



## RESEARCH ARTICLE

10.1029/2021WR030443

# Application of Recursive Estimation to Heat Tracing for Groundwater/Surface-Water Exchange

W. Anderson McAliley<sup>1,2</sup> , Frederick D. Day-Lewis<sup>3</sup> , David Rey<sup>4</sup> , Martin A. Briggs<sup>5</sup> , Allen M. Shapiro<sup>6</sup>, and Dale Werkema<sup>7</sup> 

### Key Points:

- Recursive filtering applied to heat tracing enables real-time estimation of groundwater/surface-water exchange
- Recursive filtering and smoothing applied to heat tracing improve estimation of groundwater/surface-water exchange
- Recursive filtering applied to heat tracing allows quantification of uncertainty in groundwater/surface water exchange estimates

<sup>1</sup>Student Contractor to the U.S. Geological Survey, Golden, CO, USA, <sup>2</sup>Now at U.S. Geological Survey, Water Mission Area, Golden, CO, USA, <sup>3</sup>Pacific Northwest National Laboratory, Earth Systems Science Division, Environmental Subsurface Science Group, Richland, WA, USA, <sup>4</sup>U.S. Geological Survey, Water Mission Area, Denver Federal Center, Denver, CO, USA, <sup>5</sup>U.S. Geological Survey, Water Mission Area, Storrs, CT, USA, <sup>6</sup>U.S. Geological Survey, Water Mission Area, Reston, VA, USA, <sup>7</sup>Environmental Protection Agency, Office of Research and Development, Center for Public Health and Environmental Assessment, Pacific Ecological Systems Division, Pacific Coast Ecology Branch, Newport, OR, USA

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

F. D. Day-Lewis,  
[frederick.day-lewis@pnnl.gov](mailto:frederick.day-lewis@pnnl.gov)

### Citation:

McAliley, W. A., Day-Lewis, F. D., Rey, D., Briggs, M. A., Shapiro, A. M., & Werkema, D. (2022). Application of recursive estimation to heat tracing for groundwater/surface-water exchange. *Water Resources Research*, 58, e2021WR030443. <https://doi.org/10.1029/2021WR030443>

Received 25 MAY 2021  
 Accepted 21 MAY 2022

### Author Contributions:

**Conceptualization:** Frederick D. Day-Lewis, Allen M. Shapiro  
**Data curation:** W. Anderson McAliley, David Rey  
**Formal analysis:** W. Anderson McAliley, Frederick D. Day-Lewis, Martin A. Briggs  
**Funding acquisition:** Frederick D. Day-Lewis, Dale Werkema

**Abstract** We present and demonstrate a recursive-estimation framework to infer groundwater/surface-water exchange based on temperature time series collected at different vertical depths below the sediment/water interface. We formulate the heat-transport problem as a state-space model (SSM), in which the spatial derivatives in the convection/conduction equation are approximated using finite differences. The SSM is calibrated to estimate time-varying specific discharge using the Extended Kalman Filter (EKF) and Extended Rauch-Tung-Striebel Smoother (ERTSS). Whereas the EKF is suited to real-time (“online”) applications and uses only the past and current measurements for estimation (filtering), the ERTSS is intended for near-real time or batch-processing (“offline”) applications and uses a window of data for batch estimation (smoothing). The two algorithms are demonstrated with synthetic and field-experimental data and are shown to be efficient and rapid for the estimation of time-varying flux over seasonal periods; further, the recursive approaches are effective in the presence of rapidly changing flux and (or) nonperiodic thermal boundary conditions, both of which are problematic for existing approaches to heat tracing of time-varying groundwater/surface-water exchange.

**Plain Language Summary** We present and demonstrate a new approach to infer groundwater/surface-water exchange based on temperature time series collected at different depths below the sediment/water interface. Algorithms for real-time (“online”) and “offline” applications are presented. The new algorithms are effective in the presence of rapidly changing flow across the sediment/water interface, which has posed challenges to existing approaches.

## 1. Introduction

Heat has been used as a tracer to estimate groundwater/surface-water exchange and hydraulic properties since pioneering work by Stallman in the 1960s (Anderson, 2005; Constantz, 2008; Stallman, 1965). The measurement technology (e.g., thermistors and thermocouples) is relatively inexpensive, long-term field installations are possible, and data analysis is straightforward. The increasing interest in hyporheic, hypolentic, and hypopaludal processes over the last two decades has led to burgeoning applications of heat tracing below the sediment water interface (e.g., Anderson, 2005; Briggs et al., 2013; Constantz, 2008; Rau et al., 2014). Tools for data analysis are well established and some are available within the public domain (Gordon et al., 2012; Koch et al., 2016; Voytek et al., 2014). Data analysis commonly involves (a) calibration of numerical or analytical models of heat transport to infer specific discharge associated with groundwater/surface-water exchange (Koch et al., 2016; Lapham, 1989; Voytek et al., 2014) or (b) analysis of the amplitude phase of signals extracted from temperature time series, which are then related to specific discharge through analytical models (Gordon et al., 2012; Schneidewind et al., 2016). We briefly review each of these two data analysis approaches as formulated within the existing tools *IDTempPro* and *VFLUX*.

The *IDTempPro* software package (Koch et al., 2016; Voytek et al., 2014) was written in C# for the Windows platform and provides a framework to calibrate a numerical model of heat transport to temperature time series collected at different depths below the sediment/water interface; this code allows for consideration of heterogeneity of hydraulic and thermal properties, complex (non-sinusoidal) boundary conditions, and abrupt changes

© 2022 Battelle Memorial Institute. This article has been contributed to by U.S. Government employees and their work is in the public domain in the USA. This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs License](https://creativecommons.org/licenses/by/4.0/), which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

**Investigation:** W. Anderson McAliley, Frederick D. Day-Lewis, Martin A. Briggs, Dale Werkema

**Methodology:** W. Anderson McAliley, Frederick D. Day-Lewis, David Rey, Allen M. Shapiro

**Project Administration:** Frederick D. Day-Lewis

**Resources:** Frederick D. Day-Lewis, Dale Werkema

**Software:** W. Anderson McAliley, Frederick D. Day-Lewis

**Supervision:** Frederick D. Day-Lewis, David Rey

**Validation:** W. Anderson McAliley, Martin A. Briggs

**Visualization:** W. Anderson McAliley, Frederick D. Day-Lewis, David Rey, Martin A. Briggs

**Writing – original draft:** W. Anderson McAliley, Frederick D. Day-Lewis, David Rey, Martin A. Briggs, Allen M. Shapiro, Dale Werkema

**Writing – review & editing:** W. Anderson McAliley, Frederick D. Day-Lewis, David Rey, Allen M. Shapiro

in exchange rate. Such abrupt changes can result from extreme weather driving rapid stage fluctuations, dam control on stream flows, or any flashy system. Model calibration is performed manually or by nonlinear regression. Estimation of time-varying discharge in *IDTempPro* is cumbersome and commonly entails the independent analysis of manually windowed subsets (days to week) of long-term seasonal data sets. *IDTempPro* provides a graphical interface for the *VS2DH* model. The finite difference scheme used by *VS2DH* is described in the following section.

The *VFLUX* MATLAB-based software package (Gordon et al., 2012; Irvine et al., 2015) provides a framework to extract signal amplitude and phase from temperature time series collected at different depths using dynamic harmonic regression (DHR). In DHR, a Kalman Filter (KF) algorithm (Young et al., 1999, 2004) is used to identify the time-varying amplitude and phase of a finite number of sinusoids that, superposed, optimize the fit to observed temperatures. Assuming an analytical model for heat transport, the extracted time-varying amplitude decay and (or) phase versus depth are related to discharge. Algorithms have been developed to use amplitude or phase or both amplitude and phase (e.g., Irvine et al., 2015). Compared to *IDTempPro*, *VFLUX* and related signal-extraction methods (e.g., Hatch et al., 2006; Keery et al., 2007; Schneidewind et al., 2016) are better suited to the estimation of time-varying discharge. Signal extraction, however, can prove problematic in the presence of episodic forcing (e.g., storm events) and during times of year when sinusoidal signals are naturally weak or when the vertical temperature gradient is small (Irvine et al., 2015; Lutz, 2012; Rau et al., 2015). In groundwater-upwelling zones, signals are also strongly attenuated with depth and may not be identified with DHR below approximately 15 cm under most conditions using typical temperature sensors (Briggs et al., 2016). In practice, DHR is subject to leakage and may impede recovery of abrupt changes in flux (Glose et al., 2021; Koch et al., 2016), but multifrequency signal extraction, for example, van Kampen et al. (2022) can mitigate these limitations. Reliance on simple analytical models to relate amplitude and (or) phase to discharge precludes consideration of heterogeneity, artificial heating sources, and the dependence of physical properties on temperature.

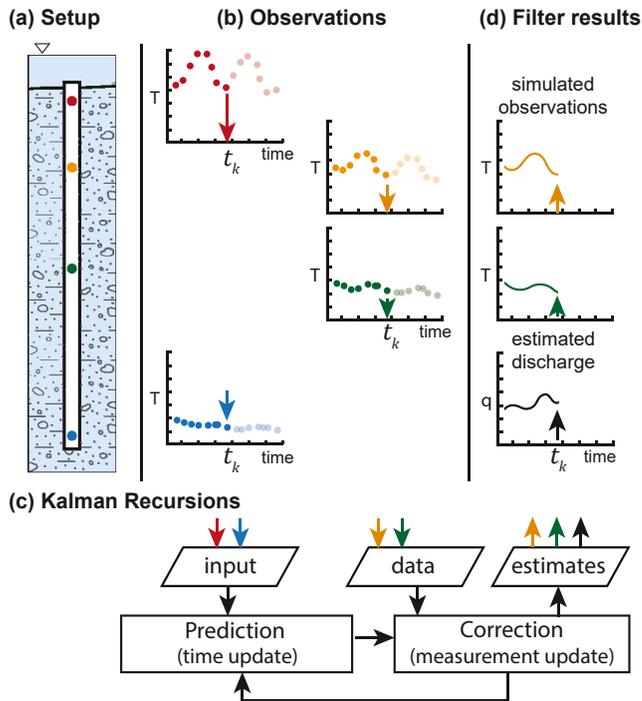
Here, we propose a recursive-estimation framework (Figure 1) to identify time-varying specific discharge associated with groundwater/surface-water exchange. This approach uses: (a) a state-space model (SSM) of heat transport and (b) Kalman-based algorithms commonly used in control engineering (Brown & Hwang, 2012; Gelb, 1974; Simon, 2006) and time series analysis (Durbin & Koopman, 2012; Harvey, 1990). In the former context, Kalman-based approaches recursively predict the future states of a dynamical system (e.g., a circuit model or kinematic model of a vehicle's motion) and their covariance using both a process model and measurements. In the latter context, the same mathematical procedure is used for a time series model (e.g., ARIMA or Box-Jenkins) rather than a physics-based process model. Our approach therefore can be viewed as a hybrid between calibration of rigorous process models (e.g., *IDTempPro*) and digital signal processing (e.g., *VFLUX*), and subsequent analysis with analytical models. Indeed, our formulation offers the strengths of both: (a) computational efficiency for estimation of time-varying discharge, discretized based on the measurement interval; (b) resolution of abrupt changes in discharge; (c) consideration of complicated time-varying boundary conditions; and (d) quantitative assessment of estimation uncertainty offered by the posterior state covariance matrix in the Kalman-based approach.

