Energy and Resources Research Institute

Faculty of Engineering

UNIVERSITY OF LEEDS

THE ROLE OF SENSITIVITY ANALYSIS IN MODEL IMPROVEMENT

Alison S. Tomlin

Michael Davis, Rex Skodje, Frédérique Battin-LeClerc, Maximilien Cord

Use of complex kinetic mechanisms

UNIVERSITY OF LEEDS

• Many examples of areas where complex kinetic mechanisms are used in engineering and environmental design and control:

design of efficient, clean combustion devices

safety applications for range of fuels and hydrocarbons

atmospheric response to pollution control measures

systems biology

drug design

• In practical applications, complex kinetics linked to detailed models of fluid flow and other physical processes.

Development of complex kinetic mechanisms

UNIVERSITY OF LEEDS

Complex chemical mechanisms built by:

1. proposing a set of rules for the interaction between species – **mechanism protocol**
2. developing **effective parameterisations** for the kinetics described within the mechanism.

- Our ability to specify protocols is well developed in e.g. hydrocarbon oxidation.
- Large comprehensive mechanisms e.g. biodiesel surrogate methyl decanoate: 3012 species and 8820 reactions (Herbinet et al., 2008).

That's a lot of parameters!

- Many have to be estimated using rules related to chemical structure.
- **Does this lead to a robust mechanism? How can we check?**

Evaluation of kinetic mechanisms

UNIVERSITY OF LEEDS

- Comparison of model with experiment for simple to complex scenarios.

Then what?

- Agreement for the right reasons? Confidence in simulations?
- If discrepancies, then how do we find the contributing causes?
- Sensitivity and uncertainty analysis can help to answer these questions.

BUT!

- Do we have enough fundamental experiments to cover the conditions experienced in practical devices for all fuels of interest?
- Raises questions for the optimisation of complex fuel mechanisms – ill conditioned problem for larger fuels.
- Can carry out sensitivity analysis over any modelled conditions.

Need strong feedback loop between model evaluation and methods for model improvement.

Typical methodology (?)

UNIVERSITY OF LEEDS

- Develop mechanism using protocols.
- Compare mechanism against experiments for key targets such as ignition delays, species profiles, flame speeds, etc.
- Maybe using local sensitivity analysis try tweaking some parameters to improve fit with experiment OR optimise against target data sets (much less common).
- Publish comparisons and mechanism (fully documented??)
- Linear sensitivities can certainly be useful but if simulation time was not an issue we could do much better by exploring the full feasible input space for parameters.

Which may of course be huge!

Screening methods

UNIVERSITY OF LEEDS

• Methods such as linear sensitivity analysis or the global Morris method can be used for screening out unimportant parameters before more complex global sensitivity methods are used.

• Often the parameter space to be investigated is enormous:

- large no. of parameters *n*
- large uncertainty ranges.

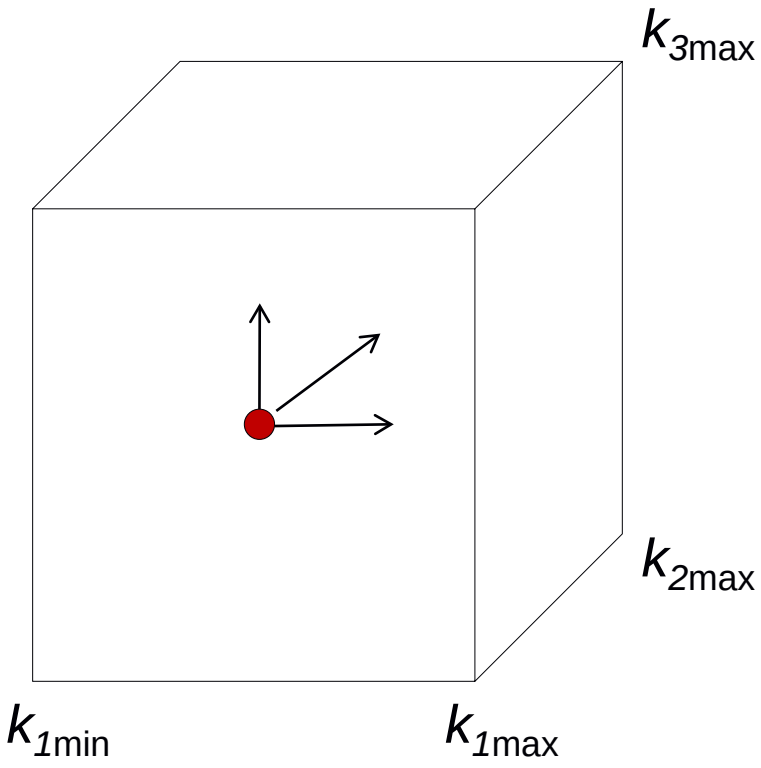
• In a **linear brute force** method each parameter is changed in turn by a small amount (5-25%) and the model response recorded.
• The parameters are then ranked according to effects on the model response.
• **Global screening** explores wider input space requiring more runs.

Comparison of sampling methods for 3 parameter system

UNIVERSITY OF LEEDS

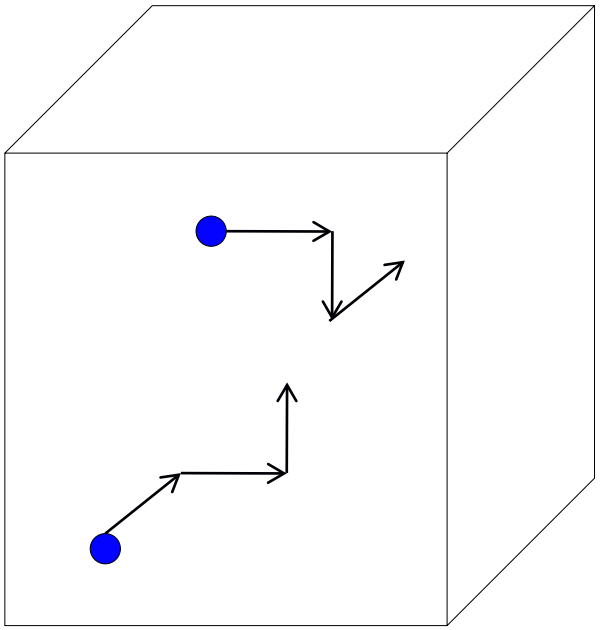
Local (● nominal values)

Cost 1 OR N_p



Morris global screening (two trajectories *r*)

Cost $(N_p+1)\ x\ r$



etc...

Assessing results (Morris)

UNIVERSITY OF LEEDS

Elementary effect of parameter k_j on variable c_i given by:

$$d_{ij}(\mathrm{k}) = \frac{c_i(k_1,\ldots,k_{j-1},k_j \pm \Delta,k_{j+1},\ldots,k_m) - c_i(\mathrm{k})}{\Delta}$$

Mean effect of factor k_j on variable c_i :

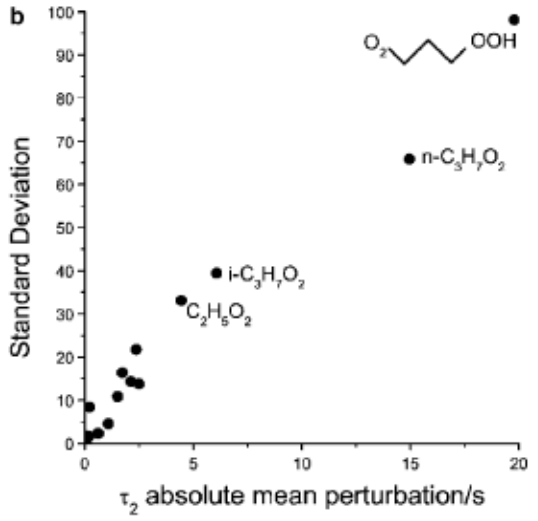
$$\overline{d}_{ij} = \frac{\sum_{l=1}^{r} \left| d_{ij}^{l} \right|}{r}$$

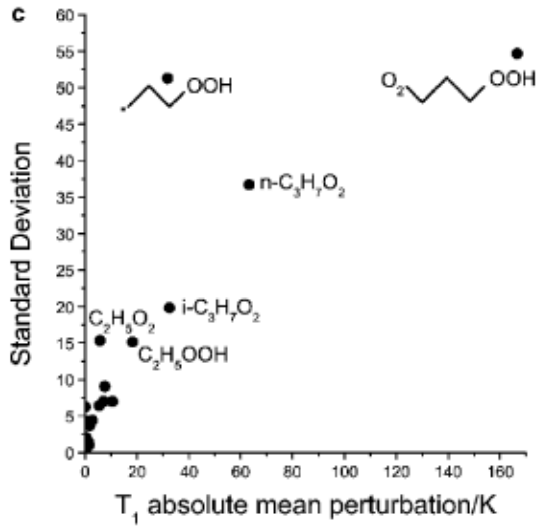
Variance of effect:

$$\sigma^2(d_{ij}) = \frac{r\sum_{l=1}^{r}(d_{ij}^{l})^2 - \left(\sum_{l=1}^{r} d_{ij}^{l}\right)^2}{r(r-1)}$$

Example from propane ignition study: (Hughes PCCP 2006, 593K 101.3kPa)

UNIVERSITY OF LEEDS





Response of time to ignition and cool flame temperature to changes in enthalpy of formation of species.

