

## Slide 1

This document is intended to be a companion to the presentation prepared as a class project for the course Quantifying Uncertainty (12.S990, 2012). The presentation is called “[A] Probabilistic collocation method (PCM) for modeling [the] response of GEOS-Chem simulations [of benzo[a]pyrene] to model parameter uncertainties.” Although this particular study involves the simulations of benzo[a]pyrene (hereafter “BaP” or “bap” for convenience), it is also meant to be a testing of the waters for use of the PCM on the entire collection of GEOS-Chem tracers whose simulations rely on uncertain parameters. This work describes (1) the motivation behind, (2) a vague introduction to, and (3) the results of, using the PCM on GEOS-Chem parameters and output.

\*Note: For clarity, throughout this document I use pairs of apostrophes to bookend 'strings of words which represent a single object or entity but may not clearly denote a single object without the apostrophes'.

## Slide 2

GEOS-Chem's three current persistent organic pollutant (POP) simulations, which include the chemicals phenanthrene, pyrene, and benzo[a]pyrene, are very new developments which were introduced by Carey Friedman and Noelle Selin (a submitted 2012 paper). BaP in particular is the focus of this project, and is of special interest for Arctic life, because it is a toxic chemical that accumulates in the polar region and has been implicated as a carcinogen. Due to the novelty of the POPs modules, and the fact that long-term measurements of the substances are very sparse in the Arctic, simulation output for the Arctic is a prime candidate for the uncertainty analysis that is performed below.

## Slide 3

To explicitly state the problem addressed by this project, the uncertainty in the Arctic annual mean total atmospheric concentration of benzo[a]pyrene is estimated based on the given uncertainties in the chemical parameters that are included in the GEOS-Chem model of BaP. The problem is then one of uncertainty propagation through the somewhat complex and non-linear calculation that is a GEOS-Chem model run.

## Slide 4

To solve the problem stated above, the probabilistic collocation method (PCM) is used, which involves the fitting of 'the model's response to changes in its parameters' to a polynomial which can then be sampled to get some statistical knowledge about the uncertainty due to these parameters. In more

detail, PCM is based on polynomial chaos expansion (PCE), which is the fitting of a response to a set of orthogonal polynomials. The PCM provides the criteria to select collocation points, which are sets of parameter values which will best (most efficiently) allow for estimating the PCE fit of the model response. Each set of parameter values (i.e. each collocation point) is then fed into the model and the model is run to get the output. This is repeated until a complete collection of 'parameter value set' / 'output' pairs is compiled. A complete collection in this case is when there are an equal number of 'pairs' and 'terms in the polynomial to be fit'. The number of terms in the polynomial depends on the order of expansion used, with the number increasing rapidly with order and number of parameters. (e.g. For 6 parameters the second-order expansion has 28 terms and the third-order has 84.)

Once the model runs are complete, the polynomial, the collocation points, and the output form a system of equations that can be solved to find the coefficients of the polynomial and thus obtain the estimator of the model response. The model response can then be sampled using the polynomial expansion instead of the much more computationally intensive full model.

#### Slide 5

In a 1997 paper by Pan et al. (JGR, 102, 21915-21924), PCM was used on some relatively simple models of atmospheric radiative forcing. These simple models were, themselves, easy to sample and therefore appropriate to use to test the performance of PCM, since a true distribution of uncertainty could be generated to compare to the estimated distribution. To test the implementation of PCM used in this project, one of these radiative forcing calculations was used (and slightly simplified). A comparison of the “true result” - that attained by sampling the given function itself – and the estimated results (the PCEs of 2<sup>nd</sup> through 5<sup>th</sup> order) was performed and is shown on the next slide.

#### Slide 6

With the true distribution in black and the estimated in blue, it is easy to see how well the method performs, even at low order. Since the 4<sup>th</sup> and 5<sup>th</sup> order approximations will be difficult to perform with GEOS-Chem output due to time constraints, it is important that the PCE has success at lower order.

#### Slide 7

Moving on to the application to GEOS-Chem itself, the uncertain chemical parameters are listed on this slide. The octanol-air partition coefficient is important for both the deposition of bap onto vegetated surfaces (atmospheric removal) and onto organic carbon particulate matter (which will affect its atmospheric lifetime). The black carbon-air partition coefficient also determines bap affinity for

particulate matter, particularly black carbon aerosol.

