

Applying Consistency and Optimization Tools to Atmospheric Datasets: An Upper Atmosphere Ozone/Hydroxyl Example

Gregory P. Smith, *SRI International*
Michael Frenklach, Ryan Feeley, Andrew
Packard & Peter Seiler, *UC Berkeley*

Support from NASA Geoscience, NSF Aeronomy,
and NSF Chemical and Thermal Systems programs



-
- The problem and setup for upper stratosphere/mesosphere photochemistry (H,O)
 - Results obtainable from current sensitivity, optimization, and data collaboration tools
 - Conclusions and uses
 - J. Geophys. Res. (A) 111: D23301 (2006)



-
- Limited set of sounding, balloon, and satellite measurements of O_3 , OH, HO_2 at ≥ 40 km, by IR or μ -wave emission or UV solar limb absorption or emission.
 - Simple photochemistry: O_x photolysis, OH + O & O_3 , H + O_2 , O & OH + HO_2
 - Models underpredict O_3 (ozone deficit), and HO_2 .
 - Models overpredict mesospheric OH but are inconsistently low at 40 km (HO_x dilemma)
 - Ad hoc suggested alterations: \downarrow O + HO_2 , may \uparrow OH + O, \uparrow (or \downarrow) OH + HO_2
 - Systematic examination required, new solutions unlikely
 - Relevant lab/reactor experiments should be included
 - ☞ How do we systematically determine which atmospheric measurements are consistent with each other, and with the measured photochemical rate constants?
 - ☞ What optimized combination of rate constants, within their uncertainties, gives the best predictability?