Larger absolute mean – larger effect.

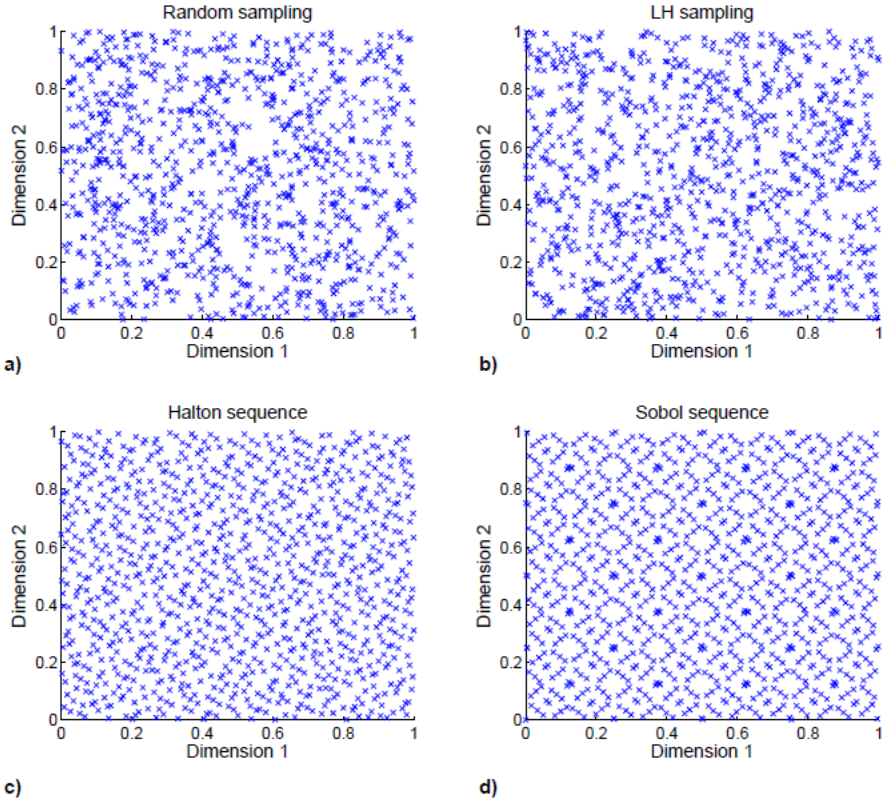
Larger standard deviation – larger nonlinear/interactive effects.

Sampling based methods

UNIVERSITY OF LEEDS

- Conceptually straightforward.
- Based on random or quasi random sampling of input parameter space.

- Perform many simulations until output mean/variances converge.
- No. of necessary runs depends on number of **important** parameters.
- Unlike Morris, MC methods may not increase in cost with input space dimension.
- Cost may still be prohibitive especially if interactive effects between parameters are present.



How do we deal with the tyranny of parameters?

UNIVERSITY OF LEEDS

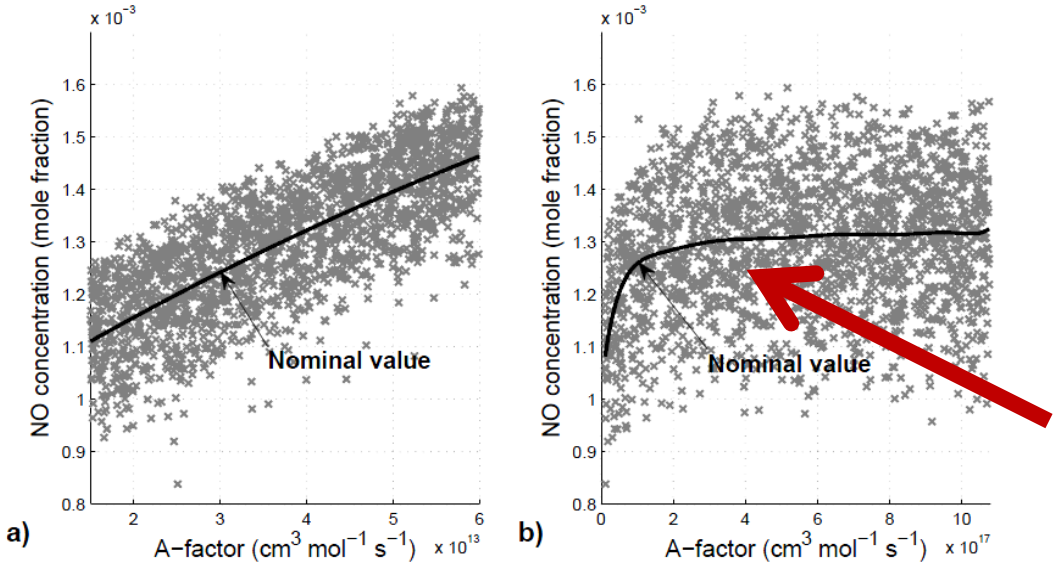
- Choice of sensitivity/uncertainty methods:

Partial derivative - linear	Brute force linear	Global Screening	Full Global
Cheap ↑	~ N_p ↕	Expensive ↓	Expensive ↓
Restricted to chosen values ↓	Restricted to chosen values ↓	Explores full input space ↑	Explores full input space ↑
No interactions ↓	No interactions ↓	Non-linear but no interactions	Parameter interactions ↑
Not always directly related to targets ↓	Relates to targets ↑	Relates to targets ↑	Relates to targets ↑

Monte Carlo (MC) simulations

UNIVERSITY OF LEEDS

- Interpretation of results difficult for large input space.
- Scatter plots used for each parameter to see overall effect.
- Large scatter often obscures mean effect of individual parameter.
- Linear effects can be shown using Pearson correlations, non-linear effects using rank correlation (Spearman correlations).
- Calculation of full sensitivity coefficients **VERY** expensive!



Example from flame calculation: NOx prediction.

Highly nonlinear

High Dimensional Model Representations (HDMR)

UNIVERSITY OF LEEDS

- Developed to reduce the sampling effort required for full global analysis.
- Output is expressed as a finite hierarchical function expansion:

$$f(\mathbf{x}) \equiv f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{1 \le i \le j \le n} f_{ij}(x_i, x_j) + \ldots + f_{12\ldots n}(x_1, x_2, \ldots, x_n)$$

- Usually second-order expression provides satisfactory results.
- Model replacement built using quasi random sample and approximation of component functions by orthonormal polynomials.
- Model replacement can be used to generate full Monte Carlo statistics.
- 1st & 2nd order sensitivity indices easily calculated from polynomial coefficients.

Required sample size determined by accuracy of model fit.

Requirements of the method

UNIVERSITY OF LEEDS

- Feasible input ranges for the parameters under investigation.

