Uncertainty quantification in data-driven solar physics simulations

A new project demonstrating an information-theoretic approach to quantifying uncertainty due to simulation start time for data-driven magnetofrictional simulations with memory.

Eric Hall (ehall001@dundee.ac.uk) & Karen Meyer (kmeyer001@dundee.ac.uk), Mathematics Division, School of Science and Engineering, University of Dundee

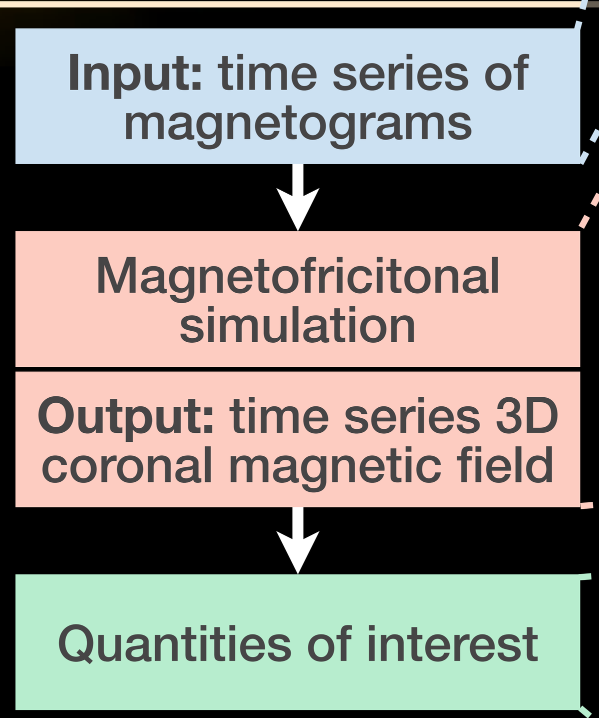
University of Dundee

Why care about uncertainty in solar physics simulations?

- Reliable and accurate predictions of solar eruptions are vital to mitigating the consequences of **severe space weather**.
- Events that lead to solar eruptions are challenging to measure directly.
- Solar physicists **rely on simulations of mathematical models** to understand the evolution of the 3D coronal magnetic configuration for a solar active region.
- Understanding uncertainty in predictions is **critical for decision support** and useful for **identifying physical mechanisms** driving active region behaviour.

Modelling the solar corona

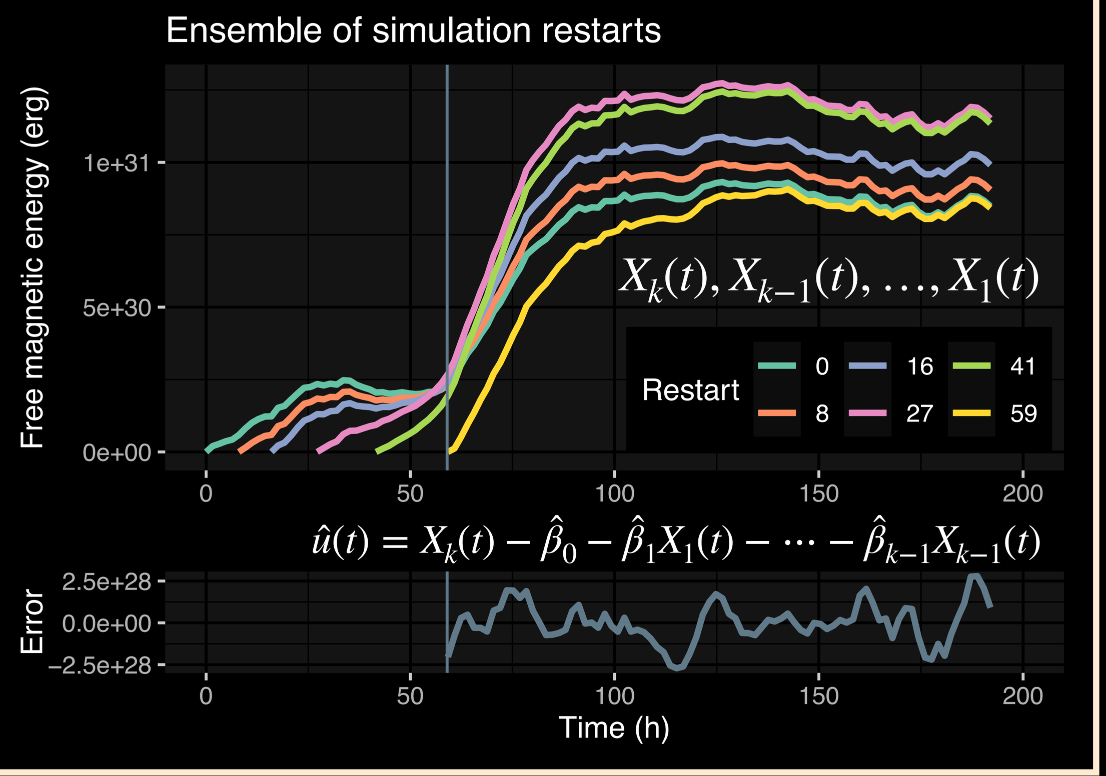
- Magnetofrictional simulation for coronal magnetic field above a solar active region (3D domain) through a sequence of related nonlinear force-free equilibria.
- Nonlinear force-free field (NLFFF) involves **stressing and relaxing of coronal magnetic fields**.
- Key element: **simulation maintains a memory of previous flux connectivities and currents**.



