Supplement to "Objective frequentist uncertainty quantification for atmospheric CO_2 retrievals"

Pratik Patil[1,2], Mikael Kuusela[1] and Jonathan Hobbs[3]

[1]Department of Statistics and Data Science, Carnegie Mellon University
[2]Machine Learning Department, Carnegie Mellon University
[3]Jet Propulsion Laboratory, California Institute of Technology

S1 Supplement to Section 2.1

Figure S1 shows an example sounding for OCO-2.

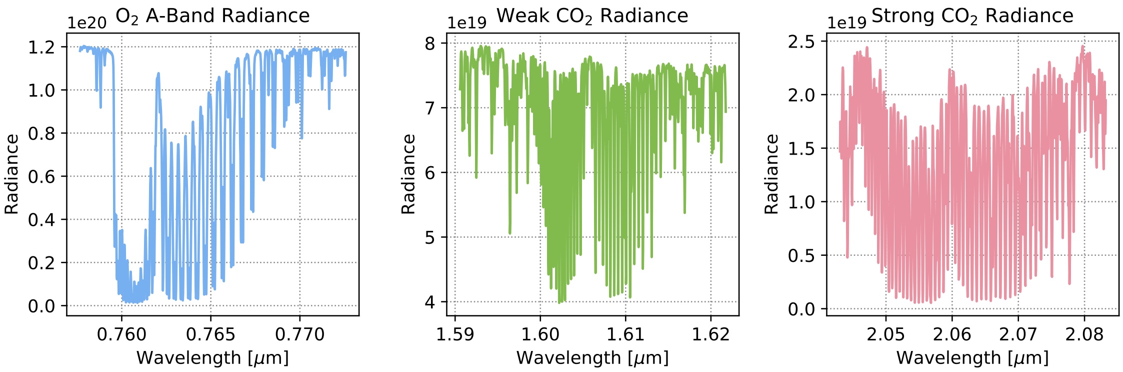


Figure S1: Example sounding from OCO-2. A sounding includes 1016 radiances in each of the three infrared spectral bands.

S2 Supplement to Section 5.1

S2.1 Description of state vector elements

The specific elements of the state vector $\boldsymbol{x} \in \mathbb{R}^{39}$ are:

- Variables $x_1, \ldots, x_{20}$ are the dry-air mole fractions of atmospheric CO_2, i.e., the number of moles of CO_2 per one mole of dry air, in parts per million (ppm) at 20 fixed pressure levels. In the sequel, we denote these as levels 1 to 20, with level 1 being highest in the atmosphere (pressure ~ 0.1 hPa) and level 20 being the surface.

- Variable x_{21} is the surface air pressure in hPa. It corresponds to the total weight of the air molecules in the atmospheric column.

- Variables $x_{22}, \ldots, x_{27}$ relate to surface albedo, i.e., the fraction of total incoming solar radiation reflected off the Earth's surface. Albedo varies across the three OCO-2 spectral bands and also within each band. In the surrogate model, albedo is modeled as a linear function within each spectral band; see Section B.2 in [1]. The albedo for each band is therefore parameterized by an intercept and a slope which enter the state vector as nuisance variables (x_{22}, x_{24}, and x_{26} are the three intercepts and x_{23}, x_{25}, and x_{27} are the slopes).

- Variables $x_{28}, \ldots, x_{39}$ parameterize the atmospheric aerosol concentrations and distributions. The surrogate model assumes that there are 4 aerosol types, which have distinct absorption and scattering properties. The first two are location-dependent composite species. For our investigation, these are sulfate and dust. The latter two are two cloud species, one for ice clouds and another for liquid water clouds [2, 1]. Each type is parameterized by 3 parameters corresponding to the aerosol optical depth (AOD, i.e., the opaqueness of the aerosol species measured as the natural logarithm of the ratio of incoming to transmitted radiation) as well as the altitude and thickness of each aerosol layer. Variables x_{28}, x_{31}, x_{34}, and x_{37} are the log-AOD values, variables x_{29}, x_{32}, x_{35}, and x_{38} are the altitudes, and variables x_{30}, x_{33}, x_{36}, and x_{39} are the log-thicknesses.

S2.2 Forward model singular values

Figure S2 shows the singular values of the linearized forward model $\boldsymbol{K}$.