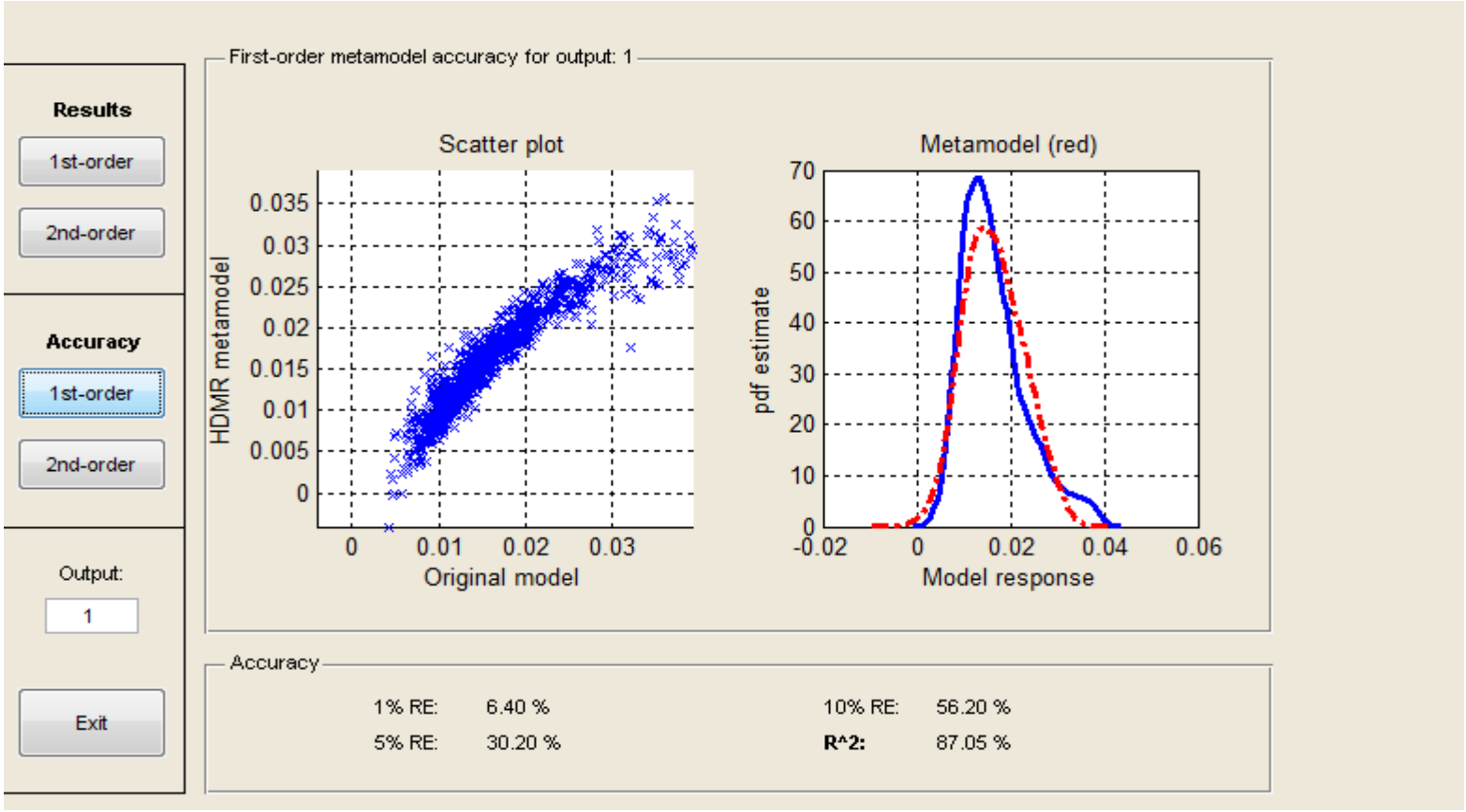
Can these be provided with the mechanisms?

- Understanding of correlations – transformations have to be made to deal with these.
- Quasi-random number sequence.
- Model simulations over the quasi-random sample.
- Usually boot-strap until simulated target output distribution and sensitivity coefficients converge.
- Higher order terms usually require much bigger sample size.

HDMR fit is usually quick – simulations may not be.

Examples from HDMR code: butane mole frac in JSR: 750 K.

UNIVERSITY OF LEEDS

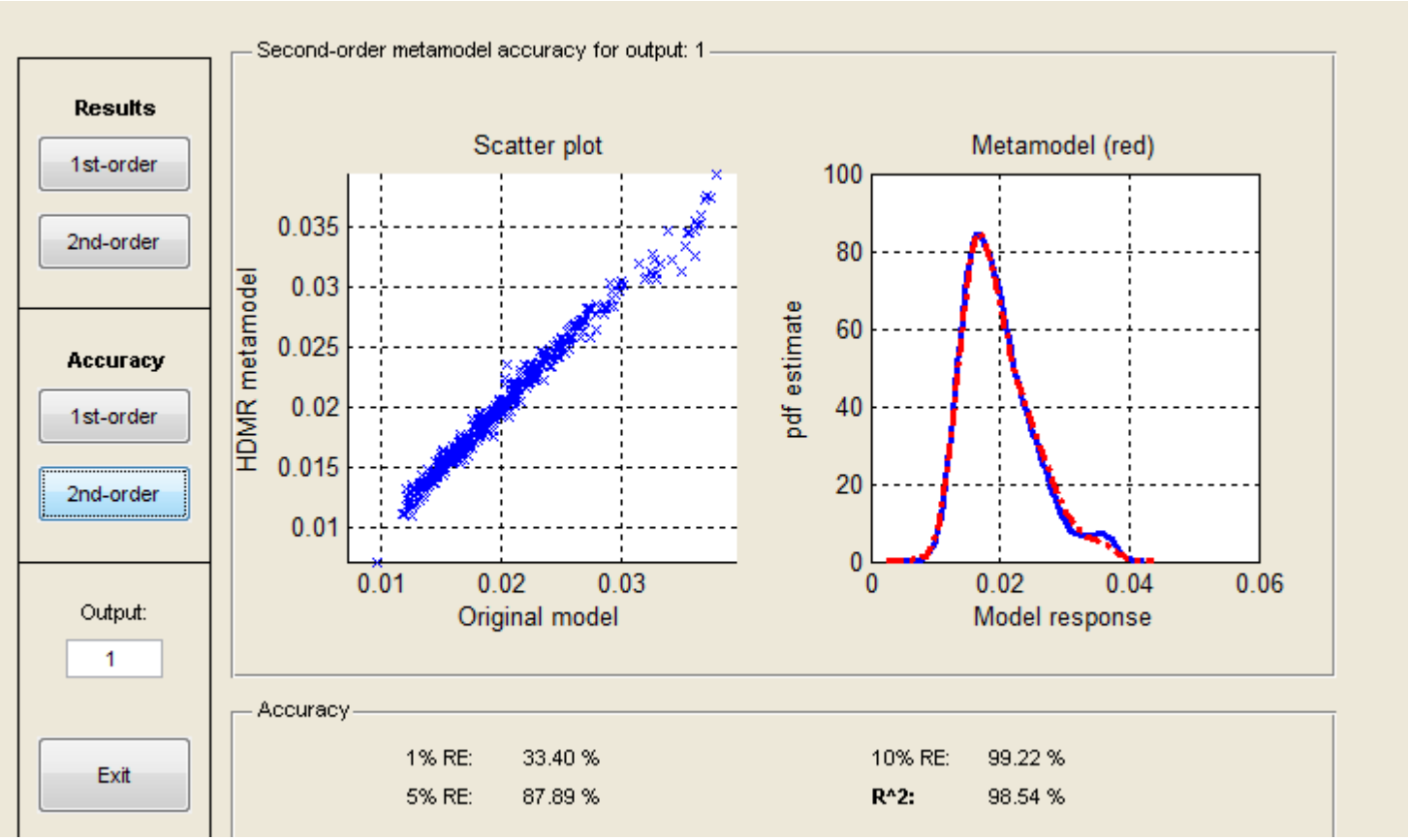


The right shows broad pdf of simulated concentration.

In this case the 1st order model is not a perfect fit to the data.

2nd order effects

UNIVERSITY OF LEEDS



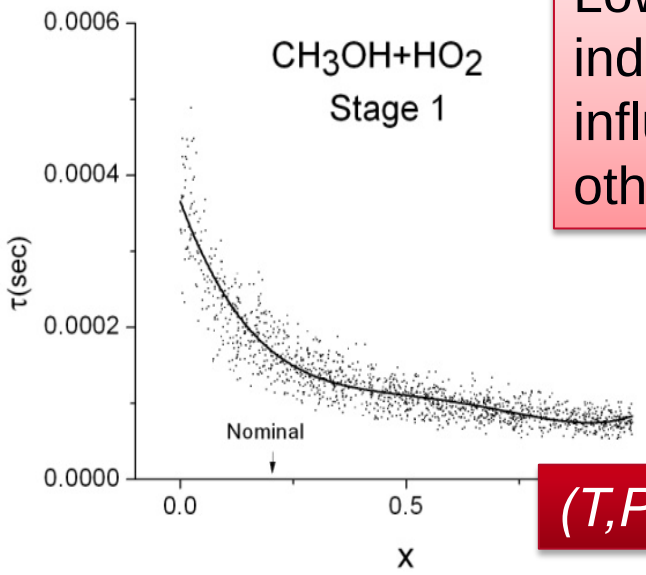
Including second order effects improves the model fit and the overall accuracy of the calculated sensitivity indices.