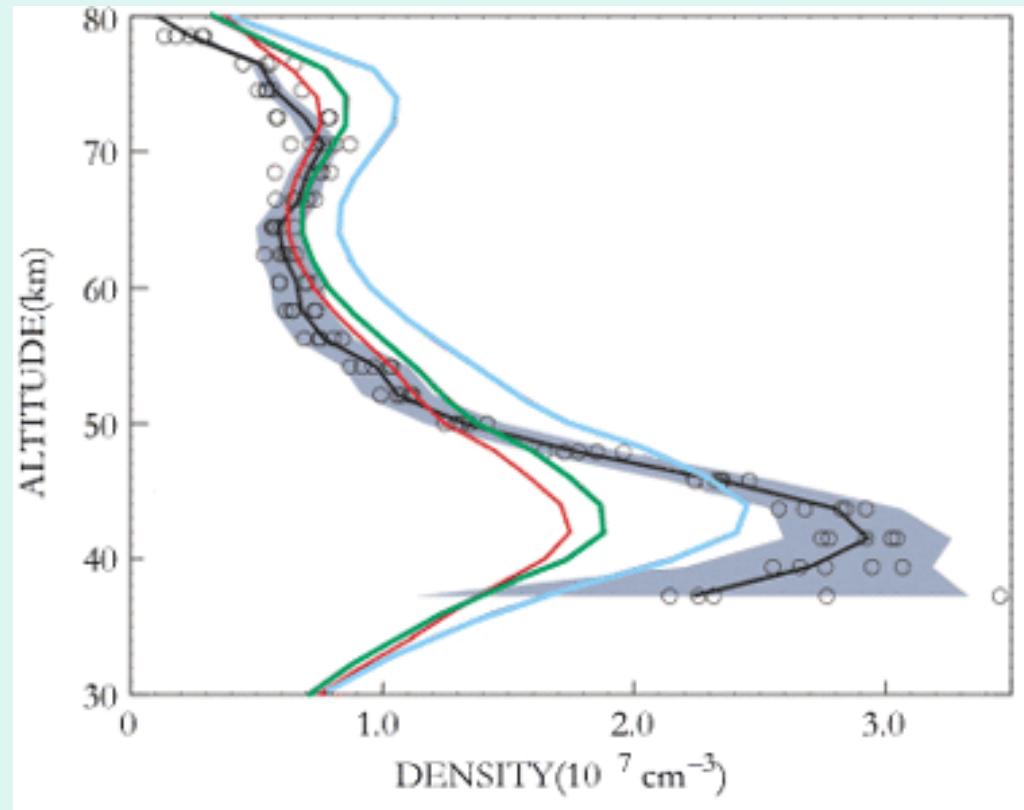


MAHRSI OH data 8/15/97 54N noon Conway et al GRL 27, 2613

JPL-97

-50% O + HO₂

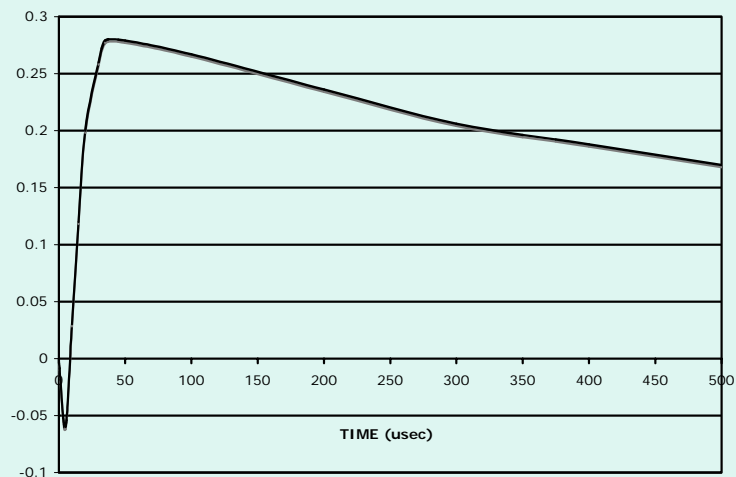
-20% O + HO₂,
+30%OH + HO₂



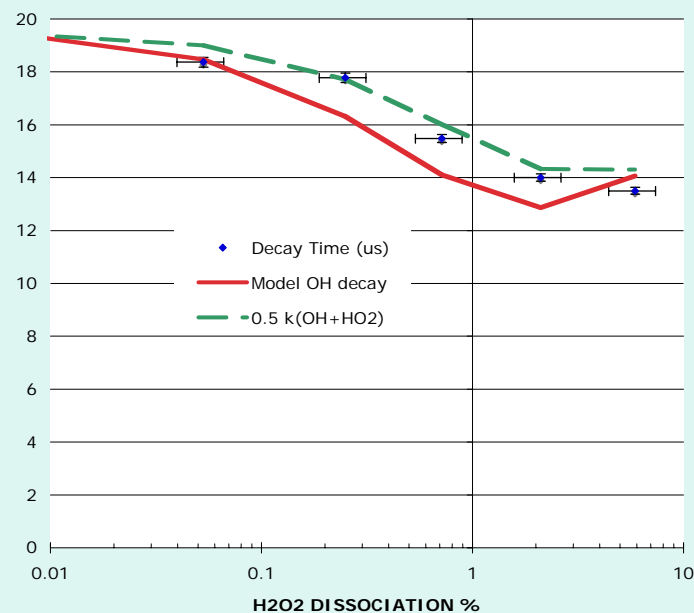


Lab Reactor data targets

OH decay (IR) after 2.3mJ/cm² photolysis of .03O₃, .36H₂, .48CO₂, .12Ar @ 25torr 243K (Nizkorodov et al. JPC A 104, 3964 (2002));
OH , HO₂ + O₃



Photolyze <5% of 1 torr H₂O₂ in 40 torr N₂, measure OH decays (SRI);
OH + H₂O₂, HO₂





Computational Model Construction

- Select representative results from atmospheric data, as **targets**. Include relevant lab reactor results. Find UARS-HALOE values for ozone and water if not given.
- Obtain corresponding results from LLNL 2-D diurnally varying model (thanks to Peter Connell); perform box model sensitivity analysis to identify controlling rate parameters for targets. $S_{ij} = dX_i/X_i / dk_j/k_j$
- Evaluate liberal uncertainty limits for both observations and rate constants. Include model uncertainties (water, irradiation, optical thickness).
- Rerun the box model to compute target predictions for any allowable combinations of sensitive rate parameters; generate a polynomial function to describe this **response surface**.
- To optimize, minimize **objective function** $\Phi = \sum(\Lambda) \delta^{-2} [1 - X_C/X_T]^2$, where the X are the modeled and target concentrations, the weight δ is the uncertainty. (Options, such as $\Lambda = 0$ or 1 depending on whether $|1 - X_C/X_T| < \delta$.)



