# Automatic Relevance Determination for Gaussian Processes with Functional Inputs

**Luis Damiano**[1], Margaret Johnson[2], Joaquim Teixeira[2], Max D. Morris[3], Jarad Niemi[1]

[1]Department of Statistics, Iowa State University [2]Jet Propulsion Laboratory, California Institute of Technology [3]Departments of Statistics, and Industrial and Manufacturing Systems Engineering, Iowa State University

## Automatic Dynamic Relevance Determination

A **framework for Gaussian Processes with functional inputs** [2]. We generalize previous work with indexed inputs [5, 7, 8] and incorporate the ideas behind automatic relevance determination for vectors [9].

Let $X(t) \in \mathcal{X} = \{X : [0,1] \to \mathbb{R}\}$, $t \in [0,1]$ be a functional input, $y \in \mathbb{R}$ a scalar output, and $f : \mathcal{X} \to \mathcal{Y}$ an unknown function with evaluations $y_i = f(X_i)$, $i = 1, \ldots, N \in \mathbb{N}$. We place a GP prior on $f$.

$$\mathbf{y} \sim \mathcal{N}\left(\mathbf{m}_f, \sigma_f^2\, \mathbf{R}_f + \sigma_\varepsilon^2 \mathbf{I}\right) \tag{1}$$

$$(\mathbf{m}_f)_i = m_f(X_i) \tag{2}$$

$$(\mathbf{R}_f)_{ij} = \exp\left\{-0.5\phi^{-2}\, d_f(X_i, X_j)\right\} \tag{3}$$

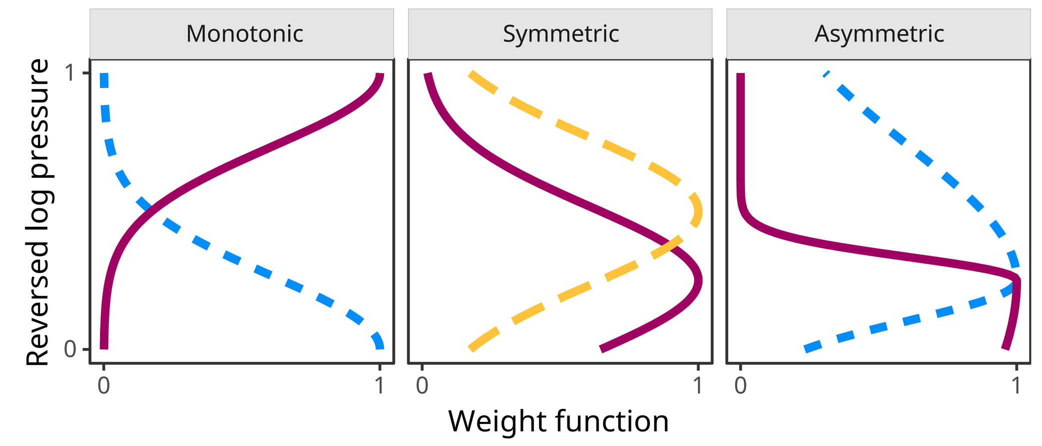
$$d_f(X_i, X_j) = \int_{\mathcal{T}} \omega(t)\big(X_i(t) - X_j(t)\big)^2 dt \tag{4}$$

$$\omega(t) : \mathcal{T} \to \mathbb{R}^+ \tag{5}$$

Space: $\sigma_\varepsilon^2 > 0$, $\sigma_f^2 > 0$, $\phi > 0$, and $m_f(\cdot)$ is left unspecified w.l.o.g.

Priors: $\phi \sim$ InvGamma, and $\sigma_f, \sigma_\varepsilon \sim$ N$^+$

## Asymmetric Squared Exponential weight function




$$\omega(t) = \begin{cases} \exp\left[-(t-\tau)^2 \lambda \kappa^{-1}\right] & \text{for } t \leq \tau \\ \exp\left[-(t-\tau)^2 \lambda \kappa\right] & \text{for } t > \tau \end{cases} \tag{9}$$

Space: $\omega(t) : \mathcal{T} = [0,1] \to (0,1]$, $\tau \in [0,1]$, $\lambda > 0$, $\kappa > 0$

Priors: $\tau \sim$ Beta, $\lambda \sim$ N$^+$, $\log(\kappa) \sim$ N

## Learning & validation

Fully Bayesian inference on the unknown quantities $\boldsymbol{\theta}$. One long chain [10] with $M \in \mathbb{N}$ post warm-up samples generated via the NUTS algorithm [3],

$$\log p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{X}) = -\frac{1}{2}(\mathbf{y} - \mathbf{m}_y)^\top {\mathbf{S}_y}^{-1}(\mathbf{y} - \mathbf{m}_y) - \frac{1}{2}\log|\mathbf{S}_y| + \log p(\boldsymbol{\theta}) \tag{6}$$

where $(\mathbf{X}, \mathbf{y})$ training input and output, $(\mathbf{m}_y, \mathbf{S}_y)$ mean and covariance.

