Author response to all referees' comments on "Mapping the drivers of uncertainty in atmospheric selenium deposition with global sensitivity analysis"

We would like to thank the three reviewers for their time and helpful comments about the manuscript. We appreciate their interest in the manuscript and think that they have raised pertinent points about our analysis. We have considered these comments carefully and present our responses below, as well as changes to the main text (reviewer comment are in blue and author responses are in black).

Response to Referee #1

While it is likely not a factor for the main conclusions, the boundary conditions for the model represent the year 2000. How might inter-annual differences or recent trends in these assumed conditions propagate to outputs like atmospheric Se lifetime? For example, via particle abundance, meteorology, etc.

We agree with the reviewer's point that this is important to consider. Since we have only tested the sensitivity of model parameters in a certain year (2000), we are not able to comment on how the meteorology of different years would affect the atmospheric Se lifetime. We believe that the largest interannual change to consider is precipitation shifts and their effect on Se deposition, 80% of which is through wet deposition. For example, there was a La Niña event in the year 2000 (`https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php`), which would cause higher precipitation than normal over the Western Pacific and various teleconnections. We restricted the scope of our sensitivity analysis to Se cycle input parameters, which already cover a 34-dimensional space. To avoid computational expense, we did not investigate the dependencies of model output on climate-related parameters in this sensitivity analysis. However, we intend to consider the effects of inter-annual variability and long-term climate change in future studies. We added this caveat to the discussion, P32L10–12:

"It must be noted that our simulations were performed only for the year 2000 and focused on uncertainties in the Se-related input parameters, neglecting variations in the Se lifetime and deposition due to interannual variability in meteorology and sulfate aerosol properties. In future studies we intend to investigate how the Se cycle varies under different climate conditions."

In the absence of better information, the authors scale Se emissions spatially using the spatial distribution of S emissions. This is certainly a reasonable starting point. Given the large uncertainties in emissions magnitudes, how would the spatial uncertainty in S:Se ratio compare? The authors mention that the magnitude and spatial distribution of emissions are the most important uncertainties to constrain, but how do those two compare to each other?

This is a very good point. We followed a similar thought process as the reviewer, and considered this assumption to be a reasonable starting point for our sensitivity analysis. There are likely spatial uncertainties in the S:Se ratio, however the number of parameters we could consider in our sensitivity analysis was limited due to computational expense. Since the results showed that emissions are the most important to constrain for predicting Se deposition maps, we intend to investigate these spatial uncertainties in future studies. For natural sources, this is very difficult, due to a lack of studies. For example, there

is only one oceanic field study that has concurrently measured both S and Se fluxes (Amouroux et al., 2001). A current study is underway in our group to measure marine Se and S fluxes in various locations, which will improve our knowledge about these processes. In terms of anthropogenic variations of Se:S ratios, Lee et al. (2015) suggest that although the Se content of Chinese coal is double that of U.S. coal, S varies in a similar proportion and thus the Se:S ratios of both types of coal are similar. Nevertheless, there needs to be further research into whether regionally varying coal cleaning methods and pollution control technologies can alter the Se:S ratio of emissions.

The reactions modeled are assumed to have no temperature dependence (rates for 298K used). Most of the atmosphere, however, is significantly colder than this. How would the unquantified temperature dependence affect the given results?

As mentioned in the paper, we could not find any studies that measured the temperature dependence of the Se reaction rates, and therefore did not include this in the sensitivity analysis. We looked at the temperature dependence of the analogous S reactions (Burkholder et al., 2015) between 298 K and 240 K (covering most conditions of the lower troposphere, where we expect gaseous Se species to occur). Between these temperatures, the DMS + OH reaction rate decreases by 20%, DMS + NO_3 increases by 54%, H_2S + OH decreases by 6%, and OCS + OH decreases by 58%. These changes are all smaller than the uncertainty range that we assume for these reaction rates, spanning orders of magnitude. Therefore, we do not think that the unquantified temperature dependence would have a strong effect on our sensitivity analysis.

The range of photolysis frequencies are scaled from 0-2 times with a uniform distribution. Photolysis frequencies in general vary substantially between chemical species; why are these rates treated on a relatively narrower range than the reaction rates, which do vary over orders of magnitude?

As we mention in the paper, there is a lack of information about the uncertainties involved in the photolysis rates. Since the quantum yield of these photolysis reactions could theoretically be 0, we set the lower bound of the photolysis scaling factor as 0. To yield a range that is symmetrical around a unity scaling factor, we set the upper bound as 2. Since there is generally very little sensitivity of deposition fluxes to reaction rates (which we varied by orders of magnitude), we do not think that using a wider range for photolysis rates would cause these parameters to become influential.

Response to Referee #2

Thank you for your comment. The ranges used in this sensitivity analysis for Se emissions from different sources were listed in Table 3. We have now included a summary of past studies estimates in the supplementary material (Table S1).

The sources of each Se species are listed in Table 1. We vary the amount of emissions for each species in the sensitivity analysis, between the ranges shown in Table 3. Therefore, in each simulation there can be a different amount of each species emitted.

We base this conclusion on Figure 11, which shows the total Sobol' index averaged over different land types. Over the entire globe, marine emissions have the largest influence (green bars), due to oceans covering 70% of the Earth. However, if we only average over pasture (purple bars) or cropland (red bars) areas, then all emission sources show similar mean Sobol' indices between 0.15–0.30. We now clarify this on P28L15:

"All four emission parameters show similar levels of importance for agricultural regions, i.e. showing similar mean Sobol indices for pasture and cropland areas."

This point was brought up by Referee #1 as well, which we addressed above. We have now included in the discussion a caveat that our sensitivity analyses do not cover variations of the atmospheric Se lifetime due to interannual variability in climate and emissions.

Response to Referee #3

Additional information is required on the methods carried out to convince me that the surrogate model is suitable for the sensitivity analysis, particularly for Se Lifetime.

Thank you for your feedback, which made us consider again how we validated the surrogate models. We decided to run an additional validation set using the full SOCOL-AER model, to verify the suitability of our surrogate models. We chose the parameters for the 50 additional validation runs by enriching the existing training set so that a pseudo-Latin hypercube of 450 runs was formed (using the UQLAB function uq_LHSify). This ensures that our validation set is testing new regions of the parameter space compared to the training set.

In the figure below, we show the results of the validation runs compared to the predictions of the surrogate models for global Se burden, deposition flux, and lifetime. The compared values are shown as blue circles and the 1:1 line as a red line. The global lifetime surrogate model is calculated by dividing the PCE of the global burden by the PCE of the global deposition flux. We had found the LOO error of the burden PCE to be around 0.02 and the LOO of the deposition flux PCE to be on the order of 10^{-6}. We can compare these LOO errors to the validation errors, which are calculated from the independent validation set (Marelli and Sudret, 2019):

$$\text{Error}_{val} = \frac{N-1}{N} \left[\frac{\sum\limits_{i=1}^{N} \left(\mathcal{M}(\mathbf{x}_{val}^{(i)}) - \mathcal{M}^{PCE}(\mathbf{x}_{val}^{(i)}) \right)^2}{\sum\limits_{i=1}^{N} \left(\mathcal{M}(\mathbf{x}_{val}^{(i)}) - \hat{\mu}_{Y_{val}} \right)^2} \right]$$

where N is the number of validation runs, $\mathcal{M}(\mathbf{x}_{val})$ is the SOCOL-AER model output of the validation runs, $\mathcal{M}^{PCE}(\mathbf{x}_{val})$ is the prediction of the surrogate model for the sample points in the validation set, and $\hat{\mu}_{Y_{val}}$ is the sample mean of validation set output.
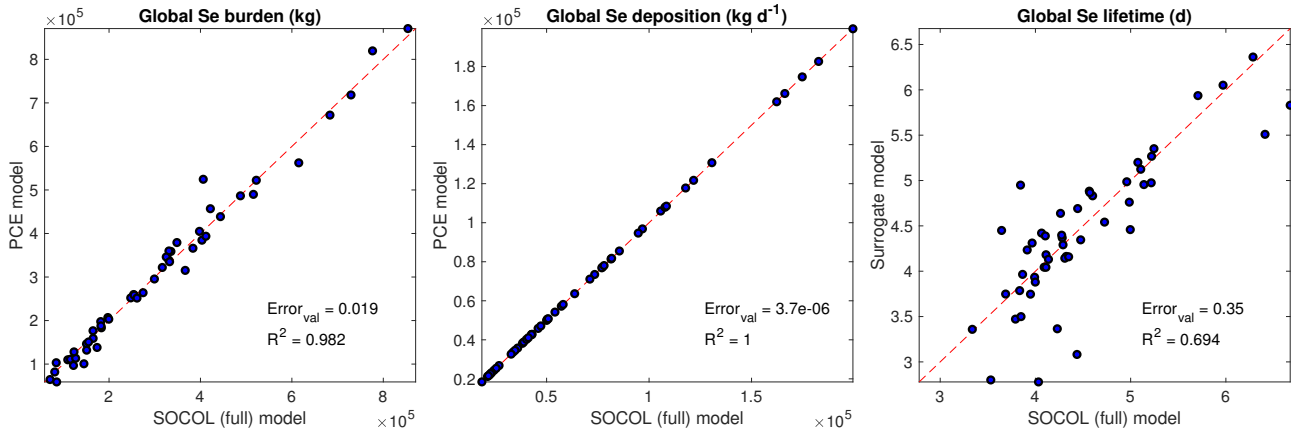
The validation errors match the LOO errors for global Se burden and Se deposition, showing that the LOO error is behaving correctly. However, as the referee suspected, the surrogate model for the global lifetime shows a relatively high validation error of around 0.35. The global selenium lifetime is clearly a

difficult parameter to emulate, likely requiring much more than 400 training runs to be accurately modelled. However, we think that the general results of the lifetime sensitivity analysis would not change, especially the identification of the most important parameters (OCSe reaction rates and dummy aerosols). We still think that the discussion of the selenium lifetime sensitivity analysis is valuable in the paper, especially as an introduction to the reader about a sensitivity analysis with one model output.

In the second sensitivity analysis of the paper, we looked into the factors that affect Se deposition in each grid box. The independent validation dataset verifies the accuracy of the deposition flux surrogate models. We compare the simulated deposition in the 50 validation runs at all 8192 horizontal grid boxes with the PCE-predicted values, i.e. $50 \times 8192 = 409\,600$ points. The left plot shows the results in linear space and the right plot in logarithmic space, since in the linear plot the smaller deposition fluxes collapse around 0. Most points fall around the 1:1 line, with the validation error being 0.013. Therefore, these surrogate models have much better accuracy than the lifetime surrogate model and are not affected by the issues brought up by the reviewer.
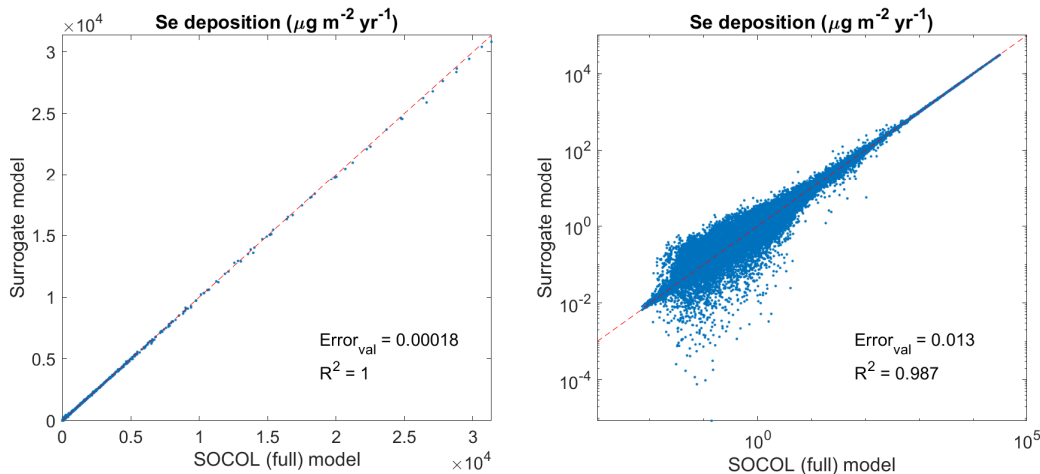
We have added these results to the supplementary information of the paper. In the main paper, we have added in the methods section how we produced the validation dataset (P18L9–13):

"The cross-validation approach would usually remove the need for an independent validation dataset, saving computational expense. However, to evaluate the post-processing steps applied to the surrogate models, we also produced an independent validation dataset of 50 SOCOL-AER runs. The parameters for these runs were chosen by enriching the training experimental design to create a pseudo-Latin Hypercube of 450 runs, ensuring that the distance between the validation runs and existing training runs is maximal."

We also now acknowledge the accuracy of the lifetime surrogate model compared to the validation dataset in the main paper (P21L24–26):

"We derive a surrogate model for the atmospheric Se lifetime by dividing the Se burden PCE model by the Se deposition PCE model (Sect. 3.3). This surrogate model shows a higher error (0.35) than the burden and deposition flux PCE models, which would only be reduced by running more training runs (Fig. S2)."

There are some choices made in the application of the polynomial chaos expansion that are not explained and it's difficult to know whether they are rule of thumb decisions or made specifically for this application. How was q=0.75 decided upon and what other choices were there to least-angle regression?

The choice of q=0.75 is arbitrary and based on our own experience. The range [0.7-0.8] seems to be a good compromise between computational expense and accuracy of the PCE model. We now rephrase this on P16L27:

"Based on previous experience, we selected a q value of 0.75".

I don't follow the logic of explaining LOO validation and then using a model that doesn't actually leave any of the training runs out. Can you explain why this is appropriate and how the validation might change if you had used LOO? Is it appropriate to continue to call it LOO validation with a full model? It may be explained in one of the given references but I'd like to see some explanation here.

To calculate the LOO error, the process of leaving out one training point and calculating a new PCE model is repeated N times, where N is the size of the training dataset. At each stage of the LOO error calculation, the same basis set of polynomials is used to calculate all PCEs. The final basis set that is selected by the algorithm is the one that minimizes the LOO error, ensuring that the PCE calculated with the full training set does not overfit the data. Previous studies have found that the LOO error performs well in predicting the generalization error of a statistical model (Molinaro et al., 2005; Blatman and Sudret, 2010), which we also found when comparing the LOO errors to our independent validation set errors. The LOO error for a PCE model would be calculated using Eq. 7 (in the paper), which requires the calculation of N PCE models. An equivalent form of Eq. 7 is used (Eq. 8), which requires the calculation of a single PCE model.
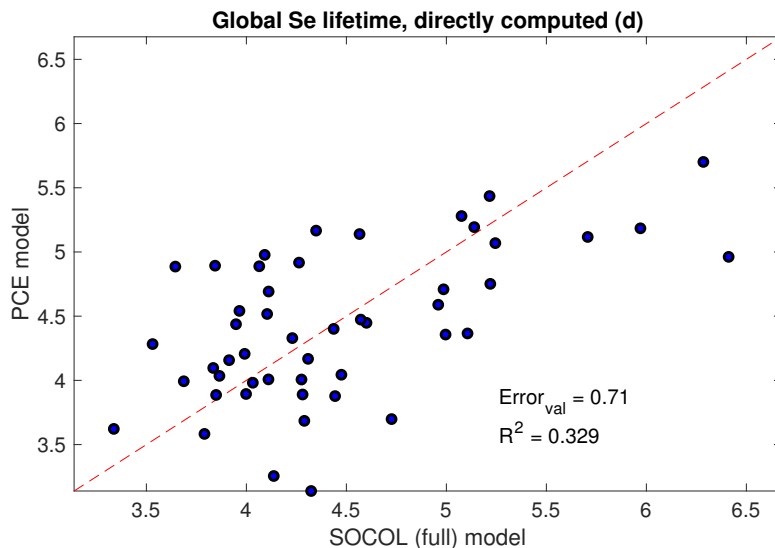
You choose the degree based on whether the validation doesn't decrease in the next step. Why and how does the validation increase? I would expect the validation to always decrease to some extent with extra terms. If a threshold is applied, what is it?

Technically, the error should decrease as the degree increases. However, in practice the sample size is given and does not change as we increase the degree. The number of coefficients to estimate does increase together with the polynomial degree but in the process we do not bring any new information (no additional training points). Therefore, it becomes increasingly difficult to estimate the coefficients with the limited available information. On top of that, the least-square solution involves inverting a matrix which leads to a larger numerical error as the matrix size increases. For these reasons, the error starts to increase at a certain degree. The algorithm implemented in UQLab is aware of these issues and thus stops the algorithm when it realizes that the error is not decreasing despite increasing the degree for two successive iterations.

The output of interest, the atmospheric Se lifetime, is of the form $Y = \frac{a}{b}$, where a is the global Se burden and b is the deposition flux. If a and b are linear or can easily be approximated by polynomials it is better to do so than trying to approximate a function Y, which is possibly of degree order $-1, -2$, etc. (which would eventually be possible but would require a larger experimental design).

We can confirm this approach with our new independent validation set. We produced a PCE of the Se lifetime directly, and compared this PCE's predictions with the 50 validation runs (figure shown below). The validation error, 0.71, is worse than the validation error we found when computing PCEs for a and b separately and calculating the lifetime from these two PCEs, which had a validation error 0.35 (compare to lower right figure on page 4). Therefore, this approach of calculating separate surrogate models for burden and deposition succeeded in improving the accuracy of the surrogate lifetime model.

Global Se lifetime, directly computed (d)

$\text{Error}_{val} = 0.71$

$R^2 = 0.329$

(y-axis: PCE model; x-axis: SOCOL (full) model)

We decided to include a general description of how the sensitivity indices can be derived from the PCE, while for a more mathematical description the reader can refer to the Sudret (2008) paper. We now include on P20L5–9:

"As shown by Sudret (2008), Sobol sensitivity indices are a function of the calculated coefficients in the PCE model (Eq. 4), due to the similarity of the PCE decomposition with variance decomposition. First order Sobol' indices can be calculated as $D_i = \sum_{\alpha \in \mathcal{A}_i} y_\alpha^2$, where $\mathcal{A}_i$ is the set of polynomial terms involving only variable i. Similarly, higher-order Sobol' indices can be calculated by summing the squares of the coefficients of the polynomial terms that include the variables of interest."

I'm not convinced that double counting interactions because they are less than 0.05 is a good idea. 0.05 is quite a large fraction of the 0.15 that is the largest main effect. Can you find a way to investigate the effect this is having?

As we acknowledge in the paper (P20L16–17), it is an upper limit for the overall dummy aerosol effect. In order to calculate the effect this would have, we would have to sum the individual Sobol' indices. We are limited in this by numerical issues when the Sobol' indices are calculated by Monte Carlo methods. However, we can calculate this when the surrogate model is a PCE, since the Sobol' indices can be calculated analytically. For illustrative purposes, we use a PCE that was directly trained by the global Se lifetime (i.e. showing a higher validation error than the surrogate model we use in the paper). For this PCE, we can accurately account for all individual indices, and find a total dummy aerosol Sobol' index of 0.20. This can be compared to the aggregated index that we would have calculated from summing the total indices of the aerosol parameters, 0.23. We do not find a large shift in the total index, and we believe that our approach in the paper is appropriate for summarizing the overall effect of the dummy aerosols, given the surrogate models that we use.

In Figure 5, the main effects are quite low and nowhere near adding up to 1 – was the amount of interaction in this model expected?

We did expect that many non-linear terms involving multiple input parameters could be involved in the Se lifetime, due to the structure of the model. There are many examples that we could provide of interactive effects that could occur. For example if there are low emissions of a certain substance, its rate constant will not matter in the sensitivity analysis. If there are higher emissions of a certain substance, its rate constant could affect the lifetime more, depending on its order of magnitude. Additionally, the dummy aerosol parameters of radius and emissions also would contribute to the variance interactively, as discussed in the paper.

Still with Figure 5, in previous experience seeing interactions that are large and consistent between multiple variables is a sign that the model fit is actually poor. It's not clear because of the way you have carried out LOO and added the interaction terms whether this is indicating poor model fit or whether these are real interactions.

As we showed above, the surrogate model for lifetime does not perfectly capture the model behaviour, however it still shows reasonable accuracy. The large interaction effect is not necessarily a sign of poor model fit, but it could be an intrinsic reason why it is difficult to produce a surrogate model of the lifetime with the quantity of training runs we ran (400, more than the usual rule of thumb of 10 times the number of parameters). More training runs are likely needed to fully capture the interactive behaviour involved in modelling the Se lifetime.

We have shown above the validation error of our surrogate model, which is not insignificant (0.35). This error could be reduced by running more training runs for the PCE models, however it is not clear how quickly the error will decrease with additional runs. Given that our size of the training set is already quite large and computationally expensive, we had to accept the current accuracy of the surrogate model and proceed with our analysis. In the paper, the sensitivity analysis of the global Se lifetime focuses on the distribution of the simulated lifetime, the ranking of the most important parameters, and visualizing relationships between the lifetime and the input parameters. We do not believe that the main conclusions of our lifetime analysis would change with an expanded training set, even though the Sobol' indices would be refined. The sensitivity analysis of Se deposition would not be affected by these issues, since the LOO and validation errors are, in general, on the order of 10^{-2}.

References

Amouroux, D., Liss, P. S., Tessier, E., Hamren-Larsson, M., and Donard, O. F.: Role of oceans as biogenic sources of selenium, Earth and Planetary Science Letters, 189, 277–283, URL https://doi.org/10.1016/S0012-821X(01)00370-3, 2001.

Blatman, G. and Sudret, B.: An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis, Probabilistic Engineering Mechanics, 25, 183–197, URL https://doi.org/10.1016/j.probengmech.2009.10.003, 2010.