Optimize rate parameters to predict 3 lab targets and 36 O₃, OH, HO₂ representative observations at 38 - 73 km.

Unsatisfactory results: fail for 16 targets.

Which modeled targets are inconsistent and should be reexamined or excluded?

Minimize slack variable t (extra uncertainty) subject to

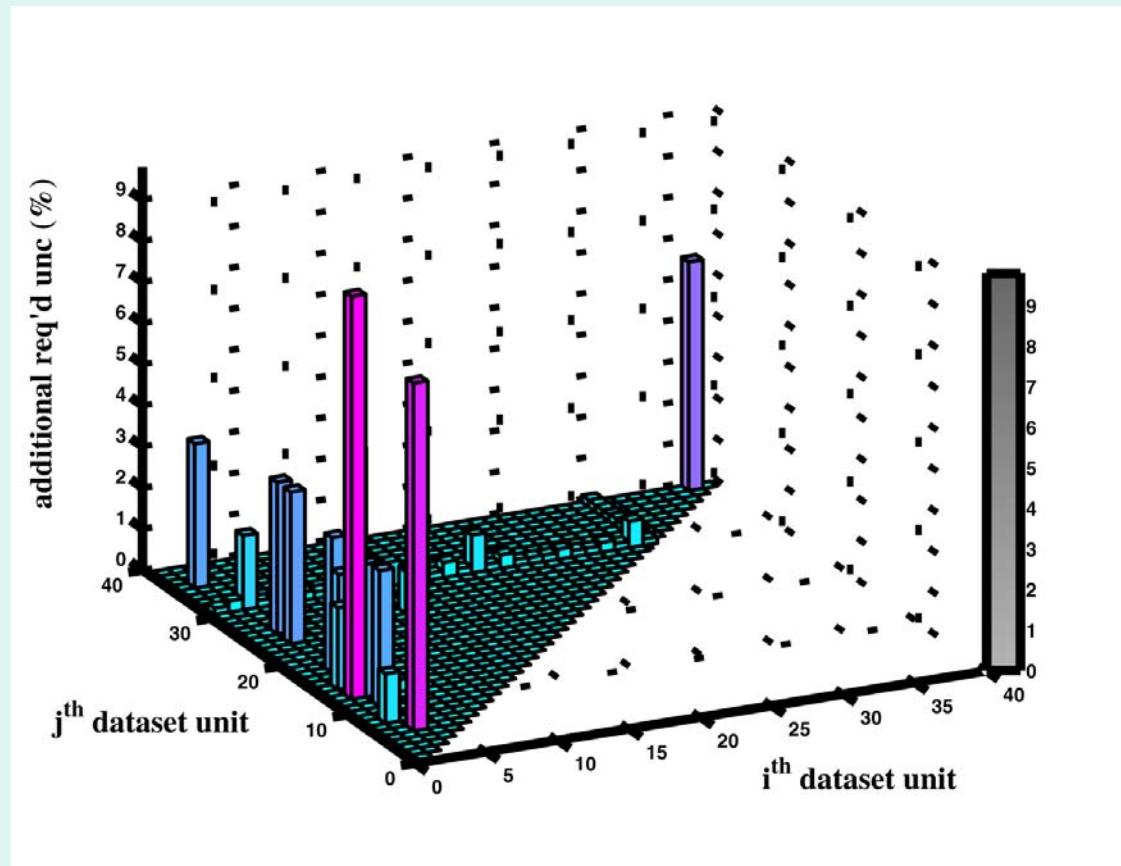
$$|1 - X_{ci}(k)/X_{Ti}| < \delta_i + t.$$

We want to find a feasible set k such that $t < 0$.

The target set i can be all 39, or 2 at a time.