## 2. Methods

Heat transport below the sediment/water interface is commonly described using the one-dimensional (1D) convection/conduction partial differential equation (PDE) (after Constantz, 2008):

$$C_s \frac{\partial T}{\partial t} = K_T \frac{\partial^2 T}{\partial z^2} - C_w q \frac{\partial T}{\partial z}, \quad (1)$$

where  $K_T$  is thermal conductivity of bulk sediment (W/(m °C)),  $C_w$  is volumetric heat capacity of water (J/(m<sup>3</sup> °C)),  $C_s$  is the volumetric heat capacity of the bulk sediment (J/(m<sup>3</sup> °C)),  $T$  is temperature (°C),  $t$  is time (s),  $z$  is vertical position (m), and  $q$  is specific fluid discharge (m/day). We note that Equation 1 neglects thermal dispersion per common practice although consideration of dispersion is straightforward. Analytical solutions to the PDE are available for idealized boundary conditions and homogeneous properties (e.g., Luce et al., 2017). We note also that the 1D assumption is commonly used and justified (e.g., Lutz, 2010; Reeves & Hatch, 2016)



**Figure 1.** Schematic diagram illustrating heat tracing for groundwater/surface-water exchange: (a) field-experimental setup with temperature sensors indicated by colored circles; (b) observed temperatures used for model boundary conditions and calibration data; (c) the two-step filtering procedure involving prediction and correction; and (d) the filter results, including estimated temperature and discharge states. The filter steps through data assimilating new measurements at time  $t_k$  to produce state estimates for time  $k$ .

for heat tracing although such simplifications can, in general, introduce error in other contexts (Sinsbeck & Tartakovsky, 2015). Numerical approaches (e.g., finite-difference or finite-element) allow for consideration of complex boundary conditions and (or) heterogeneous properties. Although Equation 1 is written for homogenous and temperature-independent properties, this is not a requirement for numerical models. For the heat tracing problem considered here, the model domain is within the sediment layer with boundary conditions assigned based on temperature measurements.

In the following sections, we first formulate the process model for solution of Equation 1 and subsequently formulate filtering and smoothing algorithms to infer specific discharge.

### 2.1. Process Model in Continuous Time

To facilitate recursive estimation, we reformulate Equation 1 as an SSM. First, we discretize the PDE's spatial terms using a finite-difference approximation consistent with that of the *VS2DH* model (Healy & Ronan, 1996), which is used by *IDTempPro*. For the finite-difference cell denoted by the subscript  $i$ :

$$\dot{T}_i(t) = \frac{K_T}{C_s} \frac{T_{i+1}(t) - 2T_i(t) + T_{i-1}(t)}{\Delta z^2} - \frac{C_w q(t)}{C_s} \frac{T_i(t) - T_{i-1}(t)}{\Delta z} \quad (2)$$

where  $\Delta z$  is the spatial discretization, which is assumed to be constant,  $T_i(t)$  is the temperature at the center of the finite-difference cell  $i$  at time  $t$ , and the subscripts  $i + 1$  and  $i - 1$  denote adjacent finite-difference cells. Although we have assumed homogeneity of physical properties in Equations 1 and 2, this assumption is not necessary, and heterogeneity could be considered in the SSM as it is in *VS2DH*. Rearranging Equation 2 and accounting for boundaries adjacent to cells 1 and  $n$ :

$$\dot{T}_i(t) = \left( \frac{K_T}{C_s \Delta z^2} \right) T_{i+1}(t) + \left( \frac{-2K_T}{C_s \Delta z^2} - \frac{C_w q(t)}{C_s \Delta z} \right) T_i(t) + \left( \frac{K_T}{C_s \Delta z^2} + \frac{C_w q(t)}{C_s \Delta z} \right) T_{i-1}(t), \quad i = 2, \dots, n-1 \quad (3a)$$

$$\dot{T}_1(t) = \left( \frac{K_T}{C_s \Delta z^2} \right) T_2(t) + \left( \frac{-2K_T}{C_s \Delta z^2} - \frac{C_w q(t)}{C_s \Delta z} \right) T_1(t) + \left( \frac{K_T}{C_s \Delta z^2} + \frac{C_w q(t)}{C_s \Delta z} \right) u_1(t) \quad (3b)$$

$$\dot{T}_n(t) = \left( \frac{K_T}{C_s \Delta z^2} \right) u_2(t) + \left( \frac{-2K_T}{C_s \Delta z^2} - \frac{C_w q(t)}{C_s \Delta z} \right) T_n(t) + \left( \frac{K_T}{C_s \Delta z^2} + \frac{C_w q(t)}{C_s \Delta z} \right) T_{n-1}(t) \quad (3c)$$

where  $u_1(t)$  and  $u_2(t)$  are the temperatures at the top and bottom boundaries, respectively. Equations 3a–3c constitutes a system of  $n$ -coupled first-order ordinary differential equations. The coefficients in Equations 3a–3c are time-varying because two terms contain  $q(t)$ . To simplify the system, we approximate  $q(t)$  as constant over each measurement interval; this approximation is well justified, given sampling rates are commonly fast (on the order of minutes) compared to the expected changes in flux and thermal response (on the order of hours or days). We note that the time derivatives in Equation 1 remain derivatives in Equations 2 and 3 and hence the dot notation. The use of finite-difference discretization in space but not time is often referred to as the “method of lines” (Schuesser, 1991).

We can now express Equations 3a–3c as a linear time-varying SSM transition equation:

$$\dot{T}(t) = A(t)T(t) + B(t)u(t), \quad (4)$$

where at time  $t$ ,  $A(t)$  is the  $n$ -by- $n$  state-transition matrix, containing the coefficients of the finite-difference cell temperatures in Equations 3a–3c;  $B(t)$  is the  $n$ -by-2 input matrix, containing the coefficients of the two boundary temperatures in Equations 3b and 3c; and  $u(t)$  is a 2-by-1 vector of system inputs that is the boundary temperatures. Consistent with Koch et al. (2016) and many other studies, thermal boundary conditions at the top

$(u_1(t))$  and bottom  $(u_2(t))$  of the model are assigned based on measured temperature time series at the shallowest and deepest measurement locations, respectively.

The solution to Equation 4 to advance the system states from sample time  $t_{k-1}$  to  $t_k$  is

$$T(t_k) = e^{A(t_k-t_{k-1})\Delta t} T(t_{k-1}) + \int_{t_{k-1}}^{t_k} e^{A(t_k-\tau)} \mathbf{B}(\tau) \mathbf{u}(\tau) d\tau, \quad (5)$$

where  $\Delta t$  is the sample interval. In Equation 5, the integration is performed assuming the temperatures at the boundaries (i.e.,  $\mathbf{u}(\tau)$ ) vary linearly between measurements at time  $t_{k-1}$  and  $t_k$  (i.e., a first-order hold).

Our use of a finite-difference spatial discretization to reduce the PDE to a system of ordinary differential equations and the subsequent use of the matrix exponential for temporal solution are rare but not unique. The procedure was first proposed in groundwater hydrology by Umari and Gorelick (1986) for problems in advective-dispersive transport and groundwater flow. Here, the procedure allows us to formulate the process model as an SSM (Equation 4), thus facilitating a straightforward application of well-established and powerful system-identification tools for state estimation (Brown & Hwang, 2012; Gelb, 1974; Särkkä, 2013; Simon, 2006).

We note that this discretization scheme results in a first-order spatial difference for the advective term, which results in a convergence rate on the order of  $\Delta z$ . Such a differencing scheme would lead to stability issues for a spatial and temporal finite-differencing approach; however, since we use the matrix exponential in place of temporal discretization, the solution is not subject to the same stability issues.

## 2.2. Recursive Estimation in Discrete Time

Applications of recursive estimation in hydrology are numerous (Sun et al., 2016) but rarely involve the type of real-time (“online”) state-estimation applications for which the methods were originally developed and intended (Kalman, 1960); rather, Kalman-based approaches are more commonly used for offline (i.e., not real time) parameter-estimation problems, in which the Kalman recursions sequentially assimilate previously acquired time series measurements. We refer the interested reader to Sun et al. (2016) for a review of Kalman-based applications in hydrology and to Shapiro et al. (2022), Shapiro and Day-Lewis (2021), Kang et al. (2018), Kitanidis (2015), Ng et al. (2014), and Li et al. (2015) for representative hydrologic applications.

Whereas the heat-transport process is simulated in continuous time, the estimation problem is solved in discrete time; that is the recursive estimation is performed at the times where measurements are acquired. This continuous/discrete scheme is common in the literature (e.g., Brown & Hwang, 2012) and provides for the accurate and stable process modeling and efficient assimilation of time series data. The state vector for the discrete-time estimation problem,  $\mathbf{x}_k$ , at measurement time  $t_k$  is augmented to include specific discharge as a state:

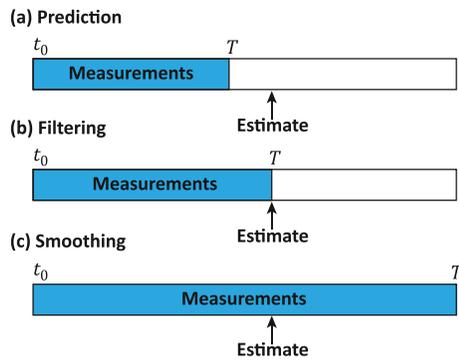
$$\mathbf{x}_k = [T_1(t_k) \ \dots \ T_n(t_k) \ q(t_k)]^T. \quad (6)$$

Converting the state transition equation (Equation 4) from continuous to discrete time, including the measurement equation, using the new augmented state vector,  $\mathbf{x}_k$ , and applying a first-order hold for the input, yields

$$\mathbf{x}_k = \Phi_{k-1} \mathbf{x}_{k-1} + \mathbf{G}_{k-1}^0 \mathbf{u}_{k-1} + \mathbf{G}_{k-1}^1 \mathbf{u}_k + \mathbf{w}_{k-1}, \text{ and} \quad (7a)$$

$$\mathbf{y}_k = \mathbf{H} \mathbf{x}_k + \mathbf{v}_k, \quad (7b)$$

where  $\Phi_{k-1}$  is the discrete-time transition matrix calculated at time  $k-1$ ;  $\mathbf{G}_{k-1}^0$  and  $\mathbf{G}_{k-1}^1$  are the discrete-time input matrices calculated at time  $k-1$ , discussed subsequently;  $\mathbf{u}_{k-1}$  is a 2-by-1 vector containing the temperatures measured at the top and bottom boundaries;  $\mathbf{w}_{k-1}$  is the process-noise vector with each element corresponding to one of the augmented  $n+1$  states;  $\mathbf{y}_k$  is a  $m$ -by-1 vector of simulated observations; and  $\mathbf{v}_k$  is the measurement-noise vector with each element corresponding to one of the augmented  $m$  measurements. For our problem, the state-space measurement equation (Equation 7b) simply extracts or interpolates from  $\mathbf{x}_k$  the temperatures  $\mathbf{y}_k$  at measurement locations. For each observation, the corresponding row of  $\mathbf{H}$  comprises weights that linearly interpolate the temperature at the observation location using the closest two finite-difference cell centers. Here, we assume that both the process and measurement noise are independent Gaussian white noise.