Mackay, Green, van Ballegooijen. Astrophysical Journal, 729:97 (2011).

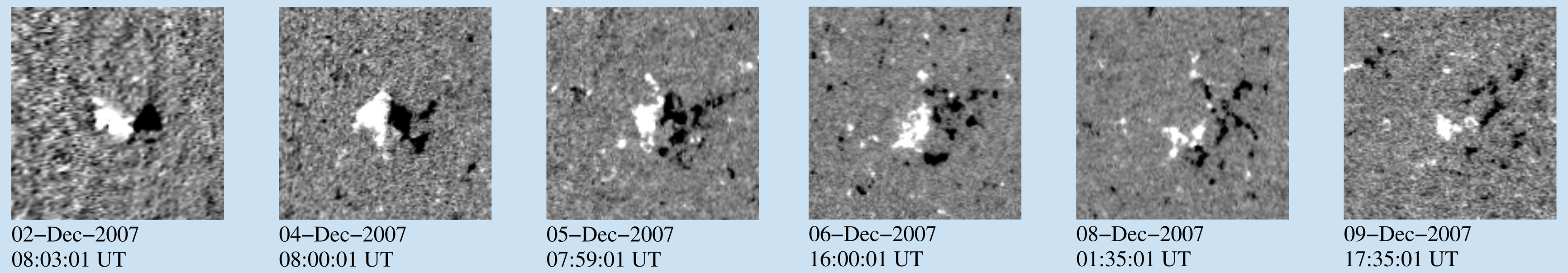
Simulation start time is a source of model-form uncertainty

- **Epistemic uncertainty**, but not reducible: data collection limited to visible regions and by availability of sensors.
- Change in simulation start is a **nonparametric/nonlocal** perturbation.
- Simulation restarts are **cointegrated**[†] - i.e. stochastic error model - not colinear.



[†]Granger, Newbold. J Econometrics, 2 (1974).

Line-of-sight magnetograms of AR10977



AR10977 exhibits an emergent phase followed by decay, with significant flux cancellation. White/black indicates positive/negative magnetic polarities. Our simulations use 120 magnetograms at a cadence of 96 minutes.

Gibb, Mackay, Green, Meyer. Astrophysical Journal, 782:71 (2014).

Magneto frictional simulation

- 3D magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$ evolved by ideal induction eq.: $\frac{\partial \mathbf{A}}{\partial t} = \mathbf{v} \times \mathbf{B}$.
- Simulate plasma experiencing a frictional force as it moves with respect to reference frame: plasma velocity is given by $\mathbf{v} = \frac{\mathbf{j} \times \mathbf{B}}{\nu'}$, $\mathbf{j} = \nabla \times \mathbf{B}$.
- As the coronal field is perturbed via boundary motions, it evolves through a series of quasi-static NLFFF distributions, satisfying $\mathbf{j} \times \mathbf{B} = 0$.
- Observed magnetograms yield a sequence of lower boundary conditions for $\mathbf{A}$ (horizontal components A_{xb}, A_{yb}) and the initial condition (via extrapolation).

