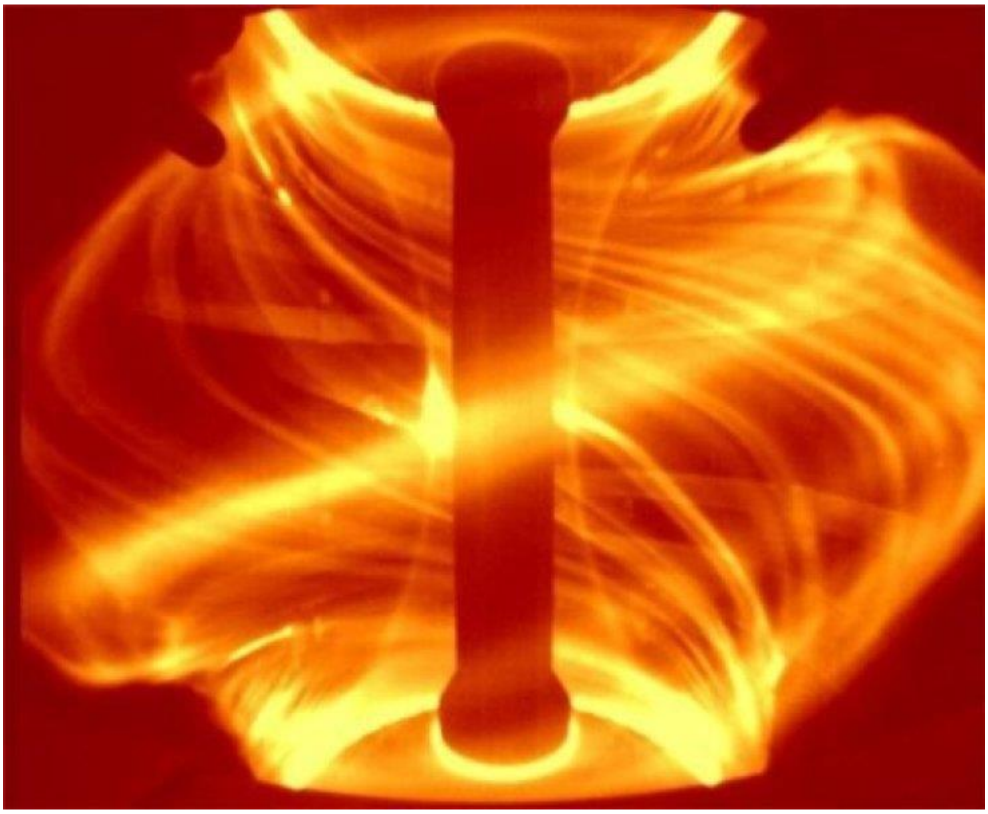
EPSRC Centre for Doctoral Training in Industrially Focused Mathematical Modelling



Techniques for initialising simple data assimilation calculations for plasma models

Ana Osojnik

Contents

1. Introduction 2

Background 2

Glossary 3

2. Mathematical model and observations 3

Model of plasma instabilities 3

Observations 3

3. Data assimilation 4

Ensemble Kalman filter 4

4. Parameter (and initial state) estimation 4

Choice of solvers 5

Methodology 5

5. Optimisation results 5

Performance profiles 5

Sampling rates and noise levels 6

6. Data assimilation results 6

7. Discussion, Conclusions & Recommendations 7

8. Potential Impact 7

References 7

1. Introduction

Background

Mathematical models are used to describe complex physical phenomena in disciplines from natural sciences to engineering. Their purpose is to provide a framework for systematically studying various effects and, ultimately, for making predictions about future events. Mathematical models are often built around equations composed of variables and parameters, and describe relationships between the variables. Solutions of these equations, which are dependent on parameters and initial state of variables/system components, determine behaviour of the system over time. No mathematical model is a perfect description of the described physical system, and thus models are of little value if they are not adjusted over time to better fit the reality of system behaviour.

The goal of data assimilation is to describe physical systems over time, by incorporating measurements of system variables into a model estimation of the unknown *true state* of the system

Data assimilation is a dynamic process in which observations (measurements) of components of the physical system are incorporated into a model estimation/prediction of the unknown *true state* of the system. The wide-ranging applicability of such dynamic model-fitting methods make them of interest to the Culham Centre for Fusion Energy (CCFE), who use mathematical models to understand the behaviour of plasma in tokamaks. Tokamaks are one of the most developed magnetic confinement devices, which are needed to confine the hot plasma in a shape of a torus. A plasma is a mixture of negatively and positively charged particles with properties of a gas. Fusion reactions between types of hydrogen atoms, called deuterium and tritium, take place inside the plasma at high temperatures. Large amounts of energy are released from these reactions, and exploiting this energy source is the practical issue underlying the development of thermonuclear fusion power.

