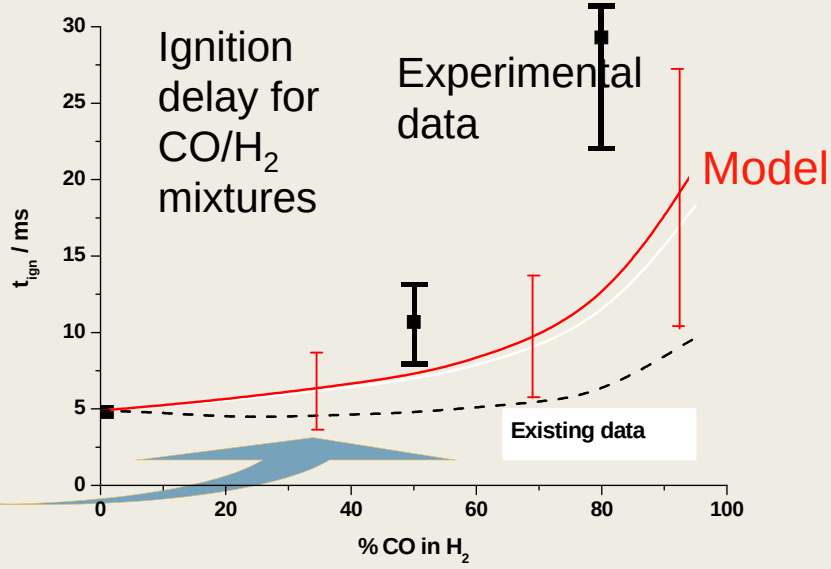
TUESDAY 2-2 (ALISON TOMLIN)

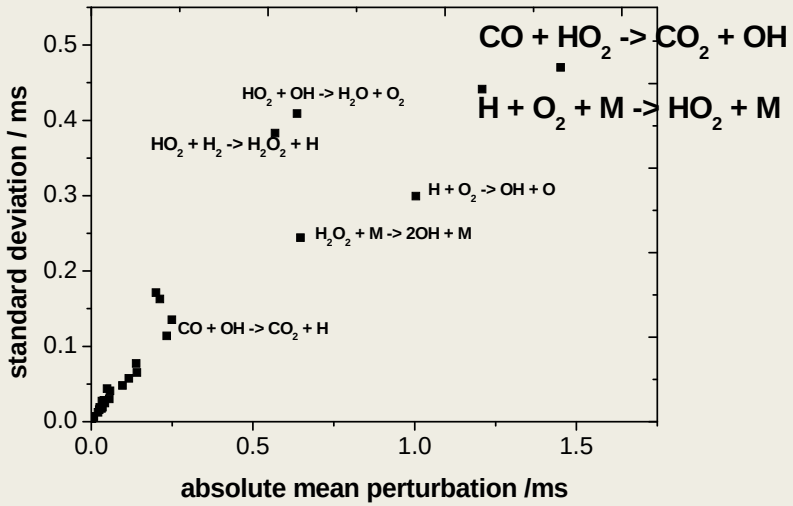
Global uncertainty analysis 1

Global uncertainty and sensitivity methods

Sensitivity and uncertainty analysis

- Uncertainty analysis (UA) estimates the overall predictive uncertainty of a model given the state/or lack of knowledge about its input parameters.
- UA puts error bars on predictions.



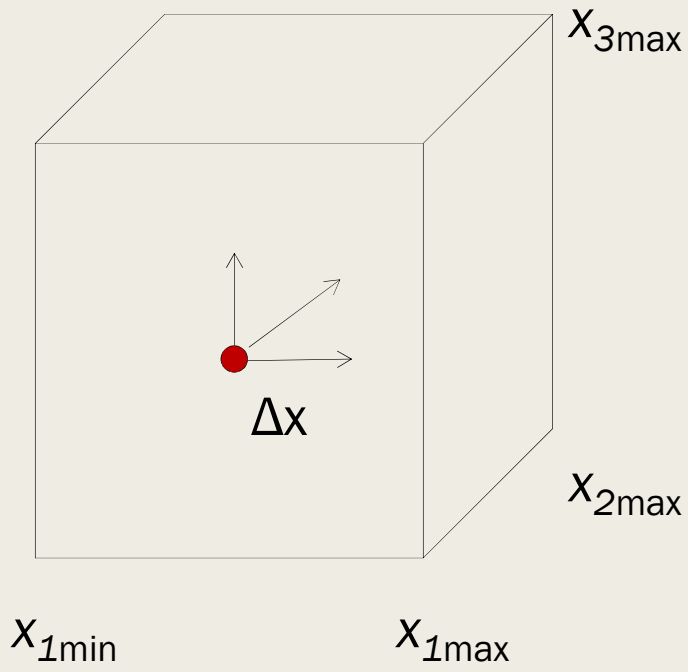


Sensitivity analysis (SA) determines how much each input parameter contributes to the output uncertainty (usually variance).

LOCAL VS GLOBAL

Local sensitivity coefficients

- nominal values



Local first-order sensitivity coefficient

$$S_{ij} = \frac{\partial Y_i}{\partial x_j}$$

Normalised first-order sensitivity coefficient

$$S'_{ij} = \frac{x_j}{Y_i} \frac{\partial Y_i}{\partial x_j}$$

Contributions to uncertainty?

• If $\sigma(x_j)$ are known then we can estimate overall uncertainty:

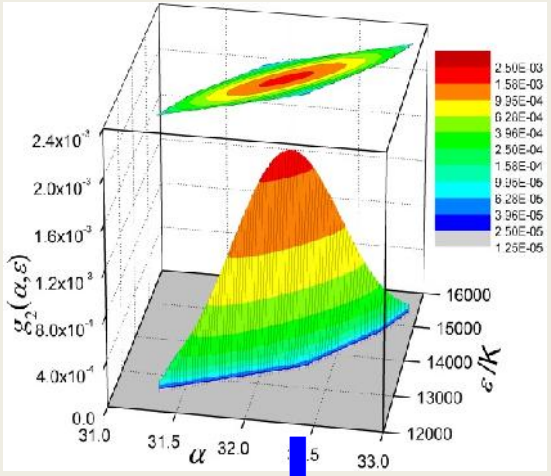
$$\pm^2(\mathbf{Y}) = \sum_j (\mathbf{S}'_j)^2 \frac{\pm^2(x_j)}{x^2_j}$$

• The **fractional contribution** of each parameter to this **uncertainty** can be estimated.

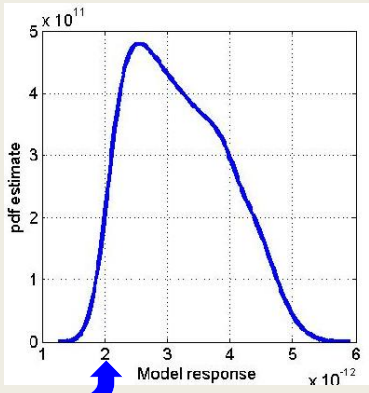
• Gives a better measure of parameter importance than S'_{ij} alone.

• Tells us how better quantification of each parameter could reduce overall modelling uncertainty.

Local vs global methods



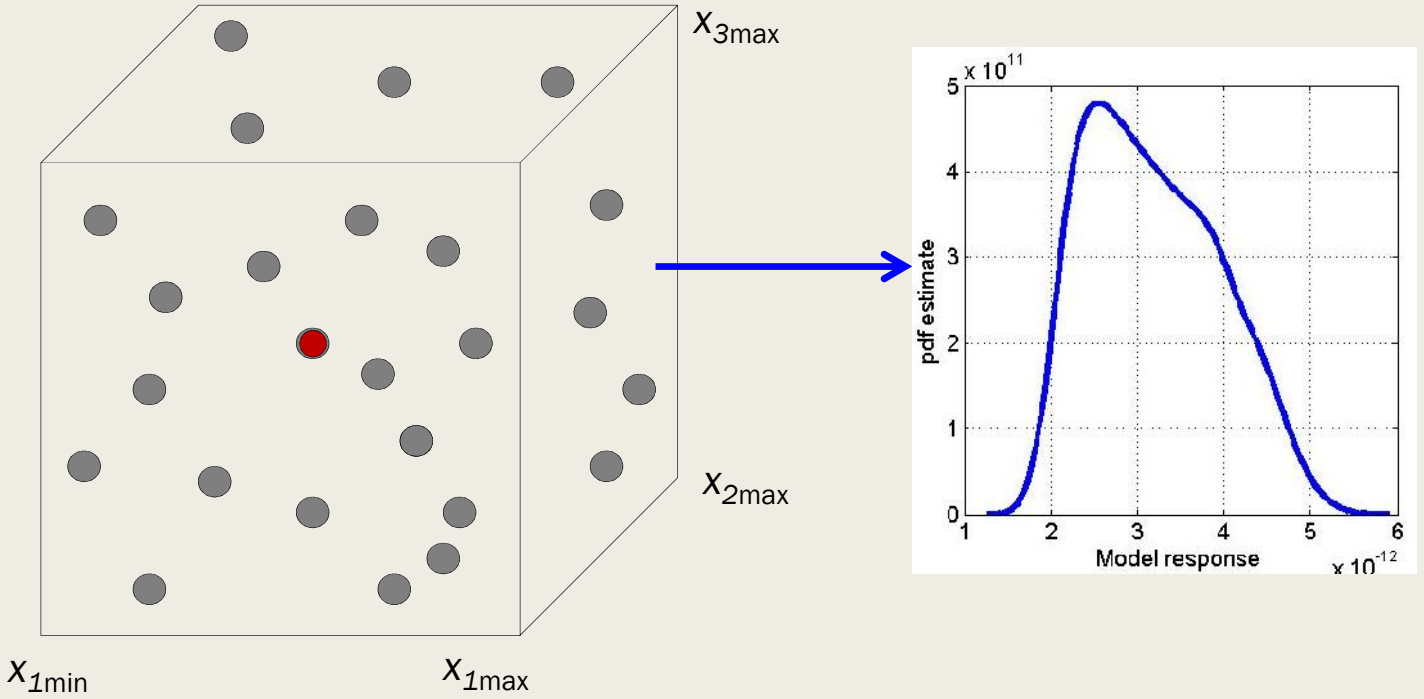
Probability distribution of predicted Y_i



