Statistical modelling for energy system planning

Chris Dent
chris.dent@durham.ac.uk
DTU Compute Seminar
11 May 2015

Particular thanks to Stan Zachary (Heriot-Watt University), Antony Lawson, Amy Wilson, Meng Xu and Michael Goldstein (Durham University), National Grid, Ofgem, DECC

Durham University
Durham Energy Institute

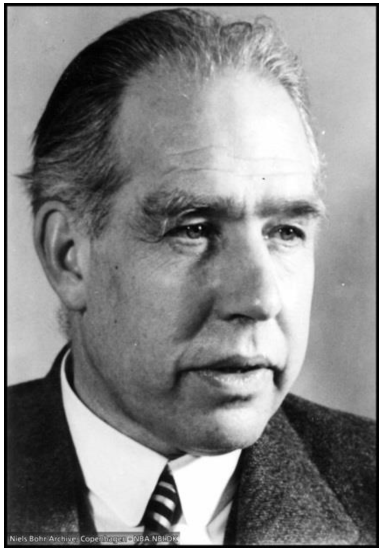
Durham University

Durham University
Durham Energy Institute

- 'Risk and Reliability Modelling of Energy Systems' day, each November, Google *Durham Risk Day*
- Durham Energy Institute promotes interdisciplinary research

Uncertainty in modelling

- Some well know quotations about modelling…
  - *All models are wrong, but some are useful* (George Box)
  - *Simple models for insight, complex models for quantification* (Ben Hobbs)
  - *Prediction is very difficult, especially about the future* (Niels Bohr)
  - *Don't be too proud of this technological terror you've constructed* (Darth Vader to Death Star commander, Star Wars Episode IV)



Durham University

Durham Energy Institute

Content

- Example – adequacy assessment
- General formulation of modelling uncertainty
- Interlude – climate projection
- Decision problems (network and generation investment)

- What do I mean by statistical modelling?
  - I am talking about quantifying uncertainties in planning (i.e. capital investment decisions not fixed, wider range of uncertainties than ops)
  - I do not just mean estimation of parameters from numerical data (with confidence intervals)
  - I also mean general issues of management of uncertainty, combining traditional data with expert judgement
  - In this talk the maths is not most the important aspect
  - Heavily influenced by Michael Goldstein and Jonty Rougier (see reading list), text on slides my own except for “MG” slides

EXAMPLE: ADEQUACY ASSESSMENT

Durham University

Durham Energy Institute

Risk of absolute supply shortages

Durham University

Durham Energy Institute

Adequacy assessment: formulation

- Snapshot margin of available generating capacity over demand

$$Z = X + Y - D$$

  - X, Y: available existing (conventional) and additional (e.g. wind) generating capacity, D: demand
  - Loss of Load Probability:

$$[\mathrm{LOLP}] = \mathrm{P}(Z < 0)$$

- Unified framework for annual peak and whole season calculations
  - X, Y, D: demand and available capacity at a randomly chosen time
  - Expectation values conditional on assumed state of knowledge

$$[\mathrm{LOLE}] = \sum_t \mathrm{P}(Z_t < 0) = n_{\mathrm{periods}}[\mathrm{LOLP}]$$

Adequacy assessment: notes

- X usually assumed independent of (Y, D)
  - Good reason to suspect that Y and D not independent if Y is renewables
- Hindcast: use empirical joint wind/demand distribution in calculation

$$[\mathrm{LOLE}] \propto \sum_t \mathrm{P}(X + w_t < d_t)$$

  - (w_t, d_t) appropriately scaled to future system background under study
  - Included within unified snapshot picture by taking an instance of (Y, D) to be a random sample from the historic data (w_t, d_t)
- Capacity value metrics visualise additional generation's contribution within adequacy calculations
  - Effective Load Carrying Capability (ELCC)

$$\mathrm{P}(X < D) = P(X + Y < D + \nu_Y^{\mathrm{ELCC}})$$

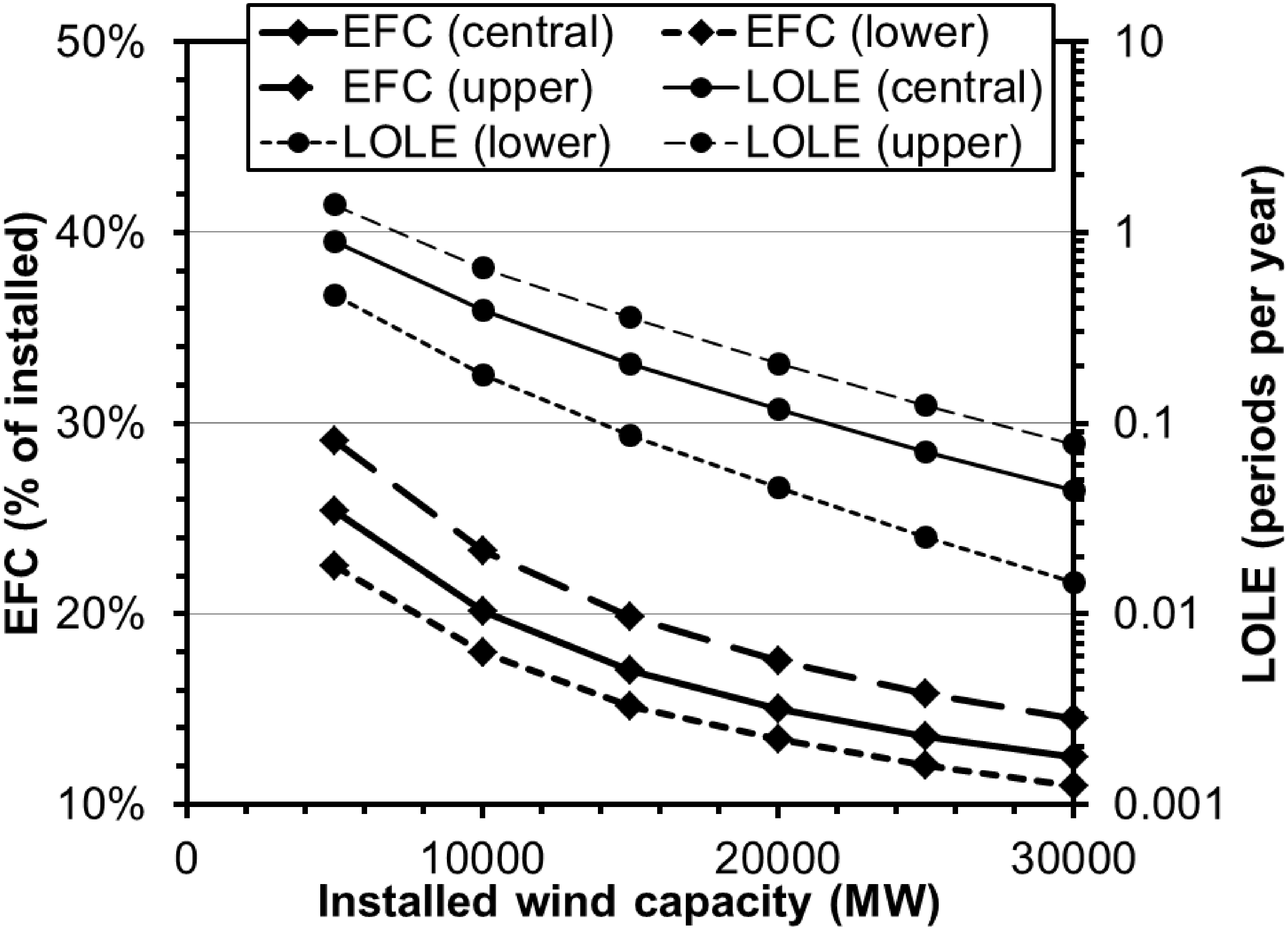
  - Equivalent Firm Capacity (EFC)

