Using multilevel Monte Carlo for uncertainty quantification in flood forecasting

Alastair Gregory

Imperial College London

Joint work with Colin Cotter and the Firedrake team at Imperial College London

a.gregory14@imperial.ac.uk

Maths Foresees Annual Conference, Edinburgh

07/09/2016

Overview

1. Motivation
2. Numerical Simulation
3. Multilevel Monte Carlo
4. Flood Demonstration
5. Conclusion

1 Motivation

2 Numerical Simulation

3 Multilevel Monte Carlo

4 Flood Demonstration

5 Conclusion

Uncertainty Quantification

- One can quantify the uncertainty in solutions to dyanamical systems that have random properties or components, such as initial conditions or dynamics, by ensemble methods.
- This is especially important in nonlinear problems, where a small uncertainty in the initial state of a system can lead to widespread and unknown probabilistic forecast distributions in the future.

Uncertainty Quantification for the Shallow Water Equations

- The shallow water equations, in 2 dimensions, with depth h, momentum (m_u, m_v), gravity g, topography b and source S are:

$$\begin{bmatrix} \frac{\partial h}{\partial t} \\ \frac{\partial m_u}{\partial t} \\ \frac{\partial m_v}{\partial t} \end{bmatrix} + \begin{bmatrix} \frac{\partial m_u}{\partial x} \\ \frac{\partial \left(\frac{m_u^2}{h} + \frac{g}{2} h^2 \right)}{\partial x} \\ \frac{\partial \left(\frac{m_u m_v}{h} \right)}{\partial x} \end{bmatrix} + \begin{bmatrix} \frac{\partial m_v}{\partial y} \\ \frac{\partial \left(\frac{m_u m_v}{h} \right)}{\partial y} \\ \frac{\partial \left(\frac{m_v^2}{h} + \frac{g}{2} h^2 \right)}{\partial y} \end{bmatrix} = \begin{bmatrix} S \\ -gh\frac{\partial b}{\partial x} \\ -gh\frac{\partial b}{\partial y} \end{bmatrix}$$

- Uncertainty quantification can be used in flood forecasting by creating an ensemble of realisations of shallow water model simulations with some component of randomness.

Uncertainty Quantification for the Shallow Water Equations

- Random components: Initial state of the system, the topography, the source (such as rainfall) and absorption properties of the land bed.
- Could reduce computational cost: *Extreme value theory* or *importance sampling* when timescales are long.
- Importance sampling: Draw from a different distribution to the actual forecast prior, weighting them in importance with respect to this distribution. E.g. only sample seasons / months of the year when rainfall is expected to be at its highest, to forecast the likelihood of a large flood.

Monte Carlo Estimation

- Monte Carlo: Estimate quantities of interest from an ensemble of realisations of the (random) shallow water equations. E.g. in flood forecasting these quantities could be the mean amount of flooding in a certain region of the modelled domain.
  If $W = \left(h, m_u, m_v\right)$ then a Monte Carlo estimator using an ensemble of N realisations (W_i) for the expectation of a functional g of W ($\mathbb{E}[g(W)]$) is:

$$\widehat{g(W)} = \frac{1}{N}\sum_{i=1}^{N} g(W_i)$$

- This estimator has a very slow rate of convergence, $N^{-\frac{1}{2}}$, making the application of it for uncertainty quantification of the shallow water equations infeasbile, especially for high resolution.

1 Motivation

2 Numerical Simulation

3 Multilevel Monte Carlo

4 Flood Demonstration

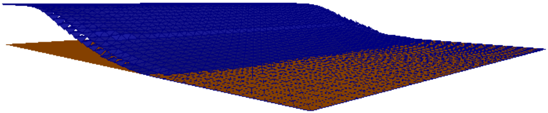
5 Conclusion

Flooddrake

- Alongside Firedrake, we have developed a discontinuous Galerkin Finite Element model of the shallow water equations (based on Ern [EPD08]) [https://github.com/firedrakeproject/flooddrake].
- Due to the ease and user-friendliness of Firedrake, it is very simple to use in Python.

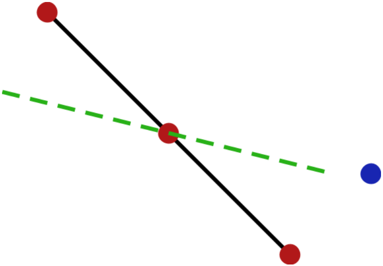
Flooddrake

- The model has wetting and drying capabilities, as well as being well balanced.
- Mean-preserving slope limiting is done by the Kuzmin [Kuz13] methodology.
- Adaptive time-stepping dependent on Courant conditions.
- Two key parameters: ϵ, the negligible depth of water that anything below this can be set to 0 and u_{bnd}, the upper bound for the norm of velocity to avoid spurious shocks.