Global sensitivity/uncertainty methods

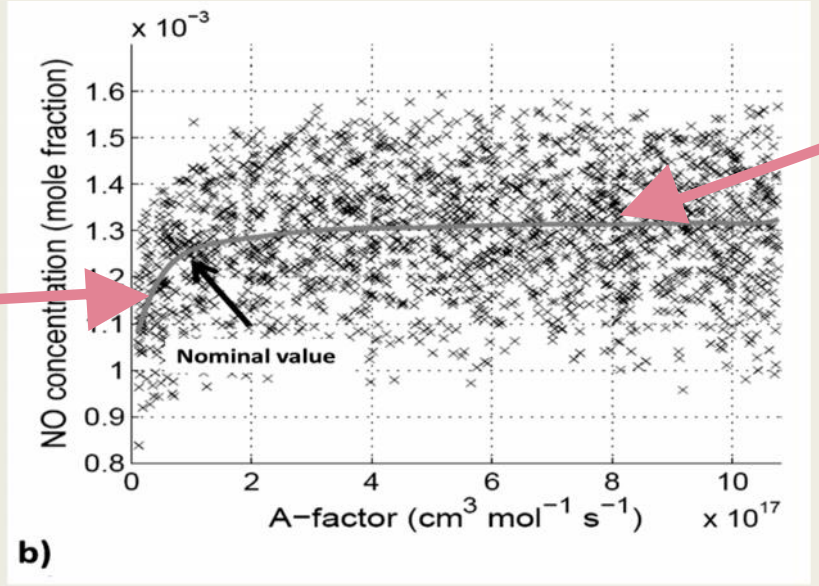
Global - attempts to cover whole input space using a sampling method.

Najm, Wang, Frenkach, Sheen, Tomlin, Turanyi etc.



Why use global methods?

- Local sensitivity and uncertainty methods are usually based on a single (best estimate) value of the parameters.
- If the sensitivity of the output changes depending on the values of the parameters then local methods could be inaccurate.
- Particularly important for non-linear models and models with large uncertainties.



High sensitivity

Low sensitivity

Disadvantages of global methods

- In order to cover the regions of parameter uncertainty, *sampling based methods* need to be used and therefore a *large number of model runs* is needed instead of the single run required for local sensitivity analysis using e.g. decoupled direct method.
- The methods also require prior knowledge of the input parameter distributions.
- Methods are then required to interpret the data from a large number of samples to determine the sensitivity indices (see next lecture).
- Whilst global methods based on sampling can be applied to large parameter systems the issue of sample sparsity can be an issue.
- Screening methods are therefore often first applied to identify unimportant parameters which do not need to be varied in the full global approach.

SCREENING METHODS

Why use screening methods?

- To identify parameters which do not contribute greatly to the output variance and therefore do not need to feature in a full global method.
- Still need to cover sufficiently wide area of input uncertainty space.
- Several choices:
  - *Most commonly used is **local sensitivity analysis**.*
  - ***Morris Method.** A one at a time method which attempts to screen the whole input uncertainty space. Can be expensive for large parameter systems.*

Local sensitivity analysis as screening method.

- Problematic unless it is applied at various values of the nominal parameters e.g. recommended value, recommended value x 2, recommended value x 0.5.
- Example: Ignition delay simulation problem.
  - *Computing sensitivities to ignition delays cannot simply be done using methods such as the decoupled direct method.*
  - *Brute force methods have to be used which require one run per parameter. EXPENSIVE!*
  - *Often, a surrogate, such as temperature at a given time point is used – which can be computed using a single time point in Chemkin.*
  - *Based on the assumption that reactions that lead to sensitivities in temperatures also lead to sensitivities in ignition delays.*

Morris “One at a time” Methods

The value of each parameter x_j is modified within the range $\left[x_j^{\min}, x_j^{\max}\right]$ by a fixed amount Δ that is determined in the following way. A vector

$$\left\{0, \frac{1}{q-1}, \frac{2}{q-1}, \frac{3}{q-1}, \ldots, 1\right\}$$

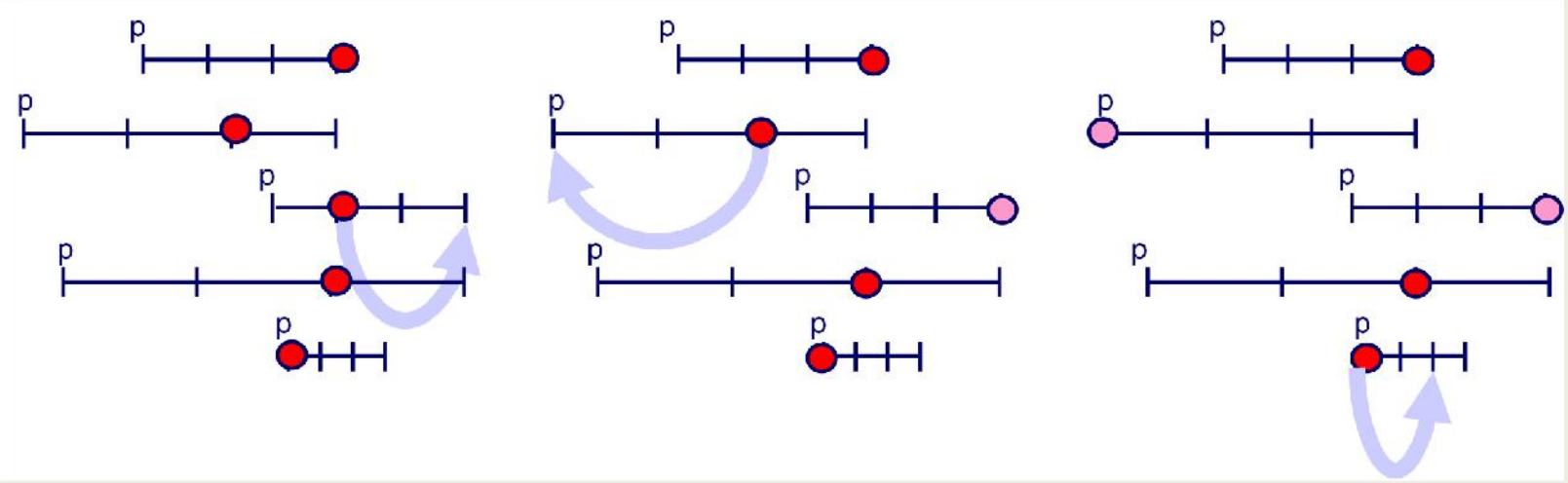
is generated using a small even number q selected by the user. Then, zero and one are assigned to x_j^{min} and x_j^{max} respectively. The other parameter values are scaled accordingly.

- The first parameter set is the selected randomly from this range.
- The next parameter set is identical to the previous one, except for the value of a single parameter which is changed randomly.
- In each run the parameters are ordered randomly and changed one at a time in this way until all the parameters have been changed once.