**Figure 2.** Schematic diagram explaining (a) prediction, where estimation is performed at times beyond the interval over which data are available,  $[t_0, T]$ ; (b) filtering, where estimation is performed at the time,  $T$ , of the current measurement; and (c) smoothing, where estimation is performed for times within the interval over which data are available (after Särkkä (2013)).

The matrices  $\Phi_{k-1}$ ,  $G_{k-1}^0$ , and  $G_{k-1}^1$  are found at each sample time assuming a first-order hold for the input following Virtanen et al. (2020). The SSM is augmented a second time to include the inputs from times  $k-1$  and  $k$ , and the matrix exponential of the augmented system's transition matrix is calculated, producing:

$$e \begin{bmatrix} A(t_{k-1}) \Delta t & B(t_{k-1}) \Delta t & 0 \\ 0 & 0 & I \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \Phi_{k-1} & G_{k-1}^0 + G_{k-1}^1 & G_{k-1}^1 \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix}. \quad (8)$$

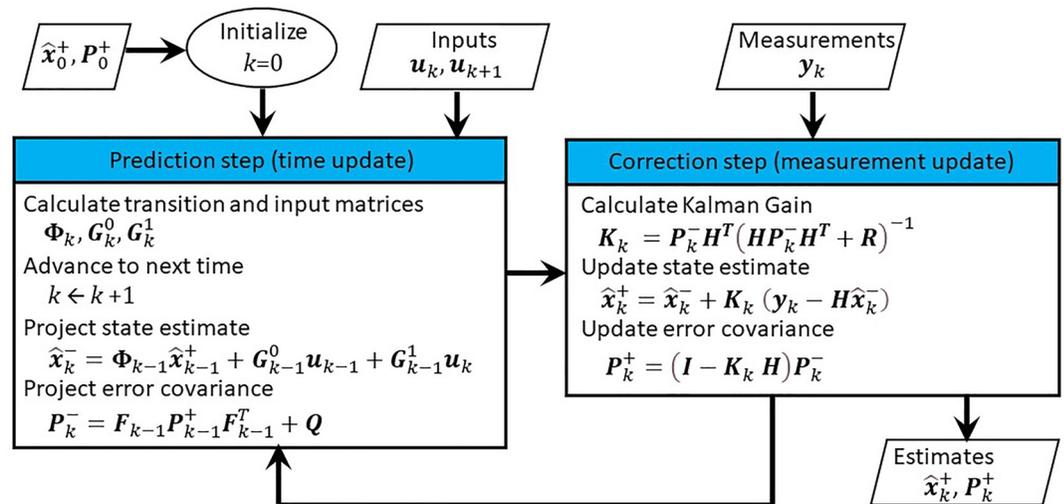
The relevant matrix blocks are extracted from the result of this operation, from which  $\Phi_{k-1}$ ,  $G_{k-1}^0$ , and  $G_{k-1}^1$  are calculated.

The last row of the linear system of Equation 7a describes the transition of  $q(t)$ . The true transition of  $q(t)$  is governed by hydraulic forcing, which is unknown in practice; hence, a stochastic process must be assumed for the transition of this state. Here,  $q(t)$  is modeled as a random walk (RW), a parsimonious model commonly used for the variation of unobserved states. The RW model in discrete time, from time step  $k-1$  to  $k$ , is

$$q_k = q_{k-1} + w_{k-1}, \quad (9)$$

where  $w_{k-1}$  is zero-mean Gaussian white-noise. Identification of the variance is discussed in Appendices A and B and demonstrated in Supporting Information S1 (Sections 3–4).

The goal of filtering is to estimate a system's current states using information acquired up to and including the current time (Figure 2), that is, real-time (online) estimation. Similar to the linear KF, the EKF assimilates each new measurement using a two-step procedure, prediction, and correction, illustrated schematically in Figure 3. The filter is initialized with a prior state estimate ( $\hat{x}_0^+$ ) and prior covariance ( $P_0^+$ ). The prediction step (time update) projects the system states and their covariance from time  $t_{k-1}$  to  $t_k$ . The projection of the states (Equation 10) is performed based on Equation 7a, and the projection of the covariance (Equation 11) is based on a linearization involving the state Jacobian,  $F_k$ , calculated at step  $k$  using a central finite-difference scheme and perturbation around the current state estimate:



**Figure 3.** Schematic diagram showing the workflow for the two-step Extended Kalman Filter (EKF) recursion and associated calculations. The Extended Rauch-Tung-Striebel Smoother workflow (not shown) includes forward and backward EKF passes over the measurements, using the results of the first pass to initialize the second.

$$\hat{\mathbf{x}}_k^- = \Phi_{k-1} \hat{\mathbf{x}}_{k-1}^+ + \mathbf{G}_{k-1}^0 \mathbf{u}_{k-1} + \mathbf{G}_{k-1}^1 \mathbf{u}_k, \quad (10)$$

$$\mathbf{P}_k^- = \mathbf{F}_{k-1} \mathbf{P}_{k-1}^+ \mathbf{F}_{k-1}^T + \mathbf{Q}, \quad (11)$$

where  $\mathbf{Q}$  is the process noise covariance, assumed here to be time-invariant. The state Jacobian  $\mathbf{F}_k$  differs from the discrete-time transition matrix  $\Phi_k$  because  $\Phi_k$  depends on discharge, which is included in the state vector. We further assume that the process noise for all states is independent zero-mean Gaussian white noise and constant across time with one variance common to all temperature states and a different variance for the RW of the specific discharge, consistent with  $w$  (Equation 7a).

The correction step (measurement update) assimilates new measurements at time  $t_k$ , based on a measurement Jacobian,  $\mathbf{H}$ , which, here, is time-invariant because the predicted measurements  $\mathbf{y}_k$  depend linearly on the temperatures in the current state  $\mathbf{x}_k$ . The Kalman gain,  $\mathbf{K}_k$  (Equation 12), is used in the correction step to weight how the new measurements affect the state estimate (Equation 13) and covariance (Equation 14):

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}^T (\mathbf{H} \mathbf{P}_k^- \mathbf{H}^T + \mathbf{R})^{-1}, \quad (12)$$

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H} \hat{\mathbf{x}}_k^-), \quad \text{and} \quad (13)$$

$$\mathbf{P}_k^+ = (\mathbf{I} - \mathbf{K}_k \mathbf{H}) \mathbf{P}_k^-, \quad (14)$$

where  $\mathbf{R}$  is the measurement noise covariance at time  $k$ . Here, we assume the measurement noise is time-invariant independent zero-mean Gaussian white noise; thus,  $\mathbf{R}$  is a diagonal matrix with elements being measurement error variances, consistent with  $v$  in Equation 7b.

The EKF algorithm thus advances through the time series data, predicting and correcting at each sample time (Figure 3). Our EKF recursions follow well-established procedures for the filter as described in both modern (Brown & Hwang, 2012; Särkkä, 2013; Simon, 2006) and classic texts (Gelb, 1974) on optimal state estimation.

Whereas the goal of filtering is to estimate current states based on measurements up to the present time, smoothing uses all available information. The goal of smoothing is to estimate a system's state using information from the past and present with respect to the times being estimated, that is, en bloc (offline) estimation (Figure 2). Just as the linear Rauch-Tung-Striebel Smoother (RTSS) is the smoothing algorithm based on the linear KF, the ERTSS is the smoothing algorithm based on the EKF (Särkkä, 2013). Here, we consider fixed-interval smoothing (FIS), which produces estimates over a fixed time interval that is identical to the interval over which measurements are available. The ERTSS FIS entails by first running an EKF forward through the time series measurements to obtain forward-pass filter estimates. The state and covariance estimates for each step of the forward pass are then used in a second, backward pass through the measurements, producing the smoothed result. For a linear problem, the RTSS can be shown to produce optimal estimates (Rauch et al., 1965), but the same is not guaranteed for the ERTSS applied to nonlinear problems (Särkkä, 2013). Of course, optimality depends also on the appropriate selection of  $\mathbf{Q}$  and  $\mathbf{R}$ , which control the reliance on the prediction relative to the reliance on measurements. See Appendices A and B for automated selection methods, including the Discrepancy Principle (Constable et al., 1987; Hansen, 1997), which is demonstrated in Section 3. Furthermore, the convergence to optimal estimates and the stability of the filter depend also on the selection of  $\hat{\mathbf{x}}_0^+$  and  $\mathbf{P}_0^+$ .

In this work, we focus on the estimation of specific discharge and assume that the values of thermal properties ( $K_T$ ,  $C_w$ , and  $C_s$ ) are known. Compared to hydraulic conductivity, which varies by approximately six orders of magnitude in natural sediments, thermal properties are more tightly constrained (Stonestrom & Constantz, 2003). Furthermore, in situ direct measurement of the thermal properties of bed materials is possible using inexpensive handheld field probes (e.g., the Decagon KD-2 or the Meter Group TEMPOS) as in, for example, Briggs et al. (2014).

The limitations of our approach include: (a) the 1D assumption, (b) the assumption of constant discharge between sample times, (c) assumption of known thermal properties, and (d) the linearization and Gaussian assumption underlying the EKF and ERTSS. The first limitation is common to many approaches to heat tracing. The second limitation is mitigated by using modern measurement technology, which is capable of logging measurements at minute intervals. The third can be addressed by in situ measurement as noted above or assessed using the sensi-

tivity analysis (see Section 3.1). The third and fourth could also be addressed in future extensions of this work involving stochastic Kalman-based frameworks, such as the Ensemble Kalman Filter (EnKF) and Particle Filter (PF) (e.g., Särkkä, 2013; Simon, 2006); these filters (and associated smoothers) are better suited than the EKF to strongly nonlinear and non-Gaussian problems.

### 2.3. Python Implementation

Three Jupyter notebooks, available from McAliley et al. (2022) and also reported in Supporting Information S1, set up and run analyses of synthetic and field-experimental data (Section 3). The spatial discretization of the convection/conduction equation, construction of the SSM matrices, matrix-exponential solution, and filtering and smoothing algorithms are implemented in a Python library, *tempest1d*, also available from McAliley et al. (2022); this library contains two classes that are needed to implement the filter. The first class, *ModelProperties*, is a convenience container to hold spatial discretization parameters and thermal properties. The second class, *EKF*, extends *filterpy*'s *ExtendedKalmanFilter* class to form the Jacobian of the system dynamics function and hold problem-specific parameters. Code dependencies include *filterpy* (Labbe, 2018) and, in turn, *filterpy*'s dependencies (e.g., *numpy* and *scipy*). In addition, the Jupyter notebooks use *matplotlib* for graphics. The use of Jupyter notebooks in this work facilitates: (a) reproducibility of research, (b) application to other data sets, (c) extension by others, and (d) integration with other software and interfaces.