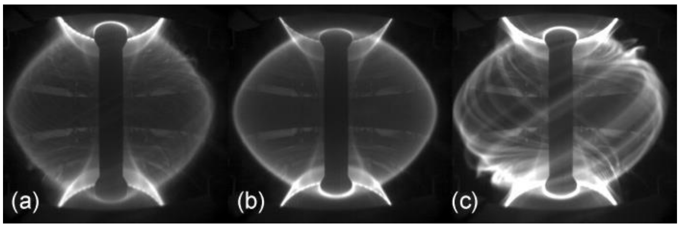


Figure 1: (a) and (b) show plasma inside a tokamak at low (a) and high (b) confinement modes, during which the plasma exhibits stable behaviour. (c) shows the unstable rearrangement of the plasma during occurrence of edge-localised modes.

A fundamental barrier to power production is the occurrence of plasma instabilities, such as sawteeth or edge-localised modes (ELMs), when large electric currents are passed through the plasma to heat it up. In tokamak experiments, these instability events are large fluctuations in electron and temperature density during large energy relaxations in the plasma, followed by particle and heat pulses/eruptions at the plasma boundary (see Figure 1 (c)). Since such instabilities are damaging to the tokamak device, there is a need to understand them better through mathematical modelling in combination with analysis of experimental data.

When working with real data and complex systems, it is difficult to identify parameters which characterise the model we are going to assimilate, and to determine the initial state of the system due to observation noise. Our aim is to consider the feasibility of using a data assimilation technique called the Ensemble Kalman Filter to improve the parameterisation of a simplified model designed to reproduce the behaviour of sawteeth and edge-localised modes.

Glossary

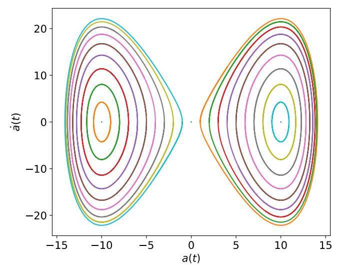
- True state: a real physical state (unknown in real systems, known in synthetic ones).
- Phase portrait: a geometric representation of the trajectories of a dynamical system in a coordinate system with axes being the values of system variables.
- $\dot{a}(t)$: the rate of change of the variable $a(t)$ with respect to time.
- Reference solution: a model solution used to generate synthetic measurements.
- Zero-mean Gaussian distribution: a description of a set of possible values in terms of their probabilities, with values close to zero being more probable.
- Sampling rate of observations: number of measurements per period of oscillation.
- Ensemble: a collection of states representing the distribution of state uncertainty.

2. Mathematical model and observations

Model of plasma instabilities

A simplified oscillatory model for plasma instabilities is introduced

Simulations of large-scale plasma behaviour have proved difficult due to the range of plasma instabilities that tend to have nonlinear interactions, and the variety of spatial and time scales present in plasma models. Simplifications of models arising from symmetry considerations of the torus-shaped tokamak enable the study of sawteeth and edge-localised modes via simpler ordinary differential equation models [1]. The purpose of these models is to describe how system components, such as electron density, temperature, or magnetic field, evolve and interact over time. Our model comprises two coupled ordinary second-order nonlinear differential equations for system variables a and b, in which there are five model parameters. The model can exhibit very different behaviours for small changes in parameters or the initial state. In the system, the equation for a is decoupled from the equation for b. Solving the equation for a, we find that the solution exhibits typical oscillating behaviour as shown in phase portraits in Figure 2 (left). In Figure 2 (right), we show solution trajectories to the full problem, given different starting points.



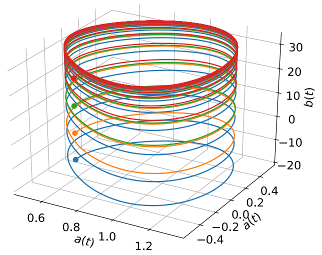


Figure 2: Phase portraits showing time-dependent paths of stable system solutions to our model for different initial states. (left) a vs $\dot{a}$; (right) a vs $\dot{a}$ vs b.