The oxidation rate parameter is the rate constant for reaction with hydroxyl radical, the only atmospheric removal reaction for bap. Henry's Law constant affects the fraction of gas-phase bap that will be removed by wet processes such as rainout and washout. The enthalpy of phase transfer from gas to organic carbon is used for temperature-dependent partitioning to particulates in the atmosphere, and the gas-aqueous enthalpy is important in wet deposition.

### Slides 8-10

The next three slides are plots of (estimates of) the uncertainty distributions of the parameters described above. All of the distributions are log-normal, and correspond to highly uncertain parameters, with the wings of the distributions extending outside of a three order of magnitude range. This is due to a combination of the scarcity of measurements of these parameters and the large spread in measured values when they are measured. In all of the figures, the green line shows the value of the parameter that was chosen for use in the model for prior studies, and the red lines show other reported values for the parameter. Notice that each parameter falls into either the 'not well-measured' category ( $K_{BC}$ ,  $\Delta H$ ,  $\Delta H_w$ ) or the 'large spread in measured values' category ( $K_{OA}$ ,  $H^*$ ) except  $k_{OH}$ , which is actually better constrained.

### Slide 11

Aside from values for the chemical parameters, the GEOS-Chem runs require prescribed meteorology and emissions of bap. For the meteorology, the GEOS-5 data set is used at  $4^\circ \times 5^\circ$  spatial resolution. Emissions are spatially resolved and based on bottom-up inventories, but represent a very important non-parameter uncertainty in bap simulations. The bap runs are performed "offline", which means that the other atmospheric species are prescribed (from a prior "full chemistry" simulation by GEOS-Chem) and therefore have the same evolution for all bap runs. This greatly decreases the computation time for the simulations, with the implicit assumption that atmospheric bap has no noticeable effect on the concentrations of the other atmospheric constituents. This is likely a good assumption since the atmospheric concentration of bap is tiny ( $<1 \text{ ng m}^{-3}$ ). For each model run used in this project (each with a unique set of parameter values), a 1 year spin-up is performed, followed by a simulation of 1 year (2005). The spin-up prevents the transient behaviour associated with temporal discontinuities in parameters from being included in the output used. The model output used was the logarithm of the Arctic-average annual mean total atmospheric bap concentration. (i.e. The time and space average of 'gas- and particle-phase' bap concentration). In total, 112 runs were performed: 28 for the 2<sup>nd</sup> order PCE fit and 84 more for the 3<sup>rd</sup> order. Since the model output spanned orders of magnitude, the PCE was used to estimate the logarithm of the output instead of the output itself, to prevent predictions of negative concentrations.

### Slides 12-13

These slides show the results of sampling the 2<sup>nd</sup> order and 3<sup>rd</sup> order polynomials using the parameter distributions shown earlier. Both resulting distributions show similar behaviour, with the 3<sup>rd</sup> order estimate having a sharper peak and a more asymmetric structure. The 3<sup>rd</sup> order distribution shows that the average Arctic concentration is very likely between about 0.0003 and 0.01 ng m<sup>-3</sup>.

### Slide 14

For those that believe that the uncertainty distributions of the parameters cannot possibly be as wide as those shown above, the 3<sup>rd</sup> order estimate was also sampled using reduced-width parameter distributions (the standard deviation of the log-normal distributions is halved for all but  $k_{OH}$ ). The distribution predictably sharpens, and the secondary bump on the more positive side of the peak disappears in this case, giving a likely range of around 0.001 and 0.01 ng m<sup>-3</sup>.

### Slide 15

This slide is unnecessary for those reading this document.

### Slide 16

The above results show that the uncertainty in Arctic concentration due to the large uncertainty in the model parameters is actually relatively low. This means that the uncertainty or errors in the emissions inventory used is potentially a very important contributor to the overall model uncertainty.

### Slide 17

To increase confidence in the result, another similar method that makes use of the same data could be used for comparison. The best candidate is a regression-based surface response estimate, which would use more collocation point / output pairs than terms in the polynomial, and use singular value decomposition to get the polynomial coefficients. At the time of writing, this method had not yet been tested.

To conclude, the result of the PCM's application to GEOS-Chem was a success. The model's response distribution is of the kind that would be expected and was relatively stable between 2<sup>nd</sup> and 3<sup>rd</sup> order

estimations. The resulting distribution gives information which will be very useful going forward with the simulation of BaP using GEOS-Chem. This information, in combination with a study of the output uncertainty due to emissions, will provide the necessary foundation for an assessment of which areas of the model will yield the greatest improvement from continuing work.

MIT OpenCourseWare  
<http://ocw.mit.edu>

12.S990 Quantifying Uncertainty  
Fall 2012

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.