$$\mathrm{P}(X + Y < D) = P(X + \nu_Y^{\mathrm{EFC}} < D)$$

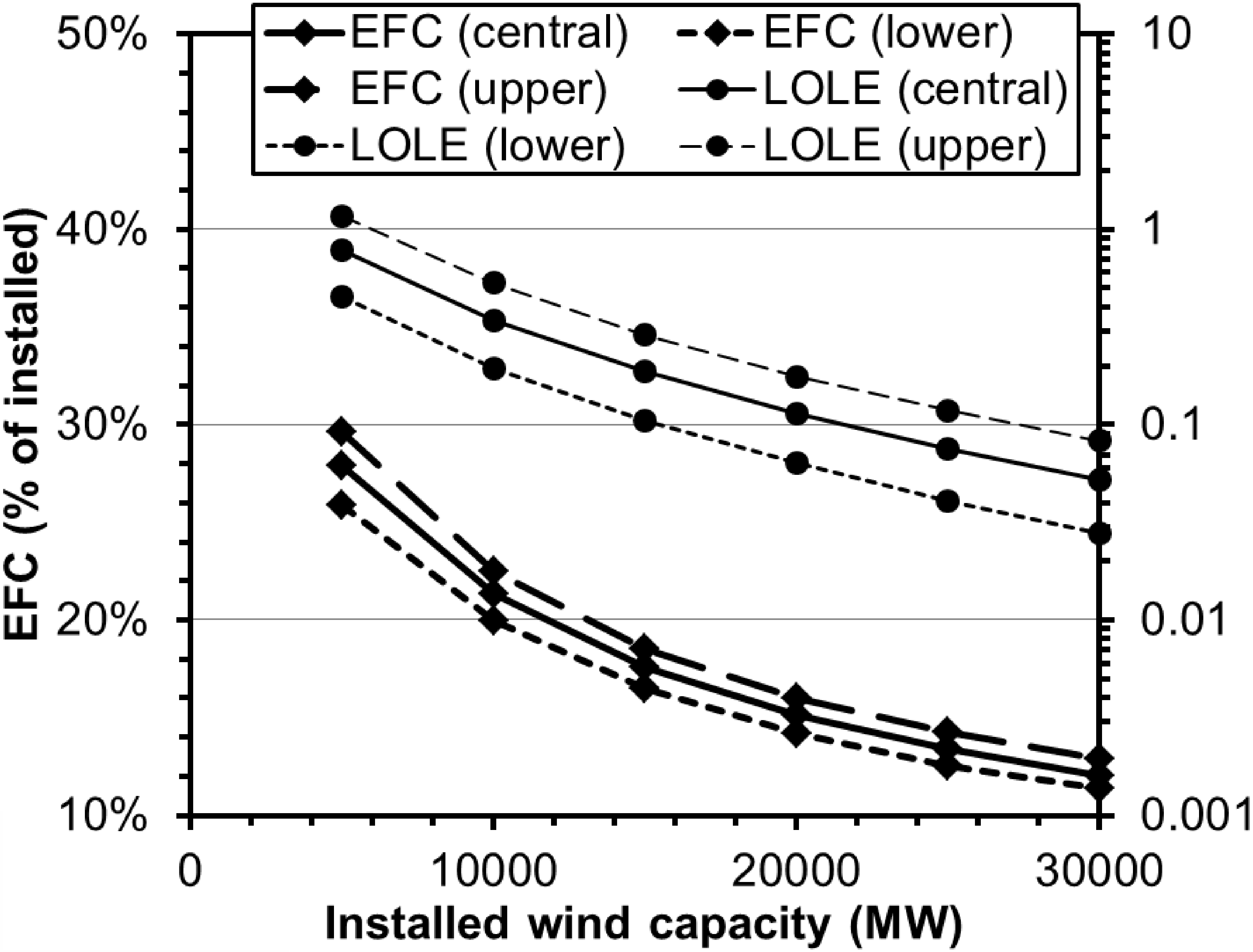
Durham University

Durham Energy Institute

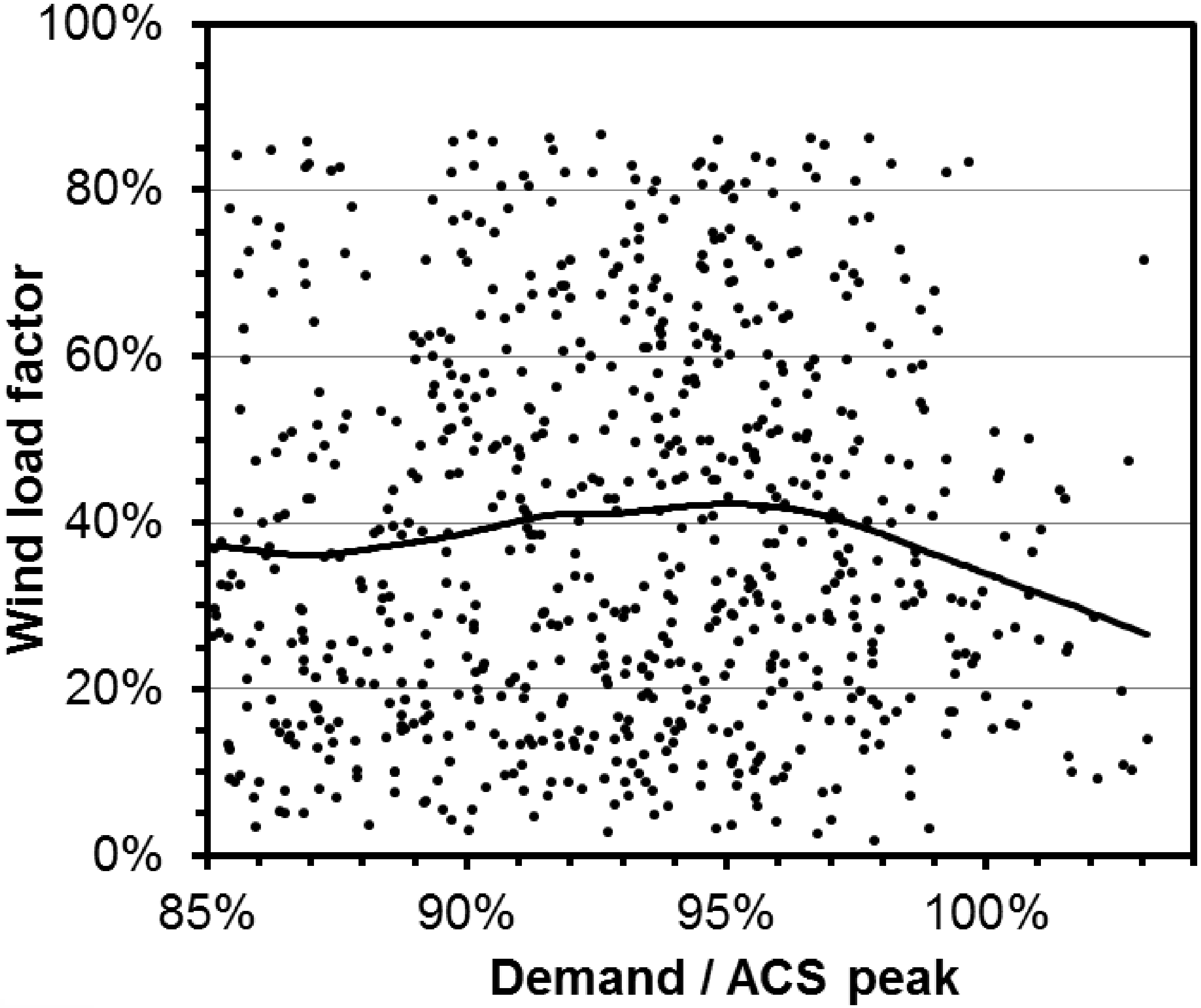
Results: hindcast



Results: $Y - D$ independence



$Y - D$ independence?!



- Likely too strong an assumption
- Exchanges sampling uncertainty for structural uncertainty

Durham University
Durham Energy Institute

MODELLING UNCERTAINTIES: GENERAL ISSUES

Durham University
Durham Energy Institute

Classification of uncertainties

- Aleatoric uncertainty (From Latin *aleator*, gambler; *alea*, dice)
  - Things which are different each time we run an experiment, even if we knew system structure and parameters perfectly
  - e.g. outcome of tossing a coin, …
  - or which conventional generating units are broken down
- Epistemic uncertainty (From Greek *episteme*, knowledge)
  - Uncertainties which can in principle be reduced if we obtain better information
  - e.g. forecast of underlying energy demand growth
  - Sometimes it is possible to reduce epistemic uncertainties greatly (e.g. probability of a head), sometimes not (e.g. whether CCS will ever work on a commercial scale, or if data are inevitably very limited)

Descriptive versus inferential

- Descriptive statistics
  - Simply a representation or summary of measured data
  - Might be a sample mean, SD, …
  - or a graphical representation
- Inferential statistics
  - Drawing conclusions about a whole population based on a sample which is smaller than the population – estimates depend on both underlying population and data one happens to sample
  - In a power systems context, may involve drawing conclusions about probabilities of future events based on historic data – what uncertainties does this introduce?
  - Limited sample size; indirect relevance of data such as historic/future demand patterns, climate variability and change, …

Nature of uncertainties in adequacy

- Given all probability distributions, out-turn of 'trial' is uncertain