Validation statistic posterior expectation approximated using a thinned posterior parameter sample with $\tilde{M} \in \mathbb{N}$ draws,

$$\hat{v}_1 = \tilde{M}^{-1} \sum_{\tilde{m}=1}^{\tilde{M}} N^{-\frac{1}{2}} \left\| \mathrm{E}\left\langle \boldsymbol{y}_* | \boldsymbol{X}, \boldsymbol{X}_*, \boldsymbol{y}, \boldsymbol{\theta}_{\tilde{m}} \right\rangle - \boldsymbol{y}_* \right\| \tag{7}$$

$$\hat{v}_2 = \tilde{M}^{-1} \sum_{\tilde{m}=1}^{\tilde{M}} p(\boldsymbol{y}_* | \boldsymbol{X}, \boldsymbol{X}_*, \boldsymbol{y}, \boldsymbol{\theta}_{\tilde{m}}) \tag{8}$$

where $(\mathbf{X}_*, \mathbf{y}_*)$ test input and output.

## Input relevance statistics

Most relevant location

$$\tau = \underset{t \in \mathcal{T}}{\arg\max}\, \omega(t) \tag{10}$$

Global relevance

$$\Omega = \phi^{-2} \int_{\mathcal{T}} \omega(t)\, \mathrm{d}t \tag{11}$$

## Choice of weight function

Dynamic input predictive relevance varies over the index space

Smoothness irons out erratic relevance patterns over the index space

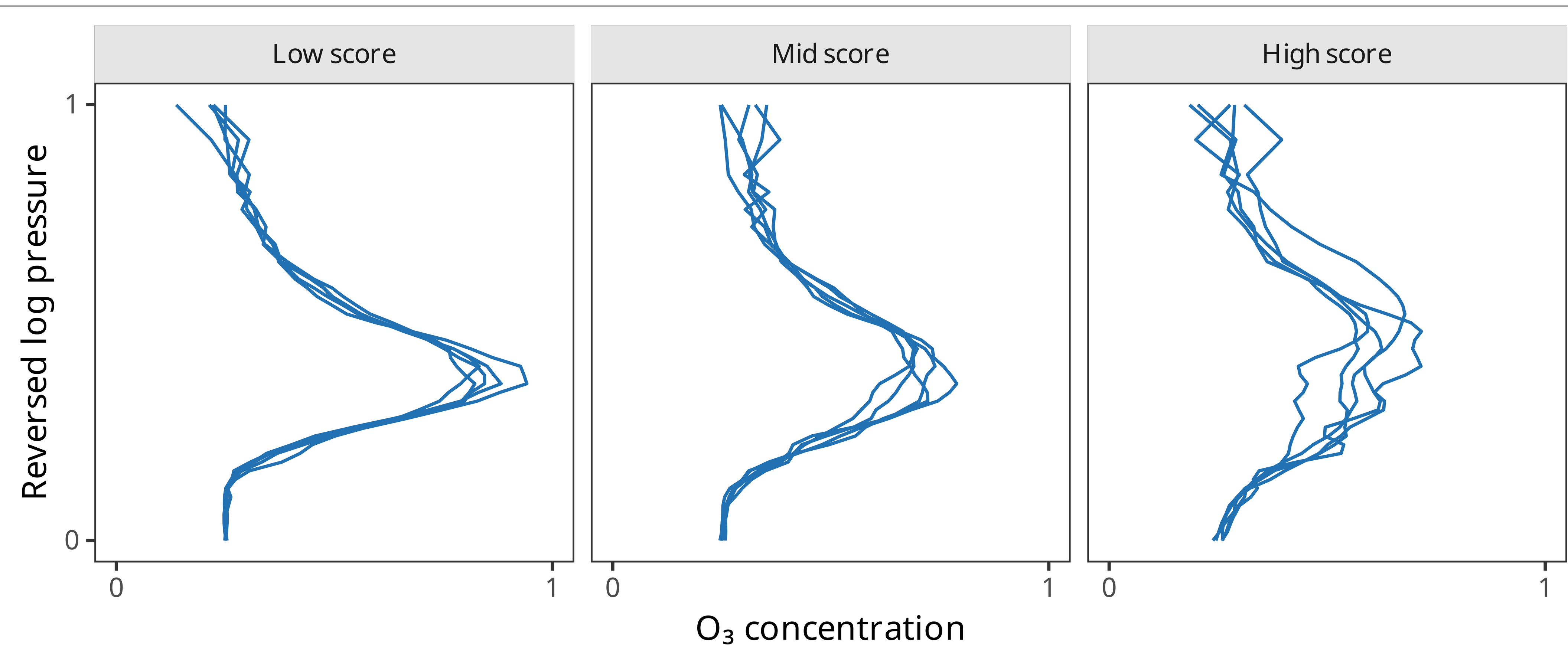
Parsimony fewer parameters than vector representations