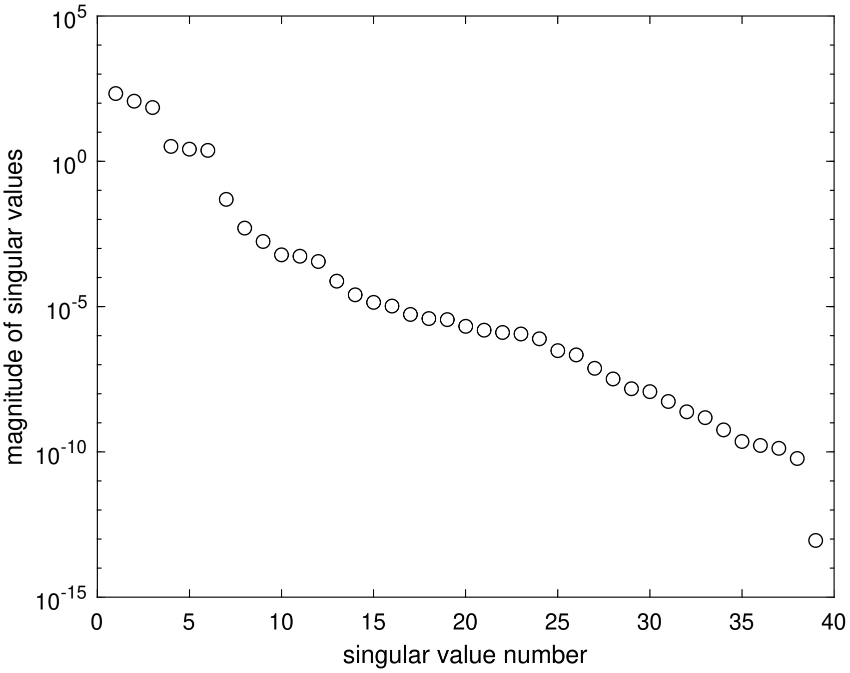
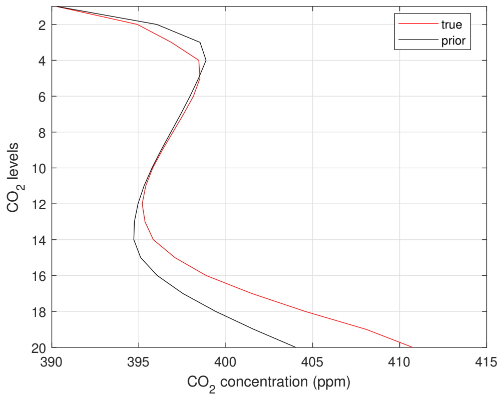


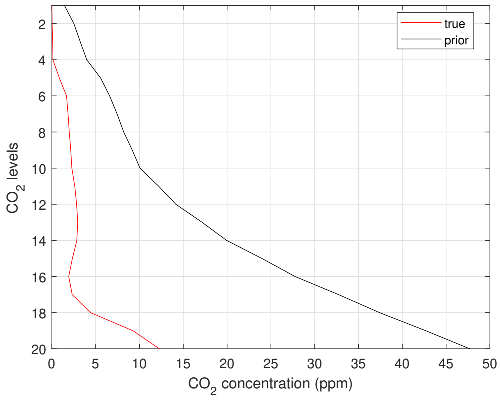
Figure S2: Singular value decay of the linearized forward model $\boldsymbol{K}$ near Lamont, OK, in October 2015.

S2.3 Exploratory analysis of the prior and generative models

In this section, we visualize and describe various components of the generative model and compare those to the prior model. We begin by comparing the prior mean and standard deviation to the true generative model mean and standard deviation for the CO_2 part of the state vector



(a) Mean misspecification



(b) Standard deviation misspecification

Figure S3: Visualization of the misspecification of the mean and the standard deviation for the CO_2 profile between the true generative process and the prior. Figure (a) visualizes the misspecification in the mean and Figure (b) shows the misspecification in the standard deviation.

