



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Post-hoc Uncertainty Quantification for Remote Sensing Observing Systems

Amy Braverman, Jon Hobbs, Joaquim Teixeira, and Michael Gunson
Jet Propulsion Laboratory, California Institute of Technology

October 9, 2020



- ▶ Introduction
- ▶ Remote sensing observing systems
- ▶ Orbiting Carbon Observatory 2 (OCO-2) mission, science, and data
- ▶ Statistical model of the observing system and uncertainty
- ▶ Methodology
- ▶ Application to OCO-2
- ▶ Summary and discussion points



Remote sensing levels of data processing:

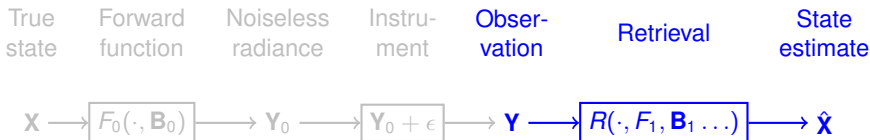
- ▶ Level 0: raw photon counts direct from satellite
- ▶ Level 1: georectified and calibrated radiances
- ▶ **Level 2: estimates of geophysical state**
- ▶ Level 3: "statistical summaries" of Level 2 on uniform space-time grid
- ▶ Level 4: output of models or data assimilation

Level 2 "data" aren't "data"; they are inferences!

How do we calculate uncertainties of the point estimates provided by remote sensing observing systems?



Remote sensing observing system:



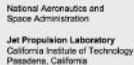
Retrieval is an inference problem: estimate $\mathbf{X}$ when you only get to see $\mathbf{Y}$.

F_0 = nature's true forward function; $\mathbf{B}_0$ = other true quantities.

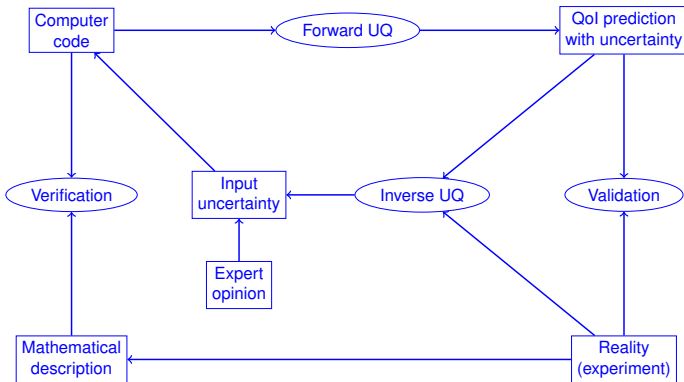
F_1 = forward model used in retrieval, R ; $\mathbf{B}_1$ = other retrieval inputs.

ϵ = instrument measurement error.

$\dots$ = other retrieval algorithm inputs.



Uncertainty quantification:



Adapted from Wu et al, (2018). DOI: 10.1016/j.nucengdes.2018.06.004.

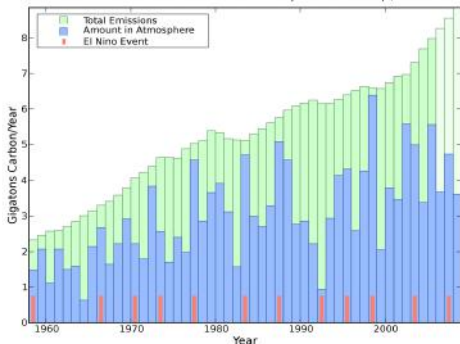


- ▶ Uncertainty quantification (UQ) provides a formalism for understanding uncertainty in the output of computational models.
- ▶ Remote sensing data processing utilizes complex computational models.
- ▶ But:
 - ▶ we lack information about uncertainty of known inputs
 - ▶ there are many unknown unknowns (e.g., interaction effects)
 - ▶ computational artifacts must be taken into account
 - ▶ analysis must be performed after the fact and be computationally feasible



OCO-2 mission and science

Fossil Fuel Emissions of CO₂ and Atmospheric Buildup, 1958-2008



Graphic courtesy of Annmarie
Eldering.

- ▶ Less than half of the carbon is staying in the atmosphere.
- ▶ Where are the sinks that are absorbing more than half of this CO₂?
- ▶ Why does CO₂ build-up vary from year to year with nearly uniform emission rates?
- ▶ How will CO₂ sinks respond to climate change?



OCO-2 mission and science

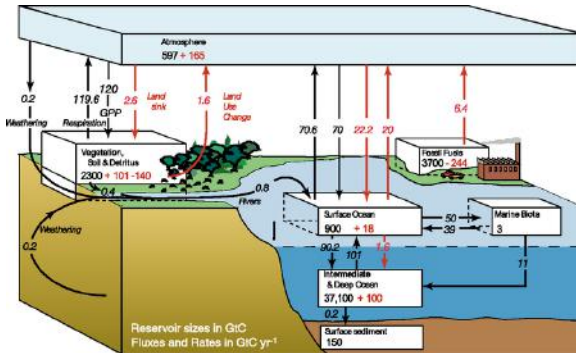


Figure 7.3 from *Climate Change 2007: Working Group I: The Physical Science Basis*.

