A statistical finite element approach to nonlinear PDEs

Connor Duffin[1] Edward Cripps[1] Thomas Stemler[1,2] Mark Girolami[3,4]

[1]Maths & Stats, UWA
[2]Complex Systems Group, UWA
[3]Engineering, University of Cambridge
[4]Alan Turing Institute

February 25, 2021

Motivation

- FEM is the most popular way of numerically solving PDEs, across science and engineering.
- However, no statistically coherent way of getting data into FEM simulations.
- How to combine our knowledge (FEM model/physics) with observations?
- Moreover, how to reconcile a mismatched model with observations, that isn't inversion.

Talk outline

1. FEM: introducing some notation.
2. The linear statFEM construction, getting data into an FEM simulation, Poisson's equation in 1D.
3. Extending to the nonlinear, time-dependent case, building on nonlinear state-space modelling/DA.
4. Case study: solitons, the Korteweg-de Vries equation, and applying the filtering methods.

Section 1

The statistical finite element method (statFEM)

FEM

- Most widely-used method of discretizing PDEs. Classic linear, steady-state example (Poisson equation in $\mathbb{R}^d$):

$$\begin{cases} -\Delta u = f, \text{ in } \Omega, \; u = g \text{ on } \partial\Omega, \\ u := u(x), \; x \in \Omega \subset \mathbb{R}^d, \\ u : \Omega \to \mathbb{R}, \; f : \Omega \to \mathbb{R}. \end{cases}$$

- Multiply the above by testing functions $v \in \mathcal{V}(\Omega)$ and integrate to give the weak form (after getting rid of BCs):

$$\int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x = \int_\Omega f \cdot v \, \mathrm{d}x.$$

  Which we can write as the shorthand $\mathcal{A}(u, v) = \langle f, v \rangle$.

FEM

Now project to finite dimensional space $\mathcal{V}_h \subset \mathcal{V}$, which is defined as $\mathcal{V}_h = \mathrm{span}\{\phi_i\}_{i=1}^M$. Note finite dimensional. Choose basis functions ϕ_i to be interpolating polynomials of the form $u_h(x) = \sum_{i=1}^M u_i \phi_i(x)$. Then the weak form gives the system

$$\sum_{i=1}^{M} u_i \, \mathcal{A}(\phi_i, \phi_j) = \langle f, \phi_j \rangle \quad j = 1, \ldots, M.$$

For judicious choice of ϕ_i we get a sparse linear system which can be solved by the various methods available.

A key advantage of FEM is that the choice of basis functions can be defined over a *mesh* — a discretization of Ω — that can be very complex (e.g. heatsinks, turbines, engine cylinder blocks).

The statFEM construction

Case study: 1D Poisson equation with constant RHS, Dirichlet boundaries:[1]

$$\begin{cases} -\partial_x^2 u = 1 + \xi \\ u := u(x), \ x \in [0,1] \\ u(0) = u(1) = 0 \end{cases}$$

We give ξ a Gaussian process prior: $\xi \sim \mathcal{GP}(0, K)$, K problem specific and can be chosen as e.g. square-exponential.
This implicitly defines a probability measure over the solution space in which we are looking at (some function space, e.g. $H_0^1(\Omega)$, under regularity conditions on ξ, e.g. $\xi \in L^2(\Omega)$).

[1] Mark Girolami et al. "The Statistical Finite Element Method (statFEM) for Coherent Synthesis of Observation Data and Model Predictions". en. In: *Computer Methods in Applied Mechanics and Engineering* 375 (Mar. 2021), p. 113533. ISSN: 0045-7825. DOI: 10.1016/j.cma.2020.113533.