Total magnetic energy	Free magnetic energy	Relative magnetic helicity
$E_{TOT} = \int \mathbf{B}^2 \, dV$	$E = \frac{1}{8\pi} \int (\mathbf{B}^2 - \mathbf{B}_p^2) \, dV$	$H_r = \int (\mathbf{A} \cdot \mathbf{B}) \, dV - \int (\mathbf{A}_p \cdot \mathbf{B}_p) \, dV$
Energy in NLFFF magnetic field.	Energy potentially available for release.	Twistedness of NLFFF magnetic field lines relative to potential.

Information-theoretic dissimilarity in time series

- Conditional MI for higher-order effects*: $I(X:Y \mid Z) = \sum_z p(z) \sum_{y,x} p(x,y \mid z) \log\left(\frac{p(x,y \mid z)}{p(x \mid z) p(y \mid z)}\right)$
- Sensitivity index quantifying impact of X on Y_t taking account the history of Y:

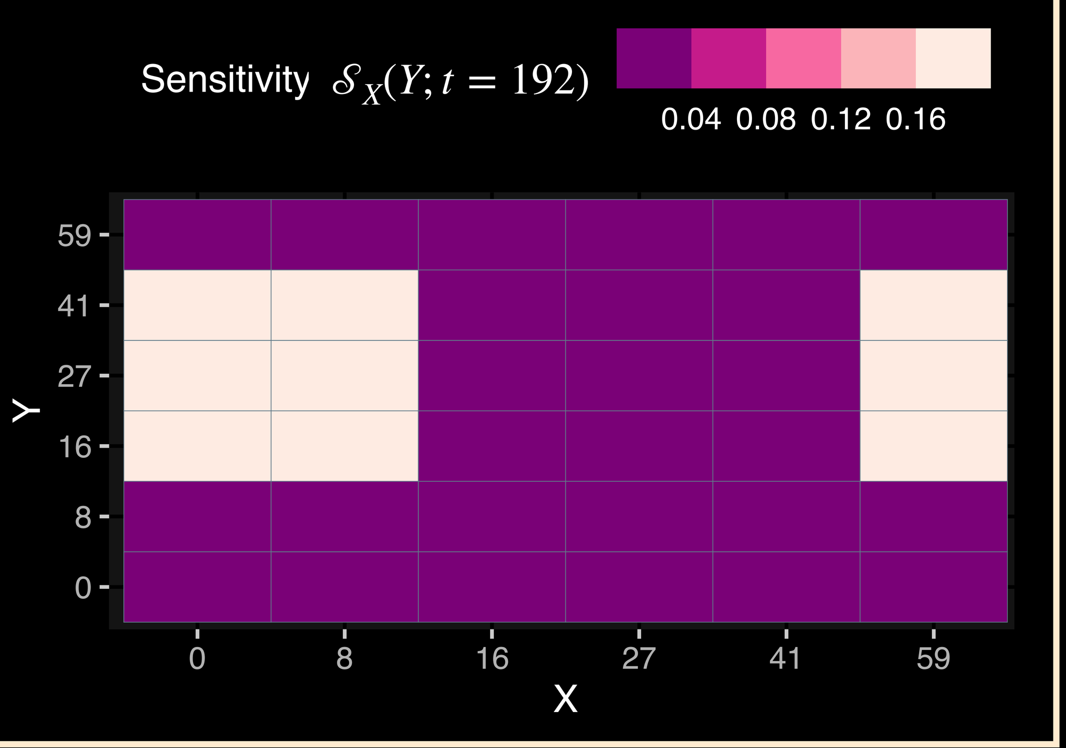
$$\mathcal{S}_X(Y;t) := \mathrm{I}\left(Y_t : \mathbf{X}_{t-1}^{(l)} \mid \mathbf{Y}_{t-1}^{(k)}\right) = \sum p(y_t, y_{t-1}^{(k)}, x_{t-1}^{(l)}) \log \frac{p(y_t \mid y_{t-1}^{(k)}, x_{t-1}^{(l)})}{p(y_t \mid y_{t-1}^{(k)})}$$

- Known as transfer entropy $T_{X \to Y}^{(k,l)}$ in other contexts[‡] (econ, nonlinear time series analysis) where the goal is to pick apart **causal relationships**.

