inputs

A suite of models is used to translate raw temperatures, wind speeds, solar irradiance and streamflows into power system inputs, including multivariate regression (wind and solar power, electricity demand, system imports/exports) and hydrologic mass-balance operational models of reservoirs (hydropower). Coupled with our stochastic weather and streamflow generation techniques, these models yield realistic time

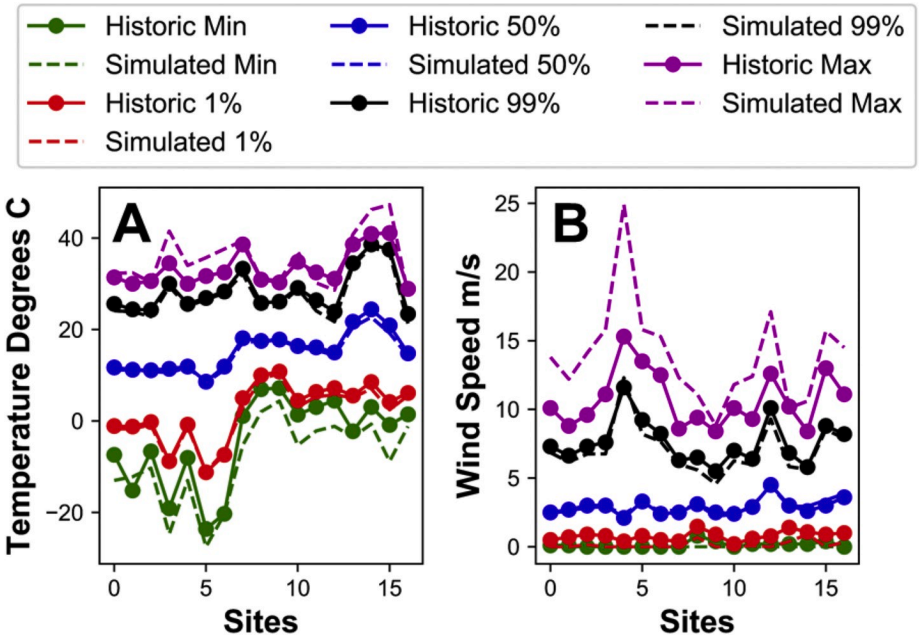

Fig. 7. Historical and simulated temperatures and wind speeds across the 17 GHCN stations, distinguished by percentile (1st, 50th, and 99th) and min/max value.

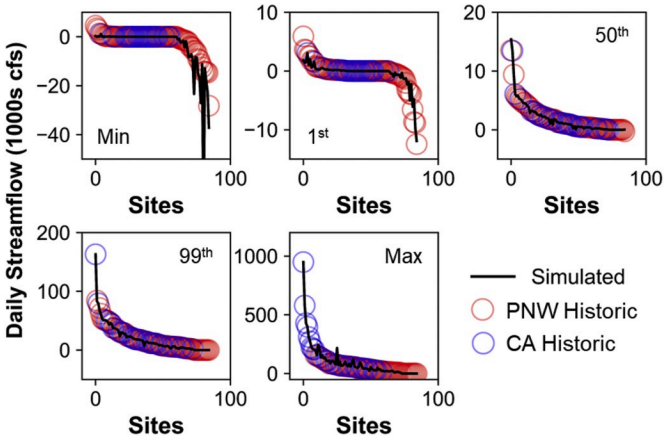

Fig. 8. Historical and simulated streamflows across the 85 stream gauges considered, distinguished by percentile (1st, 50th, and 99th) and min/max values.

series of power system inputs that mimic historical data on seasonal, daily and hour time scales (Table 2).

For example, Fig. 9 (panel A) shows historical (blue) and simulated (red) seasonality in wind power "capacity factor" (a unitless number between 0 and 1 corresponding to the average hourly output of a wind farm as a fraction of installed capacity), aggregated for the entire CAISO system. The simulated data is produced by coupling stochastically generated wind speeds at GHCN stations with a multivariate regression model of system-wide wind power availability based on wind speeds (Table 2), and then adding in a record of synthetic residuals (model errors). Results indicate alignment with historical data on a monthly basis, with highest capacity factors occurring in the summer and lowest during winter.

