1   **A Machine Learning Methodology for the Generation of a Parameterization of the Hydroxyl Radical**

3   Daniel C. Anderson[1,2], Melanie B. Follette-Cook[2,3], Sarah A. Strode[2,3], Julie M. Nicely[2,4], Junhua Liu[2,3],
4   Peter D. Ivatt[2,4], Bryan N. Duncan[2]
5
6   [1] GESTAR II, University of Maryland Baltimore County, Baltimore, MD, USA
7   [2] Atmospheric Chemistry and Dynamics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD,
8   USA
9   [3] GESTAR II, Morgan State University, Baltimore, MD, USA
10   [4] Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA
11
12   *Correspondence to*: Daniel C. Anderson (daniel.c.anderson@nasa.gov)

13   **Abstract**

14   We present a methodology that uses gradient boosted regression trees (a machine learning
15   technique) and a full-chemistry simulation (i.e., training dataset) from a chemistry climate model (CCM)
16   to efficiently generate a parameterization of tropospheric hydroxyl radical (OH) that is a function of
17   chemical, dynamical, and solar irradiance variables. This surrogate model of OH is designed to be
18   integrated into a CCM and allow for computationally-efficient simulation of nonlinear feedbacks
19   between OH and tropospheric constituents that have loss by reaction with OH as their primary sinks
20   (e.g., carbon monoxide (CO), methane (CH_4), volatile organic compounds (VOCs)). Such a model
21   framework is advantageous for studies that require multi-decadal simulations of CH_4 or multi-year
22   sensitivity simulations to understand the causes of trends and variations of CO and CH_4. To allow the
23   user to easily target the training dataset towards a desired application, we are outlining a methodology
24   to generate a parameterization of OH and not presenting an "off the shelf" version of a
25   parameterization to be incorporated into a CCM. This provides for the relatively easy creation of a new
26   parameterization in response to, for example, changes in research goals or the underlying CCM
27   chemistry and/or dynamics schemes. We show that a sample parameterization of OH generated from a
28   CCM simulation is able to reproduce OH concentrations with a normalized root mean square error of
29   approximately 5%, as well as capturing the global mean methane lifetime within approximately 1%. Our
30   calculated accuracy of the parameterization assumes inputs being within the bounds of the training
31   dataset. Large excursions from these bounds will likely decrease the overall accuracy. However, we
32   show that the sample parameterization predicts large deviations in OH for an El Niño event that was not
33   part of the training dataset, and that the spatial distribution and strength of these deviations are
34   consistent with the event. This result gives confidence in the fidelity of a parameterization developed
35   with our methodology to simulate the spatial and temporal responses of OH to perturbations from large
36   variations in the chemical, dynamical, and solar irradiance drivers of OH. In addition, we discuss how
37   two machine learning metrics, Gain feature importance and SHAP values, indicate that the behavior of a
38   parameterization of OH generally accords with our understanding of OH chemistry, even though there
39   are no physics- or chemistry-based constraints on the parameterization.

40

41   **1.0 Introduction**
42   The hydroxyl radical (OH) is the dominant tropospheric oxidant. It removes numerous species from
43   the atmosphere, including carbon monoxide (CO) and methane (CH_4), which are the largest OH sinks on
44   a global scale (Spivakovsky et al., 2000;Spivakovsky et al., 1990). Recent trends in CH_4, the second most
45   important anthropogenic greenhouse gas, can potentially be explained by changes in OH abundance
46   (Rigby et al., 2017), although changes in emissions are also a likely contributor (Turner et al., 2017).