Morris “One at a time” Methods

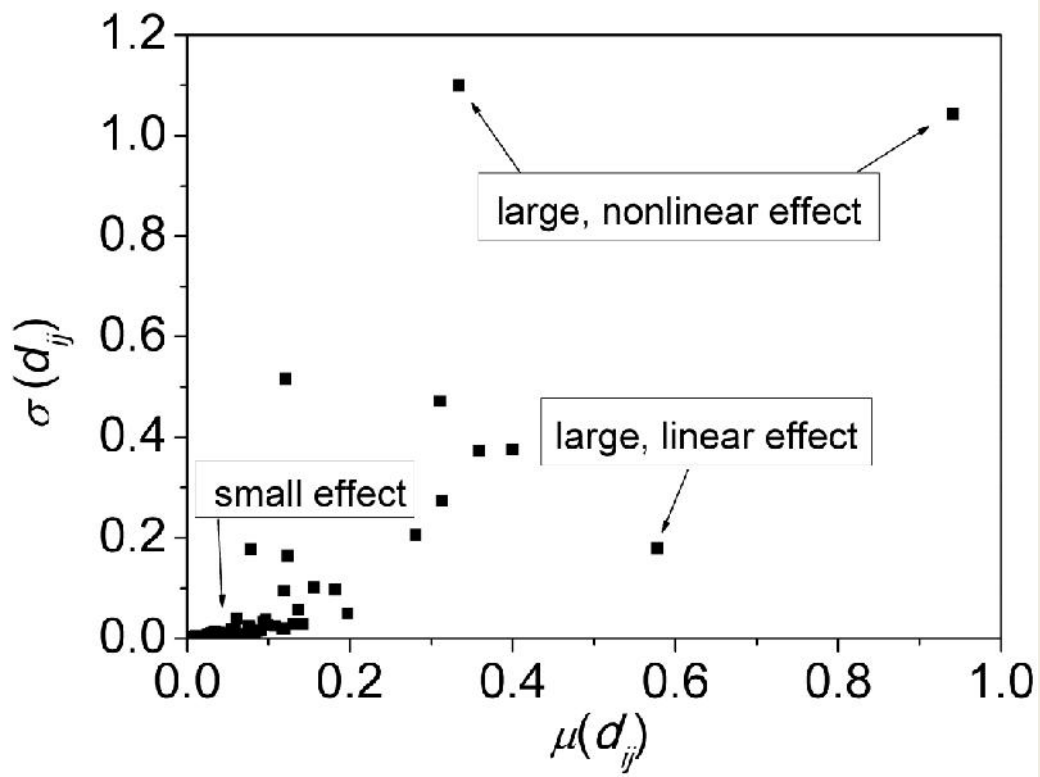
Several runs are performed, each starting with a new random set and parameter ordering and d_{ij} shows the effect of changing parameter x_j on model result Y_i at arbitrary values of all other parameters:

$$d_{ij} = \frac{Y_i\left(x_1^z, x_2^z, \ldots, x_j^z + \Delta, \ldots, x_m^z\right) - Y_i\left(\mathbf{x}^{z-1}\right)}{|\Delta|}$$

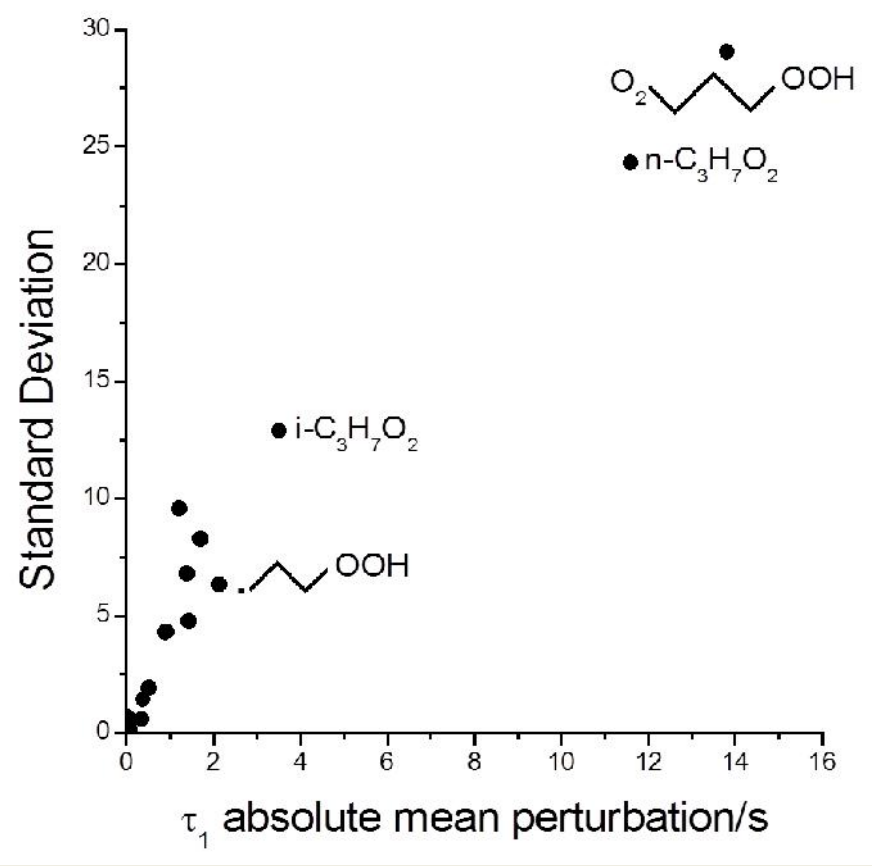


Morris methods

- Around r = 10-20 runs may be needed for the method to converge.
- Sampling effort required is therefore $r(m+1)$ where m is the number of parameters. Not a cheap method for large parameter systems!
- Statistical analysis of the d_{ij} values obtained gives the expected value $\mu(d_{ij})$ and variance $\pm(d_{ij})$ of changing parameter x_j on model result Y_i.



Examples of application



- Note the high standard deviation of the outputs compared to the mean.
- Very nonlinear responses requiring large sample size to converge.

Morris analysis for species ΔH_f^o with respect to time to cool flame for propane oxidation. T= 593 K, equimolar C_3H_8+O_2 at 53.4 kPa, diluted by N_2 to 101.3 kPa (Hughes et al. 2006)

Advantages and disadvantages of Morris methods

+ ve

- Represent the whole parameter space rather than a nominal value point like local methods.
- Can highlight nonlinearities in response.
- Give a clear visual presentation of important and unimportant parameters.

- ve

- Can be computationally expensive for large parameter, nonlinear systems.
- Do not highlight the influence of parameter interactions – only general nonlinearities.
- Do no explicitly provide sensitivity indices.

FULLY GLOBAL SAMPLING METHODS

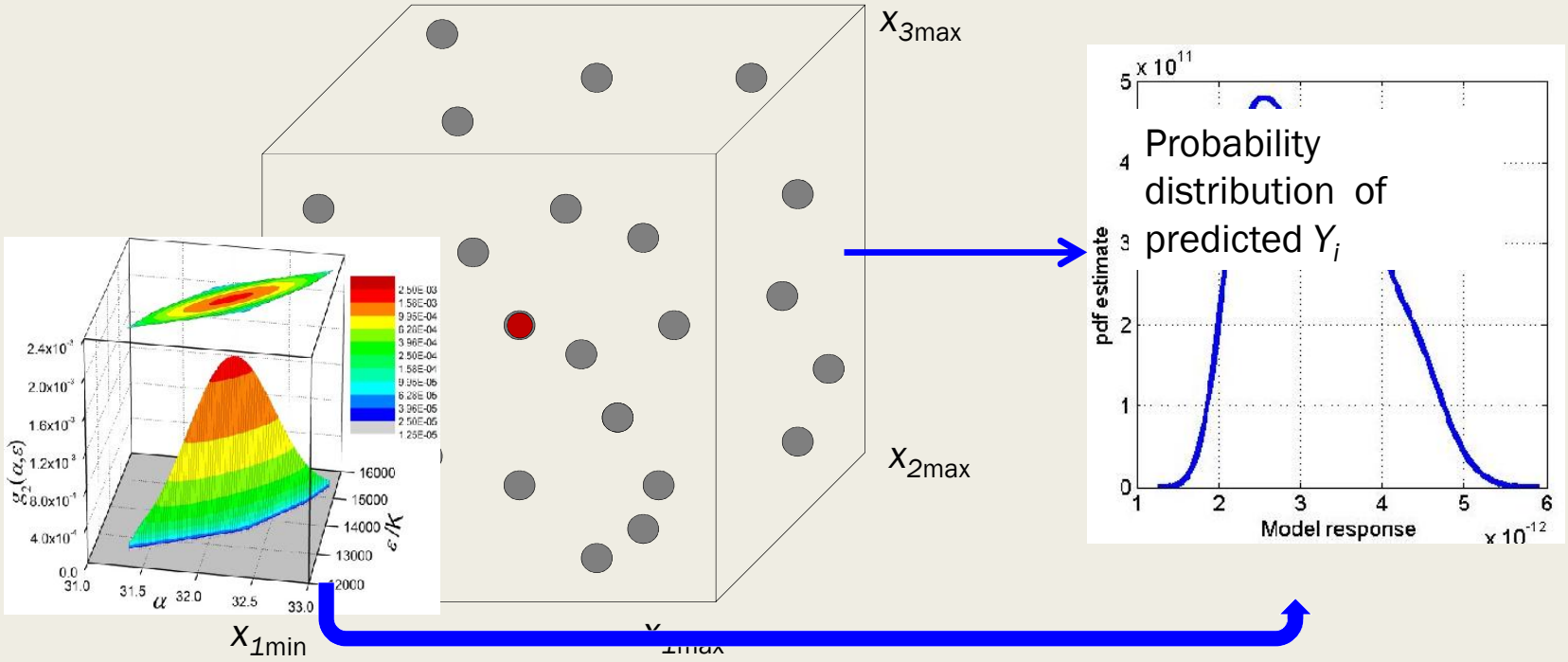
Are they necessary?