Deriving the prior

As with the deterministic case multiply by testing functions ϕ_j, for $j = 1, \dots, M$, and expand to give $u_h(x) = \sum_{i=1}^{M} u_i \phi_i(x)$:

$$\sum_{i=1}^{M} u_i \, \mathcal{A}(\phi_i, \phi_j) = \langle f, \phi_j \rangle + \langle \xi, \phi_j \rangle \quad j = 1, \dots, M.$$

From which we can write out the associated finite-dimensional measure:

$$\mathbf{u} \sim \mathcal{N}\left(\mathbf{A}^{-1}\mathbf{f}, \mathbf{A}^{-1}\mathbf{G}_\theta \mathbf{A}^{-\top}\right)$$

where

- $\mathbf{u} = (u_1, \dots, u_M)^\top \in \mathbb{R}^M$ vector (FEM coefs).
- $\mathbf{f} = (\langle f, \phi_1 \rangle, \dots, \langle f, \phi_m \rangle)^\top \in \mathbb{R}^M$ vector.
- $\mathbf{A}_{ij} = \mathcal{A}(\phi_i, \phi_j) \in \mathbb{R}^{M \times M}$ matrix.
- $\mathbf{G}_{\theta,ij} = \langle \phi_i, K\phi_j \rangle \in \mathbb{R}^{M \times M}$ matrix.

Note that if $\mathbf{G}_\theta \to \mathbf{0}$ then the Gaussian collapses to a Dirac at the FEM solution.

Prior measure

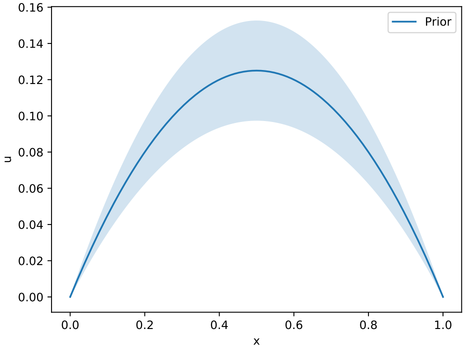


Figure: StatFEM prior: 1D Poisson example as defined by the previous slide. Mean shown as blue line, 95% probability intervals shown as blue ribbon.

- GP introduces natural uncertainty inside of the PDE solution.
- This uncertainty can be characterized by the *a priori* chosen GP parameters.
- These parameters can be chosen to represent physically relevant length/space-scales.
- At one level above (choice of covariance function): smoothness.

Combining with data

Now, suppose we have observed some (noisy, possibly mismatched) data y:

$$\mathbf{y} = \mathbf{H}\mathbf{u} + \boldsymbol{\delta} + \boldsymbol{\varepsilon},$$

- $\mathbf{y} \in \mathbb{R}^N$: observations.
- $\mathbf{u} \in \mathbb{R}^M$: statFEM model, $\mathbf{u} \sim \mathcal{N}(\mathbf{m}_u, \mathbf{C}_u)$.
- $\mathbf{H} : \mathbb{R}^M \to \mathbb{R}^N$: observation operator.
- $\boldsymbol{\delta} \sim \mathcal{GP}(0, \mathbf{K}_\delta)$: systematic model bias/discrepancy/mismatch.
- $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \sigma_y^2 \mathbf{I})$: observation noise.
- Assumed $\mathbf{u} \perp \boldsymbol{\delta} \perp \boldsymbol{\varepsilon}$.

Some basic manipulations

This gives the likelihood $p(\mathbf{y} \mid \mathbf{u})$:

$$p(\mathbf{y} \mid \mathbf{u}) = \mathcal{N}(\mathbf{Hu}, \mathbf{K}_\delta + \sigma^2 \mathbf{I}).$$

Which can be combined with the prior $p(\mathbf{u})$ to give the posterior $p(\mathbf{u} \mid \mathbf{y})$:

$$p(\mathbf{u} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{u}) p(\mathbf{u}).$$

And the marginal likelihood $p(\mathbf{y})$ (marginalize over $\mathbf{u}$):

$$p(\mathbf{y}) = \mathcal{N}(\mathbf{Hm}_u, \mathbf{HC}_u\mathbf{H}^\top + \mathbf{K}_\delta + \sigma^2 \mathbf{I}).$$

We optimize the marginal likelihood to learn the mismatch hyperparameters. From here on in I'll implicitly condition on these.