Interpretation better physical understanding and statistical modeling

## Microwave Limb Sounder

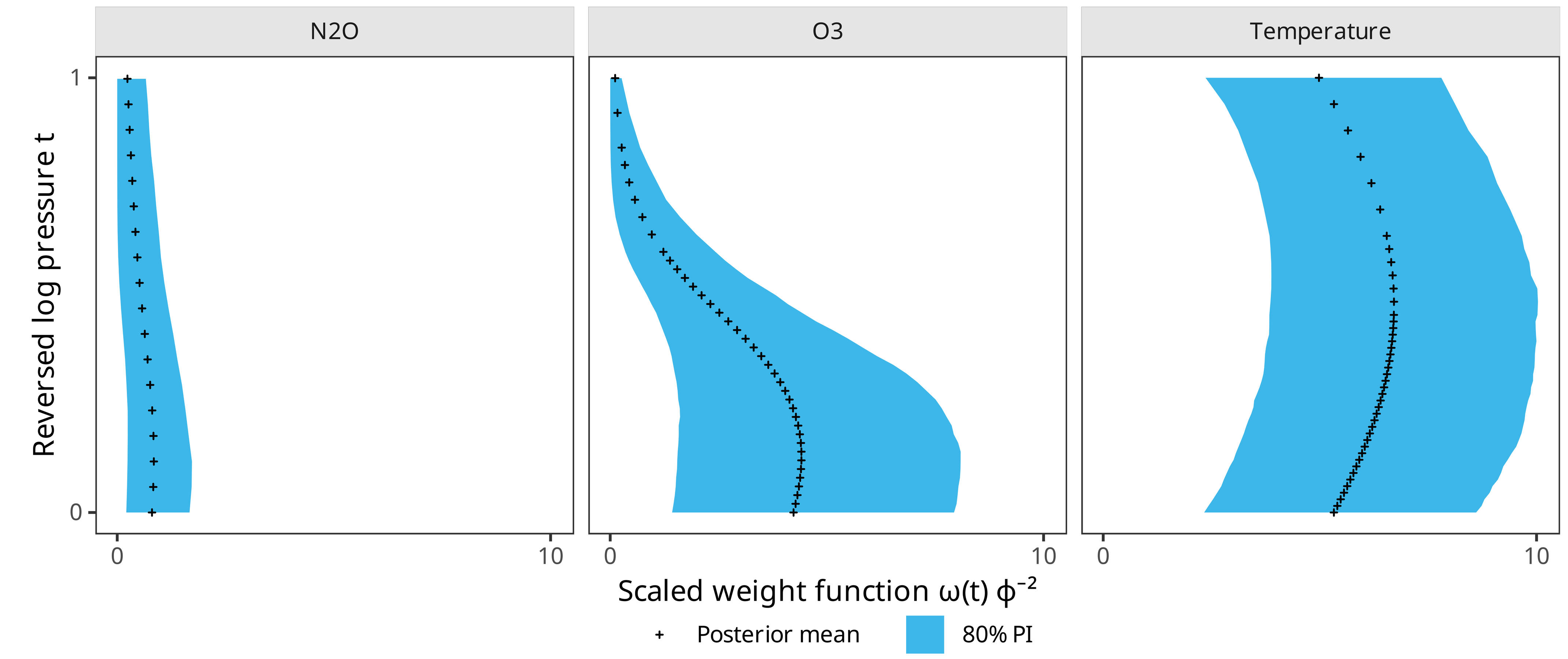
- Computer model: forward model [11, 12, 14] estimates, or *retrieves*, geophysical variables from electromagnetic radiation
- Planetwide, daily data products [6] and uncertainty experiments [1, 13] rely on a myriad of runs
- Output: score for reflected sunlight around 190GHz [4]
- Functional input: atmospheric profiles over a vertical grid
- We consider some species in pressure regions expected to be well-informed by the measurements [6]