Burkholder, J., Sander, S., Abbatt, J., Barker, J., Huie, R., Kolb, C., Kurylo, M., Orkin, V., Wilmouth, D., and Wine, P.: Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies, Evaluation No. 18, Tech. Rep. 15-10, Jet Propulsion Laboratory, Pasadena, URL https://jpldataeval.jpl.nasa.gov/pdf/JPL_Publication_15-10.pdf, last access: 30 August 2019, 2015.

Lee, K., Hong, S.-B., Lee, J., Chung, J., Hur, S.-D., and Hong, S.: Seasonal variation in the input of atmospheric selenium to northwestern Greenland snow, Science of the Total Environment, 526, 49–57, URL http://doi.org/10.1016/j.scitotenv.2015.04.082, 2015.

Marelli, S. and Sudret, B.: UQLab user manual – Polynomial chaos expansions, Tech. rep., Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich, report # UQLab-V1.3-10, 2019.

Molinaro, A. M., Simon, R., and Pfeiffer, R. M.: Prediction error estimation: a comparison of resampling methods, Bioinformatics, 21, 3301–3307, URL https://doi.org/10.1093/bioinformatics/bti499, 2005.

Sudret, B.: Global sensitivity analysis using polynomial chaos expansions, Reliability Engineering & System Safety, 93, 964–979, URL `https://doi.org/10.1016/j.ress.2007.04.002`, 2008.

Mapping the drivers of uncertainty in atmospheric selenium deposition with global sensitivity analysis

Aryeh Feinberg[1,2,3], Moustapha Maliki[4], Andrea Stenke[1], Bruno Sudret[4], Thomas Peter[1], and Lenny H. E. Winkel[2,3]

[1]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[2]Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, Zurich, Switzerland
[3]Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland
[4]Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich, Zurich, Switzerland

Correspondence: Aryeh Feinberg (aryeh.feinberg@env.ethz.ch)

Abstract. An estimated 0.5–1 billion people globally have inadequate intakes of selenium (Se), due to a lack of bioavailable Se in agricultural soils. Deposition from the atmosphere, especially through precipitation, is an important source of Se to soils. However, very little is known about the atmospheric cycling of Se. It has therefore been difficult to predict how far Se travels in the atmosphere and where it deposits. To answer these questions, we have built the first global atmospheric Se model by implementing Se chemistry into an aerosol–chemistry–climate model, SOCOL-AER. In the model, we include information from the literature about the emissions, speciation, and chemical transformation of atmospheric Se. Natural processes and anthropogenic activities emit volatile Se compounds, which oxidize quickly and partition to the particulate phase. Our model tracks the transport and deposition of Se in 7 gas-phase species and 41 aerosol tracers. However, there are large uncertainties associated with many of the model's input parameters. In order to identify which model uncertainties are the most important for understanding the atmospheric Se cycle, we conducted a global sensitivity analysis with 34 input parameters related to Se chemistry, Se emissions, and the interaction of Se with aerosols. In the first bottom-up estimate of its kind, we have calculated a median global atmospheric lifetime of 4.4 d (days), ranging from 2.9–6.4 d (2nd–98th percentile) given the uncertainties of the input parameters. The uncertainty in the Se lifetime is mainly driven by the uncertainty in the carbonyl selenide (OCSe) oxidation rate and the lack of tropospheric aerosol species other than sulfate aerosols in SOCOL-AER. In contrast to uncertainties in Se lifetime, the uncertainty in deposition flux maps are governed by Se emission factors, with all four Se sources (volcanic, marine biosphere, terrestrial biosphere, and anthropogenic emissions) contributing equally to the uncertainty in deposition over agricultural areas. We evaluated the simulated Se wet deposition fluxes from SOCOL-AER with a compiled database of rainwater Se measurements, since wet deposition contributes around 80% of total Se deposition. Despite difficulties in comparing a global, coarse resolution model with local measurements from a range of time periods, past Se wet deposition measurements are within the range of the model's 2nd–98th percentile at 79% of background sites. This agreement validates the application of the SOCOL-AER model to identifying regions which are at risk of low atmospheric Se inputs. In order to constrain the uncertainty in Se deposition fluxes over agricultural soils we should prioritize field campaigns measuring Se emissions, rather than laboratory measurements of Se rate constants.

1 Introduction

Selenium (Se) is an essential dietary trace element for humans and animals, with the recommended intake ranging between 30 and $900\,\mu g\,d^{-1}$ (Fairweather-Tait et al., 2011). The amount of Se in crops depends on the amount of bio-available Se in the soils where the crops are grown (Winkel et al., 2015). Levels of Se in soils, as well as Se dietary intakes, vary strongly around the world (Jones et al., 2017). Selenium deficiency is considered a more widespread issue than Se toxicity, as around 0.5 to 1 billion people are estimated to have insufficient Se intakes (Fordyce, 2013)(Combs, 2001; Fordyce, 2013).

Atmospheric deposition is an important source of Se to soils. In several regions, Se concentrations in soils were found to correlate with precipitation (Låg and Steinnes, 1978; Blazina et al., 2014). As well, several studies have attributed an increase in soil Se concentrations to regional anthropogenic Se emissions to the atmosphere (Haygarth et al., 1993; Dinh et al., 2018), suggesting a link between atmospheric Se inputs and soil concentrations. However, apart from some budget studies in the 1980's (Ross, 1985; Mosher and Duce, 1987), there has been a lack of research into global-scale atmospheric cycling of Se and the spatial patterns of Se deposition.

1.1 Atmospheric Se cycle

Since Se and sulfur (S) are in the same group on the periodic table, they share chemical properties and their biogeochemical cycles are similar. Like S, Se is emitted to the atmosphere by both natural and anthropogenic sources, with the total annual emissions estimated between 13 and $19\,Gg\,Se\,yr^{-1}$. Natural sources of atmospheric Se include volatilization by the marine and terrestrial biospheres, volcanic degassing and eruptions, and minor contributions from sea salt and mineral dust. Anthropogenic Se is emitted during coal and oil combustion, metal smelting, and biomass burning. However, very few in-situ measurements of Se emissions fluxes and speciation exist. Once in the atmosphere, volatile Se species are oxidized, eventually forming species like elemental Se and SeO_2 (Wen and Carignan, 2007). These oxidized species are expected to partition to the particulate phase; previous measurements have found that 75–95% of Se is in particulates (Mosher and Duce, 1983, 1987). The fate of atmospheric Se is dry and wet deposition, with wet deposition accounting for an estimated 80% of total deposition globally (Wen and Carignan, 2007).

Atmospheric chemistry modeling studies have been applied to other trace elements to predict atmospheric lifetimes and spatial patterns of deposition. For example, atmospheric mercury models were developed more than two decades ago (Petersen and Munthe, 1995), and now there are around 16 global and regional atmospheric mercury models (Ariya et al., 2015). A recent mechanistic modeling paper has advanced the understanding of atmospheric arsenic cycling (Wai et al., 2016). To our knowledge, Se chemistry has never previously been included in an atmospheric chemistry–climate model (CCM), and thus many questions surrounding atmospheric Se transport remain unanswered. For example, the atmospheric lifetime of Se and thus the scales at which it can be transported (local, regional, hemispheric, global) are not well constrained.

1.2 Global sensitivity analysis

Since the atmospheric Se cycle has been investigated only by a limited number of studies, it is essential that we consider the relevant parametric uncertainties when building an atmospheric Se model. The reaction rate coefficients of Se compounds have either only been measured by one laboratory study, or no laboratory measurement exists and these rate coefficients need to be estimated. Selenium emission fluxes from certain sources have been measured, however it remains difficult to extrapolate these measurements to global fluxes due to the high degree of spatial and temporal variability. Atmospheric Se modeling can only be considered trustworthy when combined with full accounting of input parameter uncertainties and their propagation through the model. Through "global sensitivity analysis" (Saltelli et al., 2008) we can identify which input uncertainties are the most important for the uncertainty in the model output. A sensitivity analysis is called *global* when the sensitivity is evaluated over the entire input parameter space, as opposed to *local* methods that test sensitivity only at a certain reference point in the space (i.e., based on the gradient of the output at this reference point). Sensitivity analysis provides a framework to prioritize which model inputs should be further constrained in order to reduce the uncertainty in the model output.

Until recently, most sensitivity analyses of atmospheric chemistry models consisted of local methods, principally the one-at-a-time approach (OAT). In OAT, the model is initially run with a set of default parameters to yield a "reference" simulation. Multiple sensitivity simulations are then conducted, so that for each simulation one parameter is perturbed from the reference set at a time. The influence of these perturbations on the model output of interest would then be analyzed. However, this approach may be flawed because it only considers the first order response of the model to each parameter, ignoring interactions that might exist between parameters (Saltelli et al., 2008; Lee et al., 2011). As well, the uncertainty ranges of the input parameters are rarely quantified and reported; much of the possible parameter space often remains unexplored. Global methods such as variance-based sensitivity analysis allow the uncertainty in model output to be apportioned to each input variable. Sobol' indices, which represent the fraction of model variance that one input variable explains, provide a ranking system for the importance of input variables (Sobol, 1993). The benefits of global sensitivity analysis include: 1) identifying the most influential input variables, i.e., the ones that should be further constrained to yield the biggest reduction in model uncertainty; 2) identifying input variables that do not play any role in the model output variance; this could represent a route to simplify the model, since the process involving these input variables can be neglected; 3) understanding the behavior of the model, for example how the output depends on interactions between input variables; 4) identifying possible model bugs or discontinuities, since the model will be tested with a wide range of input parameter values (Saltelli et al., 2000; Ryan et al., 2018).

There are several recent examples of atmospheric chemistry studies that included a global sensitivity analysis. The sensitivity of cloud condensation nuclei number density to input parameters in an aerosol model was investigated at the local (geographic) scale (Lee et al., 2011) and at the global scale (Lee et al., 2012, 2013), revealing regional differences in parameter rankings. Revell et al. (2018) investigated the sensitivity of the tropospheric ozone columns to emission and chemical parameters, to identify which processes are responsible for the bias in modeled tropospheric ozone. Marshall et al. (2019) employed global sensitivity analysis methods to identify how radiative forcing responds to volcanic emission parameters. In these examples, surrogate modeling techniques (also known as emulation) were employed to replace a process-oriented, computationally ex-

pensive model with an approximative statistical model. The statistical model has the advantage that it is quicker to evaluate; therefore, it can be used to calculate the model output throughout the parametric space (Lee et al., 2011). The examples given above all used Gaussian process emulation, however other surrogate modeling techniques exist, including polynomial chaos expansions (PCE) (Ghanem and Spanos, 2003). The PCE approach is well-suited to sensitivity analysis, since the Sobol' sen-

5   sitivity indices can be extracted analytically from the constructed PCE, with no need to evaluate the surrogate model through Monte Carlo sampling (Sudret, 2008). This can greatly reduce the computational time required to conduct the sensitivity analysis, especially when one is interested in conducting a separate sensitivity analysis for many model grid boxes.

1.3   Outline

This study focuses on the construction of the first global atmospheric Se model and the insights that this model reveals into

10   which Se cycle uncertainties would be the most important to constrain. Section 2 describes the SOCOL-AER model and the implementation of Se chemistry in the SOCOL-AER model. The SOCOL-AER model is a suitable tool to model the Se cycle, since it successfully describes the major properties of the atmospheric S cycle (Sheng et al., 2015; Feinberg et al., 2019b). The statistical methods that we use to conduct the sensitivity analysis are discussed in Sect. 3. Section 4 details the methods used to compile a database of measured wet Se deposition fluxes, which we use to evaluate the model. The results of the

15   sensitivity analyses are presented in Sect. 5.1 for the atmospheric Se lifetime and Sect. 5.2 for the Se deposition patterns. Section 5.3 illustrates the comparison between the compiled Se deposition measurements and simulated results. A discussion of both sensitivity analyses follows in Sect. 6, and the paper is concluded in Sect. 7.

2   Model Description

2.1   SOCOL-AER

20   SOCOL-AERv2 is a global CCM that includes a coupled sulfate aerosol microphysical scheme (Sheng et al., 2015; Feinberg et al., 2019b). The base CCM, SOCOLv3 (Stenke et al., 2013), is a combination of the general circulation model ECHAM5 (Roeckner et al., 2003) and the chemical model MEZON (Egorova et al., 2003). The MEZON submodel comprises a comprehensive atmospheric chemistry scheme, with 89 gas-phase chemical species, 60 photolysis reactions, 239 gas-phase reactions, and 16 heterogeneous reactions. Chemical tracers are advected in the model using the Lin and Rood (1996) semi-Lagrangian

25   method. Photolysis rates are calculated using a look-up table approach based on the simulated overhead ozone and oxygen columns. The MEZON model solves the system of differential equations representing chemical reactions with a Newton-Raphson iterative method for short-lived chemical species and an Euler method for long-lived species.

     The sulfate aerosol model AER (Weisenstein et al., 1997) was first coupled to SOCOL by Sheng et al. (2015). SOCOL-AER includes gas-phase S chemistry and 40 sulfate aerosol tracers, ranging in dry radius size from 0.39 nm to 3.2 μm. SOCOL-

30   AER simulates microphysical processes that affect the aerosol size distribution, including binary homogeneous nucleation, condensation of H_2SO_4 and H_2O, coagulation, evaporation, and sedimentation. SOCOL-AER was extended in Feinberg et al.

(2019b) to include interactive wet and dry deposition schemes. The wet deposition scheme, based on Tost et al. (2006), calculates scavenging of gas-phase species depending on their Henry's law coefficients and aerosol species depending on the particle diameter. The wet removal of tracers is coupled to the grid cell simulated properties of clouds and precipitation. The dry deposition scheme is based on Kerkweg et al. (2006, 2009), which uses the surface resistance approach of Wesely (1989).

5 In addition to surface type and meteorology, the calculated dry deposition velocities depend on reactivity and solubility for gas-phase compounds and size for aerosol species. SOCOL-AER uses an operator splitting approach, wherein the model time step is 2 hours for chemistry and radiation and 15 minutes for dynamics and deposition. Aerosol microphysical routines use a sub-time step of 6 minutes.

For the simulations in this study we use boundary conditions for the year 2000. Sea ice coverage and sea surface temperatures

10 are prescribed from the Hadley Centre dataset (Rayner et al., 2003). Year 2000 concentrations of the most relevant greenhouse gases (CO_2, CH_4, and N_2O) derive from NOAA observations (Eyring et al., 2008). Anthropogenic CO and NO_x emissions are based on the RETRO dataset (Schultz and Rast, 2007), while natural emissions are taken from Horowitz et al. (2003). Sulfur dioxide (SO_2) emissions from anthropogenic sources follow the year 2000 inventory from Lamarque et al. (2010) and Smith et al. (2011). Volcanic degassing SO_2 emissions are assigned to surface grid boxes where volcanoes are located (Andres and

15 Kasgnoc, 1998; Dentener et al., 2006). Atmospheric emissions of dimethyl sulfide (DMS) are calculated using a wind-based parametrization (Nightingale et al., 2000) and a marine DMS climatology (Lana et al., 2011). To represent mean conditions for photolysis, the look-up table for photolysis rates is averaged over two solar cycles (1977–1998).

2.2 Implementing Se emissions and chemistry in SOCOL-AER

2.2.1 Selenium species overview

20 We included the Se cycle in SOCOL-AER (Fig. 1) based on the existing literature on atmospheric Se (Ross, 1985; Wen and Carignan, 2007). Seven Se gas-phase tracers have been added to SOCOL-AER (Table 1): carbonyl selenide (OCSe), thiocarbonyl selenide (CSSe), carbon diselenide (CSe_2), dimethyl selenide (DMSe), hydrogen selenide (H_2Se), oxidized inorganic Se (OX_Se_IN), and oxidized organic Se (OX_Se_OR). The oxidized inorganic Se tracer represents species such as elemental Se, selenium dioxide (SeO_2), selenous acid (H_2SeO_3), and selenic acid (H_2SeO_4). Very little is known about the kinetics of inter-

25 conversion between the oxidized inorganic Se species (Wen and Carignan, 2007), and therefore in our model they are treated as one tracer. However, these species all have very low vapor pressures under atmospheric conditions (Rumble, 2017) and likely partition to the particulate phase. Oxidized organic Se species include dimethyl selenoxide (DMSeO) and methylseleninic acid (CH_3SeO_2H), which form after oxidation of DMSe (Atkinson et al., 1990; Rael et al., 1996). Similar to the oxidized inorganic Se compounds, oxidized organic Se species also partition to the particulate phase due to their low volatilities (Rael et al., 1996).

30 ### 2.2.2 Selenium emissions

To determine which Se compounds are emitted by the different sources, we have reviewed studies that investigated the speciation of Se emissions. Thermodynamic modeling and in situ measurements of combustion exhaust gases have detected the
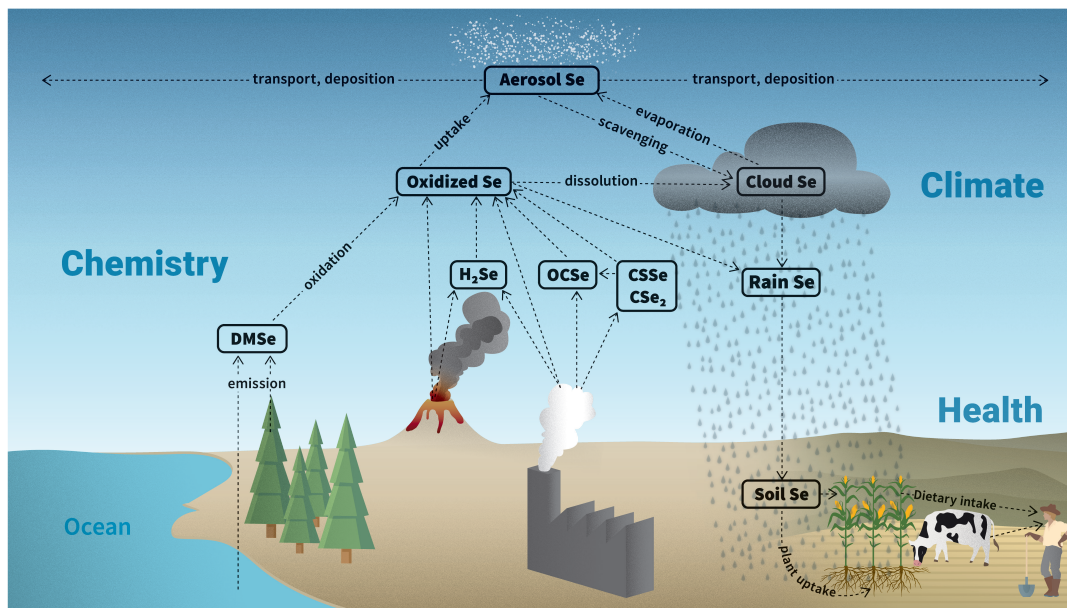
Figure 1. Scheme of the atmospheric Se cycle in SOCOL-AER, based on information in Ross (1985); Wen and Carignan (2007). For simplicity the two oxidized Se tracers in SOCOL-AER are represented with a single box. The impact on agriculture and human health is also shown, since it motivates the study of atmospheric Se.

following Se species in anthropogenic emissions: oxidized inorganic Se, H_2Se, OCSe, CSe_2, and CSSe (Yan et al., 2000, 2004; Pavageau et al., 2002, 2004). Oxidized inorganic Se species and minor amounts of H_2Se are expected to be emitted from volcanic degassing (Symonds and Reed, 1993; Wen and Carignan, 2007). A variety of methylated Se species have been observed from biogenic emissions, including DMSe, dimethyl diselenide (DMDSe), dimethyl selenylsulfide (DMSeS), and methane se-

5 lenol (MeSeH) (Amouroux and Donard, 1996, 1997; Amouroux et al., 2001; Wen and Carignan, 2007). Since DMSe is usually the dominant emitted compound, and little is known about the oxidation kinetics of the other methylated species, DMSe is the only species emitted by marine and terrestrial biogenic emissions in SOCOL-AER.

Atmospheric S emissions have been measured more extensively than Se emissions, so we scale inventories of S emissions to yield the spatial distribution of emitted Se. For the sensitivity analysis we assume that the Se emissions have the same

10 distribution as S emissions, and we focus on the uncertainties in the global scaling factors for each source. The range in scaling factors will be discussed in Sect. 3.1.4. The spatial distribution of anthropogenic and biomass burning Se emissions comes from the SO_2 inventory for the year 2000 (Lamarque et al., 2010; Smith et al., 2011). For sea surface DMSe concentrations, we scale the DMS climatology calculated by Lana et al. (2011), and calculate DMSe emissions using a wind-driven parametrization (Nightingale et al., 2000). The locations and strength of background degassing volcanic emissions are taken from the GEIA

15 inventory (Andres and Kasgnoc, 1998; Dentener et al., 2006). Since little is known about both terrestrial biogenic Se emissions and terrestrial S emissions (Pham et al., 1995), we assume that terrestrial Se is emitted in all land surface grid boxes, excluding glaciated locations.