47 Likewise, the large increase in CH_4 during 2020 has been attributed to decreases in OH resulting from
48 COVID-19 related changes in NO_X (NO_X = NO + NO_2) abundance (Laughner et al., 2021). Understanding
49 the non-linear chemistry of the drivers of OH and feedbacks among these species is therefore important
50 for characterizing past and present changes in the atmosphere as well as in projecting future climate
51 scenarios.
52
53   Chemistry-climate models (CCMs) with detailed chemical mechanisms are used to investigate the
54 complex, non-linear chemistry between these species and their impacts on the atmosphere (e.g.,Fiore et
55 al., 2006;Voulgarakis et al., 2015;Gaubert et al., 2017;Holmes, 2018). The utility of CCMs for this
56 purpose is limited, however, by the large computational expense of a CCM with a full representation of
57 $O_3 - NO_X - VOC$ (Ozone, NO_X, Volatile Organic Compound) chemistry combined with the need to model
58 over decadal time scales in order to let CH_4 perturbations fully evolve (Prather, 1996). Because of this
59 computational expense, simulations are necessarily limited to a short time frame, performed at coarse
60 horizontal resolutions, and/or limited in the number of sensitivity runs that can be performed (e.g., Fiore
61 et al., 2006;Holmes, 2018;Voulgarakis et al., 2015).
62
63   There are several alternatives (i.e., surrogate models) to running a full chemical mechanism that
64 capture some of the relationship between OH and trace gases, such as CO and CH_4, and are less
65 computationally expensive. Prescribed OH fields, either static or annually-varying, from a full chemistry
66 simulation or a climatology have been used for decades to simulate and understand trends in CO and
67 CH_4 in a computationally-efficient way (e.g., Saito et al., 2013;Wang et al., 2004). However, this method
68 linearizes CO and CH_4 chemistry with OH, preventing the simulation of nonlinear feedbacks in changes in
69 CO and CH_4 on OH, and thus could bias, for instance, interannual CH_4 changes (Chen and Prinn, 2006).
70
71   Spivakovsky et al. (1990) developed a parameterization of OH, later updated by Duncan et al. (2000),
72 that captures many of the nonlinear feedbacks between OH and tropospheric constituents (e.g., CO,
73 CH_4, VOCs) that have loss by reaction with OH as their primary sinks. The method to generate the
74 parameterization uses higher order polynomials with various chemical species, meteorological variables,
75 and variables related to solar irradiance as inputs. The degree of the nonlinear impacts simulated by the
76 parameterization of OH depends on the method used to populate these inputs. For instance, many of
77 the meteorological and solar irradiance variables may be provided by the model at run time. The
78 chemical variables that are not all simulated explicitly in the surrogate model may be provided as
79 climatological or monthly means from a full chemistry simulation. Duncan et al. (2007a) and Duncan and
80 Logan (2008) used this parameterization of OH in an atmospheric model of CO to elucidate the causes of
81 trends and interannual variations in CO from 1988-1997 on regional to global scales as well as those
82 observed by individual *in situ* monitors around the world.
83
84   Building on the CO-OH studies of Duncan et al. (2007a) and Duncan and Logan (2008), Elshorbany et
85 al. (2016) developed the computationally Efficient CH4–CO–OH (ECCOH) chemistry module, which
86 captures many of the nonlinearities and feedbacks of the CH_4-CO-OH system without the computational
87 expense of a full chemistry simulation. ECCOH calculates 24-hour averaged OH from a combination of
88 archived (e.g., multiple VOCs, NO_X) and online (e.g., pressure, temperature, cloud albedo) chemical,
89 meteorological, and solar irradiance variables. Despite the partial reliance of the parameterization of
90 OH in ECCOH on archived fields, its strength lies in the ability to calculate OH at a significantly reduced
91 computational expense (Duncan et al., 2000;Elshorbany et al., 2016). ECCOH has been successfully
92 implemented in the NASA Goddard Earth Observing System (GEOS) general circulation model (GCM).
93

