

Bayesian Inference of Parameters for Numerical Storm Surge Models

Student Name: Benjamin Russell

Supervisor Name: Tobias Weinzierl

Submitted as part of the degree of MEng Computer Science to the
Board of Examiners in the Department of Computer Sciences, Durham University

Abstract —

Context/Background - When modelling storm surges using numerical methods, it is important that the model is correctly calibrated to ensure accurate results. Doing this without empirical measurements of the site being simulated is a hard and time consuming problem, and for this reason there exist many different methods for estimating correct model parameters.

Aims - This paper aims to compare a variety of different techniques for uncertainty quantification to infer correct parameters. This will allow a comparison between the time taken to estimate the parameters and their accuracy when compared with real data for storm surges. This comparison will highlight which techniques are more appropriate in real world scenarios.

Method - The methodology used in this paper was to implement three different techniques one of which is a hybrid of the other two. These were then used to infer and calibrate the model parameter values and the results were then tested in the model on a variety of real world storm surges with their accuracies and runtimes compared.

Keywords — Storm surge, uncertainty quantification, inverse problem, Bayesian inference, Monte-Carlo, Surrogate Model, kriging, polynomial chaos

I INTRODUCTION

When severe storms hit coastal regions, the high winds and atmospheric pressure can cause a sea level rise. This affect is known as a storm surge. Storm surges pose a grave danger to coastal communities both from an economic and a human perspective. From 1963 to 2012 49% of all deaths caused by hurricanes in the US were due to the subsequent storm surge (REF DEATH PAPER). It has also been estimated that if coastal flood defences are not improved in US coastal regions then by 2050 then yearly losses could exceed US\$1trillion (REF COST PAPER). To address this issue it is necessary to plan and implement better coastal flood defences. One such way to plan this new infrastructure is to use numerical modelling techniques to simulate hypothetical storm surge events, induced by hurricanes, and study the simulated sea level rise to assess coastal regions most at risk to loss of human life and loss of property. This poses a problem as storm surge models are very complex and their accuracy depends inherently on how well calibrated the input parameters are. To address this, without having accurate measurements of the underlying phenomena governing these the parameters, it is necessary to estimate them and calibrate the model accordingly. Unfortunately due to the complex nature of the numerical models this is time-prohibited and therefore more sophisticated methods of estimation and calibration are used to infer these parameters and quantify the uncertainty of them.

In this paper we focus on using a range of different techniques that fall under the umbrella of surrogate modelling. This is the use of less accurate models that are less computationally expensive to run than the full model (REF QUIRANTE). The remainder of this section will give an overview of storm surge modelling, surrogate modelling, inverse uncertainty quantification, and model calibration as well as the motivations for comparing the various techniques.

A Background

Simulating storm surges is often done using 2D shallow water equations given as (REF EQUATION PAPER).

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{U})}{\partial x} + \frac{\partial \mathbf{G}(\mathbf{U})}{\partial y} = \mathbf{S}_\tau + \mathbf{S}_b \quad (1)$$

where $\mathbf{U} = (h, hu, hv)^T$ this represents the conserved variables according to the three-dimensional conservation laws, with h being the z coordinate (flow depth) and u, v representing the velocity vector averaged by depth. With flux functions \mathbf{F}, \mathbf{G} defined as:

$$\mathbf{F} = \left(hu, hu^2 + \frac{1}{2}g_\psi h^2, huv \right) \quad (2)$$

$$\mathbf{G} = \left(hv, huv, hv^2 + \frac{1}{2}g_\psi h^2 \right) \quad (3)$$

where $g_\psi = g \cos^2 \psi$ with ψ being the cosine direction of the seabed's normal. \mathbf{S}_τ describes the frictional forces caused by the seabed and is parameterised according to Manning's friction law and an empirically determined n called a Manning roughness coefficient (REF MANNING PAPER). These equations are then solved using finite volume methods to calculate how water depth changes over time against a discretised spatial mesh.

The Manning roughness n must be determined empirically using real world measurements. This is infeasible for a general model for simulating storm surge as it would require these experimental values for n to be determined at each location you would like to simulate. For this reason we use inverse uncertainty quantification to perform parameter calibration. This is done using a Bayesian model starting with a prior distribution of roughness coefficients. This allows for the simulated results to be compared with the real-world results and update the parameters accordingly until they have been estimated to a desired level of confidence. The precise formulation of this Bayesian approach will be explained in a later section. In order to infer the final model parameters the Bayesian model must be sampled repeatedly and the sampled values for n are then used in the simulation with the new results used to further refine the Bayesian model. This sampling is usually performed using Markov chain Monte Carlo (MCMC) (REF MCMC POPULAR), this is, however, computationally expensive. The requirement of running the simulation multiple times in this approach is massively time prohibitive when using complex models such as those for storm surge simulation, which is why surrogate modelling is used as "grey-box" substitute.