  - However LOLE or EFC are deterministic functions of the data – i.e. given probability distributions, these model outputs are fixed
  - 'We have uncertainty' – usually where people stop in adequacy
- We have very limited directly relevant data
  - e.g. even under strong (and tenuous) assumptions of stationary climate, considerable uncertainty over what distribution of net demand is
  - 'We have uncertainty in the uncertainty'
- There is inevitably subjective judgment involved in carrying out the adequacy assessment in these circumstances
  - Reasonable people will differ on quantification of this uncertainty
  - 'We have uncertainty in the uncertainty in the uncertainty'!
- Uncertainty in relationship model structure/real system relationship
  - Remember: investigating sensitivity to parameters is not the whole story
- Subjective Bayesian picture: 'We have uncertainty'!

Model structure - Hot weather

France imports UK electricity as plants shut

Robin Pagnamenta, Energy and Environment Editor
Published at 12:00AM, July 3 2009

France is being forced to import electricity from Britain to cope with a summer heatwave that has helped to put a third of its nuclear power stations out of action.

With temperatures across much of France surging above 30C this week, EDF's reactors are generating the lowest level of electricity in six years, forcing the state-owned utility to turn to Britain for additional capacity.

Fourteen of France's 19 nuclear power stations are located inland and use river water rather than seawater for cooling. When water temperatures rise, EDF is forced to shut down the reactors to prevent their casings from exceeding 50C.

A spokesman for National Grid said that electricity flows from Britain to France during the peak demand yesterday morning were as high as 1,000MW — roughly equivalent to the output of Dungeness nuclear power station on the Kent coast.

Nick Campbell, an energy trader at Inenco, the consultancy, said: "We have been exporting continuously from this morning and the picture won't change through peak hours, right up until 4pm."

