Correspondence to:

W. J. Pringle,
wpringle@anl.gov

Coupled Lake-Atmosphere-Land Physics Uncertainties in a Great Lakes Regional Climate Model

William J. Pringle[1], Chenfu Huang[2], Pengfei Xue[1,2], Jiali Wang[1], Khachik Sargsyan[3], Miraj B. Kayastha[2], T. C. Chakraborty[4], Zhao Yang[4], Yun Qian[4], and Robert D. Hetland[4]

[1]Environmental Science Division, Argonne National Laboratory, Lemont, IL, USA, [2]Department of Civil, Environmental and Geospatial Engineering, Michigan Technological University, Houghton, MI, USA, [3]Sandia National Laboratories, Livermore, CA, USA, [4]Pacific Northwest National Laboratory, Richland, WA, USA

Abstract This study develops a surrogate-based method to assess the uncertainty within a convective permitting integrated modeling system of the Great Lakes region, arising from interacting physics parameterizations across the lake, atmosphere, and land surface. Perturbed physics ensembles of the model during the 2018 summer are used to train a neural network surrogate model to predict lake surface temperature (LST) and near-surface air temperature (T2m). Average physics uncertainties are determined to be 1.5°C for LST and T2m over land, and 1.9°C for T2m over lake, but these have significant spatiotemporal variations. We find that atmospheric physics parameterizations alone are the dominant sources of uncertainty (45%–53%), while lake and land parameterizations account for 33% and 38% of the uncertainty of LST and T2m over land respectively. Interactions of atmosphere physics parameterizations with those of the land and lake contribute to an additional 13%–17% of the total variance. LST and T2m over the lake are more uncertain in the deeper northern lakes, particularly during the rapid warming phase that occurs in late spring/early summer. The LST uncertainty increases with sensitivity to the lake model's surface wind stress scheme. T2m over land is more uncertain over forested areas in the north, where it is most sensitive to the land surface model, than the more agricultural land in the south, where it is most sensitive to the atmospheric planetary boundary and surface layer scheme. Uncertainty also increases in the southwest during multiday temperature declines with higher sensitivity to the land surface model.

Plain Language Summary Regional climate models couple together lake, atmosphere, and land surface components, each containing several simplifications of complex small-scale physics, known as parameterizations. In this study, we explore the uncertainty arising from the choice of interacting parameterizations and associated parameters across coupled lake-atmosphere-land components of a Great Lakes regional climate model. To do this, we train a machine learning surrogate model on the climate model outputs of lake and air surface temperatures during the 2018 summer. The surrogate model is then rapidly queried thousands of times to find the uncertainty range and which physics parameterizations contribute to it. We find that atmospheric physics parameterizations are the dominant sources of uncertainty and that another 15% comes from its interactions with lake and land physics. The surface temperatures are more uncertain over the deeper northern lakes and forested northern land areas than over the shallower lakes and more agricultural land in the south. Uncertainty increases on and over the lakes during a time of rapid warming in the late spring-early summer, and over the southwest land area during periods where the temperature drops over multiple days.

1. Introduction

Uncertainty about the physical processes in the atmosphere, water bodies, and the land surface can contribute to biases and large spread in weather and climate models (Bellprat et al., 2012b; Eidhammer et al., 2024; Ricciuto et al., 2018; Zanna et al., 2019). This is most often attributable to unresolved physics requiring parameterizations that, while often based on theory, are necessarily simplifications requiring several assumptions and free empirical parameters. To appropriately deal with the inevitable presence of this epistemic uncertainty, the weather and climate modeling fields have employed the use of Perturbed Physics/Parameter Ensembles (PPE; Bellprat et al., 2012a; Eidhammer et al., 2024). Several PPE-based studies have investigated processes in the atmosphere (Bellprat et al., 2012a; Qian et al., 2015, 2018, 2024), land surface processes (Ricciuto et al., 2018; Xu et al., 2022), and oceans (Huber & Zanna, 2017). Land-atmosphere interactions were investigated in C. Wang et al. (2021).

Supervision: Pengfei Xue, Jiali Wang, Yun Qian
Visualization: William J. Pringle, Yun Qian
Writing – original draft: William J. Pringle
Writing – review & editing: William J. Pringle, Pengfei Xue, Jiali Wang, Khachik Sargsyan, T. C. Chakraborty, Zhao Yang, Yun Qian

Prior PPE studies have rarely focused on complex coastal systems, where interactions between the atmosphere, land and water bodies can collectively modulate these uncertainties. One region where these interactions are important and not well-assessed by the climate modeling communities is the Great Lakes of North America (Sharma et al., 2018; J. Wang et al., 2022), the world's largest surface freshwater system. Many climate models have used a simplified representation of the lake surface, and thus may not adequately capture the impact of the lakes on the regional climate (Briley et al., 2021). However, several recent studies have now incorporated coupled three-dimensional (3-D) hydrodynamic processes of the lakes into regional climate models focused on the Great Lakes region (GLR; Kayastha et al., 2023; Sun et al., 2020; Xue et al., 2017, 2022). To date, investigation of physics parameterizations in GLR regional climate models has been limited to ad-hoc performance evaluation experiments without the 3-D lake (Notaro et al., 2021), and without a formal PPE analysis.

One of the reasons for the lack of in-depth analysis of parameterizations of unresolved physics across the lakes, atmosphere, and land is that running coupled models with perturbations across all model components is computationally costly. To mitigate this issue, PPE studies have adopted the use of surrogate models (Ricciuto et al., 2018). Surrogate models are extremely inexpensive to query, often based on linear models or polynomials (Bellprat et al., 2012b; Qian et al., 2018; Ricciuto et al., 2018), so that the entire uncertainty space can be explored. However, this requires the generation of an adequate surrogate model in the first place. Further, we are typically interested in how the uncertainties and interactions vary across space and time, hence we need a way to deal with the large dimensionality of the model.

This paper focuses on addressing the problem of assessing physics uncertainties across a complex atmosphere-land-lake system. We assume that the choice of a set of physics parameterizations are equally reasonable a priori, and the range of model outputs represent an inherent uncertainty of the model due to incomplete physics. To efficiently assess the uncertainty we develop a surrogate model-based framework that is, in principle, generalizable to all regional climate models and apply it to simulations of the GLR with a newly developed coupled modeling system (Kayastha et al., 2023). We investigate the sensitivity of (near-)surface temperature, a critical component of the surface energy budget, to variations in parameterizations/parameters. We quantify the contribution of each parameterization/parameter to this uncertainty, including the effects of interactions between parameterization/parameters, which is made possible through the presented training methodology of the nonlinear surrogate model. The variation of this physics uncertainty across space and time and the importance of the coupled effects are highlighted.

2. Methods and Data

2.1. Model Description and Experimental Design

2.1.1. Coupled Lake-Atmosphere Model Configuration

We use a two-way coupled atmosphere and 3-D hydrodynamic lake modeling setup developed by Kayastha et al. (2023) for our PPE analysis. The atmosphere component is the Weather Research and Forecasting (WRF) model v4.2.2 (Skamarock et al., 2021) with the Advanced Research WRF dynamic core (Skamarock & Klemp, 2008). The hydrodynamic lake component is based off of the Finite Volume Community Ocean Model (FVCOM; Chen et al., 2003) v4.1 (see more descriptions below). The coupled atmosphere-lake model domain is centered at 45.5°N and 85.0°W and has dimensions of 543×484 grid points in the west-east and south-north. Grid spacing is 4 km, covering the GLR (Figure 1). There are 50 stretched vertical levels topped at 50 hPa. The initial and boundary conditions are from 3-hourly 0.25° European Centre for Medium-Range Weather Forecasts atmospheric reanalysis of the global climate, version 5 (ERA5; Hersbach et al., 2020).

Three commonly used physics options each for cloud microphysics (MP), longwave and shortwave radiation (LW + SW RA), and planetary boundary layer and surface layer (PBL + SFC), and two land surface models (LSM) are used in this study, as summarized in Table 1. The default configuration from our previous studies (Kayastha et al., 2023; J. Wang et al., 2022) is indicated in the table. In this study, we modified the source codes of the SFC scheme in WRF so that each SFC scheme can use one of any three chosen parameterizations of surface roughness length over water ($z0_w$): constant Charnock coefficient of 0.0185, the Coupled Ocean–Atmosphere Response Experiment (COARE) 3.5 algorithm (Edson et al., 2013), or the depth-dependent over shallow waters (<100 m) scheme (Jiménez & Dudhia, 2018). Our hypothesis is that the depth-dependent scheme should improve surface variables over the Great Lakes as a large portion is <100 m deep (Figure 1). Note that $z0_w$ is used
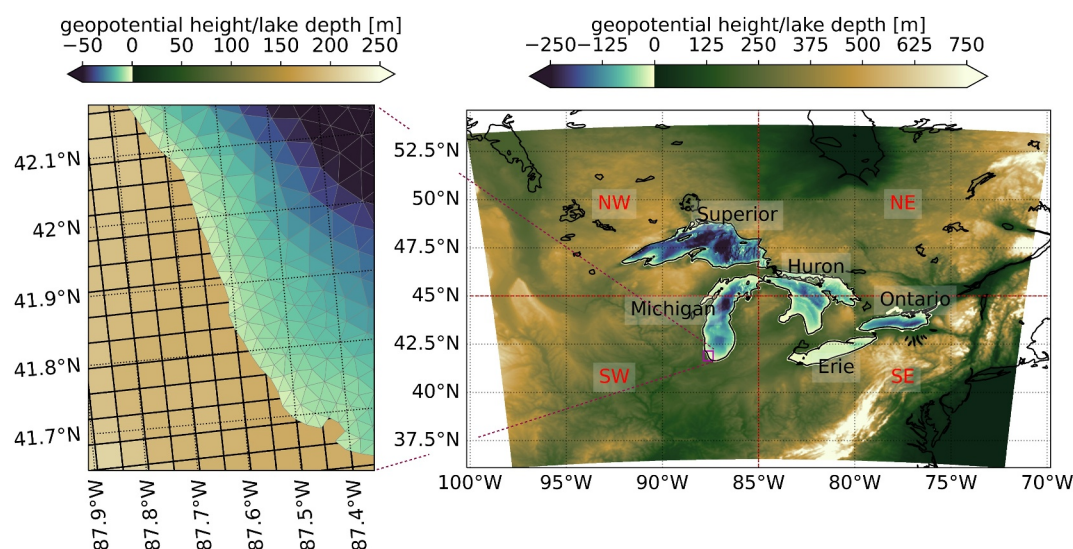
Figure 1. Computational domain of the Weather Research and Forecasting-Finite Volume Community Ocean Model (WRF-FVCOM) model with zoom-in view of grid sizes near the Chicago coastline. Shading shows the geopotential height in the WRF atmosphere model and bathymetric depths in the FVCOM lake model. Labels denote lake names and land area quadrants.

directly as the momentum roughness length and is modified internally by the SFC scheme to obtain different temperature and moisture roughness lengths. This modification was found to improve the modeling latent and sensible heat fluxes over the Great Lakes (Charusombat et al., 2018).

