Case Study

# Need for Process Based Empirical Models for Water Quality Management: Salinity Management in the Delaware River Basin

Eliot S. Meyer[1]; Daniel P. Sheer, M.ASCE[2]; Paul V. Rush[3]; Richard M. Vogel, Dist.M.ASCE[4]; and Hannah E. Billian[5]

**Abstract:** Managing salinity in the Upper Delaware Estuary is an important operational goal within the Delaware River Basin (DRB). High salinity concentrations can create water quality and operational challenges which increase treatment costs for downstream water utilities and cause ecological damage. This study reviews the advantages and limitations of process based empirical models (PBEM) as an alternative to complex hydrodynamic models or statistical models (i.e., multivariate regression) for salinity management. PBEMs involve choosing a parsimonious form of equation(s) that logically reproduces important physical relationships. A PBEM was developed to model specific conductivity (SC) (proxy for salinity) at three locations within the DRB more than 50 years. The resulting models explain most of the variations in historic SC and give comparable performance to a much more complex hydrodynamic model. The PBEM was then combined with streamflow, tidal forecasts, and an error model to develop an operational tool for assessing salinity impacts of potential reservoir releases and for generating ensemble forecasts of chlorinity. The authors also document how such ensemble forecasts can be employed to generate probabilistic forecasts of future salinity levels under various water resource system operating assumptions. **DOI: 10.1061/(ASCE)WR.1943-5452.0001260.** 



## Introduction

More than 15 million people rely on water resources from the Delaware River Basin (DRB), and the mainstem of the Delaware River is the longest undammed river east of the Mississippi in the United States (Rupert 2014). In managing this large and important watershed, the Delaware River Basin Commission (DRBC) must contend with salinity concerns at the outflow of the Delaware River into the Delaware Bay.

Due to a wide range of operational, financial, and ecological concerns, it is useful to be able to predict one component of salinity (chlorinity) in the Delaware Estuary. Whereas overall water quality has improved since the enactment of the Clean Water Act in 1972, managing operations of upstream reservoirs to improve water quality is still an important objective (Kauffman et al. 2011; Rupert 2014). Elevated source water chlorinity can cause taste concerns and corrosivity challenges for drinking water utilities including Philadelphia and its suburbs, and increase the treatment and maintenance costs for municipal and industrial water suppliers (Crittenden 2005). High chlorinity can also affect the maintenance cost for facilities that use the Delaware Estuary water for cooling. Periods of high chlorinity in the Delaware Estuary can have substantial environmental impacts, including damage to sensitive ecosystems (Bernhard et al. 2005; Gallegos and Jordan 2002; Paerl 1988; Powell et al. 1992). Preventing the chlorinity at Philadelphia Water's Delaware River intake from exceeding the EPA's secondary drinking water standard of 250 mg/L chloride is an important management objective. Rules that direct additional releases from upstream reservoirs when high chlorinity conditions exist are used to ensure that the objective is met (EPA 2018).

Releases from reservoirs in the DRB made to reduce chlorinity can impact the ability of these reservoirs to meet other management goals both within and external to the basin. Reservoir releases can result in reductions in water supply reliability, not only for Philadelphia, but also for upstream communities, including New York City, whose three reservoirs in the upper branches of the DRB supply up to 50% of the city's water needs (Rupert 2014). Water releases made to reduce downstream chlorinity also lower reservoir levels, impacting important cold-water fisheries and recreational boating by reducing the ability of water managers to maintain instream flows during appropriate times of the year. The ability to predict the impacts of reservoir releases on chlorinity will help promote the efficient use of available reservoir storage for salinity management.

A wide range of models exist for predicting chlorinity in rivers, ranging from one-, two-, and three-dimensional hydrodynamic models (EPA 2013; Ji 2008; Kim and Johnson 1998), to very simplistic bivariate (Vogel et al. 2005) and multivariate statistical models (Cohn et al. 1989; Helsel and Hirsch 2002). A hydrodynamic model is defined as a set of equations based on Newtonian continuum mechanics which describe the physical movement in time and space of water, its properties, and constituent parts (Hodges 2014). The goal of such models, much like analogous weather and climate models, is to describe the behavior of their subjects (i.e., a body of water, lake, river) in detail, given boundary conditions and a numerical approximation

[1]Principal Scientist, Hazen and Sawyer, One South Broad St., Suite 900, Philadelphia, PA 19107 (corresponding author). ORCID: https://orcid.org/0000-0002-2160-2364. Email: emeyer@hazenandsawyer.com

[2]President Emeritus, HydroLogics, Inc., 6098 Loventree Rd., Columbia, MD 21044.

[3]Deputy Commissioner, New York City Dept. of Environmental Protection, 7870 State Rd. 42, Grahamsville, NY 12740.

[4]Professor Emeritus, Dept. of Civil and Environmental Engineering, Tufts Univ., 200 College Ave., Medford, MA 02155.

[5]Hydraulic Modeler, HDR, 13621 Deerwater Dr., Germantown, MD 20874.

Note. This manuscript was submitted on March 5, 2019; approved on February 25, 2020; published online on July 3, 2020. Discussion period open until December 3, 2020; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Water Resources Planning and Management***, © ASCE, ISSN 0733-9496.

of the physical processes under examination. Because such models strive to explicitly model all relevant processes, they are often constrained by computational resources.

From a management and planning standpoint, there are several undesirable characteristics of hydrodynamic models: (1) calibration and validation difficulties, (2) long model run times, (3) the difficulty in using such models to develop short-term forecasts, and (4) the high cost of development. Hydrodynamic models are challenging to calibrate and validate because the data required is large (Hodges 2014), determining boundary conditions is difficult, and measurements rarely, if ever, span the wide-ranging, high-resolution output results produced by such models. Hydrodynamic models are so computationally intensive that the time it takes to run such models make their use in planning and operations difficult. For instance, the 1D model of the DRB has a run time on the order of 12 h. These long model run times and boundary condition uncertainties also prevent the use of hydrodynamic models for real time operational decisions incorporating short-term forecasts. Finally, hydrodynamic models are costly to develop because of their complexity.

Alternatively, *naïve* multivariate regression approaches can process many observed values to develop statistical relationships between any number of variables. Serago and Vogel (2018) document numerous advantages of regression approaches over other modeling approaches, including their ease of application, opportunities for graphical display, flexibility for fitting nonlinear relationships, ease of constructing confidence intervals for model predictions, and opportunities for uncertainty analysis. However, statistical regression can lead to spurious relationships between variables of interest if they are not accompanied by an understanding of the underlying physical processes that lead to the statistical relationships, and unlike physical models, these regressions cannot be extrapolated outside the range of values used for calibration (Hahn 1977).

Unlike hydrodynamic and nonphysical regression models, creating a process-based empirical model (PBEM) involves the development of a simple set of equations that logically describes the *shape* (functional form) of the overarching relationship between variables, and then calibrating the parameters of those equations to fit empirical data. Two defining characteristics of such models, which lead to many of their desirable attributes, are their parsimony and logical necessity (Boulding 1980; Quine 1966). PBEMs ensure logical necessity by matching the form of the model to the scientific understanding of the underlying behavior being modeled. These models are not new to water quality modeling: Henry's Law, D'Arcy's Law, and the Streeter–Phelps equation are just a few examples (Rinaldi and Soncini-Sessa 1978; Streeter and Phelps 1925). A more complex example is the SPARROW watershed modeling tool developed by the US Geological Survey intended for regional interpretation of water quality monitoring data, using a hybrid statistical/process-based approach to estimate pollutant sources and contaminant transport in watersheds and surface waters (Smith et al. 1997).

In contrast to hydrodynamic models, PBEMs are simple to calibrate, have short run times, can be more easily adapted for use in planning and generating forecasts, do not suffer from initial condition uncertainty, and cost little to develop. The run time of a parsimonious PBEM can be on the order of 15 s. This computational efficiency enables a PBEM to integrate into larger water resources management models to test alternative reservoir operations and simulate chlorinity over the hydrological record (80+ years). Because PBEMs can be limited to state variables that are monitored in real or near real time, such measurements can serve as the initial conditions for PBEM based forecasts. These forecasts are important for the potential use of PBEMs in ensemble predictions of future chlorinity given current conditions, allowing water managers to know quickly the potential chlorinity levels downstream of reservoir releases. In a hydrodynamic model, the existence of model error will cause errors in the model's estimate of initial conditions, which will introduce bias and otherwise affect the chlorinity ensembles. A PBEM whose state conditions are dependent only on variables that are monitored when the forecasts are made will have measurement errors already included in the initial conditions. Presumably, the measurement errors will be smaller than the model errors and result in the PBEM based ensembles performing better than the ensembles based on the hydrodynamic model. Finally, the cost for developing PBEMs is an order of magnitude lower than the cost of developing a hydrodynamic model, provided enough data are available; fortunately, such data are available for the Delaware Estuary. These motivating factors provide the impetus for developing PBEMs of chlorinity for several points in the Delaware River where long-term, semicontinuous monitoring data are available.