94   Through manipulation of the input parameters (i.e., chemical, meteorological, and solar irradiance
95   variables) to the parameterization of OH, as well as emissions and dynamics, ECCOH can help
96   deconvolve the causes of local to global trends and variations in OH, CO, and CH4. For example, Strode
97   et al. (2015) used the ECCOH module to investigate the effects of different model biases in GEOS on
98   simulated OH. To do this, they performed multiple sensitivity simulations, adjusting tropospheric water
99   vapor, ozone, and NO_x to match satellite observations, to understand the impacts on OH and CH_4
100   lifetime. Similarly, Elshorbany et al. (2016) investigated the impacts of varying CH_4 and CO emissions on
101   the growth rate of atmospheric methane concentrations through multiple sensitivity runs. One
102   limitation of ECCOH in the configuration used in Strode et al. (2015) and Elshorbany et al. (2016),
103   however, is the difficulty in updating the parameterization to reflect advances in atmospheric chemistry.
104
105   Machine learning algorithms are one potential method to quickly and accurately generate a new
106   parameterization of OH. A variety of machine learning techniques, such as neural networks (Nicely et
107   al., 2017;Nicely et al., 2020;Kelp et al., 2020), ridge regression (Nowack et al., 2018), random forest
108   regression (Keller and Evans, 2019;Sherwen et al., 2019), and gradient boosted regression trees (GBRTs)
109   (Ivatt and Evans, 2020;Stirnberg et al., 2020) have been successfully used in atmospheric chemistry
110   applications. In particular, GBRT models (Elith et al., 2008;Chen and Guestrin, 2016) use an ensemble of
111   decision trees to predict the value of a target based on multiple inputs and have been used to predict
112   surface OH using a combination of satellite observations and output from 3-dimensional models (Zhu et
113   al., 2022). Decision trees are created sequentially, with each new tree minimizing a cost function based
114   on the results of the previous tree (Elith et al., 2008;Stirnberg et al., 2020). Unlike some other machine
115   learning algorithms, such as neural networks, regression tree methods have easily interpretable metrics
116   that help highlight the influence of the different input variables (Yan et al., 2016). These metrics can
117   help further understanding of the model behavior in ways other machine learning techniques cannot.
118   GBRT models are also relatively quick to generate and can capture the highly non-linear relationships
119   that describe tropospheric chemistry (Ivatt and Evans, 2020).
120
121   We present a new method for the efficient generation of a parameterization of OH using GBRTs and
122   a full chemistry simulation from a CCM, which serves as the training dataset. We illustrate our method
123   by generating a parameterization of OH for the ECCOH module (Elshorbany et al., 2016), which captures
124   many of the nonlinearities and feedbacks of the CH_4-CO-OH system, as implemented into the NASA
125   GEOS GCM. Our methodology allows for the parameterization to be easily and rapidly regenerated in
126   response to changes in, for instance, the underlying model chemical mechanism (e.g., updates to the
127   chemical rate constants or absorption cross-sections) or model dynamics, which affect many of the
128   variables that influence OH (e.g., Anderson et al., 2021). Likewise, the parameterization can be modified
129   to include new input variables. Users can and should retrain the parameterization with datasets that are
130   appropriate for the intended application. That is, we are not offering a parameterization for "off the
131   shelf" use but a methodology by which a user can easily create a parameterization suitable for their
132   needs. In Section 2, we outline the methodology used to develop the parameterization of OH, while in
133   Section 3, we evaluate performance of the parameterization. Finally, in Section 4, we summarize the
134   results and discuss implications for scientific research.
135
136   **2.0 Description of the Methodology to Generate a Parameterization of OH**
137   In this section, we outline the methodology to generate a parameterization of OH that may be used
138   in research studies as discussed above. Specifically, we illustrate the methodology by describing the
139   creation of a sample parameterization of OH for the ECCOH module that predicts daily averaged OH. In
140   Section 2.1, we present the creation of the training dataset, and in Section 2.2, we outline the
141   methodology used to create the parameterization of OH.

142
2.1 Creation of the Training Dataset for a Parameterization

143 **2.1 Creation of the Training Dataset for a Parameterization**

144   We created the training dataset using output from a 40-year (1980 -2019) GEOS CCM simulation,
145 consistent with our intent to integrate the parameterization into the ECCOH modeling framework. This
146 simulation, called MERRA2 GMI (https://acd-ext.gsfc.nasa.gov/Projects/GEOSCCM/MERRA2GMI/), was
147 run in replay mode (Orbe et al., 2017) with MERRA2 (Modern Era Retrospective analysis for Research
148 and Applications) meteorology (Gelaro et al., 2017) and the Global Modeling Initiative (GMI) chemical
149 mechanism (Duncan et al., 2007b;Strahan et al., 2007). Aerosols were calculated with the Goddard
150 Chemistry Aerosol Radiation and Transport (GOCART) module (Chin et al., 2002;Colarco et al., 2010).
151 The model was run at a resolution of c180 on the cubed sphere (approximately 0.625° longitude by 0.5°
152 latitude) with 72 vertical layers. In this analysis, we use only tropospheric output at daily and monthly
153 resolutions. The GMI chemical mechanism includes approximately 120 species and 400 reactions,
154 characterizing the photochemistry of the troposphere and stratosphere. Further simulation details,
155 including a description of the emissions, are available elsewhere (Anderson et al., 2021;Strode et al.,
156 2019).