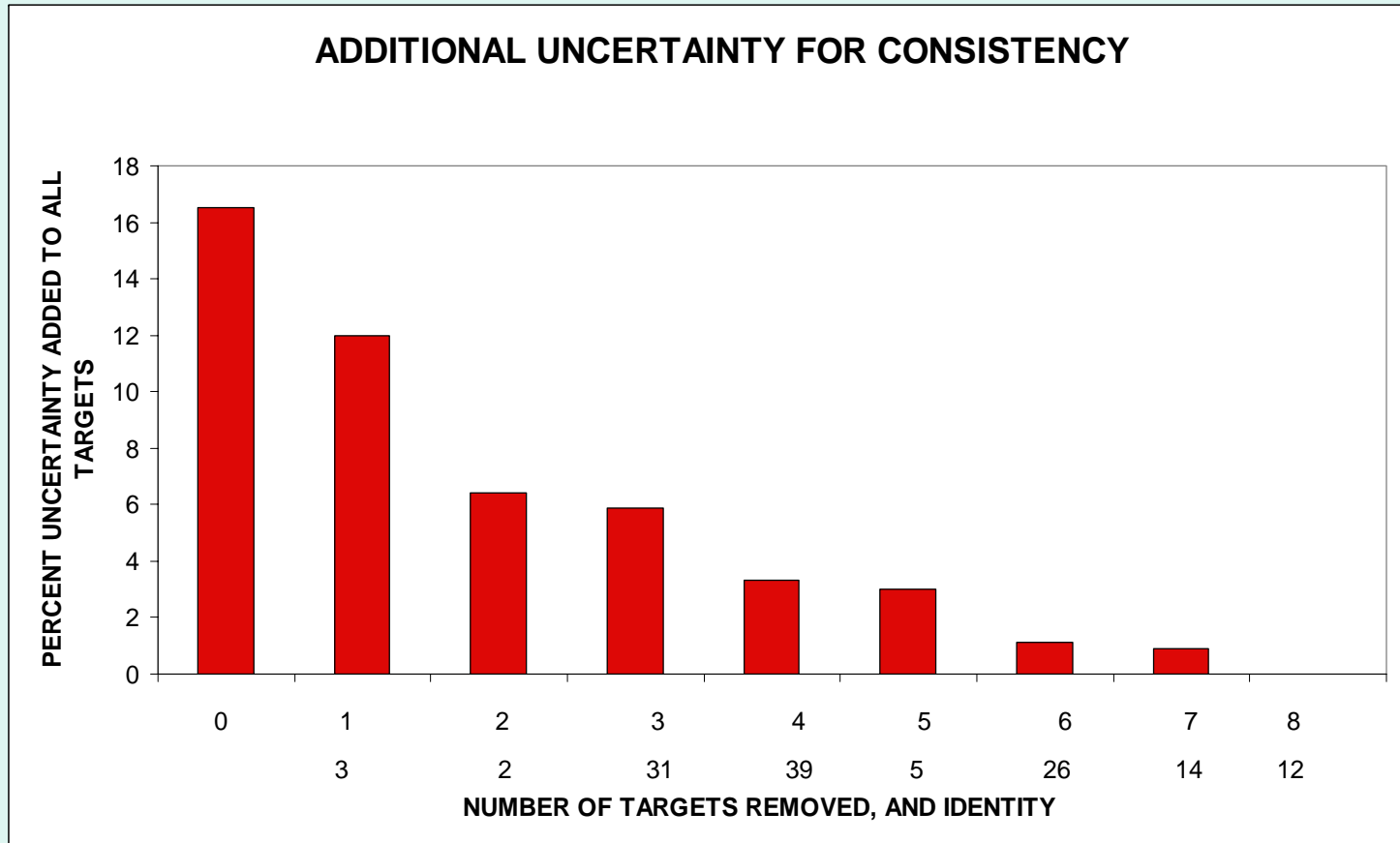
1. Pairwise Inconsistent Data:

How much uncertainty must be added to a pair of targets to find a feasible mechanism?