Table 1. Description of Se tracers included in SOCOL-AER

Abbreviation	Name	Sources	Sinks
OCSe	Carbonyl selenide	Anthropogenic emissions, chemical production	Chemical loss
CSSe	Thiocarbonyl selenide	Anthropogenic emissions	Chemical loss
CSe_2	Carbon diselenide	Anthropogenic emissions	Chemical loss
DMSe	Dimethyl selenide	Marine and terrestrial emissions	Chemical loss
H_2Se	Hydrogen selenide	Anthropogenic and volcanic emissions	Chemical loss
OX_Se_IN	Oxidized inorganic Se	Anthropogenic and volcanic emissions, chemical production	Dry and wet deposition
OX_Se_OR	Oxidized organic Se	Anthropogenic and volcanic emissions, chemical production	Dry and wet deposition
-	Se in S aerosol (40 tracers)	Uptake of gas-phase oxidized Se	Dry and wet deposition
-	Se in dummy aerosol	Uptake of gas-phase oxidized Se	Dry and wet deposition

2.2.3 Chemistry of Se species

We conducted a literature review to develop the model's chemical scheme of the Se cycle. Reactions of Se species that have been measured by laboratory studies are compiled in Table 2. We neglect any temperature dependency in the Se reaction rates, since the Se reactions have only been studied at around 298 K. For all compiled reactions, atmospheric Se compounds react much quicker than the analogous S compounds, due to the stronger bonds that S forms with carbon and hydrogen than Se (Rumble, 2017). In addition, there are reactions that are known to occur for the analogous S compound, but have never been studied for the Se compound (OCSe + OH, CSSe + OH, CSe_2 + OH, and H_2Se + OH). These reaction rate constants were estimated in Fig. 2, which shows the ratio of the Se compounds' rates to analogous S rates (i.e., an enhancement factor for replacing an S atom with Se) plotted versus the S reaction rate. For S reactions which have a fast reaction rate (e.g., DMS + Cl, $k = 1.8 \times 10^{-10}$ $\mathrm{cm^3\,molec^{-1}\,s^{-1}}$), replacing S with Se does not yield a large difference in measured rates (DMSe + Cl, $k = 5.0 \times 10^{-10}$ $\mathrm{cm^3\,molec^{-1}\,s^{-1}}$). This is because these reactions are already close to the collision-controlled limit, and thus lowering the activation energy by substituting a Se atom for S has little impact on the overall rate. On the other hand, slow reactions like DMS + O_3 are sped up by more than four orders of magnitude when Se is substituted for S. We used the log-log relationship in Fig. 2 to predict the reaction rates for OCSe, CSSe, and H_2Se with OH (blue stars). The CSe_2 reaction with OH is calculated from the CSSe reaction, assuming a similar enhancement for the substitution of a second Se atom as between the measured CSe_2 + O and CSSe + O reaction rates (Li et al., 2005). The branching ratio for the CSSe + OH reaction products was assumed to be 30% OCSe and 70% OX_Se_IN, the same as the measured CSSe + O branching ratio (Li et al., 2005). We recognize that these estimates are inherently uncertain, and therefore address these uncertainties in our sensitivity analysis (Sect. 3.1.2).

The photolysis of gas-phase Se compounds was included using absorption cross sections of H_2Se (Goodeve and Stein, 1931) and OCSe (Finn and King, 1975). The absorption cross section of CSe_2 (King and Srikameswaran, 1969) has been measured, however in too low resolution to be incorporated into the model. Therefore, we assume that CSe_2 and CSSe have the same

Table 2. Rate constants of Se compound gas-phase reactions at around 298 K and the corresponding rate constant of the analogous S compound. All S reaction rates are from Burkholder et al. (2015), except the DMS + O_3 reaction rate which is taken from Wang et al. (2007). No corresponding rate constants for CSe_2 reactions are listed, since CSe_2 is obtained from doubly substituting Se in CS_2.

Reaction	Se rate constant $(cm^3 \, molec^{-1} \, s^{-1})$	Corresponding S constant $(cm^3 \, molec^{-1} \, s^{-1})$	Reference for Se rate constant
Measured reactions			
$OCSe + O \rightarrow CO + OX\_Se\_IN$	2.4×10^{-11}	1.3×10^{-14}	Li et al. (2005)
$CSe_2 + O \rightarrow OCSe + OX\_Se\_IN$ (32%) $\rightarrow 2\,OX\_Se\_IN$ (68%)	1.4×10^{-10}	-	Li et al. (2005)
$CSSe + O \rightarrow OCSe$ (30%) $\rightarrow OX\_Se\_IN$ (70%)	2.8×10^{-11}	3.6×10^{-12}	Li et al. (2005)
$DMSe + OH \rightarrow OX\_Se\_OR$	6.8×10^{-11}	6.7×10^{-12}	Atkinson et al. (1990)
$DMSe + NO_3 \rightarrow OX\_Se\_OR$	1.4×10^{-11}	1.1×10^{-12}	Atkinson et al. (1990)
$DMSe + O_3 \rightarrow OX\_Se\_OR$	6.8×10^{-17}	2.2×10^{-21}	Atkinson et al. (1990)
$DMSe + Cl \rightarrow OX\_Se\_OR$	5.0×10^{-10}	1.8×10^{-10}	Thompson et al. (2002)
$H_2Se + O \rightarrow OX\_Se\_IN$	2.1×10^{-12}	2.2×10^{-14}	Agrawalla and Setser (1987)
$H_2Se + Cl \rightarrow OX\_Se\_IN$	5.5×10^{-10}	7.4×10^{-11}	Agrawalla and Setser (1986)
$H_2Se + O_3 \rightarrow OX\_Se\_IN$	3.2×10^{-16}	$< 2.0 \times 10^{-20}$	Belyaev et al. (2012)
Estimated reactions			
$OCSe + OH \rightarrow OX\_Se\_IN$	5.8×10^{-13}	2.0×10^{-15}	Estimated, see text
$CSe_2 + OH \rightarrow OCSe + OX\_Se\_IN$	1.5×10^{-10}	-	Estimated, see text
$CSSe + OH \rightarrow OX\_Se\_IN$ (70%) $\rightarrow OCSe$ (30%)	3.0×10^{-11}	1.2×10^{-12}	Estimated, see text
$H_2Se + OH \rightarrow OX\_Se\_IN$	7.2×10^{-11}	4.7×10^{-12}	Estimated, see text

cross section as CS_2 (Burkholder et al., 2015). Given the lack of available information, quantum yields for all Se photolysis reactions were assumed to be 1.

2.2.4 Condensation of Se on preexisting aerosol particles

As nonvolatile species, oxidized inorganic and organic Se would condense on available atmospheric surfaces. In the SOCOL-AER model, the uptake of these oxidized Se species by sulfate aerosols is calculated similarly to the existing scheme of gas-phase H_2SO_4 uptake on sulfate particles (Sheng et al., 2015). We track the size distribution of Se in the aerosol phase with 40 tracers, one for each sulfate aerosol size bin. The sulfate aerosol size distribution changes through processes like growth, evaporation, and coagulation. We track how these microphysical processes change the size distribution of condensed Se through mass-conserving schemes. Evaporation of condensed Se only occurs when the smallest sulfate aerosol bin evaporates, releasing
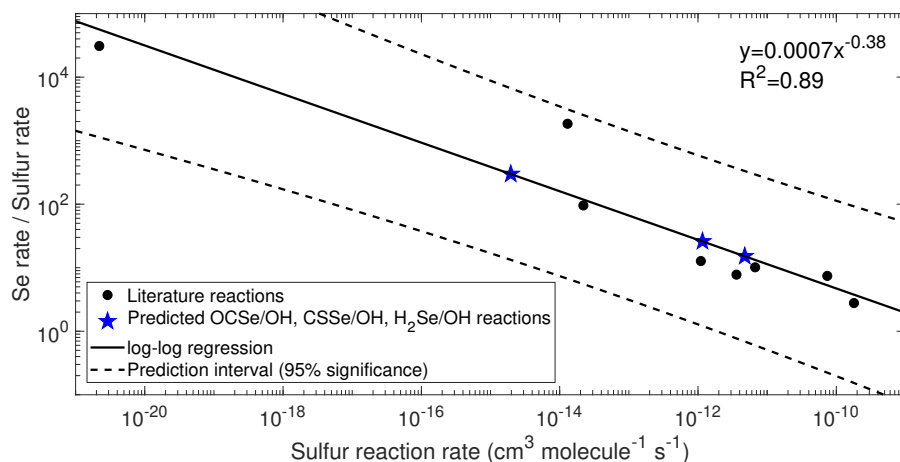
Figure 2. Estimation of unknown Se reaction rates from the analogous S reaction rate. A power regression is performed, with statistics shown in the upper right corner of the plot. For the $H_2S + O_3$ reaction only an upper limit estimate was available, and therefore it was not included in the analysis.

the Se stored in that bin as gas-phase inorganic oxidized Se. Sedimentation and deposition of the host sulfate particles are concurrent with sedimentation and deposition of the condensed Se tracers. Gas-phase oxidized Se tracers are also removed by dry and wet deposition, with the assumption that they have the same Henry's law constant as gas-phase H_2SO_4.

One limitation of using SOCOL-AER for the Se cycle is that the model only includes online sulfate aerosols. This means
5   that the transport of Se on other aerosols, including dust, sea salt, and organic aerosols, would be neglected. This may not be a poor assumption, since Se and S are often co-emitted and have been found to be highly correlated in atmospheric aerosol measurements (e.g., Eldred, 1997; Weller et al., 2008). Nevertheless, we included a "dummy" aerosol tracer to test the effect of missing aerosol species in SOCOL-AER. The dummy aerosol tracer represents monodisperse particles that are emitted in a latitudinal band in the model and undergo Se uptake, sedimentation, and wet and dry deposition. This dummy aerosol tracer is
10   clearly a simplification of true atmospheric processes, as in reality other aerosols are distributed in size and can coagulate with sulfate aerosol particles. However, by varying the radius, location, and emission magnitude of these particles (Sect. 3.1.7), we can determine whether missing aerosols affect atmospheric Se cycling.

3   Statistical methods

To conduct the sensitivity analysis of our Se model, we first need to select the input parameters that would be included in
15   the sensitivity analysis. The probability distributions of these input parameters' uncertainties were determined by reviewing literature sources and using our best judgement. Variance-based sensitivity analysis methods usually require 10^4 to 10^6 model runs, which would be prohibitively expensive for the full SOCOL-AER model. We therefore replace the SOCOL-AER model

with a surrogate PCE model. The SOCOL-AER model is run at carefully selected points within the parameter space, creating a set of "training" runs. The training runs are used to produce a surrogate PCE model, which approximates the outputs of the full SOCOL-AER model throughout the input parameter space. Sensitivity indices can then be derived from the surrogate model. All statistical methods presented in this section are available in UQLAB, an open-source MATLAB-based framework for uncertainty quantification (Marelli and Sudret, 2014).

3.1 Uncertainty ranges of input parameters

We restricted the scope of our sensitivity analysis to parameters that have been implemented in the model as part of the Se cycle. We neglect other model ~~uncertainties~~parameters, including those related to ~~the wet deposition parametrization~~ climate, deposition parametrizations, or the emissions of sulfate aerosol precursors. The focus of our sensitivity analysis is to prioritize which Se-related parameters should be further constrained. Since we do not vary all other model parameters, the uncertainties of output quantities may be underestimated. However, given the large dimension of our parameter space with 34 Se-related input parameters, including additional non-Se-related parameters would be challenging. In the following section, we will discuss the uncertainty distributions for each of the 34 input parameters included in our study. Due to the lack of detailed information available in literature about the parameter distributions, we chose loguniform or uniform distributions for all but one of the parameters. This follows the conservative approach recommended by the Maximum Entropy Principle, as the uniform and loguniform distributions maximize entropy while fulfilling the data constraints (Kapur, 1989). The uncertainty distributions of all input parameters are listed in Table 3.

3.1.1 Measured rate constants (k_1–k_{12})

The Se reactions studied in the literature have each only been measured by one laboratory group (Table 2). Since only one measurement technique has been applied, the reported measurement uncertainties may underestimate the true uncertainties of these rate constants. To approximate an uncertainty distribution for these rate constants, we reviewed S compound reactions that have been studied by multiple research groups. The reaction that had the largest spread in reported rate constants at $\sim$298 K was OCS + OH, which has been measured in six studies (Atkinson et al., 1978; Kurylo, 1978; Cox and Sheppard, 1980; Leu and Smith, 1981; Cheng and Lee, 1986; Wahner and Ravishankara, 1987; Burkholder et al., 2015). The measured reaction rate constant varied over multiple orders of magnitude; therefore, we calculated its variability on a logarithmic scale. The coefficient of variation (standard deviation divided by mean) of this reaction rate in logarithmic space was around 6%. We assumed that this maximum S coefficient of variation would apply to the measured Se reaction rates. The bounds were calculated as 88% and 112% (± 2 standard deviations) of the available measured rate constant in logarithmic space, i.e.:

$$\text{Bounds} = k^{1 \pm 0.12} \tag{1}$$

where k is the measured rate constant and "Bounds" are its upper and lower bounds, all expressed in $\mathrm{cm^3\,molec^{-1}\,s^{-1}}$. The maximum upper bound was set to 1.0×10^{-9} $\mathrm{cm^3\,molec^{-1}\,s^{-1}}$, since at this order of magnitude the Se reaction rates are collision-limited (Seinfeld and Pandis, 2016).

Table 3. Probability distributions of the model input parameters selected for the sensitivity analysis.

Input parameter	Description	Lower bound	Upper bound	Distribution
Measured reaction rate coefficients ($cm^3\,molec^{-1}\,s^{-1}$)				
k_1	$OCSe + O \rightarrow CO + OX\_Se\_IN$	1.3×10^{-12}	4.5×10^{-10}	loguniform
k_2	$CSe_2 + O \rightarrow OCSe + OX\_Se\_IN$	2.6×10^{-12}	7.9×10^{-10}	loguniform
k_3	$CSe_2 + O \rightarrow 2\,OX\_Se\_IN$	6.0×10^{-12}	1.0×10^{-9}	loguniform
k_4	$CSSe + O \rightarrow OCSe$	4.0×10^{-13}	1.8×10^{-10}	loguniform
k_5	$CSSe + O \rightarrow OX\_Se\_IN$	1.0×10^{-12}	3.9×10^{-10}	loguniform
k_6	$DMSe + OH \rightarrow OX\_Se\_OR$	4.1×10^{-12}	1.0×10^{-9}	loguniform
k_7	$DMSe + NO_3 \rightarrow OX\_Se\_OR$	7.0×10^{-13}	2.8×10^{-10}	loguniform
k_8	$DMSe + O_3 \rightarrow OX\_Se\_OR$	7.8×10^{-19}	5.9×10^{-15}	loguniform
k_9	$DMSe + Cl \rightarrow OX\_Se\_OR$	3.8×10^{-11}	1.0×10^{-9}	loguniform
k_{10}	$H_2Se + O \rightarrow OX\_Se\_IN$	8.3×10^{-14}	5.3×10^{-11}	loguniform
k_{11}	$H_2Se + Cl \rightarrow OX\_Se\_IN$	4.6×10^{-11}	1.0×10^{-9}	loguniform
k_{12}	$H_2Se + O_3 \rightarrow OX\_Se\_IN$	4.4×10^{-18}	2.3×10^{-14}	loguniform
Estimated reaction rate coefficients ($cm^3\,molec^{-1}\,s^{-1}$)				
k_{13}	$OCSe + OH \rightarrow OX\_Se\_IN$	2.7×10^{-14}	2.4×10^{-11}	loguniform
k_{14}	$CSe_2 + OH \rightarrow OCSe + OX\_Se\_IN$	7.8×10^{-12}	1.0×10^{-9}	loguniform
k_{15}	$CSSe + OH \rightarrow OX\_Se\_IN$	1.1×10^{-12}	9.3×10^{-10}	loguniform
k_{16}	$CSSe + OH \rightarrow OCSe$	4.7×10^{-13}	4.0×10^{-10}	loguniform
k_{17}	$H_2Se + OH \rightarrow OX\_Se\_IN$	3.7×10^{-12}	1.0×10^{-9}	loguniform
Scaling factors for photolysis rates				
fk_{18}	$OCSe + h\nu \rightarrow OX\_Se\_IN$	0	2	uniform
fk_{19}	$CSe_2 + h\nu \rightarrow 2\,OX\_Se\_IN$	0	2	uniform
fk_{20}	$CSSe + h\nu \rightarrow OX\_Se\_IN$	0	2	uniform
fk_{21}	$H_2Se + h\nu \rightarrow OX\_Se\_IN$	0	2	uniform
Global emission by source category ($Gg\,Se\,yr^{-1}$)				
$emiss_{mar}$	Marine biogenic Se emissions	0.4	35	loguniform
$emiss_{terr}$	Terrestrial biogenic Se emissions	0.15	5.25	uniform
$emiss_{ant}$	Anthropogenic Se emissions	3	9.6	uniform
$emiss_{vol}$	Volcanic Se emissions	0.076	49.1	loguniform
Speciation of emissions (%)				
$\%OCSe_{ant}$	OCSe fraction in anthropogenic emissions	0	6	uniform
$\%CSe_{2\,ant}$	CSe_2 fraction in anthropogenic emissions	0	6	uniform
$\%CSSe_{ant}$	CSSe fraction in anthropogenic emissions	0	6	uniform
$\%H_2Se_{ant}$	H_2Se fraction in anthropogenic emissions	0	6	uniform
$\%H_2Se_{vol}$	H_2Se fraction in volcanic emissions	0	13	uniform
Accommodation coefficient				
acc_{coeff}	Oxidized Se uptake on aerosols	0.02	1	uniform
Dummy aerosol parameters				
r_{aer}	Radius of missing aerosol (µm)	0.01	3	loguniform
$emiss_{aer}$	Emission magnitude of missing aerosol ($kg\,yr^{-1}$)	-	-	lognormal (see text)
lat_{aer}	Latitudinal band of aerosol emission	90° to 80° S	80° to 90° N	uniform

3.1.2 Estimated rate constants (k_{13}–k_{17})

Five Se reaction rate constants have not been measured previously in the laboratory and were estimated based on their relationship with analogous S rate constants (Fig. 2). We calculated the uncertainty bounds of estimated rate constants using the 95% error interval of prediction with a linear regression (Wackerly et al., 2014):

$$\text{Bounds} = 10 \char`\^ \left[x + \hat{Y} \pm t_{0.025} \sqrt{ \frac{\text{SSE}}{n-2} \left(1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{S_{xx}} \right) } \right] \tag{2}$$

where x is the logarithm (to the base 10) of the corresponding S rate constant, $\hat{Y}$ is the logarithm of the predicted ratio of the Se rate to the S rate, n is the number of data points in the regression, $t_{0.025}$ is the 2.5[th] percentile value of the Student's t distribution for $n-2$ degrees of freedom, $\bar{x}$ is the mean of the logarithm of S rate constants in Fig. 2, SSE is the sum of squares of the residuals, and S_{xx} is the variance of the S rate constants in Fig. 2. All rate constants are in units of $\mathrm{cm^3\,molec^{-1}\,s^{-1}}$.

3.1.3 Photolysis rate scaling factors

Uncertainties in our calculated Se photolysis rates arise from uncertainties in the measured OCSe and H_2Se cross sections, the assumption that CSe_2 and $CSSe$ have the same cross section as CS_2, the quantum yields of photolysis reactions, and the look-up table approach that SOCOL-AER applies to calculate photolysis rates. Given the lack of specific information about these uncertainties, we apply a scaling factor on the calculated photolysis rates ranging from 0 to 2 (Table 3).

3.1.4 Global emissions from source categories

For the sensitivity analysis, we do not alter the spatial distribution of Se emissions from anthropogenic, marine biogenic, terrestrial biogenic, and volcanic sources (Fig. 3). The parameters that we varied are the scaling factors for each map, i.e., the global total emissions from each source. We reviewed atmospheric Se budget estimates to determine the range in total emissions for different sources (Table S1). The estimates for global DMSe emissions ranged from the lower limit value of $0.4\ \mathrm{Gg\,Se\,yr^{-1}}$ in Nriagu (1989) to $35\ \mathrm{Gg\,Se\,yr^{-1}}$ in Amouroux et al. (2001). DMSe emissions are calculated online in the model from a marine DMSe concentration map. Based on the results of a previous simulation, we normalize the marine DMSe concentration map in Fig. 3a so that it leads to $1\ \mathrm{Gg\,Se\,yr^{-1}}$ emissions globally. All other maps are also normalized to $1\ \mathrm{Gg\,Se\,yr^{-1}}$ emissions, so that we can directly apply the total global emissions as a scaling factor. The widest uncertainty range of terrestrial emissions was given by Nriagu (1989), from 0.15–$5.25\ \mathrm{Gg\,Se\,yr^{-1}}$. Total global anthropogenic Se emissions in 1983 were estimated between 3.0 and $9.6\ \mathrm{Gg\,Se\,yr^{-1}}$ (Nriagu and Pacyna, 1988). We applied the same uncertainty range to the total anthropogenic emissions in the year 2000, because it is unclear how global Se emissions have changed during this period. To estimate global Se emissions from degassing volcanoes, we reviewed measurements of Se to S ratios in volcanic emissions, extending the studies reviewed in Floor and Román-Ross (2012) (Table S1S2). There is a high degree of variability in the emitted Se to S ratios between different volcanoes, and even temporally and spatially for the same volcano (Floor and Román-Ross, 2012). The Se to S ratio in volcanic emissions ranges from 6×10^{-6} for White Island, New Zealand (Wardell et al., 2008) to 3.9×10^{-3} for Merapi Volcano, Indonesia (Symonds et al., 1987). By multiplying this range in ratios with the
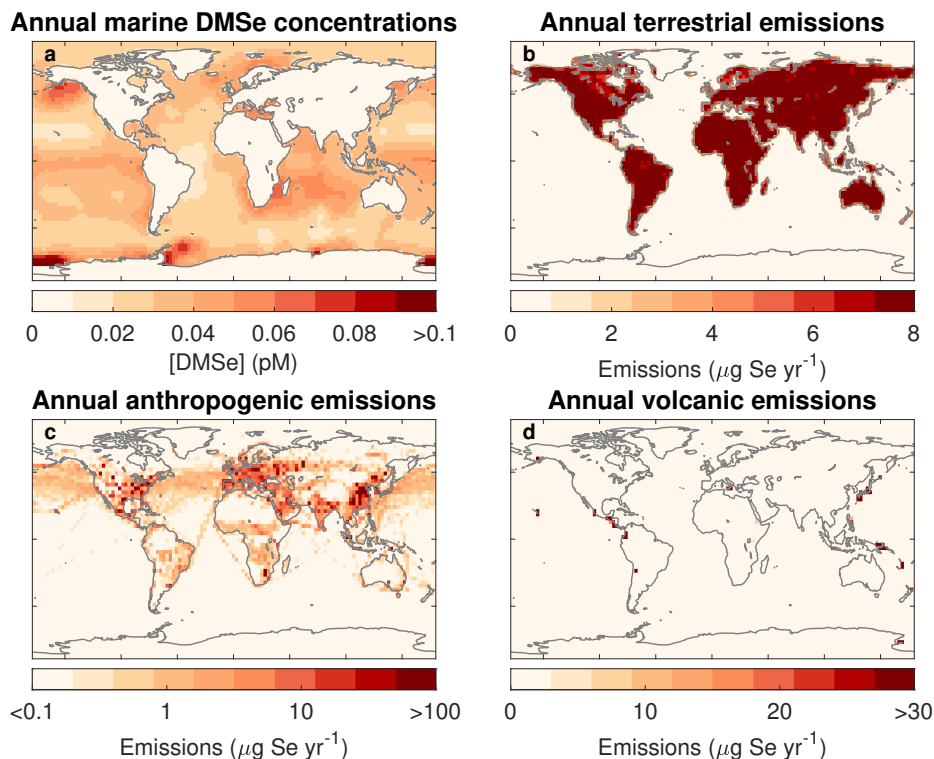
Figure 3. Spatial distribution of Se sources. Each source map is normalized so that $1\,\mathrm{Gg\,Se\,yr^{-1}}$ would be emitted globally. This is scaled during the sensitivity analysis.

global mean total degassing SO_2 emissions from Andres and Kasgnoc (1998) and Dentener et al. (2006), $12.6\,\mathrm{Tg\,S\,yr^{-1}}$, we calculate an uncertainty range for total volcanic Se emissions: 0.076–$49.1\,\mathrm{Gg\,Se\,yr^{-1}}$. Loguniform distributions were used for the source types whose total emission ranges vary by more than 2 orders of magnitude (volcanic and marine biogenic), whereas uniform distributions were used for the terrestrial and anthropogenic Se emissions.