Posterior

Solve Poisson to give $\mathbf{u}_0$. Then, generate some fake data by scaling $\mathbf{u}_0$, simulating $\boldsymbol{\delta}$ and ε: $\mathbf{y} = 0.8\mathbf{H}\mathbf{u}_0 + \boldsymbol{\delta} + \varepsilon$, with parameters $\sigma_\delta = 0.1$, $\ell_\delta = 1$, and $\sigma_y = 0.05$.
Put this data into our statFEM model to give the posterior, $p(\mathbf{u} \mid \mathbf{y})$.

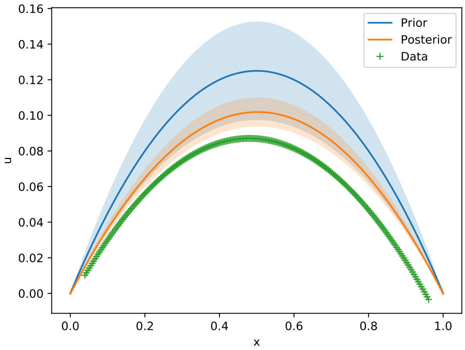


Figure: StatFEM posterior: Means shown as lines, 95% CI shown as transluscent ribbons.

- The posterior over the model, $p(\mathbf{u} \mid \mathbf{y})$ is the main object of interest.
- Posterior is a compromise between the model and the data: retains known physical properties of the system at hand (e.g. BCs), with UQ.

Section 2

Nonlinear problems

Nonlinear PDEs

- Burgers equation: $u_t + uu_x - \delta u_{xx} = 0$ (fluid mechanics, traffic flow),
- KdV: $u_t + uu_x + u_{xxx} = 0$ (internal waves, plasma physics, integrable systems),
- Nonlinear Schrödinger: $iu_t + \frac{1}{2}u_{xx} - \kappa\,|u|^2\,u = 0$ (optics, acoustics, integrable systems),
- Navier-Stokes: $u_t + (u \cdot \nabla)u - \nu\nabla^2 u = -\nabla w + g$ (fluid dynamics).

When extending to nonlinear PDEs we then have the following problems:

1. Not Gaussian anymore (discretized PDE operator no longer linear).
2. In general not available in closed form.
3. Most nonlinear systems are also time-dependent - we need to deal with this too.

So, need to build a general method for nonlinear/time-dependent PDEs that combines our model with the data.

The nonlinear statFEM construction

Now:

$$\begin{cases} \partial_t u + \mathcal{L}u + \mathcal{F}(u) + \dot{\xi} = 0, \\ u := u(x,t), \quad \dot{\xi} := \dot{\xi}(x,t), \\ x \in \Omega \subset \mathbb{R}^d, \quad t \in [0,T], \\ u, \dot{\xi} : \Omega \times [0,T] \to \mathbb{R}. \end{cases}$$

- $\mathcal{L}$ is a linear differential operator.
- $\mathcal{F}$ is nonlinear (possibly differential) operator.
- $\dot{\xi}$ is delta-correlated in time with spatial correlations from K $\dot{\xi}(x,t) \sim \mathcal{N}(0, \delta(t-t')K(x,x'))$.

The nonlinear statFEM construction

Some notation:

- Let $u_h \approx \sum_{i=1}^{M} u_i(t)\phi_i(x)$ from FEM.
- $\mathbf{u}(t) = (u_1(t), \ldots, u_M(t))^\top$, and $\mathbf{u}_n := \mathbf{u}(n\Delta_t)$ (stepsize Δ_t).
- $(\langle \xi_n - \xi_{n-1}, \phi_i \rangle)_{i=1}^{M} = \mathbf{e}_{n-1} \sim \mathcal{N}(\mathbf{0}, \Delta_t \mathbf{G}_\theta)$ is the discretized increments of a Brownian motion process.