Surrogate modelling is the use of a simplified model that approximates the outcome of the actual model. In the context of this paper it is a simplified model to approximate the outcome of the storm surge simulation. The use of such a surrogate model is necessary due to the long runtime of an actual simulation run (REF Tsunami paper). Designing a good surrogate model is difficult and must be done in way that approximates well otherwise the runtime benefit is not

worth it and will result in inaccurate parameter calibration, discussed previously. There are many different approaches to surrogate modelling such as kriging (REF Kriging paper) and support vector machines. In this paper we consider kriging, an alternative method, and a hybrid of the two.

B Motivation

As previously discussed the use of full numerical models for parameter calibration and inference is infeasible due to the runtime cost of repeated model runs when performing Bayesian inference. For this reason the primary way of improving parameter calibration is to create more accurate and faster surrogate models. It is also possible to perform parameter calibration using the *adjoint method* in which an analytical inverse problem is formulated, however this can be extremely complicated for large numerical models such as storm surge simulation and is therefore beyond the scope of this paper. For this reason in this paper we focus on the surrogate modelling approach to parameter calibration.

The commonly used methods for surrogate modelling draw from a wide range of mathematical disciplines such statistics and even machine learning. In this paper we focus on the statistical methods for surrogate modelling of storm surge phenomena in order to determine their efficacy for this problem.

C Objectives

In this paper we aim to answer the following research question: *How can surrogate modelling be used to calibrate storm surge model parameters?* We have divided the outcomes of the research into three categories: minimum, intermediate, and advanced.

- **Minimum:** Implement and analyse the efficacy of kriging to approximate the result of a numerical storm surge simulation. This includes implementing a storm surge simulation as well using a pre-existing software library.
- **Intermediate:** Implement and analyse the efficacy of polynomial chaos (REF PolyChaos) as a surrogate modelling technique and compare the results to those of kriging.
- **Advanced:** Implement the proposed hybrid technique *polynomial chaos kriging* proposed by (REF PC-Kriging) and compare the results to the previous methods. As well as analyse the contributions posed by the constituent methods (kriging, and polynomial chaos) to the final result of the hybrid method. Finally we will conclude whether or not the hybrids result is dominated by one of the constituent methods giving only a negligible increase in either accuracy or decrease in runtime when compared to a single method.

The layout of the rest of this paper is as follows: section II presents a survey of related and background literature on surrogate modelling for Bayesian inference in the context of parameter calibration, section III describes the proposed solution and methodology to answer the proposed research question, and section IV addresses how the proposed solution and literature review have been carried out in way that ensures any findings drawn from this research are valid and reproducible.

II RELATED WORK

In this section we provide an overview of the existing literature around Bayesian inference and surrogate modelling in the domain of numerical simulation of complex physical phenomena. There are a very broad range of approaches to solving this problem therefore since our paper focuses on the statistical approaches rather than analytical approaches the literature review will reflect this.

A *Gaussian Process Emulation*

One of the more common methods of surrogate modelling is the aforementioned kriging. This method requires the construction of a Gaussian process emulator (GPE). A GPE operates under the assumption that the outputs of the model follows some multivariate Gaussian distribution. A set of known inputs and outputs can be used to train the GPE such that a set of basis functions are produced that can be easily evaluated to get an approximation of the model output (REF GPE BOOK). It has been shown that in complex models such as those for flood inundation, the empirical variation between predicted model outputs in the GPE is greater than the theoretical difference (REF Flood Paper) meaning that unexplained inaccuracies can be present in a GPE. However (REF Flood Paper) did also show that using a GPE allows for Bayesian inference to be used for model calibration in highly complex hydrodynamic models. The variation in GPE outputs when used for Bayesian calibration was demonstrated by (REF CFD Paper) to be of comparable accuracy to performing full model runs with a significant improvement in runtime, although this was done in the domain of particulate flow simulation which is still a complex numerical problem it may not be directly comparable to storm surge. (REF VALID PAPER) presented a series of diagnostics that can be applied to a GPE in order to determine its effectiveness at approximating model outputs. One of these metrics is the Mahalanobis Distance which allows measuring the number of standard deviations a point is away from the mean in a multi-dimensional case (REF Mahalanobis). These metrics are significant as they compare the GPE output variation with that of the actual model rather than reality, this is important for our application as our GPE should approximate the model specifically rather than the underlying physical processes. As an individual method GPE does have the benefit that it gives the sensitivity and uncertainty caused by the emulation process when compared to the full model (REF GPE sens), this is a useful property as it allows for a detailed sensitivity analysis of the emulator compared to the original model to be carried out.