2. Add uncertainty to entire dataset until feasible mechanism exists; eliminate the most sensitive target to this 'slack variable' parameter and repeat; data collaboration finds maximum self-consistent group



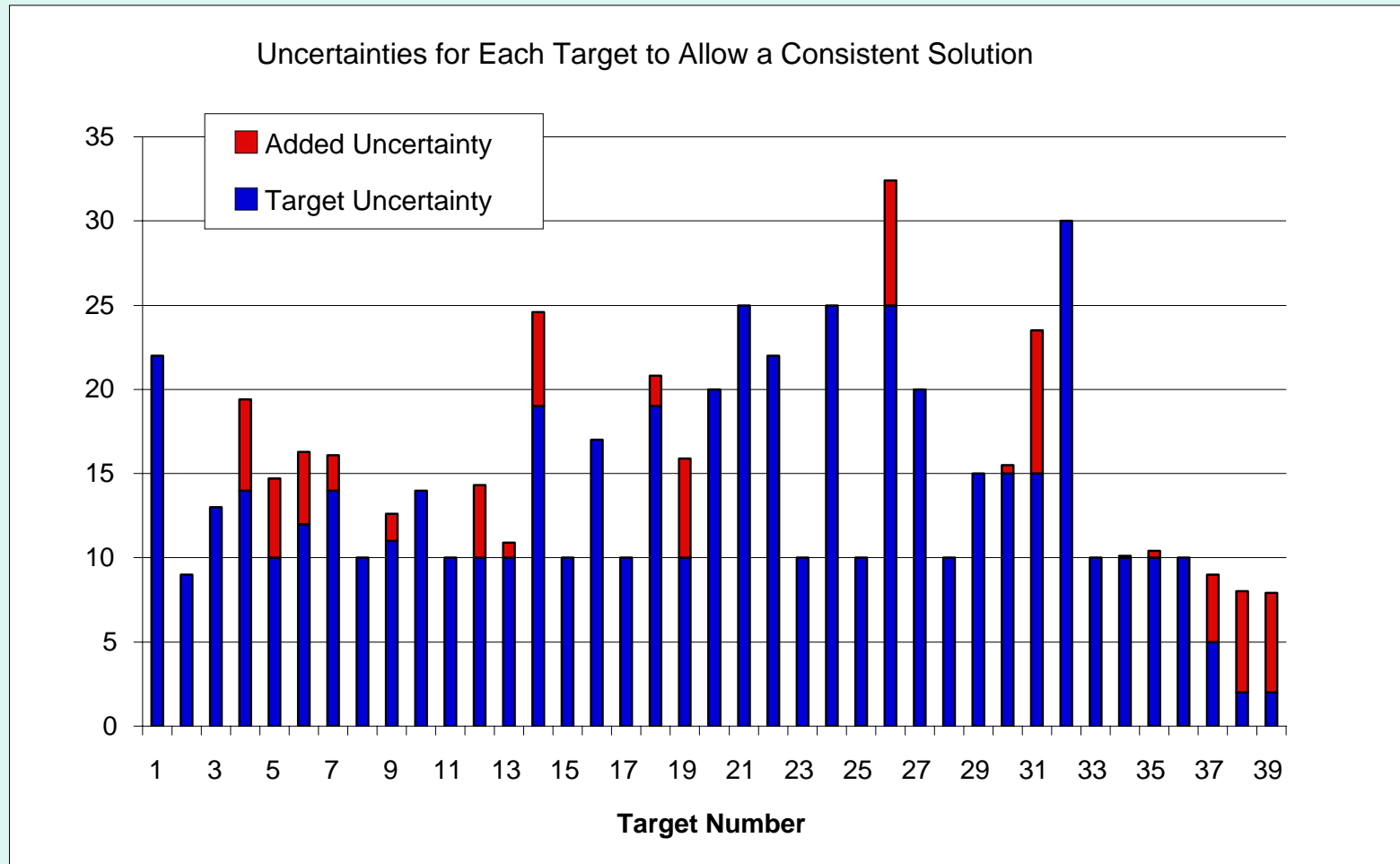


Slack uncertainty variable sensitivities after eliminating 3 targets (3, 31, 39)

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.