## Data

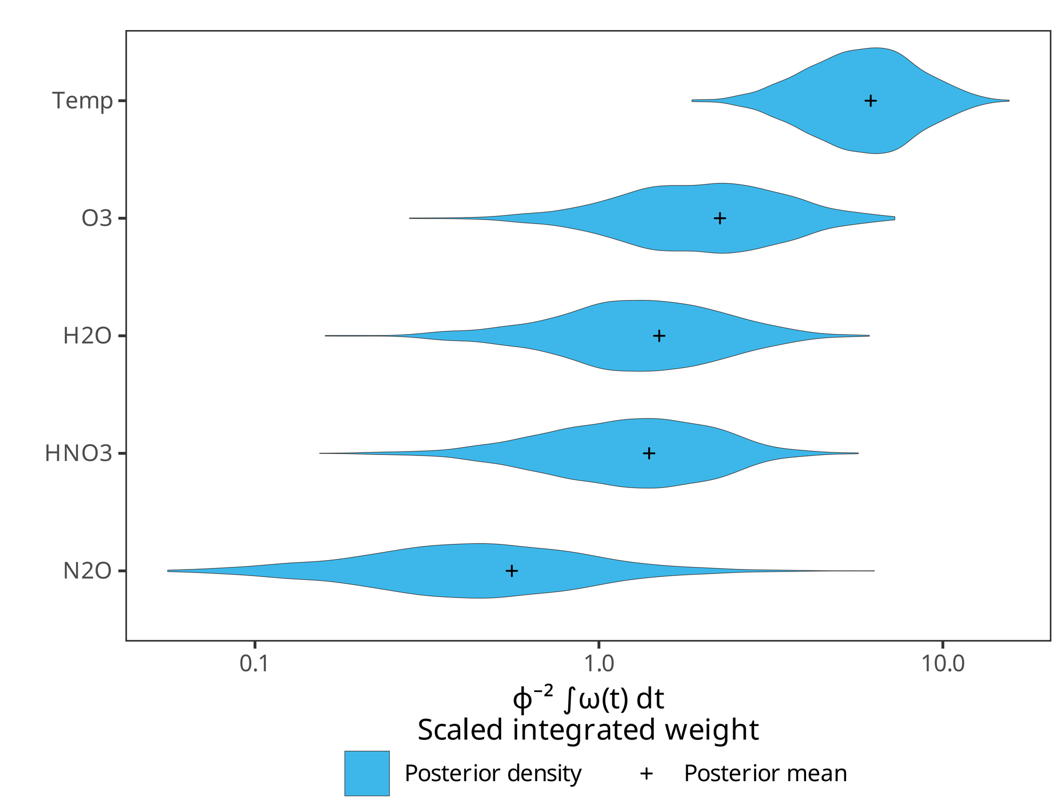



Radiance score variability seems associated with the tropopause Ozone concentration

## Are inputs equally relevant over the atmosphere?




## Are all the inputs equally relevant?




## How does ADRD compare to ARD predictionwise?

| | RMSE | $R^2$ | nCRPS | nPPLD | Cov 95% |
|---|---|---|---|---|---|
| Asym dbl exp | .18 | .97 | .10 | -.03 | .96 |
| Asym sqd exp | .18 | .97 | .10 | -.03 | .96 |
| FPCA (99%) | .34 | .88 | .19 | .03 | .95 |
| Vector SE | .23 | .95 | .13 | -.01 | .98 |
| Vector ARD | .19 | .97 | .10 | -.03 | .97 |

Underline: best in class over 8 data splits using $t_{7,.80} \approx 1.41$.