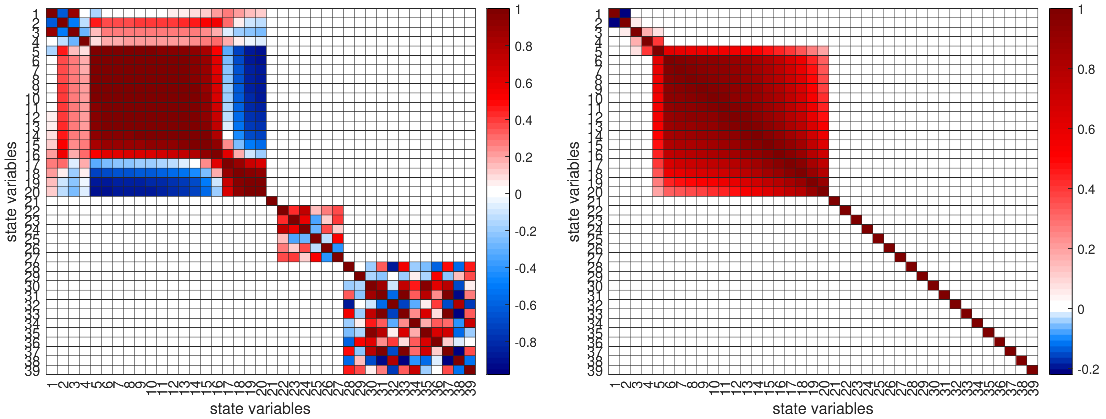
Table S1: Comparison of the true and prior mean and standard deviation for the nuisance variables $x_i, i = 21, \dots, 39$. A detailed description of these variables is given in Section S2.1.

i	true mean (sd)	prior mean (sd)
	surface pressure	
21	972.3235 (1.7508)	970.6240 (4.000)
	albedo	
22	0.1267 (0.0204)	0.1267 (1.0000)
23	0.0001 (0.0001)	0.0000 (0.0005)
24	0.2484 (0.0044)	0.2484 (1.0000)
25	-0.0001 (0.0000)	0.0000 (0.0005)
26	0.2027 (0.0026)	0.2027 (1.0000)
27	-0.0000 (0.0000)	0.0000 (0.0005)
	aerosols	
28	-3.8786 (0.3380)	-3.7643 (2.0000)
29	0.8187 (0.0683)	0.9000 (0.2000)
30	-2.4492 (0.0409)	-2.9957 (0.1823)
31	-6.1959 (0.8492)	-4.7370 (2.0000)
32	0.3255 (0.0101)	0.9000 (0.2000)
33	-3.9219 (0.0201)	-2.9957 (0.1823)
34	-4.3980 (0.2477)	-4.3820 (1.8000)
35	-0.0087 (0.0355)	0.3000 (0.2000)
36	-3.2080 (0.0112)	-3.2189 (0.2231)
37	-5.6803 (0.2517)	-4.3820 (1.8000)
38	1.0917 (0.0870)	0.7500 (0.4000)
39	-2.3052 (0.0004)	-2.3026 (0.0953)

in Figure S3. We observe that both the mean and the standard deviation have the largest misspecification near the surface. Table S1 contains the same information for the nuisance variables $x_{21}, \dots, x_{39}$. Generally speaking, the prior standard deviation is by design larger than that of the true process, which provides some protection against the prior mean misspecification.

We next visualize the correlation structure in the generative process and the prior. Figure S4(a) shows a heat map of the correlation matrix for the true generative process. We observe that the state vector consists of four independent subgroups of variables corresponding to the CO_2 profile, surface pressure, albedo variables and aerosol variables. While variables across these groups are uncorrelated, there are large within-group correlations. Figure S4(b) shows a heat map of the correlation matrix for the prior process. Again, the CO_2 variables are independent of the nuisance variables, but in this case, there are no correlations between the nuisance variables. Notice also the differences in the correlation structure within the CO_2 profile between the generative process and the prior.

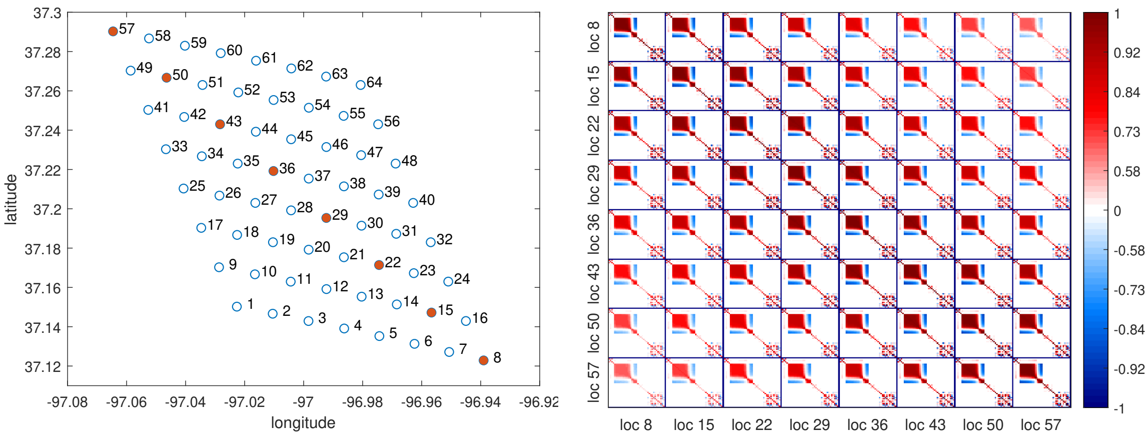
Lastly, we visualize the spatial correlation structure for the generative process over an 8×8