3.1.5 Speciation of Se emission sources

The available speciation information for Se emissions is largely qualitative; possible emission species have been identified but not quantified (Sect. 2.2.2). To estimate the uncertainty range for Se emission speciation, we use estimates of speciation from an atmospheric S budget (Watts, 2000). The second most important anthropogenic S species after SO_2 is H_2S. Watts (2000) estimates the anthropogenic emissions of H_2S are $3.1 \pm 0.3\,\mathrm{Tg\,S\,yr^{-1}}$, compared to an anthropogenic SO_2 emission total of $53.2\,\mathrm{Tg\,S\,yr^{-1}}$ (Lamarque et al., 2010; Smith et al., 2011) in the year 2000. H_2S therefore contributes at most 6% of total S emissions. Since $OCSe$, CSe_2, $CSSe$, and H_2Se are in general less stable than the analogous S species (Table 2), we consider 6% to be a maximum value for the mass fraction of anthropogenic Se emissions that come from each of these

13

species. The anthropogenic speciation fractions for OCSe, CSe$_2$, CSSe, and H$_2$Se are varied between 0 and 6%. The rest of the anthropogenic emissions, 76–100%, are attributed to OX_Se_IN, representing species such as Se and SeO$_2$.

Regarding the speciation of volcanic S emissions, Watts (2000) estimates that $0.99 \pm 0.88 \, \mathrm{Tg\,S\,yr^{-1}}$ is in the form of H$_2$S. Comparing this to the estimate for volcanic SO$_2$ emissions, $12.6 \, \mathrm{Tg\,S\,yr^{-1}}$ (Andres and Kasgnoc, 1998; Dentener et al., 2006), H$_2$S contributes at most 13% of volcanic S emissions. Therefore, in our sensitivity analysis the percentage of volcanic Se emissions in the form of H$_2$Se ranges from 0–13%. Conversely, the percentage of OX_Se_IN in volcanic emissions ranges from 87–100%.

3.1.6 Accommodation coefficient of oxidized Se

The accommodation coefficient represents the probability that a gas-phase oxidized Se molecule will stick to an aerosol upon collision. No laboratory studies have investigated the accommodation coefficient of oxidized Se on aerosol surfaces. However, review papers suggest that Se efficiently partitions to the aerosol phase upon oxidation (Mosher and Duce, 1987; Wen and Carignan, 2007), indicating a high accommodation coefficient. Due to the lack of further information, we assume an uncertainty range of 0.02–1 for the accommodation coefficient, as selected by Lee et al. (2011) for H$_2$SO$_4$. This accommodation coefficient applies to uptake of Se on sulfate and dummy aerosols.

3.1.7 Dummy aerosol tracers

SOCOL-AER only includes sulfate aerosol, lacking other aerosol types (e.g., dust, sea salt, organic aerosol, etc.) that may also transport Se. To test how these other aerosol types might affect Se transport and deposition, in the sensitivity analysis we vary the emission location, particle radius, and emission magnitude of the dummy aerosol tracer (introduced in Sect. 2.2.4). These input parameters are different from other uncertainties, in that they are intended to investigate a lack in model completeness rather than uncertainties in measurable quantities. In our experiments the aerosols are emitted in one of 18 latitude bands, ranging from 90°–80° S to 80°–90° N. The latitude parameter can demonstrate whether Se deposition is affected by these missing aerosol types only in certain latitudinal bands. The emission radius affects the transport of Se on these missing particles, since the atmospheric lifetime of these particles depends on radius (Seinfeld and Pandis, 2016). For example, additional coarse particles (radius > 1 μm) could inhibit far-range Se transport, since they sediment and are removed quicker than the accumulation mode sulfate particles. We vary the emission radius between 0.01 and 3 μm, as this covers the range of atmospherically relevant particle sizes.

To determine a reasonable range for the emission magnitude of additional aerosol particles, we analyzed the particle emission inventories from the AEROCOM I project (Dentener et al., 2006). Aerosol types are segregated into size classes based on their effective radius, and the emission magnitude for 10° latitude bands is calculated for each size class. The mass emission of particles correlates with the particle size, with generally larger mass emissions for larger particle sizes (Fig. S1). Therefore, the value of emission magnitude in our experiments depends lognormally on the particle radius (r), with mean $\mu = 2.54r + 19.0$ and standard deviation $\sigma = 2.88$. We acknowledge that in the real atmosphere particles are emitted globally as size distributions, not monodisperse particles in one latitudinal band. Nevertheless, by including these input parameters as part of the sensitivity

analysis, we can identify whether the lack of tropospheric aerosols other than sulfate impacts the simulated Se deposition maps in SOCOL-AER.

3.2 Experimental setup and model outputs

The creation of surrogate models requires a set of training runs with the full SOCOL-AER model. The values of the input parameters are varied simultaneously between training runs, so that interactions between parameters can also be detected. A Latin hypercube design is used to draw N samples from an M-dimensional input parameter space. In Latin hypercube sampling, each parameter range is divided into N equally probable sub-intervals. N values for each parameter are drawn randomly from within each sub-interval. The selected values for all input parameters are then matched randomly with each other, to yield N points that cover the parameter space better than purely random Monte Carlo sampling (McKay et al., 1979). The general rule of thumb is to select around $10M$ training simulations, to adequately cover the sample space (Loeppky et al., 2009). We ran $N = 400$ training simulations in our sensitivity analysis with 34 input parameters. Producing training simulations from the full SOCOL-AER model is the most computationally expensive step in the uncertainty and sensitivity analysis. With 48 cores each training run takes around 14 hours, corresponding to around 10^9 core-seconds for 400 training simulations.

The initial conditions file for the training runs was created from a previous 10 year spinup of the model under year 2000 conditions. The atmospheric mixing ratios of Se tracers are initially set to 0 in each simulation. The simulations are 18 months long, with the first 6 months considered to be an equilibration period for the Se tracers. Therefore, we analyze only the last 12 months of the 18-month simulation. The model is run with T42 horizontal resolution (approximately $2.8° \times 2.8°$ or $300 \text{ km} \times 300 \text{ km}$ in latitude and longitude) with 39 vertical levels up to 0.01 hPa ($\sim$80 km). The interactions between chemical species (i.e., greenhouse gases and aerosols) and radiation are decoupled in our simulations. The decoupling between chemistry and climate prevents meteorological differences between training runs with different input parameters, eliminating the influence of precipitation variability on deposition. Deposition flux differences in these relatively short simulations can then be more easily attributed to changes in the input parameters.

All relevant Se cycle fluxes and reservoir burdens are outputted by SOCOL-AER as monthly averages. For the sensitivity analysis, the target outputs are annual mean values of the global atmospheric Se lifetime and Se surface deposition fluxes. The global Se lifetime is calculated as the total atmospheric Se burden divided by total deposition (Seinfeld and Pandis, 2016). Deposition fluxes of Se were calculated by summing up the dry and wet deposition fluxes of aerosol and gas-phase Se species. Since the model is run in T42 horizontal resolution, there are 8192 surface grid boxes representing geographical coordinates on the globe. We therefore have 8192 model outputs for Se deposition and we construct a PCE model and conduct a sensitivity analysis for each one. It can be argued that the computational cost can further be reduced by dimensionality reduction techniques (e.g., Blatman and Sudret, 2011b, 2013; Ryan et al., 2018), like building the PCE models on a reduced output set coming from a principal component analysis (PCA). We did not consider such an approach in this work because the cost of building the 8192 PCE models is marginal compared to that of evaluating the full SOCOL-AER model. In summary, the 400 SOCOL-AER training runs yield 400 values for atmospheric Se lifetimes and 400×8192 values for deposition fluxes.

3.3 Surrogate modeling with PCE

Sudret (2008) provided a detailed description of using PCE to build surrogate models, which was updated by Le Gratiet et al. (2017). We will summarize the most important features of the approach here. For this study, a certain output of the SOCOL-AER model (Y) can be thought of as a finite variance random variable which is a function of the $M = 34$ input parameters ($\mathbf{X} = [X_1, X_2, \ldots, X_{34}]$):

$$Y = \mathcal{M}(\mathbf{X}) \tag{3}$$

The input $\mathbf{X} \in \mathbb{R}^M$ is modeled by a joint probability density function (PDF) $f_{\mathbf{X}}$ whose marginals are assumed independent, i.e., $f_{\mathbf{X}} = \prod_{i=1}^{M} f_{X_i}(x_i)$.

In polynomial chaos decomposition, the output variable Y is decomposed into the following infinite series (Ghanem and Spanos, 2003):

$$Y = \sum_{\boldsymbol{\alpha} = \mathbb{N}^M} y_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}(\mathbf{X}) \tag{4}$$

where $y_{\boldsymbol{\alpha}}$ are coefficients to be determined and $\boldsymbol{\alpha}$ is a multi-index set that defines the degree of the multivariate orthonormal polynomial $\psi_{\boldsymbol{\alpha}}(\mathbf{X}) = \prod_{i=1}^{M} \psi_{\alpha_i}(X_i)$. The latter belongs to a family of polynomials that are orthogonal with respect to the marginal PDF f_{X_i}. For example, univariate uniform probability distributions correspond to the family of Legendre polynomials and Gaussian probability distributions correspond to the family of Hermite polynomials (Xiu and Karniadakis, 2002). Multivariate orthogonal polynomials can be constructed through multiplying the relevant univariate polynomials. The *order* of a polynomial term is defined as the total number of variables included in the polynomial term. The *degree* of a polynomial term represents the sum of the exponents of all variables appearing in the term. The number of coefficients to estimate grows exponentially with both the dimension and the degree. To allow calculation of the coefficients, the terms in Eq. 4 are truncated by restricting the maximum degree of the polynomials. Advanced truncation schemes can also be used to reduce the number of terms and thus the computational budget. Here we consider hyperbolic truncation which removes high order interaction terms from the PCE, while retaining high degree polynomials of a single variable (Blatman and Sudret, 2011a). Hyperbolic truncation is based on selecting only the terms that satisfy:

$$\mathcal{A} = \left\{ \boldsymbol{\alpha} : \left(\sum_{i=1}^{M} \alpha_i^q \right)^{1/q} \leq p \right\} \tag{5}$$

where q is a value selected between 0 and 1, and p is a value selected as the maximum degree of the PCE. Setting $q = 1$ corresponds to the standard truncation scheme where all terms with degree below p are selected. The lower the value of q, the more higher order interaction terms are removed from the PCE. In our caseBased on previous experience, we selected a q value of 0.75.

PCE coefficients are generally calculated by least-square regression (Berveiller et al., 2006) using an *experimental design* which consists of uniformly sampled realizations of the input variables $\mathcal{X} = \{\mathbf{x}^{(1)} \ldots \mathbf{x}^{(N)}\}$ and corresponding model evalua-

tions $\{\mathcal{M}(\mathbf{x}^{(1)})\ldots\mathcal{M}(\mathbf{x}^{(N)})\}$, i.e.:

$$y_{\boldsymbol{\alpha}} = \arg\min_{y_{\boldsymbol{\alpha}}\in\mathbb{R}^{\text{card}\mathcal{A}}} \frac{1}{N}\sum_{i=1}^{N}\left(\mathcal{M}(\mathbf{x}^{(i)}) - \sum_{\boldsymbol{\alpha}\in\mathcal{A}} y_{\boldsymbol{\alpha}}\Psi_{\boldsymbol{\alpha}}(\mathbf{x}^{(i)})\right)^2 \tag{6}$$

When the dimension is large, regression techniques that allow for sparsity, i.e., by forcing some coefficients to be zero, are favored. In this work, we consider least-angle regression as proposed in Blatman and Sudret (2011a) and follow the implementation in the UQLAB PCE module (Marelli et al., 2019).

The accuracy of the PCE in representing the full SOCOL-AER model is evaluated with a cross-validation metric named the leave-one-out (LOO) error, ϵ_{LOO} (Blatman and Sudret, 2010a). The ~~cross-validation approach removes the need for an independent validation dataset, saving computational expense. The~~ leave-one-out procedure consists of creating a PCE using all but one of the training simulations, $\mathcal{M}^{\text{PC}\backslash i}$. This PCE is then used to predict the output value of the model at the removed training simulation point, $\mathbf{x}^{(i)}$. The process is repeated for all N training points so that a residual sum of squares can be calculated. This residual is then divided by the output variance in the training dataset, yielding the LOO error:

$$\epsilon_{\text{LOO}} = \frac{\sum_{i=1}^{N}(\mathcal{M}(\mathbf{x}^{(i)}) - \mathcal{M}^{\text{PC}\backslash i}(\mathbf{x}^{(i)}))^2}{\sum_{i=1}^{N}(\mathcal{M}(\mathbf{x}^{(i)}) - \hat{\mu}_Y)^2} \tag{7}$$

where $\hat{\mu}_Y$ is the sample mean of the model output in the training dataset. In practice, the LOO error is estimated using a *single* PCE model considering the entire experimental design (Marelli and Sudret, 2014; Blatman and Sudret, 2010b), to avoid the procedure of creating N PCE models. Equation 7 is evaluated as:

$$\epsilon_{\text{LOO}} = \sum_{i=1}^{N}\left(\frac{\mathcal{M}(\mathbf{x}^{(i)}) - \mathcal{M}^{\text{PC}}(\mathbf{x}^{(i)})}{1 - h_i}\right)^2 \bigg/ \sum_{i=1}^{N}\left(\mathcal{M}(\mathbf{x}^{(i)}) - \hat{\mu}_Y\right)^2 \tag{8}$$

where h_i is the ith component of the vector defined by:

$$\mathbf{h} = \text{diag}\left(\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\right) \tag{9}$$

and $\mathbf{A}$ is the experimental matrix whose components read:

$$A_{ij} = \Psi_j\left(\mathbf{x}^{(i)}\right) \qquad i = 1,\ldots,n; \quad j = 0,\ldots,P-1 \tag{10}$$

where $P \equiv \text{card}\mathcal{A}$ is the number of terms in the PCE. In our study, we use a degree-adaptive scheme to construct the PCE models sequentially from degree 1 to maximum degree 13. If the LOO error does not decrease over two steps in degree, the algorithm is stopped and the maximum degree is selected as the one with the lowest LOO error. This method reduces the risk of overfitting the training set. The maximum degree of 13 was selected because the PCE becomes too computationally expensive to calculate above this degree. In any case, almost all PCEs calculated in this study are below degree 10.

To improve the accuracy of the approximation, we applied post-processing steps to the construction of PCE models (Table 4). The global atmospheric Se lifetime is a function of the atmospheric Se burden and the total Se deposition. We found improved overall accuracy when separate PCE models were created for the atmospheric burden and total deposition, rather

Table 4. Summary of methods to construct surrogate models and calculate Sobol' sensitivity indices of the SOCOL-AER output parameters.

Output parameter	Surrogate modeling method	Sensitivity analysis method
Global atmospheric Se lifetime	Constructed a PCE model of atmospheric Se burden and a PCE model of Se deposition flux. Divided the burden PCE by the deposition PCE.	Monte Carlo estimation of Sobol' sensitivity indices
Deposition flux of Se (7319 grid boxes)	Constructed PCE models of deposition fluxes in each grid box (LOO error was less than 0.05).	Sobol' sensitivity indices are extracted directly from deposition PCE models
Deposition flux of Se (873 grid boxes with LOO > 0.05)	Constructed PCE models of log-transformed deposition fluxes in each grid box. Exponential of the PCE models yields surrogate models for deposition fluxes.	Monte Carlo estimation of Sobol' sensitivity indices

than first calculating the lifetime for each simulation and constructing a PCE model of the lifetime. To calculate a surrogate model of the Se lifetime, we divided the PCE model of the atmospheric Se burden by the PCE model of total deposition. When calculating the PCE models of Se deposition fluxes, we first used the deposition fluxes directly from the training simulations. However, 873 of the 8192 grid boxes showed LOO errors higher than 0.05, which was selected as the acceptable threshold for

5 this study. Improved LOO errors were achieved when the simulated deposition fluxes were log-transformed before constructing the PCE model. However, after log-transformation the sensitivity indices for deposition cannot be extracted analytically from the PCE (Sect. 3.4), increasing computational expense. Therefore, we only log-transformed the data from these 873 grid boxes before creating their PCE models. Surrogate models for deposition fluxes in these 873 grid boxes are calculated by taking the exponential of the PCE model of log-transformed data.

10 The cross-validation approach would usually remove the need for an independent validation dataset, saving computational expense. However, to evaluate the post-processing steps applied to the surrogate models, we also produced an independent validation dataset of 50 SOCOL-AER runs. The parameters for these runs were chosen by enriching the training experimental design to create a pseudo-Latin Hypercube of 450 runs, ensuring that the distance between the validation runs and existing training runs is maximal.

15 ## 3.4 Sensitivity analysis

A Sobol' sensitivity index represents the fraction of model variance caused by the parametric uncertainty of a certain input variable or the interaction between multiple variables (Sobol, 1993; Marelli et al., 2019). It is a global measure, meaning that the sensitivity index considers the entire parametric uncertainty space. Let us consider a model $\mathcal{M}$ whose inputs, defined over a domain $\mathcal{D}_{\mathbf{x}}$, are assumed independent. It can be shown that it admits the following so-called Sobol' decomposition:

20 $$\mathcal{M}(x) = \mathcal{M}_0 + \sum_{i=0}^{M} \mathcal{M}_i(x_i) + \sum_{1 \leq i \leq j \leq M}^{M} \mathcal{M}_{ij}(x_i, x_j) + ... + \mathcal{M}_{1,2,...,M}(x_1, ..., x_M) \quad (11)$$

where $\mathcal{M}_0$ is a constant and the other summands are defined such that their integrals with respect to any of their arguments is 0, i.e.:

$$\int_{\mathcal{D}_\mathbf{x}} M_{i_1,...,i_s}(x_{i_1},...,x_{i_s}) f_{\mathbf{x}_{i_k}} \mathrm{d}x_{i_k} = 0, \qquad 1 \le k \le s \tag{12}$$

Following this decomposition, it can be shown that the variance of the random variable $Y = \mathcal{M}(\mathbf{X})$ can be cast as (Saltelli et al., 2008):

$$D = \mathrm{Var}[\mathcal{M}(\mathbf{X})] = \sum_{i=1}^{M} D_i + \sum_{1 \le i \le j \le M}^{M} D_{ij} + ... + D_{12...M} \tag{13}$$

where a partial variance with respect to several variables $X_{i_1}, \ldots, X_{i_s}$ can be calculated as:

$$D_{i_1,...,i_s} = \int_{\mathcal{D}_\mathbf{x}} \mathcal{M}^2_{i_1,...,is}(x_{i_1},...,x_{i_s}) f_{X_{i_1}}(x_{i_1})...f_{X_{i_s}}(x_{i_s}) \mathrm{d}x_{i_1}...\mathrm{d}x_{i_s} \tag{14}$$

The Sobol' indices ($S_{i_1,...,i_s}$) for a subset of input parameters are eventually obtained by normalizing the corresponding variance, i.e.:

$$S_{i_1,...,i_s} = \frac{D_{i_1,...,i_s}}{D} \tag{15}$$

The number of variables involved in a Sobol' index determines its order. A first order Sobol' index ($S_i = \frac{D_i}{D}$) apportions the variance in the model output to the effect of a single variable, X_i. Second order indices ($S_{ij} = \frac{D_{ij}}{D}$) represent the impact of the interaction of two variables (e.g., X_i and X_j, $i \ne j$) on the model output variance, not already accounted for by S_i and S_j. Higher order indices can be calculated as well. The summation of all individual Sobol' indices yields a value of 1, i.e., accounting for all of the variance in the output. However, the number of higher order indices to calculate can become very large:

$$\text{Number of n}^{\text{th}} \text{ order indices} = \binom{n}{M} = \frac{n!}{M!(n-M)!} \tag{16}$$

It therefore becomes computationally demanding in our case with $M = 34$ input parameters to calculate sensitivity indices with order higher than 2. Furthermore, by the *sparsity-of-effect principle* (Goupy and Creighton, 2006), it is expected that the model is primarily driven by first order effects and low order interactions. We therefore only calculate individually the first and second order Sobol' indices, which is common practice in global sensitivity analysis.