- ▶ Benefit of space-based greenhouse-gas measurement is coverage and resolution.
- ▶ Challenge is the need for high-precision measurements to detect small changes in flux against large background variations.

- ▶ A primary objective of OCO-2 is to find and study the natural sinks.



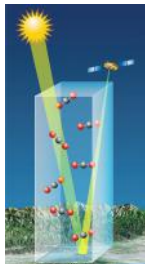
OCO-2 mission and science



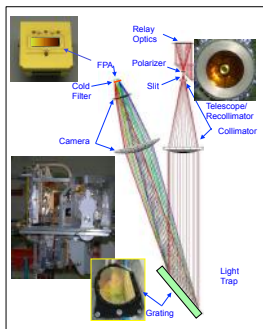
Ground tracks

OCO-2 in polar orbit at 705 km.
16-day ground-track repeat cycle
(14.57 orbits per day).

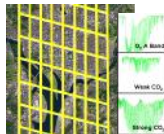
Longitude offset between orbits =
25 degrees; 1.5 degrees between
tracks after 16 days.



Measurement principle



Measurement technology



Ground footprints

Eight 1.29×2.25 km
footprints cross-track.

Three co-bore-sighted
high-resolution grating
spectrometers,
observing in oxygen-A,
weak CO₂ and strong
CO₂ bands (3024
wavelengths).

- For each footprint, OCO-2 measures a 3048×1 vector of radiances, $\mathbf{Y}$ (Crisp, 2015).
- The retrieval uses $\mathbf{Y}$ to estimate the 45×1 atmospheric state vector, $\mathbf{X}$, which includes CO₂ dry air mole-fraction at 20 altitudes.

Greatest interest lies in the scalar quantities,

$$\hat{X} \approx E(\mathbf{h}'\hat{\mathbf{X}}_{1:20}|\mathbf{Y}), \quad \hat{S} \approx \text{var}(\mathbf{h}'\hat{\mathbf{X}}_{1:20}|\mathbf{Y}),$$

where $\mathbf{X}_{1:20}$ is the first 20 elements of the OCO-2 state vector (the CO₂ vertical profile), and $\mathbf{h}$ is the 20-element pressure weighting function.

OCO-2 uses the Optimal Estimation (OE; Rodgers 2000) method to “retrieve” the maximum a posteriori estimate, $\hat{X}$, and then computes $\hat{S}$ as a linear function of $\hat{X}$, all under Gaussian assumptions.

- ▶ $\hat{X}$ is treated as the posterior mean total column mole-fraction of CO₂.
- ▶ $\hat{S}$ is treated as the posterior variance of total column mole-fraction of CO₂.



True state Forward function Noiseless radiance Instru- ment Obser- vation Retrieval State estimate

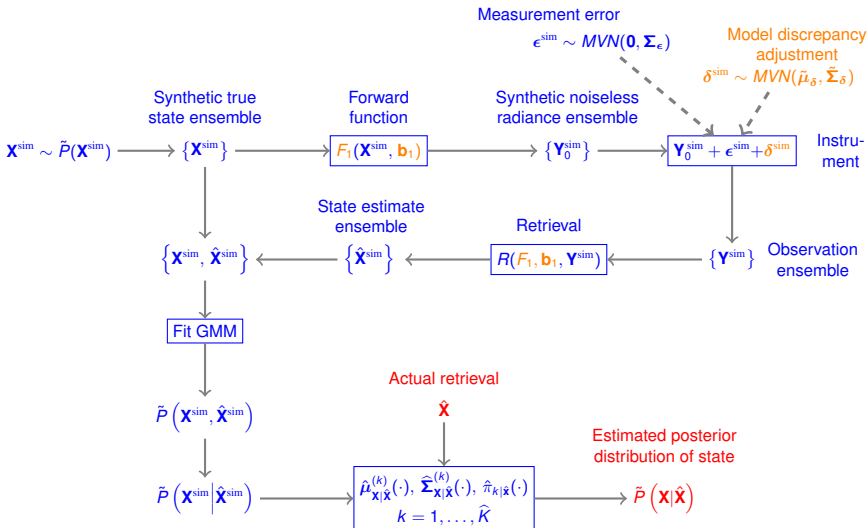


Uncertainty is quantified by conditional distributions.

- ▶ OE: $P(\mathbf{X}|\mathbf{Y})$; only source of uncertainty is ϵ .
- ▶ Our view: $P(\mathbf{X}|\hat{\mathbf{X}})$; uncertainty induced by mismatches of F_1 to F_0 , $\mathbf{B}_1$ to $\mathbf{B}_0$, non-Gaussianity, computational artifacts (e.g., discretization, definition of vertical grid, etc.), and unknown interactions among all these sources.
- ▶ How to proceed given that we can't even enumerate all sources of uncertainty?