Mean-Preserving Slope Limiter

From Kuzmin [Kuz13], vertex-based slope limiting can be used to prevent shocks.



Here, the vertices (red spots) of a cell in a DG1 function (black line) under vertex-based slope limiting (green line) get shifted, whilst preserving the mean of the cell, to keep each vertex within limits defined by the neighbouring cell averages (blue spot).

Animation of Flooddrake Simulation

- This is an animation of the wetting and drying capabilities of the Flooddrake numerical method, showing the moment a dam breaks in a reflective-wall square model lake flooding dry land, before then receding.

1 Motivation

2 Numerical Simulation

3 Multilevel Monte Carlo

4 Flood Demonstration

5 Conclusion

Multilevel Monte Carlo Estimation

- Mike Giles [Gil08] introduced the novel multilevel Monte Carlo as a tool to reduce the computational cost of a Monte Carlo estimate, with a certain root-mean-square-error (RMSE).
- Idea is a hierarchy of sub-ensembles, a combination of larger lower resolution ones and smaller higher resolution ones.
- Plenty of literature published recently, extending this method on to probability distributions / CDFs, data assimilation / filtering estimators [GCR16] and MCMC estimators.

Multilevel Monte Carlo Estimation

- This feasibility study applied 'vanilla' multilevel Monte Carlo to estimate a mean quantity of the shallow water simulations to get a feel of the speed-up offered by this framework over the infeasible Monte Carlo method.
- A Python package that uses Firedrake has been developed that can construct multilevel Monte Carlo estimators of fields / quantities of interests from random high dimensional systems [https://github.com/firedrakeproject/firedrake-mlmc].

Multilevel Monte Carlo Estimation

- Given that $\widehat{g(W_l)}$ corresponds to a Monte Carlo estimator of the shallow water equations using the flooddrake numerical model with resolution $\Delta x_l = \Delta x_0 2^{-l}$, $l \in [0, L]$, the 'vanilla' multilevel Monte Carlo estimator of $\mathbb{E}[g(W_L)]$ is given by,

$$\widehat{g(W_L)}^{ML} = \widehat{g(W_0)} + \sum_{l=1}^{L} \left(\widehat{g(W_l)} - \widehat{g(W_{l-1})}\right).$$

- This returns an estimator, that is unbiased from it's standard Monte Carlo counterpart, of the expectation of the functional g of W_L, on the finest level.

Multilevel Monte Carlo Estimation

- Using algorithm in [Gil08], one can find the optimal sample sizes of each estimator (N_l) and the finest level L on-the-go. This is optimal in terms of minimising computational cost whilst bounding RMSE.
- In particular, the sample sizes should be decreasing as the sub-ensembles increase in resolution (leading to a combination of large but low accuracy and small but high accuracy ensembles).

Multilevel Monte Carlo Estimation

- A crucial aspect of the multilevel Monte Carlo framework is being able to correlate fine and coarse 'pairs' of ensemble members in the sub-ensembles. This is sometimes quite trivial. E.g. when randomising initial conditions, use the same initial conditions for coarse and fine 'pairs'.
- In otherwords, the ensemble sample mean and variance (V_l) of,

$$\big(g(W_l) - g(W_{l-1})\big),$$

should decay to 0 as $l \to \infty$.

Computational Cost Reductions

- Require $V_l = O(\Delta x_l^{\beta})$, $\beta > 0$.
- Cost of one sample on level l is $O(\Delta x_l^{-\gamma})$.
- Optimal computational cost reductions if $\beta > \gamma$.