EDF warned last month that France might need to import up to 8,000MW of electricity from other countries by mid-July — enough to power Paris — because of the combined impact of hot weather, a

Climate change puts nuclear energy into hot water

By James Kanter
Published: Sunday, May 20, 2007

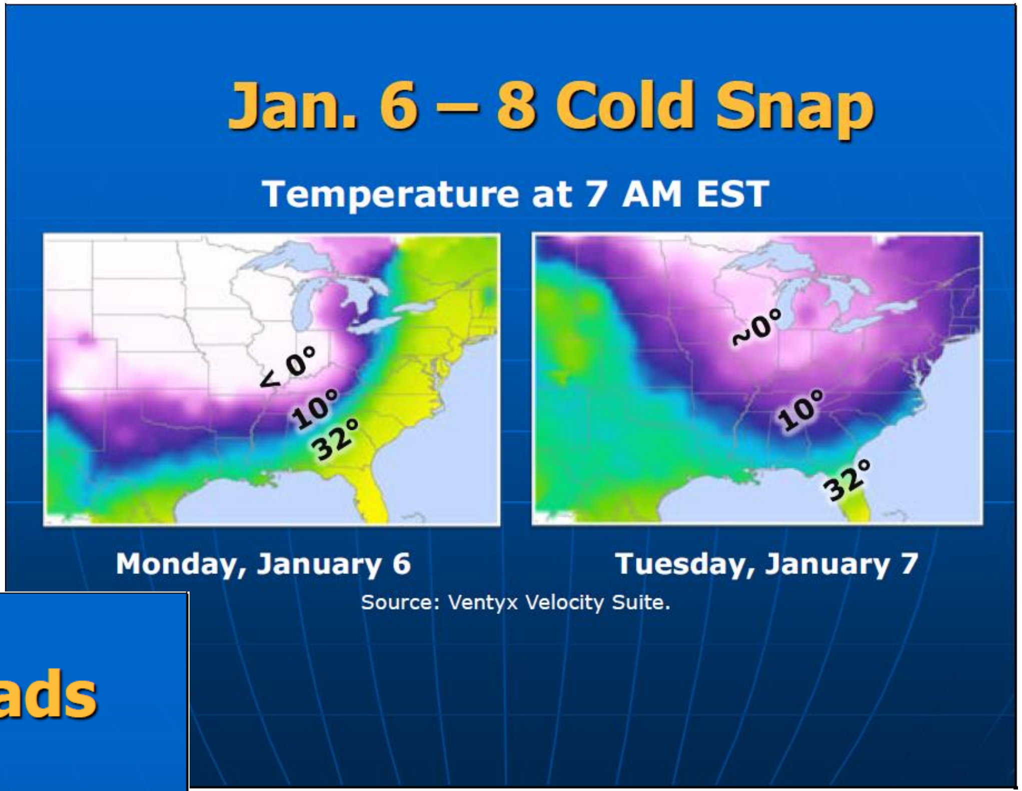
PARIS — Could climate change be the latest jinx on nuclear power?

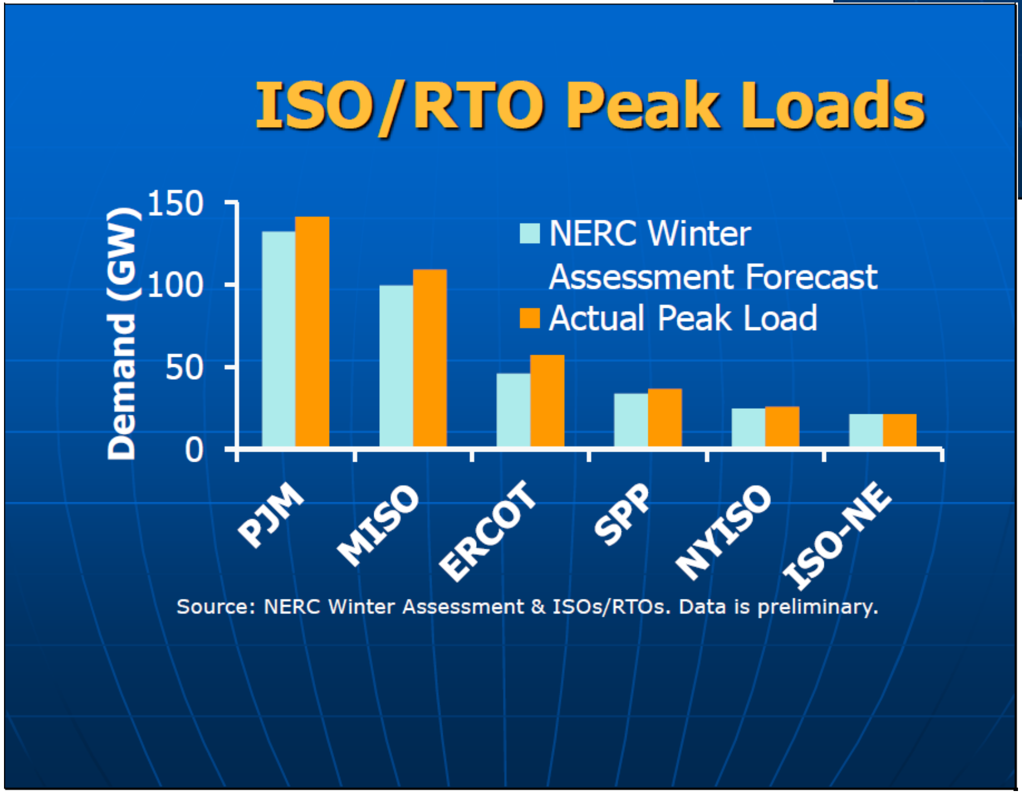
Long regarded with suspicion because of radioactivity, nuclear power suddenly has a revived image, thanks to the idea that many more plants could be built without worsening global warming. Unlike power plants fired by coal and natural gas, nuclear fission produces no carbon dioxide, the main greenhouse gas.

But there is a less well-known side of nuclear power: It requires great amounts of cool water to keep reactors operating at safe temperatures. That is worrying if the rivers and reservoirs which many power plants rely on for water are hot or depleted because of steadily rising air temperatures.