- Both local sensitivity methods and global screening methods can provide useful information on parameter importance in models.
- For large systems, as a first cut they may be all that can be afforded computationally.
- However, to obtain a full picture of the model output distributions global sampling methods are required.
- These sample the *full input space* (uniform or probabilistic) and thus provide a clear picture of the output distributions.
- Since a full model run has to be simulated for each parameter set in the sample, we need a sampling method that *converges* as quickly as possible.

Global sensitivity/uncertainty methods

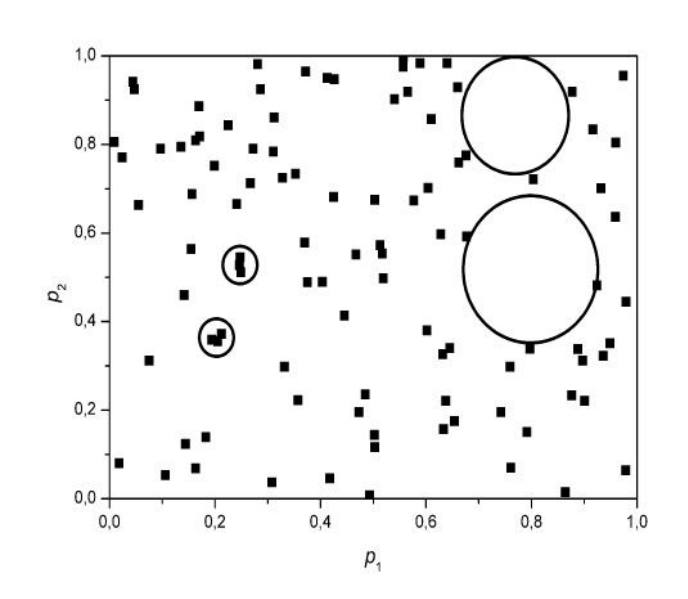
Global - attempts to cover whole input space using a sampling method.

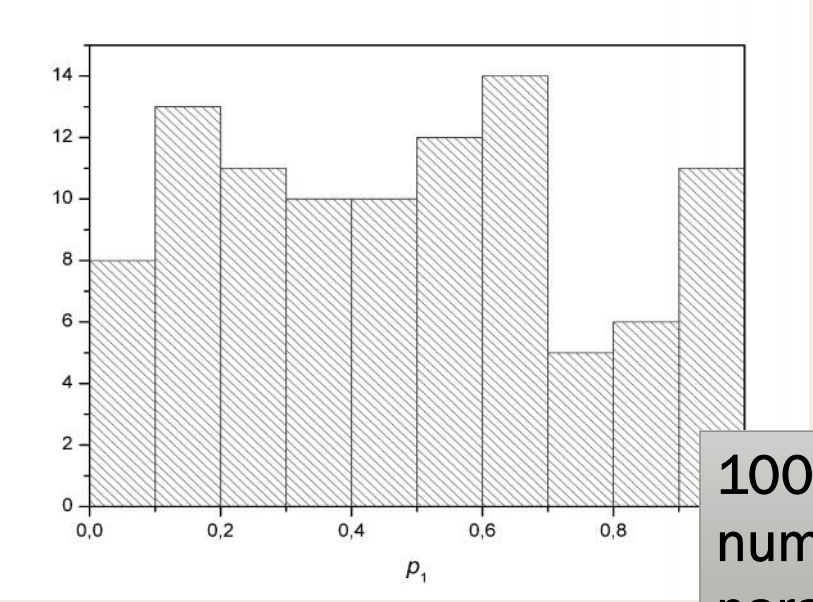
Najm, Wang, Frenkach, Sheen, Tomlin, Turanyi etc.



Monte Carlo random sampling

- For a global sampling method it is important to get good coverage of the input parameter space – which may be high dimensional.
- Typical random sampling methods can lead to clustering and holes.

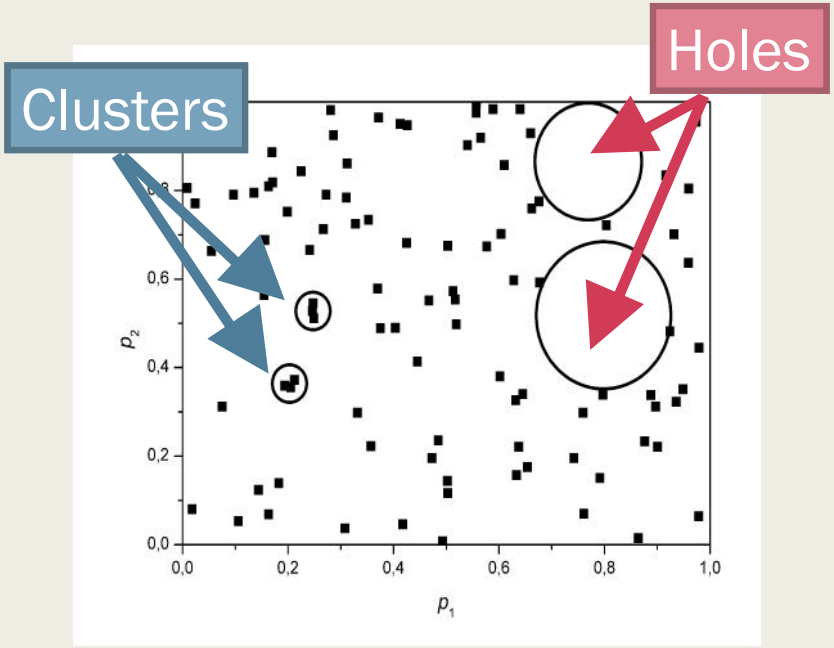


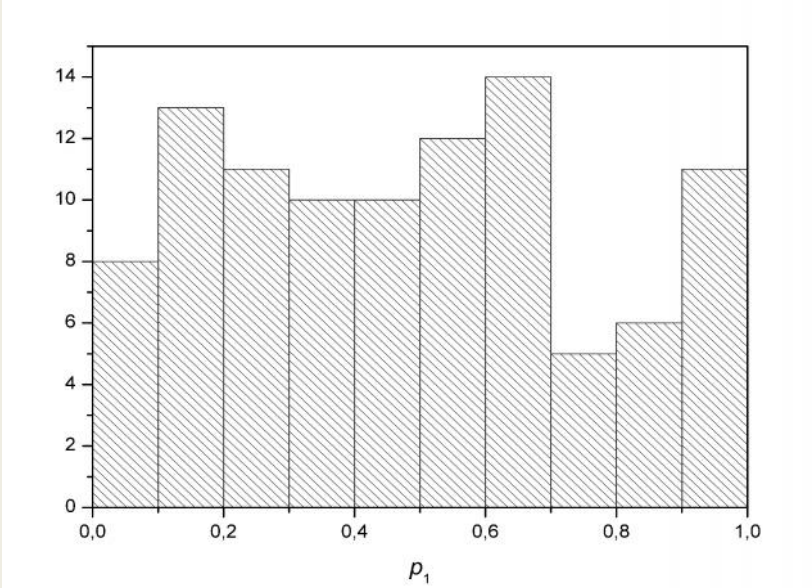


100 numbers – 2 parameters

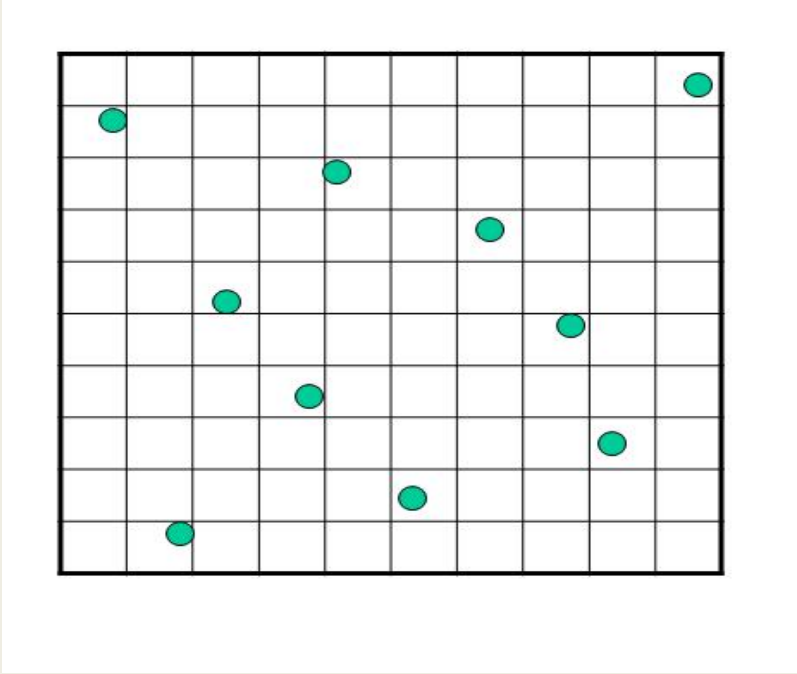
Monte Carlo random sampling

- For a global sampling method it is important to get good coverage of the input parameter space – which may be high dimensional.
- Typical random sampling methods can lead to clustering and holes.