The hydrodynamic lake component is based off of the Finite Volume Community Ocean Model (FVCOM; Chen et al., 2003) v4.1. Horizontal resolution on an unstructured triangular mesh varies from 2 to 4 km offshore to 1–2 km along the coasts (Figure 1), and the model has 40 vertical sigma layers (Kayastha et al., 2023). The model is initialized from a quiescent state in the winter when the lakes are unstratified with an initial uniform lake temperature of 2°C. As in Kayastha et al. (2023), FVCOM is run simultaneously with WRF, including a two-way information exchange between them at 1-hr intervals using the OASIS3-MCT coupler (Craig et al., 2017). Here, the lake surface temperature (LST) and ice cover are dynamically calculated by FVCOM and are provided to WRF as overlake surface boundary conditions using an inverse distance weighting interpolation. Only WRF cells that are designated water by its land-water grid mask and are within the FVCOM domain receive information from FVCOM. In turn, the atmospheric forcings required by FVCOM are dynamically calculated and provided by WRF using bilinear interpolation regardless of the WRF land-water grid mask value.

FVCOM parameterization options and parameters are summarized in Table 1. Vertical mixing is modeled with either the Mellor-Yamada Level 2.5 (MY-2.5) closure (Mellor & Yamada, 1982) or the General Ocean Turbulence Model implementation of the k-ϵ closure (Stips et al., 2002). We use the Large and Pond (1981) or Andreas et al. (2012) bulk wind stress (WS) parameterization to impart momentum from the atmosphere into the lake at the surface. Lastly, we vary two uncertain parameters: the turbulent Prandtl number (Pr$_t$) and R: the fraction of the downward shortwave flux associated with the longer wavelength irradiance. We vary Pr$_t$ from 0.1 to 1 and R from 0.74 to 0.78, the latter representing a range from relatively clear Type IA water to lower clarity Type III water (Paulson & Simpson, 1977). We also assign values to the two other variables in the shortwave radiation absorption formulation dependent on R: the attenuation depth for the longer wavelength component of shortwave irradiance, $\zeta_1 = 1.7 + 0.3\alpha_R$ and the attenuation depth for shorter wavelength component of shortwave irradiance, $\zeta_2 = 6 + 9.7\alpha_R$, in which $\alpha_R = (0.74 - R)/0.04$, based on values from Paulson and Simpson (1977); Chen et al. (2011).

2.2. Model Uncertainty Analysis

We use a surrogate model framework for the forward uncertainty propagation of the expensive coupled lake-atmosphere-land physics model. In this framework, physics parameterizations and parameters are assigned a

Table 1
Uncertain Physics Parameterizations and Parameters in the Coupled WRF-FVCOM Model

Parameter(ization)	Definition	Options	Priors
PBL + SFC	WRF planetary boundary layer and surface scheme	YSU + MM5$_{rev}$[a] MYJ + MOJ MYNN-2.5 + MYNN	$\mathcal{U}_d[1,3]$
MP	WRF cloud microphysics scheme	Morrison Thompson[a] WSM6	$\mathcal{U}_d[1,3]$
LW + SW Rad	WRF longwave and shortwave radiation scheme	RRTM + CAM3 RRTMG + RRTMG[a] Goddard + Goddard	$\mathcal{U}_d[1,3]$
z0$_w$	WRF surface roughness length over water scheme	COARE 3.5 Charnock = 0.0185[a] Depth-dependent	$\mathcal{U}_d[1,3]$
LSM	WRF land surface model	Noah[a] Noah-MP	$\mathcal{U}_d[1,2]$
VM	FVCOM vertical mixing scheme	MY-2.5[a] GOTM k-ϵ	$\mathcal{U}_d[1,2]$
WS	FVCOM bulk wind stress parameterization	Large and Pond (1981)[a] Andreas et al. (2012)	$\mathcal{U}_d[1,2]$
Pr$_t$	FVCOM turbulent Prandtl number	Value from 0.1 to 1[a]	$\mathcal{U}[0.1,1]$
R	FVCOM shortwave radiation absorption fraction	Value from 0.74 to 0.78[a]	$\mathcal{U}[0.74,0.78]$

Note. Detailed descriptions of WRF parameterizations can be found in Skamarock et al. (2021). See text for description of FVCOM parameterizations and parameters. YSU: Yonsei University PBL, MYJ: Mellor-Yamada-Janjic PBL, MYNN-2.5: Mellor-Yamada-Nakanishi-Niino Level 2.5 PBL, MM5$_{rev}$: revised fifth-generation PSU–NCAR Mesoscale Model SFC, MYJ: Monin-Obukhov-Janjic SFC, MYNN: Mellor-Yamada-Nakanishi-Niino SFC, Morrison: Morrison double-moment 6-class MP, Thompson: Thompson double-moment 6-class MP, WSM6: WRF single-moment 6-class MP, RRTM: Rapid Radiative Transfer Model RAD, RRTMG: Rapid Radiative Transfer Model for General Circulation Models RAD, CAM3: NCAR Community Atmosphere Model 3.0 RAD, Goddard: New NASA Goddard RAD, Noah: Unified Noah LSM, Noah-MP: Noah-multiparameterization LSM, MY-2.5: Mellor-Yamada Level 2.5 VM, GOTM k-ϵ: General Ocean Turbulence Model k-ϵ VM.
[a]Indicates settings of the default setup from Kayastha et al. (2023).

statistical distribution (typically uniform or Gaussian) a priori by the modeler, hereafter termed the "parametric priors" and denoted by λ. For the physics parameterizations, a discrete uniform distribution is chosen, $\mathcal{U}_d[1,N_s]$, where the value of the integer corresponds to a specific scheme in the range from 1 to the number of chosen schemes, N_s. As the order of these integers is unimportant, one hot encoding is used to covert to binary representations before surrogate model training. For the parameters, a continuous uniform distribution, $\mathcal{U}[a,b]$, is chosen to range between deemed plausible values a and b. The parametric priors for the coupled lake-atmosphere-land model are summarized in Table 1.

To generate the surrogate model for the uncertainty analysis, a training set is required from samples of the physics model. To form this training set, the joint distribution of the parametric priors must be sampled efficiently to minimize the number of expensive coupled model simulations, typically using Quasi-Monte Carlo (QMC) or Latin hypercube sampling (Qian et al., 2015). The sampling method used in this study (QMC) is detailed in Section 2.3. From this training set, a surrogate model is generated to approximate the quantity of interest (QoI),

$$Z = f(\lambda, \boldsymbol{x}, t) \approx g(\lambda, \boldsymbol{x}, t) \tag{1}$$

where Z is the spatiotemporally (x indicates spatial location and t is time) varying modeled QoI as a function of the parametric priors λ, and can be approximated by the surrogate model, g. In this study, we use a multilayer perceptron artificial neural network (NN) surrogate model with a Rectified Linear Unit (ReLu) activation function, implemented with *PyTorch* (Paszke et al., 2019). We also implemented a Polynomial Chaos (Sargsyan et al., 2014) surrogate model in the code distributed with this study. However, it did not perform as well as the NN model, so we do not show those results here. This is expected for polynomials, as they are not well-suited to discrete inputs that we use for physics parameterizations, but we included it in the code for potential future use in an analysis with only continuous parameter inputs (e.g., Ricciuto et al., 2018).

Based on the formation of g, a global sensitivity analysis (GSA) and the determination of the QoI distribution can be performed rapidly through Monte Carlo sampling of the surrogate models. For the GSA, we use variance-based Sobol sensitivity indices to indicate the fractional contributions of each parameter or group of parameters to the total variance of the QoI (Sargsyan, 2017). The indices are defined as,

$$S_i = \frac{\mathbb{V}[\mathbb{E}(f(\lambda|\lambda_i)]}{\mathbb{V}[(f(\lambda)]} \tag{2}$$

$$S_i^T = 1 - \frac{\mathbb{V}[\mathbb{E}(f(\lambda|\lambda_{-i})]}{\mathbb{V}[(f(\lambda)]} \tag{3}$$

$$S_{ij}^J = \frac{\mathbb{V}[\mathbb{E}(f(\lambda|\lambda_i,\lambda_j)]}{\mathbb{V}[(f(\lambda)]} - S_i - S_j \tag{4}$$

where $\mathbb{V}$ is the variance with respect to the fixed parameter (λ_i, (λ_i,λ_j) pair, or the rest of the parameters, λ_{-i}), and $\mathbb{E}$ is the expectation with respect to the rest of the parameters (Sargsyan et al., 2014). S_i is the main effect sensitivity index of parameter i, which measures the first-order effect; S_i^T is the total effect index of parameter i, which measures the total effect of a parameter including the interactions with all other variables; and S_{ij}^J is the joint-effect index of parameter i interacting with parameter j. In this work, we use the Saltelli et al. (2010) sampling-based approach to compute Equations 2–4. This method requires generating two random samples each of size M and $(N_d + 2)M$ corresponding surrogate model evaluations (Sargsyan, 2017). We determined that $M = 10,000$ is sufficient for convergence in the current study. The GSA samples are also used to calculate the uncertainty, which we define as the 90% interval, that is, the range between the 5th and 95th percentiles of the distribution.