This approach is also able to reproduce hourly and daily time series characteristics for wind power production. Fig. 9 (panel B) shows close agreement between historical and simulated daily autocorrelation in wind power production, suggesting the model does an adequate job preserving any statistically significant "memory" in daily wind power production.

Fig. 9 (panel C) shows historical and simulated seasonality in solar power capacity for the CAISO system. The simulated data is produced by coupling stochastically generated solar irradiance (minus cloud effects) at seven NSRDB sites with a multivariate regression model of system-wide solar power availability based on site-specific irradiance. Results indicate alignment with historical data on a monthly basis, again with highest capacity factors occurring in the summer months and lowest during winter. This approach is also able to reproduce hourly and daily time series characteristics for solar power production. Fig. 9 (panel D) compares hourly capacity factors produced using historical irradiance data for a week in Summer 2006 alongside stochastically generated solar power data for the same calendar week (with differences being due to simulated cloud effects).

Consideration was also given to volume of simulations required to achieve statistical "convergence" between historical and simulated power system inputs. A primary motivating factor in developing the underlying framework of the CAPOW model is to explore the impacts of hydrometeorological uncertainty, especially extreme events, on power systems and electricity markets. To be useful in this regard, the stochastic engine of CAPOW, as well as the UC/ED model, must be run over a sufficiently large number of years to produce the kind of low probability, high magnitude "tail" events that are concerning to grid participants (e.g., episodes of extreme shortfalls or overabundance in supply). Considering the high computational requirements of the UC/ED model, which relies on mixed integer programming, a relevant question is "how many years are enough"?

Fig. 10 explores this question for the CAPOW model. Each panel shows data for a different input in the CAISO system: hydropower production, wind power production, load (electricity demand), and "net load", defined here as load minus total renewable energy (wind, solar, and hydropower) and resources considered to be "must run", like nuclear and geothermal. Net demand is an important metric because it represents the amount of electricity that would need to be met by dispatchable generators (coal and natural gas).

The colored lines measure the absolute difference between the historical record and synthetically generated values as a function of simulation volume. For example, in the bottom left panel (load), the red line tracks the difference between the historical record and stochastically simulated values, in terms of the 99th percentile of hourly electricity demand. At low simulation volumes, this difference starts at

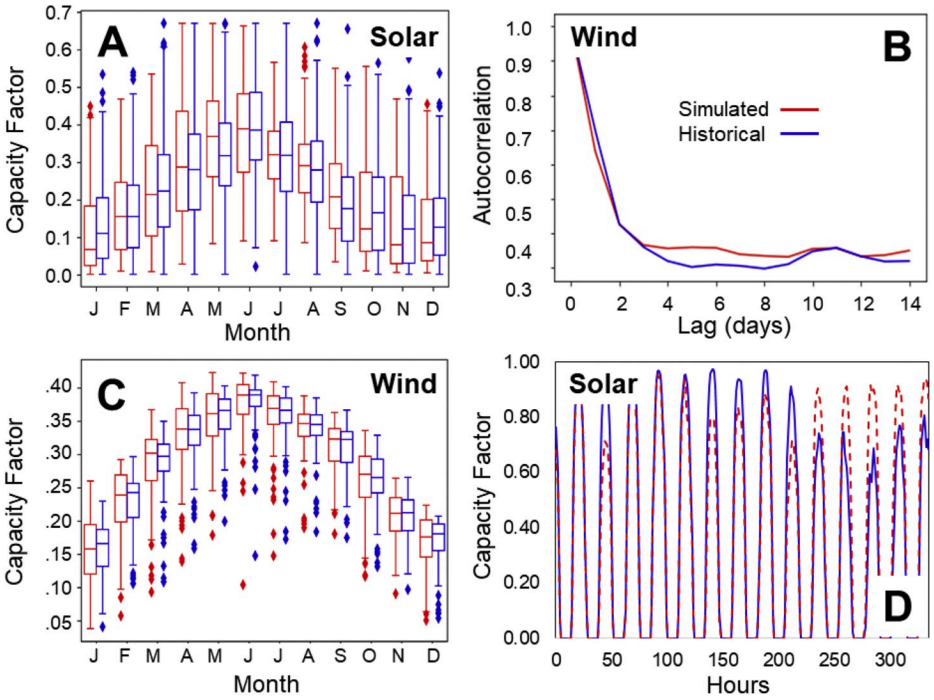

Fig. 9. (A) Capacity factors for aggregate wind power production in the CAISO market; (B) daily autocorrelation in daily wind power production in the CAISO market; (C) capacity factors for aggregate solar power production in the CAISO market; (D) hourly capacity factors for a sample period in the CAISO market. Red = simulated; blue = historical. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