A postprocessor that adds stochastic error into the model output improves the results of a PBEM. To obtain model output that is truly representative of the data used to calibrate a model, it is necessary to add model error to the output of the model. This fact is discussed in detail by Farmer and and Vogel (2016), who show the important consequences of not adding model error to watershed simulation model output. In general, when error is not added to simulation model output, the resulting output will have lower variance (and all other upper moment ratios) than the observations used to calibrate the model. When the variance of the model output is too low, then predictions of extremes will generally be systematically biased, with the large (i.e., flood) predictions being systematically too small, and the small (i.e., drought) predictions being systematically too large. Vogel (2017) documents a generalized approach to converting a deterministic simulation model into a stochastic simulation model. It is now common practice in the field of flood forecasting to add forecast model error to forecast model outputs (Baatz et al. 2015; Bogner and Pappenberger 2011). Such postprocessing methods are reviewed in Li et al. (2017) and Vannitsem et al. (2018).

The overall objective of this study is to document the numerous advantages of a PBEM over alternative more complex hydrodynamic and less complex multivariate statistical models, for the purpose of water quality management. To achieve this overall objective, a PBEM is developed, implemented, and compared with alternative hydrodynamic models for the purpose of managing salinity in the Delaware Estuary. The overall approach is to: (1) develop a PBEM for specific conductivity (SC) (as a proxy for salinity) at three locations (Chester, Ben Franklin Bridge, and Reedy Island) in the Delaware Estuary, (2) fit the PBEM using historical SC, streamflow, and tide levels at the three locations, (3) assess and compare the performance of the model with a hydrodynamic model and a multivariate linear regression model, (4) develop an analysis of the model errors (innovations) to ensure model predictions reproduce the variance of the observations used to calibrate the model, and (5) incorporate the PBEM into a larger water resources management model of the Delaware River to aid in the development and testing of management alternatives for operation of the water resources of the Delaware Estuary, including the use of such models in making ensemble forecasts.

## Methods

### Data

To build the PBEM, sources of data for SC, inflows, tide levels, and wind speed and direction are identified. While not used in the final PBEM, the wind data are incorporated into a multivariate linear

regression model as a comparison with the PBEM and the hydrodynamic model (DYNHYD).

SC data for Ben Franklin Bridge, Chester, and Reedy Island are available from NOAA starting on November 8, 1963, October 1, 1963, and October 3, 1963, respectively (NOAA 2018). The monitoring stations are turned off in the winter (due to the possibility of ice), and there are other missing values, but otherwise, the data are largely complete. Fig. 1 shows the locations of the three main stations.

SC is related to the parameter of interest (chlorinity) by a nearly linear relationship (Cox et al. 1967), especially at higher levels of chlorinity where the majority of the conductivity is the result of ocean water diluting fresh water inflows. SC is also a function of other ions in the water, and for nearly fresh water these may be important. However, at the chlorine concentrations of interest for this model, SC is an acceptable surrogate. Model output in units of SC can be converted to chlorinity by linear interpolation of the table independently developed by the DRBC, which reflects the SC–chlorinity relationship from Cox et al. (1967).

Streamflow data are from USGS gages at the Schuylkill River at Philadelphia (USGS 01474500) and the Delaware River at Trenton, NJ (USGS 01463500) (USGS 2016). For the Reedy Island salinity model, additional incremental inflows downstream of the Schuylkill


**Fig. 1.** (Color) Map of Delaware Bay with locations of specific conductivity data. [Base map by Esri, HERE, Garmin, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), © OpenStreetMap contributors, and the GIS User Community.].

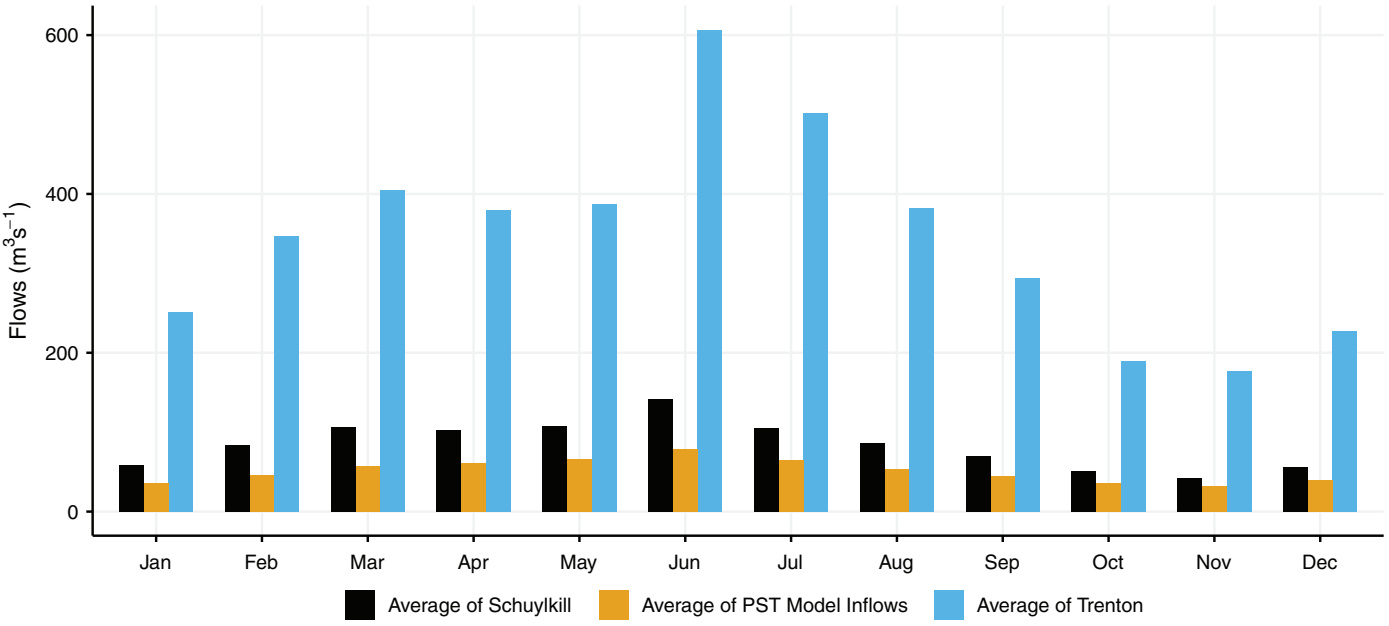


**Fig. 2.** (Color) Average monthly inflows of three sources of inflows considered in the PBEM and multivariate linear regression models.

River are added; the inflow record is taken from the Delaware OASIS water management model. Fig. 2 illustrates the relative contribution of each source of flow. The monthly flows from the Delaware River at Trenton, NJ (gray) dominate the flows, especially during the summer months. The *Model Inflows* are representative of the incremental inflow downstream of the confluence between the Schuylkill and the Delaware Rivers and upstream of Reedy Island and are derived from an existing DRBC model called the Planning Support Tool (PST) (Delaware River Basin Commission 2015).

To model the tidal impacts on SC, hourly water level data at Reedy Island is used as a proxy for the relative strength of the tide on any given day. These hourly water-level data extend from 1980 to present. Because higher than average high tides and higher than average low tides results in higher SC, aggregating water levels to the daily time step using the mean daily value captures the tidal behavior that influences SC. The relationship between water levels at Reedy Island and Atlantic City, NJ (the closest station to Reedy Island with long-term daily water-level data) was used to extend the water-level data before 1980. This linear relationship between the two locations explains more than 67% of the variance in Reedy Island mean water level.

Wind speed and direction were also investigated as possible contributors to changes in salinity. Wind speed along the estuary has been shown to affect salinity over longer temporal scales (i.e., monthly) (Ross et al. 2015). Historical hourly wind speed and direction at the Philadelphia airport from 1965–2005 were averaged, and the estuary component obtained; however, when compared with model goodness-of-fit without a wind component, the relationship was not statistically significant for any of the stations. Therefore, to maintain model parsimony, wind impacts were considered negligible. However, wind was included as a variable in the multivariate linear regression model discussed below and was found to be statistically significant.

### PBEM Description and Development

In creating a PBEM, the objective is to ensure that the form of the model being fit reflects the expected large-scale relationships between parameters based on a broad understanding of the dynamics of the underlying system. This differs from simple empirical models where the form of the model to be fit is often arbitrary (e.g., linear, exponential, log-linear). The process of creating a useful and credible PBEM of the relationship between inflow and SC begins by defining the shape of the relationship which results from the following physical logical processes which the PBEM must reproduce:

1. SC at a point in the estuary is the result of dilution of seawater by freshwater.
2. If there were no inflow, the SC would be that of seawater.
3. At high enough flows, the freshwater advection will overwhelm the tidal and molecular upstream transport of salt and the SC will match that of the inflows.
4. If inflow were to be constant for a long period, then an *equilibrium* would develop, and the SC would be constant.
5. This equilibrium concentration would be a function of flow.
6. When actual concentration differs from the equilibrium concentration on a given day, the actual concentration changes to decrease the difference, likely in a manner described by exponential decay of the difference.

One through six comprise most of the logical necessity needed for this PBEM. If the Delaware Estuary is thought of as a reactor with two inflows (freshwater and seawater) and an overflow, the resulting salinity and water mass balance leads to the conclusion that:

1. The SC/inflow relationship for the equilibrium concentration should approximate an S curve with asymptotes at the two extremes of seawater chlorinity and freshwater chlorinity.

Furthermore, two processes are mostly responsible for moving water in the system—inflow and tide (molecular diffusion energy is almost certainly negligible in comparison). The tidal influence is dwarfed by streamflow when flows are high. This indicates that the forcing functions for the S curve are different at high flows and low flows, and that:

2. The S curve should not be symmetric; thus, a two or three parameter function is not likely to be adequate to match the data over the entire range.