The total Sobol' index (S_i^T) summarizes all sensitivity indices which include the effect of a given input variable:

$$S_i^T = S_i + \sum_{j \ne i} S_{ij} + \sum_{j \ne i} \sum_{\substack{k \ne i \\ k \ne j}} S_{ijk} + ... \tag{17}$$

In practice, it is possible to calculate the total Sobol' index without individually calculating all of the higher order indices (Marelli et al., 2019). The total Sobol' indices can be used to rank the influence of input variables on the variability in the

model output. It should be noted that the sum of all total Sobol' indices will be greater than 1 if the model is non-additive, since interaction terms would be counted multiple times.

Other studies have emphasized the computational expense of conducting a sensitivity analysis for all grid boxes in a chemistry–climate model (Ryan et al., 2018). By using PCE as the surrogate model, we can reduce the computational expense since it is possible to compute the Sobol' sensitivity indices *analytically*. As shown by Sudret (2008), Sobol' sensitivity indices are a function of the calculated coefficients in the PCE model (Eq. 4). ~~However, this is~~, due to the similarity of the PCE decomposition with variance decomposition. First order Sobol' indices can be calculated as $D_i = \sum_{\alpha \in \mathcal{A}_i} y_\alpha^2$, where $\mathcal{A}_i$ is the set of polynomial terms involving only variable i. Similarly, higher-order Sobol' indices can be calculated by summing the squares of the coefficients of the polynomial terms that include the variables of interest. The analytical solution of Sobol' indices is no longer possible when post-processing steps are applied to the original PCE models to create surrogate models. For example: 1) the surrogate model for the atmospheric lifetime is calculated from two PCE models, one for total Se burden and one for total Se deposition; 2) Se deposition in 873 grid boxes is calculated using the exponential of a PCE model (see Table 4). In these cases, it is still possible to calculate the Sobol' indices through Monte Carlo estimation (Marelli et al., 2019). To accomplish this, the surrogate models are sampled around 10^6 times, which remains tractable.

To summarize categories of input variables, we aggregate the Sobol' indices of several input parameters to yield a total Sobol' index for that category. For example, we summarize the total dummy aerosol effect by summing the total Sobol' indices of the dummy aerosol radius, emission magnitude, and latitude. It would anyways be difficult to separate the effects of the dummy aerosol input parameters since they are correlated inputs in the experimental design. The second order indices involving two dummy aerosol input parameters may be double-counted with this method. However, since these indices are small (< 0.05) we do not expect a large error in the aggregated total Sobol' index (Sect. 5.1).

3.5 Resampling of surrogate models

In order to estimate distribution statistics (mean, standard deviation, quantiles) of the output variable, we resample each surrogate model 40 000 times. We also use these 40 000 samples of the parameter space to calculate relationships between input parameters and output variables. To visualize marginal relationships between a certain input parameter and the output, we replace the value of the input parameter in the 40 000 samples by a fixed value and calculate the mean and variance of the surrogate model output. Repeating this step with other evenly spaced values in the input parameter range, we can produce the univariate relationship between the model output and the input parameter.

4 Compilation of Se wet deposition flux data

We decided to compare SOCOL-AER results with measurements of Se wet deposition, since wet deposition contributes an estimated 80% of total deposition (Wen and Carignan, 2007). We conducted a systematic literature review to assemble a dataset of measured wet Se deposition fluxes, extending an earlier review from Conde and Sanz Alaejos (1997). We searched in Web of Science (Clarivate Analytics) for combinations of the following criteria: "Selenium", "Se", "rain", "precipitation", "wet

deposition", "trace element". The last search was completed in July 2019, yielding a total of 672 search results. We screened these search results for studies that measured Se concentrations in rainwater for at least one month, neglecting studies that measured bulk (wet and dry) deposition or that extrapolated wet deposition fluxes from aerosol measurements. The compiled dataset, which is available in the Supplement, consists of 29 papers and data from two measurement networks, the European

5   Monitoring and Evaluation Programme (EMEP) and Canadian National Atmospheric Chemistry (NAtChem) database, for a total of 73 measurement sites (Table 5). From these studies, we extracted the annual mean Se deposition fluxes, if available, or we used rainwater volume-mean weighted Se concentrations combined with the annual precipitation depths to calculate the deposition flux. If the paper did not provide the annual precipitation depth, we calculated the mean annual precipitation depth for the time period and location of the study from the historical Multi-Source Weighted-Ensemble Precipitation (MSWEP)

10   dataset (Beck et al., 2019). The calculated annual deposition fluxes for 9 sites are extrapolated from less than 12 months of rainwater Se measurements; the majority of sites (64) were measured for longer than a year. For studies that spanned multiple years, a multi-annual mean deposition flux was calculated to compare with the model. Additional metadata was extracted from the papers, including geographic coordinates, the time period, collection methods, sampling frequency, and analytical methods. Despite the fact that the model is based on year 2000 conditions for emission maps and meteorology, we compare the

15   model with data from all measurement years, since the dataset is relatively small. We created surrogate models of the Se wet deposition flux in the grid boxes where measurements were made (as in Sect. 3.3 for the total deposition fluxes). We can then compare the resampled distribution of Se wet deposition in these model grid boxes with the corresponding observation.

5   Results

5.1   Atmospheric Se lifetime

20   From the 400 training runs of SOCOL-AER, we created PCE models of the global and annual mean total Se burden and deposition flux. The LOO error of the global burden is around 0.02 and the LOO error of the deposition flux is on the order of 10^{-6}. The LOO error is low for total deposition since for a mass conserving model total deposition should equal the sum of the emission input parameters. Indeed, all 400 training runs showed very good Se mass conservation: total annual Se deposition fluxes were 98–102% of total emission fluxes. We derive a surrogate model for the atmospheric Se lifetime by dividing the

25   Se burden PCE model by the Se deposition PCE model (Sect. 3.3). This surrogate model shows a higher error (0.35) than the burden and deposition flux PCE models, which would only be reduced by running more training runs (Fig. S2). The surrogate model for the atmospheric lifetime is resampled to calculate the probability distribution ~~for the atmospheric lifetime~~of this output, given our assumptions for the uncertainty ranges of all 34 parameters (Sect. 3.5). The histogram of the atmospheric lifetime (Fig. 4) shows a narrow range for the atmospheric Se lifetime, with median of 4.4 d (days), 2nd percentile value of 2.9

30   d, and 98th percentile value of 6.4 d. Despite the large uncertainty ranges for the reaction rate constants of Se, spanning multiple orders of magnitude (Table 3), the uncertainty of the simulated atmospheric lifetime is less than one order of magnitude.

In order to identify the input parameters that drive the variability in the simulated Se lifetime, we apportion the variance into contributions from the most important parameters (Fig. 5). The most important variable is k_{13}, which is the rate constant for

Table 5. Previous studies measuring Se wet deposition fluxes.

Reference	Location	Reference	Location
Suzuki et al. (1981)	Tokyo, Japan	Arimoto et al. (1985)	Marshall Islands
Cutter and Church (1986)	Delaware and Bermuda	Arimoto et al. (1987)	American Samoa
Dasch and Wolff (1989)	Massachusetts, USA	Heaton et al. (1990)	Rhode Island, USA
Cutter (1993)	Bermuda	Scudlark et al. (1994)	Maryland, USA (2 sites)
Al-Momani et al. (1997)	Antalya, Turkey	Cutter and Cutter (1998)	Mace Head, Ireland
Lawson and Mason (2001)	Maryland, USA	De Gregori et al. (2002)	Valparaiso, Chile
Scudlark et al. (2005)	Maryland, USA	Sakata et al. (2006)	Japan (10 sites)
Shimamura et al. (2007)	Tokyo, Japan	Landing et al. (2010)	Florida, USA
Zhou et al. (2012)	Mt. Heng, China	Liu et al. (2012)	Shigatse, Tibet
Gratz et al. (2013)	Illinois, USA (4 sites)	Pan and Wang (2015)	China (10 sites)
Lynam et al. (2015)	Alberta, Canada	Xing et al. (2017)	Jiaozhou Bay, China
Nie et al. (2017)	Mt. Lushan, China	Blazina et al. (2017)	Plynlimmon, UK
Savage et al. (2017)	Bangladesh and Sri Lanka	Uchiyama et al. (2019)	Tokyo, Japan
Suess et al. (2019)	Pic du Midi, France	Pearson et al. (2019)	Alaska, USA (3 sites)
Savage et al. (2019)	Northern Ireland, UK	NAtChem database	Ontario, Canada (9 sites)
EMEP network	9 sites in Germany and Czech Republic		
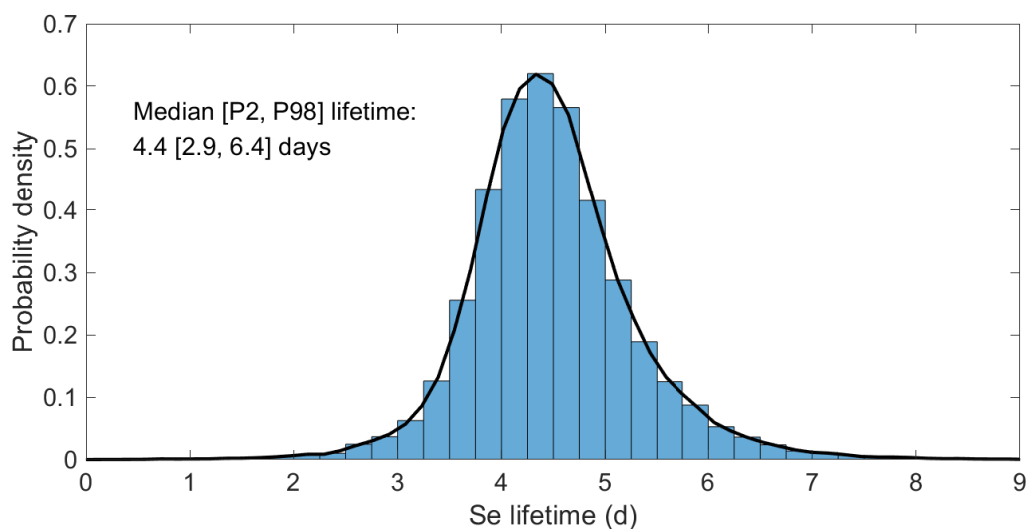
Figure 4. Distribution of the atmospheric Se lifetime, resampled from the surrogate model. Summary statistics (median, 2nd percentile, and 98th percentile value) are listed on the plot.
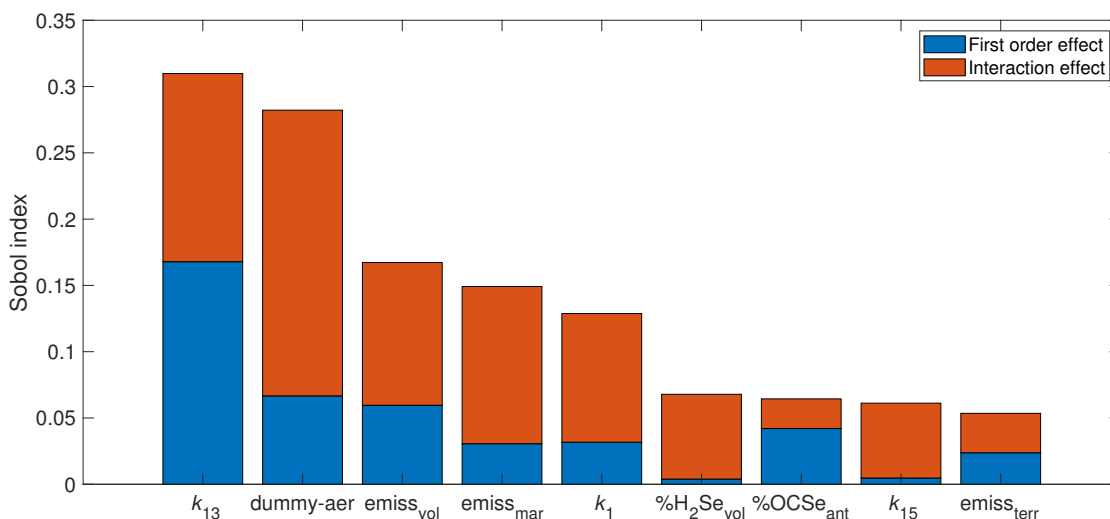
Figure 5. Sensitivity indices of the most important parameters ($S_i^T > 0.05$) for the atmospheric Se lifetime. Total Sobol' indices are partitioned into first order and interaction effects. The total Sobol' indices for the dummy aerosol parameters are aggregated, since it is difficult to separate the effects of the correlated input parameters (aerosol radius and emission magnitude).

the OCSe + OH reaction, followed by the dummy aerosol input parameters. Nonlinearities are also clearly important for the global Se lifetime, since all first order terms only account for 53% of the variance in the Se lifetime, with interaction terms accounting for the other 47%. However, the interaction contribution is made up of many small individual interaction terms. Only two interaction terms account for more than 5% of the total variance in the Se lifetime: the interaction between k_1 and k_{13}

5 (5.3% of variance) and the interaction between the dummy aerosol radius and dummy aerosol emissions (5.0% of variance). Through resampling the surrogate model for the Se lifetime, we investigate the relationships between the Se lifetime and input parameters in Figs. 6 and 7.

Several of the most influential input parameters for the Se lifetime are related to OCSe processes (Fig. 6a–c). Given the median estimates for the reaction rate constants in Table 2, OCSe has the longest lifetime of any gas-phase Se species. Therefore,

10 chemical reaction rates and emissions of OCSe have a strong effect on the overall Se burden and Se lifetime. Slower reaction rates of OCSe with OH (k_{13}) and O (k_1) lead to longer Se lifetimes. Since OH is more prevalent in the lower atmosphere than O, the dependence of the Se lifetime on k_{13} is stronger than the dependence of the lifetime on k_1. The influence of k_{13} on the Se lifetime is mostly saturated above 10^{-12} cm^3 molec^{-1} s^{-1}. Above this threshold in k_{13}, the OCSe burden becomes a minor part of the total Se burden (accounting for less than 2% of total Se), and therefore the Se lifetime is not affected by OCSe pro-

15 cesses. The second order interaction between k_{13} and k_1 is also shown in Fig. 6b. When the reaction of OCSe with OH is fast ($k_{13} > 10^{-12.5}$ cm^3 molec^{-1} s^{-1}), the value of k_1 is not important for the Se lifetime since almost all OCSe will react with OH. In cases when the OCSe + OH reaction is slow ($k_{13} < 10^{-12.5}$ cm^3 molec^{-1} s^{-1}), the value of k_1 has a stronger effect

23

on the Se lifetime since not all OCSe has reacted with OH. The Se lifetime increases for higher fractions of anthropogenic Se emitted as OCSe, again showing that higher OCSe burdens prolong the atmospheric Se lifetime. OCSe has been associated with anthropogenic emissions by only one study, which inferred its presence from a similarity in boiling point with a detected Se species (Pavageau et al., 2002). OCSe has never been identified in the ambient atmosphere; on the other hand, SOCOL-AER predicts maximum OCSe concentrations of sub-ppt levels, which would be difficult to measure analytically. Still, processes related to OCSe, a highly speculative atmospheric Se compound, contribute to the model's variability in the Se lifetime.

After OCSe, the most important input parameters for the atmospheric Se lifetime are related to dummy aerosols. It is difficult to interpret the univariate effects of aerosol emissions and radius (Fig. 7a, b), since by design the two inputs are correlated. We hypothesize that the influence of dummy aerosols on the Se lifetime is related to the uptake rate on these dummy aerosols, which would be proportional to the surface area density. The mass emissions of the aerosols are proportional to volume (r^3), which when divided by dummy aerosol radius yields a metric that is proportional to the emitted surface area (r^2). The Se lifetime increases monotonically with the surface area of emitted dummy aerosols, although the response is not linear (Fig. 7c). We compare this response with the range of available surface area of sulfate aerosol for $10°$ latitude bands, shaded in yellow. For lower dummy aerosol surface areas, the available sulfate aerosol dominates the uptake of gas-phase Se and additional dummy aerosols do not play an important role. The dummy aerosols have a stronger effect on Se lifetime at the upper limit of the sulfate aerosol surface area range. The absence of other aerosols than sulfate leads to a lower Se lifetime in SOCOL-AER. However, the effect of dummy aerosols is not drastic, only increasing the mean Se lifetime from 4.3 to 4.8 d (Fig. 7c).

The other inputs have minor impacts on the global Se lifetime. Stronger emissions of volcanic Se lead to shorter overall global Se lifetimes, while emissions of marine and terrestrial biogenic Se lead to longer Se lifetimes (Fig. 6d–f). In SOCOL-AER, biogenic sources emit DMSe, which is not removed by wet and dry deposition, while volcanic sources emit mainly oxidized Se species, which can be removed by wet and dry deposition. Biogenic emissions must first be oxidized before deposition can occur, which can prolong the Se lifetime. The influence of the two other terms with $S_i^T > 0.05$, the reaction rate constant of CSSe + OH (k_{15}) and the fraction of volcanic emissions emitted as H_2Se, is mainly through interaction terms, and therefore the Se lifetime responds only weakly to univariate variations in these variables (Fig. 6g, h).

5.2 Spatial patterns of Se deposition

Surrogate models for Se deposition in all surface grid boxes were calculated according to Sect. 3.3. After log-transforming Se deposition in 873 grid boxes, the LOO error is below 0.1 in all grid boxes, but still above 0.05 in 354 grid boxes, mainly in polar regions (Fig. S2). However, due to limitations in computational time to run more training runs, we proceeded with the sensitivity analyses of all grid boxes. By resampling the surrogate models, we calculate maps of mean Se deposition and its coefficient of variation of over all input uncertainties (Fig. 8). Selenium deposition is highest in areas close to anthropogenic and volcanic emissions (Fig. 3c, d), which are point sources. The deposition pattern is clearly affected by precipitation: dry areas (e.g., eastern portion of ocean basins, polar regions, and the Sahara Desert) show the lowest Se deposition fluxes globally. Other interesting features of the Se deposition patterns, for example identifying the regions influenced by long-range transport, will be investigated in upcoming studies. In this study, we focus on determining the most important sources of uncertainty to the
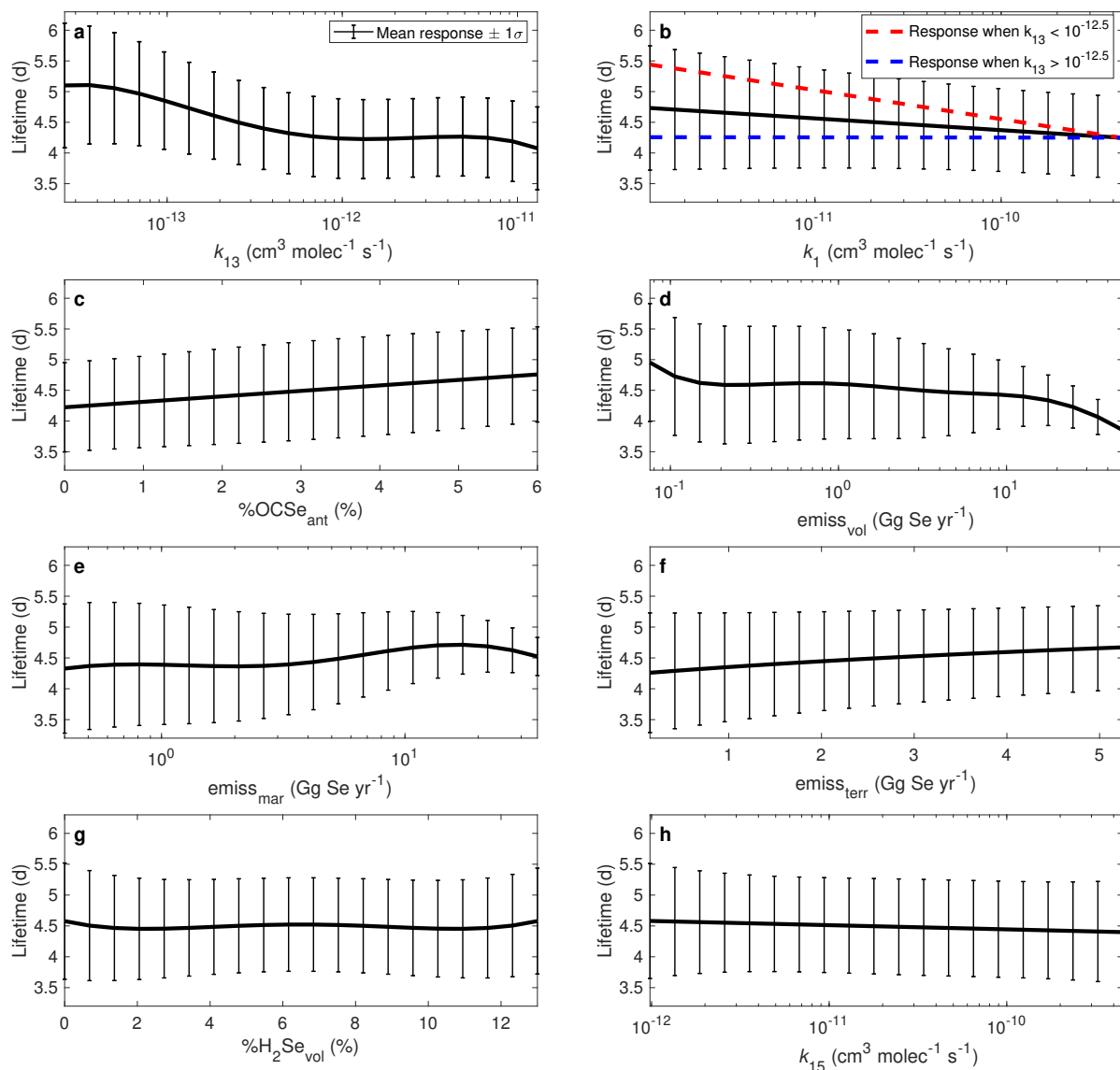
Figure 6. Relationships between the atmospheric Se lifetime and the most important input parameters from Fig. 5. Using resampling techniques (Sect. 3.5), we calculate the mean and standard deviation Se lifetime over the range of input parameters. Interaction effects are illustrated in the k_1 plot (*b*), by grouping the samples into cases where k_{13} is high and low.