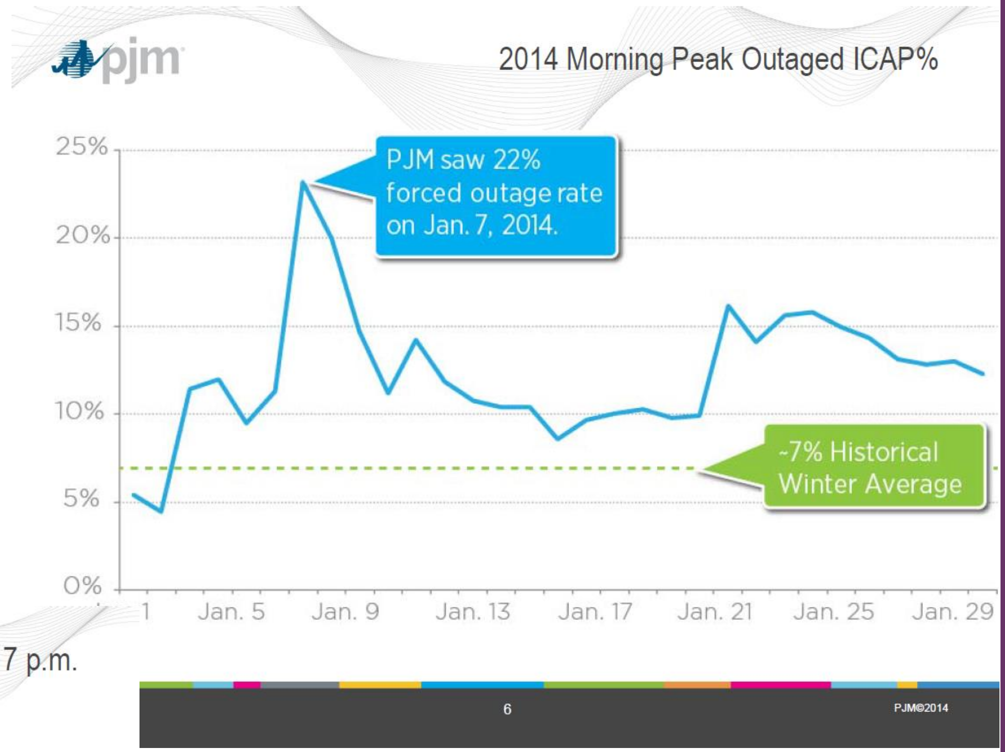
If temperatures soar above average this summer - let alone steadily increase in years to come, as many scientists predict - many nuclear plants could face a dilemma: Either cut output or break environmental rules, in either case hurting their reputation with customers and the public.

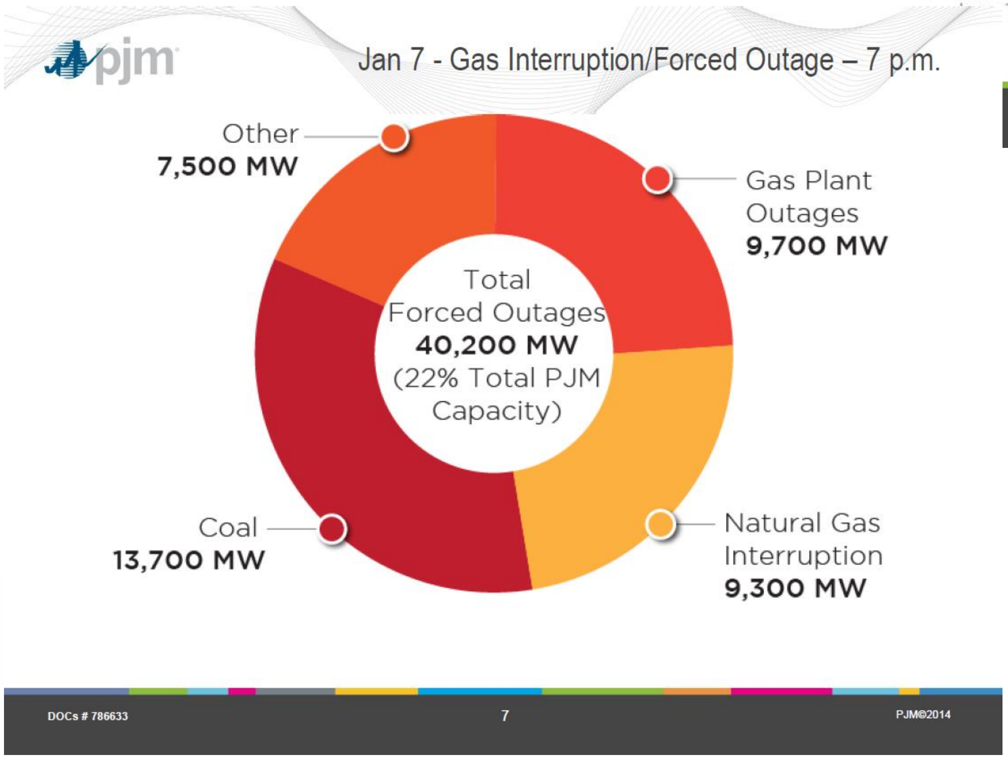
Cold weather





Cold weather





Adequacy: modelling uncertainties

- Available conventional capacity
  - Assumed independent of all else
  - Individual units independent of each other
- Treatment of VG – demand relationship
  - Quantity of data
  - Assumptions about wind-demand (in)dependence
- Demand evolution
  - Underlying growth
  - Change of technologies connected
  - Small number of extreme years in datasets
- Network effects
  - Internal to system
  - Interconnection to other systems
- And…
  - Expected value indices – one shot at a winter

Durham University
Durham Energy Institute

UNCERTAINTY IN COMPLEX COMPUTER MODELS: GENERAL FORMULATION

Durham University

Durham Energy Institute

General formulation

- Typically we may regard a computer model as a function
$$y = f(x)$$
  - x is input data and parameter choices
  - y is model outputs
- Constraint costs (costs of required re-dispatch of generation due to finite network capacity)
  - x is network capacities, generation locations, availability properties and costs, demand profile, etc etc
  - y is constraint cost
- Generation investment
  - x is system background inc demand growth, fuel prices, possibly paramterisation of companies' decision criteria, etc etc
  - y is investment outcome, or perhaps adequacy risk
- However we are not sure what x should be
  - Nor whether, for a particular choice of x, f gets the consequences right

Sources of uncertainty (MG)

- Parametric uncertainty (each model requires a, typically high dimensional, parametric specification)
- Condition uncertainty (uncertainty as to boundary conditions, initial conditions, and forcing functions)
- Functional uncertainty (model evaluations take a long time, so the function is unknown almost everywhere)
- Stochastic uncertainty (either the model is stochastic, or it should be)
- Solution uncertainty (as the system equations can only be solved to some necessary level of approximation)
- Structural uncertainty (the model only approximates the physical system)
- Measurement uncertainty (as the model is calibrated against system data all of which is measured with error)
- Multi-model uncertainty (usually we have not one but many models related to the physical system)
- Decision uncertainty (to use the model to influence real world outcomes, we need to relate things in the world that we can influence to inputs to the simulator and through outputs to actual impacts. These links are uncertain.)

