*Water Resources Research*

Supporting Information for

**A multi-model ensemble of empirical and process-based models improves the predictive skill of near-term lake forecasts**

Freya Olsson1\*, Tadhg Moore1,2, Cayelan C. Carey1, Adrienne Breef-Pilz1, R. Quinn Thomas1,3

1Department of Biological Sciences, Virginia Tech, Blacksburg, Virginia, USA 24061

2Independent researcher

3Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, Virginia, USA 24061

\*Corresponding author (freyao@vt.edu)

**Contents of this file**

Text S1 to S2

Figures S1 to S4

Tables S1 to S2

**Introduction**

This supporting information (SI) file provides additional detailed methods and results to support the primary manuscript. SI method material includes detailed information on process-model parameter calibration and observational uncertainty calculations used to estimate forecast shadowing time. SI results material includes individual model forecasts at three horizons over the full two-year forecast period, depth-disaggregated forecast performance metrics, and a summary figure of the best and worst performing models aggregated at each horizon for forecasts at 1 m and 8 m depths.

Text S1. Parameter tuning for individual process-based models

Each process model parameters were set to model defaults with the exception of identified sensitive parameters which were included in the data assimilation step and allowed to vary (see Figure 1). The sensitive parameters are listed in Table S1 for each individual process model along with their initial parameter ranges, based on lakes of similar sizes in the published literature and a preliminary parameter sensitivity analysis.

Text S2. Calculating observation uncertainty

Observational uncertainty was calculated using sensor data collected at Falling Creek Reservoir (see Methods; Carey et al., 2023). Using these high-frequency data (collected every 10 minutes), the standard deviation in water temperatures was calculated individually at each depth and then averaged (Figure S2).

A graph of a model

Description automatically generatedFigure Figure S1. Evolution of process-model parameter values during the spin-up period (October 2020 - March 2021). Each process model tuned two different parameters: GLM - mean sediment temperature (zone1temp, zone2temp), short-wave radiation factor (sw\_factor); GOTM - short-wave radiation scaling factor (swr\_scale\_factor), wind speed scaling factor (wind\_scale\_factor); and Simstrat - wind fitting parameter (f\_wind), short-wave radiation fitting parameter (p\_sw\_water). Note the differing y-axis for each parameter.

A graph with numbers and lines

Description automatically generated

**Figure S2.**Observational uncertainty (standard deviation, in oC) at each observed depth, calculated from data collected between July 2018 - October 2020 at the deepest point of Falling Creek Reservoir throughout the water column. These values represent the mean standard deviation within each day. The mean daily standard deviation for all depths is shown at the bottom right.

A graph of different colored lines

Description automatically generated

**Figure S3.** Time series of water temperature forecasts generated by the two baseline models (persistence, climatology) and the three individual process models (PM1, PM2, PM3) at 1, 7, and 14 days-ahead for 1 m and 8 m depths. The solid line shows the median forecast and dotted lines show the corresponding 95% confidence intervals. The black points are the observed water temperatures.

A graph of different colors

Description automatically generated**Figure S4.** Proportion of forecast for which each model forecast represents the best forecast (left column; a, b) or worst forecast (right column; c, d) at 1 m (top row) and 8 m (bottom row) across the 1-14 day forecast horizon. The two multi-model ensembles (MMEs), individual baseline models, and individual process models were ranked based on the ignorance score for each individual forecast and the proportion of best (rank=1) and worst (rank=7) calculated at each depth and horizon.

Table S1. Process model (PM) parameters selected for tuning within the FLARE framework during the spin-up period (October 2020 - March 2021) to provide the parameter values used in the first analysed forecasts. Both the native model parameter name and its FLARE equivalent are given, with the starting range of parameter values set at the beginning of the model run.

|  |  |  |  |
| --- | --- | --- | --- |
| Process Model | Model parameter name | FLARE parameter name | Starting range |
| GLM (General Lake Model; PM1) | sediment/sed\_temp\_mean  meteorology/sw\_factor | zone1temp/zone2temp  sw\_factor | 10 to 12  0.5 to 1.5 |
| GOTM (General Ocean Turbulence Model; PM2) | surface/meteo/swr/scale\_factor  surface/meteo/u10/scale\_factor | swr\_scale\_factor  wind\_scale\_factor | 0.5 to 1.2  0.5 to 1.2 |
| Simstrat (PM3) | ModelParameters/f\_wind  ModelParameters/p\_sw\_water | F\_wind  p\_sw\_water | 0.5 to 1.0  0.5 to 1.5 |

**Table S2.** Mean ignorance score (IGN), bias, and standard deviation (SD) across all forecast horizons at 1 m and 8 m for each forecast model (PM = process model; MME = multi-model ensemble).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Depth** | **Model** | **IGN** | **Bias** | **SD** |
| 1 m | Climatology | 2.06 | -0.07 | 2.25 |
| Full MME | 1.89 | 0.41 | 2.32 |
| Persistence | 2.12 | -0.10 | 1.98 |
| PM1 | 2.30 | 1.39 | 154 |
| PM2 | 1.99 | -0.25 | 1.51 |
| PM3 | 2.26 | 1.12 | 1.47 |
| PM MME | 1.98 | 0.76 | 1.76 |
| 8 m | Climatology | 1.13 | 0.01 | 0.81 |
| Full MME | 1.16 | -0.38 | 1.37 |
| Persistence | 2.15 | -0.02 | 0.53 |
| PM1 | 1.37 | -0.27 | 1.36 |
| PM2 | 1.66 | -1.01 | 1.43 |
| PM3 | 1.55 | -0.59 | 1.47 |
| PM MME | 1.50 | -0.63 | 1.51 |