2.3. Computational Design and Surrogate Model Construction

Our simulations target the analysis of 2018 summer season (JJA), with May 2018 used as the spinup period (Kayastha et al., 2023). During this period, the lakes go through a rapid spring warming phase due to radiatively driven convection (Austin, 2019) and become strongly stratified with stable lake temperatures for the rest of the season. The warming phase occurs earliest in May-June for the southern shallow lakes (Erie and Ontario) and latest in June–July for the northern deep lakes (Superior, Huron, and Michigan). This season thus presents an interesting period of time to investigate the lake-atmosphere coupling and its sensitivity to different forms of physics parameterizations and parameter quantities. The model performance has been evaluated compared to observations in Kayastha et al. (2023).

An ensemble of simulations is performed by sampling the $9(=N_d)$ parametric priors (Table 1) with a QMC low-discrepancy Korobov sequence from *chaospy*. A total of 18 ensemble members ($=N_e$) are requested for training. This number was chosen because, (a) close to 90% (actually 89.5%) of the distribution of a continuous variable is sampled using the Korobov sequence (W. J. Pringle et al., 2023), and (b) the discrete variables have either 2 or 3 unique values, so 18 samples each parameterization option an equal number of times (6 or 9), which is a requirement of uniform design (Fang et al., 2000), a method akin to the present approach given our assumption of uniformly distributed priors. A test set of nine ensemble members ($=N_{et}$) generated using Latin hypercube sampling is also produced for validation purposes only.

The QoIs chosen for the sensitivity assessment are daily mean 2-m air temperature (T2m) and LST. Hourly outputs from the WRF-FVCOM at each available grid point are processed into the daily quantities. Analysis is

performed on the full spatiotemporal data set for each QoI. In total we have $N_t = 92$ days with $N_p = 262,812$ WRF grid points for T2m and $N_p = 35,749$ FVCOM grid points for LST. $N_e = 18$ ensemble members, producing a $[N_e \times N_t \times N_p]$ matrix, which is reshaped to $[N_e \times N_t * N_p]$ for generality. However, training a surrogate model for each point in space-time $(N_t * N_p)$ is computationally costly, so we first apply a dimensionality reduction of the problem using Karhunen-Loève expansions (KLE). Using this decomposition, the QoI approximation can be written as,

$$Z = f(\lambda, \boldsymbol{x}, t) \approx \bar{f}(\boldsymbol{x}, t) + \sum_{j=1}^{L} \xi_j(\lambda) \sqrt{\mu_j} \phi_j(\boldsymbol{x}, t) \tag{5}$$

in terms of uncorrelated, zero-mean, unit-variance random variables $\xi_j(\lambda)$ and eigenvalue-eigenfunction pairs $(\mu_j, \phi_j(\boldsymbol{x}, t))$ of the covariance, truncated at eigenvalue L that explains a user-defined level of variance. $\bar{f}(\boldsymbol{x}, t)$ indicates the ensemble mean. In the current problem, we truncate to $L = N_d$ modes of variation that require surrogate approximation. This explains 91% and 93% variance for LST and T2m training sets, respectively. Truncating here produces inverse KLE transform errors of similar magnitude between the training and test sets, limiting overfitting to the training set.

We make a prediction for all the eigenmodes using a single NN surrogate model, minimizing the following weighted smooth L1 loss function,

$$\mathcal{L} = \frac{\sum_{n=1}^{N_b} \sum_{j=1}^{L} l_{j,n} \mu_j}{N_b \sum_{j=1}^{\infty} \mu_j} \tag{6}$$

where,

$$l_{j,n} = \begin{cases} 0.5(\hat{\xi}_{j,n} - \xi_{j,n})^2, & \text{if } |\hat{\xi}_{j,n} - \xi_{j,n}| < 1 \\ |\hat{\xi}_{j,n} - \xi_{j,n}| - 0.5, & \text{otherwise.} \end{cases} \tag{7}$$

in which $\hat{\xi}_{j,n}$ is the predicted value for the eigenmode, j of the ensemble member n in the batch of size, N_b ($\subseteq N_e$). The training set is divided into three equal-sized batches for training. Ensemble members that are not included in each batch are used as the validation set for that batch. The NN model uses two hidden layers with 15 neurons in each, which is the midpoint between the size of the input layer (=20 after one hot encoding) and the output layer ($L = 9$). A small dropout level (=0.001), is used to reduce overfitting to the small training set (a higher dropout level produces surrogates with unreasonably high joint effect sensitivity indices). The process above is repeated nine times with different random seeds (111, 222, …, 999) to provide an uncertainty range for surrogate prediction that can be incorporated into the uncertainty analysis.

The surrogate approximations in KLE space can be converted back to space-time dimensions through the inverse KLE transform. As the full space-time dimension is very large, for our analysis we manipulate $\phi_j(\boldsymbol{x}, t)$ and $\bar{f}(\boldsymbol{x}, t)$ for a subset of interest such as by averaging these functions across $\boldsymbol{x}$ and/or t, and by selecting a subset of $\boldsymbol{x}$ corresponding to a specific lake or area of land. By averaging across both $\boldsymbol{x}$ and t, we obtain the spatiotemporal average for a set of input parametric priors, $\bar{f}(\lambda)$, and use this to summarize the surrogate model errors (Figure 2). From this summary perspective, the surrogate models for LST and T2m are shown to be able to predict the test set with moderate to strong correlation, respectively. The weaker correlation for LST could indicate that some nonlinear physical mechanisms in the lake or atmosphere-lake exchange are not fully represented by the parametric priors.

3. Results

3.1. Overall

The overall physics uncertainties are estimated as 1.50°C, 1.87°C and 1.47°C for LST, T2m over lake, and T2m over land, respectively, in terms of the 90% interval of $\bar{f}(\lambda)$ according to the surrogate model (Figure 3). This
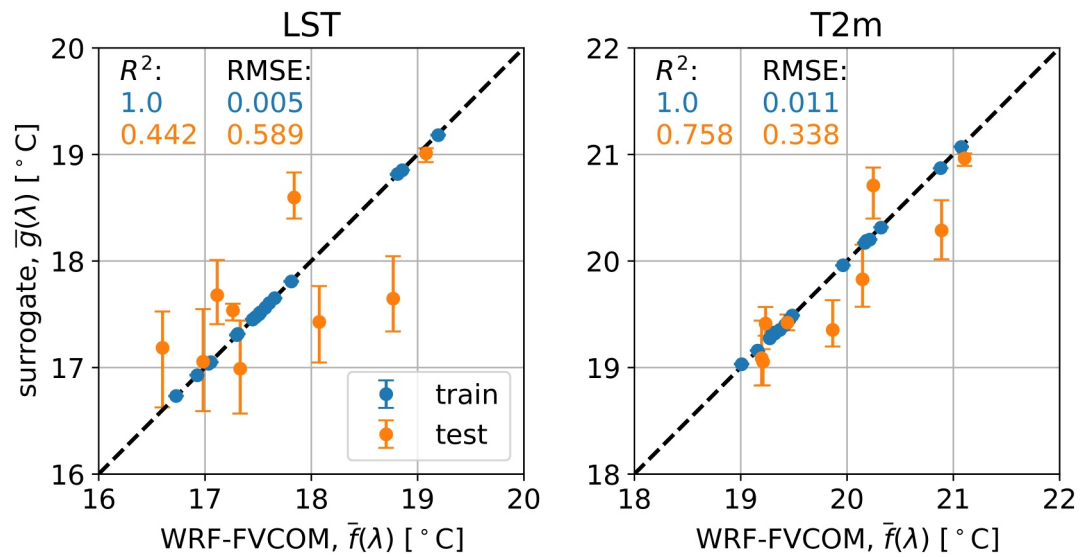
Figure 2. Accuracy of surrogate model predictions of the spatiotemporally averaged lake surface temperature and T2m for the training set ($N_e = 18$) and test set ($N_{et} = 9$) with error bars showing the mean and range across the nine random seeds.