Finally, in the real world, estuaries are effective integrators of fluctuations in daily flow. Therefore, if we are to use the abstract concept of an equilibrium concentration as a part of a PBEM:

3. The equilibrium will need to be a function of a moving average of antecedent flows.

Although the functional form should logically be a nonsymmetrical S-curve, the SC data show that the highest SC ever recorded at the most downstream station is about 20,000 $\mu$S/cm, approximately half the SC of seawater (Miller et al. 1988). Thus, the concentrations fall entirely (or nearly entirely) on the concave section of the flow/SC relationship. Further, the maximum at the two other stations are 4,500 and 1,200 $\mu$S/cm. Thus, a concave curve (e.g., a power function with a negative exponent and a constant to scale the curve) is sufficient. The asymptote will not be zero, so a constant to shift the asymptote is required as well. This behavior is demonstrated visually in Fig. 3.

The effort to reproduce the behavior of the SC data also suggests that there is different behavior at low flows and high flows for all three stations. Therefore, a second power curve is incorporated to allow for different parameters at high and low flows and the two functions are interpolated at intermediate flows. The range of flows over which the interpolation occurs is different for each station. There are now 14 parameters to be chosen for each station. The PBEM model can be described by the following equations:

$$q_{n,t} = \frac{1}{n}\sum_{j=t-n}^{t} q_j \tag{1}$$

$$\kappa_{s,t} = \kappa_{s,t-1} + \min\left\{L_s, d_s \times \left[a_s \times (q_{s,t-1})^{P_s} + C_s - \kappa_{s,t-1}\right] + (\tau_s/q_{l,t}) \times T_t\right\} \tag{2}$$

$$\kappa_{l,t} = \kappa_{l,t-1} + \min\left\{L_l, d_l \times \left[a_l \times (q_{l,t-1})^{P_l} + C_l - \kappa_{l,t-1}\right] + (\tau_l/q_{l,t}) \times T_t\right\} \tag{3}$$

$$K_t = \begin{cases} \kappa_{s,t}, & q_{l,t} \le Q_l \\ \kappa_{s,t} \times \dfrac{(q_{l,t} - Q_l)}{(Q_2 - Q_l)} + \kappa_{l,t} \times \dfrac{(Q_2 - q_{l,t})}{(Q_2 - Q_l)}, & Q_l < q_{l,t} \le Q_2 \\ \kappa_{l,t}, & q_{l,t} \ge Q_2 \end{cases} \tag{4}$$

where, $q_j$ = sum of Trenton, Schuylkill, and PST model inflow on day j; $K_t$ = total predicted SC at time t; $\kappa_{s,t}$ = short-term moving average flow SC at time t; $\kappa_{l,t}$ = long-term moving average flow SC at time t; $q_{s,t}$ = short-term moving average flow at time t; $q_{l,t}$ = long-term moving average flow at time t; $Q_l$ = maximum flow at which $\kappa_{s,t}$ contributes to $K_t$; $Q_2$ = minimum flow at which $\kappa_{l,t}$ contributes to $K_t$; $L_s$ = short-term maximum 1-day change in SC; $L_l$ = long-term maximum 1-day change in SC; $d_s$ = decay coefficient for short-term moving average flow; $d_l$ = decay coefficient for long-term moving average flow; $a_s$ = constant for short-term moving average flow relationship; $a_l$ = constant for long-term moving average flow relationship; $P_s$ = power coefficient for short-term moving average flow relationship; $P_l$ = power coefficient for long-term moving average flow relationship; $C_s$ = offset for short-term moving average flow relationship; $C_l$ = offset for long-term moving average flow relationship; $\tau_s$ = tidal coefficient for short term moving average flow relationship; and $\tau_l$ = tidal coefficient for long term moving average flow relationship.

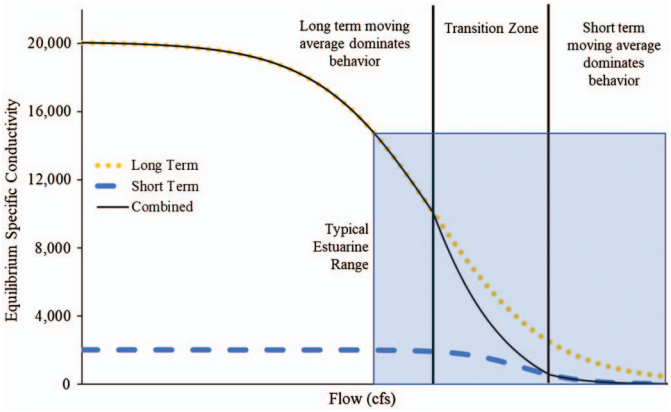


**Fig. 3.** (Color) Conceptual depiction of modeled salinity behavior related to flow which the PBEM attempts to reproduce.

### Hydrodynamic Model

In the DRB, the USACE, academic researchers, individual stakeholders, and the Delaware River Basin Commission (DRBC) have developed several hydrodynamic models. A nonexhaustive but representative list of the DRB models available include: (1) DYNHYD: a one-dimensional (1D) model based on the EPA modeling framework Water Quality Analysis Simulation Program (WASP5) (Ambrose et al. 1993), (2) a USACE three-dimensional hydrodynamic model (Kim and Johnson 1998), (3) Philadelphia Water Department (PWD) three-dimensional hydrodynamic model (Philadelphia Water Department 2015), and (4) DRBC two- and three-dimensional models based on the Environmental Fluid Dynamics Code (EFDC). These models are under various states of development and availability.

To compare the PBEM developed here with a representative hydrodynamic model, output from the 1D DYNHYD are used. This model was developed to evaluate the effectiveness of control strategies for polychlorinated biphenyls (PCBs) in the basin and used to develop total maximum daily loads (TMDLs) for the DRBC (Ambrose et al. 1993; EPA 2013). DYNHYD has also been used to evaluate impacts of management strategies on chlorinity.

The process of developing and calibrating a hydrodynamic model for water quality requires iteratively calibrating the hydrodynamic model of water movement (i.e., DYNHYD), linking it to a water quality model (i.e., TOXI5) (Suk and Collier 2003), calibrating the water quality model, and then iteratively fitting these two components. TOXI5 is part of the water quality analysis simulation program (WASP5) developed for the EPA (Ambrose et al. 1993). WASP5 is a modeling system with a modular structure to model the fate and transport of pollutants, and was designed to link with DYNHYD (Shoemaker et al. 1997).

The primary governing equations for DYNHYD are derived from conservation of momentum [Eq. (5)] and conservation of volume [Eq. (6)] at each node (Suk and Collier 2003). This model is calibrated using tidal heights in the Delaware Estuary.

$$\frac{\partial U}{\partial t} = -U\frac{\partial U}{\partial x} + a_{g,\lambda} + a_f + a_{w,\lambda} \tag{5}$$

$$\frac{\partial H}{\partial t} = -\frac{1}{B}\frac{\partial Q}{\partial x} \tag{6}$$

where, $\frac{\partial U}{\partial t}$ = local inertia term (m/s$^2$); = time (s); $U\frac{\partial U}{\partial x}$ = Benoulli acceleration (m/s$^2$); $x$ = channel distance (m); $a_{g,\lambda}$ = gravitational acceleration along dimension $\lambda$ (m/s$^2$); $a_f$ = frictional acceleration (m/s$^2$); $a_{w,\lambda}$ = wind stress acceleration along dimension $\lambda$ (m/s$^2$);

$\lambda$ = longitudinal axis (m); $H$ = elevation of water surface (m); B= channel width (m); and Q= flow rate ($m^3/s$).

Once this hydrodynamic model is fit, it is linked to the water quality model (TOXI5) to find dispersion coefficients and the appropriate time step. Because SC has extensive data records and varies temporally and spatially, it is used as the calibration parameter for linking to TOXI5. The main governing equations of TOXI5 are the mass balance equation for an infinitesimally small fluid volume [Eq. (7)], which is applied computationally using a finite-difference form of the equation [Eq. (8)] (Ambrose et al. 1993).

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x}(U_x C) + \frac{\partial}{\partial y}(U_y C) + \frac{\partial}{\partial z}(U_z C) + \frac{\partial}{\partial x}\left(E_x \frac{\partial C}{\partial}\right) + \frac{\partial}{\partial y}\left(E_y \frac{\partial C}{\partial y}\right) + \frac{\partial}{\partial z}\left(E_z \frac{\partial C}{\partial z}\right) + S_L + S_B + S_K \quad (7)$$

$$\frac{\partial AC}{\partial t} = \frac{\partial}{\partial x}\left(U_x AC + E_x A \frac{\partial C}{\partial x}\right) + A(S_L + S_B + S_K) \quad (8)$$

where: $C$ = concentration of the water quality constituent ($g/m^3$); $t$ = time (days); $U_x$, $U_y$, $U_z$ = longitudinal, lateral, and vertical advective velocities (m/day); $E_x$, $E_y$, $E_z$ = longitudinal, lateral, and vertical diffusion coefficients ($m^2$/day); $S_L$ = direct and diffuse loading rate ($g/m^3$-day); $S_B$ = boundary loading rate ($g/m^3$-day); $S_K$ = total kinetic transformation rate ($g/m^3$-day); and $A$ = cross-sectional area ($m^2$).