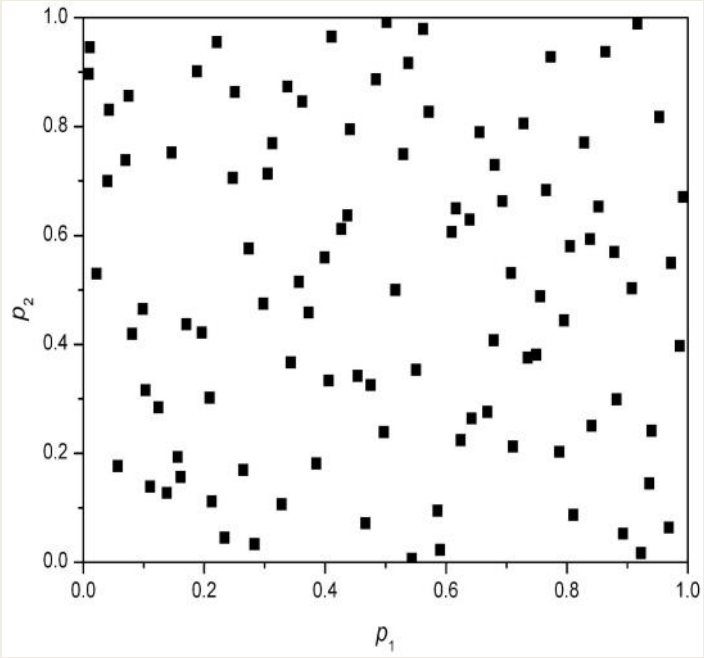
Structured sampling: Latin Hypercube

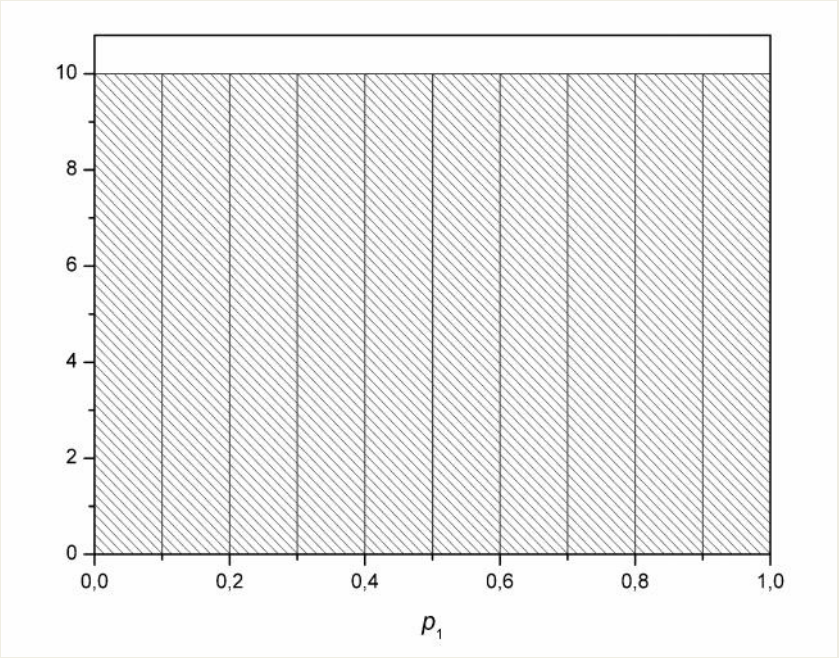


- Points generated by Latin hypercube sampling according to a uniform distribution.
- Each horizontal and vertical stratus contains a single point, while the location of the point is random in the corresponding small square.

Structured sampling

- Latin Hypercube, 100 points, 2 parameters.
- More even sample.





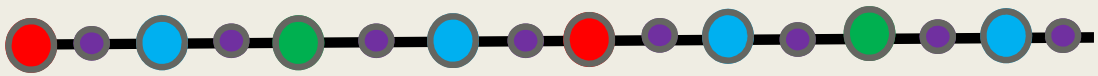
Low discrepancy sequences

- Latin Hypercube sampling can become expensive as the number of parameters increases – unless the resolution of the grid is sacrificed.
- Low discrepancy sequences offer the best distributed sample for a given sample size and parameter space dimensionality.
- Successive sample points are added to positions as far away as possible from existing sample points so that clustering can be avoided.
- The best known low discrepancy sequences include those of Halton (1960), Faure (1992), Sobol' (1967) and Niederreiter (1988).
- The Halton sequence is based on numbers generated within different power sequences for each dimension of the parameter space.
- The Sobol sequence uses only base 2.

Sobol sequence

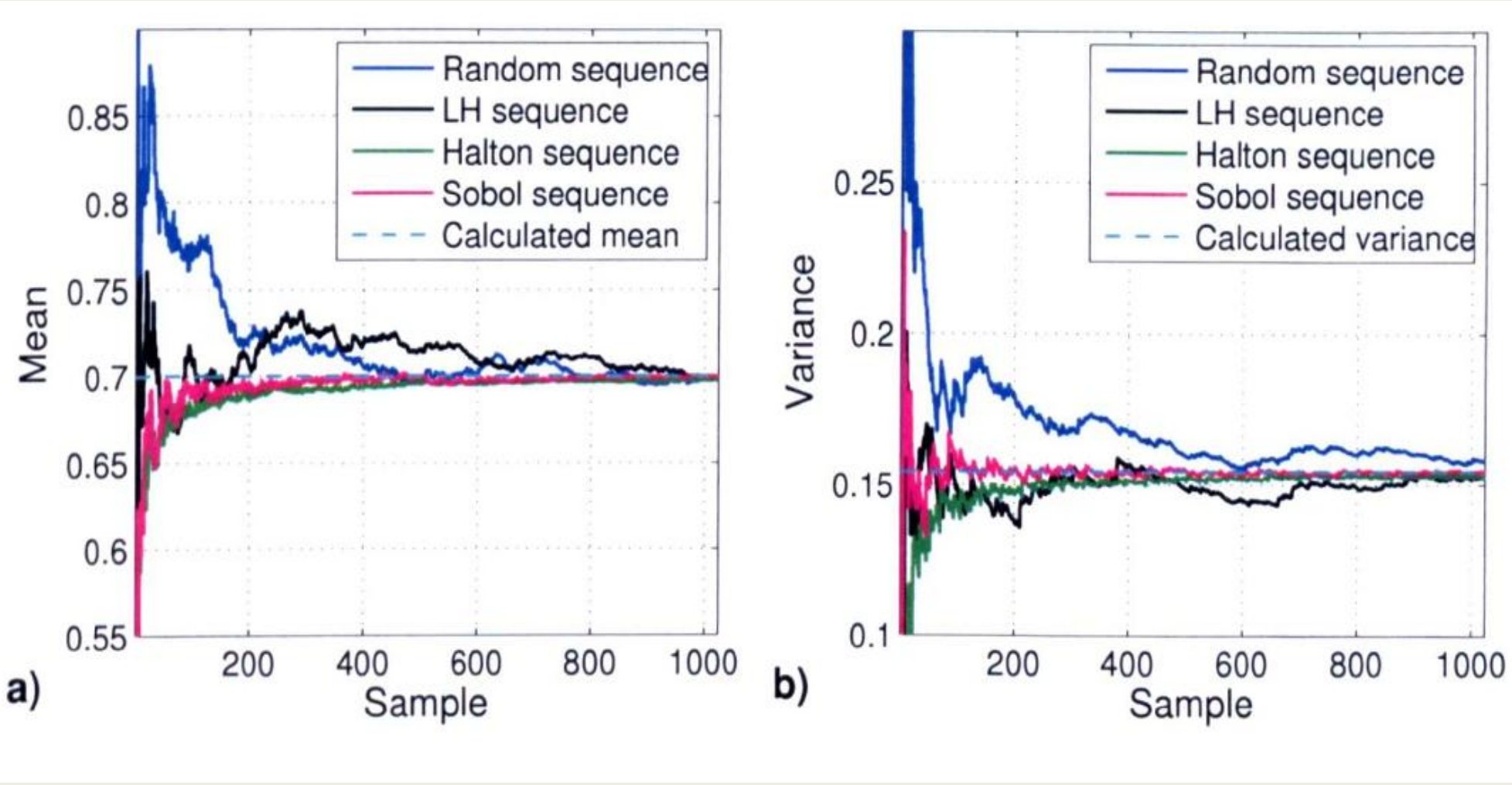
Sobol sequences use a base of two to form successively finer uniform partitions of the unit interval and then reorder the coordinates in each dimension.