157

158   We created a dataset of training targets, representing the full range of simulated OH values, for
159 each month. We generate parameterizations for each month instead of one, yearly parameterization to
160 increase computational efficiency of the generation of the parameterization. The spatiotemporal
161 variability in the abundance and emissions of OH drivers on the yearly scale would necessitate a far
162 larger dataset and more complicated sampling procedures to ensure representativeness of both OH and
163 the input variables. As demonstrated in Section 3.0, the adopted monthly approach accurately captures
164 OH while limiting the size of the training dataset.

165

166   We generated the training dataset using daily averaged data. For each day of a month, we divided
167 all simulated tropospheric OH concentrations from the 40-year simulation into 20 equally-sized
168 percentile bins (i.e., $0 - 5^{th}$ percentile, $5^{th} - 10^{th}$ percentile, etc.). Then, we randomly selected 200,000
169 values from each bin, resulting in 4,000,000 training targets for each day of the month. We also
170 included the maximum and minimum OH values of the entire dataset to represent the full range of
171 values. We then combined training targets to form one large dataset with 120,000,000 values (for a 30-
172 day month), encompassing the full range of OH concentrations from each day of the month. To limit the
173 size of the training dataset, we then subsampled these targets, again randomly selecting 200,000 values
174 from equally-sized percentile bins of OH concentration. The procedure resulted in a dataset with
175 4,000,000 training targets that span all days within a given month. A schematic of this process is shown
176 in Figure S1. We omitted data from 4 years (1985, 1995, 2005, 2015) from the training dataset for
177 model evaluation and from an additional year, 2016, for an El Niño case study discussed in Section 3.3.
178 We also created a training dataset for monthly-averaged output, discussed in Sect. 4.0, using an
179 analogous process.

180

181   Finally, for each OH target, we extracted the inputs for the regression tree parameterization from
182 the MERRA2 GMI simulation from the corresponding model grid box. We list parameterization inputs in
183 Table 1. The parameterizations of Spivakovsky et al. (2000), Duncan et al. (2007a) and Elshorbany et al.
184 (2016), along with expert knowledge of OH chemistry, informed our choice of inputs. The relative
185 location of a particular OH target is indicated with the latitude and pressure variables. As discussed in
186 the next section, NO_2 serves as a sufficient proxy for the impact of NO_X and NO_y on OH, so NO_2 is the
187 only reactive nitrogen species included as an input parameter. For both ice and water cloud as well as
188 aerosol optical depths, we include the optical depth above and below each datapoint as separate inputs.
189 We use aerosol optical depth (AOD) at 550 nm, calculated from the GOCART aerosol module. We took

190 all 27 inputs from the MERRA2 GMI simulation except surface UV albedo, which we took from the Ozone
191 Monitoring Instrument (OMI) climatology of Kleipool et al. (2008).
192
193 *Table 1:* *Inputs to the machine learning parameterization of OH. UV Albedo is the value at the surface. Cloud fraction*
194 *is the fraction at a given model level. C4 & C5 alkanes are one input as they originate from a lumped variable in the*
195 *GMI mechanism.*

Chemical Inputs		Meteorological/Radiative Inputs	
NO_2	Formaldehyde (HCHO)	Temperature	Stratospheric O_3 Column
CO	Hydrogen peroxide (H_2O_2)	Cloud Fraction	Aerosol Optical Depth above
CH_4	Methyl hydroperoxide (CH_3OOH; MHP)	Latitude	Aerosol Optical Depth below
O_3	Acetone (CH_3COCH_3)	UV Albedo	Water Cloud Optical Depth above
Isoprene (C_5H_8)	C4 & C5 Alkanes	Water Vapor	Water Cloud Optical Depth below
Propene (C_3H_6)	Ethane (C_2H_6)	Pressure	Ice Cloud Optical Depth above
Propane (C_3H_8)		Solar Zenith Angle	Ice Cloud Optical Depth below