### Multivariate Linear Regression

Another computationally efficient strategy for modeling the type of environmental problems is a multivariate linear regression, which was developed to compare with the PBEM and the hydrodynamic model. A multivariate linear regression model was built for each of the three stations using the same data as for the PBEM, including Schuylkill and Delaware River flows, water level data at Reedy Island, and wind data at the Philadelphia airport as the independent variables, and SC at Chester, Ben Franklin Bridge, and Reedy Island as the dependent variables. The general form of the equation [Eq. (9)] uses several of the same variables of the PBEM for simplicity and to enable a more direct comparison between the fits of the linear regression model and the PBEM. Because streamflow and SC, like many water quality constituents, are lognormally distributed, these variables were log transformed. The coefficients are fit using ordinary least squares (OLS)

$$\ln(K_t) = c_{q_s} \ln(q_s) + c_{q_l} \ln(q_l) + c_T T + c_W W \quad (9)$$

where, $K_t$ = specific conductivity on day t; $q_s$ = short-term moving average flow; $q_l$ = long-term moving average flow; $T$ = daily average water level at Reedy Island (representative of tide) (ft); and $W$ = wind at the Philadelphia airport.

### Statistical Model of Innovation Ratios

To add error to simulation model output, a power-law regression model for predicting the model errors is constructed by exploiting their high level of persistence. To improve model performance, we developed a statistical model of the historical innovation ratios

$$I_t = \frac{S_t}{O_t} \quad (10)$$

where $O_t$ and $S_t$ = salinity observations and salinity simulations at time $t$, respectively, and $I_t$ represents the historical innovation ratio. Correcting the salinity model output using the innovation ratios ensures that the model will reproduce the statistical characteristics of the observations ($O_t$) upon which it is based (calibrated). Innovation ratios are used instead of the model error differences (i.e., residuals) because, in the authors' experience, innovation ratios exhibit symmetric or nonskewed distributions as opposed to error differences which typically exhibit highly skewed distributions.

A corrected daily salinity value on day t, denoted $\tilde{S}_t$, is calculated by dividing salinity model simulation output $S_t$, which represents the conditional mean salinity on day t, by an estimate of the innovation ratio for that day, $\hat{I}_t$:

$$\tilde{S}_t = \frac{S_t}{\hat{I}_t} \quad (11)$$

To implement Eq. (11), a model for the historical innovation ratios [Eq. (12)] is required. There are many ways in which one can develop a model for predicting the historical innovation ratios, including most of the approaches outlined in the recent review article by Li et al. (2017). In this study, the innovation ratios are predicted using an autoregressive model:

$$\ln(I_t) = \beta_0 + \sum_{i=1}^{p} \beta_i \ln(I_{t-i}) + \varepsilon_t \quad (12)$$

Importantly, innovation ratios and differences also tend to exhibit a high degree of heteroscedasticity, which is in part why we fit a power-law or log-log model in Eq. (12) to ensure approximately homoscedastic residuals ($\varepsilon_t$). Up to three lagged values of the innovation ratios are included, because in the authors' experience their inclusion enabled, to a first approximation, a very high level of explanatory power needed to simulate future ensembles of innovation ratios.

Multivariate OLS regression is used to estimate values of the model parameters in Eq. (12) by minimizing the sum of squares of the model residuals ($\varepsilon_t$). Ignoring the relatively small model error term, exponentiation of Eq. (12) leads to the following regression model estimates of the historical innovations, denoted $\hat{I}_t$, where the hats over variables denote that they have been estimated from the historical data:

$$\hat{I}_t = e^{\beta_0} I_{t-1}^{\beta_1} I_{t-2}^{\beta_2} I_{t-3}^{\beta_3} \quad (13)$$

Next, Eq. (13) is used to estimate the daily innovation ratios, denoted $\hat{I}_t$, which are then substituted into Eq. (11) to obtain corrected daily simulation values $\tilde{S}_t$. The net result is that the corrected simulation values $\tilde{S}_t$ should exhibit statistical characteristics which more closely resemble the observations $O_t$ than did the conditional mean daily simulations $S_t$ as is shown below.

### Ensemble Forecasts using PBEM

To demonstrate the ability of PBEMs to generate ensemble predictions, the model was integrated into a water resources systems model called the NYC Operations Support Tool (OST). New York City uses this tool to support operators managing the reservoirs in the NYC water supply system (Porter et al. 2015). The New York City Department of Environmental Protection (NYC-DEP) originally developed the OST as a method of controlling turbidity in the water supply (Weiss et al. 2013). This tool has been called "one of the most advanced and complex support tools for water supply operations in the world" (National Academies of Sciences, Engineering 2018), and it is integrated within the larger Delaware basin model maintained by the Delaware River Basin Commission. Because of the complexity of the system, and various flow, water

quality, and ancillary service demands, the OST must integrate a large amount of hydrological, meteorological, water quality, and other data. Furthermore, it must be computationally efficient enough to be used in real time. The OST model is used to manage daily operations, and in the preparation of drought, water quality events, and other events (Porter et al. 2015).

Whereas this integration of models is only included for demonstration purposes, a PBEM for salinity or other water quality measure would enable managers to make more informed operational decisions through forecasts. To integrate the PBEM, including the error correction (innovation) model, into the OST, the model was converted into operations control language (OCL), which is used by the underlying water resources model of the OST (National Academies of Sciences, Engineering 2018). The OST is then able to produce an ensemble forecast of salinity, limited to 52 traces of salinity based on ensemble forecasts of river flow from the National Weather Service for the upper basin and using the Hirsch forecast for the lower basin (Hirsch 1981; National Academies of Sciences, Engineering 2018). A set of ensemble forecasts for each of the three locations was developed using this flow forecast technique used by the OST, combined with the SC PBEM and the relationship between SC and chlorinity.

**Table 1.** Estimated PBEM model parameters for the three sites

| | SC station | | |
|---|---|---|---|
| Variable | Ben Franklin | Chester | Reedy |
| $N_s$ | 10 | 15 | 15 |
| $N_l$ | 30 | 45 | 90 |
| $Q_1$ ($ft^3/s$) | 2,800 | 4,000 | 5,000 |
| $Q_2$ ($ft^3/s$) | 4,800 | 5,250 | 9,000 |
| $a_s$ | 355,000 | $2.03 \times 10^{14}$ | 425,000 |
| $P_s$ | −0.8998 | −3.042 | −0.3484 |
| $C_s$ | 132.08 | 198.74 | −8,722.22 |
| $d_s$ | 0.19 | 0.96 | 0.99 |
| $L_s$ | 3 | 6 | 966.23 |
| $\tau_s$ | 35,312.54 | 2,517.34 | 1,000 |
| $a_l$ | 2,600,000 | $1.9 \times 10^{12}$ | 985,000 |
| $P_l$ | −1.06 | −2.54 | −0.46 |
| $C_l$ | 118.2 | 230 | −5,112.09 |
| $d_l$ | 0.0426 | 0.99 | 0.99 |
| $L_l$ | 5.5 | 28 | 1,500 |
| $\tau_l$ | −2,544.23 | 908.45 | −525,605 |

## Results and Discussion

### PBEM and Linear Regression Calibration

After the functional form of the PBEM model has been determined, the number of days and flow ranges for the short-term moving average and the long-term moving averages were determined as described in step 9 of the PBEM development. Then, the 10 model parameters were fit initially with a two-part sum of squared errors minimization, one minimization for the low flow model, and one minimization for the high flow model (Fylstra et al. 1998). After adjusting the fit in an ad hoc manner to further minimize errors, particularly for high SC events, a final goodness-of-fit analysis was performed using model residuals as described in the next section. Table 1 provides a summary of the fitted model parameters.

Similarly, the multivariate linear regression model used as a baseline of model performance was fit using OLS (Table 2). The moving average flow values were logarithmically transformed to account for their approximately lognormal distribution, as was SC, except for the Reedy Island model, because at Reedy Island SC is closer to normally distributed than lognormally distributed.

### Model Goodness-of-Fit Comparisons

The overall goodness-of-fit of the fitted models is quantified in several ways. The Nash-Sutcliffe Efficiency (NSE), the coefficient of determination ($R^2$) and the percent bias were all used to quantify model fit. In addition to these quantitative measures, examination of time series plots and scatter plots of model innovations elucidate the quality of the model. For comparison, model fit characteristics of a run of the calibrated DYNHYD model are also included. Table 3 summarizes the quantitative measures of model fit, which

**Table 2.** Regression table for multivariate linear regression

| | *Dependent variable:* | | |
|---|---|---|---|
| | $\ln(SC_{obs})$ | | $SC_{obs}$ |
| | Ben Franklin | Chester | Reedy Island |
| $\ln(Flow_{10d})$ | −0.253*** | — | — |
| | (0.004) | — | — |
| $\ln(Flow_{15d})$ | — | −0.154*** | −3,950.533*** |
| | — | (0.011) | (36.182) |
| $\ln(Flow_{30d})$ | −0.117*** | −0.557*** | — |
| | (0.004) | (0.012) | — |
| $\ln(Flow_{90d})$ | — | — | −2,606.514*** |
| | — | — | (44.655) |
| Wind | −0.00001*** | −0.00001*** | −0.080*** |
| | (0.00000) | (0.00000) | (0.016) |
| Tide | −0.010*** | −0.028*** | 137.255*** |
| | (0.002) | (0.04) | (28.485) |
| Constant | 8.887*** | 12.574*** | 69,191.430*** |
| | (0.020) | (0.042) | (331.692) |
| Observations | 13,175 | 13,188 | 15,623 |
| $R^2$ | 0.749 | 0.699 | 0.748 |
| Adjusted $R^2$ | 0.749 | 0.699 | 0.748 |
| Residual standard error | 0.153 (df = 13,170) | 0.322 (df = 13,183) | 2,359.125 (df = 15,618) |
| F statistic | 9,816.039*** (df = 4; 13,170) | 7,652.451*** (df = 4; 13,183) | 11,613.510*** (df = 4; 15,618) |

Note: df = degrees of freedom. *p < 0.1; **p < 0.05; and ***p < 0.01.