Durham University

Durham Energy Institute

Emulators

- Full computer model

$$y = f(x)$$

- Emulator

$$\tilde{y} = \tilde{f}(x)$$

  - Encodes our state of knowledge about y for each x
  - y only known precisely for values of x at which we have evaluated f
  - Elsewhere uncertainty in y represented as a probability distribution, i.e. $\tilde{f}(x)$ is a random variable
  - e.g. mean of $\tilde{f}(x)$ is an interpolator between evaluations of f
  - SD of $\tilde{f}(x)$ quantifies our uncertainty in f for values of x at which we have not evaluated
  - Covariance between $\tilde{f}(x)$ and $\tilde{f}(x')$ encodes how much an evaluation of f at x' is telling us about $f(x)$
- Constructing emulator based on evaluations of f is a statistical inference / estimation problem
  - Might do this by building an emulator of a faster model (e.g. old version or coarse grid) then updating based on limited number of runs of full model

Durham University

Durham Energy Institute

INTERLUDE: CLIMATE PROJECTION MODELS

Durham University
Durham Energy Institute

Climate models

- Probably the most prominent use of large scale mathematical models in guiding public policy
  - Physics-based models of atmosphere and oceans (CFD)
  - Calibrate against limited history
  - Physical processes may not be represented microscopically (sub grid scale in time or space)
  - Parameters may be tuned or derived from finer mesh models
  - Limited data to assimilate for initial conditions
  - Numerical convergence (size of time steps) – and bugs! (think how complex these models are, no one person can claim “ownership”)
- General attitude in climate model community is to use faster computers to include more physical detail
  - Alternative is to do multiple runs to explore uncertainties discussed previously

Climate models

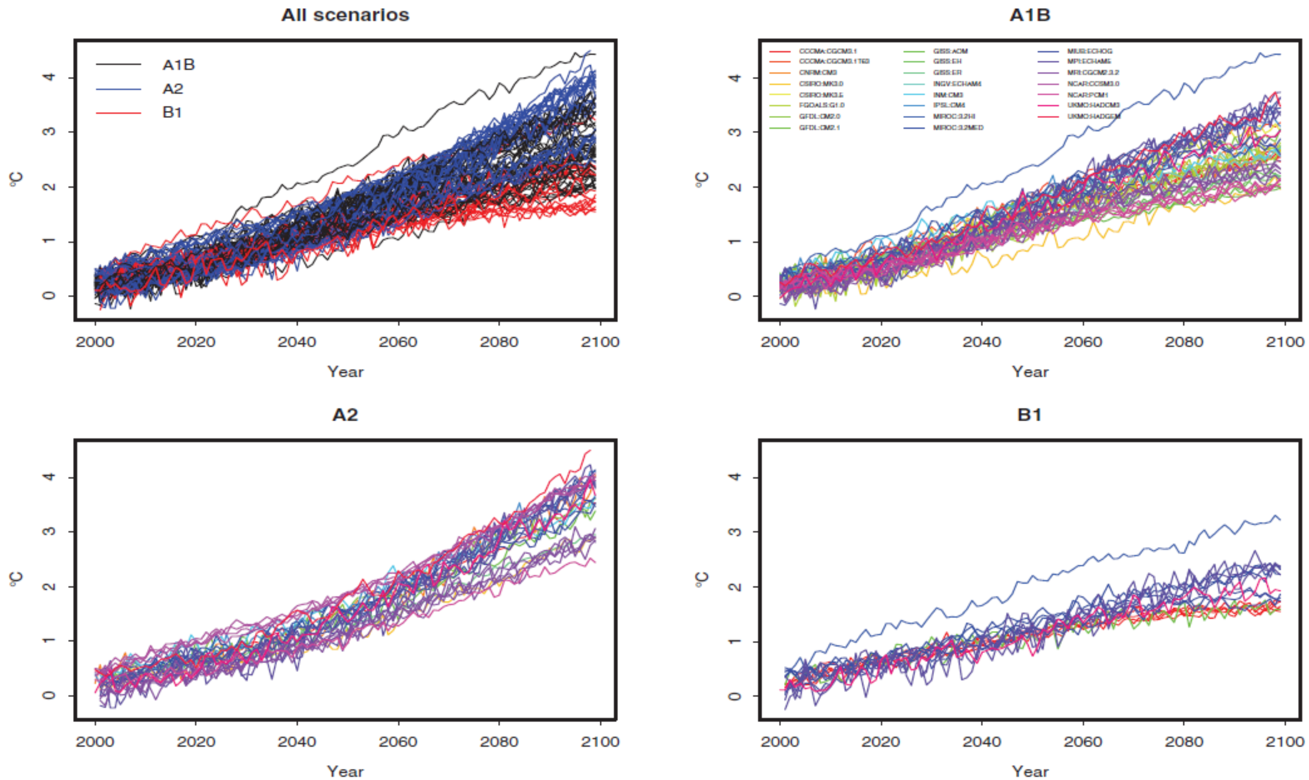


Figure 3. Projections of annually averaged global mean temperature for the 21st century, derived from 23 different GCMs as used in the IPCC AR4 (data obtained from the *KNMI Climate Explorer* at `http://climexp.knmi.nl/`). Projections are shown as changes from the 1980–99 mean. The first panel shows all of the available simulations from the A1B, A2 and B1 emissions scenarios (see Box 2). The remaining panels show the simulations from each scenario individually, with colours corresponding to individual GCMs. Not all GCMs were run for all scenarios. Conversely, some GCMs were run more than once per scenario with different initial configurations of the prognostic variables

- Results of different models often taken as uncertainty in outputs
  - However for each of these traces there is uncertainty
  - Might all the models be wrong in the same way?