Um, Hall, Katsoulakis, Tartakovsky. J Comput Phys, 394 (2019).
*Taverniers, Hall, Katsoulakis, Tartakovsky. J Comput Phys, 444 (2021).
[‡]Schreiber. Phys Rev Lett, 85:2 (2000).

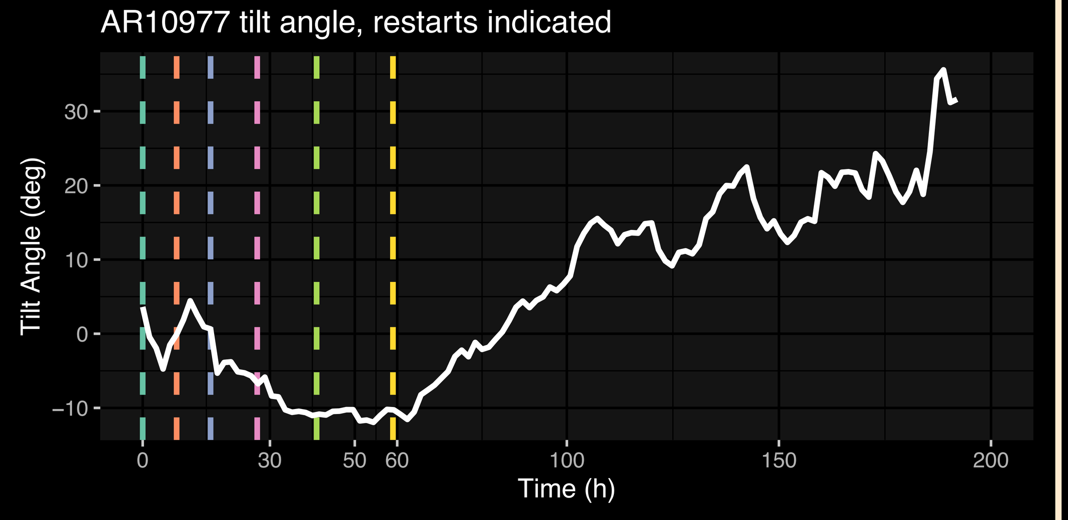
Global Sensitivity Analysis (GSA) for systems with memory

- $\mathcal{S}_X(Y;t)$ quantifies uncertainty about Y_t resolved by past states of X & Y, in excess of uncertainty about Y resolved by its own history.
- Interpretation: **uncertainty in Y due to simulation start time** (relative to X).
- Dynamic quantity calculated over rolling windows; small sample bias shuffle correction.



Validation of GSA: variation in tilt angle

- Pathwise **GSA identifies restarts at 16, 27 and 41 as having high uncertainty** relative to 0, 8 and 59 at simulation time 192 h.
- **Physical explanation:** active region tilt angle decreases (clockwise rotation, injecting negative helicity), levels off between 30-60 hr, then increases (counterclockwise rotation, injecting positive helicity and 'unshearing' the field/making it 'less energised').



Restarts at 16, 27, and 41 experience diminishing clockwise rotation before the counterclockwise rotation period. Restarts at 27 and 41 do not undergo an 'unshearing' phase removing energy. Restart at 59 begins after the active region starts to decay, so less energy is injected overall.

Conclusions

- Introduce conditional mutual information as a pathwise Global Sensitivity Analsys (GSA) method for complex physics-based simulations with memory.
- Apply our GSA approach to **quantify uncertainty due to simulation start times** in an ensemble of NLFFF simulations of the solar coronal magnetic field, which is **validated by underlying physics**.
- Presently applying these methods to a curated data set of active regions.