**Table 3.** Summary of goodness-of-fit statistics for the process-based empirical model (PBEM) and a comparable hydrodynamic model (DYNHYD) and multivariate linear regression model (MLR) of salinity

| | Ben Franklin | | | Chester | | | Reedy Island | | |
|---|---|---|---|---|---|---|---|---|---|
| | PBEM | DYNHYD | MLR | PBEM | DYNHYD | MLR | PBEM | DYNHYD | MLR |
| NSE | 0.834 | 0.458 | 0.631 | 0.837 | 0.775 | 0.435 | 0.706 | 0.684 | 0.748 |
| $R^2$ | 0.834 | 0.566 | 0.632 | 0.838 | 0.858 | 0.436 | 0.714 | 0.774 | 0.748 |
| Percent bias | −0.1 | 11.7 | 0 | −0.8 | 25.1 | 0 | 3 | 12.3 | 0 |

illustrate that the simple processed based empirical model (PBEM) accurately fits the historic observed SC, especially at the Ben Franklin and Chester SC stations (NSE of 0.834 and 0.837, respectively). Because of the much wider range of SCs at Reedy Island, the PBEM model yields lower goodness-of-fit (NSE of 0.706), but this fit still represents a slight improvement over the DYNHYD Model (NSE of 0.684). The $R^2$ values of each PBEM model are comparable to the NSE values. In terms of percent bias, both the Ben Franklin and the Chester PBEM models have slight negative bias (−0.10 and −0.80, respectively), yet both led to much smaller bias than that of the DYNHYD model and the Reedy PBEM.

### *Comparisons Among Modeled Time Series*

Whereas the observed record of SC spans more than 50 years, two different periods of two-year duration are representative of the behavior of each model at all three locations (Fig. 4). Considering the Ben Franklin and Chester sites, all three models exhibit high salinity conditions at the beginning of the period of record, from November 1, 1964 to November 1, 1966. This period corresponded to an extremely low-flow period on the Delaware River. Furthermore, tide events associated with tropical storm Dora also coincide with these particularly high salinity events (Gordon 1965). The time series of Reedy Island salinity demonstrates the variability of salinity closer to the mouth of the Delaware Bay, and the increase in average salinity. Despite this complex variability, the PBEM still manages to capture the overall behavior of salinity over time.

A period of average flows and SC (November 1997 to November 1999) demonstrates how each model performs in a more common flow regime. At the Ben Franklin station, the PBEM and MLR models fit the observed low salinity values more accurately than DYNHYD, but even relatively minor high SC events are not captured by the multivariate linear regression (MLR) model.

In comparison with the DYNHYD model, the time series plots in Fig. 4 show that the performance of the PBEM is comparable

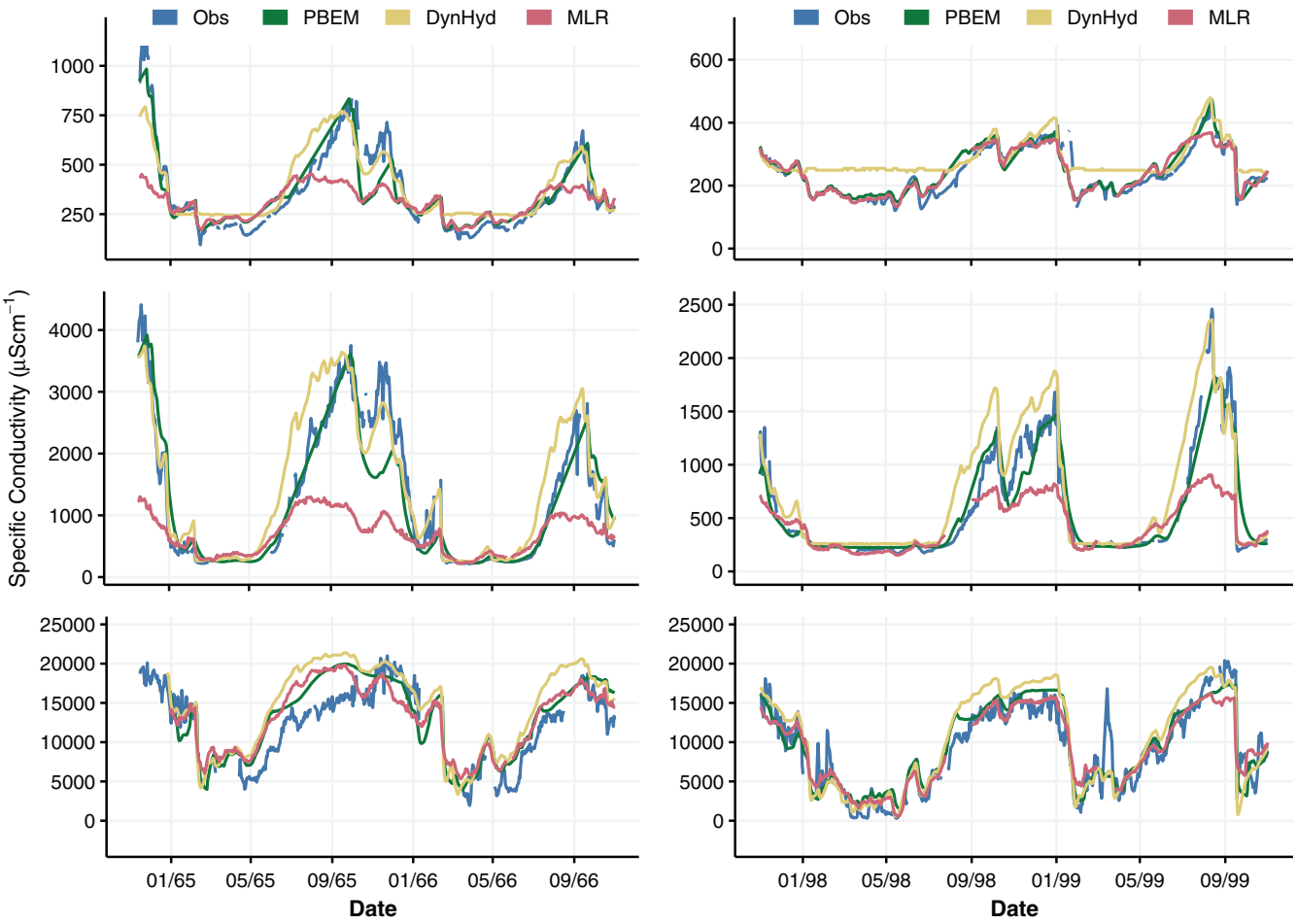


**Fig. 4.** (Color) Time series of specific conductivity of Ben Franklin (top), Chester (middle), and Reedy Island (bottom) for two different two-year periods.

with the more computationally complex hydrodynamic model. The Ben Franklin PBEM performs much better relative to the DYNHYD model. The DYNHYD model performs poorly under low salinity conditions, largely as a result of being calibrated for low flow events (i.e., higher salinities). In general, the PBEM and DYNHYD models are both able to match the overall timing and magnitude of SC events.

### Behavior of Innovation Ratios

Overall goodness-of-fit of the PBEM model ignores the importance of fit at high SC values. Due to operational concerns related to high salinity levels in the Delaware Estuary, it is of paramount importance for the model to predict high salinity events, especially salinities above 250 mg/L at Chester and Ben Franklin Bridge (1,120 $\mu$S/cm in SC).

As described in the previous section, characterizing the innovation ratios created by the model is an important part of utilizing the output of the model for operations and forecasting. In general, for an unbiased model, one expects innovation ratios to have expectation equal to unity over all conditions in which the model is needed. Fig. 5 illustrates the relationship between the innovation ratios and SC for the three stations and also illustrates a line at unity along with a nonparametric LOWESS smooth (Helsel and Hirsch 2002) fit to the innovation ratios. As expected, the average value of the innovation ratios is approximately equal to unity over nearly all SC values, and particularly so for the high SC values which are so important to reservoir operations. In Fig. 5, an innovation ratio above 1 indicates under prediction (S < O), whereas an innovation ratio below 1 indicates an over prediction (S > O). All three sets of innovation ratios exhibit similar overall patterns, with high innovations at low salinity, and correspondingly lower innovation ratios at high SC. However, the PBEM performs noticeably better in this regard at the Ben Franklin station, with a maximum innovation ratio of less than 2 and a minimum of approximately 0.5, indicating that the simulated values range from −50%–200% of the observed values at their most extreme. At the most consequential simulated values (high SC), this range is much narrower, approximately varying from −10%–20% of the observed SC. The PBEM performs similarly at Chester and Reedy Island, although at lower SC values the maximum innovation ratios for these two models are higher.