Methodology





- ▶ Model $P(\mathbf{X}|\hat{\mathbf{X}})$ as a Gaussian mixture model (GMM; McLachlan and Peel, 2000) derived from $P(\mathbf{X}, \hat{\mathbf{X}})$.
- ▶ Estimate the parameters of $P(\mathbf{X}|\hat{\mathbf{X}})$ from realistic, simulated data.
- ▶ Estimate the (conditional) GMM for a new (actual) $\hat{\mathbf{X}}$:
 - ▶ plug the new $\hat{\mathbf{X}}$ into the formulas for regression means and variances for all components in the GMM
 - ▶ simulate from the conditional GMM



The Gaussian mixture density for a multivariate random vector $\mathbf{V}$ is,

$$f_{\mathbf{V}}(\mathbf{v}) = \sum_{k=1}^K \pi_k \phi(\mathbf{v}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \quad \sum_{k=1}^K \pi_k = 1,$$

where:

$\phi(\mathbf{v}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ = the multivariate normal density function with mean vector $\boldsymbol{\mu}_k$ and covariance matrix $\boldsymbol{\Sigma}_k$, evaluated at $\mathbf{v}$;

π_k = the (mixing) weight of component k ,

K = the total number of components.

We abbreviate this density by,

$$\mathbf{V} \sim \text{GMM} \left(K, \{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \pi_k\}_{k=1}^K \right).$$



The parameters of this model are,

$$\left\{ \hat{K}, \hat{\mu}_1, \dots, \hat{\mu}_{\hat{K}}, \hat{\Sigma}_1, \dots, \hat{\Sigma}_{\hat{K}}, \hat{\pi}_1, \dots, \hat{\pi}_{\hat{K}} \right\}.$$

We use the R package mclust (Scrucca et al., 2016) to estimate them.



Partition $\mathbf{V} = (\mathbf{W}', \mathbf{U}')$.

The **regression mean function** for component k evaluated at $\mathbf{U} = \mathbf{u}$ is,

$$\hat{\mu}_{\mathbf{W}|\mathbf{U}}^{(k)}(\mathbf{u}) = \hat{\mu}_{\mathbf{W}}^{(k)} + \hat{\Sigma}_{\mathbf{WU}}^{(k)} \left(\hat{\Sigma}_{\mathbf{UU}}^{(k)} \right)^{-1} \left[\mathbf{u} - \hat{\mu}_{\mathbf{U}}^{(k)} \right],$$

where $\hat{\Sigma}_{\mathbf{V}}^{(k)}$ is the estimated covariance matrix for component k , which is partitioned as,

$$\hat{\Sigma}_{\mathbf{V}}^{(k)} = \begin{bmatrix} \hat{\Sigma}_{\mathbf{WW}}^{(k)} & \hat{\Sigma}_{\mathbf{WU}}^{(k)} \\ \hat{\Sigma}_{\mathbf{UW}}^{(k)} & \hat{\Sigma}_{\mathbf{UU}}^{(k)} \end{bmatrix}.$$

The regression covariance function for component k evaluated at $\mathbf{U} = \mathbf{u}$ is,

$$\hat{\Sigma}_{\mathbf{w}|\mathbf{u}}^{(k)}(\mathbf{u}) = \hat{\Sigma}_{\mathbf{w}\mathbf{w}}^{(k)} - \hat{\Sigma}_{\mathbf{w}\mathbf{u}}^{(k)} \left(\hat{\Sigma}_{\mathbf{u}\mathbf{u}}^{(k)} \right)^{-1} \hat{\Sigma}_{\mathbf{u}\mathbf{w}}^{(k)}.$$

The (posterior) mixing weights are,

$$\hat{\pi}_{k|\mathbf{u}} = \tilde{P}(\kappa = k | \mathbf{U} = \mathbf{u}) = \frac{\hat{\pi}_k \phi \left(\mathbf{u}; \hat{\mu}_{\mathbf{w}|\mathbf{u}}^{(k)}(\mathbf{u}), \hat{\Sigma}_{\mathbf{w}|\mathbf{u}}^{(k)}(\mathbf{u}) \right)}{\sum_{l=1}^{\hat{K}} \hat{\pi}_l \phi \left(\mathbf{u}; \hat{\mu}_{\mathbf{w}|\mathbf{u}}^{(l)}(\mathbf{u}), \hat{\Sigma}_{\mathbf{w}|\mathbf{u}}^{(l)}(\mathbf{u}) \right)},$$

for $k = 1, \dots, \hat{K}$.

κ indicates component membership: the probability that $\mathbf{V}$ comes from component k is $\hat{\pi}_k$ before observing $\mathbf{U}$, and $\hat{\pi}_{k|\mathbf{u}}$ after seeing $\mathbf{U} = \mathbf{u}$.