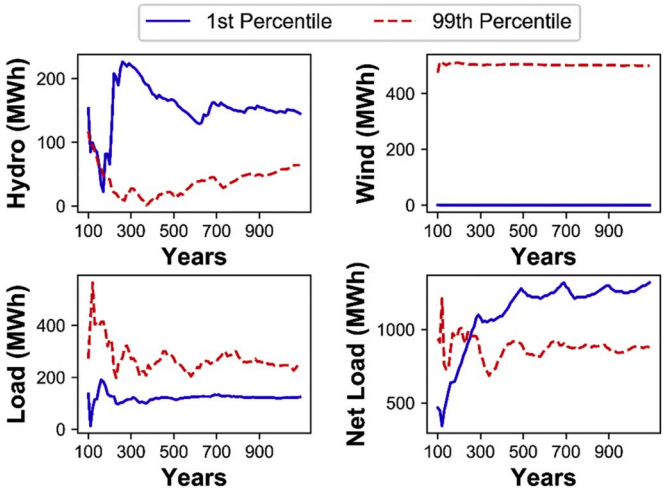

Fig. 10. Absolution deviations between historical and simulated inputs to the CAISO system in terms of their 1st and 99th percentile values, tracked as a function of the volume of simulation years.

around 280 MWh (average hourly demand in the CAISO market is more than 25,000 MWh, indicating an error of less than 1%). As the number of simulated years increases, the absolute difference first increases but then stabilizes, appearing to asymptotically approach a value close to 220 MWh. Stabilization occurs when increasing the number of simulation years has a negligible impact on the difference between historical and simulated values. Fig. 10 shows that simulations from CAPOW's stochastic engine tend to converge statistically after about 1000 years, suggesting this would be a reasonable lower bound on simulation volume to run through the UC/ED model.

Overall, our results suggest that CAPOW's stochastic engine is able to reproduce historical statistical characteristics across multiple hydrometeorological variables and power system inputs, needing approximately 1000 simulation years to achieve stable distributions. A final validation step is to evaluate whether the stochastic engine creates an expanded distribution of system states—in other words, does simulation over 1000 years cause extreme events outside the historical record to emerge from joint uncertainties in individual system processes? Without directly running the UC/ED model, a preliminary analysis of this kind can be conducted using net load as a metric of interest, since this typically correlates strongly with electricity prices and would be a key indicator of the potential for system shortfalls (extremely high net demand) and oversupply (extremely low net load).

Fig. 11 evaluates net load in the CAISO system under different scenarios. The shaded areas show the distribution of net load over the period 1953–2008, simulated using historical hydrometeorological

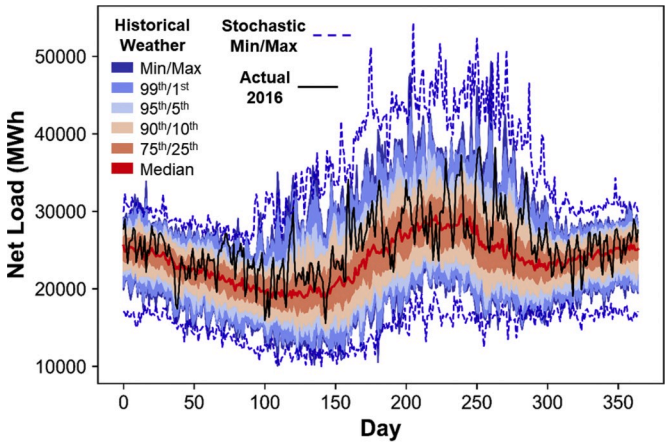

Fig. 11. Simulated net demand for the California wholesale market. Shaded areas represent uncertainty driven by historical (1953–2008) hydrometeorological time series. Actual historical net demand for a single year (2016) is also shown in black. Enveloping the simulations forced by historical hydrometeorology are minimum and maximum values acquired from 1000 synthetic runs produced by the stochastic model.