Methanol oxidation

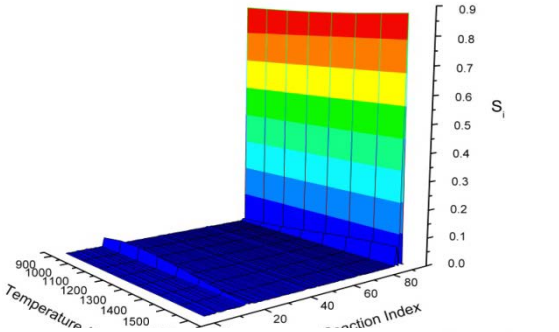
Sensitivity of ignition delays

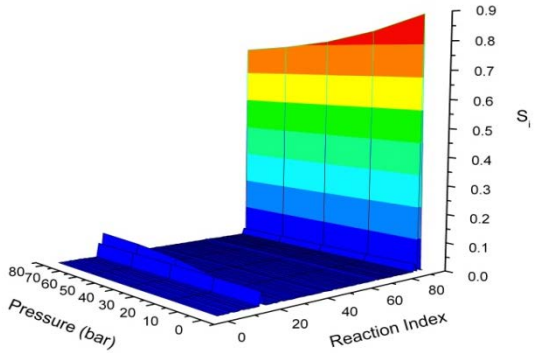
UNIVERSITY OF LEEDS

- Mechanism - Li et al. (2007); 18 species, 93 reactions.
- Target output - ignition delay time (τ) for stoichiometric mixtures of methanol and oxygen over a range of temperatures and pressures.
- Enthalpies of formation and A-factors varied over random sample.
- Using initial ranges one reaction dominated (up to 90% of total output variance).



Low scatter indicates low influence of all other parameters





(T,P,φ)=(1150K,5bar,1)

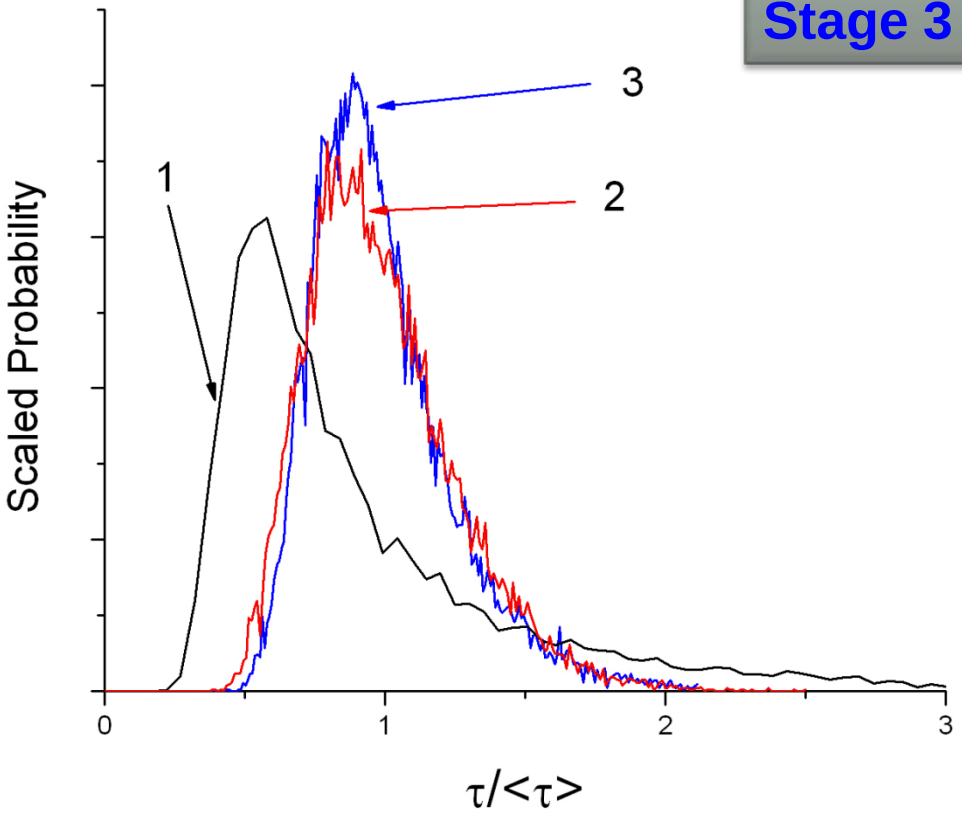
Results of model updates

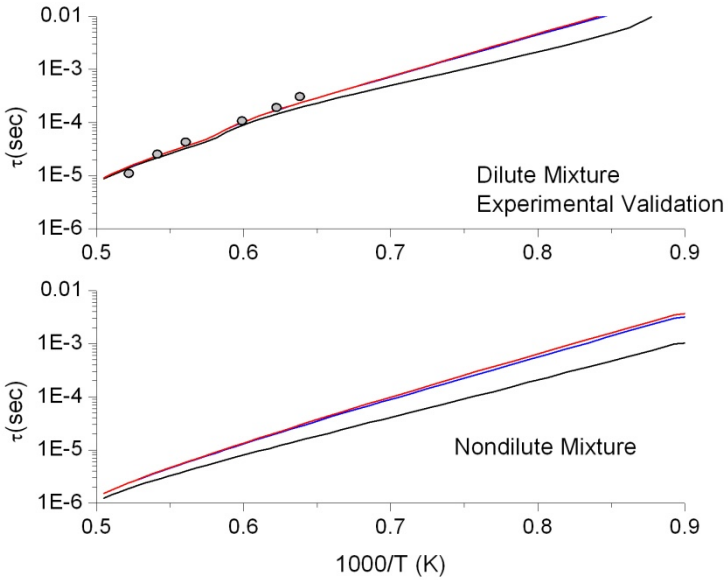
UNIVERSITY OF LEEDS

Stage 1 – Li mech

Stage 2 – $CH_3OH + HO_2$ updated TST

Stage 3 - $CH_3OH + O_2$ updated TST





(T,P,φ)=(1150K,5bar,1)

P = 1.5 bar

Butane oxidation in a jet stirred reactor

Performance of EXGAS mech.