B *Polynomial Chaos Expansion*

Polynomial chaos expansion is another popular method for surrogate modelling. In this method the model is represented as a quantity of interest that can be approximated by an infinite of polynomial coefficients are elements of an orthogonal probability space (REF Xiu Paper). (REF Xiu2 Paper) shows that it is necessary to truncate this infinite series and discusses proofs of the convergence of the series for different selections of polynomials. (REF tsunami paper) used a polynomial chaos surrogate (PC surrogate) to simulate sea elevation rises during tsunamis, constructing a fourth order PC surrogate that only required 125 full model runs to create, and with the maximum error (when compared to the full model) of only 1%. The computational savings from a PC surrogate can be huge which is highly beneficial when performing MCMC

(REF PC Cost Paper) this paper demonstrated up to a 30,000 times reduction in computational cost when using a PC surrogate to perform Bayesian inference for parameter calibration. The paper did find that as the truncation order of the PC series increases then the speedup gets worse, however as (REF tsunami) showed for sea elevation an order 4 truncation is sufficient. A method for deriving coefficients for PC surrogates called least angle regression (LAR) (REF LAR Paper) has been shown to address the problem of high truncation orders, this method allows for a PC surrogate to be constructed to a high order and then truncated by only selecting the coefficients that heavily contribute to the output, this is called a sparse PC approximation. A further study into using PC surrogates for calibration of Manning coefficients showed that the RMSE of the surrogate was close to zero at all points except for a few outliers and even then it was less than one standard deviation of error (REF PC RMSE PAPER). As a surrogate method polynomial chaos has some advantages (REF PC adv paper) found some advantages for it: the number of model runs for training is relatively low, uncertainty in the model parameters can easily be incorporated into the model. (REF PC adv paper) also showed that if the parameter uncertainty is assumed to be distributed as a standard normal then there exists a very efficient training procedure for the surrogate.

C Hybrid Models

Due to the success of the previous two methods recent developments have been made in the literature to try and create hybrid surrogate models to capture the advantages of the different techniques. The obvious hybrid is that of polynomial chaos and kriging, known as PC-Kriging (REF PC-Kriging paper). PC-Kriging uses a polynomial chaos expansion to approximate the global behaviour of the model and kriging (GPE) is used to simulate the local variations that can occur. The results obtained in the paper suggest that PC-Kriging is at least as accurate as the constituent surrogates. A similar approach was also presented by (REF PC-kriging 2), their approach was to use LAR to select the best polynomials for a PC surrogate and then use these polynomials during the regression process in the GPE. This paper supported the findings that the hybrid method outperforms the constituent methods, however it was noted that the improvements varied widely depending on what application it was being applied to, this is motivation to apply this technique to our storm surge model. An extension was proposed that uses an additional Bayesian framework to create a sparse representation of model outputs to further improve accuracy (REF PC-K ext). The results show that the accuracy is vastly improved when performing fewer evaluations of the original model. However as previously stated for tsunami sea level models only 125 model evaluations were required (REF tsunami) to get high accuracy and at this point the Bayesian framework extension from (REF PC-K ext) does not provide any greater accuracy than plain PC-Kriging. PC-Kriging and its extensions have not been applied to storm surge modelling in the literature with most approaches focusing on the individual methods and seeking improvements from the formulation of the original model or Bayesian inference procedure instead, for this reason we have decided to apply PC-Kriging to parameter calibration in this paper. Other hybrid techniques exist in the literature for surrogate modelling, however these usually make assumptions and are specific to the physical process being simulated. Due to the lack of these specialised methods for storm surge modelling we have not addressed these.

III SOLUTION

This section provides a description of the proposed solution and methodology in this project. This includes the formulation of the surrogate models and the proposed experiments and analysis that will be carried out on them. First a brief description of the datasets, programming languages, and other software packages to be used is presented.