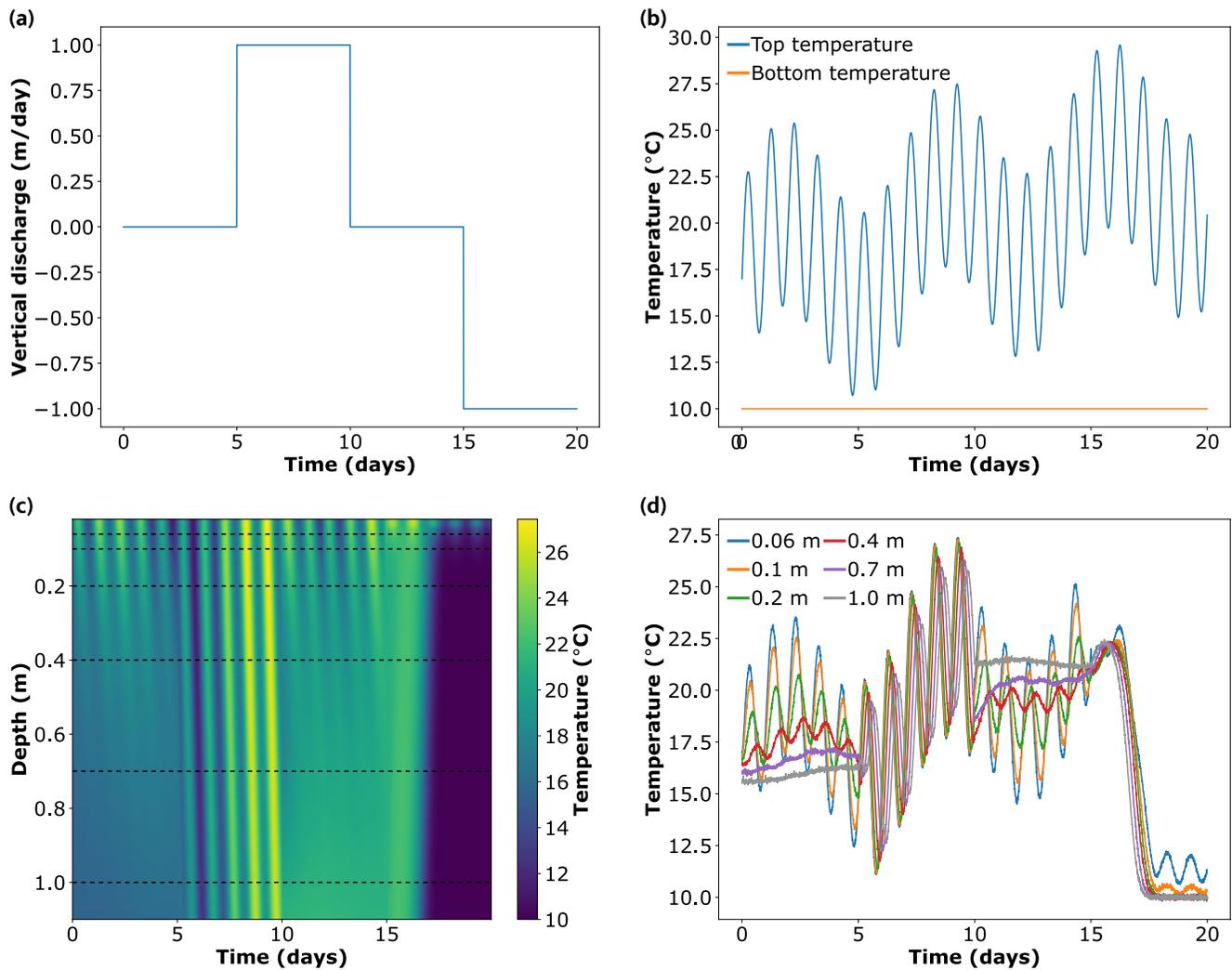
## 3. Examples

We demonstrate the EKF and the ERTSS in two examples. First, we consider a synthetic data set with step changes in specific discharge; this example provides an assessment of the algorithms' ability to resolve extremely abrupt changes in the specific discharge. The application to synthetic data allows for a straightforward comparison of the performance of the EKF, ERTSS, and the commonly used discharge-estimation algorithms of *VFLUX*, for a problem where the true specific discharge is known. In the second example, we consider field-experimental data from the Upper Neversink watershed in New York, representing a realistic application where the true physical parameters and discharge are unknown.

### 3.1. Example 1: Synthetic Problem

The synthetic scenario spans 20 days and features a vertical discharge with four stages of constant discharge, each stage lasting 5 days. The four values of discharge for the stages are 0, 1, 0, and  $-1$  m/day as shown in Figure 4a. To generate synthetic data, we assume that the sediment/water interface has sinusoidal diurnal and weekly periodic components with amplitudes of  $5^{\circ}\text{C}$  and  $3^{\circ}\text{C}$ , respectively, and an upward linear trend of  $0.3^{\circ}\text{C}$  per day as shown in Figure 4b; this signal is used to prescribe the Dirichlet boundary condition for the top of the synthetic model. We place the bottom boundary at 5 m below the surface, where we assume a constant temperature of  $10^{\circ}\text{C}$ , giving the boundary condition for the base of the model. The values of hydraulic and thermal properties and model specifications are given in Table 1. The numerically computed temperatures of the topmost meter over time are shown in Figure 4c; from these temperatures, the synthetic data are taken at 0.06, 0.1, 0.2, 0.4, 0.7, and 1 m depths below the sediment/water interface at a 10-min interval. The synthetic data were corrupted with zero mean Gaussian noise of standard deviation  $0.0625^{\circ}\text{C}$ , the resolution of the commonly used *iButton* temperature sensor (Figure 4d).

The EKF and ERTSS require a process covariance matrix,  $\mathbf{Q}$ , as explained in Section 2. For this example, we assume a diagonal  $\mathbf{Q}$ . Each diagonal entry represents the variance of the corresponding state. The last diagonal entry corresponds to the temporal variance of the discharge over  $\Delta t$ ; the assumed value for this covariance controls the magnitude of temporal variation in the discharge estimate and thus strongly affects the performance of the filter. We tested values ranging across several orders of magnitude, choosing a value of  $7 \times 10^{-5} (\text{m/day})^2$  (see Appendices A and B, Supporting Information S1). This choice corresponds to an expected standard deviation of  $0.0086$  m/day from the start of a 10-min time interval to the end. The first  $n$  diagonal entries of the process covariance matrix correspond to the temperatures in each model cell; these values could be set to zero for the EKF, but such a choice results in numerical instability in the ERTSS. Therefore, the remaining diagonal entries were set to  $10^{-4} (^{\circ}\text{C})^2$ , a small value, reflecting the assumption that the heat transport model is highly certain.



**Figure 4.** Synthetic 20-day data with (a) known specific discharge; (b) time-varying temperature boundary conditions to generate synthetic data; (c) simulated temperatures with dashed horizontal black lines showing the measurement locations; and (d) temperature time series measurements with additive Gaussian noise.

**Table 1**  
*Physical Properties and Model Specifications for Example 1*

Physical property	Value
$K_T$	2 W/(m°C)
$C_w$	$4.182 \times 10^6$ J/(m <sup>3</sup> °C)
$C_s$	$2 \times 10^6$ J/(m <sup>3</sup> °C)
Model specification	Value
Initial estimate for $q$	-0.864 m/day
Initial, prior standard deviation for temperature states	5°C
Initial, prior standard deviation for $q$	1.002 m/day
Process standard deviation for temperature states	$1 \times 10^{-2}$ °C
Process standard deviation for $q$	0.0086 m/day
Standard measurement error	0.0625°C

Both the EKF and ERTSS estimates of discharge are close to the true discharge (Figure 5); however, the ERTSS improves the estimates over the EKF estimates in three respects. First, the ERTSS estimates are smoother in time as each estimate is informed by more measurements and thus less affected by measurement noise. Second, it decreases the uncertainty associated with the estimates as indicated by the tighter confidence intervals. Third, it better localizes the step changes in time because it is using information from before and after the changes to inform all estimates. Figures 5c and 5e demonstrate this third effect. For both the EKF and ERTSS, the full response to the step changes is approximately 0.2 days (<5 hr).

As a basis for comparison, we analyzed the synthetic data using *VFLUX* (version 2) (Figure 5e). The “Hatch Amplitude” analytical model of *VFLUX* was applied using the known synthetic thermal parameters and the 0.1 and 0.2 m depth pair. *VFLUX* does not resolve the step changes as well as either the ERTSS or EKF. The *VFLUX* DHR algorithm requires approximately 2 days to fully respond to step changes. Furthermore, the quality of discharge estimates is strongly degraded during the period of strong upwelling from days 15–20 likely due to the highly attenuated diurnal signal at 0.2 m depth during this time, highlighting the limitations of the signal-based analytical models in zones of groundwater discharge.

We also investigated the effect of using an incorrect value for the thermal conductivity. We perturbed the true thermal conductivity by 10% in either direction. The RMSE of the discharge estimates increased from 0.052 m/day to 0.065 and 0.064 m/day because of using a thermal conductivity value that was 10% too low and 10% too high, respectively. The mean of the change in estimated discharge over the change in thermal conductivity was  $-0.0022$  m/day per  $W/m^{\circ}C$ ; thus, the discharge estimates are not strongly sensitive to reasonable uncertainty in thermal conductivity in this example.

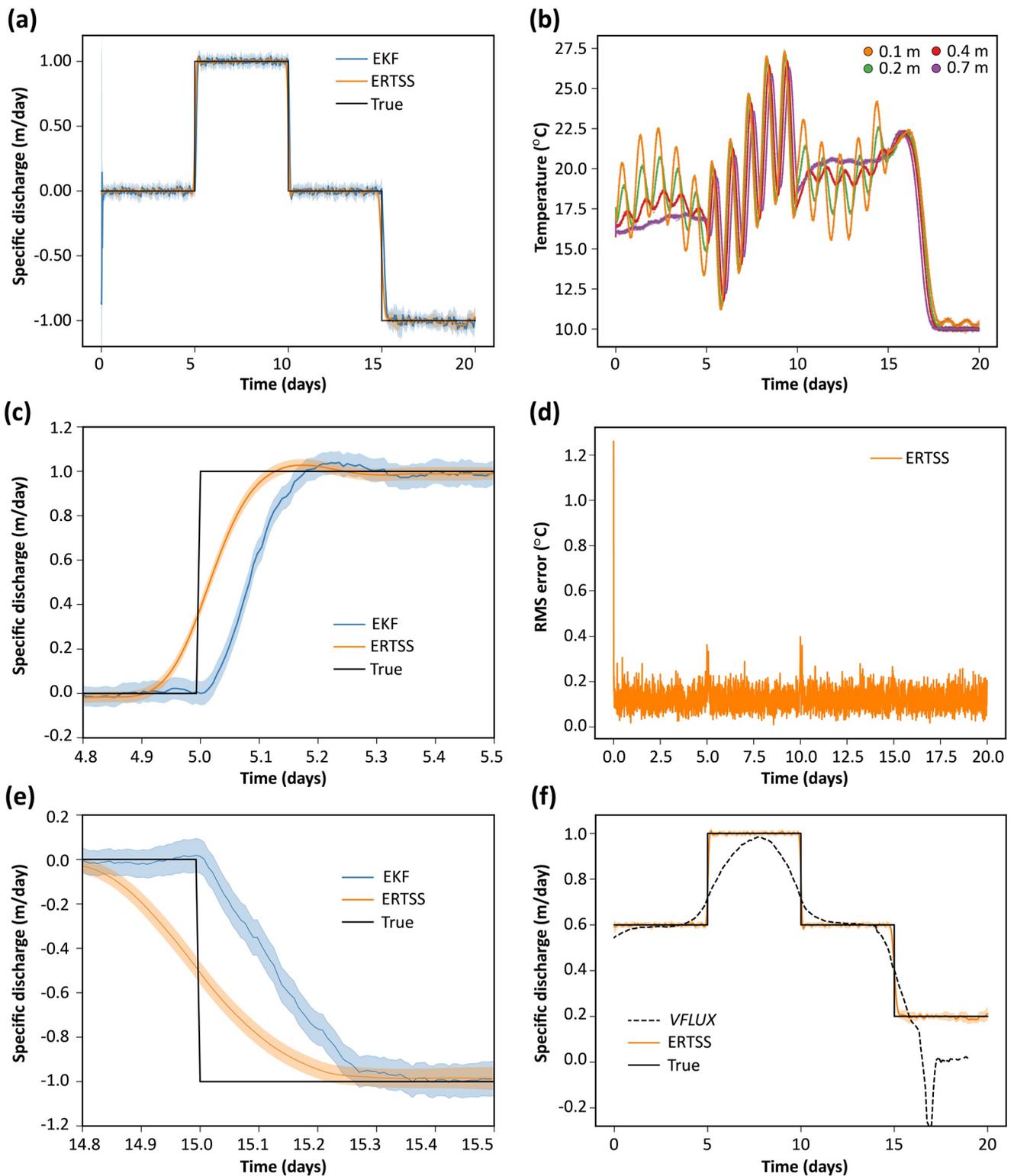
### 3.2. Example 2: Field-Experimental Data From the Upper Neversink Subwatershed