simulated deposition maps. The relative uncertainty in simulated deposition, illustrated by the coefficient of variation (Fig. 8b), is highest in areas affected by marine and volcanic emissions because these emission sources have wider uncertainty ranges compared to anthropogenic and terrestrial emissions (Table 3). In the following global sensitivity analysis we can identify the input factors that contribute to the variation in Se deposition in each grid box.
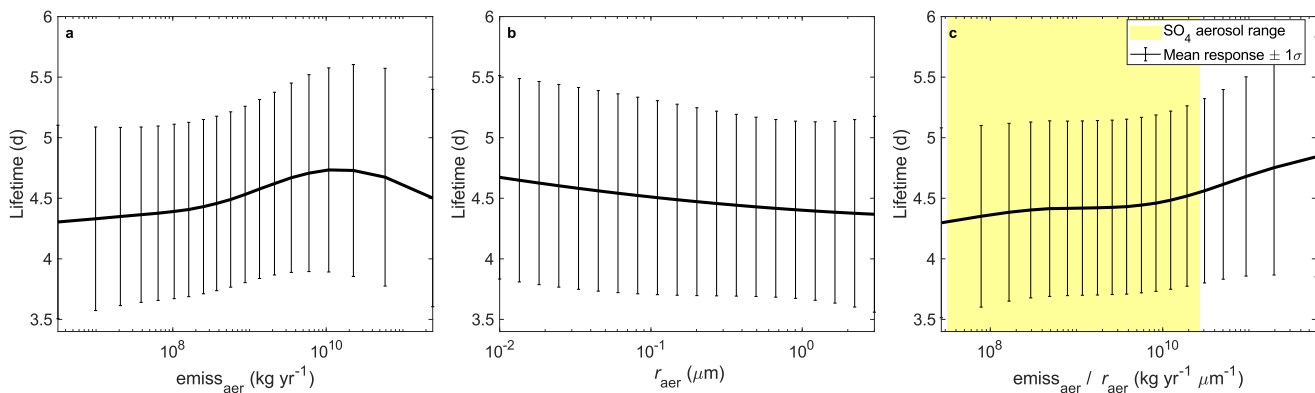
Figure 7. Relationship between the atmospheric Se lifetime and the dummy aerosol input parameters (*a-b*). A quantity related to surface area is shown in (*c*), which also includes the comparable range for sulfate aerosol as yellow shading. For sulfate, this is calculated by dividing the deposited mass of sulfate (equal to emitted mass in steady state) in each 10° latitudinal band by the effective sulfate radius in that latitudinal band. Since aerosol mass emissions are a lognormally distributed quantity, the bounds are infinite and thus we include only the 5th to 95th percentile range in (*a, c*).
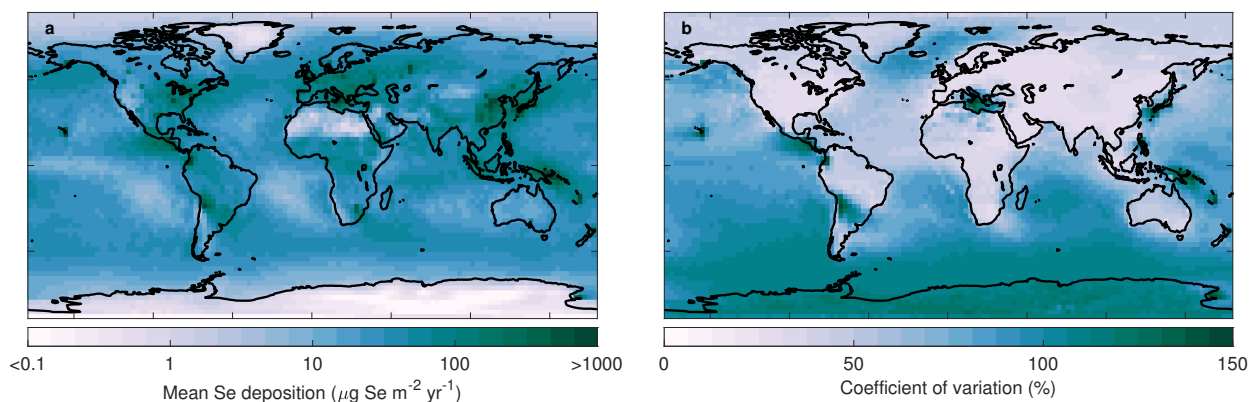
Figure 8. Map of the mean Se deposition (*a*) and associated coefficient of variation (*b*), calculated by resampling surrogate models of deposition in each grid box.

Figures 9 and 10 illustrate the spatial variation in the importance of input parameters. We chose example grid boxes (indicated by blue circles) to illustrate the marginal relationships between Se deposition and the input parameters. Overall, the most important input parameters are the total emissions from each source. In the example grid boxes shown in Fig. 9a–h, Se deposition increases linearly with increasing the emissions from the different source categories. The linear relationship is logical since deposition balances emission in the steady state. In areas that are more remote from emission regions (Sahara, Antarctic, and Arctic), other factors become more important but are still minor compared to the emission inputs. The aerosol accommodation
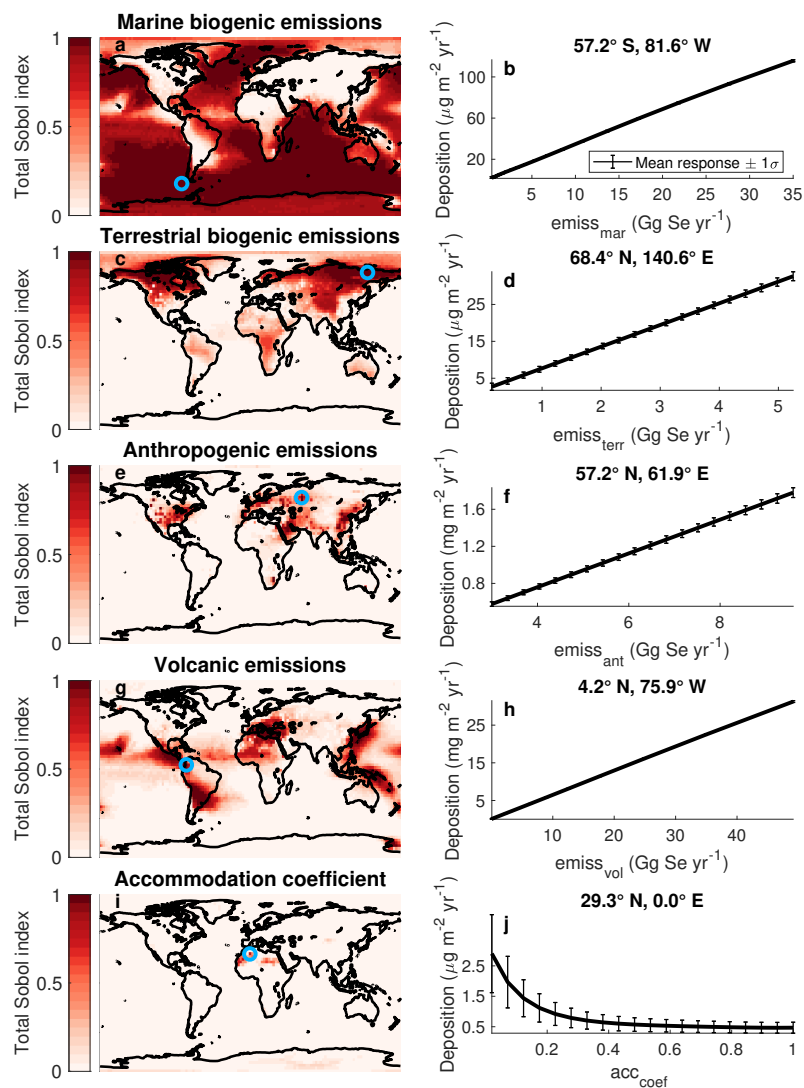
Figure 9. Maps of the total Sobol' indices of emission parameters and the accommodation coefficient for total Se deposition (*left column*). A blue circle indicates the grid box where the total Sobol' index is maximum. The relationship between total Se deposition and the input parameters in that grid box, calculated by resampling the surrogate model for deposition (*right column*). Note that the magnitude of deposition (*y*-axis) varies in each plot, depending on the grid box shown.

coefficient affects areas where the precipitation is very low, for example in the Saharan Desert (Fig. 9i). In these dry regions, total Se deposition is dominated by dry deposition. When the model is run with low accommodation coefficients, less oxidized Se partitions to the particulate phase and more remains in the gas phase. Dry deposition of particles in the 0.1–1 μm diameter size range, within the range of sulfate and dummy aerosols, is slower than gas compounds due to the slower Brownian motion

5   of particles (Seinfeld and Pandis, 2016). Chemical reaction rate constants, specifically the reaction rates of DMSe, impact Se

deposition in polar regions. Slower reaction rates of DMSe with OH (k_6) and O_3 (k_8) enhance deposition over the example Antarctic grid box (Fig. 10a-c). Longer DMSe lifetimes allow more marine Se emissions to reach polar regions, which have little local Se emission sources. The chemical rate coefficients are more important in the Antarctic than the Arctic since there is more O_3 and OH in the Northern Hemisphere than the Southern Hemisphere, meaning that the DMSe lifetime is longer in
5   the Southern Hemisphere than the Northern Hemisphere. Dummy aerosol parameters are only important for Se deposition in the Arctic (Fig. 10d–f). We calculate the relationship between Se deposition and dummy aerosols using the same quantity for emitted surface area of the dummy aerosols as in Fig. 7c. With increasing dummy aerosol surface area, the Se deposition in the Arctic increases, but only after surpassing the threshold of the available sulfate aerosol surface area in that latitude band (Fig. 10e). The dummy aerosols also have a stronger effect on Se deposition when they are emitted in a latitude band closer
10   to the example grid box at 76.7° N (Fig. 10f). Attachment of oxidized Se to dummy aerosols increases the overall lifetime of Se (Sect. 5.1), leading to enhanced transport of Se to the Arctic region. The transport of Se on dummy aerosols does not lead to higher deposition in the Antarctic, perhaps because wet deposition in the Antarctic circumpolar storm track impedes the transport of aerosol poleward.

It is also important to note which input parameters do not influence Se deposition in any of the grid boxes. Variations in the
15   speciation of emissions, photolysis rates, and 15 of the Se reaction rate constants have a negligible influence on deposition.

Although other parameters may play a role in certain grid boxes, the emission parameters are most important on the global scale, evidenced by their higher mean total Sobol' index (Fig. 11). Figure 9 illustrates which regions are affected by different emission sources. Variations of marine emissions impact the most grid boxes; however, their influence is mainly confined to the oceans, coastal areas, and Southern Hemispheric continents. Since the motivation of studying Se deposition is to understand
20   its impact on agricultural soils, we also calculated the mean influence of parameters in pasture and cropland areas (Fig. 11), using maps from Ramankutty et al. (2008). The importance of anthropogenic emissions increases when looking only at pasture and cropland areas, since agricultural areas coincide with human settlement. All four emission parameters show similar levels of importance for agricultural regions, i.e. showing similar mean Sobol' indices for pasture and cropland areas. Therefore, further work in understanding any of these emission processes would be valuable to reducing the uncertainty in deposition
25   fluxes.

5.3   Comparison with deposition flux measurements

With surrogate models of wet Se deposition, we can estimate the modeled wet Se deposition throughout the parametric uncertainty space. By comparing these modeled distributions of wet deposition with observed values, one could constrain input parameters to which deposition is sensitive. However, we do not attempt at this stage to calibrate the parameters to existing
30   measurements because of several challenges in comparing the compiled measurement dataset with the simulations in this study. Firstly, the emissions and meteorology in this study are representative of the year 2000, whereas the measurements were made between 1975 and 2017. Secondly, Se is a difficult element to measure at environmental concentrations, which might lead to inaccurate reported deposition fluxes. The most popular analytical method is inductively coupled plasma mass spectrometry (ICP-MS), which was used for 58 of the 73 sites in the database. However, it is difficult to measure Se with ICP-MS due
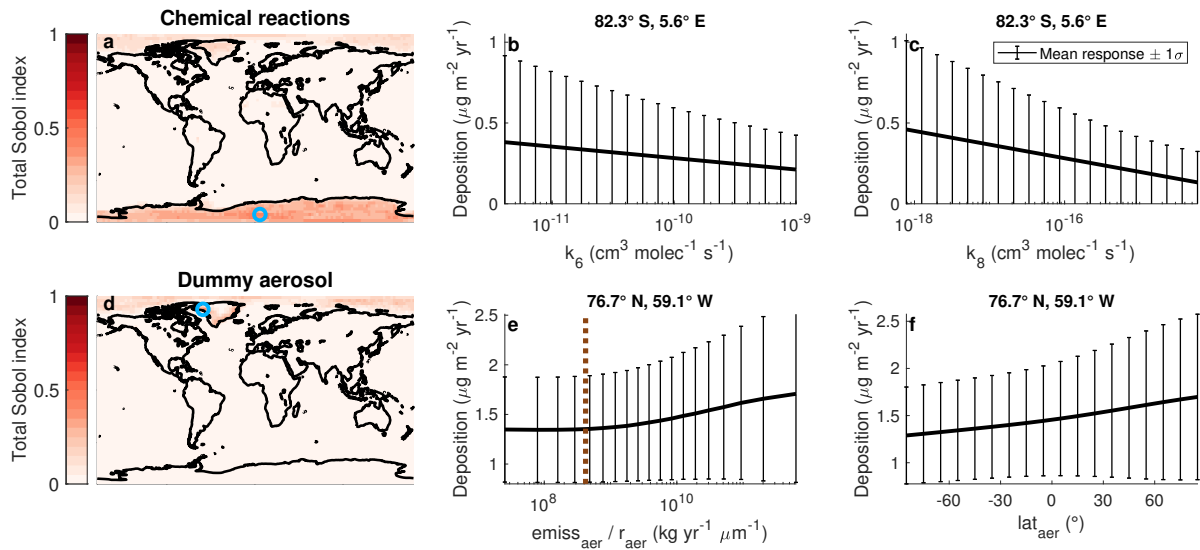
Figure 10. Maps of the total Sobol' indices of reaction rate constants and dummy aerosol parameters for total Se deposition (*left column*). A blue circle indicates the grid box where the aggregated total Sobol' index is maximum. The relationship between total Se deposition and the multiple relevant input parameters within the aggregated index for that grid box is shown (*center and right columns*). The emitted dummy aerosol surface area is compared to the corresponding sulfate quantity in the latitude band of the grid box, shown as a dashed brown line in (*e*).
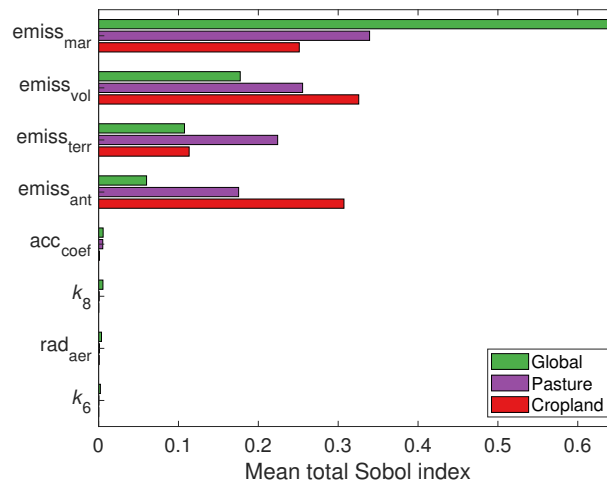
Figure 11. Bar plot summarizing the importance of the input parameters to total Se deposition globally and in agricultural areas. For the cropland and pasture means, Sobol' indices are averaged over grid boxes that are covered by more than 25% cropland or pasture area in the Ramankutty et al. (2008) database.

to the low ionization of Se, the Se signal being split on the five stable isotopes, and especially mass interferences (Winkel et al., 2012). Several studies reported that Se concentrations in rainwater samples were often below the detection limit of the analytical method (e.g., Arimoto et al., 1987; Gratz et al., 2013). Unfortunately, other studies often do not explicitly report the detection limit, the fraction of samples under the detection limit, and how these samples are treated statistically. Thirdly, many

5   of the measurement sites were located in urban locations close to point-source emissions. Due its coarse resolution ($2.8° \times 2.8°$), the model would have difficulty reproducing point values for Se concentrations and deposition fluxes.

Figure 12 compares the measured Se wet deposition fluxes with the resampled median deposition fluxes from the surrogate models, also showing the likely range predicted by the models (defined with bounds of the 2nd and 98th percentile values). With the results from the deposition sensitivity analysis (Fig. 9), we categorize each measurement location by the input pa-

10   rameter that predominates the uncertainty in modeled deposition, indicated by the symbol in Fig. 12. The compiled data and measurements show good agreement at the lower end of deposition values, where the measurement sites are more remote from point-source emissions. The agreement worsens at higher values of observed Se deposition, which correspond to more urban measurement sites. As discussed before, it is not surprising that the model has difficulty matching the measurements for higher observed values of deposition, since the model is coarse resolution and the simulation year may be mismatched from

15   the measurement year. Indeed, the model underestimates several Se deposition measurements in urban sites from East Asia after 2005. Anthropogenic SO_2 emissions, an analogue of anthropogenic Se emissions, has increased in East Asia since 2000 (Smith et al., 2011). Natural emission factors dominate the variability in several of these locations in the year 2000 simulations, likely because the input anthropogenic emission maps do not correspond to the measurement time period. Nevertheless, we find overall that 53% of existing measurements are within the likely range of the model's prediction. The agreement improves to

20   79% when comparing the model range to measurements in background locations, defined as having observed deposition values below 150 $\mu g\,Se\,m^{-2}$. These results provide confidence to SOCOL-AER's predictions of Se deposition fluxes in non-urban locations.