We chose to use the three surrogate modelling methods outlined below as there has been no direct comparison in the context of complex hydrodynamic simulations such as hurricane storm surge, and there has been no application of PC-Kriging to such simulations either.

A *Software Tools and Datasets*

The full storm surge model will be created using the GeoClaw package (REF Geoclaw)(REF Leveque)(REF ClawPack). This software was chosen due to its adaptive mesh refinement algorithms that allow for more accurate and faster forward model runs. To formulate the GPE we will use the GPy Python package, this package allows easy construction of a GPE and application of kriging. The PC surrogate and PC-Kriging hybrid model will be implemented in C due to its fast execution speed. Data visualisation and plotting will be performed using the Matplotlib Python package due to its flexible nature.

For this project we will evaluate the model and surrogates on data obtained on Hurricane Ike (2008). The storm data will be sourced from the Automated tropical Cyclone Forecasting System (ATCF) and the regions bathymetry and measured sea elevation will be sourced from the National Oceanic and Atmospheric Administration (NOAA).

B *Parameter Inference Procedure*

The Bayesian inference procedure used is based on (REF tsunami)(REF Bayes book). This involved first formulating the full model as the product of conditional probability distributions and a likelihood function. We can then describe the model parameters in terms of probability distributions, the priors are constructed using whatever information we have available at the start. Then using model or surrogate runs we can compare results with real data and employ Bayes' rule to calculate the posterior distributions until we have the parameters to a desired degree of accuracy. The full procedure is as follows:

η is the vector of real sea level rises for each sensor from the NOAA data. Let n be the vector of model parameters (n_1, n_2, n_3) , the three friction coefficients at varying water depths. Finally let M represent the forward model evaluation as a function such that $\eta \approx M(n)$. Bayes' rule tells us that:

$$P(n|\eta) \propto L(\eta|n)q(n) \quad (4)$$

where L is the likelihood function of elevations η given n and $q(n)$ is the prior distribution of n . L can be formulated using the discrepancy between the real sea level rise and rise predicted by the model i.e. $\epsilon = \eta - M$. All errors at each sea level sensor are assumed to be i.i.d. with their own probability density function P_ϵ . Therefore L can be described as the product of probability for each sensor:

$$L(\eta|n) = \prod_i P_\epsilon(\eta_i - M_i(n)) \quad (5)$$

As in (REF tsunami) we assume all ϵ_i are normally distributed such that $\epsilon \sim N(0, \sigma^2)$. Therefore:

$$L(\eta|n) = \frac{1}{\sqrt{2\pi\sigma^2}} \prod_i \exp\left(-\frac{(\eta_i - M_i(n))^2}{2\sigma^2}\right) \quad (6)$$

It is unreasonable to assume that the variance of errors will be the same at all the sensors therefore Bayes' rule must be applied separately at each sensor with its own σ^2 . The variance is also unknown and therefore will be used as a model parameter like n . This gives a full formulation of Bayes' rule as:

$$P(n_1, n_2, n_3, \sigma_i^2 | \eta_i) \propto \frac{1}{\sqrt{2\pi\sigma^2}} \prod_i \exp\left(-\frac{(\eta_i - M_i(n))^2}{2\sigma^2}\right) q(\sigma^2)q(n_1)q(n_2)q(n_3) \quad (7)$$

For prior distributions all n_i are assumed to be uniformly distributed and the variance is assumed to be a Jeffrey's prior (REF Jeff Paper) as we do not have any useful information about it (REF Sraj paper). In order to sample the posterior $P(n_1, n_2, n_3, \sigma_i^2 | \eta_i)$ we use the MCMC algorithm mentioned in section I.A

C Surrogate Models

C.1 Gaussian Process Emulator

Our GPE is based on the approach taken by (REF PHD). A more thorough description is given in (REF GPE BOOK). Let \mathbf{X} be a set of training model inputs and \mathbf{f} be the set of true model outputs for \mathbf{X} with \mathbf{f}_* be the predicted model outputs from the GPE then:

$$\mathbf{f}_* | \mathbf{f} \sim \mathcal{N}(\boldsymbol{\mu}_* + \Sigma_*^T \Sigma^{-1}(\mathbf{f} - \boldsymbol{\mu}), \Sigma_{**} - \Sigma_*^T \Sigma^{-1} \Sigma_*) \quad (8)$$