To demonstrate the applicability of our approach to field data, we use data from the Upper Neversink watershed, which falls within the Delaware River Basin, one of the U.S. Geological Survey's Integrated Water Science basins. The Neversink is an area of ongoing intensive hydrologic monitoring, including new infrastructure for networked real-time monitoring of vertical temperature profiles in streambeds and exposed groundwater seeps. We present the data available from a vertical temperature installation in a river-bank seep at the hillslope break that typically contributes the observed (with thermal infrared) discharge to the channel below (Figure 6).

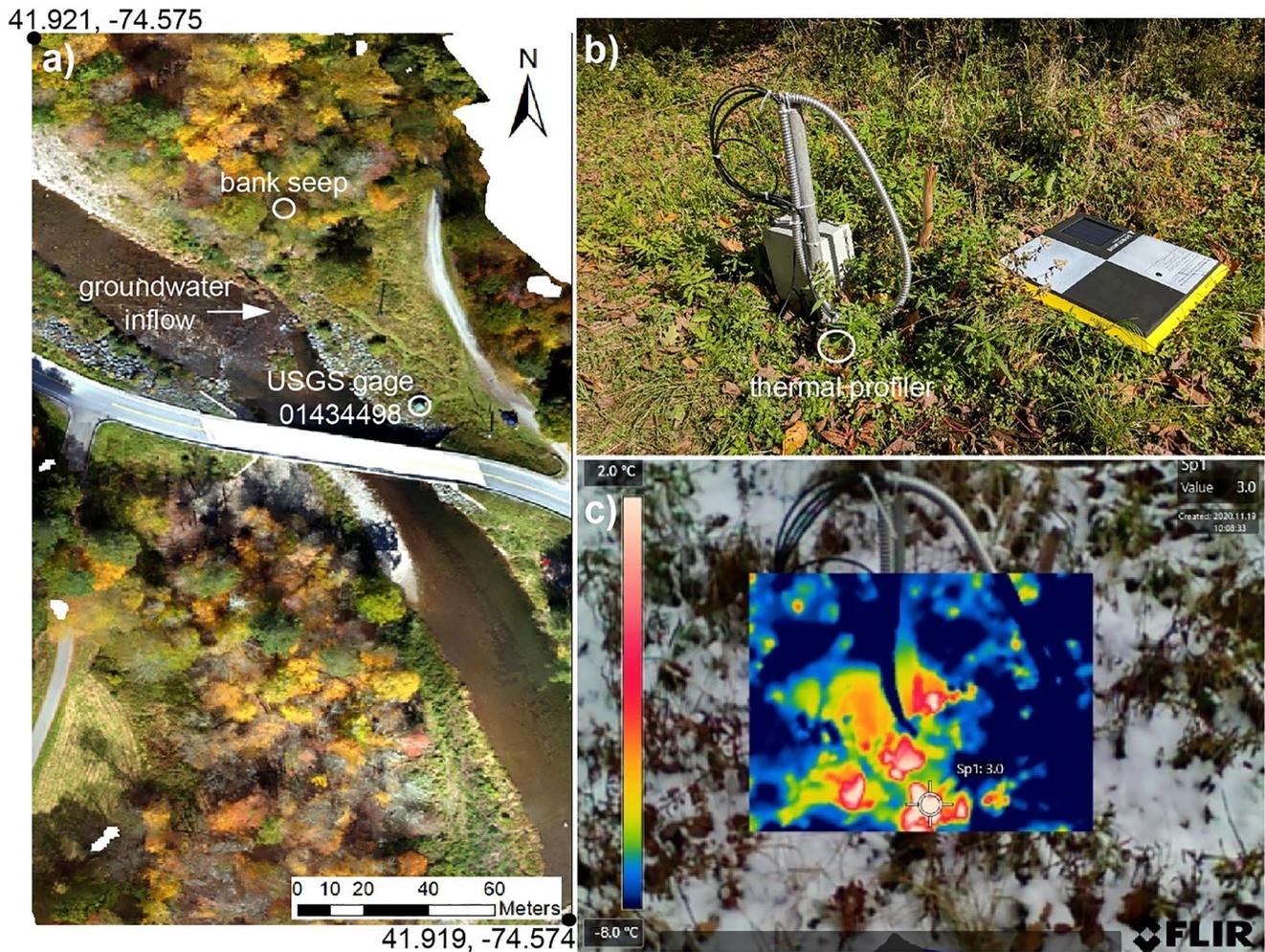
Temperatures were recorded using a Campbell Scientific CS230 temperature profiler with thermistors located at 0, 0.05, 0.1, 0.3, and 0.5 m below the sediment/water interface (Figure 7a). We neglect data from the thermistor at 0 m as data from the sediment/water interface are commonly noisy and may prove problematic if stage drops, and the thermistor is exposed to air. We use the temperature time series from the thermistors at 0.05 and 0.3 m to define Dirichlet boundary conditions for the SSM. Measurements from the thermistor at 0.1 m are used for calibration. Data from the thermistor at 0.5 m are not used because the depth to bedrock is less than 1 m, resulting in increasing lateral heat transport with increasing depth. Furthermore, the *VFLUX* analysis used only data from the thermistors at 0.05 and 0.1 m due to extended times of highly attenuated diurnal signals at deeper sensors. The EKF requires a minimum of three thermistors, so using the thermistors at 0.05 m, 0.1 m, and 0.3 m allows the closest comparison of the estimates by both methods.

The covariance matrices for the EKF were formed as follows. First, we form the measurement noise covariance matrix, which expresses the estimated measurement error. The manufacturer-specified resolution of the thermistors is  $0.0078^{\circ}C$ . However, based on experience with the instrument, we estimate that the standard deviation due to instrument error is  $0.03^{\circ}C$ . Furthermore, the temperature data as available online from USGS are rounded to the nearest  $0.1^{\circ}C$ . Assuming that rounding errors are uniformly distributed between  $-0.05^{\circ}C$  and  $0.05^{\circ}C$ , the standard deviation of the rounding error is  $0.1/\sqrt{12}^{\circ}C = 0.0289^{\circ}C$ . The total measurement noise standard deviation is the square root of the sum of the squares of these two values,  $0.042^{\circ}C$ . The square of this value is placed along the diagonal of the measurement covariance matrix, and the off-diagonal terms are set to zero.

We use a diagonal process covariance matrix, just as in the synthetic example. A value of  $0.00086$  m/day was chosen for the discharge standard deviation using the discrepancy principle (see Appendices A and B, Supporting Information S1, and Section 4). We again set the remaining diagonal entries of the process covariance matrix to the value of  $10^{-4} (^{\circ}C)^2$  to ensure stability for the ERTSS. This value corresponds to a process standard deviation



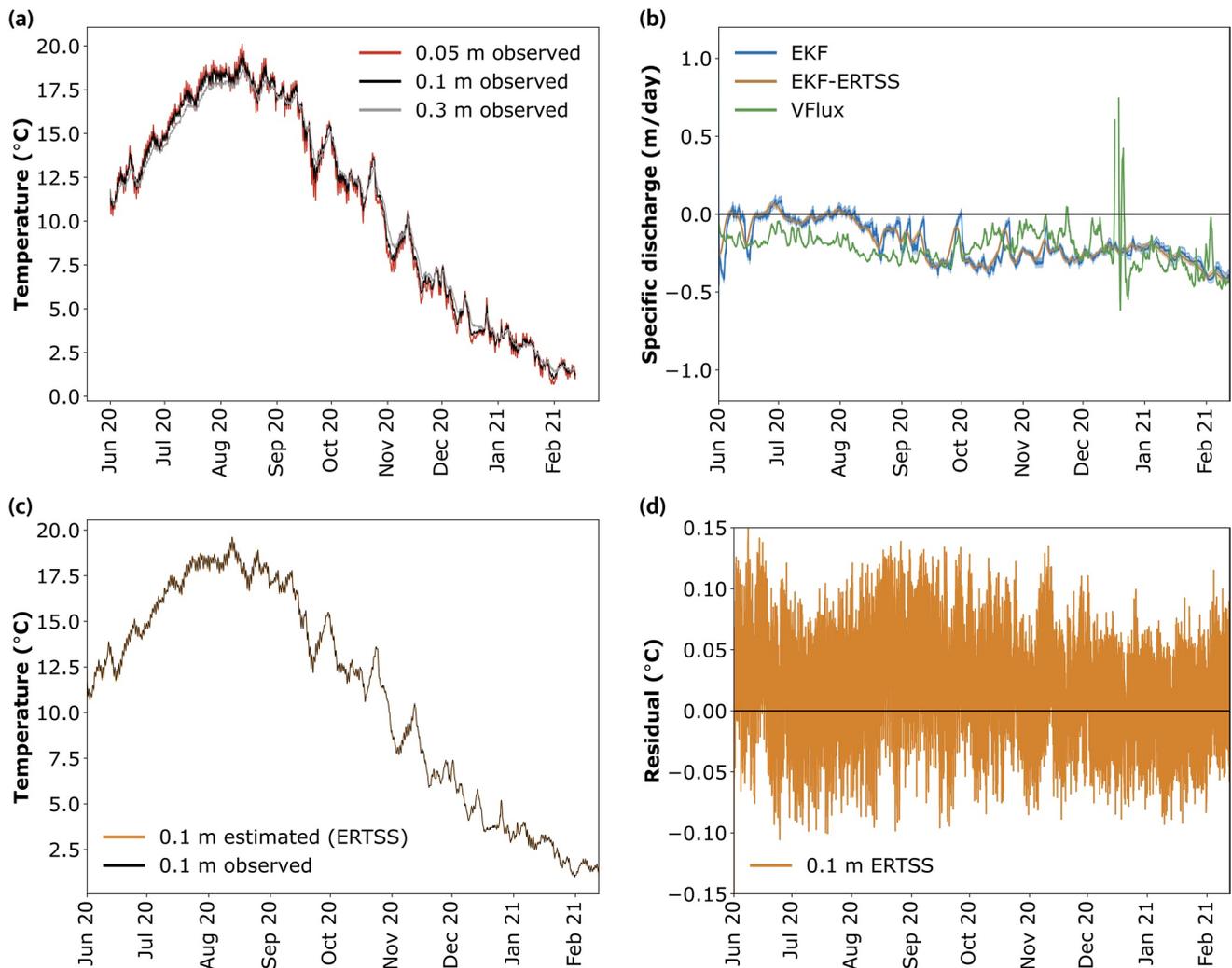
**Figure 5.** (a) Estimated discharge for Example 1, obtained using the Extended Kalman Filter (EKF) and Extended Rauch-Tung-Striebel Smoother (ERTSS), with 95% confidence intervals shaded; (b) measured temperatures at the middle four thermistors, plotted as circles, and temperatures predicted by the EKF with ERTSS; (c) estimated discharge during the first step change; (d) RMS error in temperature at each time step; (e) estimated discharge during the last step change; and (f) Comparison of the ERTSS and *VFLUX* estimates.