(a) True correlation (b) Prior correlation

Figure S4: Visualization of the misspecification of the correlation structure between the true generative process and the prior. Figure (a) displays the correlation between the state vector elements in the true generative process. Figure (b) shows the same for the prior process.

spatial grid near Lamont, OK. These 64 sounding locations are shown in Figure S5(a) and Figure S5(b) displays the correlations across the locations along the diagonal in Figure S5(a). Nearby state vectors are strongly spatially correlated, but there is a fair amount of decorrelation when moving across the grid.



(a) Grid locations (b) Grid correlation structure

Figure S5: Visualization of the correlation structure between the state vector elements for the generative process over an 8×8 spatial grid. Figure (a) shows the coordinates of the sounding locations near Lamont, OK, with the numbers giving an index for each location. Figure (b) visualizes the correlation structure between the state vector elements across the marked locations along the diagonal in Figure (a). Notice the nonlinear color scale in Figure (b).

S3 Supplement to Section 5.2: Illustrative instances

We pick some illustrative realizations to visualize the state vectors and confidence intervals produced by the two methods. We choose three representative cases corresponding to the minimum, nominal and maximum coverage for the operational method in Figure 3(b) which are also rows 1, 8 and 10 in Table 1, respectively.

Figure S6(a) shows the CO_2 part of the $\boldsymbol{x}$ instance having the smallest operational coverage. We observe, consistent with Section 5.2.1, that the operational interval is biased upward. It is overoptimistic about the amount of uncertainty, leading it to miss the true X_{CO2} value for this particular $\boldsymbol{\varepsilon}$ realization. From Table 1, we know that this happens for roughly 21% of $\boldsymbol{\varepsilon}$ realizations when the noncoverage probability should ideally be only 5%. The proposed interval, on the other hand, is wider and as a result ends up covering the true X_{CO2} value for the same $\boldsymbol{x}$ and $\boldsymbol{\varepsilon}$ realizations. From Table 1, we know that the noncoverage probability for this interval is 5%, as it should be.

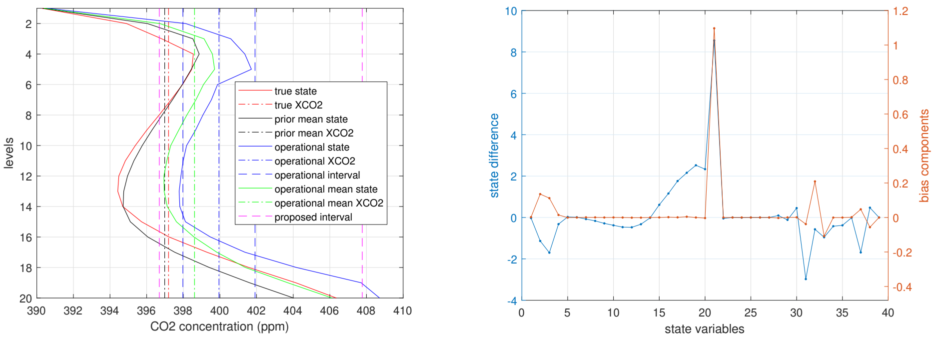
It is quite insightful to investigate the source of the upward bias for this $\boldsymbol{x}$ realization. Contrary to what one might at first imagine, it is not caused by a misspecification of the CO_2 profile in the prior. In fact, as shown in Figure S6(a), the prior CO_2 profile is very similar to the true CO_2 profile in $\boldsymbol{x}$ and the X_{CO2} value implied by the prior is almost the same as the true X_{CO2} value. Instead, it turns out that the bias is primarily caused by an upward fluctuation in the surface pressure variable x_{21}, as shown by Figure S6(b). The large positive difference in x_{21} between the true state and the prior gets multiplied by the relatively large positive bias multiplier of this variable (Figure 2(a); see also Section 5.2.1) leading to a large positive contribution to the overall bias. This results in a positively biased operational X_{CO2} point estimate and a miscalibrated confidence interval.