Durham University

Durham Energy Institute

DECISION PROBLEMS

Model-based decision making (MG)

- In practice, it is extremely rare to find a serious quantification of the total uncertainty about a complex system arising from the all of the uncertainties in the model analysis.
- Therefore, for all applications, no-one really knows the reliability of the model based analysis. Therefore, there is no sound basis for identifying appropriate real world decisions based on such model analyses.
- This is because modellers/scientists don’t think about total uncertainty this way… nor do most statisticians.
- Policy (decision) makers don’t know how to frame the right questions for the modellers
- There are few funding mechanisms to support this activity (CJD: costs as much as building model, lots more runs, makes interpretation and use of results SEEM harder, allows less definitive conclusions)
- And it is hard!

Simple decision problem

- Suppose: (random variables in caps, fixed quantities lower case)
  - we are making a capital investment decision z (say in network)
  - uncertainty in system background is represented by a random variable X
  - the cost (initial capital + future operations) is $Y = f(X, z)$
- If we knew everything about f, and were risk-neutral (i.e. seek to minimise expected cost)

$$\min_z \int \mathrm{d}x\, \rho_X(x)\, f(x, z)$$

  - $\rho_X(x)$ is the pdf representing uncertainty in system background X
- Can introduce risk aversion through utility function

$$\min_z \int \mathrm{d}x\, \rho_X(x)\, g(f(x, z))$$

Durham University

Durham Energy Institute

Decision problem with emulator

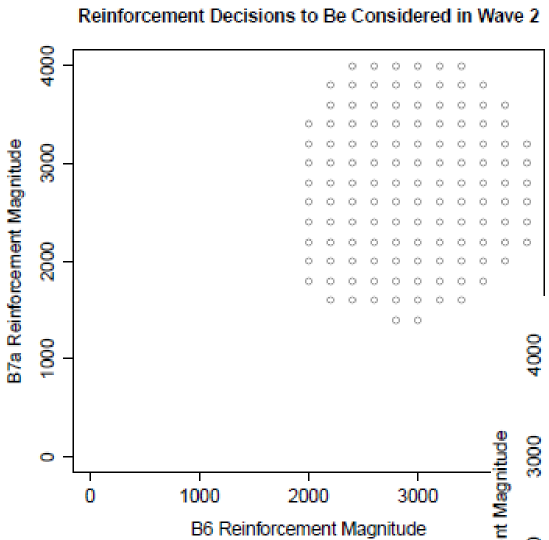
$$\min_z \int \mathrm{d}x \, \rho_X(x) \, f(x,z)$$

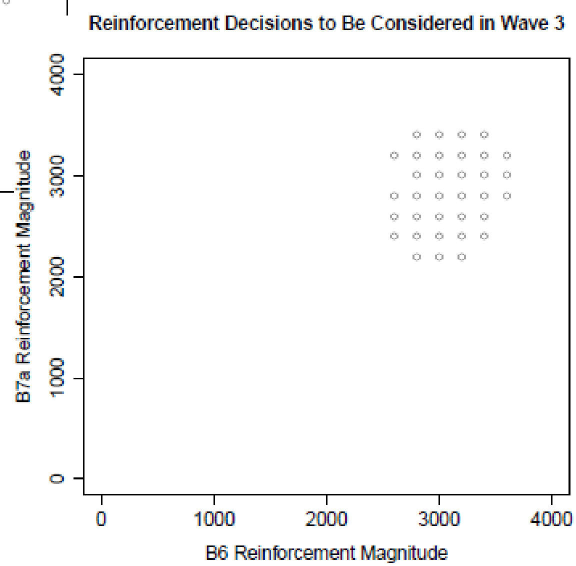
- If we have the emulator $\tilde{y} = \tilde{f}(x)$ representing issue that we have not evaluated f everywhere, and thus have uncertainty over value of f for almost all x, we might write

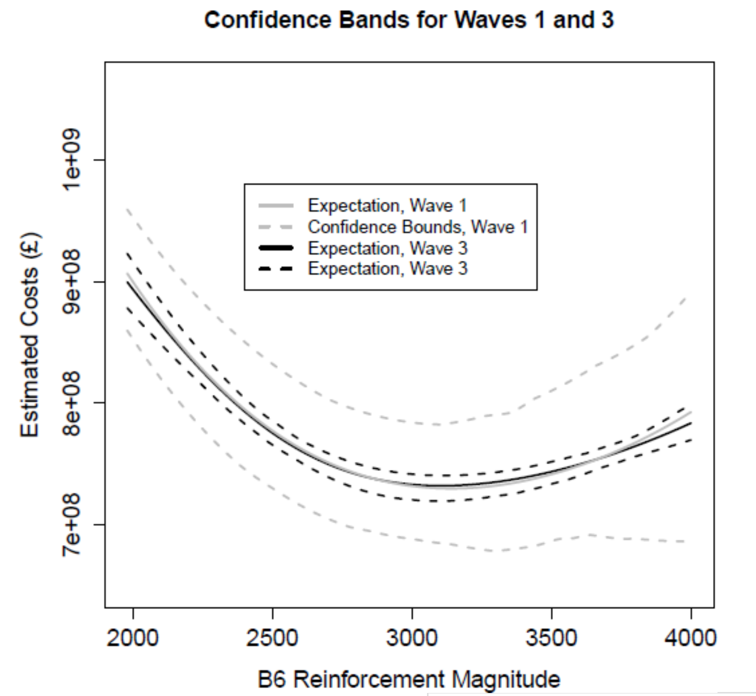
$$\min_z \int \mathrm{d}x \, \rho_X(x) \, \tilde{f}(x,z)$$

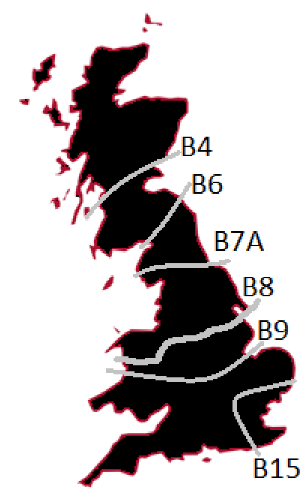
  - Problem – what does this mean as the value of the objective $\int \mathrm{d}x \, \rho_X(x) \, \tilde{f}(x,z)$ is not known precisely
  - Can only determine window in which optimum might lie
  - Then by narrowing down decision space can produce better emulator (make sampled points in relevant region more dense)
- Slight diversion – decision depends on quantification of uncertainty
  - Look at sensitivity of "optimal" decision to precise quantification
    - If not sensitive, just do it
    - If sensitive, think carefully