data. Colors correspond to different percentiles of net load (ranging from 1st to 99th) as well as the min/max values for this time period. Net load simulated using hydrometeorological data from 1953 to 2008 is then compared alongside actual historical net load recorded for a recent year, 2016, which is represented with a black line. For the most part, actual net load for 2016 is enveloped by the distribution of values simulated using 1953–2008 hydrometeorological data. Fig. 11 also shows minimum and maximum values acquired from 1000 years of synthetic runs produced by the stochastic engine of CAPOW (blue dotted lines). Min/max values produced by the stochastic engine suggest that the CAPOW model, by exploring joint uncertainties in hydrometeorolgical variables at sufficiently high simulation values, is able to access rare extreme events outside the historical record. The additional information provided by stochastic modeling appears to be especially valuable during late summer, when net load is the highest and the stochastic model produces maximum values that are considerably larger than the highest values simulated using weather and hydrology from 1953 to 2008. These more extreme synthetic values are likely to include rare but plausible "compound" events in which combinations of high electricity demand, and low renewable energy availability create extremely high net load, with associated risks for reliability and high market prices.

4. Conclusions

Despite growing interest in the potential vulnerabilities of bulk electric power systems to hydrometeorological variability (and extremes), there are few (if any) open source modelling packages capable of exploring this issue in a comprehensive manner. This paper presents a new model, CAPOW, which we specifically designed to explore the influence of joint uncertainties in temperatures, wind speeds, solar irradiance and streamflow on bulk power systems and wholesale electricity markets. CAPOW couples synthetic generation of hydrometeorological variables with simulation models of relevant infrastructure (dams, power plants), allowing for in depth exploration of the role of weather and hydrology on system outcomes. The model is free and downloadable via public online repositories.

The CAPOW model uses a topological representation of the conterminous U.S. West Coast power system to form a unit commitment and economic dispatch (UC/ED) model that simulates system operations and tracks performance (system costs, prices, etc.) on an hourly basis. When using historical weather and streamflow data as inputs to the model, it is able to capture 75% of the variability in daily electricity prices in the CAISO market. Although designed specifically with the U.S. West Coast in mind, the steps taken to construct CAPOW, as well as much of the code base, can be extended to other systems of interest. However, some critical functionalities may need to be added. For example, CAPOW does not currently represent thermal power plant curtailments due to inadequate cooling water supplies caused by low streamflows and high temperatures.

When run in stochastic mode, CAPOW couples the UC/ED model with a "stochastic engine" that creates synthetic records of temperatures, wind speeds, solar irradiance and streamflow for a group of 17 meteorological stations, 7 solar resource assessment sites, and 85 stream gauges distributed throughout the West Coast. Stochastically generated hydrometeorological variables are used to predict electricity demand (via temperatures, wind speeds), wind power production (via wind speeds), solar power production (via irradiance) and hydropower availability (via streamflows), which then drive the UC/ED model. The statistical properties (moments, cross correlations, time series characteristics) of synthetic data produced mirror those of the historical record, while also allowing for the generation of more extreme (but plausible) events. Exploring the joint uncertainty in relevant hydrometeorological variables is computationally tractable, with the statistics of stochastic simulations converging with the historical record after approximately 1000 simulation years. Overall, our framework –which is also easily transferrable across systems and geographic areas—simulates the operations of bulk electric power systems and wholesale markets at sufficient scales and resolutions to simulate system operations in a realistic way, and over sufficient time horizons to explore joint uncertainty across multiple, correlated variables of interest. As such, it should prove to be a valuable future resource for direct grid participants as well as the research community, particularly in answering questions related to the vulnerability of the grid to future changes in hydroclimate, as well as the sensitivity of variable renewable energy dominated grids to stationary hydrometeorological uncertainty.

Software and data availability

All code and data required to run the CAPOW model, as well as some documentation of the model, is available at https://github.com/romulus97/CAPOW_PY36 under the MIT free software license.