Observations

Synthetic observations are generated by adding random errors from *zero-mean Gaussian distribution* to the model solution

The aim of this project is to test hypothetical situations which use artificially generated measurements of quantities that describe the system. We obtain these measurements by first solving the model for some parameters and initial conditions. As system measurements are assumed not to be completely accurate, due to rounding errors and other unaccounted sources of noise, we add small numbers to the model solution at specific points in time to obtain a set of observations. We pick the error values randomly, using uniform and Gaussian distributions, so that most error values lie centred around zero and are sufficiently small. In real data, for system variables measured at predetermined points in time, to be descriptive of the system behaviour the spacing of measurements must be small enough. Therefore, we introduce the term *sampling rate of observations*, which describes observations by their number of samples per period of oscillation.

3. Data assimilation

Data assimilation involves combining a dynamical model with time-dependent measurements in some optimal way to approximate the true state of the system.

Ensemble Kalman filter

Data assimilation techniques attempt to combine of model forecasts with measurements in some optimal way

The Linear Kalman Filter is an iterative statistical data assimilation technique which can be used in cases where all terms in the model are linear. Suppose we have a sequence of system observations measured at successive points in time, along with model forecasts of these measurements. At every measurement time, we would like to find the *optimal approximation* to the true state of the system, that is, the most probable state given the model forecasts, and observations up to (and including) that point in time. The estimate of the true state is a weighted mean of the model forecast at the current time and the current observation. The weighting factors are derived from statistical uncertainty estimates, utilising statistical information about the mean and variance of the model forecast and observation noise.

The Ensemble Kalman Filter is an iterative statistical data assimilation method

A more computationally efficient version of the Kalman filter, applicable to nonlinear models, is the Ensemble Kalman Filter, where uncertainty statistics can be easily and intuitively estimated at every time step from a collection of state vectors, called *ensemble* members. These multiple system states represent the distribution of model forecasts with respect to uncertainty assumptions about the model and measurements. Ensemble members are iteratively evolved from one time step to the next via the original nonlinear model, and thus nonlinear evolution of errors, which the ensemble represents, is preserved.

We explore data assimilation on our simplified model of instabilities by assimilating over variables a and $\dot{a}$ using the Ensemble Kalman Filter, which we initialise with true (observation-generating) parameter values and initial conditions, as well as noisy ones. With true initialisation values, we find that the filter tracks the *reference solution* well, when 12 observations of Gaussian noise level 0.5 are assimilated per oscillation. We find that error estimates increase as the system variable a approaches zero. We believe that this difficulty is observed due to an interaction between the system dynamics and observation noise, which pulls the estimate of the true state away from the reference solution.

Initialisation of the Ensemble Kalman Filter requires good knowledge of model parameters, and an approximate guess of the initial state

Further tests confirm that the performance of data assimilation is sensitive to poor estimates of the unknown parameters in the assimilated model. There is a version of the Ensemble Kalman Filter which incorporates parameters as system variables and estimates their true values via assimilation, as for other system variables. Over time, the parameters should converge to their true values. We find that this estimation is more inclined to exhibit the following two difficulties: (i) mismatch between the most probable solution and the mean of the possible solutions found by the algorithm, and (ii) filter divergence. Therefore, there is a need for alternative parameter and initial state estimation methods, to be used as an aid for initialisation of the Ensemble Kalman Filter.

4. Parameter (and initial state) estimation

We use an optimisation approach to parameter and initial state estimation, as this has previously been shown to work for nonlinear models. We formulate a measure that

> Estimation of parameters and the initial state is posed as a miniminisation problem of the model error from initial observations

quantifies the total error between the model predictions and the system observations over the first period of oscillation, where the period is estimated from observations. We call this measure the error function. Our goal is to find the parameters and initial state that define the model, for which the error function takes the smallest possible value.

Choice of solvers

We consider three optimisation algorithms in order to find the parameters and initial state for which the error function is minimal: two local solvers for nonlinear least squares problems, DFOGN and DFO-LS [2], and a stochastic search method called Covariance Matrix Adaptation Evolutionary Strategy (CMA-ES). Our aim is to explore variations in performance of these algorithms, and their appropriateness for finding the solution to our problem.