Transmission system planning (Lawson)







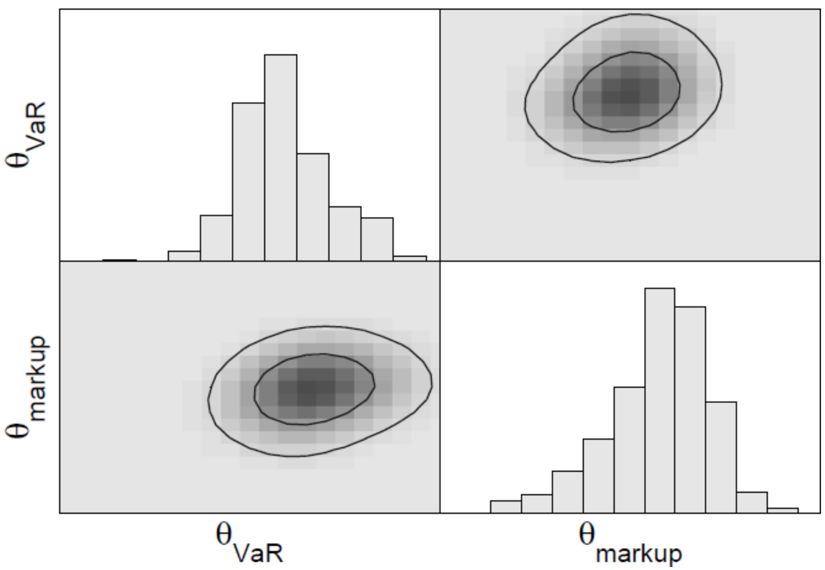


- Initially explore full range of possible reinforcements on two boundaries in GB system (with emulation)
  - Quantify uncertainty in model inputs
  - Narrow down, sample more densely (waves)
  - Introduce risk aversion

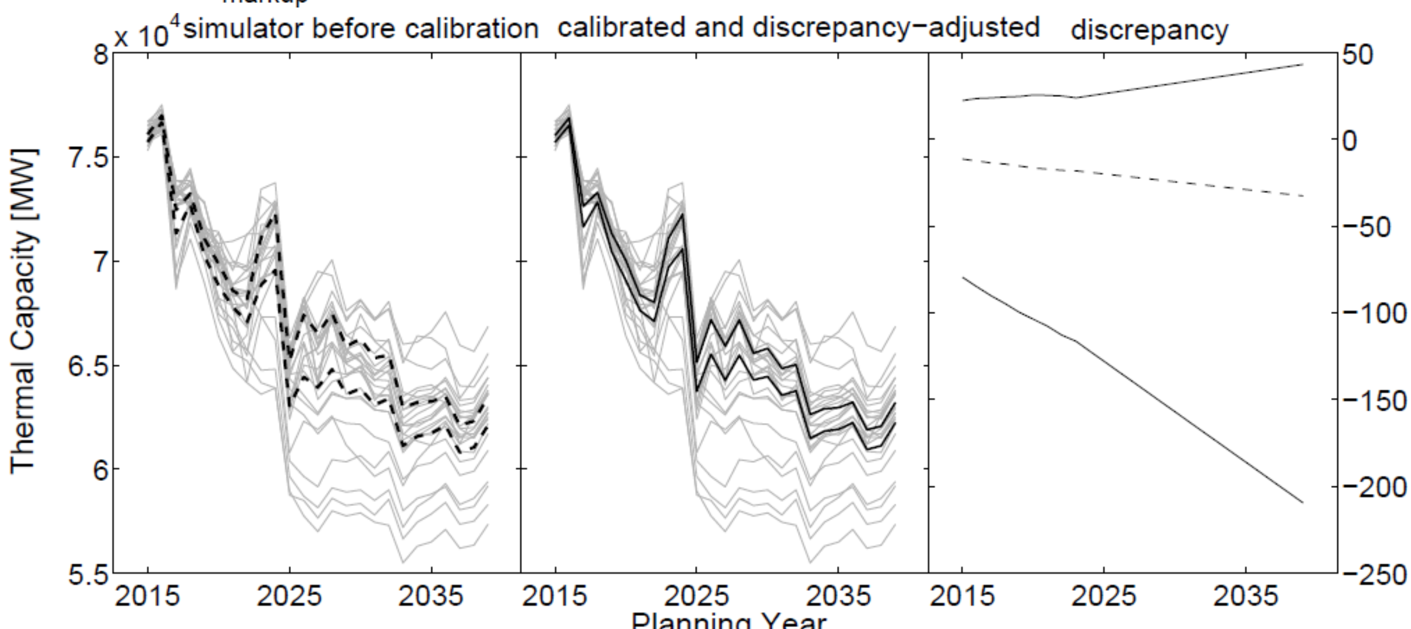
Durham University

Durham Energy Institute

Generation investment projection (Xu)



- Class of model heavily used in EMR
  - Example of calibration of simple model (after Eager et al)
  - Requires emulation due to substantial run time
  - Current work with DECC on model of full complexity, V high dimensional inputs



Durham University

Durham Energy Institute

Conclusions

- Mathematical/computer modelling is widely used in making capital and policy planning decisions
- If decisions are to be taken systematically on the basis of evidence from a computer model, uncertainty in the relationship between the model and the real world must be assessed
  - Similar calibration methods to "grey box modelling" (Henrik Madsen) except for computational complexity
  - Many applications including climate, oil reservoirs, cosmology, epidemiology
  - Not widespread in energy system planning and economics
  - Opportunities in energy systems integration where the scope of modelling (geographic, across sectors) is greater

Reading list

- Zachary, S. & Dent, C.J. (2014). Estimation of Joint Distribution of Demand and Available Renewables for Generation Adequacy Assessment. Available from http://dro.dur.ac.uk/cgi/search/author, search for author Dent
- math.nist.gov/IFIP-UQSC-2011/slides/Goldstein.pdf
- http://www.maths.bris.ac.uk/~mazjcr/climPolUnc.pdf
- http://www.significancemagazine.org/details/magazine/868617/Climate-change.html
- www.mucm.ac.uk (Managing Uncertainty in Complex Models project)