This gives the evolution equation after discretizing in space with FEM and in time with implicit/explicit Euler or Crank-Nicolson:

$$\mathcal{M}(\mathbf{u}_n, \mathbf{u}_{n-1}) + \mathbf{e}_{n-1} = 0, \quad n = 1, 2, \ldots, n_t.$$

Conditioning on data

State-space ideas

$$\mathcal{M}(\mathbf{u}_n, \mathbf{u}_{n-1}) + \mathbf{e}_{n-1} = 0,$$

$$\mathbf{y}_n = \mathbf{H}_n \mathbf{u}_n + \varepsilon_n, \quad n = 1, 2, \ldots, n_t.$$

Recalling that

- $\mathbf{e}_{n-1} \sim \mathcal{N}(\mathbf{0}, \Delta_t \mathbf{G}_\theta)$, model error *inside* the governing equations.
- $\varepsilon_n \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I})$: observation noise.
- Assumed $\mathbf{u}_n \perp \varepsilon_n$ for each n.

Computation via extended/ensemble Kalman filter, to give $p(\mathbf{u}_n \mid \mathbf{y}_{1:n})$ — the posterior over the model.[2]

[2] Connor Duffin et al. "Statistical Finite Elements for Misspecified Models". en. In: *Proceedings of the National Academy of Sciences* 118.2 (Jan. 2021). ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.2015006118.

Section 3

Internal waves: case study

Case study: waves in a tub

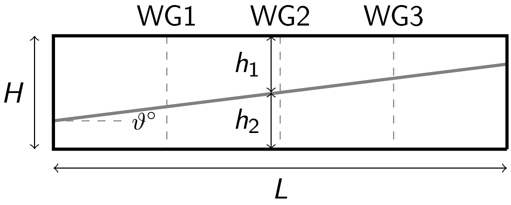
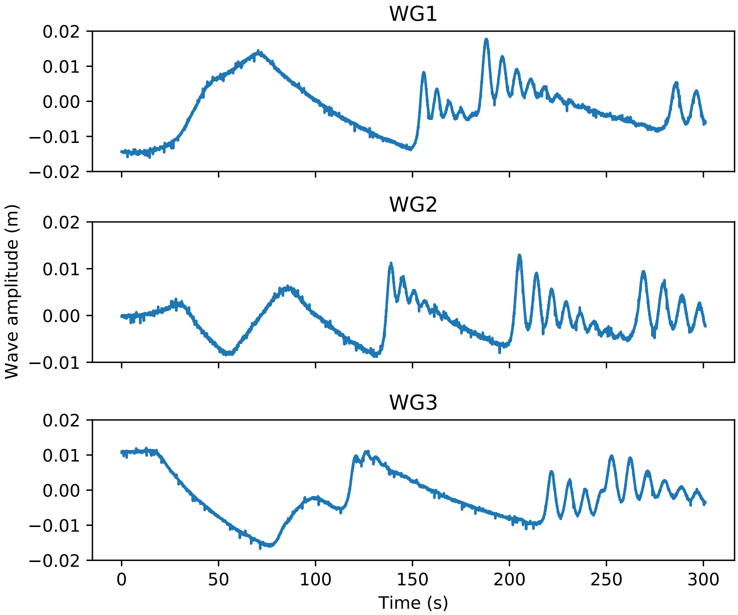


Figure: The experimental apparatus. Wave-gauges: WG1, WG2, and WG3. Initial conditions are shown as a grey line, labelled with initial angle ϑ°.

- Internal waves: waves between layers of water density.
- Classic situation: more dense water on the bottom, less dense on top (2-layer system).
- Can be modelled by the KdV equation:

$$u_t + cu_x + \alpha uu_x + \beta u_{xxx} + \nu u = 0.$$

The data



WG1
0.02
0.01
0.00
−0.01
−0.02
WG2
0.02
0.01
0.00
−0.01
WG3
0.02
0.01
0.00
−0.01
−0.02
Wave amplitude (m)
0
50
100
150
200
250
300
Time (s)

Model specificaton + setup

- Assume dynamics are modelled by the KdV equation:

$$u_t + cu_x + \alpha u u_x + \beta u_{xxx} + \nu u + \dot{\xi} = 0.$$

  Solving over $x \in [0, 6]$ m, $t \in [0, 300]$ s.