> Local and stochastic solvers are chosen for testing of robustness of the estimation method under different starting guesses, observation noise levels and sampling rates

Local solvers perform the search for the global minimum point iteratively by zooming in on error function values in the close neighbourhood of the current *best* point held (the one with the smallest function value); new best-point candidates are taken deterministically to be the best neighbouring points. In particular, solvers DFOGN and DFO-LS choose the next iterate based on the approximation of the function in the neighbouring region using a linear model, since this is cheap to evaluate and applicable to black-box or noisy functions.

On the other hand, stochastic search methods use an element of randomness in searching for new best candidates that minimise the error function. CMA-ES searches the neighbourhood of the current best minimum point by sorting between samples from a normal distribution centred at this point. At the next iteration, the distribution is moved and reshaped in the direction of best samples.

Methodology

We pick 100 different starting guesses of parameter and initial state values, and assess how each of the solvers performs. In real problems, the observed data may provide a very good estimate of the initial system state. To imitate and test such situations, we perturb the true initial state with a small error of up to 0.05, and fix it at this value. Optimisation is then carried out with the same set of 100 starting parameters. We test the robustness of solvers for these problems to noisy or sparse data sets by varying the noise levels and sampling rates of the synthetic observations.

5. Optimisation results

Performance profiles

> The local solver DFOGN is recognized as the best-performing solver. Some stagnation at local minima was observed

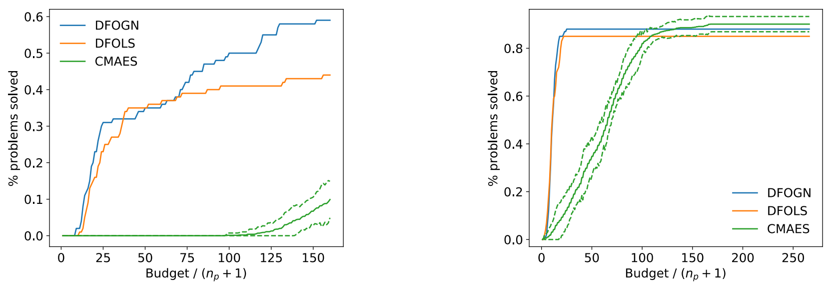


Figure 3: Performance profiles for combined parameter and initial state estimation (left) and parameter estimation (right) with $\nu = 12$ and Gaussian noise $\sigma = 0.5$. The dashed green lines indicate one standard deviation away from the mean profile for CMA-ES.

A key measure of the effectiveness of an algorithm is how the proportion of problems solved varies with the scaled number of function evaluations (scaled budget). We present our results in Figure 3. We see that, for combined parameter and initial state estimation with and without noise, DFOGN is the best-performing solver overall, since from Figure 3 (left) we can see that the increase in the number of problems solved to high accuracy was the steepest and the final proportion reached is the highest at about 60%. DFOGN is followed closely by DFO-LS, whereas CMA-ES consistently performs considerably worse, when averaged. One issue that we encountered is that these local solvers stagnate at local function minima. We find that increase in sampling rates of observations decrease the number of points local solvers get stuck at.

All solvers give high performance rates when used to estimate the parameters

We also consider parameter estimation, as in Figure 3 (right), where we see very encouraging results for all solvers, since within 500 function evaluations the solvers find the minimum for about 90% of starting guesses. DFOGN and DFO-LS are again consistently the fastest solvers over CMA-ES, although we also observe that CMA-ES can solve comparable proportions of problems if the budget is sufficiently large.

Sampling rates and noise levels

When varying noise levels and sampling rates (denoted as ν) with DFOGN solver, we find that larger sampling rates seem to mitigate large noise levels. There is a clear trade-off visible between the two when estimating parameters and the initial state, as shown in Figure 4 (left). When only estimating the parameters, this trend is less strong, and we see in Figure 4 (right) that the average error over the runs for different starting guesses stays constant. This is a positive result pointing to errors in parameter estimation occurring mostly due to small noise in the known initial state.

Larger noise levels in observations require higher frequency of observations

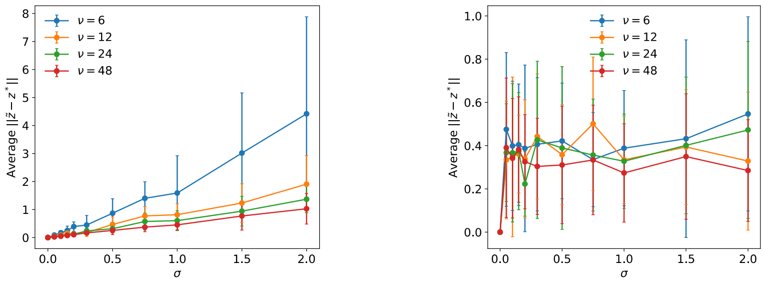


Figure 4: Plots of average error of estimates for combined parameters and initial state case (left) and just parameters (right). ν is the sampling rate and σ is the Gaussian noise level.