Conditional distribution of $\mathbf{W}$ given $\mathbf{U} = \mathbf{u}^*$:

$$[\mathbf{W}|\mathbf{U} = \mathbf{u}^*] \sim \text{GMM} \left(\hat{K}^{\text{sim}}, \left\{ \hat{\mu}_{\mathbf{W}|\mathbf{U}}^{(k)}(\mathbf{u}^*), \hat{\Sigma}_{\mathbf{W}|\mathbf{U}}^{(k)}(\mathbf{u}^*), \hat{\pi}_{k|\mathbf{u}^*} \right\}_{k=1}^K \right).$$

Approximate this distribution by simulating from it B times:

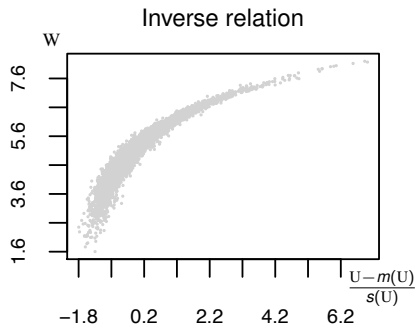
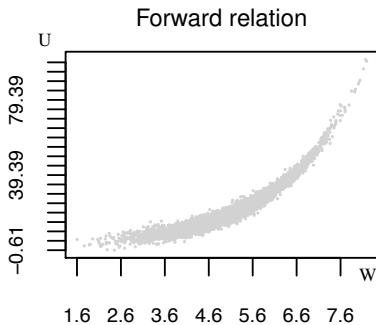
1. Let κ_b be a univariate random variable taking values in the set $\{1, \dots, \hat{K}^{\text{sim}}\}$ with,

$$P(\kappa_b = k) = \hat{\pi}_{k|\mathbf{u}^*}, \quad k \in \{1, \dots, \hat{K}^{\text{sim}}\}.$$

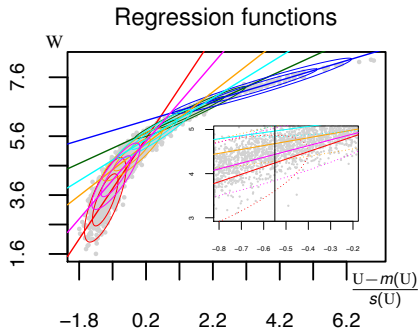
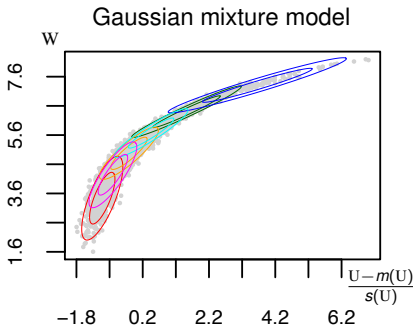
2. Draw B random variables,

$$\mathbf{W}_b^* \sim \text{N} \left(\hat{\mu}_{\mathbf{W}|\mathbf{U}}^{(\kappa_b)}(\mathbf{u}^*), \hat{\Sigma}_{\mathbf{W}|\mathbf{U}}^{(\kappa_b)}(\mathbf{u}^*) \right), \quad b = 1, \dots, B.$$

3. Fit kernel density estimate to $(\mathbf{W}_1^*, \dots, \mathbf{W}_B^*)$.



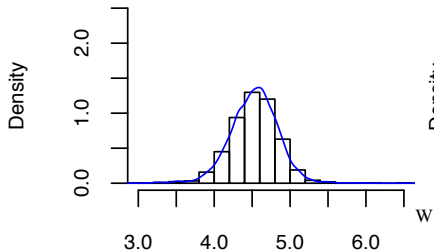
$$W \sim N(5, 1), \quad U = (1.75)^W + \epsilon, \quad \epsilon \sim N(1, 2).$$



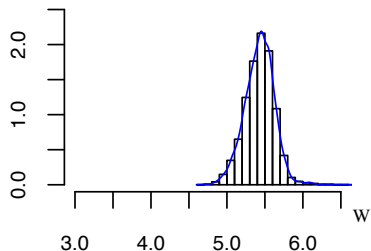
For any $U = u^*$, obtain mean, variance, and mixing weight for each component from its regression functions, and simulate.