highlights the important role that the Great Lakes play on the atmosphere by contributing to greater uncertainty of T2m over the lakes even though the average air temperature over lakes is smaller. Comparing to the distribution of the training set, we find that the 90% interval is reduced in the surrogate model, mostly as a result of a reduction in the 95th percentile value. On the other hand, the surrogate model produces a wider absolute range (maximum minus minimum) and is especially less constrained at the low end of the distribution than the training set.
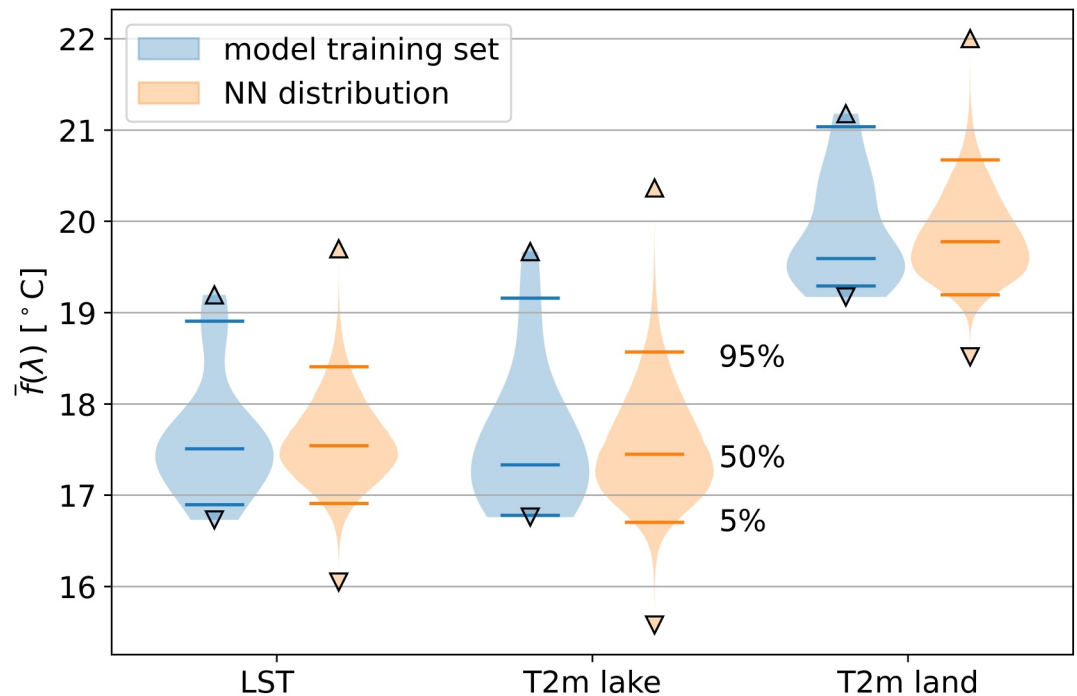

Figure 3. Comparison of the distribution of the spatiotemporally averaged lake surface temperature and T2m over the lake and land surface, $\bar{f}(\lambda)$, between the surrogate model and the physical model training set. Surrogate model distribution is the concatenation across the nine random seeds. Triangle markers indicate maximum/minimum values. The range between the 95% and 5% lines are used to define the overall uncertainty in this study.
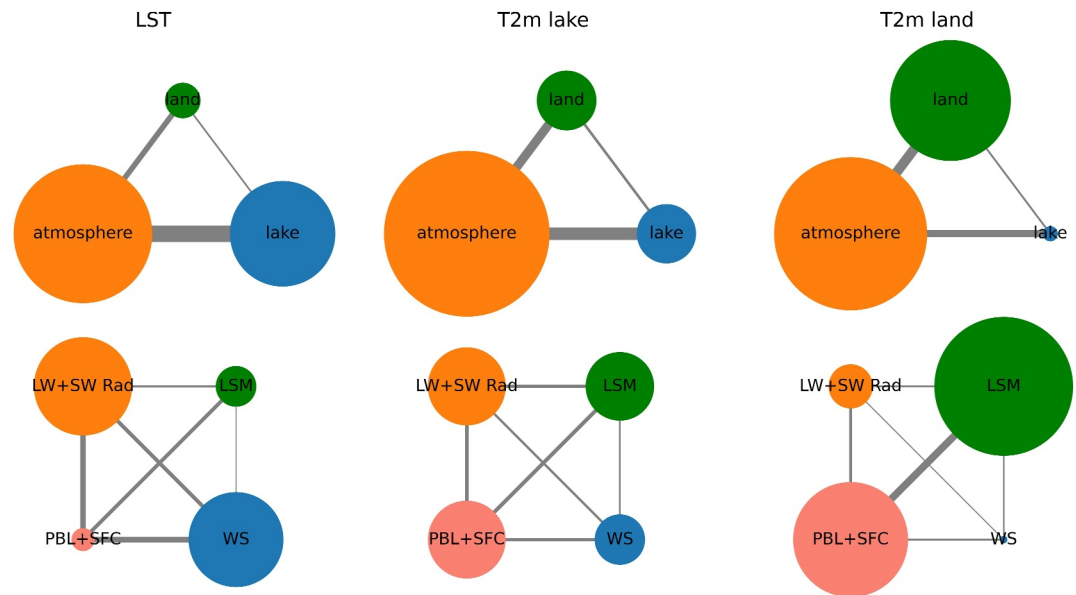
Figure 4. Relative importance of parameterizations/parameters and their interactions to lake surface temperature and T2m over the lake and land according to the surrogate model. Top: Combined atmosphere, land, and lake parameterizations/parameters. Bottom: Top four important parameterizations (refer Figure 5). The radius of circle corresponds to main effect sensitivity while the width of the gray line corresponds to joint effect sensitivity (same scaling).

A plurality of the physics uncertainty is driven by parameterizations in the atmosphere for both LST and T2m (45%–53%, Figure 4). Depending on the QoI, the impact of lake and land physics and the three-way interactions are also important. On average, LST is more sensitive to lake physics (33%) than land physics (11%) and vice versa for T2m over land (38% vs. 5%), but note that spatial variations exist (see Section 3.2). For T2m over lake we find that the relative importance of the land and the lake are near equal. Furthermore, atmosphere-lake physics interactions (13%) are found to be more important than land physics (11%) for LST. For T2m over the lake, both atmosphere-lake (9%) and atmosphere-land (7%) interactions are significant, approximately half the main effect sensitivity of land and lake physics. For T2m over land, atmosphere-lake interaction sensitivity (5%) is the same as the main effect lake sensitivity and only slightly less than the atmosphere-land interaction effect (7%). The land-lake interaction sensitivity is understandably very small for all QoIs (1%–2%) as there is only an indirect connection through the atmosphere.

For both LST and T2m, we find that the four most significant parameterizations are LW + SW Rad, PBL + SFC, LSM, and WS (Figures 4 and 5). This makes physical sense, as the radiative heat fluxes are the main energy source, and the other parameterizations control how the heat is distributed at the surface. LSM is totally representative of the land contribution, and WS is mostly representative of the total lake contribution. While the importance of atmospheric physics is similar across the QoIs, the contribution from LW + SW Rad and PBL + SFC schemes changes. For LST, the LW + SW Rad is dominant and the effect of PBL + SFC is small, since WS is more important for control of the lake surface. Here, both LW + SW Rad and WS account for 30%–40% of the total variance. The opposite is true for T2m over land where the LSM and PBL + SFC scheme dominate, accounting for approximately 40% and 50% of the total variance, respectively. This also makes physical sense as T2m is a quantity interpolated between skin temperature and the lowest atmospheric layer, which are both modulated by LSM and PBL + SFC. While for T2m over lake, the PBL + SFC and LW + SW Rad schemes have comparable impacts of around 30% explained variance. Here, WS explains just below 20% of the variance and LSM explains just above 20%.

The most important joint effect sensitivities for LST are between LW + SW Rad, PBL + SFC and WS as these interactions control the available heat flux into the lake. The same is mostly true for T2m over lake except these sensitivities are reduced while the interaction effect between PBL + SFC and LSM is increased. For T2m over land, this interaction between LSM and PBL + SFC becomes the most significant joint effect sensitivity. The other WRF schemes, PBL + SFC, MP, $z0_w$ and the LSM all have similar magnitudes of sensitivity that explain
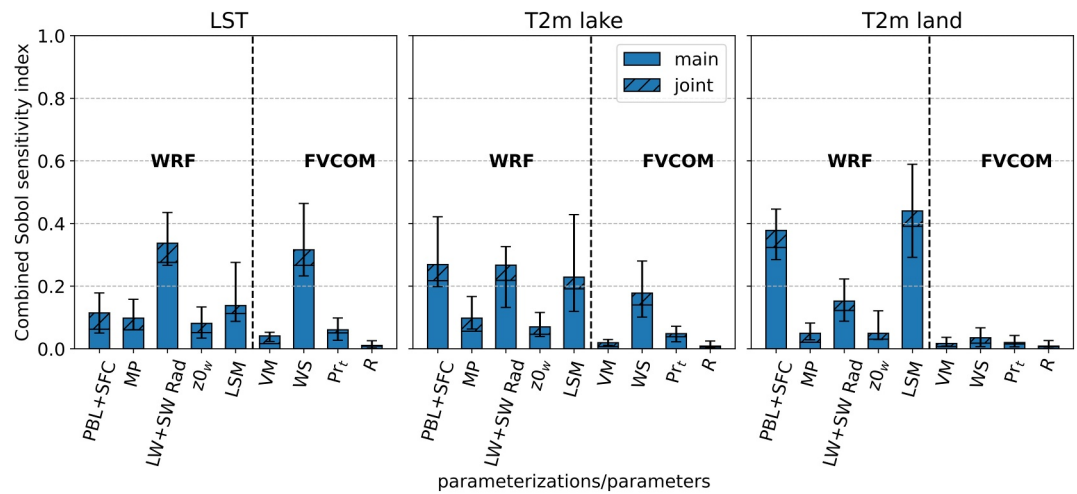
Figure 5. Relative importance (contributions to the total variance, split into main and joint effect Sobol sensitivity components) of all Weather Research and Forecasting and Forecasting-Finite Volume Community Ocean Model parameterizations/parameters (see Table 1 for explanations) to lake surface temperature and T2m over the lake and land according to the surrogate model. Error bars show the range across the nine random seeds.

between 10% and 20% of the variance. Much of this is composed of joint effect sensitivity (except for LSM), meaning a large part of that uncertainty arises from their interactions across different combinations because the coupling does not directly involve the output from these physical parameterizations.

3.2. Spatial Variations

The uncertainty of LST is shown to be largest in the deeper central areas of Lake Superior (Figure 6). Indeed, the uncertainty tends to correlate with the size of the lake, with Lakes Erie and Ontario having the smallest uncertainty on average. This emphasizes the importance of being able to realistically model large and deep lakes with a 3-D hydrodynamic model such as that used in this study.