Funding sources

This research was supported by the National Science Foundation INFEWS programs, awards #1639268 (T2) and #1700082 (T1).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envsoft.2020.104667.

References

Andreadis, K., Clark, E., Wood, A., Hamlet, A., Lettenmaier, D., 2005. Twentieth-century drought in the conterminous United States. Am. Meteorol. Soc. 6, 985–1001.
Bain, D., Acker, T., 2018. Hydropower impacts on electrical system production costs in the Southwest United States. Energies 11, 368.
Boogert, A., Dupont, D., 2005. The nature of supply side effects on electricity prices: the impact of water temperature. Econ. Lett. 88, 121–125.
Brown, C., Lund, J., Cai, X., Reed, P., Zagona, E., Ostfeld, A., Hall, J., Characklis, G., Yu, W., Brekke, L., 2015. The future of water resources systems analysis: toward a scientific framework for sustainable water management. Water Resour. Res. 51, 6110–6124.
Chowdhury, K., Kern, J., Dang, T., Galelli, S., 2019. PowNet: a power systems analysis model for large-scale water-energy nexus studies. J. Open Softw. (in review).
Collins, S., Deane, P., Gallachoir, B., Pfenninger, S., Staffell, I., 2018. Impacts of inter-annual wind and solar variations on the European power system. Joule 2, 2076–2090.
Forster, H., Lilliestam, J., 2011. Modeling thermoelectric power generation in view of climate change. Reg. Environ. Change 4, 327–338.
Foster, B., Kern, J., Characklis, G., 2015. Mitigating hydrologic financial risk in hydropower generation using index-based financial instruments. Water Resour. Econ. 10, 45–67.
Franco, G., Sanstad, A., 2008. Climate change and electricity demand in California. Climatic Change 87.
Gleick, P., 2017. Impacts of California's Five-Year (2012-2016) Drought on Hydroelectricity Generation. Oakland, CA.
Hardin, E., AghaKouchak, A., Qomi, M., Madani, K., Tarroja, B., Zhou, Y., Yang, T., Samuselsen, S., 2017. California drought increases CO2 footprint of energy production. Sustain. Cities Soc. 28, 450–452.
Harto, C., Yan, Y., Demissie, Y., Elcock, D., Tidwell, V., Hallett, K., Macknick, J., Wigmosta, M., Tesfa, T., 2011. Analysis of Drought Impacts on Electricity Production in the Western and Texas Interconnections of the United States.
Henry, C., Pratson, L., 2016. Effects of environmental temperature change on the efficiency of coal- and natural gas-fired power plants. Environ. Sci. Technol. 50, 9764–9772.
Herman, J., Cohen, J., 2019. The operation of reservoirs in California (ORCA) model [WWW Document]. URL. https://github.com/jscohen4/orca. accessed 2.5.19.
Ho, J., Hobbs, B., Donohoo-Vallett, P., Xu, Q., Kasina, S., Park, S., Ouyang, Y., 2016. Planning Transmission for Uncertainty: Applications and Lessons for the Western Interconnection.
Jimenez, P., Vila-Guerau de Arellano, J., Gonzalez-Rouco, J., Navarro, J., Montavez, J., Garcia-Bustamante, E., Dudhia, J., 2011. The effect of heat waves and drought on surface wind circulations in the Northeast of the Iberian Peninsula during the summer of 2003. J. Clim. 24, 5416–5422.