Density of $W|U = 13.76$

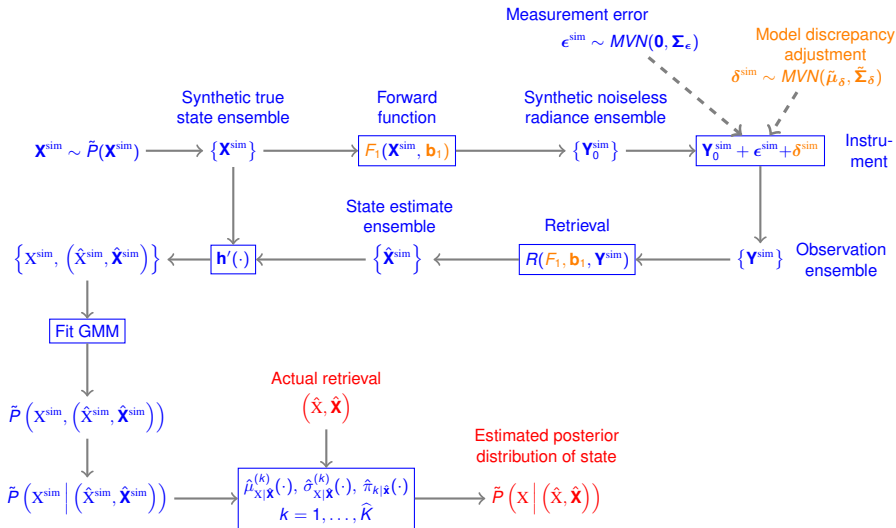


Density of $W|U = 22.41$





Application to OCO-2





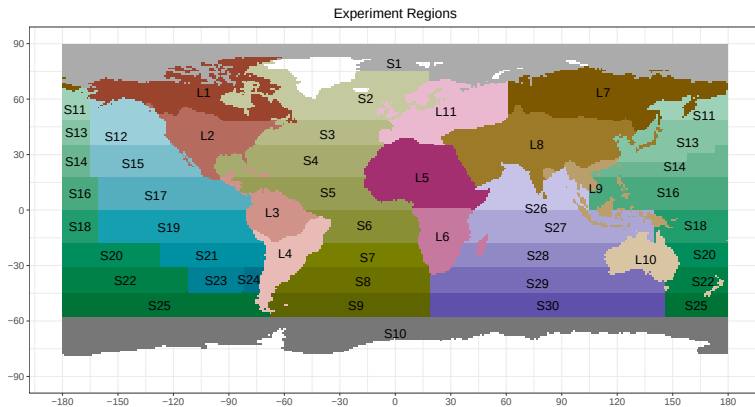
Synthetic true state ensemble, $\tilde{P}(\mathbf{X}^{\text{sim}})$:

- ▶ Ocean and land treated separately— land discussed here.
- ▶ Actual OCO-2 retrievals partitioned into 11 land regions by week-long time periods.
- ▶ Gaussian mixture model fit (by mclust) separately to each region-week (“template”).
- ▶ Make 5000 iid draws from each template GMM.

The strategy is similar to a semi-parametric bootstrap.



Application to OCO-2





How well does this work?

- ▶ Evaluate simulation-based distributions using “ground-truth” (Total Carbon Column Observing Network; TCCON, Wunch et al., 2017) where it exists.
- ▶ Consider OCO-2 footprints that contain TCCON sites during four months: August 2015, November 2015, February 2016, and May 2016.
- ▶ Also evaluate operationally derived distributions: $N(\hat{X}, \hat{S})$.
- ▶ Two metrics: centrality of the TCCON value in the reported distribution, and bias of the distribution mean relative to the TCCON value.



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

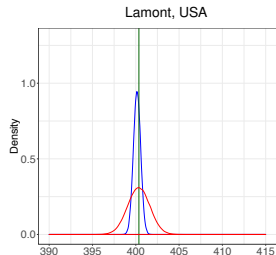
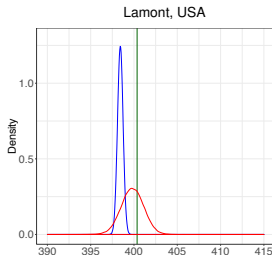
Implementation and results

Total Carbon Column Observing Network (TCCON)





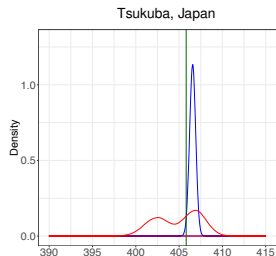
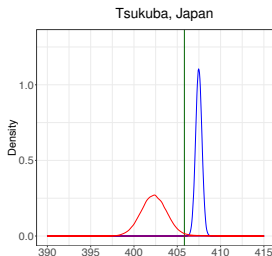
Application to OCO-2



TCCON value

Operational
distribution

Simulated
distribution



Metrics for evaluating simulated and operational distributions.

Metric 1:

$$G_{op}(X_T) = P_{op}(X \leq X_T) \approx \Phi(X_T; \hat{X}, \hat{S}),$$

$$G_s(X_T) = P_s(X \leq X_T) \approx \frac{1}{B} \sum_{b=1}^B \mathcal{I}(W_b^* \leq X_T).$$

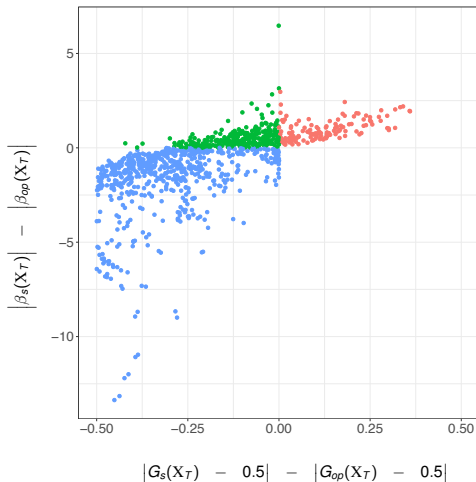
Metric 2:

$$\beta_{op}(X_T) = (\hat{X} - X_T) \quad \text{and} \quad \beta_s(X_T) = \left(\frac{1}{B} \sum_{b=1}^B W_b^* \right) - X_T.$$

X_T denotes the TCCON value, subscripts op and s denote operational and simulated quantities, respectively.