The higher uncertainty of LST over central Lake Superior is not attributable to any one particular parameterization, but the sensitivity from most of the parameterizations are somewhat elevated here (Figure 7). In general, though, offshore regions are more sensitive to radiation while coastal regions in the northern larger lakes are more sensitive to the lake WS parameterization. This could be explained by the cyclonic summer circulation patterns of the lakes, where currents are generally faster along coastal areas than offshore (Bai et al., 2013). The PBL + SFC scheme is shown to produce higher sensitivity values in specific shallow areas of lakes such as western Lake Erie and the western tip of Lake Huron, where current speeds are very low (Bai et al., 2013). Southeast Lake Michigan is the one area particularly sensitive to the LSM. Although not particularly important, the effect of $z0_w$ is most significant in the shallower southern lakes, which matches the increase in surface roughness as the lake depth decreases for the depth-dependent scheme.

The spatial map shows that T2m uncertainty is highest over land near the northern Great Lakes, over central Lake Superior and in urban areas (Figure 6). The uncertainty of T2m is lowest over land in the southern half of the domain, which explains why the uncertainty of T2m over land is found to be generally lower than over lake (Figure 5). There is a strong correlation between the uncertainty of T2m and LST over Lake Superior (i.e., lower uncertainty in the eastern and western potions of Lake Superior and high uncertainty in the deep central area). T2m is especially uncertain to the west and east of central Lake Michigan and over Chicago and Detroit, which are the largest urban areas in the domain.

For T2m, we see that the radiation and lake WS parameterizations are the most important over the lakes, in agreement with LST (Figure 8). There is also some minor effect of $z0_w$ over some regions of the lakes and coastal urban areas such as Chicago and Detroit, where T2m can be modulated by lake breezes (J. Wang, Qian, et al., 2023). The LSM dominates T2m sensitivity over land around the Great Lakes and to the north where there is a high density of forested area. Noah-MP LSM incorporates several relevant augmentations to Noah, such as a
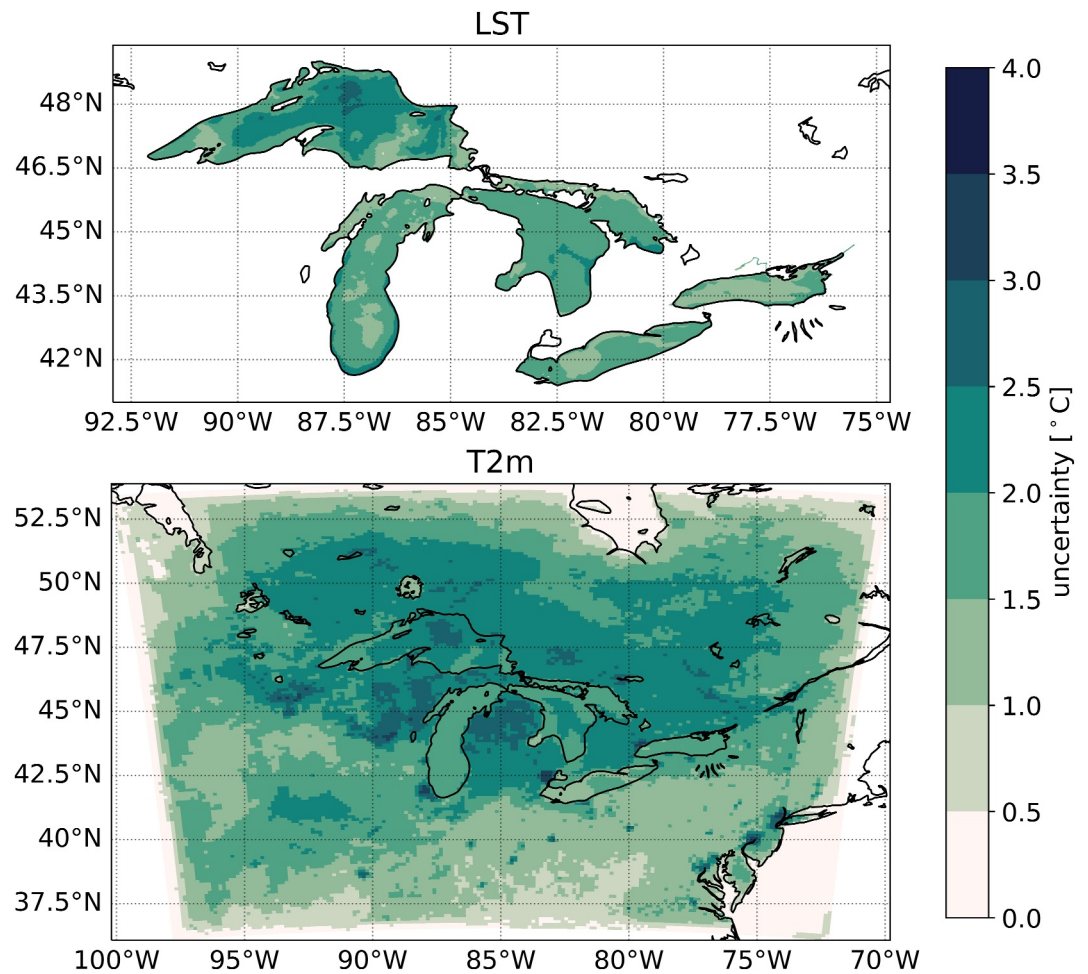
Figure 6. Spatial distribution of lake surface temperature and T2m physics uncertainty, defined as the 90% interval of the time-averaged surrogate model distribution.

separate vegetation canopy layer with canopy gaps that could produce this effect (Niu et al., 2011). On the other hand, the land around the southern lakes and further south is mainly agricultural, in which differences between the LSMs are small. PBL + SFC is dominant in the south and over urban areas. This makes physical sense due to the southwesterly prevailing winds in the summer (Bai et al., 2013) that the PBL + SFC scheme can have an impact on. Further, urban areas like Chicago are very sensitive to treatment at the surface and can be improved by coupling to an urban canopy model (J. Wang, Qian, et al., 2023).

3.3. Temporal Variations

The uncertainty of LST is shown to be fairly consistent across time for each of the lakes during the summer months studied here (Figure 9). Slightly higher uncertainty is present during the rapid warming phase, which occurs in late June–July for Lake Superior and in June (and earlier) for the other lakes. Lake Superior is shown to have the highest LST uncertainty on average ($1.77 \pm 0.52\,°C$) while Lake Ontario has the lowest ($1.51 \pm 0.31\,°C$). The high uncertainty in the rapid spring warming phase coincides with elevated sensitivity to WS across all the lakes, highlighting how parameterization of momentum flux at the surface affects the way excess heat entering the lake during that period is distributed. The importance of WS also seems to correlate with lake size or latitude, with Lake Superior and Huron showing the highest sensitivity to WS. The magnitudes of sensitivity across the different parameterizations are more equal in Lake Michigan and Ontario than in the other lakes. They are also the two lakes where we see the largest spikes in LSM sensitivity in mid-July and mid-August (although all lakes show this spike but to a lesser degree), that are spatially focused on their east coasts according to Figure 8. Both times
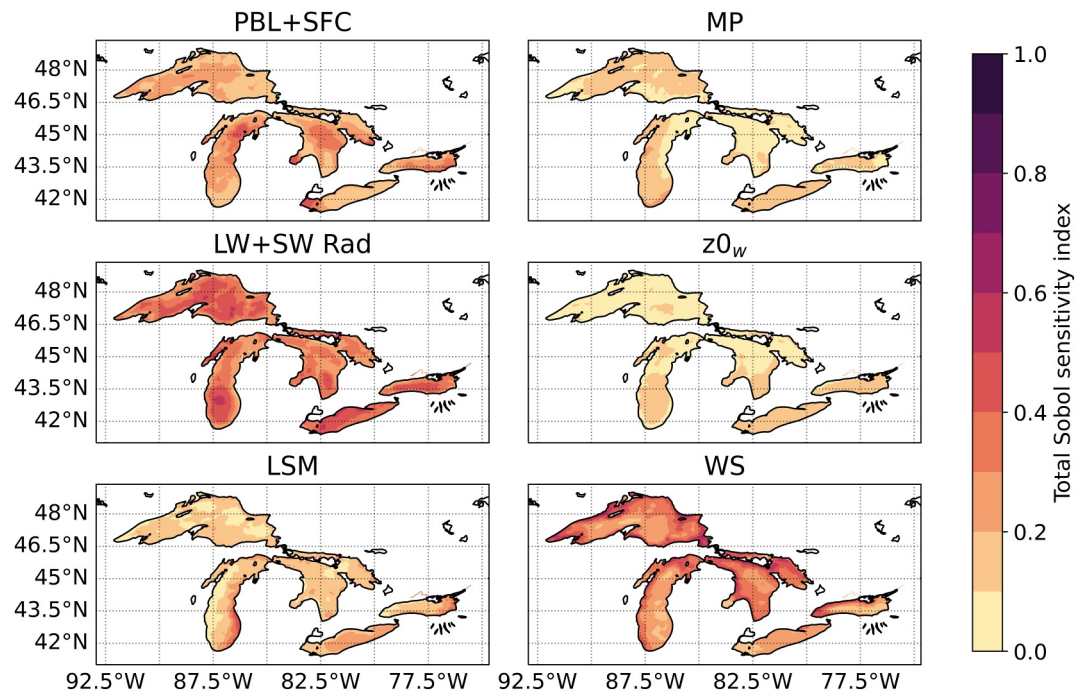
Figure 7. Spatial variation of lake surface temperature total effect Sobol sensitivity indices for the six most important physics parameterizations (refer Figure 5) according to the time-averaged surrogate model.
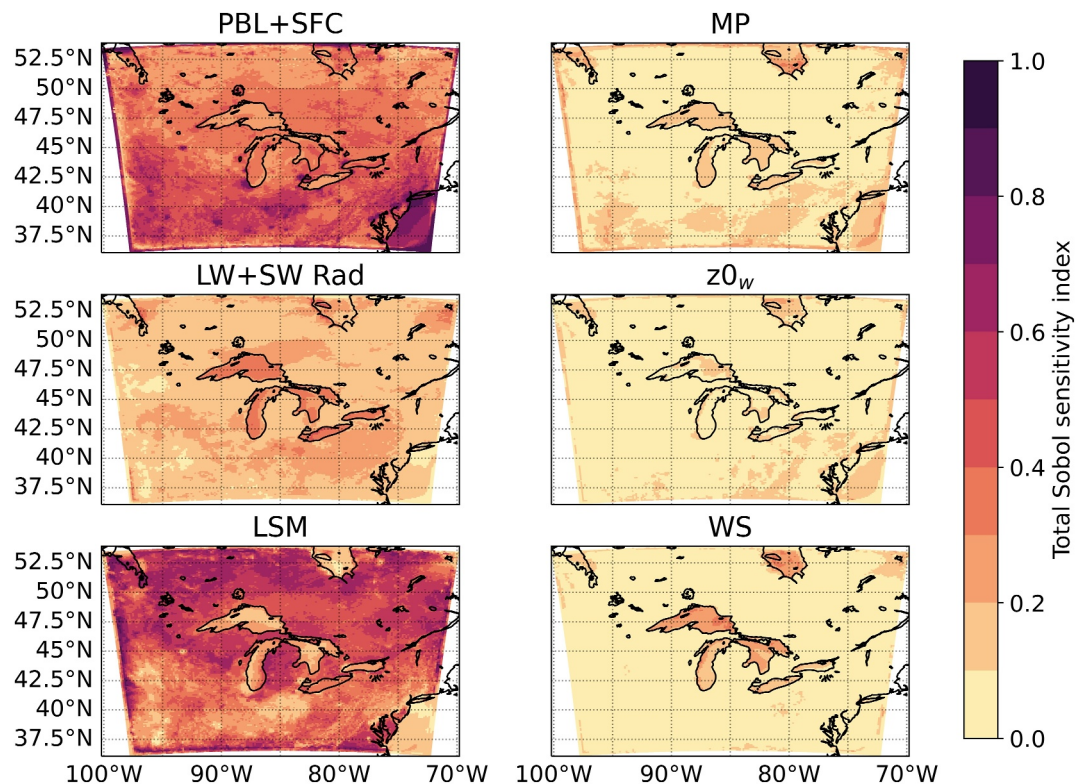