UNIVERSITY OF LEEDS

Isothermal jet-stirred reactor

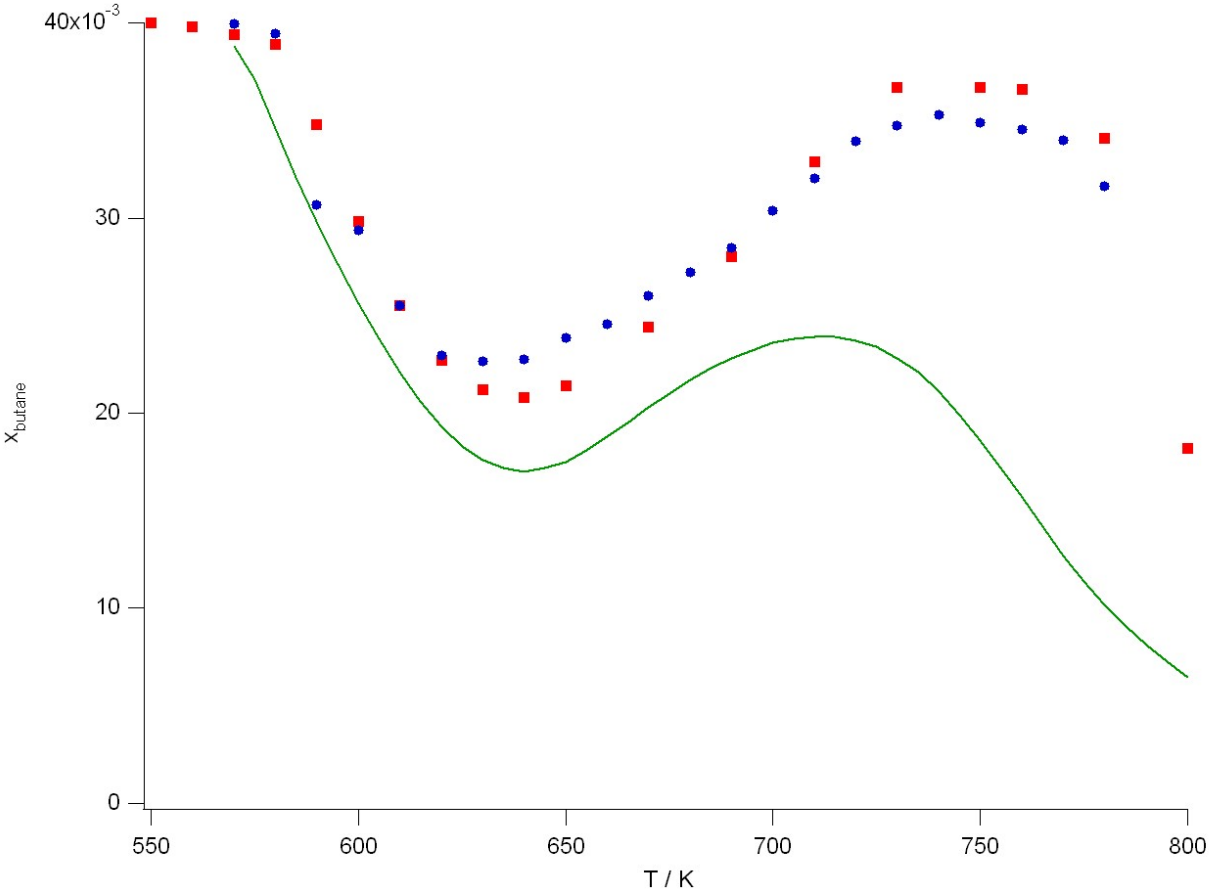
1 atmosphere

Residence time : 6 s

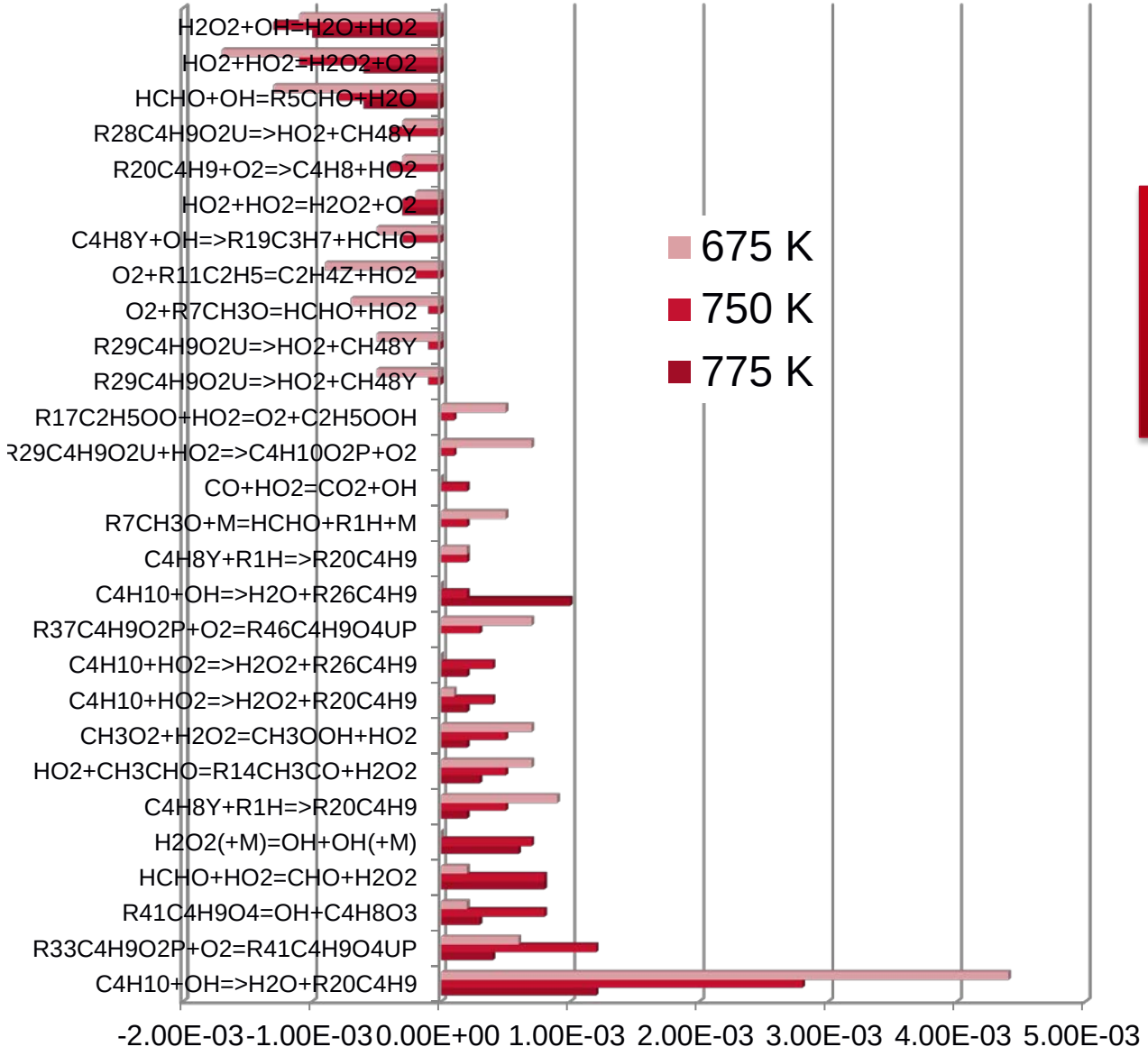
Equivalence ratio of 1

4% butane as inlet mole fraction.

EXGAS mechanism - 1304 uncertain A-factors for forward reactions studied.



UNIVERSITY OF LEEDS



Results from linear screening (25% decrease in A factors).

Global analysis (750 K)

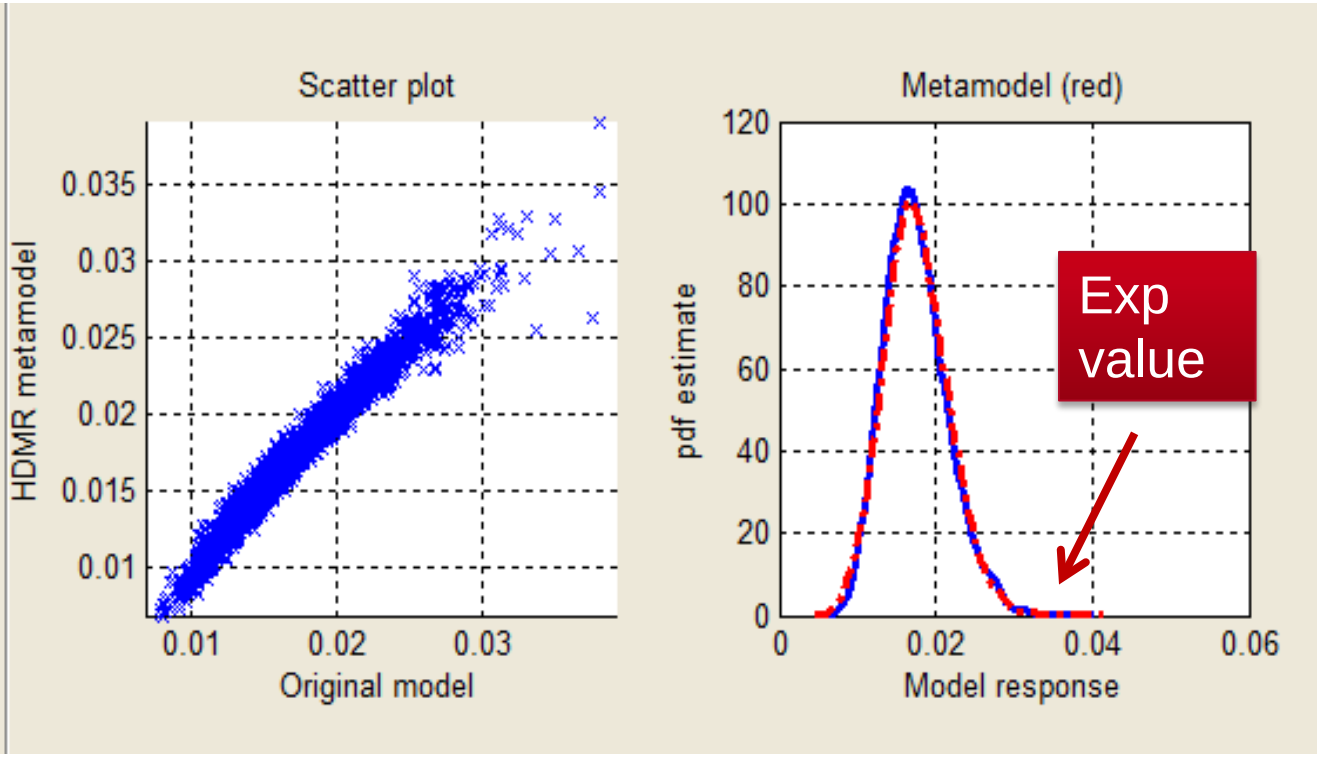
UNIVERSITY OF LEEDS

Evidence of nonlinearity, higher order parameter interactions.

512 runs sufficient to get main first order effects.

Higher order effects require several thousand runs.