0.000e+00	0.000e+00	0.000e+00
5.000e-01	5.000e-01	5.000e-01
7.500e-01	2.500e-01	7.500e-01
2.500e-01	7.500e-01	2.500e-01
3.750e-01	3.750e-01	6.250e-01
8.750e-01	8.750e-01	1.250e-01
6.250e-01	1.250e-01	3.750e-01
1.250e-01	6.250e-01	8.750e-01
1.875e-01	3.125e-01	3.125e-01
6.875e-01	8.125e-01	8.125e-01
9.375e-01	6.250e-02	5.625e-01
4.375e-01	5.625e-01	6.250e-02
3.125e-01	1.875e-01	9.375e-01
8.125e-01	6.875e-01	4.375e-01
5.625e-01	4.375e-01	1.875e-01
6.250e-02	9.375e-01	6.875e-01

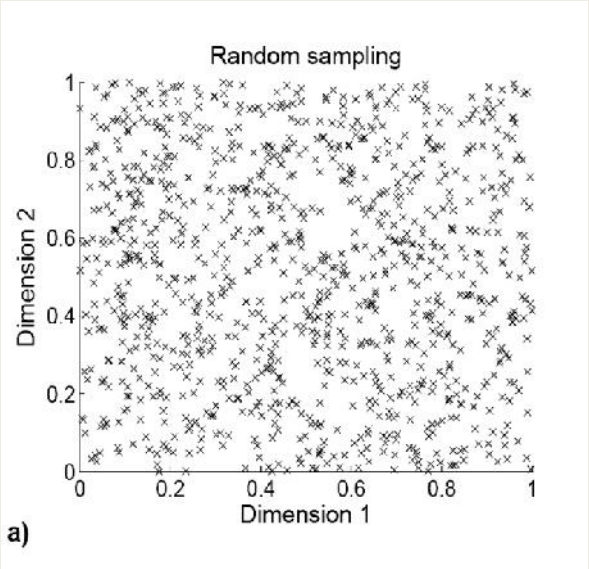


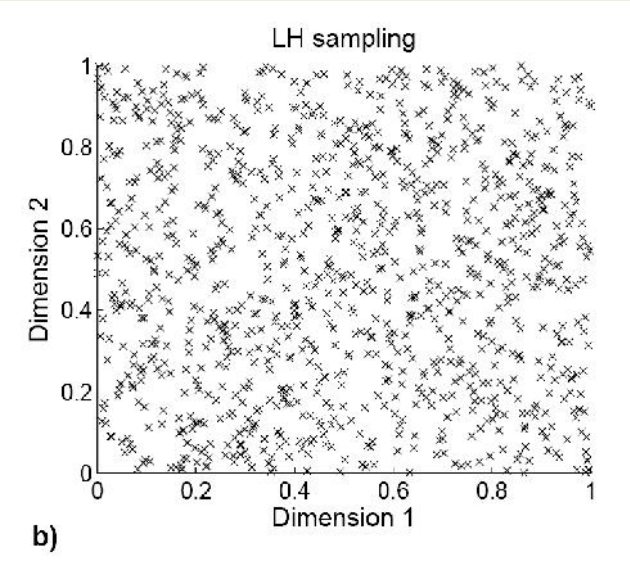
The Sobol sequence is designed to have the best convergence properties and hence can lead to savings in sampling based sensitivity and uncertainty analysis because smaller sample sizes are needed to get equivalent accuracy in the results.

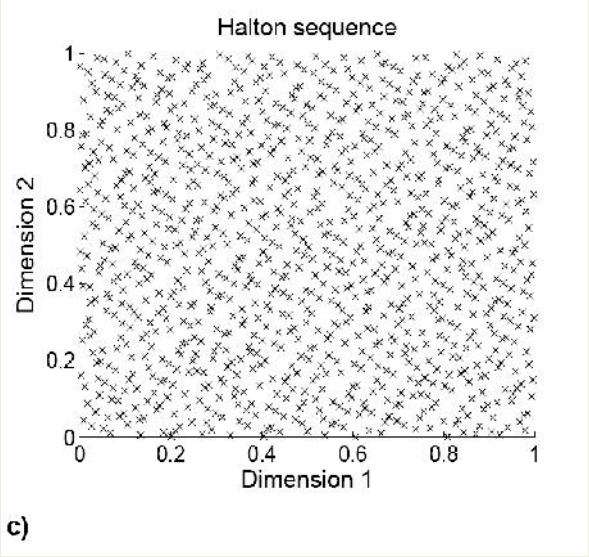
Comparison of convergence properties of different sampling strategies for a simple test model: $f(x) = x_1 + x_2^4$

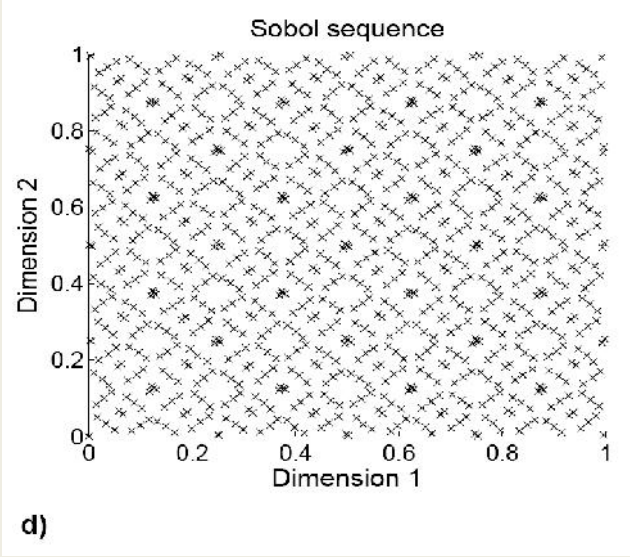


A comparison of samples produced by different sampling methods for a 2 parameter model 1024 sampling points





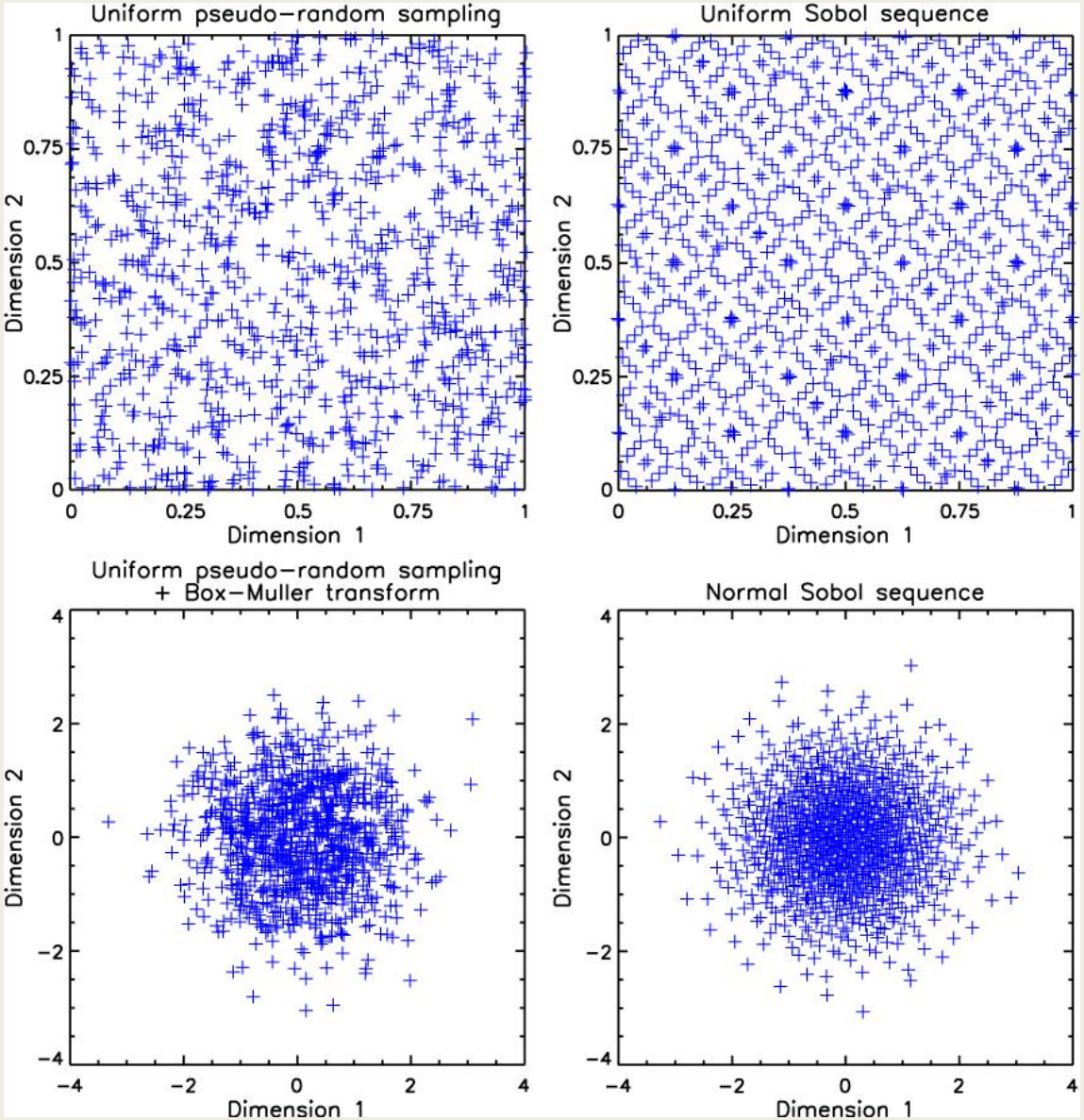




Probabilistic sampling

- If probabilistic information is known about the input parameters then we may wish to sample from this distribution e.g. *a normal distribution* based on 2σ uncertainties.
- A normal distribution of random numbers can be obtained from a uniform distribution of random numbers using the Box-Muller algorithm.
- However, a better method (Hebrard et al., 2015) may be to compute directly the *inverse normal distribution of the Sobol sequence* given its cumulative distribution function. This way we take advantage of the convergence properties of the quasi-random sample.