Adding uncertainties to specific targets to achieve consistency - a more difficult constraint to set up -- eliminate targets 2 & 3, minimize $\Sigma(\Delta U_i)^2$

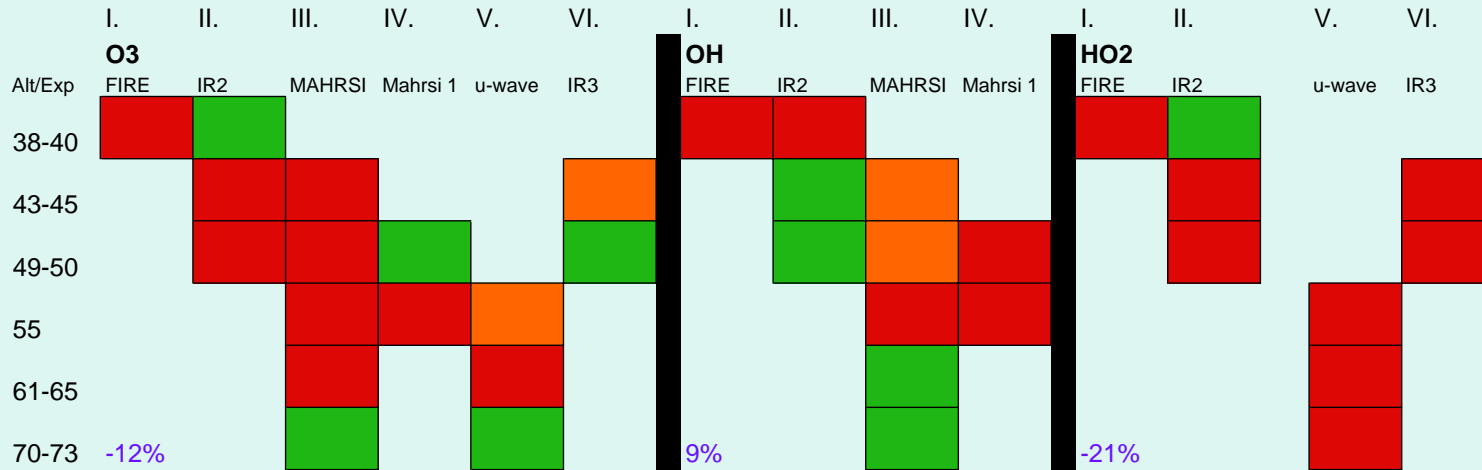




3. TARGET FITTING PATTERN FROM OPTIMIZATIONS