Application to OCO-2



$N = 944$

63.1%

25.7%

11.2%



Checking interval coverage...

- ▶ Let $Q_{\alpha/2}^s$ and $Q_{1-\alpha/2}^s$ denote the lower and upper $\alpha/2$ quantiles of the simulated distribution.
- ▶ Let $Q_{\alpha/2}^{op}$ and $Q_{1-\alpha/2}^{op}$ denote the lower and upper $\alpha/2$ quantiles of the operational (Gaussian) distribution.
- ▶ Then $[Q_{\alpha/2}^s, Q_{1-\alpha/2}^s]$ and $[Q_{\alpha/2}^{op}, Q_{1-\alpha/2}^{op}]$ are $(1 - \alpha)100\%$ confidence intervals.

- ▶ Compute

$$p_{1-\alpha}^s = \frac{1}{N_r} \sum_{n=1}^{N_r} \mathcal{I}(X_T \in [Q_{\alpha/2}^s, Q_{1-\alpha/2}^s]), \text{ and}$$

$$p_{1-\alpha}^{op} = \frac{1}{N_r} \sum_{n=1}^{N_r} \mathcal{I}(X_T \in [Q_{\alpha/2}^{op}, Q_{1-\alpha/2}^{op}]),$$

where N_r is the number of OCO-2 footprints.

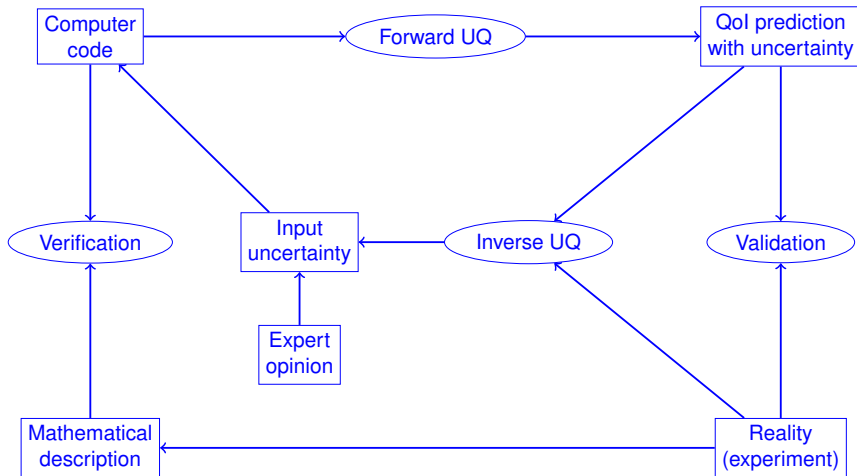


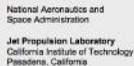
Application to OCO-2

Site	Week	$p_{.95}^{op}$	$p_{.50}^{op}$	$p_{.95}^s$	$p_{.50}^s$	N_r
Bialystok, Poland	2016-02-17	0.013	0.000	1.000	1.000	80
Darwin, Australia	2015-08-10	0.663	0.366	0.970	0.782	202
Lamont, OK USA	2015-11-02	0.288	0.143	1.000	0.909	132
Lauder, New Zealand	2016-02-29	0.449	0.170	0.932	0.441	118
Orleans, France	2015-11-02	0.627	0.322	0.746	0.661	59
Sodankyla, Finland	2015-08-20	0.533	0.133	0.600	0.267	15
Tsukuba, Japan	2016-05-13	0.390	0.169	0.974	0.818	77
Wollongong, Australia	2015-11-24	0.272	0.115	0.973	0.778	261