- 2-parameter samples, N = 1000.
- Uniform pseudo- random sample (top left)
- Sobol's quasi-random sequence sample (top right)
- Box-Muller transformation applied to an uniform pseudo-random sample (bottom left)
- Normal inverse cumulative function of a Sobol's quasi-random sequence sample (bottom right).

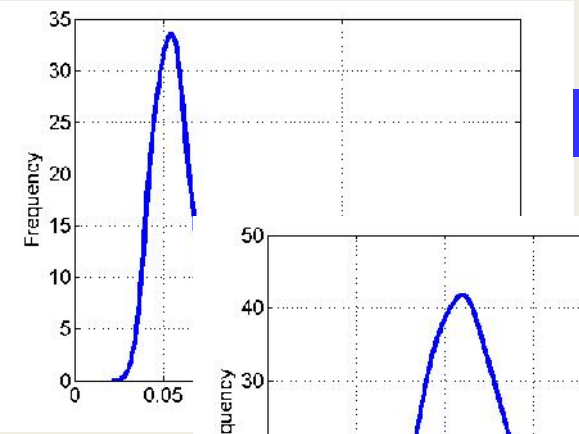


What parameters to include?

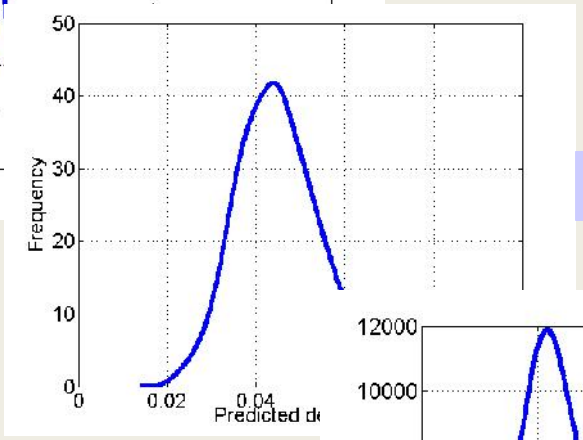
- In an ideal world:
  - *All Arrhenius parameters*
  - *Thermodynamic parameters which are used to calculate reverse reaction rates.*
  - *Species transport data*
  - *Other potential model errors*
    - Temperature profile
    - Heat transfer coefficients
    - Residence times
    - Loss rates to the walls of the reactor vessel
- In reality many of these are often ignored and a most common approach is to simply look at the A-factors for each forward reaction.
  - *Tells us something about the important reactions but does not give a full picture of uncertainties.*

INTERPRETING OUTPUT DISTRIBUTIONS

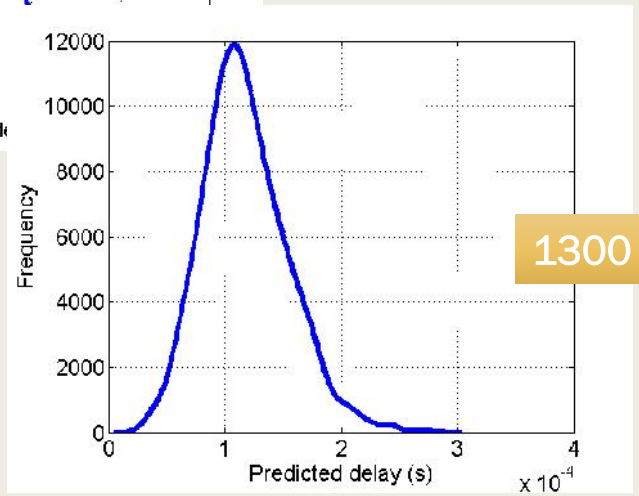
Ignition delays:
Predicted output distributions (butane model)



700 K – narrow distribution but long tail



900 K - factor of 2 uncertainty

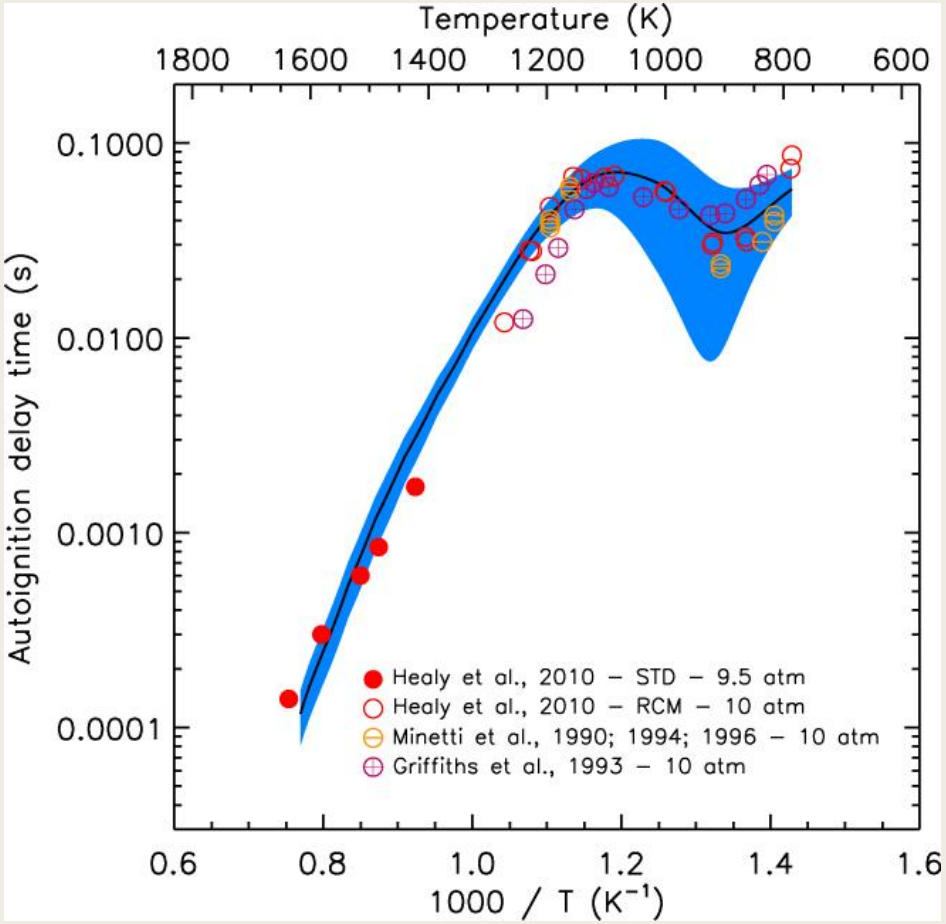


1300 K – narrows again

- 1σ used to represent output error bars.

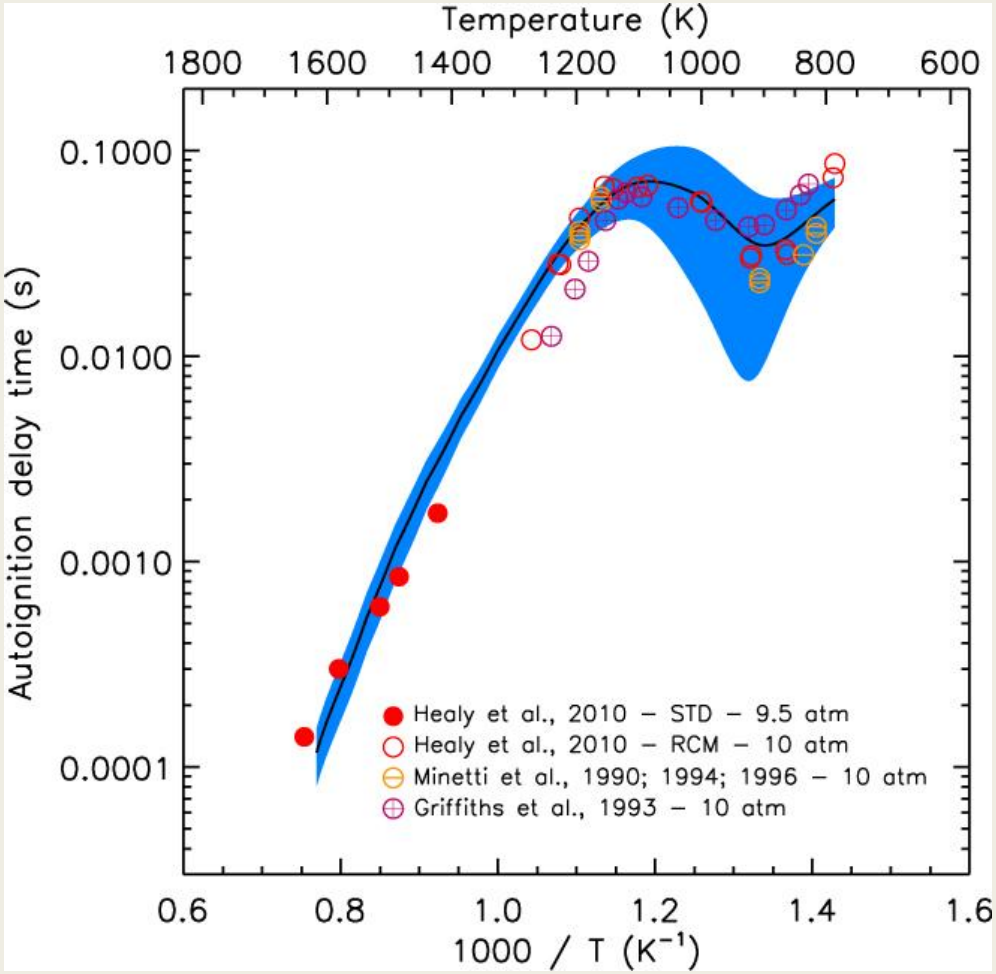
Interpreting output distributions

- Example from simulations of ignition delay time for a butane oxidation system.
- The blue shaded area represents 1σ of the outputs based on a sampled normal distribution of the input rate parameters.
- Hebrard et al. (2015)