31 A-factors selected for global runs.



Experimental value very infrequent unless significant variability in 2 butane+OH rates is allowed (>f=0.2)

First order effects + component functions

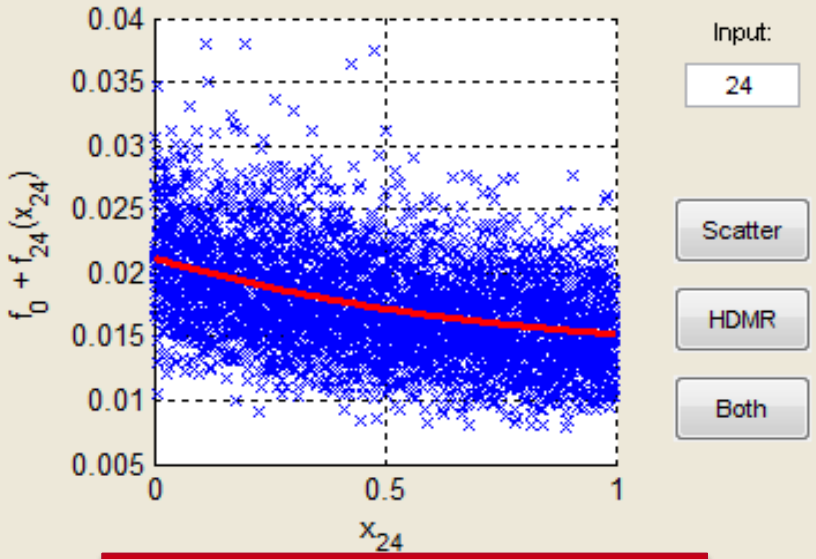
UNIVERSITY OF LEEDS

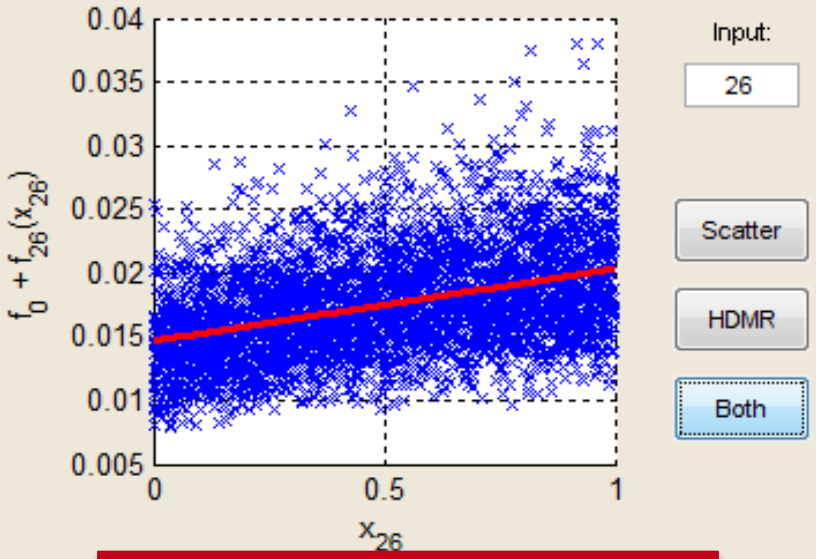
Ranking	1.	2.	3.	4.	5.
Input	24	26	29	1	30
Si	0.1764	0.1608	0.1037	0.0956	0.0842

Sum Si = 0.9463

Not really one dominant reaction

Plots





$HCHO+HO_2=CHO+H_2O_2$

Sobol Index :	0.1764 (1.)
r^2 Pearson :	0.1746 (1.)
r^2 Spearman :	0.1733 (1.)

$HO_2 + HO_2 = H_2O_2 + O_2$

Sobol Index :	0.1608 (2.)
r^2 Pearson :	0.1618 (2.)
r^2 Spearman :	0.1642 (2.)

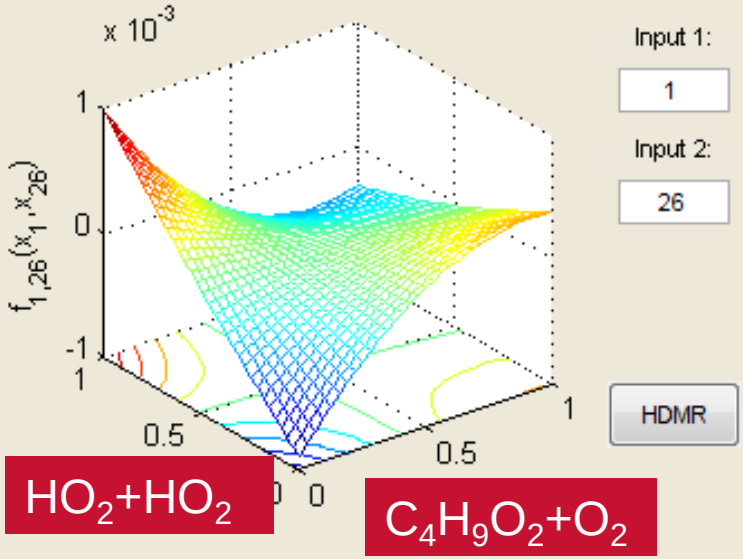
2nd-order component functions

UNIVERSITY OF LEEDS

Ranking	1.	2.	3.	4.	5.
Input	1,26	1,24	1,12	1,11	26,30
Sij	0.0031	0.0018	0.0010	0.0010	0.0008

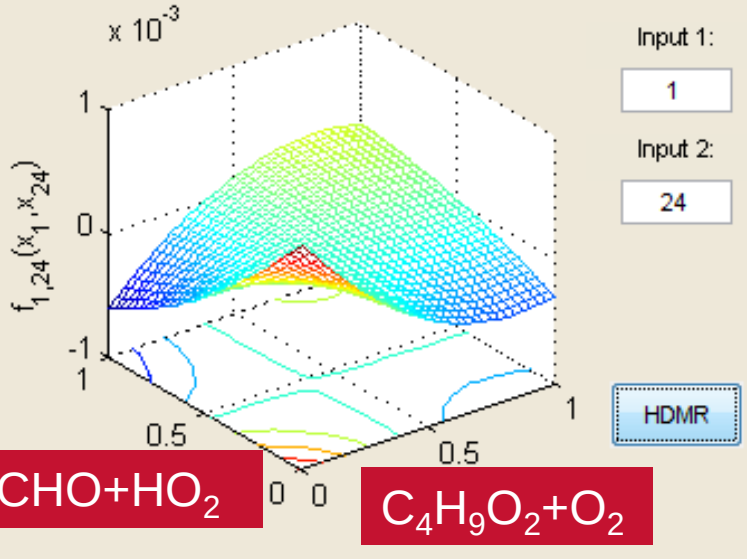
Sum Si = 0.9463
Sum Sij = 0.0162
Si + Sij = 0.9625

Plots



HO_2+HO_2

$C_4H_9O_2+O_2$



$HCHO+HO_2$

$C_4H_9O_2+O_2$

Polynomial Order :	2
Sobol Index :	0.0031 (1.)

Polynomial Order :	2
Sobol Index :	0.0018 (2.)

Reactions with high sens at high T

UNIVERSITY OF LEEDS

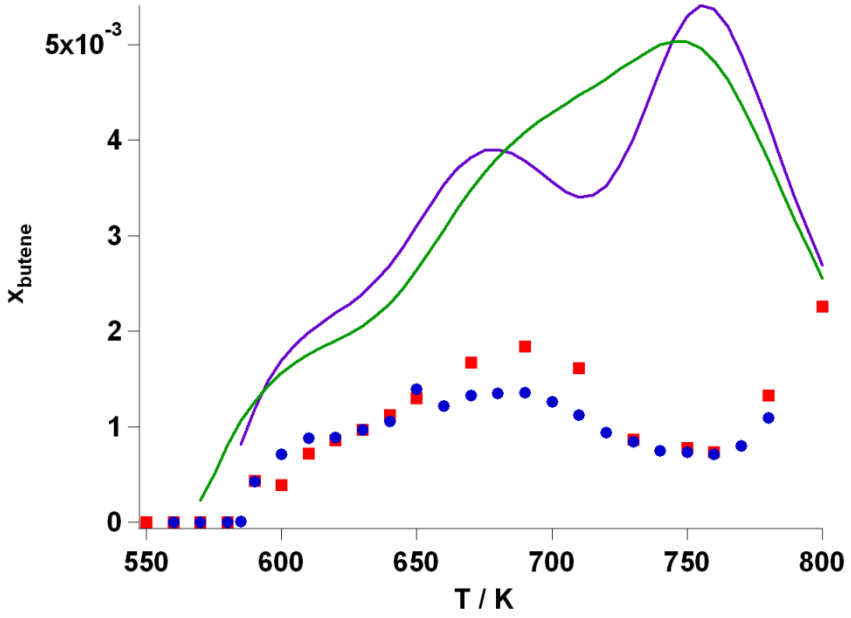
- 2nd oxygen addition reactions to form O_2QOOH species
  - ➢ also sensitivity to enthalpy of formation of these species but recent calculations have been made.
- Decomposition of O_2QOOH to OH and $C_4H_8O_3AP$.
- $HCHO + HO_2 = CHO + H_2O_2$
- $HCHO + OH = CHO + H_2O$
- H_2O_2 (+M) = OH + OH (+M) (*the third body efficiencies for this reaction vary between mechanisms from different groups for H_2O, CH_4, C_2H_6*)