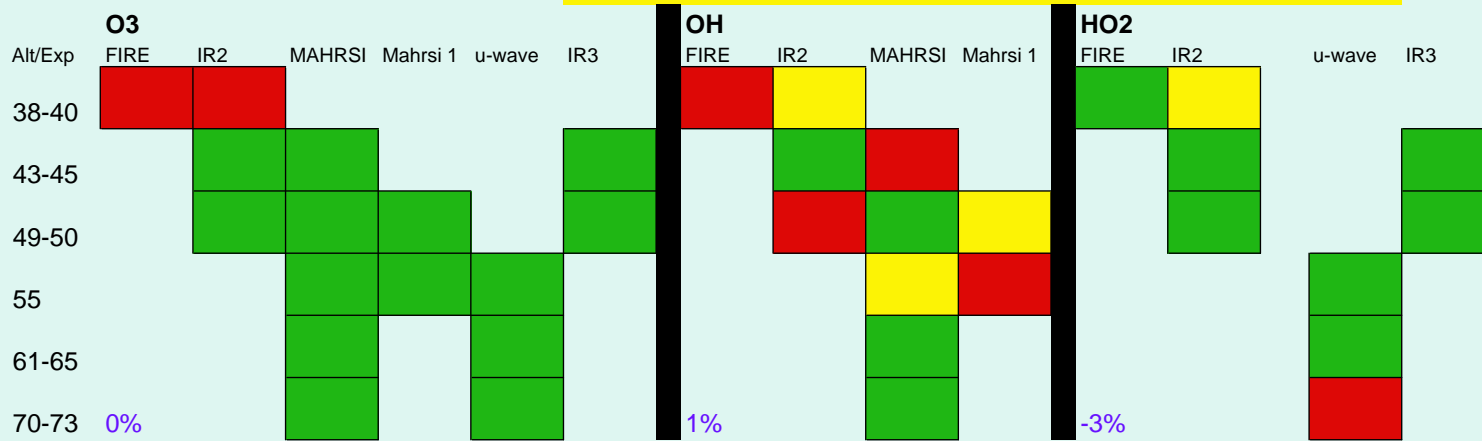
UNOPTIMIZED BASE MODEL

RED TARGET BLOCKS EXCEED UNCERTAINTY



OPTIMIZATION 1. CONSISTENT SET

YELLOW SET 2: SMALLER RATE CHANGES SOLUTION MISS





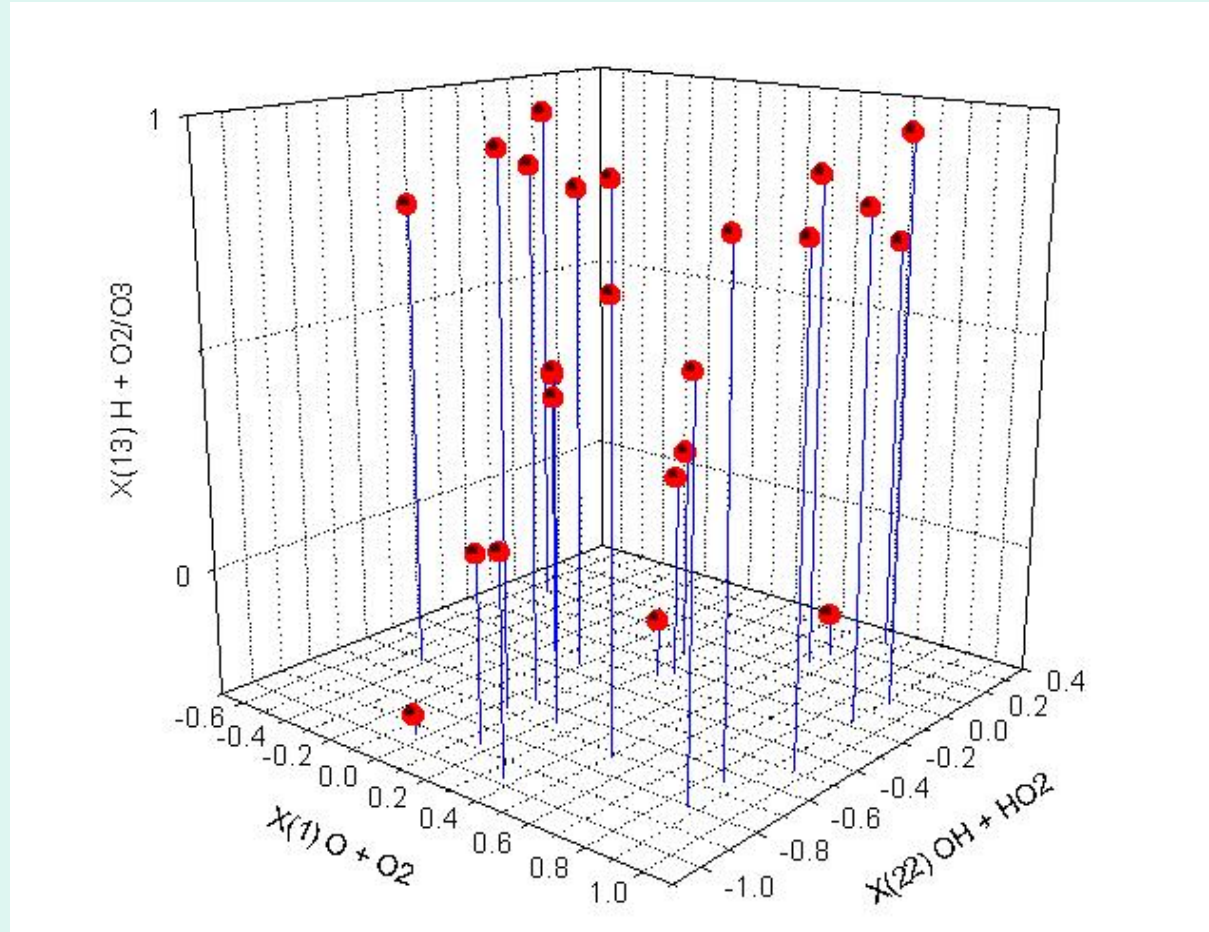
Sensitive Optimization Variables: Ranges & Changes