- Discretize with FEM in space, Crank-Nicolson in time, to give base equation:

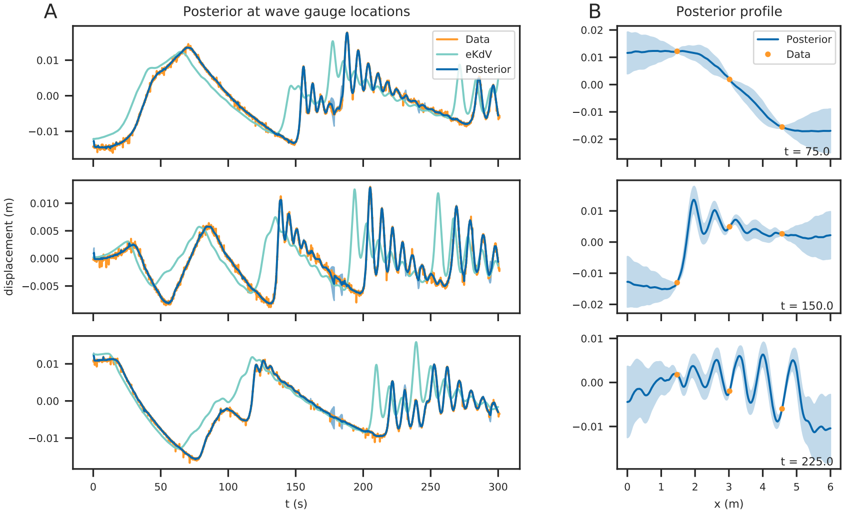
$$\mathcal{M}(\mathbf{u}_n, \mathbf{u}_{n-1}) + \mathbf{e}_{n-1} = 0, \quad n = 1, 2, \ldots, n_t.$$

- Data arrives every timestep

$$\mathbf{y}_n = \mathbf{H}_n \mathbf{u}_n + \varepsilon_n, \quad n = 1, 2, \ldots, n_t.$$

- Use EnKF to compute the posterior over the model $p(\mathbf{u}_n \mid \mathbf{y}_{1:n})$.

Results


A
Posterior at wave gauge locations
Data
eKdV
Posterior
0.01
0.00
−0.01
0.010
0.005
0.000
−0.005
displacement (m)
0
50
100
150
200
250
300
t (s)
B
Posterior profile
Posterior
Data
0.02
−0.02
t = 75.0
t = 150.0
t = 225.0
1
2
3
4
5
6
x (m)

Conclusions

StatFEM provides synthesis of data and FEM models: posterior $p(\mathbf{u} \mid \mathbf{y})$.

StatFEM methodology has now been developed for linear and nonlinear PDEs.

Full details see [2] for the linear case and [1] for the nonlinear extension.

Code available at `https://github.com/connor-duffin/statkdv-paper`.

Future work

Numerical speed-ups where possible, defining the method on the Hilbert space, possibly RKHS connections.

Applications to structural monitoring, reaction-diffusion systems (nonlinear oscillators), fluid mechanics.

Further investigation of model mismatch — more physically meaningful alternatives to Kennedy-O'Hagan?

References

Connor Duffin et al. “Statistical Finite Elements for Misspecified Models”. en. In: *Proceedings of the National Academy of Sciences* 118.2 (Jan. 2021). ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.2015006118.

Mark Girolami et al. “The Statistical Finite Element Method (statFEM) for Coherent Synthesis of Observation Data and Model Predictions”. en. In: *Computer Methods in Applied Mechanics and Engineering* 375 (Mar. 2021), p. 113533. ISSN: 0045-7825. DOI: 10.1016/j.cma.2020.113533.