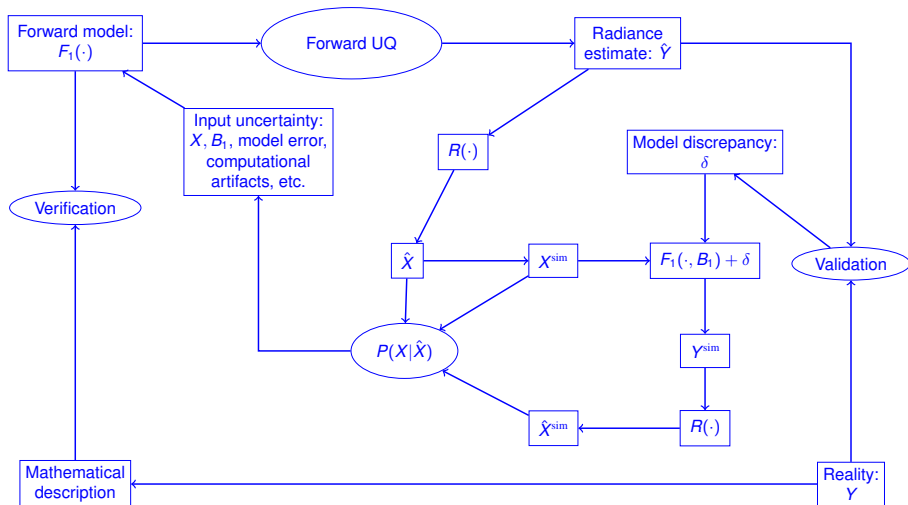


Summary and discussion





Summary and discussion





Summary and discussion points

- ▶ The challenge was to develop a practical method to assign uncertainties to every estimate produced by a remote sensing observing system.
- ▶ Uncertainty is defined by the conditional distribution of the true state given the operationally produced state estimate.
- ▶ Our method is similar to the bootstrap bias correction (Davison and Hinkley, 1997), but goes further: we correct the entire distribution (not just the mean).
- ▶ Simulation-based confidence intervals are usually valid, but not always efficient.
- ▶ Operational confidence intervals are rarely valid.



Summary and discussion points

- ▶ Method appears to perform relatively well in this study, but there are indications that we need to adjust the stratification so that the model is more globally representative.
- ▶ The most computationally intensive stage in building the model is the retrieval on the simulated radiances.
- ▶ Once the model is built, it is very fast to apply to new/actual retrieved estimates.



Crisp, David (2015). Measuring atmospheric carbon dioxide from space with the Orbiting Carbon Observatory-2 (OCO-2), *Earth Observing Systems, Proceedings of the SPIE*, Volume 9607, doi: 10.1117/12.2187291.

Davison, A.C. and Hinkley, D.V. (1997). Bootstrap Methods and Their Application, Cambridge University Press, Cambridge, UK,

McLachlan, Geoffrey and Peel, David (2000). Finite Mixture Models, Wiley Series in Probability and Statistics, New York.

Rodgers, Clive D. (2000). Inverse Methods for Atmospheric Sounding, Theory and Practice, World Scientific, Singapore.



Scrucca, Luca and Fop, Michael and Murphy, Thomas Brendan and Raftery, Adrian E. (2016). Mclust 5: clustering, classification and density estimation using Gaussian finite mixture models, *The R Journal*, Volume 8, Number 1, pages 205–233.

Wunch, Debra and co-authors (2017). Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) XCO₂ measurements with TCCON, *Atmospheric Measurement Techniques*, Volume 10, pages 2209–2238, doi: 10.5194/amt-10-2209-2017.



Acknowledgments

We would like to thank the Orbiting Carbon Observatory-2 Project for sponsoring this work.

©2020 California Institute of Technology. Government sponsorship acknowledged.



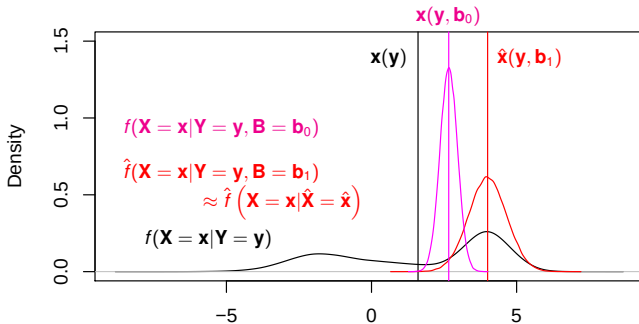
National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Backup slides



$$\mathbf{X} \sim P_{\mathbf{X}}(\mathbf{x}; \theta), \quad \mathbf{Y}_0 = F_0(\mathbf{X}, \mathbf{b}_0), \quad \mathbf{Y} = \mathbf{Y}_0 + \epsilon, \quad \epsilon \sim MVN(\mathbf{0}, \Sigma_{\epsilon}),$$
$$\theta_{\mathbf{X}} = \{\mu_{\mathbf{X}}, \Sigma_{\mathbf{X}}, \dots\}, \quad \text{and} \quad \hat{\mathbf{X}}(\mathbf{Y}, \mathbf{b}_1) = R(\mathbf{Y}, F_1, \mathbf{b}_1).$$





- ▶ “Top-down” approach: simulate the *entire* observing system, and compare retrieved states to synthetic truth over a representative ensemble of conditions.
- ▶ Treats the observing system/retrieval as an estimator and focuses on quantifying its mechanistic properties.
- ▶ Quantify mechanistic performance via $P(\mathbf{X}|\hat{\mathbf{X}})$; this distribution is the most complete description of uncertainty in $\mathbf{X}$ after observing $\hat{\mathbf{X}}$.