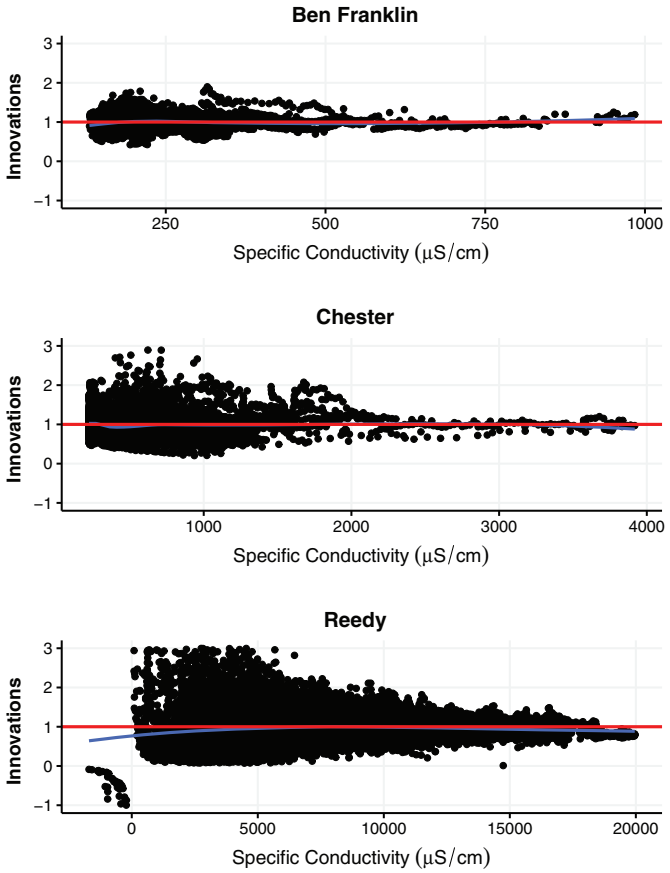


**Fig. 5.** (Color) Innovation ratios versus specific conductivity for three sites on the Delaware with a fitted LOWESS Smooth and a horizontal line at unity.

All three PBEM models of salinity can benefit from the inclusion of error modeling to improve the statistical characteristics of the model predictions. The values of the estimated model coefficients in Eqs. (12) and (13) for the three Delaware sites are given in Table 4 along with the goodness-of-fit statistic $R^2$. Chester and Reedy do not include a $\beta_3$ because the third lags were not found to be statistically significant. Heteroscedasticity associated with the innovation ratios in real space is apparent in Fig. 5, which is in part why the innovations were fit with a power law model for the innovation ratios in Eq. (12) (Helsel and Hirsch 2002). Fig. 6 illustrates the relationship between the natural logarithm of the innovation ratios and their lagged value for the three sites, providing little evidence of heteroscedasticity in the fitted power law relationships. Fig. 6 also shows the very high level of correlation and extremely good linear fit between the natural logarithm of the innovation ratios and their lagged value, for the three sites.

Evidence of improvement in the model output when integrating error analysis is documented in Table 5, which summarizes the mean and standard deviation of the observations $O_t$, conditional mean salinity simulations $S_t$, and the corrected daily simulation values, $\tilde{S}_t$. For all three sites, the conditional mean salinity simulations exhibit values of standard deviation which are generally lower than the standard deviation of the observations. For the Chester and Ben Franklin sites, the corrected daily simulation values exhibit both means and standard deviations which are much closer to the observations than the uncorrected daily simulation values. The exception is for the Reedy Island model, where the correction to the model output led to relatively poor agreement between the observations and the corrected simulations, which is likely due to two reasons: (1) the regression model for predicting the innovation ratios for the Reedy Island model has much lower goodness of fit than for the other two sites, and (2) there are numerous outlier simulation values which make it difficult to develop a reliable correction method for this site.

### Reproduction of Probability Distribution of Observations

Another aspect of the goodness-of-fit of model simulations involves the degree to which they reproduce the entire distribution

**Table 4.** Summary of model coefficients and $R^2$ for innovation regression model in Eq. (12)

| Variable | Ben Franklin | Chester | Reedy |
|---|---|---|---|
| $\hat{\beta}_0$ | −0.000795 | −0.0012 | 0.01277 |
| $\hat{\beta}_1$ | 1.3655 | 1.2118 | 1.10914 |
| $\hat{\beta}_2$ | −0.5143 | −0.2451 | −0.2123 |
| $\hat{\beta}_3$ | 0.1128 | N/A | N/A |
| $R^2$ | 94.71 | 95.13 | 84.31 |

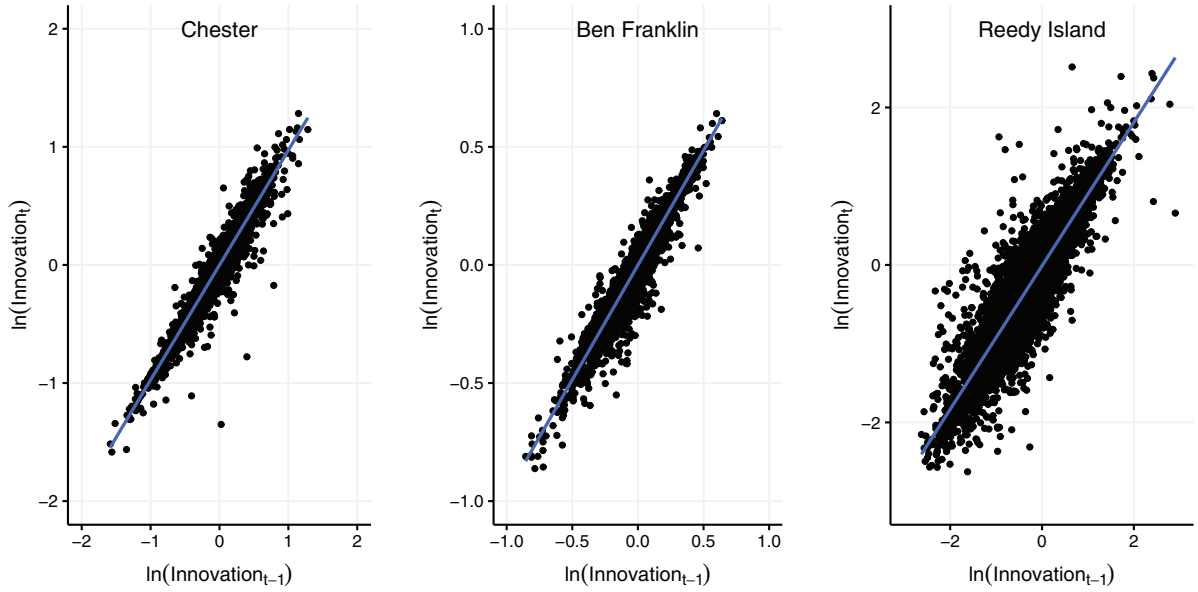


**Fig. 6.** (Color) Autoregression of innovation ratios for the Chester, Ben Franklin Bridge, and Reedy Island sites.

**Table 5.** Summary of mean and standard deviation of the observations $O_t$, conditional mean salinity simulations $S_t$, and the corrected daily simulation values, $\tilde{S}_t$

| | Ben Franklin | | Chester | | Reedy | |
|---|---|---|---|---|---|---|
| Variable | Mean | SD | Mean | SD | Mean | SD |
| Observations | 238.6 | 81.1 | 441.5 | 434 | 7,935 | 4,709 |
| Simulations without error | 230.1 | 68.5 | 409.1 | 357.1 | 8,153 | 4,231 |
| Simulations with error | 238.6 | 80.2 | 440.9 | 431.3 | 9,698 | 5,451 |

of observations upon which they are based. Fig. 7 illustrates boxplots of the model output and the observations used to fit the three models. The PBEM performs better than DYNHYD at all three locations. A pattern that emerges is that the PBEM and DYNHYD capture better the extremely high events, as demonstrated by the outliers on the high end. The linear regression model (LM) performs well in matching the median values. The boxplot of the Reedy Island models, however, demonstrates that the PBEM model underpredicts the frequency of low SC days, and slightly overpredicts the frequency of high SC days.

Importantly, Fig. 7 documents that the PBEM at each location recreates the observed distribution when corrected for errors using the innovation model, called *Corr.*in the figure. Fig. 7 demonstrates the value of using a relatively simple error model for enhancing the distributional goodness-of-fit of the model simulations to the original observations upon which each model is based.

### *Use of PBEM for Ensemble Predictions*

Using the PBEM for ensemble predictions of chlorinity demonstrates the computational efficiency of PBEMs and the utility of such PBEMs for use in water resources management. To demonstrate how the model could be used, a five-month forecast of chlorinity levels at Chester was generated for August 1, 2018 (Fig. 8), with an initial SC of 302 $\mu$S/cm (28 ppm chlorinity). Each band of color represents a range of potential outcomes. For example, the blue band represents the 90th to the 95th percentile, meaning that 5% of the traces occur in the blue region, 90% of the traces have lower chlorinity than the blue band, and 5% have a higher chlorinity. The dark green line is the median salinity for each day in the ensemble-based forecast. The black line is the observed chlorinity calculated from the observed SC at Chester, demonstrating that the forecast was able to provide a realistic estimate of future chlorinity. The probability-based forecasts illustrated in Fig. 8 for the Chester site are only possible with the computational efficiency of the underlying PBEM. Such probabilistic forecasts could provide valuable operational information for managing salinity on the Delaware River.