1 Motivation

2 Numerical Simulation

3 Multilevel Monte Carlo

4 Flood Demonstration

5 Conclusion

Problem Set-Up

- We choose our domain to be a hierarchy of regular square triangular meshes,

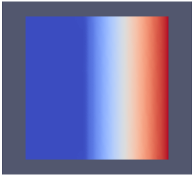
$$x \in [x_0, x_{n_l}], \quad y \in [y_0, y_{n_l}],$$

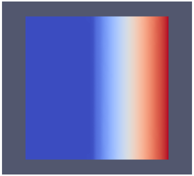
with $n_l = 5 \times 2^l$, $l \in [0, 4]$, $x_0 = y_0 = 0$ and $x_{n_l} = y_{n_l} = 50$.
- The initial conditions are $h(x, y) = m_u(x, y) = m_v(x, y) = 0$. The source (rainfall) is constant in time and uniform in space, $S(x, y) = 0.0075$.
- The randomness in the problem is introduced via the topography:

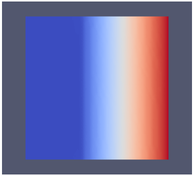
$$b(x, y) = \begin{cases} \frac{37.5}{(25-\phi)}\left(x - (25 + \phi)\right), & x > (25 + \phi), \\ 0, & x \leq (25 + \phi), \end{cases}$$

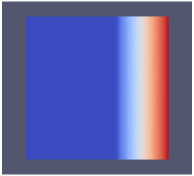
where $\phi \sim N(0, 6)$. We run our simulations until $T = 20$ and take gravity as $g = 9.8ms^{-2}$.

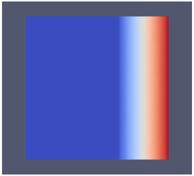
8 Random Flood Realisations

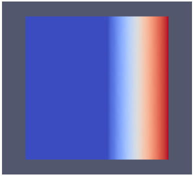

(a)

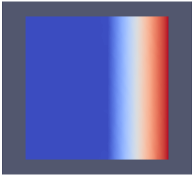

(b)

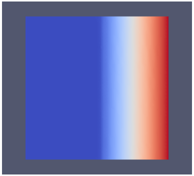

(c)


(d)


(e)

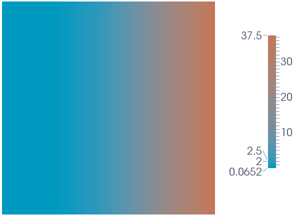

(f)


(g)


(h)

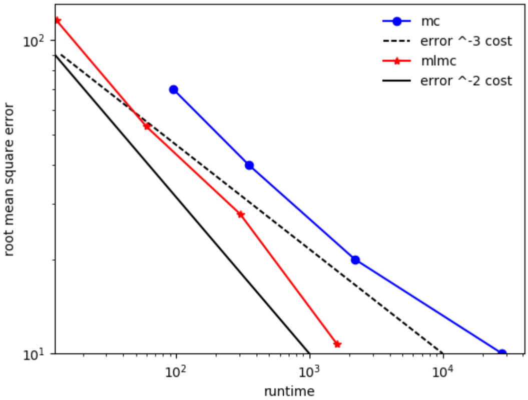
Multilevel Monte Carlo Simulation

This is the multilevel Monte Carlo estimator of the expected free surface height (the depth component of W_L plus the bedding height)



- Using the multilevel Monte Carlo Python package coupled with the Flooddrake numerical model, it would be very similar to compute this multilevel estimator on-the-go whilst propagating the ensemble members forward in time.

Speed Up Vs Monte Carlo



1 Motivation

2 Numerical Simulation

3 Multilevel Monte Carlo

4 Flood Demonstration

5 Conclusion

Conclusion and Discussion

- This feasibility demonstrates how the very simple 'vanilla' multilevel Monte Carlo can be used to significantly reduce the computational effort in quantifying the uncertainty in shallow water flood models.
- The standard Monte Carlo method has a very slow rate of convergence and is infeasible to carry out this sort of quantification, especially for high resolution forecasts.
- The multilevel Monte Carlo framework can be extended to a variation of statistical and forecasting methods, such as data assimilation and importance sampling, that could make it a very appealing technique for flood ensemble forecasting.

References

A. Ern, S. Piperno, and K. Djadel, *A well-balanced runge–kutta discontinuous galerkin method for the shallow-water equations with flooding and drying*, International Journal for Numerical Methods in Fluids **58** (2008), no. 1, 1–25.

A. Gregory, C. J. Cotter, and S. Reich, *Multilevel Ensemble Transform Particle Filtering*, SIAM Journal on Scientific Computing **38** (2016), no. 3, A1317–A1338.

M. B. Giles, *Multilevel Monte Carlo Path Simulation*, Oper. Res. **56** (2008), no. 3, 607–617.

D. Kuzmin, *Slope limiting for discontinuous galerkin approximations with a possibly non-orthogonal taylor basis*, International Journal for Numerical Methods in Fluids **71** (2013), no. 9, 1178–1190.

Thanks for listening

Contact:

a.gregory14@imperial.ac.uk
colin.cotter@imperial.ac.uk