6   Discussion

Through our consideration of model uncertainties related to Se cycling, we derived a median atmospheric Se lifetime of 4.4

25   d. This is the first bottom-up estimate for Se made with a mechanistic global atmospheric model. Our estimate for the Se lifetime matches the global arsenic lifetime calculated in the global model of Wai et al. (2016), namely 4.5 d. This agreement is due to both elements attaching to submicrometer aerosol particles; therefore, their lifetime is determined by the lifetime of these particles. Since Se and arsenic have similar atmospheric sources as well (e.g., metal smelting, coal combustion, volcanoes), it may be possible to draw analogies between their atmospheric cycles. The range of previously estimated Se

30   lifetimes from global atmospheric budgets is between 0.8 and 6 d, similar to our result (Ross, 1985; Mosher and Duce, 1987). The recent value from Mason et al. (2018) of a 0.15 yr (55 d) Se lifetime seems overestimated compared to our results and past budgets, especially since Mason et al. (2018) only consider gas-phase Se in their model, which tends to be shorter lived in the atmosphere than aerosol-bound Se. According to our sensitivity analysis results, the atmospheric Se lifetime could be
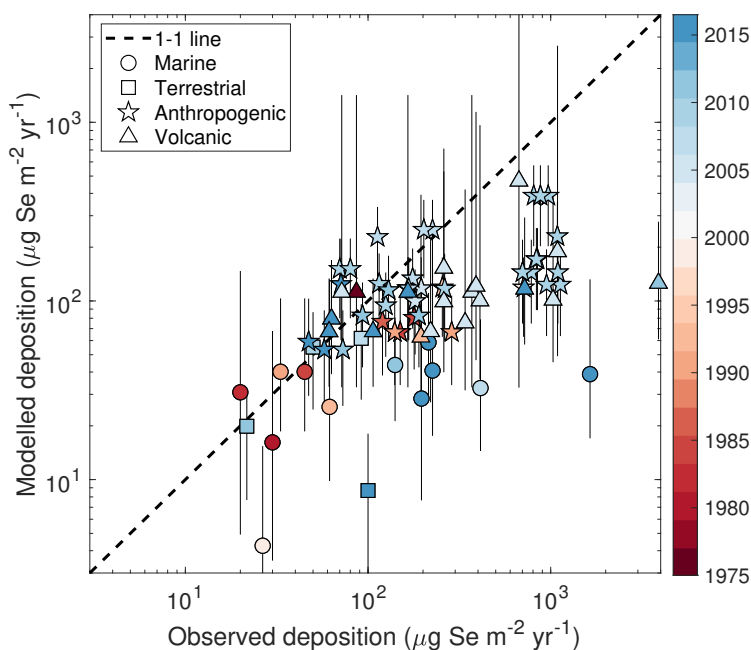
Figure 12. Comparison of wet deposition flux measurements (Table 5) with modeled fluxes. Medians for the modeled values are shown, along with vertical bars representing the 2nd and 98th percentile values. The symbols correspond to the model input parameter that is most important for deposition at the measurement location. The color in the points represents the year when the measurement was taken; multi-year measurements show the middle year.

further constrained by measuring the OCSe + OH reaction rates, and in general knowing more about whether OCSe is present in the atmosphere. Since dummy aerosols also impact the Se lifetime in our model, implementing a more complex tropospheric aerosol parametrization in SOCOL-AER would also further constrain the atmospheric lifetime of Se. However, since the main interest in Se is its atmospheric input to agricultural soils, it may be a higher priority to constrain the input parameters that affect the deposition of Se in agricultural regions rather than the Se lifetime

5   The results of the sensitivity analyses raise an obvious question: why do the input parameters that influence the atmospheric Se lifetime not appear as important for the Se deposition fluxes? One would expect that Se deposition fluxes close to areas of high emissions would be dominated by the magnitude of these emissions. One would also expect that, if anywhere, variation in the Se lifetime would play a role over remote regions, where the amount of locally emitted Se is low and thus the amount that can be transported from emission regions has a larger effect on deposition. However, the range in the atmospheric Se lifetime 10 in our simulations is relatively narrow, between 2.9 and 6.4 d if we consider the 2nd percentile and 98th percentile bounds (Fig. 4). On the other hand, emissions of various Se species can vary by orders of magnitude (Table 3). These larger variations in the amount of emitted Se have a larger impact on deposition than smaller variations in the Se lifetime, even in many remote places. Only in extremely remote areas, for example in the Arctic, do some of the parameters that affect the Se lifetime show

31

up as important, like the dummy aerosols. Parameters with regional rather than global importance for the Se lifetime, like the DMSe reaction parameters, impact deposition of Se in the Antarctic by controlling the amount of transported Se. It is not surprising that the parameter that has the largest impact on lifetime, the OCSe + OH reaction rate constant, has little impact on deposition fluxes, since emissions of OCSe are assumed to be a minor flux of Se (maximum 6% of the anthropogenic emissions flux). Like all sensitivity analyses, the results are dependent on the choice of uncertainty ranges for the different parameters; if we had selected narrower uncertainties for the Se emission sources, the uncertainties of parameters that affect Se lifetime (e.g., chemical reaction rates, dummy aerosols, etc.) may have been more important in remote regions. However, the choice of wide uncertainty ranges for the Se emissions is justified, given the variability in natural emission processes and a lack of field campaigns assessing Se emission fluxes (Sect. 3.1.4). The different results for the two types of sensitivity analyses (lifetime and deposition fluxes) highlight that the "important" parameters to constrain depend on the choice of research question.

It must be noted that our simulations were performed only for the year 2000 and focused on uncertainties in the Se-related input parameters, neglecting variations in the Se cycle due to interannual variability in meteorology and sulfate aerosol properties. In future studies we intend to investigate how the Se cycle varies under different climate conditions.

The global sensitivity analyses in this paper provides clear next steps for atmospheric Se research. The magnitude and spatial distribution of Se emissions remain the most important uncertainty to constrain, in order to improve the predictions of Se deposition patterns. Further investigations of chemical reactivity of Se species or the speciation of emissions are a lower priority, although measuring the speciation of emissions can give mechanistic insights into emission processes. The emission uncertainties could be constrained by conducting field campaigns that either measure emission fluxes of Se close to sources (e.g., Amouroux et al., 2001) or separate Se source contributions at an ambient measurement site through trajectory modeling and/or speciation measurements (e.g., Suess et al., 2019). Our model results can help identify ambient locations that would be interesting to study for field campaigns, by mapping the contribution of the Se emission sources to deposition in different regions (Fig. 9). In addition to new field measurements, we can also compile and reanalyze previously collected data from the literature to evaluate estimates of emission fluxes. Bayesian inverse modeling techniques (e.g., Stohl et al., 2009) could be employed in conjunction with the SOCOL-AER model to provide posterior estimates for Se emission fluxes. Global sensitivity analysis is an invaluable first step before such model calibration techniques, since the parameter dimensionality can be reduced by neglecting non-influential parameters. As shown in Sect. 5.3, the heterogeneity of compiled literature data represents a challenge to comparing models and measurements. Therefore, standardized measurement techniques and adequate reporting of sampling, analytical, and post-processing methods are required so that the model is not calibrated to an errant measurement.

7 Conclusions

Now that it includes Se cycling, the SOCOL-AER model can be used to predict Se transport and deposition globally. We created surrogate PCE-based models that are able to predict the output of the model throughout the uncertainty space of the input parameters. With these surrogate models, we determined that the atmospheric Se lifetime is around 4.4 d, similar to the lifetime of submicron aerosol particles in the atmosphere. Assuming that longitudinal wind speeds are around $10\,\mathrm{m\,s^{-1}}$

(Jacob, 1999), the likely Se lifetime range of 2.9–6.4 d corresponds to a distance of 2500–5000 km that Se is transported in the atmosphere. The global sensitivity analysis of Se deposition fluxes shows that reducing uncertainties in Se emissions would lead to the biggest reductions in the uncertainty of deposition maps. Field measurements that elucidate and quantify Se emission processes should be prioritized, so that model predictions of Se deposition maps can be improved. Available measurements of Se in rainwater are within the likely range of model results at 79% of background sites; remaining discrepancies may be due to the time period of the simulations in this study, the coarse resolution of the model, and analytical challenges leading to measurement inaccuracies. In a future study, SOCOL-AER can be applied to different time periods to investigate how Se deposition has changed due to variations in anthropogenic emissions and climate.

References

Agrawalla, B. and Setser, D.: Infrared chemiluminescence and laser-induced fluorescence studies of energy disposal by reactions of F and Cl atoms with H_2S (D_2S), H_2Se, H_2O (D_2O), and CH_3OH, Journal of Physical Chemistry, 90, 2450–2462, https://doi.org/10.1021/j100402a039, 1986.

5   Agrawalla, B. and Setser, D.: Vibrational distributions and rate constants from reactions of oxygen atoms with HI, GeH_4, SiH_4, H_2Se, and H_2S, The Journal of Chemical Physics, 86, 5421–5432, https://doi.org/10.1063/1.452566, 1987.

Al-Momani, I., Güllü, G., Ölmez, I., Eler, Ü., Örtel, E., Sirin, G., and Tuncel, G.: Chemical composition of eastern Mediterranean aerosol and precipitation: Indications of long-range transport, Pure and Applied Chemistry, 69, 41–46, https://doi.org/10.1351/pac199769010041, 1997.

10   Amouroux, D. and Donard, O. F.: Maritime emission of selenium to the atmosphere in eastern Mediterranean seas, Geophysical Research Letters, 23, 1777–1780, https://doi.org/10.1029/96GL01271, 1996.

Amouroux, D. and Donard, O. F.: Evasion of selenium to the atmosphere via biomethylation processes in the Gironde estuary, France, Marine Chemistry, 58, 173–188, https://doi.org/10.1016/S0304-4203(97)00033-9, 1997.

Amouroux, D., Liss, P. S., Tessier, E., Hamren-Larsson, M., and Donard, O. F.: Role of oceans as biogenic sources of selenium, Earth and
15   Planetary Science Letters, 189, 277–283, https://doi.org/10.1016/S0012-821X(01)00370-3, 2001.

Andres, R. and Kasgnoc, A.: A time-averaged inventory of subaerial volcanic sulfur emissions, Journal of Geophysical Research: Atmospheres, 103, 25 251–25 261, https://doi.org/10.1029/98JD02091, 1998.

Arimoto, R., Duce, R., Ray, B., and Unni, C.: Atmospheric trace elements at Enewetak Atoll: 2. Transport to the ocean by wet and dry deposition, Journal of Geophysical Research: Atmospheres, 90, 2391–2408, https://doi.org/10.1029/JD090iD01p02391, 1985.

20   Arimoto, R., Duce, R. A., Ray, B. J., Hewitt, A. D., and Williams, J.: Trace elements in the atmosphere of American Samoa: Concentrations and deposition to the tropical South Pacific, Journal of Geophysical Research: Atmospheres, 92, 8465–8479, https://doi.org/10.1029/JD092iD07p08465, 1987.

Ariya, P. A., Amyot, M., Dastoor, A., Deeds, D., Feinberg, A., Kos, G., Poulain, A., Ryjkov, A., Semeniuk, K., and Subir, M.: Mercury physicochemical and biogeochemical transformation in the atmosphere and at atmospheric interfaces: A review and future directions,
25   Chemical Reviews, 115, 3760–3802, https://doi.org/10.1021/cr500667e, 2015.

Atkinson, R., Perry, R., and Pitts Jr, J.: Rate constants for the reaction of OH radicals with COS, CS_2 and CH_3SCH_3 over the temperature range 299–430 K, Chemical Physics Letters, 54, 14–18, https://doi.org/10.1016/0009-2614(78)85653-X, 1978.

Atkinson, R., Aschmann, S. M., Hasegawa, D., Thompson-Eagle, E. T., and Frankenberger Jr, W. T.: Kinetics of the atmospherically important reactions of dimethyl selenide, Environmental Science & Technology, 24, 1326–1332, https://doi.org/10.1021/es00079a005, 1990.

30   Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I., McVicar, T. R., and Adler, R. F.: MSWEP V2 global 3-hourly 0.1 precipitation: methodology and quantitative assessment, Bulletin of the American Meteorological Society, 100, 473–500, https://doi.org/10.1175/BAMS-D-17-0138.1, 2019.

Belyaev, A., Sozin, A., Aleksandrov, Y., and Churbanov, M.: Homogeneous oxidation of hydrogen selenide with ozone in an argon-oxygen medium, Russian Journal of Applied Chemistry, 85, 1919–1922, https://doi.org/10.1134/S107042721212021X, 2012.

35   Berveiller, M., Sudret, B., and Lemaire, M.: Stochastic finite elements: a non intrusive approach by regression, European Journal of Computational Mechanics, 15, 81–92, https://doi.org/10.3166/remn.15.81-92, 2006.

Blatman, G. and Sudret, B.: An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis, Probabilistic Engineering Mechanics, 25, 183–197, https://doi.org/10.1016/j.probengmech.2009.10.003, 2010a.

Blatman, G. and Sudret, B.: Efficient computation of global sensitivity indices using sparse polynomial chaos expansions, Reliability Engineering & System Safety, 95, 1216 – 1229, https://doi.org/10.1016/j.ress.2010.06.015, 2010b.

5 Blatman, G. and Sudret, B.: Adaptive sparse polynomial chaos expansion based on least angle regression, Journal of Computational Physics, 230, 2345–2367, https://doi.org/10.1016/j.jcp.2010.12.021, 2011a.

Blatman, G. and Sudret, B.: Principal component analysis and Least Angle Regression in spectral stochastic finite element analysis, in: Proc. 11th Int. Conf. on Applications of Stat. and Prob. in Civil Engineering (ICASP11), Zurich, Switzerland, edited by Faber, M., Köhler, J., and Nishijima, K., 2011b.

10 Blatman, G. and Sudret, B.: Sparse polynomial chaos expansions of vector-valued response quantities, in: Proc. 11th Int. Conf. Struct. Safety and Reliability (ICOSSAR'2013), New York, USA, edited by Deodatis, G., 2013.

Blazina, T., Sun, Y., Voegelin, A., Lenz, M., Berg, M., and Winkel, L. H.: Terrestrial selenium distribution in China is potentially linked to monsoonal climate, Nature Communications, 5, 4717, https://doi.org/10.1038/ncomms5717, 2014.

Blazina, T., Läderach, A., Jones, G. D., Sodemann, H., Wernli, H., Kirchner, J. W., and Winkel, L. H.: Marine primary productivity as

15 a potential indirect source of selenium and other trace elements in atmospheric deposition, Environmental Science & Technology, 51, 108–118, https://doi.org/10.1021/acs.est.6b03063, 2017.

Burkholder, J., Sander, S., Abbatt, J., Barker, J., Huie, R., Kolb, C., Kurylo, M., Orkin, V., Wilmouth, D., and Wine, P.: Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies, Evaluation No. 18, Tech. Rep. 15-10, Jet Propulsion Laboratory, Pasadena, https://jpldataeval.jpl.nasa.gov/pdf/JPL_Publication_15-10.pdf, last access: 30 August 2019, 2015.

20 Cheng, B. and Lee, Y.: Rate constant of OH+ OCS reaction over the temperature range 255–483 K, International Journal of Chemical Kinetics, 18, 1303–1314, https://doi.org/10.1002/kin.550181202, 1986.

Combs, G. F.: Selenium in global food systems, British Journal of Nutrition, 85, 517–547, https://doi.org/10.1079/BJN2000280, 2001.

Conde, J. and Sanz Alaejos, M.: Selenium concentrations in natural and environmental waters, Chemical Reviews, 97, 1979–2004, https://doi.org/10.1021/cr960100g, 1997.

25 Cox, R. and Sheppard, D.: Reactions of OH radicals with gaseous sulphur compounds, Nature, 284, 330, https://doi.org/10.1038/284330a0, 1980.

Cutter, G. A.: Metalloids in wet deposition on Bermuda: concentrations, sources, and fluxes, Journal of Geophysical Research: Atmospheres, 98, 16 777–16 786, https://doi.org/10.1029/93JD01689, 1993.

Cutter, G. A. and Church, T. M.: Selenium in western Atlantic precipitation, Nature, 322, 720, https://doi.org/10.1038/322720a0, 1986.

30 Cutter, G. A. and Cutter, L. S.: Metalloids in the high latitude North Atlantic Ocean: sources and internal cycling, Marine Chemistry, 61, 25–36, https://doi.org/10.1016/S0304-4203(98)00005-X, 1998.

Dasch, J. M. and Wolff, G. T.: Trace inorganic species in precipitation and their potential use in source apportionment studies, Water, Air, and Soil Pollution, 43, 401–412, 1989.

De Gregori, I., Lobos, M. G., and Pinochet, H.: Selenium and its redox speciation in rainwater from sites of Valparaıso region in Chile,

35 impacted by mining activities of copper ores, Water Research, 36, 115–122, https://doi.org/10.1016/S0043-1354(01)00240-8, 2002.

Dentener, F., Kinne, S., Bond, T., Boucher, O., Cofala, J., Generoso, S., Ginoux, P., Gong, S., Hoelzemann, J., and Ito, A.: Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for AeroCom, Atmospheric Chemistry and Physics, 6, 4321–4344, https://doi.org/10.5194/acp-6-4321-2006, 2006.

Dinh, Q. T., Cui, Z., Huang, J., Tran, T. A. T., Wang, D., Yang, W., Zhou, F., Wang, M., Yu, D., and Liang, D.: Selenium distribution in the Chinese environment and its relationship with human health: a review, Environment International, 112, 294–309, https://doi.org/10.1016/j.envint.2017.12.035, 2018.

Egorova, T., Rozanov, E., Zubov, V., and Karol, I.: Model for investigating ozone trends (MEZON), Izvestiya Atmospheric and Oceanic Physics, 39, 277–292, 2003.

Eldred, R. A.: Comparison of Selenium and Sulfur at Remote Sites, Journal of the Air & Waste Management Association, 47, 204–211, https://doi.org/10.1080/10473289.1997.10464423, 1997.

Eyring, V., Chipperfield, M. P., Giorgetta, M. A., Kinnison, D. E., Manzini, E., Matthes, K., Newman, P. A., Pawson, S., Shepherd, T. G., and Waugh, D. W.: Overview of the new CCMVal reference and sensitivity simulations in support of upcoming ozone and climate assessments and the planned SPARC CCMVal report, SPARC Newsletter, 30, 20–26, http://www.pa.op.dlr.de/CCMVal/Forcings/NewCCMValSimulations_Nov2007_FINAL.pdf, last access: 9 July 2019, 2008.

Fairweather-Tait, S. J., Bao, Y., Broadley, M. R., Collings, R., Ford, D., Hesketh, J. E., and Hurst, R.: Selenium in human health and disease, Antioxidants & Redox Signaling, 14, 1337–1383, https://doi.org/10.1089/ars.2010.3275, 2011.

Feinberg, A., Moustapha, M., Stenke, A., Sudret, B., Peter, T., and Winkel, L. H. E.: Simulation data used for sensitivity analysis of atmospheric selenium in SOCOL-AER, https://doi.org/10.3929/ethz-b-000357105, 2019a.

Feinberg, A., Sukhodolov, T., Luo, B.-P., Rozanov, E., Winkel, L. H. E., Peter, T., and Stenke, A.: Improved tropospheric and stratospheric sulfur cycle in the aerosol-chemistry-climate model SOCOL-AERv2, Geoscientific Model Development Discussions, 2019, 1–39, https://doi.org/10.5194/gmd-2019-138, 2019b.

Finn, E. J. and King, G.: The absorption spectrum of carbonyl selenide: Sub-Rydberg transitions, Journal of Molecular Spectroscopy, 56, 39–51, https://doi.org/10.1016/0022-2852(75)90201-5, 1975.

Floor, G. H. and Román-Ross, G.: Selenium in volcanic environments: a review, Applied Geochemistry, 27, 517–531, https://doi.org/10.1016/j.apgeochem.2011.11.010, 2012.

Fordyce, F. M.: Selenium deficiency and toxicity in the environment, in: Essentials of Medical Geology, pp. 375–416, Springer, 2013.

Ghanem, R. and Spanos, P.: Stochastic Finite Elements: A Spectral Approach, Courier Dover Publications, Mineola, 2nd edn., 2003.

Goodeve, C. and Stein, N.: The absorption spectra and the optical dissociation of the hydrides of the oxygen group, Transactions of the Faraday Society, 27, 393–402, https://doi.org/10.1039/TF9312700393, 1931.

Goupy, J. and Creighton, L.: Introduction aux plans d'expériences, vol. 3, Dunod Paris, 2006.

Gratz, L. E., Keeler, G. J., Morishita, M., Barres, J. A., and Dvonch, J. T.: Assessing the emission sources of atmospheric mercury in wet deposition across Illinois, Science of the Total Environment, 448, 120–131, https://doi.org/10.1016/j.scitotenv.2012.11.011, 2013.

Haygarth, P., Cooke, A., Jones, K., Harrison, A., and Johnston, A.: Long-term change in the biogeochemical cycling of atmospheric selenium: Deposition to plants and soil, Journal of Geophysical Research: Atmospheres, 98, 16 769–16 776, https://doi.org/10.1029/93JD01023, 1993.

Heaton, R. W., Rahn, K. A., and Lowenthal, D. H.: Determination of trace elements, including regional tracers, in Rhode Island precipitation, Atmospheric Environment. Part A. General Topics, 24, 147–153, https://doi.org/10.1016/0960-1686(90)90450-2, 1990.

Horowitz, L. W., Walters, S., Mauzerall, D. L., Emmons, L. K., Rasch, P. J., Granier, C., Tie, X., Lamarque, J., Schultz, M. G., and Tyndall, G. S.: A global simulation of tropospheric ozone and related tracers: Description and evaluation of MOZART, version 2, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/10.1029/2002JD002853, 2003.

Jacob, D. J.: Introduction to Atmospheric Chemistry, Princeton University Press, 1999.

Jones, G. D., Droz, B., Greve, P., Gottschalk, P., Poffet, D., McGrath, S. P., Seneviratne, S. I., Smith, P., and Winkel, L. H.: Selenium deficiency risk predicted to increase under future climate change, Proceedings of the National Academy of Sciences, 114, 2848–2853, https://doi.org/10.1073/pnas.1611576114, 2017.

Kapur, J.: Maximum-entropy Models in Science and Engineering, Wiley, 1989.

5 Kerkweg, A., Buchholz, J., Ganzeveld, L., Pozzer, A., Tost, H., and Jöckel, P.: An implementation of the dry removal processes DRY DEPosition and SEDImentation in the Modular Earth Submodel System (MESSy), Atmospheric Chemistry and Physics, 6, 4617–4632, https://doi.org/10.5194/acp-6-4617-2006, 2006.

Kerkweg, A., Buchholz, J., Ganzeveld, L., Pozzer, A., Tost, H., and Jöckel, P.: Corrigendum to "Technical Note: An implementation of the dry removal processes DRY DEPosition and SEDImentation in the Modular Earth Submodel System (MESSy)" published in Atmos.

10 Chem. Phys., 6, 4617–4632, 2006, Atmospheric Chemistry and Physics, 9, 9569–9569, https://doi.org/10.5194/acp-9-9569-2009, 2009.

King, G. and Srikameswaran, K.: Analysis of the 4500 Å absorption system (R system) of carbon diselenide, Journal of Molecular Spectroscopy, 31, 269–291, https://doi.org/10.1016/0022-2852(69)90359-2, 1969.

Kurylo, M. J.: Flash photolysis resonance fluorescence investigation of the reactions of OH radicals with OCS and CS_2, Chemical Physics Letters, 58, 238–242, https://doi.org/10.1016/0009-2614(78)80285-1, 1978.

15 Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., and Owen, B.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, Atmospheric Chemistry and Physics, 10, 7017–7039, https://doi.org/10.5194/acp-10-7017-2010, 2010.