Model discrepancy

- ▶ Model discrepancy is $\delta^{\text{sim}} = F_0(\mathbf{X}, \mathbf{b}_0) - F_1(\mathbf{X}, \mathbf{b}_1)$.
- ▶ We would like to simulate from the distribution of $\delta^{\text{sim}} \sim \mathcal{N}(\tilde{\mu}_\delta, \tilde{\Sigma}_\delta)$.
- ▶ Assume this distribution is Gaussian with mean $\mathbb{E}(\delta^{\text{sim}})$ and covariance matrix $\text{cov}(\delta^{\text{sim}})$.
- ▶ But we only have access to $F_1(\hat{\mathbf{X}}, \mathbf{b}_1)$, which motivates the approximation,

$$\delta^{\text{sim}} \approx F_0(\mathbf{X}, \mathbf{b}_0) - F_1(\hat{\mathbf{X}}, \mathbf{b}_1) - \left[F_1(\mathbf{X}^{\text{sim}}, \mathbf{b}_1) - F_1(\hat{\mathbf{X}}^{\text{sim}}, \mathbf{b}_1) \right].$$



Model discrepancy

- Let $\mathbf{Y} \equiv F_0(\mathbf{X}, \mathbf{b}_0) + \epsilon$, and $\hat{\mathbf{Y}} \equiv F_1(\hat{\mathbf{X}}, \mathbf{b}_1)$. Then,

$$\begin{aligned}\delta^{\text{sim}} &\approx F_0(\mathbf{X}, \mathbf{b}_0) - F_1(\hat{\mathbf{X}}, \mathbf{b}_1) - \left[F_1(\mathbf{X}^{\text{sim}}, \mathbf{b}_1) - F_1(\hat{\mathbf{X}}^{\text{sim}}, \mathbf{b}_1) \right] . \\ &= \left(\mathbf{Y} - \epsilon - \hat{\mathbf{Y}} \right) - \left(\mathbf{Y}_0^{\text{sim}} - \hat{\mathbf{Y}}^{\text{sim}} \right) , \\ \delta^{\text{sim}} + \epsilon &\approx \left(\mathbf{Y} - \hat{\mathbf{Y}} \right) - \left(\mathbf{Y}_0^{\text{sim}} - \hat{\mathbf{Y}}^{\text{sim}} \right) .\end{aligned}$$

- Expected value:

$$\begin{aligned}\mathbb{E}(\delta^{\text{sim}} + \epsilon) &\approx \mathbb{E} \left(\mathbf{Y} - \hat{\mathbf{Y}} \right) - \mathbb{E} \left(\mathbf{Y}_0^{\text{sim}} - \hat{\mathbf{Y}}^{\text{sim}} \right) , \\ \tilde{\boldsymbol{\mu}}_{\delta} + \mathbf{0} &\approx \frac{1}{N} \sum_{n=1}^N \left(\mathbf{Y}_n - \hat{\mathbf{Y}}_n \right) - \frac{1}{M} \sum_{m=1}^M \left(\mathbf{Y}_{0,m}^{\text{sim}} - \hat{\mathbf{Y}}_m^{\text{sim}} \right) ,\end{aligned}$$

where $n = 1, \dots, N$ indexes actual OCO-2 retrievals, and $m = 1, \dots, M$ indexes trials of the simulation.



► Covariance:

$$\begin{aligned}\text{cov}(\delta^{\text{sim}} + \epsilon) &\approx \text{cov}(\mathbf{Y} - \hat{\mathbf{Y}}) + \text{cov}(\mathbf{Y}_0^{\text{sim}} - \hat{\mathbf{Y}}^{\text{sim}}) \\ &\quad - 2 \text{cov}(\mathbf{Y} - \hat{\mathbf{Y}}, \mathbf{Y}_0^{\text{sim}} - \hat{\mathbf{Y}}^{\text{sim}})\end{aligned}$$

$$\begin{aligned}\tilde{\Sigma}_{\delta} = \text{cov}(\delta^{\text{sim}}) &\leq \text{cov}(\mathbf{Y} - \hat{\mathbf{Y}}) + \text{cov}(\mathbf{Y}_0^{\text{sim}} - \hat{\mathbf{Y}}^{\text{sim}}) - \text{cov}(\epsilon), \\ &\approx \widehat{\text{cov}}(\mathbf{Y} - \hat{\mathbf{Y}}) + \widehat{\text{cov}}(\mathbf{Y}_0^{\text{sim}} - \hat{\mathbf{Y}}^{\text{sim}}) - \text{cov}(\epsilon),\end{aligned}$$

assuming $\text{cov}(\mathbf{Y} - \hat{\mathbf{Y}}, \mathbf{Y}_0^{\text{sim}} - \hat{\mathbf{Y}}^{\text{sim}}) \geq 0$, and ϵ and δ^{sim} are independent.