## In summary

- ADRD and ARD perform similarly predictionwise
- ADRD and ARD agree on the overall relevance patterns
- ADRD has $\sim 15$ times fewer parameters to learn
- ADRD rules out erratic patterns in relevance
- ADRD posterior patterns are consistent with the underlying science
  - Relevance within a chemical species over the vertical grid via the $\omega(t)$ function
  - Relevance between chemical species via the integrated weight statistic $\Omega$

## Onwards and upwards

Toward a framework for Gaussian processes with functional inputs $Y = f(X_1(u), X_2(v), \ldots, x_1, x_2, \ldots)$

- Multiple scalar‡, vector, and functional inputs
- Flexible functional weight forms: ADE, ASE, FEW‡
- Case studies: Microwave Limb Sounder, Water Erosion Prediction Project‡
- Exact integral for piecewise linear inputs‡

‡ Manuscript in preparation

## References

Braverman, A., J. Hobbs, J. Teixeira, and M. Gunson (Jan. 2021). "Post Hoc Uncertainty Quantification for Remote Sensing Observing Systems". In: *SIAM/ASA Journal on Uncertainty Quantification* 9.3, pp. 1064–1093. issn: 2166-2525. doi: 10/gm2gd2.

Damiano, L., M. Johnson, J. Teixeira, M. D. Morris, and J. Niemi (2022). "Automatic Dynamic Relevance Determination for Gaussian Process Regression with High-Dimensional Functional Inputs". In: doi: 10.48550/ARXIV.2209.00044.

Hoffman, M. D. and A. Gelman (2014). *The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo*.

Johnson, M., J. Teixeira, N. Livesey, and A. Braverman (Aug. 2020). *Forward Model Emulation for NASA's Microwave Limb Sounder*. Virtual.

Kuttubekova, G. (Jan. 2019). "Emulator for Water Erosion Prediction Project Computer Model Using Gaussian Processes with Functional Inputs". In: *Creative Components*.

Livesey, N. J. et al. (June 2020). *Earth Observing System (EOS) Aura Microwave Limb Sounder (MLS) Version 5.0x Level 2 and 3 Data Quality and Description Document*.

Morris, M. D. (Feb. 2012). "Gaussian Surrogates for Computer Models with Time-Varying Inputs and Outputs". In: *Technometrics* 54.1, pp. 42–50. issn: 0040-1706. doi: 10/ghbnds.

Muehlenstaedt, T., J. Fruth, and O. Roustant (July 2017). "Computer Experiments with Functional Inputs and Scalar Outputs by a Norm-Based Approach". In: *Statistics and Computing* 27.4, pp. 1083–1097. issn: 0960-3174, 1573-1375. doi: 10/ghbndr.

Neal, R. M. (1996). *Bayesian Learning for Neural Networks*. Vol. 118. Lecture Notes in Statistics. New York, NY: Springer New York. isbn: 978-0-387-94724-2 978-1-4612-0745-0. doi: 10.1007/978-1-4612-0745-0.

Raftery, A. E. and S. M. Lewis (Nov. 1992). "[Practical Markov Chain Monte Carlo]: Comment: One Long Run with Diagnostics: Implementation Strategies for Markov Chain Monte Carlo". In: *Statistical Science* 7.4. issn: 0883-4237. doi: 10.1214/ss/1177011143.

Read, W., Z. Shippony, M. Schwartz, N. Livesey, and W. Van Snyder (May 2006). "The Clear-Sky Unpolarized Forward Model for the EOS Aura Microwave Limb Sounder (MLS)". In: *IEEE Transactions on Geoscience and Remote Sensing* 44.5, pp. 1367–1379. issn: 0196-2892, 1558-0644. doi: 10.1109/TGRS.2006.873233.

Schwartz, M., W. Read, and W. Van Snyder (May 2006). "EOS MLS Forward Model Polarized Radiative Transfer for Zeeman-Split Oxygen Lines". In: *IEEE Transactions on Geoscience and Remote Sensing* 44.5, pp. 1182–1191. issn: 0196-2892, 1558-0644. doi: 10.1109/TGRS.2005.862267.

Turmon, M. and A. Braverman (Feb. 2019). *Uncertainty Quantification for JPL Retrievals*. Technical Report. Pasadena, CA: Jet Propulsion Laboratory, National Aeronautics and Space Administration, 2019.

Waters, J. et al. (May 2006). "The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura Satellite". In: *IEEE Transactions on Geoscience and Remote Sensing* 44.5, pp. 1075–1092. issn: 0196-2892. doi: 10.1109/TGRS.2006.873771.

**Acknowledgements**. MLS team at Jet Propulsion Laboratory for their insight in the instrument, the forward model, and other relevant atmospheric science concepts. Partially funded by Iowa State University through the Presidential Interdisciplinary Research Initiative on C-CHANGE: Science for a Changing Agriculture, and the Foundation for Food and Agriculture Research. **Manuscript**. Modified from arXiv:2209.00044