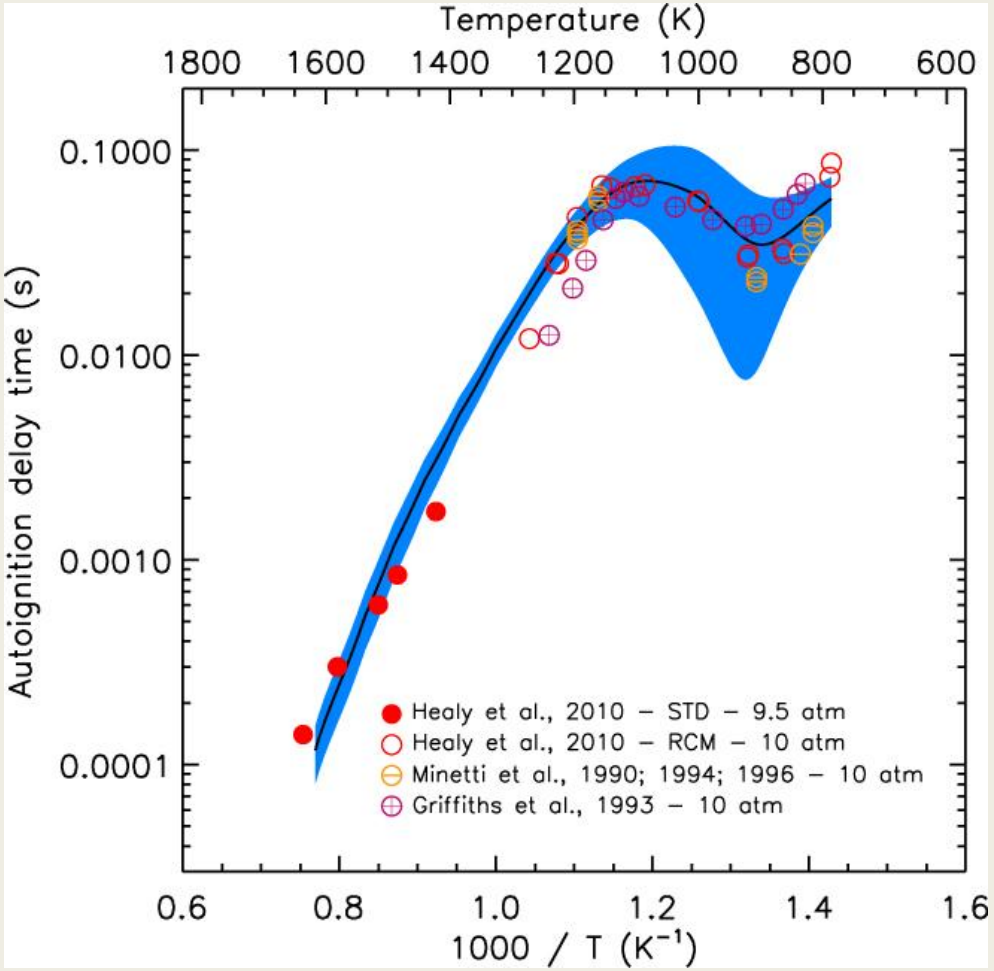
Interpreting output distributions

Reasonable agreement between model and shock tube and RCM data if uncertainties are taken into account.



Interpreting output distributions

- Lower uncertainties in high temperature region.
- Higher uncertainties in NTC region.

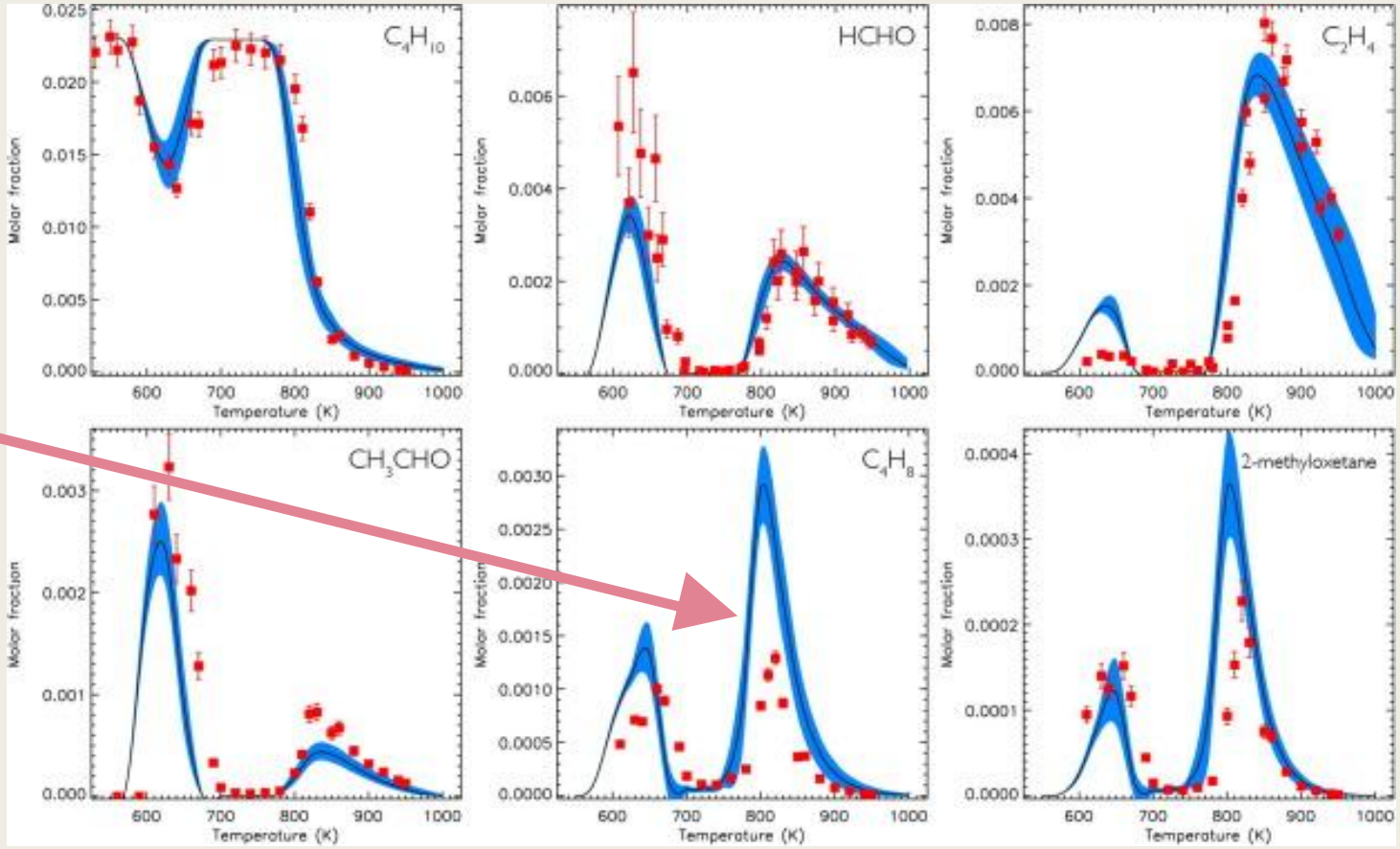


JSR data

Some discrepancies between model and experimental data even when accounting for estimated uncertainties.

Missing reaction steps?

Other uncertainties not identified?



Summary of lecture 2-2

- If we fully account for input parameter uncertainties we end up with an output distribution rather than a single value.
- Most methods represent this with some kind of variance based measure.
  - *Can be misleading if the distribution is significantly non-Gaussian.*
- Only a global sampling approach will give full probability distribution.
  - *Local methods give an estimate but only based around the nominal parameter value and this can be inaccurate if the real response is non-linear.*
- Low-discrepancy sequences have been shown to have better convergence properties compared to standard random Monte Carlo sampling and Latin Hypercube approaches.
  - *This reduces the sample size needed to estimate output variance.*
- For fuel oxidation models the Negative Temperature Coefficient regime is seen to be the most uncertain in many cases (less measurements here and more reactions contributing).