The following have high sens at both low and high T

- $H_2O_2 + OH = H_2O + HO_2$
- $HO_2 + HO_2 = H_2O_2 + O_2$
- $CH_3O_2 + H_2O_2 = CH_3OOH + HO_2$

Effects of sensitivity studies: reduce A-factor for 2nd O_2 addition by factor of 2

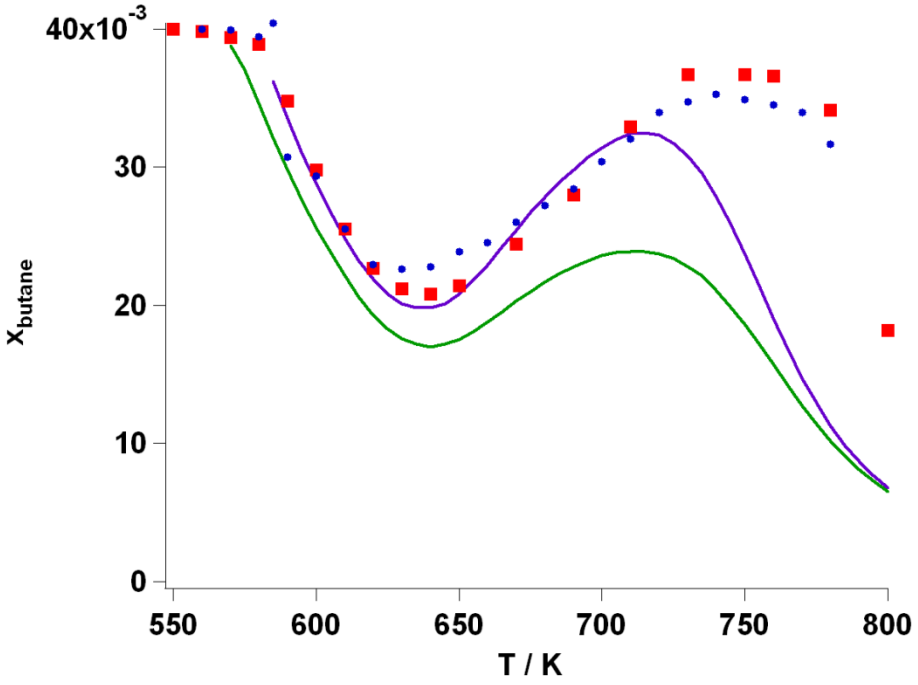
UNIVERSITY OF LEEDS



Reduction not inconsistent with recent work from Bozzelli's group.

For butene rate of reaction channels for

C_4H_8Y + OH are also very important .



Discussion

UNIVERSITY OF LEEDS

- Often only a **small number** of parameters drive output uncertainty.
- Local/global sensitivity methods provide useful step in model evaluation by identifying this parameter set and exploring feasible range of predictions.

Not always according to the experienced chemists intuition.....

- Further *ab initio* studies can then be **focussed on key parameters** improving model performance.
- Tuning should probably only be carried out with good reason and should be documented.
- Where simulations including uncertainties don't overlap with experiments – possible evidence of missing pathways/uncertainties.

Discussion 2 – requirements?

UNIVERSITY OF LEEDS

- In order to put error bars on model predictions and to compute global sensitivity coefficients requires:
  - Uncertainty ranges **AND** (joint?) pdf's for all input parameters. Sometimes available from evaluations such as Baulch but otherwise should be estimated by mechanism generators.
  - **And provided to users....**
- Information about correlations between inputs also required - at least where structural arguments have been used and Arrhenius parameters for several rates are related.
  - Should these parameters be sampled together since they come from common sources or are calculated using same methods?
- Could mechanisms be automatically extended to provide such information to users?

Open questions

UNIVERSITY OF LEEDS

- How to estimate uncertainties from for example TST calculations.
  - Apply global sensitivity analysis to these methods?
- It is wrong to restrict to only A-factors and enthalpies of formation and to ignore joint probability distributions but for how may systems do we have better information?
  - Does it matter just for key parameter identification i.e. If not using optimisation?