Next, instead of picking an adversarial $\boldsymbol{x}$, Figure S7(a) shows the $\boldsymbol{x}$ realization corresponding to the largest coverage for the operational method. For this $\boldsymbol{x}$, the operational X_{CO2} estimate is effectively unbiased and the operational interval covers the true X_{CO2} value for almost all $\boldsymbol{\varepsilon}$ realizations, resulting in substantial overcoverage. The proposed interval, on the other hand, again has the desired 95% coverage, as indicated by Table 1. For the $\boldsymbol{\varepsilon}$ realization shown in the figure, the operational and proposed intervals both cover the true X_{CO2} value.

Interestingly, the operational X_{CO2} estimator in Figure S7(a) is effectively unbiased even though the prior is badly misspecified for this CO_2 profile. Figure S7(b) explains this phenomenon. Due to the small bias multipliers of the low-altitude CO_2 values (Figure 2(a)), prior misspecification in the lower half of the CO_2 profile creates little bias, while the misspecification in the upper half of the profile is such that it creates both positive and negative contributions to the bias that cancel out. At the same time, there is a consistent positive contribution to the bias from x_{32}, which is negated by a downward fluctuation in the pressure variable x_{21} and a consistent negative bias contribution from x_{33}.

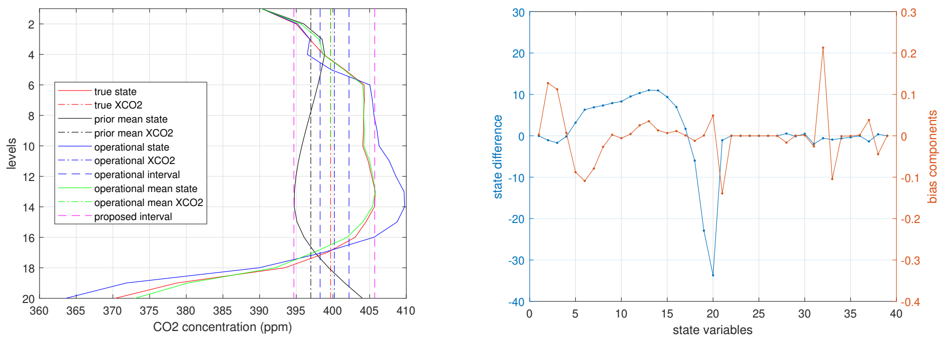
Finally, we compare in Figure S8(a) the intervals for the $\boldsymbol{x}$ realization that gives nominal coverage for the operational method in Table 1. For this $\boldsymbol{x}$, both intervals cover the true X_{CO2} value for 95% of $\boldsymbol{\varepsilon}$ realizations. The $\boldsymbol{\varepsilon}$ realization shown in the figure has both intervals covering the true X_{CO2} value. The operationally retrieved CO_2 profile has fluctuated quite substantially, but the anticorrelated fluctuations cancel out to produce a well-behaved confidence interval for X_{CO2}. Overall, the operational method is slightly positively biased for X_{CO2}, as expected based on the discussion in Section 5.2.2. Figure S8(b) reveals that the source of this bias is similar to the situation in Figure S6 in that the bias is primarily caused by the pressure variable x_{21}, albeit with a smaller magnitude. As before, the prior misspecification for the CO_2 profile near the surface (levels 18, 19 and 20) causes little harm due to the small bias multipliers of those variables (Figure 2(a)).

Overall, Figures S6–S8 show how challenging it is to understand, predict and explain the uncertainty quantification performance of the operational retrieval method. The method can



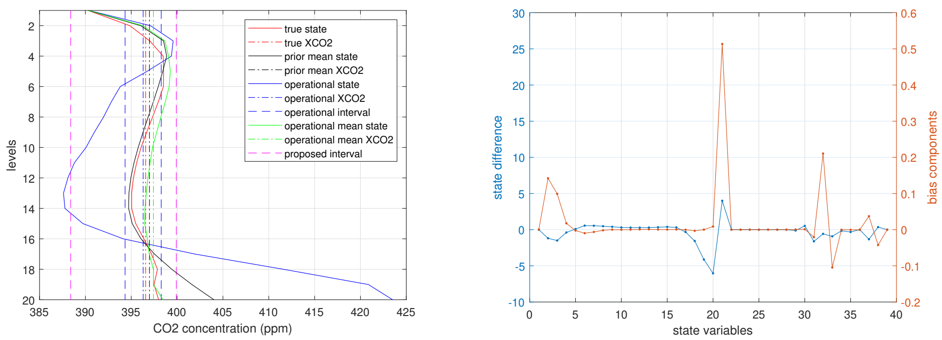
(a) Profile comparisons (b) Bias components

Figure S6: Figure (a) illustrates the operational and proposed intervals for the state realization that has the smallest coverage in Figure 3(b). Figure (b) visualizes the corresponding state differences $x_i - \mu_{a,i}$ and bias components $m_i(x_i - \mu_{a,i})$, where m_i is the ith bias multiplier.