Figure 8. Spatial variation of T2m total effect Sobol sensitivity indices for the six most important physics parameterizations (refer Figure 5) according to the time-averaged surrogate model.
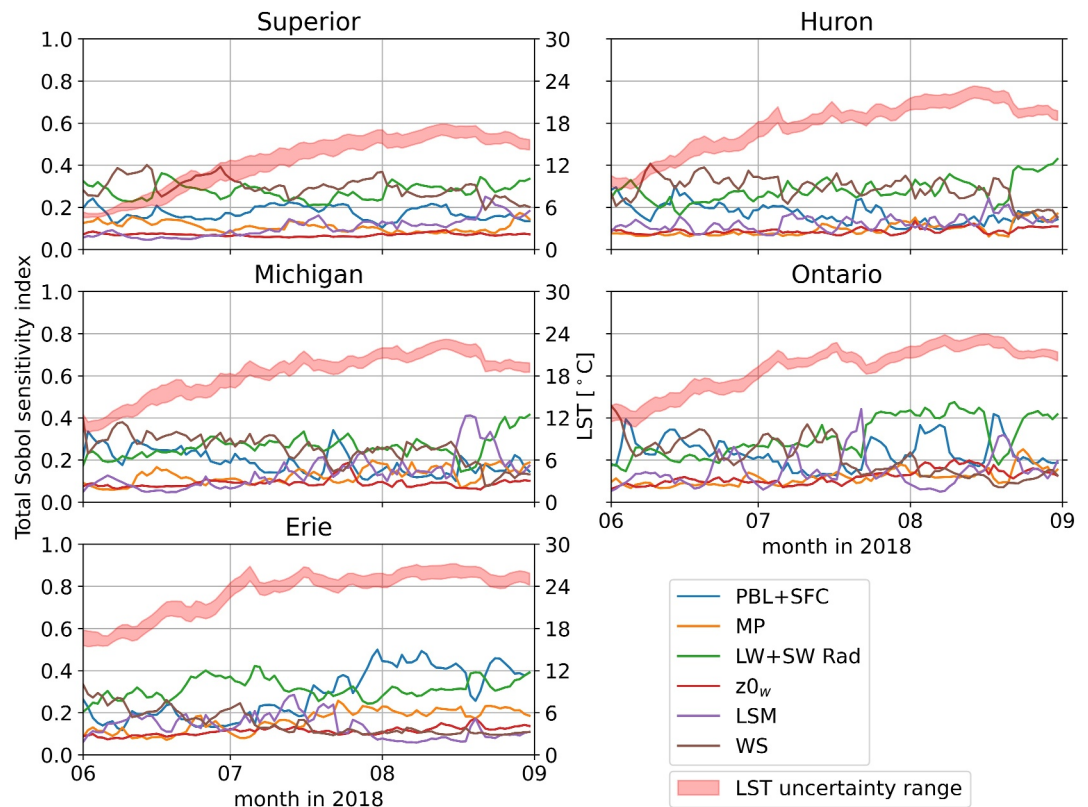
Figure 9. Temporal variation of lake surface temperature physics uncertainty (the range of the shaded area between the 5th and 95th percentiles) and total effect Sobol sensitivity indices for the six most important parameterizations (refer Figure 5) according to the spatially averaged surrogate model across each of the five Great Lakes (refer Figure 1).

correspond to a sustained cooling of LST (and T2m) over multiple days and could be related to different treatments of runoff and surface soil thermal conductivity in the LSMs (Niu et al., 2011).

The LST uncertainty tends to decrease with average lake latitude. This trend is bucked by Lake Erie, which has the second highest LST uncertainty on average ($1.70 \pm 0.20\,°C$). Erie is easily the shallowest lake and can therefore be most directly impacted by atmospheric radiation, as shown by an elevated total effect sensitivity index especially during the rapid spring warming phase. Similarly, Erie is less sensitive to WS than the other lakes. The effect of the PBL + SFC scheme from late July–August in Erie is far higher than seen for the other lakes. Spatial plots show that this sensitivity is strongest over the extremely shallow western Lake Erie (Figure 7). This influence could be coming from the nearby Detroit urban area and/or northern Indiana to the southwest that has a high sensitivity to PBL + SFC (Figure 8).

The uncertainty of T2m over the lakes is consistent with LST in that higher uncertainty is present during the rapid spring warming phase and correlates with higher latitude lakes (Figure 10). Indeed, Lake Superior has the highest average T2m uncertainty ($2.39 \pm 0.53\,°C$) while Lake Erie has the lowest ($1.77 \pm 0.52\,°C$). This latter fact is different from LST where Erie had the second highest average uncertainty. The higher T2m uncertainty in the rapid warming phase is associated with elevated sensitivity to the PBL + SFC scheme. Later in the season, starting from July, the importance of the LSM increases, particularly in the southern lakes. The impact of the radiation parameterization is fairly consistent across the lakes and throughout the season.

Over land, T2m is shown to warm up in June in the northern areas before maintaining a fairly constant temperature until late August. While in the southern areas, T2m maintains a more elevated temperature throughout the season. The average uncertainty of T2m in the northern land areas (NW = $1.80 \pm 0.40\,°C$, NE = $1.76 \pm 0.45\,°C$) is higher than in the southern areas (SW = $1.60 \pm 0.41\,°C$, SE = $1.37 \pm 0.34\,°C$), and also higher in the west than in the east. We mostly find that the uncertainty of T2m over land has no clear time variation like over the lakes, except for the SW area, where the uncertainty is significantly larger between early July and mid-August. The uncertainty of T2m
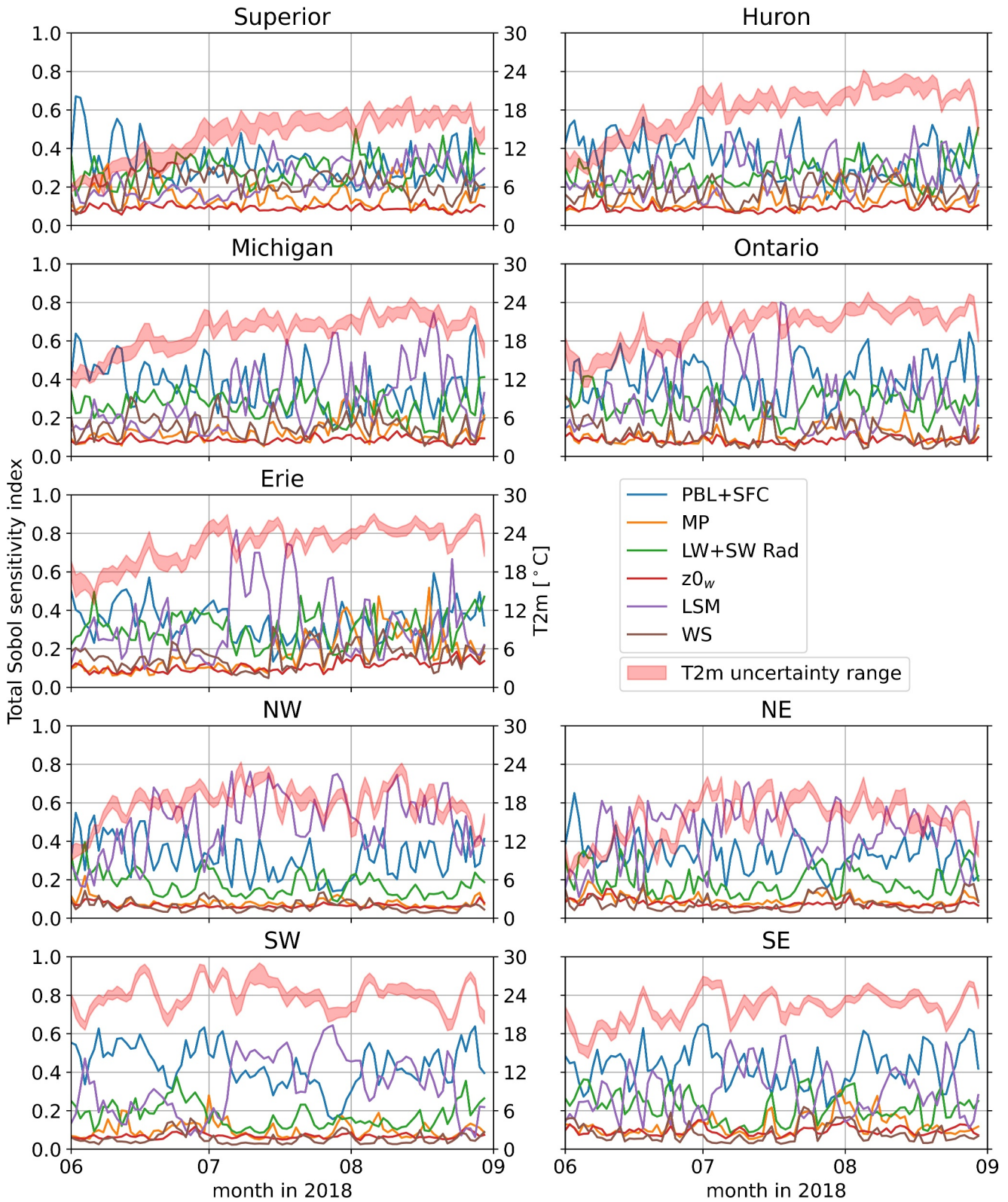
Figure 10. Temporal variation of T2m physics uncertainty (the range of the shaded area between the 5th and 95th percentiles) and total effect Sobol sensitivity indices for the six most important parameterizations (refer Figure 5) according to the spatially averaged surrogate model across each of the five Great Lakes and the four land quadrants (refer Figure 1).