## Discussion

The method for fitting the PBEM model involves several iterations to fit both overall performance and performance at high salinity values, and it is for this reason that this PBEM does not always produce unbiased predictions. However, these biases have minimal impact on the effectiveness of such a model due to the way in which the model is used. Furthermore, the limited historic record of tide data at Reedy Point likely introduces additional uncertainty, which could be quantified in a future study using stochastic sensitivity analyses.

Whereas this PBEM can give some insights into salinity behavior and the physical processes at work, there are other models that are more suited to such tasks, such as the DYNHYD hydrodynamic model considered here. For management of water resources, however, having a parsimonious, computationally efficient model which exhibits logical necessity and can reproduce important statistical properties of the empirical observations is arguably more advantageous.

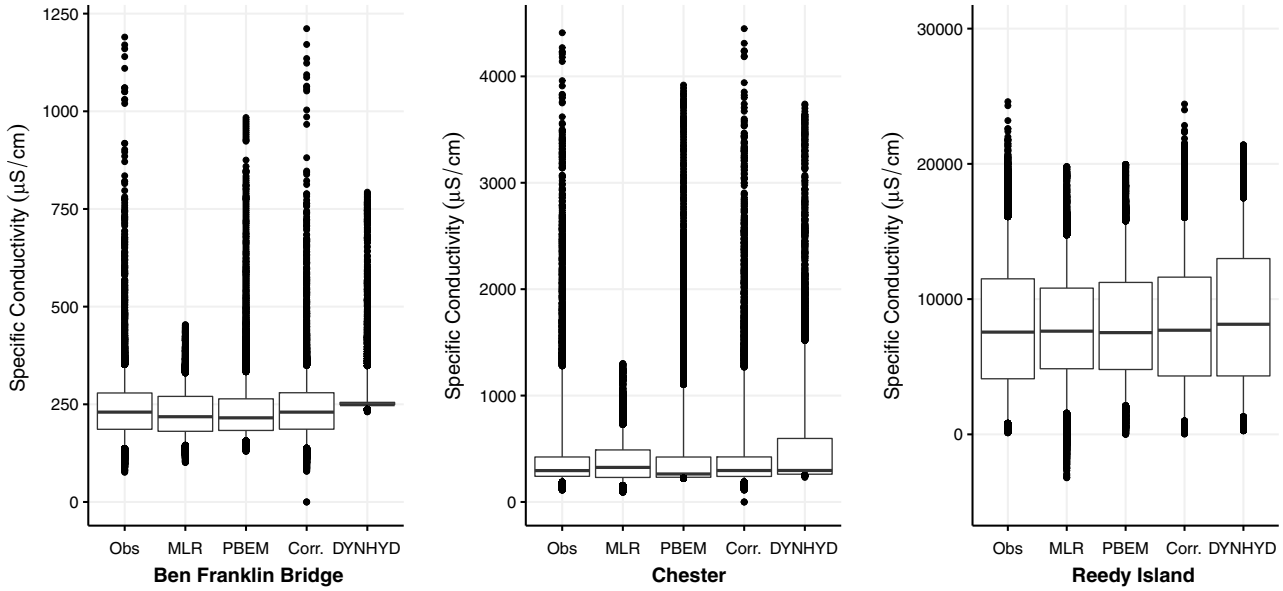


**Fig. 7.** Boxplots of modeled specific conductivity for the three sites using the MLR, PBEM, and DYNHYD models and the error corrected PBEM model denoted *Corr.*

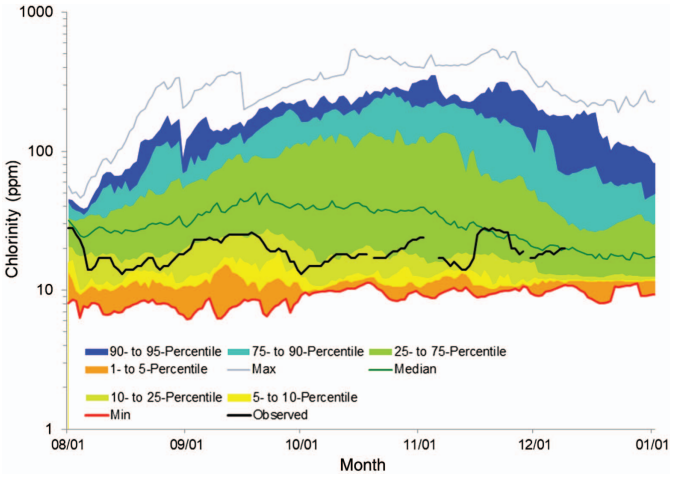


**Fig. 8.** (Color) Ensemble forecast for the Chester PBEM model beginning on August 1, 2018.

One potential application of this PBEM is for the simulation of operations of reservoirs in the Delaware River watershed over the historic hydrology. With the PBEM model incorporated into such an overall system model, alternative operations can be tested to investigate the impacts of operations on salinity and the salt front. Because salinity is a major concern within the Delaware basin, a determination of the impact of any new operational strategies on salinity concentrations should prevent detrimental operations from being enacted. Alternatively, knowledge of salinity impacts associated with reservoir releases will enable a more efficient use of the limited water resources by testing adjustment to operating protocols that impact salinity. For simulation uses of the PBEM, innovations (residuals) can be added back into the salinity results to reincorporate any exogenous influences on the salinity.

Another potential use for such a PBEM is in forecasting salinity levels using tide and flow forecasts. Such forecasts could be used to predict short-term changes in salinity and provide time for releases and changes to operations to reduce the impact of high salinity events, or alternatively to reduce the need for unnecessary releases designed for salinity mitigation. However, such forecasts would not be able to include or evaluate any potential changes to the basin geometry, and therefore the PBEM needs to be used in conjunction with more computationally intensive, hydrodynamic models that can simulate such changes.

## Conclusions

The overall goal of this study was to describe the numerous and varied benefits of a process-based empirical model (PBEM) over more detailed hydrodynamic models for the purpose of modeling SC (salinity proxy), and to understand the limitations of such models. We have reviewed the numerous advantages of a PBEM over a more detailed and physically based hydrodynamic model and a less detailed yet purely statistical model. A process-based empirical model was developed for modeling SC (a proxy for salinity) in the Delaware Estuary and applied to three stations maintained by NOAA. These three stations at Chester, Reedy Island, and Ben Franklin Bridge, are important for determining suitable operations of upstream reservoirs. Releases from these upstream reservoirs, such as Beltzville, Nockamixon, and Blue Marsh, are made in part to ensure that salinity levels at certain downstream points (e.g., Philadelphia water intakes) remain within suitable limits. Developing a parsimonious and accurate model of salinity in the Delaware Estuary that is computationally efficient is crucial to support the development of more complex operational models used to optimize decision making surrounding the timing and volume of numerous reservoir releases that are made upstream.

A framework was introduced for developing a relatively simple model for salinity management within the Delaware basin using both long-term and short-term moving average streamflow into the estuary and tidal effects expressed as average daily water levels at Reedy Point. The resulting operational model was shown to result in modeled salinity values that are equal to or better than the results of a much more complex hydrodynamic model (DYNHYD) and a multivariate linear regression model. The primary advantages of the

PBEM models include their parsimony combined with their ability to reproduce logical system constraints, which was shown to result in a computationally efficient model without sacrificing accuracy with respect to the reproduction of historical salinity values. Computational efficiency is of critical importance in operations and management problems because, when combined with a simple error model as described here, it enables the generation of ensembles of future salinity traces which are useful for obtaining probabilistic forecasts (Fig. 8). We also describe how a PBEM model can be combined with a simple stochastic error model resulting in significant improvements in the ability of the model to reproduce the statistical characteristics of the observations upon which it was based.

The PBEM model introduced in this paper compares favorably with much more complex hydrodynamic models of salinity such as the DYNHYD model and the simpler multivariate regression model. The authors hope that the salinity models introduced here will be incorporated into the future operations along the Delaware River, and that the framework for developing such parsimonious models introduced here will serve as a reminder and learning tool for developing models for other systems that can serve specific purposes. By focusing on the type of information most useful for operational and planning purposes, superfluous information can be stripped away to reduce the complexity of models. In an age of ever-increasing computational power, such a modeling process is still supremely useful for limiting the time needed to generate useful results.

## Data Availability Statement

All data used during the study are available online. Streamflow data are available from USGS (USGS 2016), SC data are available from NOAA (NOAA 2018), and tide data are available from NOAA (NOAA 2018). Additional model inflows are from the DRBC Planning Support Tool (PST) (Delaware River Basin Commission 2015). Models and code generated during the study are available from the corresponding author by request.

## References


Ambrose, R. B., T. A. Wool, and J. L. Martin. 1993. *The water quality analysis simulation program, WASP5 Part A: Model documentation*. Athens, GA: Environmental Research Laboratory.

Baatz, R., H. R. Bogena, H.-J. Franssen Hendricks, J. A. Huisman, C. Montzka, and H. Vereecken. 2015. "Dynamic aspects of soil water availability for isohydric plants: Focus on root hydraulic resistances." *Water Resour. Res.* 51 (11): 5974–5997. https://doi.org/10.1002/2014WR015608.