**Figure 6.** (a) Drone-based ortho imagery shows the location of a shallow groundwater bank seep along the west branch of the Neversink River relative to a U.S. Geological Survey Next Generation Water Observing System gage; (b) a photograph of instrumented seep in fall with noted placement of the vertical temperature profiler next to a monitoring piezometer; and (c) an infrared image of the seep in winter showing sustained discharge of relatively warm groundwater to the land surface.

of  $0.01^{\circ}\text{C}$ , which is significantly less than the measurement standard deviation of  $0.042^{\circ}\text{C}$ . Keeping this value below the measurement standard deviation ensures that the state covariance is primarily determined by the measurement standard deviation, not the temperature process standard deviation, reflecting the assumption that the process model captures the governing physics well.

The values of hydraulic and thermal properties and model specifications are given in Table 2 and were estimated based on field sediment observations. Results from the ERTSS are shown in Figure 7b. The ERTSS analysis took approximately 2 min on a laptop computer with a 2.9 GHz Quad-Core Intel Core i7 processor. Note that the use of *IDTempPro* would take at least an order of magnitude longer for these problems. As with the synthetic data test, the Hatch Amplitude model of *VFLUX* was run using the same thermal parameters as for ERTSS. *VFLUX* produces spurious estimates during a few days in December of 2020, during which temperatures stayed nearly constant. Because *VFLUX* relies on sinusoidal diurnal temperature variations, the algorithm performs poorly when there is no sinusoidal component to temperature fluctuations. Neither the EKF nor the ERTSS have this performance limitation. We also note that the SSM approach presented here uses temperature measurements to form a lower boundary condition, while *VFLUX* assumes a semi-infinite half-space. This difference might explain some of the discrepancy between *VFLUX* and ERTSS discharge estimates, especially during a rapid upward exchange. As shown in Supporting Information S1, the EKF and ERTSS are also robust in the presence



**Figure 7.** (a) Temperature data set from Upper Neversink Watershed; (b) vertical discharge estimated using Extended Kalman Filter (EKF) and Extended Rauch-Tung-Striebel Smoother (ERTSS), compared with discharge estimates obtained by applying *VFLUX* to temperatures at 0.05 and 0.1 m; (c) measured temperatures at 0.1 m compared to temperatures estimated using ERTSS; and (d) residuals between simulated and observed data based on the ERTSS results. Spurious *VFLUX* estimates in December 2020 coincide with a period of multiple days of near-constant temperatures.

of data gaps, which are a common issue for automated analyses of continuous data. Figure 7c shows that predicted temperatures track measured temperatures, and Figure 7d shows that residual errors mostly fall within 0.1°C.

#### 4. Discussion

Based on a synthetic example, the EKF/ERTSS offers improved resolution of rapid changes in flux compared to the commonly used *VFLUX* tool (Figure 5f), while offering the major benefits of signal-extraction methods, that is, computational efficiency for the estimation of time-varying flux. The ERTSS algorithm, in particular, allows for accurate temporal localization of abrupt changes. An application to field-experimental data indicates that the EKF/ERTSS may be less subject to problems arising from strong upwelling, where the diurnal signal exploited by DHR is weak or absent, and *VFLUX* produces spurious results (Figure 7b). The focus on the diurnal signal can also be problematic for DHR approaches in overcast conditions.

The choice of process variance for the discharge has a strong effect on the performance of the filter. A large discharge variance leads to discharge estimates that are governed by temperature data fit. Discharge variance

**Table 2**

*Physical Properties and Model Specifications for Analysis of the Field-Experimental Data Set From the Upper Neversink Watershed in Example 2*

Physical property	Value
$K_T$	1.68 W/(m°C)
$C_w$	$4.182 \times 10^6$ J/(m <sup>3</sup> °C)
$C_s$	$2.829 \times 10^6$ J/(m <sup>3</sup> °C)
Model specification	Value
Initial estimate for $q$	0 m/day
Initial, prior standard deviation for temperature states	5°C
Initial, prior standard deviation for $q$	1.002 m/day
Process standard deviation for temperature states	0.01°C
Process standard deviation for $q$	$8.64 \times 10^{-4}$ m/day
Standard measurement error	0.042°C

that is too large can lead to estimates that overfit the temperature data and vary unrealistically from downwelling to upwelling and back daily. A large discharge variance also results in unrealistically large uncertainties in the estimates. A small discharge variance, meanwhile, yields overly smooth discharge estimates with overconfident estimates of uncertainty and a poor data fit. We use the discrepancy principle described in Appendix A (see also Supporting Information S1, Sections 3–4) to choose the optimal process variance. By this approach, the expected measurement RMSE is estimated based on the measurement noise. Then, the optimal process variance is the value for which the resulting RMSE equals the expected. Furthermore, the variance and character of the discharge over time should be considered. For instance, in the field data example, we ensure that the discharge estimates are physically realistic not alternating between positive and negative discharge (downwelling and upwelling) on a daily basis.

The process and noise variance, as well as the initial prior covariance, also affect the stability of the EKF (and thus ERTSS) and its convergence. We find that positive diagonal entries in the process covariance matrix are necessary for ERTSS stability. Also, if the initial prior covariance is too small (i.e., excessive confidence in the initial state estimates is assumed), the filter measurement updates will not be large enough, and the filter will not converge. Conversely, stability issues can result if the prior covariance is too large. However, we find that when the filter converges stably, the ERTSS estimates are not significantly affected by the choice of initial prior covariance, and only early EKF estimates are affected.

The ERTSS analysis of the synthetic and field data required on the order of 20 and 120 s of computer processing time, respectively, using a 2.9 GHz Quad-Core Intel Core i7 processor in a laptop computer, underscoring the computational efficiency of the EKF/ERTSS approach compared to, for example, use of *1DTempPro*. The ERTSS requires approximately twice the computer processing time of the EKF and thus is tractable for practical problems in vertical heat tracing for groundwater/surface-water exchange.

## 5. Conclusions

We presented and demonstrated a new approach to estimate specific discharge associated with groundwater/surface-water exchange using temperature data collected below the sediment/water interface. Our recursive-estimation framework combines the strengths of previous approaches based on either signal extraction or model calibration: (a) computational efficiency for estimation of time-varying discharge, (b) resolution of abrupt changes in discharge, (c) consideration of arbitrary time-varying Dirichlet boundary conditions, and (d) quantitative assessment of estimation uncertainty. We also note that the new framework is effectively a multifrequency approach; it obviates the decision to analyze amplitude or phase or both and it requires no windowing procedure.

The optimal filtering approach is well suited to online, real-time data assimilation and estimation; furthermore, the Kalman-based approach is also well suited to forecasting (Figure 2), which amounts to performing the predic-

tion step without new measurements. Our SSM and the assumed model for the transition of  $q$  (Equation 9) could be modified to consider the inputs as states and to forecast inputs and discharge both. Alternatives to the RW model (e.g., structural time series, Gauss-Markov, autoregressive, or integrated RW processes) could allow for prediction of temperature states and discharge for management problems related to thermal loading to surface water, for example, maintenance of thermal refugia for temperature-sensitive aquatic species. Real-time assimilation and data reduction problems in other fields are driving intensive development and investment in libraries for machine learning, which include tools for recursive estimation, for example, *Tensorflow* (Abadi et al., 2016) and *Pyro* (Bingham et al., 2019). There is enormous potential for application of these tools in data-driven hydrology (Shapiro & Day-Lewis, 2022) to address real-time water management and capitalize on real-time hydrologic observation networks.

Future extensions to our approach and the associated Python library could include further investigation into methods for automated selection of the process covariance to identify the appropriate trade-off between model complexity and data misfit. For example, the maximum-likelihood method has been used to infer hyperparameters for process noise based on the analysis of measurement residuals (Young, 2011, p. 81). Another direction for future work is consideration of alternative filtering and smoothing methods based on, for example, the Unscented Kalman Filter (UKF), EnKF, or PF (e.g., Särkkä, 2013). Although the heat tracing problem is clearly amenable to solution using the simpler EKF as presented in this paper, the UKF could have advantages for strongly nonlinear examples, and the more computationally intensive EnKF and (or) PF could allow for simultaneous estimation of thermal and (or) hydraulic properties and stochastic assessment of parameter uncertainty. A further possible extension is the consideration of Fixed Lag Smoothing (FLS) instead of FIS. In FLS, the estimation is focused on a time prior to the most recent available measurement, that is, the estimation is delayed with respect to the measurements by a fixed lag. The FLS thus has some of the advantages of FIS, that is, improved temporal resolution of changes in states, reduced estimation uncertainty, and reduced estimation errors. Although the potential improvements with FLS come at the cost of real-time estimation, the cost may be worth a short delay (a few measurement intervals) for many problems.

Aside from improvements to the estimation framework, the process representation in our approach could easily be extended to address artificial heating. In the presence of strong upwelling, the diurnal signal that is commonly exploited in vertical heat tracing is weak or nonexistent. Active heating, for example, using electrical resistive heating elements, can generate an artificial signal to overcome this limitation of heat tracing (e.g., Briggs et al., 2016). As formulated here, our approach and associated codes can address non-sinusoidal boundary conditions, but accurate representation of common heat sources would require an axis-symmetric two-dimensional or else a three-dimensional finite-difference grid; this amounts to a straightforward extension.