over land is mostly being driven by the LSM and PBL + SFC as also shown clearly in Figure 8. The sensitivity to PBL + SFC is consistent throughout the season for the four land areas. In contrast, sensitivity to LSM is more variable with time, especially in the southern areas. The magnitude of this sensitivity increases during July and into August, the same time as the higher uncertainty in the SW. As noted previously with regards to the LST, this is correlated with a reduction in temperature over multiple days. The rate of land surface cooling computed by the LSMs can vary due to soil water content and thermal conductivity differences.

4. Discussion and Conclusions

In this study, we presented a surrogate-based approach to evaluate the physics uncertainty in a coupled lake-atmosphere-land model of the GLR. We assessed surface air and lake temperatures using a NN surrogate model that can be rapidly queried to obtain sensitivity and uncertainty information. The sensitivity information from the surrogate model agrees with physical intuition such as producing logical ranking of the relative importance of atmosphere, lake, and land contributions to LST and T2m over lake and land, and describing how atmospheric radiation and lake surface WS is the dominant control on LST. This indicates that the surrogate model applied here is robust and physically reasonable for understanding the uncertainties. Correlation of the NN model to the test set was weaker for LST than for T2m. This motivates exploring more lake and lake-atmosphere exchange parameterizations and parameters not used as a parametric prior in this study that may be having a nonlinear influence on LST.

We showed that the physics uncertainty in T2m is on average greater over the lakes than over land, although the uncertainty for T2m over land in the vicinity of the Great Lakes is highest overall. The uncertainty of both LST and T2m tends to be higher in the northern lakes and land areas than in the south. One exception to this rule is that Lake Erie LST is the second most uncertain after Lake Superior, although this higher level of uncertainty arises from different physics (PBL + SFC for Erie and WS for Superior). In regards to temporal variations, uncertainty is highest during the rapid spring warming phase for LST and T2m over the lakes, while T2m over the SW land area is higher between early July and mid-August.

The primary source of this uncertainty for surface temperatures arises from parameterizations in the atmosphere (mostly LW + SW Rad and PBL + SFC) and their interactions with lake (mostly WS) and land (LSM) physics. As expected, LW + SW Rad is important as it controls the heat available to the surface for heating, and there is relatively little spatial and temporal variation of this sensitivity. The other parameterizations are all related to surface fluxes in each of the atmosphere, lake, and land model components. Naturally, we find that WS is important for surface temperatures over the lake but unimportant over land, consistent with J. Wang et al. (2022), who found that LST affects air temperature mainly locally. This sensitivity to WS is most pronounced during the rapid spring warming phase and for the larger northern lakes. The effect of PBL + SFC is important especially to the T2m over land in the southern areas, and to LST over Lake Erie, perhaps related to the southwesterly prevailing summer winds. PBL + SFC has the highest sensitivity to T2m over urban areas which can be improved by coupling to an urban canopy model (J. Wang, Qian, et al., 2023). The effect of LSM is, in general, greatest on T2m over the northern land areas, which is more forested than the agricultural south. This makes sense based on the several canopy-based augmentations of the Noah-MP LSM. The temperatures in the southern lakes and land areas also become more sensitive to LSM from early July to mid-August, coinciding with sustained multiday temperature declines.

One hypothesis of this study was that the $z0_w$ scheme would be important for the Great Lakes as they are much shallower than the open ocean, and the depth-dependent scheme was shown to be important over the shallower continental shelves (Jiménez & Dudhia, 2018). We see here that, at least for surface temperature, the effect of $z0_w$ is relatively minor. However, this could be at least partially because the momentum flux on the atmosphere side and the lake side are calculated independently. If we instead transfer the fluxes computed in WRF (which uses $z0_w$) to FVCOM directly instead of state variables, we should see $z0_w$ assume the higher sensitivity of the WS scheme. We therefore recommend exploring this type of consistent flux-exchange coupling in future iterations of the coupled WRF-FVCOM model.

Other physics of the lake—the attenuation of the shortwave radiation, the vertical mixing scheme and the turbulent Prandtl number—were shown to be relatively unimportant to both lake and atmospheric (near-)surface temperatures. Therefore, careful treatment of atmospheric radiation and surface fluxes, as well as the numerical scheme (J. Wang, Fujisaki-Manome, et al., 2023), should be the focus of further attention in the development and

assessment of the coupled WRF-FVCOM model (Kayastha et al., 2023) or similar coupled modeling systems of the GLR (e.g., Sun et al., 2020; Xue et al., 2022). Additionally, other parameterizations should become more important for different QoIs that we intend to look at in future work, such as microphysics (MP) for precipitation (Qian et al., 2015), among others. We also aim to extend the current framework to obtain a posterior estimate of uncertainty and physics parameterizations (calibration; Lu et al., 2018; Xu et al., 2022) using observations in a subsequent study.

Data Availability Statement

The metarepository for this study is available from https://github.com/COMPASS-DOE/GreatLakes_CoupledModel_Uncertainty (W. Pringle, 2025). It contains the uncertainty analysis codes and figures as well as links to source codes and processed data.

References

Andreas, E. L., Mahrt, L., & Vickers, D. (2012). A new drag relation for aerodynamically rough flow over the ocean. *Journal of the Atmospheric Sciences*, 69(8), 2520–2537. https://doi.org/10.1175/JAS-D-11-0312.1

Austin, J. A. (2019). Observations of radiatively driven convection in a deep lake. *Limnology & Oceanography*, 64(5), 2152–2160. https://doi.org/10.1002/lno.11175

Bai, X., Wang, J., Schwab, D. J., Yang, Y., Luo, L., Leshkevich, G. A., & Liu, S. (2013). Modeling 1993-2008 climatology of seasonal general circulation and thermal structure in the Great Lakes using FVCOM. *Ocean Modelling*, 65, 40–63. https://doi.org/10.1016/j.ocemod.2013.02.003

Bellprat, O., Kotlarski, S., Lüthi, D., & Schär, C. (2012a). Exploring perturbed physics ensembles in a regional climate model. *Journal of Climate*, 25(13), 4582–4599. https://doi.org/10.1175/JCLI-D-11-00275.1

Bellprat, O., Kotlarski, S., Lüthi, D., & Schär, C. (2012b). Objective calibration of regional climate models. *Journal of Geophysical Research*, 117(23), 1–13. https://doi.org/10.1029/2012JD018262

Briley, L. J., Rood, R. B., & Notaro, M. (2021). Large lakes in climate models: A Great Lakes case study on the usability of CMIP5. *Journal of Great Lakes Research*, 47(2), 405–418. https://doi.org/10.1016/j.jglr.2021.01.010

Charusombat, U., Fujisaki-Manome, A., Gronewold, A. D., Lofgren, B. M., Anderson, E. J., Blanken, P. D., et al. (2018). Evaluating and improving modeled turbulent heat fluxes across the north american great lakes. *Hydrology and Earth System Sciences*, 22(10), 5559–5578. https://doi.org/10.5194/hess-22-5559-2018

Chen, C., Beardsley, R. C., Cowles, G., Qi, J., Lai, Z., Gao, G., et al. (2011). An unstructured-grid, finite-volume Community Ocean Model: FVCOM user manual (3rd ed.). Retrieved from http://fvcom.smast.umassd.edu/wp-content/uploads/2013/11/MITSG_12-25.pdf

Chen, C., Liu, H., & Beardsley, R. C. (2003). An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: Application to coastal ocean and estuaries. *Journal of Atmospheric and Oceanic Technology*, 20(1), 159–186. https://doi.org/10.1175/1520-0426(2003)020<0159:AUGFVT>2.0.CO;2

Craig, A., Valcke, S., & Coquart, L. (2017). Development and performance of a new version of the OASIS coupler, OASIS3-MCT-3.0. *Geoscientific Model Development*, 10(9), 3297–3308. https://doi.org/10.5194/gmd-10-3297-2017