Bernhard, A. E., T. Donn, A. E. Giblin, and D. A. Stahl. 2005. "Loss of diversity of ammonia-oxidizing bacteria correlates with increasing salinity in an estuary system" *Environ. Microbiol.* 7 (9): 1289–1297. https://doi.org/10.1111/j.1462-2920.2005.00808.x.

Bogner, K., and F. Pappenberger. 2011. "Multiscale error analysis, correction, and predictive uncertainty estimation in a flood forecasting system." *Water Resour. Res.* 47: W07524. https://doi.org/10.1029/2010WR009137.

Boulding, K. E. 1980. "Science: Our common heritage." *Science* 207 (4433): 831–836. https://doi.org/10.1126/science.6766564.

Cohn, T. A., L. L. Delong, E. J. Gilroy, R. M. Hirsch, and D. K. Wells. 1989. "Estimating constituent loads." *Water Resour. Res.* 25 (5): 937–942. https://doi.org/10.1029/WR025i005p00937.

Cox, R. A., F. Culkin, and J. P. Riley. 1967. "The electrical conductivity/chlorinity relationship in natural sea water." *Deep-Sea Res. Oceanogr. Abstr.* 14 (2): 203–220. https://doi.org/10.1016/0011-7471(67)90006-X.

Crittenden, J. C. 2005. *Water treatment principles and design*. Hoboken, NJ: Wiley.

Delaware River Basin Commission. 2015. "Delaware River basin-planning support tool (DRB-PST)." Accessed March 1, 2019. https://www.state.nj.us/drbc/programs/flow/drbpst.html.

EPA. 2013. *Water quality analysis simulation program (WASP)*. Washington, DC: EPA.

EPA. 2018. *Secondary drinking water standards: Guidance for nuisance chemicals*. Washington, DC: EPA.

Farmer, W. H., and R. M. Vogel. 2016. "On the deterministic and stochastic use of hydrologic models." *Water Resour. Res.* 52 (51): 5974–5997. https://doi.org/10.1002/2016WR018977.Received.

Fylstra, D., L. Lasdon, J. Watson, and A. Waren. 1998. "Design and use of the Microsoft Excel solver." *Interfaces* 28 (5): 29–55. https://doi.org/10.1287/inte.28.5.29.

Gallegos, C. L., and T. E. Jordan. 2002. "Impact of the spring 2000 phytoplankton bloom in Chesapeake Bay on optical properties and light penetration in the Rhode River, Maryland." *Estuaries* 25 (4): 508–518. https://doi.org/10.1007/BF02804886.

Gordon, E. D. 1965. *The Hurricane Season of 1964*. Miami: US Weather Bureau Office.

Hahn, G. J. 1977. "The hazards of extrapolation in regression analysis." *J. Qual. Technol.* 9 (4): 159–165. https://doi.org/10.1080/00224065.1977.11980791.

Helsel, B. D. R., and R. M. Hirsch. 2002. "Statistical methods in water resources." In *Hydrologic analysis and interpretation*. Amsterdam, Netherlands: Elsevier.

Hirsch, R. M. 1981. "Stochastic hydrologic model for drought management." *J. Water Resour. Plann. Manage. Div.* 107 (WR2): 303–313. https://dx.doi.org/10.1016/j.wasec.2017.06.001.

Hodges, B. R. 2014. "Hydrodynamical modeling." *Surv. Geophys.* 22 (3): 179–263. https://doi.org/10.1023/A:1013779219578.

Ji, Z.-G. 2008. *Hydrodynamics and water quality: Modeling rivers, lakes, and estuaries*. Hoboken, NJ: Wiley.

Kauffman, G. J., A. R. Homsey, A. C. Belden, and J. R. Sanchez. 2011. "Water quality trends in the Delaware River Basin (USA) from 1980 to 2005." *Environ. Monit. Assess.* 177 (1–4): 193–225. https://doi.org/10.1007/s10661-010-1628-8.

Kim, K., and B. Johnson. 1998. *Assessment of channel deepening in the Delaware River and Bay: A three-dimensional numerical model study.* Philadelphia: Ft. Belvoir Defense Technical Information Center.

Li, W., Q. Duan, C. Miao, A. Ye, W. Gong, and Z. Di. 2017. "A review on statistical postprocessing methods for hydrometeorological ensemble forecasting." *Wiley Interdiscip. Rev.: Water* 4 (6): e1246. https://doi.org/10.1002/wat2.1246.

Miller, R. L., W. L. Bradford, and N. E. Peters. 1988. *Specific conductance: Theoretical considerations and application to analytical quality control*. Washington, DC: USGS.

National Academies of Sciences, Engineering. 2018. *Review of the New York City department of environmental protection operations support tool for water supply.* Washington, DC: National Academies Press.

NOAA (National Oceanic and Atmospheric Administration). 2018. "Tides and currents." Accessed January 12, 2018. https://tidesandcurrents.noaa.gov/.

Paerl, H. W. 1988. "Nuisance phytoplankton blooms in coastal, estuarine, and inland waters." *Limnol. Oceanogr.* 33 (4): 823–847. https://doi.org/10.4319/lo.1988.33.4part2.0823.

Philadelphia Water Department. 2015. "Green City, clean waters: Tidal waters water quality mode–Bacteria and dissolved oxygen." Accessed March 1, 2019. http://phillywatersheds.org/doc/WQ_Model_Complete_Report_FinalDigital_WITHAPPENDICES.pdf.

Porter, J. H., A. H. Matonse, and A. Frei. 2015. "The New York City operations support tool (OST): Managing water for millions of people in an era of changing climate and extreme hydrological events." *J. Extreme Events* 2 (2): 1550008. https://doi.org/10.1142/S2345737615500086.

Powell, E. N., J. D. Gauthier, E. A. Wilson, A. Nelson, R. R. Fay, and J. M. Brooks. 1992. "Oyster disease and climate change: Are yearly changes in Perkinsus marinus Parasitism in oysters (*Crassostrea virginica*) controlled by climatic cycles in the Gulf of Mexico?"

*Mar. Ecol.* 13 (3): 243–270. https://doi.org/10.1111/j.1439-0485.1992.tb00354.x.

Quine, W. V. O. 1966. "On simple theories in a complex world." In *The ways of paradox and other essays*. Cambridge, MA: Harvard University Press.

Rinaldi, S., and R. Soncini-Sessa. 1978. "Sensitivity analysis of generalized Streeter-Phelps models." *Adv. Water Resour.* 1 (3): 141–146. https://doi.org/10.1016/0309-1708(78)90024-6.

Ross, A. C., R. G. Najjar, M. Li, M. E. Mann, S. E. Ford, and B. Katz. 2015. "Sea-level rise and other influences on decadal-scale salinity variability in a coastal plain estuary." *Estuarine Coastal Shelf Sci.* 157 (May): 79–92. https://doi.org/10.1016/j.ecss.2015.01.022.

Rupert, C. D. 2014. "The Delaware River Basin Commission: A unique partnership." *Water Resour. Impact* 16 (5): 3–6.

Serago, J. M., and R. M. Vogel. 2018. "Parsimonious nonstationary flood frequency analysis." *Adv. Water Resour.* 112 (Nov): 1–16. https://doi.org/10.1016/j.advwatres.2017.11.026.

Shoemaker, L., M. Lahlou, M. Bryer, D. Kumar, and K. Kratt. 1997. *Compendium of tools for watershed assessment and TMDL development*. Rep. No. EPA841-B-97-006. Washington, DC: USEPA.

Smith, R. A., G. E. Schwarz, and R. B. Alexander (1997). "Regional interpretation of water quality monitoring data." *Water Resour. Res.* 33 (12): 2781–2798. https://doi.org/10.1029/97WR02171.

Streeter, H. W., and E. B. Phelps. 1925. *A study of the pollution and natural purification of the Ohio River*. Public Health Bulletin No. 146. New York: US Department of Health, Education, and Welfare.

Suk, N. S., and C. R. Collier. 2003. "DYNHYD5 hydrodynamic model (version 2.0) and chloride water quality model for the Delaware Estuary." Accessed March 1, 2019. https://www.nj.gov/drbc/library/documents/TMDL/HydroModelRptDec2003.pdf.

USGS. 2016. "National water information system data." Accessed January 1, 2019. http://waterdata.usgs.gov/nwis/.

Vannitsem, S., D. Wilks, and J. Messner. 2018. *Statistical postprocessing of ensemble forecasts*. Amsterdam, Netherlands: Elsevier.

Vogel, R. M. 2017. "Stochastic watershed models for hydrologic risk management." *Water Secur.* 1 (Jul): 28–35. https://doi.org/10.1016/j.wasec.2017.06.001.

Vogel, R. M., B. E. Rudolph, and R. P. Hooper. 2005. "Probabilistic behavior of water-quality loads." *J. Environ. Eng.* 131 (7): 1081–1089. https://doi.org/10.1061/(ASCE)0733-9372(2005)131:7(1081).

Weiss, W. J., G. W. Pyke, W. C. Becker, D. P. Sheer, R. K. Gelda, P. V. Rush, and T. L. Johnstone. 2013. "Integrated water quality-water supply modeling to support long-term planning." *J. Am. Water Works Assoc.* 105 (4): 57–58. https://doi.org/10.5942/jawwa.2013.105.0043.