Lana, A., Bell, T., Simó, R., Vallina, S., Ballabrera-Poy, J., Kettle, A., Dachs, J., Bopp, L., Saltzman, E., and Stefels, J.: An updated climatology of surface dimethlysulfide concentrations and emission fluxes in the global ocean, Global Biogeochemical Cycles, 25,

20 https://doi.org/10.1029/2010GB003850, 2011.

Landing, W., Caffrey, J., Nolek, S., Gosnell, K., and Parker, W.: Atmospheric wet deposition of mercury and other trace elements in Pensacola, Florida, Atmospheric Chemistry and Physics, 10, 4867–4877, https://doi.org/10.5194/acp-10-4867-2010, 2010.

Lawson, N. M. and Mason, R. P.: Concentration of mercury, methylmercury, cadmium, lead, arsenic, and selenium in the rain and stream water of two contrasting watersheds in western Maryland, Water Research, 35, 4039–4052, https://doi.org/10.1016/S0043-1354(01)00140-3,

25 2001.

Le Gratiet, L., Marelli, S., and Sudret, B.: Metamodel-based sensitivity analysis: polynomial chaos expansions and Gaussian processes, chap. 38, pp. 1289–1325, Springer International Publishing, cham, Switzerland, 2017.

Lee, L., Carslaw, K., Pringle, K., Mann, G., and Spracklen, D.: Emulation of a complex global aerosol model to quantify sensitivity to uncertain parameters, Atmospheric Chemistry and Physics, 11, 12 253–12 273, https://doi.org/10.5194/acp-11-12253-2011, 2011.

30 Lee, L., Carslaw, K., Pringle, K., and Mann, G.: Mapping the uncertainty in global CCN using emulation, Atmospheric Chemistry and Physics, 12, 9739–9751, https://doi.org/10.5194/acp-12-9739-2012, 2012.

Lee, L., Pringle, K., Reddington, C., Mann, G., Stier, P., Spracklen, D., Pierce, J., and Carslaw, K.: The magnitude and causes of uncertainty in global model simulations of cloud condensation nuclei, Atmospheric Chemistry and Physics, 13, 8879–8914, https://doi.org/10.5194/acp-13-8879-2013, 2013.

35 Leu, M.-T. and Smith, R. H.: Kinetics of the gas-phase reaction between hydroxyl and carbonyl sulfide over the temperature range 300-517 K, The Journal of Physical Chemistry, 85, 2570–2575, https://doi.org/10.1021/j150617a031, 1981.

Li, S., Chwee, T. S., and Fan, W. Y.: FTIR studies of $O(^3P)$ atom reactions with CSe_2, $SCSe$, and $OCSe$, The Journal of Physical Chemistry A, 109, 11 815–11 822, https://doi.org/10.1021/jp055113p, 2005.

Lin, S.-J. and Rood, R. B.: Multidimensional flux-form semi-Lagrangian transport schemes, Monthly Weather Review, 124, 2046–2070, https://doi.org/10.1175/1520-0493(1996)124<2046:MFFSLT>2.0.CO;2, 1996.

Liu, J. Q., Yang, Y. J., Di, Y., Yang, J., Huang, W. W., Wen, T. X., and Li, Y. W.: Trace Elements of Precipitation in the Shigatse Region, Southern Tibetan Plateau, Advanced Materials Research, 518–523, 1652–1656, https://doi.org/10.4028/www.scientific.net/AMR.518-523.1652, 2012.

Loeppky, J. L., Sacks, J., and Welch, W. J.: Choosing the sample size of a computer experiment: A practical guide, Technometrics, 51, 366–376, https://doi.org/10.1198/TECH.2009.08040, 2009.

Låg, J. and Steinnes, E.: Regional distribution of selenium and arsenic in humus layers of Norwegian forest soils, Geoderma, 20, 3–14, https://doi.org/10.1016/0016-7061(78)90045-9, 1978.

Lynam, M. M., Dvonch, J. T., Barres, J. A., Morishita, M., Legge, A., and Percy, K.: Oil sands development and its impact on atmospheric wet deposition of air pollutants to the Athabasca Oil Sands Region, Alberta, Canada, Environmental Pollution, 206, 469–478, https://doi.org/10.1016/j.envpol.2015.07.032, 2015.

Marelli, S. and Sudret, B.: UQLab: A framework for uncertainty quantification in Matlab, in: Vulnerability, uncertainty, and risk: quantification, mitigation, and management, pp. 2554–2563, 2014.

Marelli, S., Lamas, C., Konakli, K., Mylonas, C., Wiederkehr, P., and Sudret, B.: UQLab user manual – Sensitivity analysis, Tech. rep., Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich, report #UQLab-V1.2-106, 2019.

Marshall, L., Johnson, J. S., Mann, G. W., Lee, L., Dhomse, S. S., Regayre, L., Yoshioka, M., Carslaw, K. S., and Schmidt, A.: Exploring How Eruption Source Parameters Affect Volcanic Radiative Forcing Using Statistical Emulation, Journal of Geophysical Research: Atmospheres, 124, 964–985, https://doi.org/10.1029/2018JD028675, 2019.

Mason, R. P., Soerensen, A. L., DiMento, B. P., and Balcom, P. H.: The Global Marine Selenium Cycle: Insights From Measurements and Modeling, Global Biogeochemical Cycles, 32, 1720–1737, https://doi.org/10.1029/2018GB006029, 2018.

McKay, M. D., Beckman, R. J., and Conover, W. J.: Comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics, 21, 239–245, https://doi.org/10.2307/1268522, 1979.

Mosher, B. W. and Duce, R. A.: Vapor phase and particulate selenium in the marine atmosphere, Journal of Geophysical Research: Oceans, 88, 6761–6768, https://doi.org/10.1029/JC088iC11p06761, 1983.

Mosher, B. W. and Duce, R. A.: A global atmospheric selenium budget, Journal of Geophysical Research: Atmospheres, 92, 13 289–13 298, https://doi.org/10.1029/JD092iD11p13289, 1987.

Nie, X., Wang, Y., Li, Y., Sun, L., Li, T., Yang, M., Yang, X., and Wang, W.: Characteristics and impacts of trace elements in atmospheric deposition at a high-elevation site, southern China, Environmental Science and Pollution Research, 24, 22 839–22 851, https://doi.org/10.1007/s11356-017-8791-1, 2017.

Nightingale, P. D., Malin, G., Law, C. S., Watson, A. J., Liss, P. S., Liddicoat, M. I., Boutin, J., and Upstill-Goddard, R. C.: In situ evaluation of air-sea gas exchange parameterizations using novel conservative and volatile tracers, Global Biogeochemical Cycles, 14, 373–387, https://doi.org/10.1029/1999GB900091, 2000.

Nriagu, J. O.: A global assessment of natural sources of atmospheric trace metals, Nature, 338, 47, https://doi.org/10.1038/338047a0, 1989.

Nriagu, J. O. and Pacyna, J. M.: Quantitative assessment of worldwide contamination of air, water and soils by trace metals, Nature, 333, 134, https://doi.org/10.1038/333134a0, 1988.

Pan, Y. and Wang, Y.: Atmospheric wet and dry deposition of trace elements at 10 sites in Northern China, Atmospheric Chemistry and Physics, 15, 951–972, https://doi.org/10.5194/acp-15-951-2015, 2015.

Pavageau, M.-P., Pecheyran, C., Krupp, E. M., Morin, A., and Donard, O. F.: Volatile metal species in coal combustion flue gas, Environmental Science & Technology, 36, 1561–1573, https://doi.org/10.1021/es015595s, 2002.

Pavageau, M.-P., Morin, A., Séby, F., Guimon, C., Krupp, E., Pécheyran, C., Poulleau, J., and Donard, O. F.: Partitioning of metal species during an enriched fuel combustion experiment. Speciation in the gaseous and particulate phases, Environmental Science & Technology, 38, 2252–2263, https://doi.org/10.1021/es034408i, 2004.

Pearson, C., Howard, D., Moore, C., and Obrist, D.: Mercury and trace metal wet deposition across five stations in Alaska: controlling factors, spatial patterns, and source regions, Atmospheric Chemistry and Physics, 19, 6913–6929, https://doi.org/10.5194/acp-19-6913-2019, 2019.

Petersen, G. and Munthe, J.: Atmospheric mercury species over central and northern Europe. Model calculations and comparison with observations from the Nordic air and precipitation network for 1987 and 1988, Atmospheric Environment, 29, 47–67, https://doi.org/10.1016/1352-2310(94)00223-8, 1995.

Pham, M., Müller, J.-F., Brasseur, G. P., Granier, C., and Mégie, G.: A three-dimensional study of the tropospheric sulfur cycle, Journal of Geophysical Research: Atmospheres, 100, 26 061–26 092, https://doi.org/10.1029/95JD02095, 1995.

Rael, R. M., Tuzaon, E. C., and Frankenberger Jr, W. T.: Gas-phase reactions of dimethyl selenide with ozone and the hydroxyl and nitrate radicals, Atmospheric Environment, 30, 1221–1232, https://doi.org/10.1016/1352-2310(95)00447-5, 1996.

Ramankutty, N., Evan, A. T., Monfreda, C., and Foley, J. A.: Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000, Global Biogeochemical Cycles, 22, https://doi.org/10.1029/2007GB002952, 2008.

Rayner, N., Parker, D. E., Horton, E., Folland, C., Alexander, L., Rowell, D., Kent, E., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/10.1029/2002JD002670, 2003.

Revell, L. E., Stenke, A., Tummon, F., Feinberg, A., Rozanov, E., Peter, T., Abraham, N. L., Akiyoshi, H., Archibald, A. T., and Butchart, N.: Tropospheric ozone in CCMI models and Gaussian process emulation to understand biases in the SOCOLv3 chemistry–climate model, Atmospheric Chemistry and Physics, 18, 16 155–16 172, https://doi.org/10.5194/acp-18-16155-2018, 2018.

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., et al.: The atmospheric general circulation model ECHAM 5. PART I: Model description, Tech. Rep. 349, Max-Planck-Institut für Meteorologie, Hamburg, http://www.mpimet.mpg.de/fileadmin/publikationen/Reports/max_scirep_349.pdf, last access: 9 July 2019, 2003.

Ross, H. B.: An atmospheric selenium budget for the region 30° N to 90° N, Tellus B, 37, 78–90, https://doi.org/10.3402/tellusb.v37i2.14999, 1985.

Rumble, J.: CRC Handbook of Chemistry and Physics, CRC Press, 2017.

Ryan, E., Wild, O., Voulgarakis, A., and Lee, L.: Fast sensitivity analysis methods for computationally expensive models with multi-dimensional output, Geoscientific Model Development, 11, 3131–3146, https://doi.org/10.5194/gmd-11-3131-2018, 2018.

Sakata, M., Marumoto, K., Narukawa, M., and Asakura, K.: Regional variations in wet and dry deposition fluxes of trace elements in Japan, Atmospheric Environment, 40, 521–531, https://doi.org/10.1016/j.atmosenv.2005.09.066, 2006.

Saltelli, A., Tarantola, S., and Campolongo, F.: Sensitivity anaysis as an ingredient of modeling, Statistical Science, 15, 377–395, https://www.jstor.org/stable/2676831, 2000.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S.: Global Sensitivity Analysis: the Primer, John Wiley & Sons, 2008.

Savage, L., Carey, M., Hossain, M., Islam, M. R., de Silva, P. M. C., Williams, P. N., and Meharg, A. A.: Elevated trimethylarsine oxide and inorganic arsenic in northern hemisphere summer monsoonal wet deposition, Environmental Science & Technology, 51, 12 210–12 218, https://doi.org/10.1021/acs.est.7b04356, 2017.

Savage, L., Carey, M., Williams, P. N., and Meharg, A. A.: Maritime Deposition of Organic and Inorganic Arsenic, Environmental Science & Technology, 53, 7288–7295, https://doi.org/10.1021/acs.est.8b06335, 2019.

Schultz, M. and Rast, S.: Emission data sets and methodologies for estimating emissions, REanalysis of the TROpospheric chemical composition over the past 40 years, A long-term global modeling study of tropospheric chemistry funded under the 5th EU framework programme, Tech. rep., EVK2-CT-2002-00170, 2007.

Scudlark, J. R., Conko, K. M., and Church, T. M.: Atmospheric wet deposition of trace elements to Chesapeake Bay: CBAD study year 1 results, Atmospheric Environment, 28, 1487–1498, https://doi.org/10.1016/1352-2310(94)90209-7, 1994.

Scudlark, J. R., Rice, K. C., Conko, K. M., Bricker, O. P., and Church, T. M.: Transmission of atmospherically derived trace elements through an undeveloped, forested Maryland watershed, Water, Air, and Soil Pollution, 163, 53–79, https://doi.org/10.1007/s11270-005-8135-5, 2005.

Seinfeld, J. H. and Pandis, S. N.: Atmospheric Chemistry and Physics: from Air Pollution to Climate Change, John Wiley & Sons, 2016.

Sheng, J., Weisenstein, D. K., Luo, B., Rozanov, E., Stenke, A., Anet, J., Bingemer, H., and Peter, T.: Global atmospheric sulfur budget under volcanically quiescent conditions: Aerosol-chemistry-climate model predictions and validation, Journal of Geophysical Research: Atmospheres, 120, 256–276, https://doi.org/10.1002/2014JD021985, 2015.

Shimamura, T., Iwashita, M., Iijima, S., Shintani, M., and Takaku, Y.: Major to ultra trace elements in rainfall collected in suburban Tokyo, Atmospheric Environment, 41, 6999–7010, https://doi.org/10.1016/j.atmosenv.2007.05.010, 2007.

Smith, S. J., Aardenne, J. v., Klimont, Z., Andres, R. J., Volke, A., and Delgado Arias, S.: Anthropogenic sulfur dioxide emissions: 1850–2005, Atmospheric Chemistry and Physics, 11, 1101–1116, https://doi.org/10.5194/acp-11-1101-2011, 2011.

Sobol, I. M.: Sensitivity estimates for nonlinear mathematical models, Mathematical Modelling and Computational Experiments, 1, 407–414, 1993.

Stenke, A., Schraner, M., Rozanov, E., Egorova, T., Luo, B., and Peter, T.: The SOCOL version 3.0 chemistry–climate model: description, evaluation, and implications from an advanced transport algorithm, Geoscientific Model Development, 6, 1407–1427, https://doi.org/10.5194/gmd-6-1407-2013, 2013.

Stohl, A., Seibert, P., Arduini, J., Eckhardt, S., Fraser, P., Greally, B. R., Lunder, C., Maione, M., Mühle, J., O'Doherty, S., Prinn, R. G., Reimann, S., Saito, T., Schmidbauer, N., Simmonds, P. G., Vollmer, M. K., Weiss, R. F., and Yokouchi, Y.: An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons, Atmospheric Chemistry and Physics, 9, 1597–1620, https://doi.org/10.5194/acp-9-1597-2009, 2009.

Sudret, B.: Global sensitivity analysis using polynomial chaos expansions, Reliability Engineering & System Safety, 93, 964–979, https://doi.org/10.1016/j.ress.2007.04.002, 2008.

Suess, E., Aemisegger, F., Sonke, J. E., Sprenger, M., Wernli, H., and Winkel, L. H. E.: Marine versus Continental Sources of Iodine and Selenium in Rainfall at Two European High-Altitude Locations, Environmental Science & Technology, 53, 1905–1917, https://doi.org/10.1021/acs.est.8b05533, 2019.

Suzuki, Y., Sugimura, Y., and Miyake, Y.: The content of selenium and its chemical form in rain water and aerosol in Tokyo, Journal of the Meteorological Society of Japan. Ser. II, 59, 405–409, https://doi.org/10.2151/jmsj1965.59.3_405, 1981.

Symonds, R. B. and Reed, M. H.: Calculation of multicomponent chemical equilibria in gas-solid-liquid systems: Calculation methods, thermochemical data, and applications to studies of high-temperature volcanic gases with examples from Mt. St. Helens, American Journal of Science, 293, https://doi.org/10.2475/ajs.293.8.758, 1993.

Symonds, R. B., Rose, W. I., Reed, M. H., Lichte, F. E., and Finnegan, D. L.: Volatilization, transport and sublimation of metallic and non-metallic elements in high temperature gases at Merapi Volcano, Indonesia, Geochimica et Cosmochimica Acta, 51, 2083–2101, https://doi.org/10.1016/0016-7037(87)90258-4, 1987.

Thompson, K., Canosa-Mas, C., and Wayne, R.: Kinetics and mechanism of the reaction between atomic chlorine and dimethyl selenide; comparison with the reaction between atomic chlorine and dimethyl sulfide, Physical Chemistry Chemical Physics, 4, 4133–4139, https://doi.org/10.1039/B204657A, 2002.

Tost, H., Jöckel, P., Kerkweg, A., Sander, R., and Lelieveld, J.: A new comprehensive SCAVenging submodel for global atmospheric chemistry modelling, Atmospheric Chemistry and Physics, 6, 565–574, https://doi.org/10.5194/acp-6-565-2006, 2006.

Uchiyama, R., Okochi, H., Ogata, H., Katsumi, N., and Nakano, T.: Characteristics of trace metal concentration and stable isotopic composition of hydrogen and oxygen in "urban-induced heavy rainfall" in downtown Tokyo, Japan; The implication of mineral/dust particles on the formation of summer heavy rainfall, Atmospheric Research, 217, 73–80, https://doi.org/10.1016/j.atmosres.2018.10.017, 2019.

Wackerly, D., Mendenhall, W., and Scheaffer, R. L.: Mathematical Statistics with Applications, Cengage Learning, 2014.

Wahner, A. and Ravishankara, A.: The kinetics of the reaction of OH with COS, Journal of Geophysical Research: Atmospheres, 92, 2189–2194, https://doi.org/10.1029/JD092iD02p02189, 1987.

Wai, K.-M., Wu, S., Li, X., Jaffe, D. A., and Perry, K. D.: Global atmospheric transport and source-receptor relationships for arsenic, Environmental Science & Technology, 50, 3714–3720, https://doi.org/10.1021/acs.est.5b05549, 2016.

Wardell, L., Kyle, P., and Counce, D.: Volcanic emissions of metals and halogens from White Island (New Zealand) and Erebus volcano (Antarctica) determined with chemical traps, Journal of Volcanology and Geothermal Research, 177, 734–742, https://doi.org/10.1016/j.jvolgeores.2007.07.007, 2008.

Watts, S. F.: The mass budgets of carbonyl sulfide, dimethyl sulfide, carbon disulfide and hydrogen sulfide, Atmospheric Environment, 34, 761–779, https://doi.org/10.1016/S1352-2310(99)00342-8, 2000.

Weisenstein, D. K., Yue, G. K., Ko, M. K., Sze, N.-D., Rodriguez, J. M., and Scott, C. J.: A two-dimensional model of sulfur species and aerosols, Journal of Geophysical Research: Atmospheres, 102, 13 019–13 035, https://doi.org/10.1029/97JD00901, 1997.

Weller, R., Wöltjen, J., Piel, C., Resenberg, R., Wagenbach, D., König-Langlo, G., and Kriews, M.: Seasonal variability of crustal and marine trace elements in the aerosol at Neumayer station, Antarctica, Tellus B: Chemical and Physical Meteorology, 60, 742–752, https://doi.org/10.1111/j.1600-0889.2008.00372.x, 2008.

Wen, H. and Carignan, J.: Reviews on atmospheric selenium: emissions, speciation and fate, Atmospheric Environment, 41, 7151–7165, https://doi.org/10.1016/j.atmosenv.2007.07.035, 2007.

Wesely, M.: Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, Atmospheric Environment (1967), 23, 1293–1304, https://doi.org/10.1016/0004-6981(89)90153-4, 1989.

Winkel, L., Vriens, B., Jones, G., Schneider, L., Pilon-Smits, E., and Bañuelos, G.: Selenium cycling across soil-plant-atmosphere interfaces: a critical review, Nutrients, 7, 4199–4239, https://doi.org/10.3390/nu7064199, 2015.

Winkel, L. H. E., Johnson, C. A., Lenz, M., Grundl, T., Leupin, O. X., Amini, M., and Charlet, L.: Environmental Selenium Research: From Microscopic Processes to Global Understanding, https://doi.org/10.1021/es203434d, https://doi.org/10.1021/es203434d, pMID: 22129299, 2012.

Xing, J., Song, J., Yuan, H., Wang, Q., Li, X., Li, N., Duan, L., and Qu, B.: Atmospheric wet deposition of dissolved trace elements to Jiaozhou Bay, North China: Fluxes, sources and potential effects on aquatic environments, Chemosphere, 174, 428–436, https://doi.org/10.1016/j.chemosphere.2017.02.004, 2017.

Xiu, D. and Karniadakis, G. E.: The Wiener–Askey polynomial chaos for stochastic differential equations, SIAM Journal on Scientific Computing, 24, 619–644, https://doi.org/10.1137/S1064827501387826, 2002.

Yan, R., Gauthier, D., and Flamant, G.: Possible interactions between As, Se, and Hg during coal combustion, Combustion and Flame, 120, 49–60, https://doi.org/10.1016/S0010-2180(99)00079-6, 2000.

Yan, R., Gauthier, D., Flamant, G., and Wang, Y.: Behavior of selenium in the combustion of coal or coke spiked with Se, Combustion and Flame, 138, 20–29, https://doi.org/10.1016/j.combustflame.2004.03.010, 2004.

Zhou, J., Wang, Y., Yue, T., Li, Y., Wai, K.-M., and Wang, W.: Origin and distribution of trace elements in high-elevation precipitation in southern China, Environmental Science and Pollution Research, 19, 3389–3399, https://doi.org/10.1007/s11356-012-0863-7, 2012.