Edson, J. B., Jampana, V., Weller, R. A., Bigorre, S. P., Plueddemann, A. J., Fairall, C. W., et al. (2013). On the exchange of momentum over the open ocean. *Journal of Physical Oceanography*, 43(8), 1589–1610. https://doi.org/10.1175/JPO-D-12-0173.1

Eidhammer, T., Gettelman, A., Thayer-Calder, K., Watson-Parris, D., Elsaesser, G., Morrison, H., et al. (2024). An extensible perturbed parameter ensemble for the Community Atmosphere Model version 6. *Geoscientific Model Development*, 17(21), 7835–7853. https://doi.org/10.5194/gmd-17-7835-2024

Fang, K.-T., Lin, D. K. J., Winker, P., & Zhang, Y. (2000). Uniform design: Theory and application. *Technometrics*, 42(3), 237–248. https://doi.org/10.1080/00401706.2000.10486045

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. https://doi.org/10.1002/qj.3803

Huber, M. B., & Zanna, L. (2017). Drivers of uncertainty in simulated ocean circulation and heat uptake. *Geophysical Research Letters*, 44(3), 1402–1413. https://doi.org/10.1002/2016GL071587

Jiménez, P. A., & Dudhia, J. (2018). On the need to modify the sea surface roughness formulation over shallow waters. *Journal of Applied Meteorology and Climatology*, 57(5), 1101–1110. https://doi.org/10.1175/JAMC-D-17-0137.1

Kayastha, M. B., Huang, C., Wang, J., Pringle, W. J., Chakraborty, T., Yang, Z., et al. (2023). Insights on simulating summer warming of the Great Lakes: Understanding the behavior of a newly developed coupled lake-atmosphere modeling system. *Journal of Advances in Modeling Earth Systems*, 15(7), e2023MS003620. https://doi.org/10.1029/2023MS003620

Large, W. G., & Pond, S. (1981). Open ocean momentum flux measurements in moderate to strong winds. *Journal of Physical Oceanography*, 11(3), 324–336. https://doi.org/10.1175/1520-0485(1981)011<0324:OOMFMI>2.0.CO;2

Lu, D., Ricciuto, D., Stoyanov, M., & Gu, L. (2018). Calibration of the E3SM land model using surrogate-based global optimization. *Journal of Advances in Modeling Earth Systems*, 10(6), 1337–1356. https://doi.org/10.1002/2017MS001134

Mellor, G. L., & Yamada, T. (1982). Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics and Space Physics*, 20(4), 851–875. https://doi.org/10.1029/RG020i004p00851

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multi-parameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research*, 116(12), 1–19. https://doi.org/10.1029/2010JD015139

Notaro, M., Zhong, Y., Xue, P., Peters-Lidard, C., Cruz, C., Kemp, E., et al. (2021). Cold season performance of the NU-WRF regional climate model in the Great Lakes region. *Journal of Hydrometeorology*, 22, 2423–2454. https://doi.org/10.1175/jhm-d-21-0025.1

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). PyTorch: An imperative style, high-performance deep learning library. In *Advances in neural information processing systems 32* (pp. 8024–8035). Curran Associates, Inc. Retrieved from http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

Paulson, C. A., & Simpson, J. J. (1977). Irradiance measurements in the Upper Ocean. *Journal of Physical Oceanography*, *7*(6), 952–956. https://doi.org/10.1175/1520-0485(1977)007<0952:IMITUO>2.0.CO;2

Pringle, W. (2025). COMPASS-DOE/GreatLakes_CoupledModel_Uncertainty (v1.0) [Workflow]. *Zenodo*. https://doi.org/10.5281/zenodo.10806950

Pringle, W. J., Burnett, Z., Sargsyan, K., Moghimi, S., & Myers, E. (2023). Efficient probabilistic prediction and uncertainty quantification of tropical cyclone-driven storm tides and inundation. *Artificial Intelligence for the Earth Systems*, *2*(2), e220040. https://doi.org/10.1175/AIES-D-22-0040.1

Qian, Y., Guo, Z., Larson, V. E., Leung, L. R., Lin, W., Ma, P. L., et al. (2024). Region and cloud regime dependence of parametric sensitivity in E3SM atmosphere model. *Climate Dynamics*, *62*(2), 1517–1533. https://doi.org/10.1007/s00382-023-06977-3

Qian, Y., Wan, H., Yang, B., Golaz, J. C., Harrop, B., Hou, Z., et al. (2018). Parametric sensitivity and uncertainty quantification in the version 1 of E3SM atmosphere model based on short perturbed parameter ensemble simulations. *Journal of Geophysical Research: Atmospheres*, *123*(23), 13046–13073. https://doi.org/10.1029/2018JD028927

Qian, Y., Yan, H., Hou, Z., Johannesson, G., Klein, S. A., Lucas, D., et al. (2015). Parametric sensitivity analysis of precipitation at global and local scales in the Community Atmosphere Model CAM5. *Journal of Advances in Modeling Earth Systems*, *7*(2), 382–411. https://doi.org/10.1002/2014MS000354

Ricciuto, D., Sargsyan, K., & Thornton, P. (2018). The impact of parametric uncertainties on biogeochemistry in the E3SM land model. *Journal of Advances in Modeling Earth Systems*, *10*(2), 297–319. https://doi.org/10.1002/2017MS000962

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

Sargsyan, K. (2017). Surrogate models for uncertainty propagation and sensitivity analysis. In R. Ghanem, D. Higdon, & H. Owhadi (Eds.), *Handbook of uncertainty quantification* (pp. 673–698). Springer. https://doi.org/10.1007/978-3-319-11259-6_22-1

Sargsyan, K., Safta, C., Najm, H. N., Debusschere, B. J., Ricciuto, D., & Thornton, P. (2014). Dimensionality reduction for complex models via Bayesian compressive sensing. *International Journal for Uncertainty Quantification*, *4*(1), 63–93. https://doi.org/10.1615/Int.J.UncertaintyQuantification.2013006821

Sharma, A., Hamlet, A. F., Fernando, H. J., Catlett, C. E., Horton, D. E., Kotamarthi, V. R., et al. (2018). The need for an integrated land-lake-atmosphere modeling system, exemplified by north America's Great Lakes region. *Earth's Future*, *6*(10), 1366–1379. https://doi.org/10.1029/2018EF000870

Skamarock, W. C., & Klemp, J. B. (2008). A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *Journal of Computational Physics*, *227*(7), 3465–3485. https://doi.org/10.1016/j.jcp.2007.01.037

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., et al. (2021). A description of the advanced research WRF model version 4.3 (Vols. TN–556+STR; Tech. Rep.). https://doi.org/10.5065/1dfh-6p97

Stips, A., Burchard, H., Bolding, K., & Eifler, W. (2002). Modelling of convective turbulence with a two-equation k-ε turbulence closure scheme. *Ocean Dynamics*, *52*(4), 153–168. https://doi.org/10.1007/s10236-002-0019-2

Sun, L., Liang, X. Z., & Xia, M. (2020). Developing the coupled CWRF-FVCOM modeling system to understand and predict atmosphere-watershed interactions over the great lakes region. *Journal of Advances in Modeling Earth Systems*, *12*(12), e2020MS002319. https://doi.org/10.1029/2020MS002319

Wang, C., Qian, Y., Duan, Q., Huang, M., Yang, Z., Berg, L. K., et al. (2021). Quantifying physical parameterization uncertainties associated with land-atmosphere interactions in the WRF model over Amazon. *Atmospheric Research*, *262*, 105761. https://doi.org/10.1016/j.atmosres.2021.105761

Wang, J., Fujisaki-Manome, A., Kessler, J., Cannon, D., & Chu, P. (2023). Inertial instability and phase error in Euler forward predictor-corrector time integration schemes: Improvement of modeling Great Lakes thermal structure and circulation using FVCOM. *Ocean Dynamics*, *73*(7), 407–429. https://doi.org/10.1007/s10236-023-01558-8

Wang, J., Qian, Y., Pringle, W. J., Chakraborty, T. C., Hetland, R., Yang, Z., & Xue, P. (2023). Contrasting effects of lake breeze and urbanization on heat stress in Chicago metropolitan area. *Urban Climate*, *48*, 101429. https://doi.org/10.1016/j.uclim.2023.101429

Wang, J., Xue, P., Pringle, W. J., Yang, Z., & Qian, Y. (2022). Impacts of lake surface temperature on the summer climate over the great lakes region. *Journal of Geophysical Research: Atmospheres*, *127*(11), e2021JD036231. https://doi.org/10.1029/2021JD036231

Xu, D., Bisht, G., Sargsyan, K., Liao, C., & Ruby Leung, L. (2022). Using a surrogate-assisted Bayesian framework to calibrate the runoff-generation scheme in the Energy Exascale Earth System Model (E3SM) v1. *Geoscientific Model Development*, *15*(12), 5021–5043. https://doi.org/10.5194/gmd-15-5021-2022

Xue, P., Pal, J. S., Ye, X., Lenters, J. D., Huang, C., & Chu, P. Y. (2017). Improving the simulation of large lakes in regional climate modeling: Two-way lake-atmosphere coupling with a 3D hydrodynamic model of the great lakes. *Journal of Climate*, *30*(5), 1605–1627. https://doi.org/10.1175/JCLI-D-16-0225.1

Xue, P., Ye, X., Pal, J. S., Chu, P. Y., Kayastha, M. B., & Huang, C. (2022). Climate projections over the Great Lakes Region: Using two-way coupling of a regional climate model with a 3-D lake model. *Geoscientific Model Development*, *15*(11), 4425–4446. https://doi.org/10.5194/gmd-15-4425-2022

Zanna, L., Brankart, J. M., Huber, M., Leroux, S., Penduff, T., & Williams, P. D. (2019). Uncertainty and scale interactions in ocean ensembles: From seasonal forecasts to multidecadal climate predictions. *Quarterly Journal of the Royal Meteorological Society*, *145*(S1), 160–175. https://doi.org/10.1002/qj.3397