where $\boldsymbol{\mu}_*$ is the mean of predicted outputs, Σ_{**} is the covariance matrix for the predicted outputs, and Σ_* is the covariance matrix between the training and predicted outputs. $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_*$ are derived from a matrix of basis functions from \mathbf{X} and a vector of regression coefficients $\boldsymbol{\beta}$ such that:

$$\boldsymbol{\mu} = g(\mathbf{X})^T \cdot \boldsymbol{\beta} \quad (9)$$

$$\boldsymbol{\mu}_* = g(\mathbf{X}_*)^T \cdot \boldsymbol{\beta} \quad (10)$$

The covariance matrices are constructed using the covariance function:

$$k(x, x') = \sigma^2 \exp\left(-\frac{(x - x')^2}{2l^2}\right) \quad (11)$$

where l is correlation length of the outputs. $\boldsymbol{\beta}$, σ^2 , l are hyperparameters in the model represented as $\boldsymbol{\theta}$. These hyperparameters are calculated by maximising the log likelihood given as:

$$\log L(\mathbf{f} | \mathbf{X}, \boldsymbol{\theta}) = -\frac{1}{2} \log |\Sigma| - \frac{1}{2} (\mathbf{f} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{f} - \boldsymbol{\mu}) - \frac{N}{2} \log 2\pi \quad (12)$$

where N is the number of basis functions. After calculating these terms they can be used in Eq. (8) to cheaply predict model outputs.

C.2 PC-Surrogate

Our PC-Surrogate is based on the formulation presented in (REF tsunami). We define U to be a function on space \mathbf{x} , time t , and ζ which is the vector of random variables representing our three friction coefficients, once again normally distributed. Therefore our PC-Surrogate is defined as:

$$U(\mathbf{x}, t, \zeta) = \sum_{k=0}^P U_k(\mathbf{x}, t) \Psi_k(\zeta) \quad (13)$$

Where Ψ_k represents the elements of the orthogonal basis of some probability space on ζ and U_k are polynomial coefficients and P is the truncated length of the series. In order to calculate the polynomial coefficients we use a non-intrusive approach called non-intrusive spectral projection (NISP) as it allows us to re-use the GeoClaw model without modification simplifying our implementation. A more complete description is given in (REF tsunami), we define U_k as:

$$U_k(\mathbf{x}, t) = \sum_q \mathcal{P}_{kq} U(\mathbf{x}, t, \zeta_q) \quad (14)$$

$$\mathcal{P}_{kq} = \frac{\Psi_k(\zeta_q) \omega_q}{\langle \Psi_k, \Psi_k \rangle} \quad (15)$$

Where $\langle \Psi_i, \Psi_j \rangle$ is the inner product of the probability space, $U(\mathbf{x}, t, \zeta_q)$ is the result of a full model run with friction coefficients set at ζ_q . ζ_q and ω_q are carefully chosen to be the points and weights of a numerical quadrature (numerical integration).

C.3 PC-Kriging hybrid

Our PC-Kriging model is based on the description given in (REF PC-Kriging paper) called Sequential PC-Kriging. This formulation is very similar to our GPE apart from instead of using Eq. (9) we define μ as:

$$\mu = \sum_{k=0}^P U_k(\mathbf{x}, t) \Psi_k(\zeta) \quad (16)$$

which is a representation of the mean using polynomial chaos. The main difference being instead of using NISP we use the LARS algorithm (REF LARS) to generate the polynomial coefficients. Eq. (9) is then plugged into our GPE model which is calibrated by once again maximising Eq. (8). This should allow for both the global behaviour of the full model to be approximated and for the local variations to be approximated too, rather than one or the other as with the two previous models.

D Proposed Analysis

To compare our three surrogate models we will be assessing them on two measures, the accuracy of the inferred friction coefficients obtained when used in the full model, and the runtime taken to perform the inference against the number of runs of the full model when training the surrogates. We will also present the inferred distribution of friction coefficients for each surrogate too.

D.1 Accuracy

The accuracy metric we have selected is the normalised root mean square error E defined as:

$$E = \sqrt{\frac{\sum_i \int (M_i - \eta_i)^2 dt}{\sum_i \int \eta_i^2 dt}} \quad (17)$$

where M_i is the sea level rise from the model and η_i is the observed rise at each NOAA sensor. The error is integrated over a 24 hour time period provided in the NOAA data.

IV VALIDITY