| Reaction | Range | Lagrange Multiplier | 1 Consistent Set | 2 Smaller Changes |
|--|----------|---------------------|------------------|-------------------|
| $O+O_2 \rightarrow O_3$ | +20% | +0.011 | +13.2% | +9.8% |
| $O+O_3 \rightarrow 2O_2$ | +50% | -0.003 | -32.8% | 0% |
| $O(^1D)+air \rightarrow O$ | +20% | -0.001 | -15.4% | 0% |
| $O(^1D)+H_2O \rightarrow 2OH$ | +30% | +0.002 | +19.5% | +20.2% |
| $H+O_2/O_3 \rightarrow HO_2/OH$ | +60% | +0.004 | +59.3% | 0% |
| $OH+O_3 \rightarrow HO_2+O_2$ | +35% | -0.001 | -18.1% | 0% |
| $OH+O \rightarrow H+O_2$ | +35% | +0.006 | +32.1% | +27.8% |
| $HO_2+O \rightarrow OH+O_2$ | +30% | -0.017 | -6.8% | -12.8% |
| $HO_2+OH \rightarrow H_2O+O_2$ | +50% | -0.006 | +23.4% | -11.0% |
| $OH+H_2O_2 \rightarrow HO_2+H_2O$ | +35% | -0.011 | -15.8% | -11.6% |
| $O_2+h\nu \rightarrow 2O$ | +17(30)% | +0.019 | +17.0% | +16.3% |
| $O_3+h\nu \rightarrow O_2+O$ or $O(^1D)$ | +30% | -0.014 | -3.5% | -13.6% |
| $H_2O+h\nu \rightarrow H+OH$ | +35% | -0.001 | +34.9% | 0% |

Lagrange Multiplier is the sensitivity of the additional dataset error parameter for obtaining consistency to changes in each specific rate parameter



4. Multiple Solutions (28+/39 targets fit) What does a marginally feasible set look like according to Monte Carlo polynomial sampling?



Most but not all solutions favor decreasing $OH + HO_2$ (X), increasing $O + O_2 \rightarrow O_3$ (Y), and increasing $H + O_2/O_3 \rightarrow HO_2/OH$ (Z)



One can also look at correlations among rate constant changes, ranges of feasible values, and targets almost always or never rejected, from the multiple solution set.

$O_3 + hv$ & $O + O_2 \rightarrow O_3$, $H_2O + hv$ & $OH + HO_2$ changes correlate.

“Redundant” data sensitive to same rates (can reduce uncertainty)

Correlation coefficients of sensitivities:

| | |
|-------------------------|-------|
| O_3 40km - O_3 65km | +0.91 |
| OH 65km - O_3 65km | +0.17 |
| HO_2 45km - OH 45km | +0.37 |
| OH 43km - OH 65km | +0.62 |

Information is gained when the feasible set is restricted by added data or smaller uncertainty.

Rate constraints formally imposed by UT/LS HO_2/OH data (+ O_3 , NO, CO): Lanzendorf et al., GRL 28: 967 (2001)



Conclusions

- A consistent US/LM HO_x data+parameter set requires liberal uncertainty limits, some exclusions, and significant rate changes.
- Optimized mechanisms for most high altitude results require increased photolytic O₃ and OH sources and a faster OH + O.
- New data (Canty et al, GRL 2006) and kinetic measurements (SRI) confirm this last point, and could be added to an automated analysis.



Questions that can be posed

- Are dataset units self-consistent? Locate and quantify.
- What models will account for all the information & error limits?
- Multiple mechanism representations can be optimized to the same dataset and compared.
- Probabilities are assignable to predictions.
- What should be measured and how well, in order to reduce an uncertainty (reaction rate, lab data, or field measurement)?
- Inferring source terms; data assimilation
- Have we used everything we know to model a scenario and its uncertainty?
- What data entry, analysis questions and tools, product plan, and resources would be necessary to approach smog model generation systemically? PRIME resources available, expandable