(a) Profile comparisons (b) Bias components

Figure S7: Figure (a) illustrates the operational and proposed intervals for the state realization that has the largest coverage in Figure 3(b). Figure (b) visualizes the corresponding state differences $x_i - \mu_{a,i}$ and bias components $m_i(x_i - \mu_{a,i})$, where m_i is the ith bias multiplier.



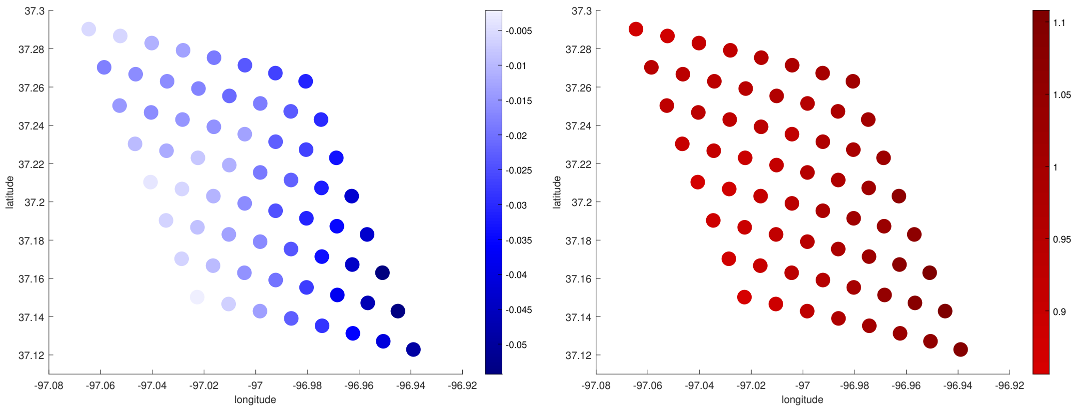
(a) Profile comparisons (b) Bias components

Figure S8: Figure (a) illustrates the operational and proposed intervals for a state realization that has nominal coverage in Figure 3(b). Figure (b) visualizes the corresponding state differences $x_i - \mu_{a,i}$ and bias components $m_i(x_i - \mu_{a,i})$, where m_i is the ith bias multiplier.

exhibit the very counterintuitive behavior where a relatively well-specified prior CO_2 profile (Figure S6) in fact has the worst coverage performance, while a badly misspecified prior CO_2 profile (Figure S7) has the highest coverage. The key to understanding the performance of the method, it turns out, is to understand the effect of the nuisance variables and how they interact with the misspecification of the CO_2 profile. Such analysis is obviously only possible when the true $\boldsymbol{x}$ is known, which would make it very difficult to perform a similar study for real-life operational retrievals. The proposed method, on the other hand, is free from these complications and exhibits consistent 95% coverage for all $\boldsymbol{x}$ realizations we have investigated.

S4 Supplement to Section 5.3

Figure S9 shows the spatial coverage and bias patterns for a case where the state vector realizations are such that all 64 intervals across the region have coverage below the nominal value.



(a) Operational coverage

(b) Operational bias

Figure S9: Operational retrieval over a grid of 8×8 soundings for an instance with undercoverage for the entire grid. Figure (a) shows the spatial coverage pattern relative to the nominal 95% in units of probability (i.e., -0.03, for example, corresponds to coverage 0.92, instead of the nominal 0.95). Figure (b) shows the corresponding bias pattern in ppm.

References

[1] Jonathan Hobbs, Amy Braverman, Noel Cressie, Robert Granat, and Michael Gunson. Simulation-based uncertainty quantification for estimating atmospheric CO_2 from satellite data. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):956–985, 2017.

[2] H. Boesch et al. *Orbiting Carbon Observatory-2 & 3: Level 2 Full Physics Retrieval Algorithm Theoretical Basis*. NASA Jet Propulsion Laboratory, January 2, 2019. OCO D-55207, Version 2.0 Rev 3.