Sensitivity + high level theory / experiment **?**

Optimisation against all available experiments **?**

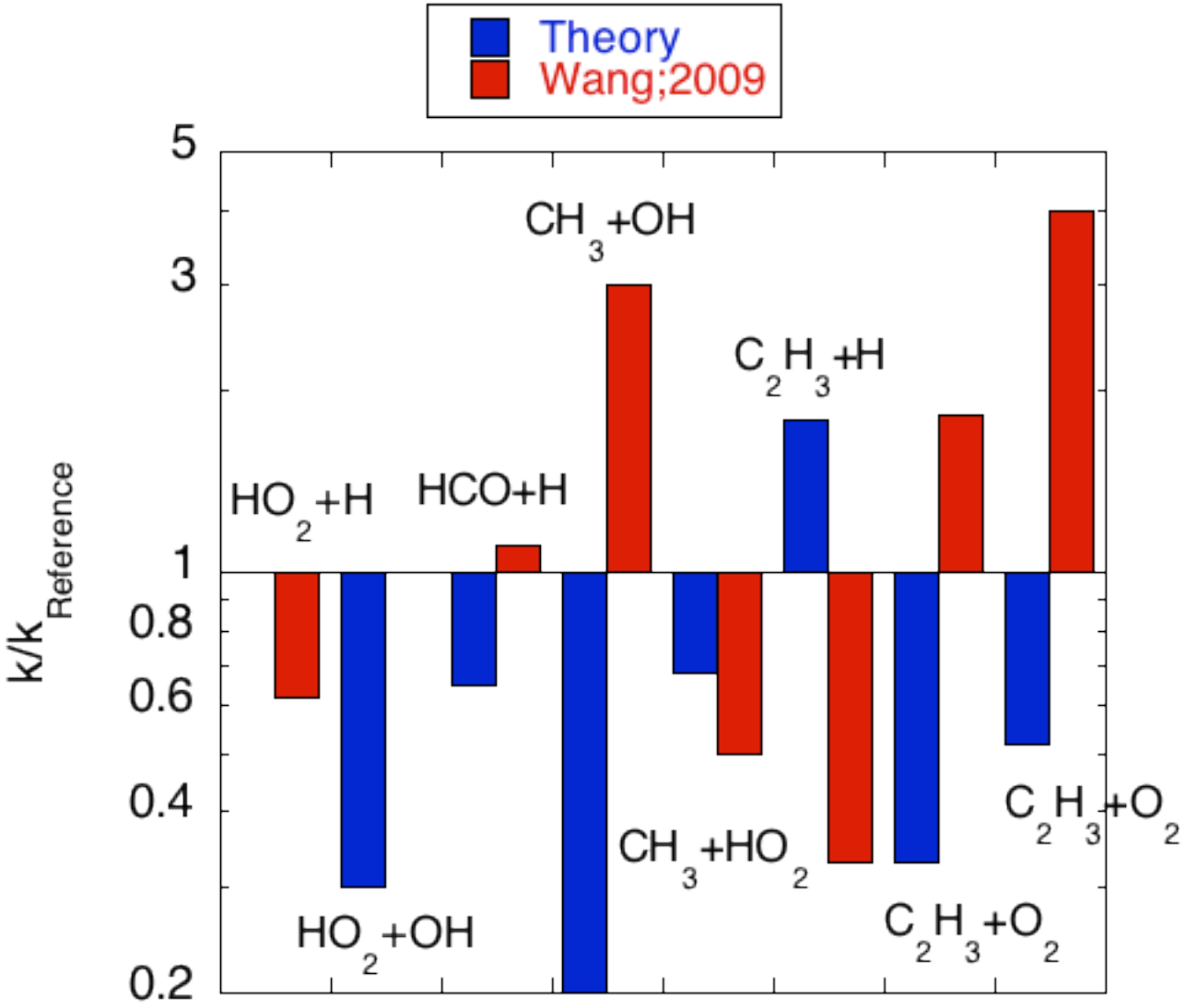
Both **?**

Model Optimisation: no cost function related to nominal value

UNIVERSITY OF LEEDS

Sheen et al. (2009) for Ethylene Combustion

This approach has now been superseded



Model Optimisation: including cost function related deviation from nominal value

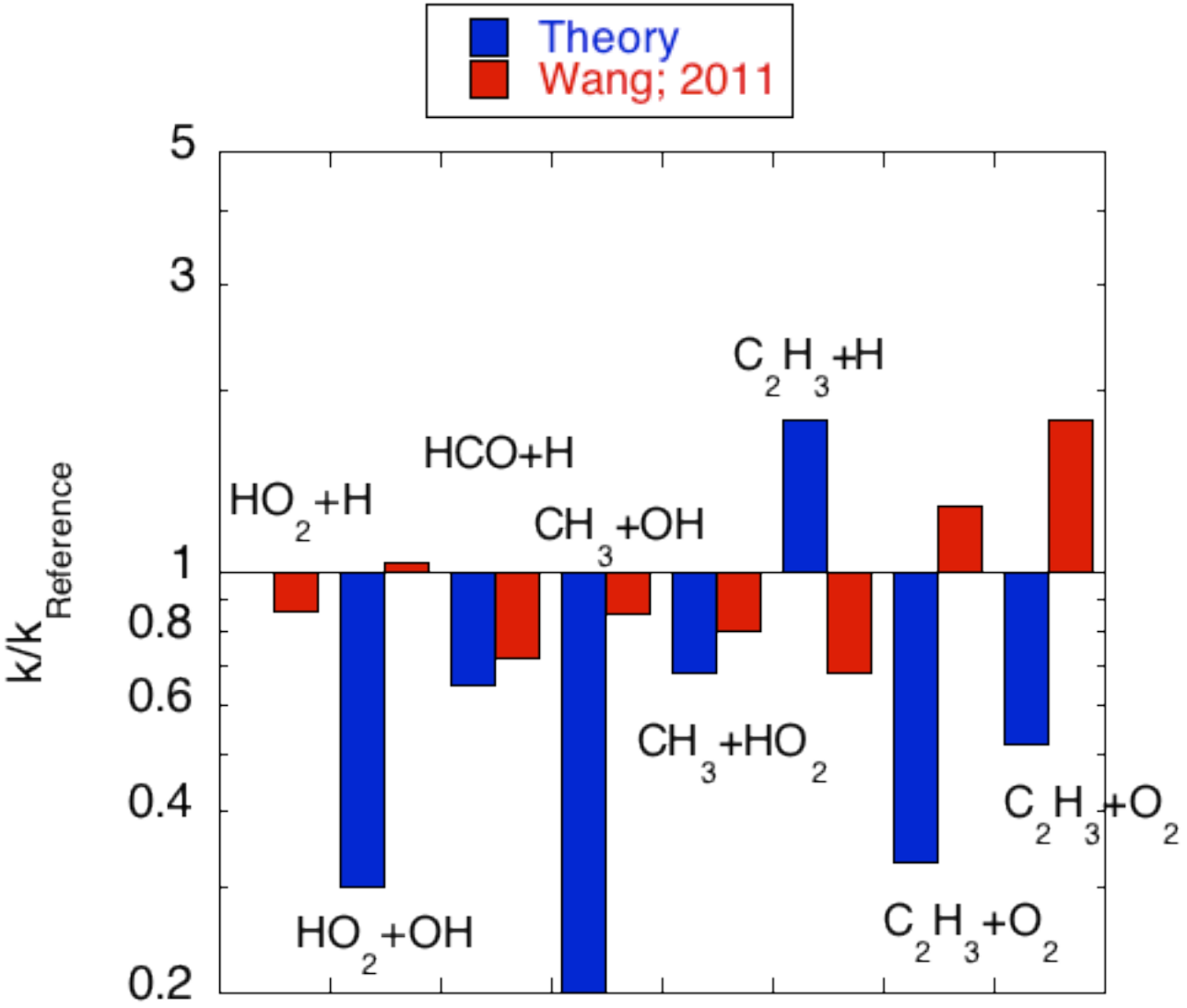
UNIVERSITY OF LEEDS

Approach used in:

Sheen et al. (2011)

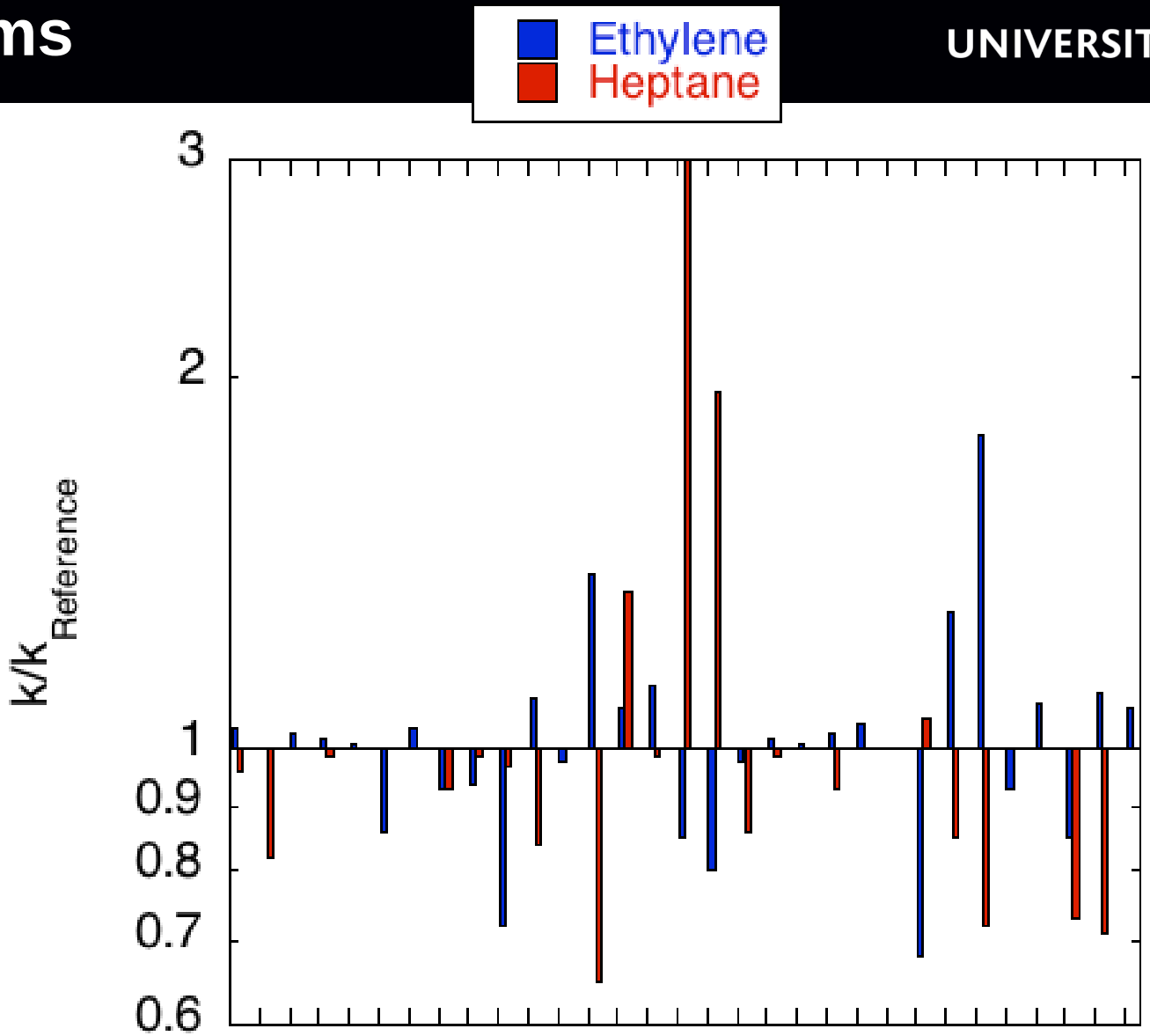
You et al. (2011)

Still sensitive to nominal value used…



Comparison for different systems

UNIVERSITY OF LEEDS



Discussion

UNIVERSITY OF LEEDS

- There should be error bars on both sets of data – this is something we should work on...
- Include theoretical values in optimization procedure?
- Would require uncertainties in both experimental and theoretical values to be available.