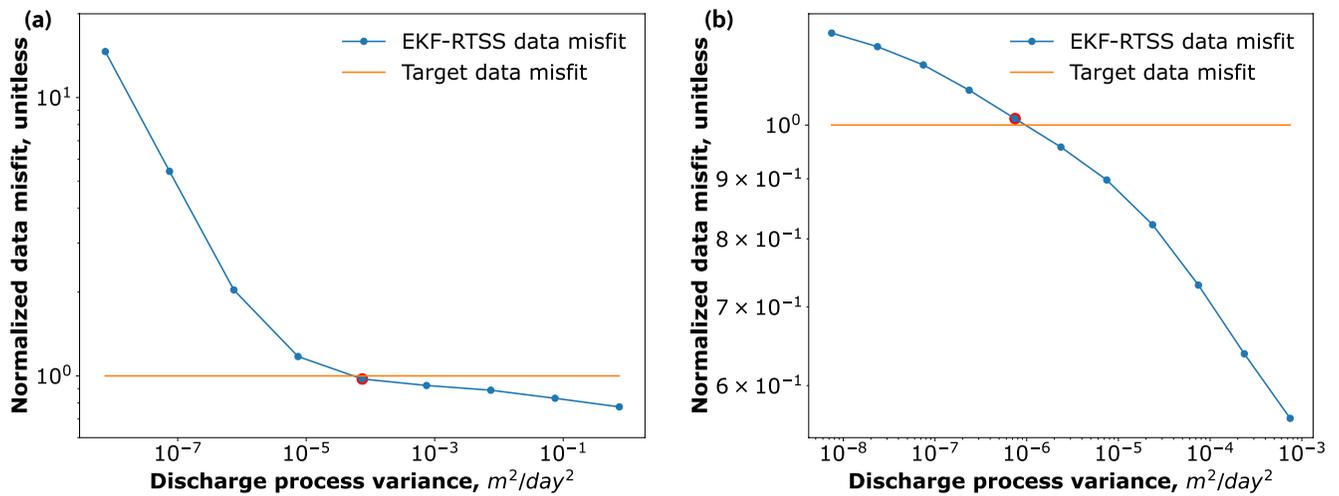
## Appendix A: Using the Discrepancy Principle to Select Discharge Process Variance

A common method in geophysics to identify covariance models in inverse problems is the discrepancy principle (Morozov, 1966). Here, we use this approach to identify the magnitude of the process variance. By the discrepancy principle, the process variance for the discharge is chosen such that the data residuals have the expected standard deviation; hence, this method requires knowledge of the standard deviation of the measurement error (see Appendix B for an alternate method that has no such requirement). Let us define the normalized data misfit as

$$\phi_d = \frac{1}{N} \sum_{k=1}^K (\mathbf{y}_k - \mathbf{H} \mathbf{x}_k^{RTS})^T \mathbf{R}^{-1} (\mathbf{y}_k - \mathbf{H} \mathbf{x}_k^{RTS}), \quad (\text{A1})$$

where  $\mathbf{x}_k^{RTS}$  is the estimated state produced by ERTSS at time  $t_k$ ,  $K$  is the index of the final time step, and  $N$  is the total number of temperature observations across all time steps. Because the measurement errors are random variables, so too the data misfit is a random variable. For independent measurement errors that follow normal distributions, the normalized data misfit follows a chi-square distribution with an expected value of 1. Thus, by the discrepancy principle, the optimal choice for the process variance will yield a normalized data misfit of 1.

We applied the discrepancy principle to the ERTSS estimates for both the synthetic example and the field data example. For the synthetic example, we used discharge process variance values ranging from  $10^{-8}$  to  $10^0$  (m/day)<sup>2</sup>.



**Figure A1.** Normalized data misfit corresponding to different choices for discharge process variance for (a) the synthetic example and (b) the Neversink River data set. The data misfit has been divided by the number of data, so the target value equals 1 (the orange line). The red circle shows the value of the discharge process variance closest to the target data misfit value.

The application of the discrepancy principle yielded a discharge process variance value of  $7.5 \times 10^{-5}$  (m/day)<sup>2</sup>. Figure A1a shows the resulting data misfit values.

We also applied the discrepancy principle to the data set from the Neversink River (Figure A1b). We used an estimate of 0.042°C for the standard deviation of the observed temperatures to determine a target data misfit. A discharge process variance of  $7.5 \times 10^{-7}$  (m/day)<sup>2</sup> results in a data misfit that is 1.015 times the expected value.

## Appendix B: Selecting Discharge Process Variance by the L-Curve Criterion

Without an accurate estimate of the measurement error, the L-curve criterion can be used (Hansen, 1992). In this method, the data misfit is plotted against a measure of model structure for many values of the process variance on a plot with logarithmically scaled axes. Often, the resulting plot resembles the letter “L.” The process variance corresponding to the point of maximum curvature, or “corner point” of the L-curve, is taken as the optimal value.

To quantify model structure, we use the norm of the first derivative in time of the discharge,

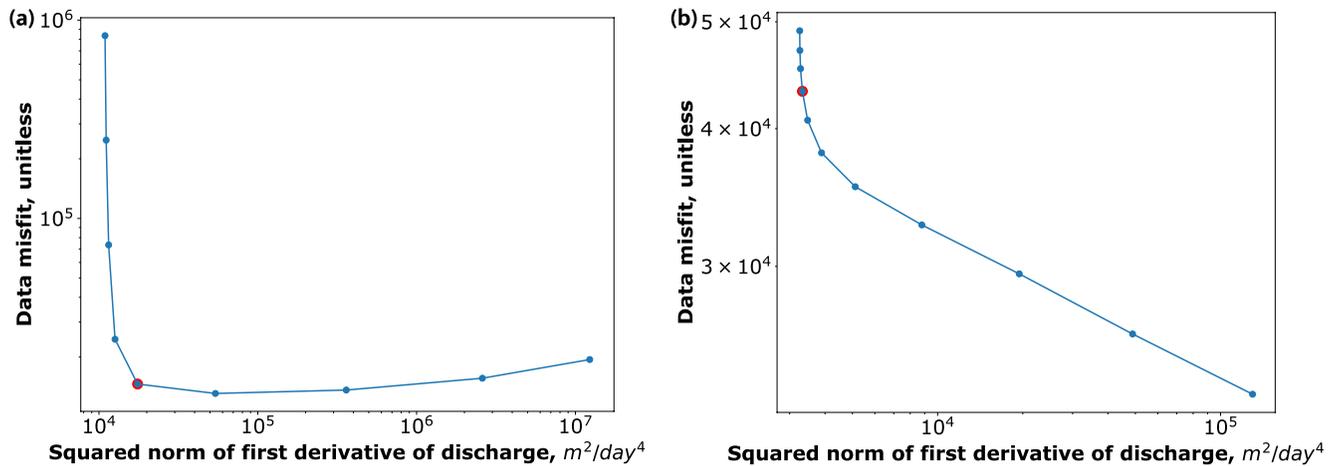
$$\phi_m = \sum_{j=1}^K \left( \frac{\partial q}{\partial t} \right)^2. \quad (\text{B1})$$

Let  $\hat{x} = \log(\phi_m)$  and  $\hat{y} = \log(\phi_d)$ , noting that  $\hat{x}$  and  $\hat{y}$  both depend on the discharge process variance. The curvature of the L-curve is computed at every inner point by finite difference as

$$\kappa = \frac{\hat{x}'\hat{y}'' - \hat{y}'\hat{x}''}{((\hat{x}')^2 + (\hat{y}')^2)^{\frac{3}{2}}} \quad (\text{B2})$$

where derivatives are taken with respect to the discharge process variance. First and second derivatives are approximated by central differencing.

We applied the L-curve criterion to both the synthetic data and the Neversink River data. Figure B1 shows the L-curve for synthetic data. We applied the method to both EKF and ERTSS results, and we found that EKF results yield a more robust and easily identified the point of maximum curvature than ERTSS results. The resulting point of maximum curvature has a corresponding discharge process variance of  $10^{-4}$  m<sup>2</sup>/day<sup>2</sup>, the same value that the discrepancy principle yielded. Figure B1 shows the L-curve for the Neversink River data. Most of the points on the L-curve have negative curvature, but there is a clear point of maximum curvature corresponding to a discharge process variance of  $2.4 \times 10^{-7}$  m<sup>2</sup>/day<sup>2</sup>.



**Figure B1.** Unnormalized data misfit (not divided by the number of data) and model structure corresponding to different choices for discharge process variance for (a) the synthetic example and (b) the Neversink River data set. The red circle in each subplot shows the point of maximum curvature. For the synthetic data set, it corresponds to a discharge process variance of  $1.0 \times 10^{-4} m^2/day^2$ . For the Neversink River data set,  $2.4 \times 10^{-7} m^2/day^2$  is the value at the point of maximum curvature.

## Data Availability Statement

The Python library *tempest1d*, Jupyter notebooks to run the examples and all data used in this work are available at <https://doi.org/10.5066/P99DBTKT>. In addition, the notebooks are provided in Portable Document Format in Supporting Information S1.

## Acknowledgments

This work was supported by the U.S. Geological Survey Water Availability and Use Science Program and the U.S. Geological Survey Toxic Substances Hydrology Program. The U.S. Environmental Protection Agency partially funded and collaborated in the research under Interagency Agreement DW-014-92497401 through its Office of Research and Development. This research was also supported by the Energy and Environment Mission Seed Program under the Laboratory Directed Research and Development (LDRD) Program at Pacific Northwest National Laboratory (PNNL). PNNL is a multiprogram national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830. The authors are grateful to Jeff Raffensperger (USGS) and Adam Mangel (PNNL) for colleague reviews of the draft manuscript, Evan Murdzek (USGS) for review of the data release associated with this paper, and three anonymous peer reviewers. The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

## References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. (2016). TensorFlow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint, 1603.04467. Retrieved from [arxiv.org/abs/1603.04467](https://arxiv.org/abs/1603.04467)
- Anderson, M. P. (2005). Heat as a ground water tracer. *Groundwater*, 43(6), 951–968. <https://doi.org/10.1111/j.1745-6584.2005.00052.x>
- Bingham, E., Chen, J. P., Jankowiak, M., Obermeyer, F., Pradhan, N., Karaletsos, T., et al. (2019). Pyro: Deep universal probabilistic programming. *Journal of Machine Learning Research*, 20, 1–6.
- Briggs, M. A., Buckley, S. F., Bagtzoglou, A. C., Werkema, D. D., & Lane, J. W. (2016). Actively heated high-resolution fiber-optic-distributed temperature sensing to quantify streambed flow dynamics in zones of strong groundwater upwelling. *Water Resources Research*, 52(7), 5179–5194. <https://doi.org/10.1002/2015WR018219>
- Briggs, M. A., Lutz, L. K., Buckley, S. F., & Lane, J. W. (2014). Practical limitations on the use of diurnal temperature signals to quantify groundwater upwelling. *Journal of Hydrology*, 519, 1739–1751. <https://doi.org/10.1016/j.jhydrol.2014.09.030>
- Briggs, M. A., Voytek, E. B., Day-Lewis, F. D., Rosenberry, D. O., & Lane, J. W. (2013). Understanding water column and streambed thermal refugia for endangered mussels in the Delaware River. *Environmental Science & Technology*, 47(20), 11423–11431. <https://doi.org/10.1021/es4018893>
- Brown, R. G., & Hwang, P. Y. C. (2012). *Introduction to random signals and applied Kalman filtering with MATLAB exercises* (p. 400). Wiley.
- Constable, S. C., Parker, R. L., & Constable, C. G. (1987). Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, 52(3), 289–300. <https://doi.org/10.1190/1.1442303>
- Constantz, J. (2008). Heat as a tracer to determine streambed water exchanges. *Water Resources Research*, 44(4), W00D10. <https://doi.org/10.1029/2008WR006996>
- Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods* (p. 368). Oxford University Press.
- Gelb, A. (1974). *Applied optimal estimation* (p. 374). MIT Press.
- Glose, T. J., Lowry, C. S., & Hausner, M. B. (2021). Examining the utility of continuously quantified Darcy fluxes through the use of periodic temperature time series. *Journal of Hydrology*, 595, 125675. <https://doi.org/10.1016/j.jhydrol.2020.125675>
- Gordon, R. P., Lutz, L. K., Briggs, M. A., & McKenzie, J. M. (2012). Automated calculation of vertical pore-water flux from field temperature time series using the VFLUX method and computer program. *Journal of Hydrology*, 420(421), 142–158. <https://doi.org/10.1016/j.jhydrol.2011.11.053>
- Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 34(4), 561–580. <https://doi.org/10.1137/1034115>
- Hansen, P. C. (1997). *Rank-deficient and discrete ill-posed problems: Numerical aspects of linear inversion*. SIAM.
- Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter* (p. 572). Cambridge University Press.
- Hatch, C. E., Fisher, A. T., Revenaugh, J. S., Constantz, J., & Ruehl, C. (2006). Quantifying surface water–groundwater interactions using time series analysis of streambed thermal records: Method development. *Water Resources Research*, 42(10), 1–14. <https://doi.org/10.1029/2005WR004787>
- Healy, R. W., & Ronan, A. D. (1996). *Documentation of computer program VS2DH for simulation of energy transport in variably saturated porous media – Modification of the U.S. Geological Survey's computer program VS2DT* (U.S. Geological Survey Water-Resources Investigations Report 96-4230, p. 36). USGS.