Kern, J., Characklis, G., 2017. Evaluating the financial vulnerability of a major electric utility in the Southeastern U.S. To drought under climate uncertainty and an evolving generation mix. Environ. Sci. Technol. 55, 8815–8823.
Kern, J.D., Characklis, G.W., Foster, B.T., 2015. Natural gas price uncertainty and the cost-effectiveness of hedging against low hydropower revenues caused by drought. Water Resour. Res. 51, 2412–2427. https://doi.org/10.1002/2014WR016533.
Logan, B., 2015. Urgency at the nexus of food, energy, and water systems. Environ. Sci. Technol. 2, 149–150.
Mazdiyasni, O., AghaKouchak, A., 2015. Substantial increase in concurrent droughts and heatwaves in the United States. Proc. Natl. Acad. Sci. 112, 11484–11489.
Miara, A., Macknick, J., Vörösmarty, C., Tidwell, V., Newmark, R., Fekete, B., 2017. Climate and water resource change impacts and adaptation potential for US power supply. Nat. Clim. Change 7, 793–798.
Mkarov, Y., Etingov, P., Zhou, N., Ma, J., Samaan, N., Diao, R., Malhara, S., Guttromson, R., Du, P., Sastry, C., 2010. Analysis Methodology for Balancing Authority Cooperation in High Penetration of Variable Generation.
Najibi, N., Devineni, N., 2017. Recent trends in frequency and duration of global floods. Earth Syst. Dynam. 9, 757–783.
Nowak, K., Prairie, J., Rajagopalan, B., Lall, U., 2010. A nonparametric stochastic approach for multisite disaggregation of annual to daily streamflow. Water Resour. Res. 46.
Null, S., Viers, J., Mount, J., 2010. Hydrologic response and watershed sensitivity to climate warming in California's Sierra Nevada. PloS One 5.
Pandzžić, H., Dvorkin, Y., Wang, Y., Qiu, T., Kirschen, D., 2014. Effect of time resolution on unit commitment decisions in systems with high wind penetration. In: IEEE PES General Meeting. IEEE, National Harbor, MD.
Reed, P., Hadka, D., Herman, J., Kasprzyk, J., Kollat, J., 2013. Evolutionary multiobjective optimization in water resources: the past, present, and future. Adv. Water Resour. 51, 438–456.
Seel, J., Mills, A., Wiser, R., Deb, S., Asokkumar, A., Hassanzadeh, M., Aarabali, A., 2018. Impacts of High Variable Renewable Energy Futures on Wholesale Electricity Prices, and on Electric-Sector Decision Making.
Sengupta, M., Xie, Y., Lopez, A., Habte, A., Maclaurin, G., Shelby, J., 2018. The National solar radiation data base (NSRDB). Renew. Sustain. Energy Rev. 89, 51–60.
Staffell, I., Pfenninger, S., 2018. The increasing impact of weather on electricity supply and demand. Energy 145, 65–78.
Su, Y., Kern, J.D., Characklis, G.W., 2017. The impact of wind power growth and hydrological uncertainty on financial losses from oversupply events in hydropower-dominated systems. Appl. Energy 194. https://doi.org/10.1016/j.apenergy.2017.02.067.
Tarroja, B., AghaKouchak, A., Samuelsen, S., 2016. Quantifying climate change impacts on hydropower generation and implications on electric grid greenhouse gas emissions and operation. Energy 111, 295–305.
Turner, S., Voisin, N., Fazio, J., Hua, D., Jourabchi, M., 2019. Compound climate events transform electrical power shortfall risk in the Pacific Northwest. Nat. Commun. 10.
US Environmental Protection Agency, 2018. Emissions and generation resource integrated database (eGRID) [WWW Document]. URL. http://www2.epa.gov/energy/egrid.
van Vliet, M., Wiberg, D., Leduc, S., Riahi, K., 2016. Power-generation system vulnerability and adaptation to changes in climate and water resources. Nat. Clim. Change 6, 375–380.
van Vliet, M., Yearsley, J., Ludwig, F., Vogel, S., Lettenmaier, D., Kabat, P., 2012. Vulnerability of US and European electricity supply to climate change. Nat. Clim. Change 2.
Voisin, N., Kintner-Meyer, M., Wu, D., Skaggs, R., Fu, T., Zhou, T., Nguyen, T., Kraucunas, I., 2018. Opportunities for joint water–energy management: sensitivity of the 2010 western U.S. Electricity grid operations to climate oscillations. Bull. Am. Meteorol. Soc. 299–312.
Western Electricity Coordinating Council, 2016. 2016 Power Supply Assessment.
Western Electricity Coordinating Council System Adequacy Planning Department, 2015. Release Notes for the WECC 2024 Common Case. Salt Lake City, UT.
Woodhouse, C., Pederson, G., Morino, K., McAfee, S., McCabe, G., 2016. Increasing influence of air temperature on upper Colorado River streamflow. Geophys. Res. Lett. 43, 2174–2181.