6. Data assimilation results

We develop a new approach for initialisation of the Ensemble Kalman Filter, which involves using observations over the first period for parameter and initial state estimation. The optimised parameter values are then used to initialise the model to be assimilated, whereas the optimised initial state is used for initialisation of the ensemble of initial model states. We show the results for one run in Figure 5, where we see that the ensemble average matches the reference solution very well, as desired.

Data assimilation preceded with parameter and initial state estimation gives encouraging results.

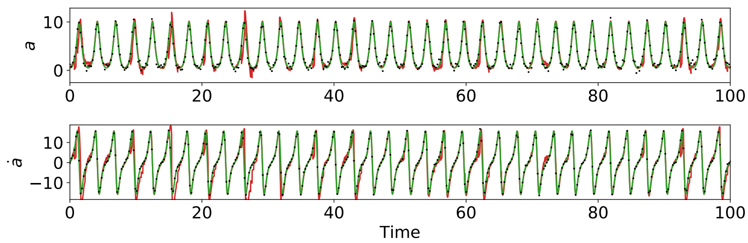


Figure 5: Ensemble average (red) with observations (black), and reference solution (green) for observation sampling rate $\nu = 12$ and Gaussian noise level $\sigma = 0.5$

7. Discussion, Conclusions & Recommendations

We have considered a statistical data assimilation method applied to the problem of identifying the parameters and initial conditions in a simple model for plasma instabilities, and using synthetic data. We found that the Ensemble Kalman Filter assimilation of the nonlinear system exhibits difficulties in certain regions of oscillation, which could potentially be improved by reducing the impact of very noisy observations or other methods that prevent ensemble members from diverging to other solutions.

To treat the initialisation difficulties we encountered, we developed a method for estimation of parameters and initial state of the system, aimed at finding sufficiently good estimates for successful initialisation of data assimilation. Tests showed the local solver DFOGN as the most favourable solver in many aspects and situations, with fast progress and small errors in a high proportion of cases when stagnation did not occur. CMA-ES was seen to overcome those difficulties, but it is thought to require high computational efforts for good accuracy estimates, which in particular makes its scalability to larger size systems questionable. We found that the final error of optimised estimates for noisy data could potentially be controlled through increased sampling rate.

The size of the parameter estimation problem does not increase with system dimension, but only with the number of parameters. When considering the problem with a slightly noisy fixed initial state, the success of the solvers we tested points to possible extensions of parameter estimation to higher dimensional systems.

Extensions to the full coupled model requires imposition of nonlinear inequality constraints on parameters and initial states

CCFE would like to perform data assimilation of the full coupled model for plasma instabilities, for which parameters and initial state are poorly known. The formulation of our estimation scheme is extendible, with the additional requirement that we will need to impose nonlinear inequality constraints on the optimised variables. This is to ensure that we search among oscillatory, non-divergent solution trajectories. We could impose constraints through a penalty term in the error function, in conjunction with the use of global solvers nonspecific to least squares problems.

Overall we have seen promising results which confirm the feasibility of parameter and initial state estimation in nonlinear plasma models, and suggest possible benefits to statistical data assimilation as well as other applications where model tuning is required.

8. Potential Impact

Sawteeth and edge-localised modes are proving a major challenge in the development of tokamak-based commercial fusion power plants, since they result in eventual damage to the device or degradation of plasma if not contained. Better understanding of these instabilities is essential, and applications of data assimilation techniques to plasma models along with data gathered from experiments with tokamak devices, such as JET and MAST in Culham, could become a valuable contribution to increasing understanding.

Wayne Arter, at CCFE, commented "*This could lead to simple models, running in real time to help control the ITER experiment to optimise power production, for eventual use in commercial reactors. These techniques should have more general application in industrial plant control, say in chemical reactors.*"

References

1. A Murari, W Arter et al (2011) *Symmetry Based Analysis of Macroscopic Instabilities in Thermonuclear Plasmas*. Available at: http://www.euro-fusionscipub.org/wp-content/uploads/2014/11/EFDP11007.pdf
2. L Roberts, C Cartis et al (2017) *A Derivative-Free Solver for Least-Squares*. To appear.