- Irvine, D. J., Lautz, L. K., Briggs, M. A., Gordon, R. P., & McKenzie, J. M. (2015). Experimental evaluation of the applicability of phase, amplitude, and combined methods to determine water flux and thermal diffusivity from temperature time series using VFLUX 2. *Journal of Hydrology*, 531, 728–737. <https://doi.org/10.1016/j.jhydrol.2015.10.054>
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1), 35–45. <https://doi.org/10.1115/1.3662552>
- Kang, X., Shi, X., Deng, Y., Revil, A., Xu, H., & Wu, J. (2018). Coupled hydrogeophysical inversion of DNAPL source zone architecture and permeability field in a 3D heterogeneous sandbox by assimilation time-lapse cross-borehole electrical resistivity data via ensemble Kalman filtering. *Journal of Hydrology*, 567, 149–164. <https://doi.org/10.1016/j.jhydrol.2018.10.019>
- Keery, J., Binley, A., Crook, N., & Smith, J. W. N. (2007). Temporal and spatial variability of groundwater–surface water fluxes: Development and application of an analytical method using temperature time series. *Journal of Hydrology*, 336(1–2), 1–16. <https://doi.org/10.1016/j.jhydrol.2006.12.003>
- Kitanidis, P. K. (2015). Compressed state Kalman filter for large systems. *Advances in Water Resources*, 76, 120–126. <https://doi.org/10.1016/j.advwatres.2014.12.010>
- Koch, F. W., Voytek, E. B., Day-Lewis, F. D., Healy, R., Briggs, M. A., Lane, J. W., & Werkema, D. (2016). 1DTempPro V2: New features for inferring groundwater/surface-water exchange. *Groundwater*, 54(3), 434–439. <https://doi.org/10.1111/gwat.12369>
- Labbe, R. (2018). Filterpy. Retrieved from <https://github.com/r1abbe/filterpy>
- Lapham, W. (1989). *Use of temperature profiles beneath streams to determine rates of vertical ground-water flow and vertical hydraulic conductivity* (U.S. Geological Survey Water Supply Paper 2337, p. 35). USGS.
- Lautz, L. K. (2010). Impacts of nonideal field conditions on vertical water velocity estimates from streambed temperature time series. *Water Resources Research*, 46(1), W01509. <https://doi.org/10.1029/2009WR007917>
- Lautz, L. K. (2012). Observing temporal patterns of vertical flux through streambed sediments using time-series analysis of temperature records. *Journal of Hydrology*, 464–465, 199–215. <https://doi.org/10.1016/j.jhydrol.2012.07.006>
- Li, J. Y., Kokkinaki, A., Ghorbanidehno, H., Darve, E. F., & Kitanidis, P. K. (2015). The compressed state Kalman filter for nonlinear state estimation: Application to large-scale reservoir monitoring. *Water Resources Research*, 51(12), 9942–9963. <https://doi.org/10.1002/2015WR017203>
- Luce, C. H., Tonina, D., Applebee, R., & DeWeese, T. (2017). Was that assumption necessary? Reconsidering boundary conditions for analytical solutions to estimate streambed fluxes. *Water Resources Research*, 53(11), 9771–9790. <https://doi.org/10.1002/2017WR020618>
- McAliley, W. A., Rey, D. M., & Day-Lewis, F. D. (2022). *Data release for tempest1d*: Recursive estimation of vertical groundwater/surface-water exchange using heat tracing. U.S. Geological Survey Data Release. <https://doi.org/10.5066/P99DBTKT>
- Morozov, V. A. (1966). Regularization of incorrectly posed problems and the choice of regularization parameter. *USSR Computational Mathematics and Mathematical Physics*, 6(1), 242–251. [https://doi.org/10.1016/0041-5553\(66\)90046-2](https://doi.org/10.1016/0041-5553(66)90046-2)
- Ng, G.-H. C., Bedford, D., & Miller, D. (2014). Development of a mechanistic model and data assimilation framework for assessing desert ecophysiology. *Water Resources Research*, 50(6), 4662–4685. <https://doi.org/10.1002/2014WR015281>
- Rau, G. C., Cuthbert, M. O., McCallum, A. M., Halloran, L. J., & Andersen, M. S. (2015). Assessing the accuracy of 1-D analytical heat tracing for estimating near-surface sediment thermal diffusivity and water flux under transient conditions. *Journal of Geophysical Research: Earth Surface*, 120(8), 1551–1573. <https://doi.org/10.1002/2015JF003466>
- Rau, G. R., Andersen, M. S., McCallum, A. M., Roshan, H., & Acworth, R. I. (2014). Heat as a tracer to quantify water flow in near-surface sediments. *Earth-Science Reviews*, 129, 40–58. <https://doi.org/10.1016/j.earscirev.2013.10.015>
- Rauch, H. E., Tung, F., & Striebel, C. T. (1965). Maximum likelihood estimates of linear dynamic systems. *American Institute of Aeronautics and Astronautics Journal*, 3(8), 1445–1450. <https://doi.org/10.2514/3.3166>
- Reeves, J., & Hatch, C. E. (2016). Impacts of three-dimensional nonuniform flow on quantification of groundwater-surface water interactions using heat as a tracer. *Water Resources Research*, 52(9), 6851–6866. <https://doi.org/10.1002/2016WR018841>
- Särkkä, S. (2013). *Bayesian filtering and smoothing* (p. 144). Cambridge University Press.
- Schuesser, W. E. (1991). *The numerical method of lines*. Academic Press.
- Schneidewind, U., van Berkel, M., Anibas, C., Vandersteen, G., Schmidt, C., Joris, I., et al. (2016). LPMLE3: A novel 1-D approach to study water flow in streambeds using heat as a tracer. *Water Resources Research*, 52(8), 6596–6610. <https://doi.org/10.1002/2015wr017453>
- Shapiro, A. M., & Day-Lewis, F. D. (2021). Estimating and forecasting time-varying groundwater recharge in fractured rock: A state-space formulation with preferential and diffuse flow to the water table. *Water Resources Research*, 57(9), e2020WR029110. <https://doi.org/10.1029/2020WR029110>
- Shapiro, A. M., & Day-Lewis, F. D. (2022). Reframing groundwater hydrology as a data-driven science. *Groundwater*. <https://doi.org/10.1111/gwat.13195>
- Shapiro, A. M., Day-Lewis, F. D., Kappel, W. M., & Williams, J. H. (2022). Incorporating snowmelt into daily estimates of recharge using a state-space model of infiltration. *Groundwater*, GWAT13206. <https://doi.org/10.1111/gwat.13206>
- Simon, D. (2006). *Optimal state estimation: Kalman, H infinity, and nonlinear approaches* (p. 552). Wiley-Interscience.
- Sinsbeck, M., & Tartakovsky, D. M. (2015). Impact of data assimilation on cost-accuracy tradeoff in multifidelity models. *SIAM/ASA Journal on Uncertainty Quantification*, 3(1), 954–968. <https://doi.org/10.1137/141001743>
- Stallman, R. W. (1965). Steady 1-dimensional fluid flow in a semiinfinite porous medium with sinusoidal surface temperature. *Journal of Geophysical Research*, 70(12), 2821–2827. <https://doi.org/10.1029/jz070i012p02821>
- Stonestrom, D. A., & Constantz, J. (2003). *Heat as a tool for studying the movement of ground water near streams*. Circular 1260, US Geological Survey.
- Sun, L., Ousmane, S., Nistor, I., & Liu, K. (2016). Review of the Kalman-type hydrological data assimilation. *Hydrological Sciences Journal*, 61(13), 2348–2366. <https://doi.org/10.1080/02626667.2015.1127376>
- Umari, A. M. J., & Gorelick, S. M. (1986). *Evaluation of the matrix exponential for use in ground-water-flow and solute-transport simulations: Theoretical framework* (Water-Resources Investigations Report 86-4096). U.S. Geological Survey.
- van Kampen, R., Schneidewind, U., Anibas, C., Bertagnoli, A., Tonina, D., Vandersteen, G., et al. (2022). LPMLEn—A frequency domain method to estimate vertical streambed fluxes and sediment thermal properties in semi-infinite and bounded domains. *Water Resources Research*, 58(3), e2021WR030886. <https://doi.org/10.1029/2021WR030886>
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0 contributors, 2020, SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Voytek, E. B., Drenkelfuss, A., Day-Lewis, F. D., Healy, R., Lane, J. W., Jr., & Werkema, D. (2014). 1DTempPro: Analyzing temperature profiles for groundwater/surface-water exchange. *Groundwater*, 52(2), 298–302. <https://doi.org/10.1111/gwat.12051>

- Young, P. C. (2011). *Recursive estimation and time-series analysis: An introduction for the student and practitioner* (p. 521).
- Young, P. C., Pegregal, D. J., & Tych, W. (1999). Dynamic harmonic regression. *Journal of Forecasting*, *18*(6), 369–394. [https://doi.org/10.1002/\(sici\)1099-131x\(199911\)18:6<369::aid-for748>3.0.co;2-k](https://doi.org/10.1002/(sici)1099-131x(199911)18:6<369::aid-for748>3.0.co;2-k)
- Young, P. C., Taylor, C. J., Tych, W., Pegregal, D. J., & McKenna, P. G. (2004). *The captain toolbox*. Centre for Research on Environmental Systems and Statistics, Lancaster University. Retrieved from [www.es.lancs.ac.uk/cres/captain](http://www.es.lancs.ac.uk/cres/captain)