196
197   While we have used the publicly-available MERRA2 GMI dataset to train the sample
198 parameterization described in this manuscript, the training data could come from any simulation or
199 combination of self-consistent simulations that has output of the variables outlined in Table 1. These
200 training datasets could come from existing simulations, which would greatly reduce computational
201 expense, or from a training dataset tailored for the purposes of a given study. Even though we use
202 daily-averaged training data for ECCOH, a user could train the parameterization with a dataset at any
203 temporal resolution in order to make the parameterization compatible with a specific modeling platform
204 or research goal. As discussed later, the parameterization performs best when applied to
205 photochemical environments analogous to those on which it was trained. Therefore, users should
206 carefully ensure that the training dataset reasonably encompasses the full range of photochemical
207 environments necessary for a given sensitivity test or experiment. For example, as we will discuss
208 further in Section 4, because the MERRA2 GMI training dataset only covers 1980 – 2018, it is
209 inappropriate to use this for an application exploring changes in CH_4 from the pre-industrial period to
210 2100. Instead, a new training dataset covering that time period would be required.
211
212 **2.2 Creation of the GBRT Parameterization**
213   While other machine learning methods could likely produce parameterizations with similar
214 performance as the one we describe here, we use GBRTs because of the speed in training a new
215 parameterization, their accuracy, and the interpretability of the parameterization itself. We refer to the
216 GBRT models as parameterizations to prevent confusion when referring to 3-dimensional models.
217   We used the XGBoost package (Chen and Guestrin, 2016) version 0.81 in Python version 3.6 to
218 create 12 parameterizations of OH (one for each month), training the parameterizations on the MERRA2
219 GMI datasets described in Sect. 2.1. Each parameterization outputs 24-hour averaged OH. For each
220 month, we used 80% of the dataset (3.2 million datapoints) for model training and the remainder for
221 model validation. In addition, as outlined in-depth in Sections 2.1 and 3.0, we also evaluated the model
222 on 5 years of data not included in the model training. Increasing the size of the training dataset did not
223 improve model performance but did increase model training time, so the training set was restricted to a
224 size that represented the full ranges of OH values.
225

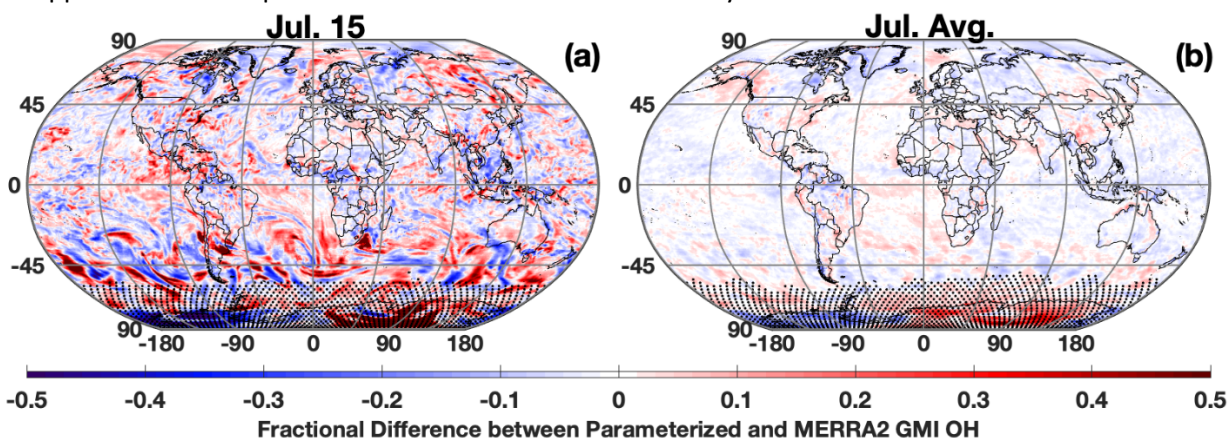
226   To maximize parameterization performance while also balancing the potential of overfitting, we
227 tuned hyperparameters, including the learning rate, the maximum tree depth, and the number of trees.
228 We chose hyperparameter values that minimized the parameterization normalized root mean square
229 error (NRMSE) (Eq. 1.) of the training dataset. In Eq. 1, N is the number of samples, y is the MERRA2 GMI
230 OH, $\hat{y}$ is the parameterized OH, and IQR is the interquartile range of the dataset. We set the learning
231 rate, which controls the magnitude of change when adding a new tree, to 0.1, while we varied the
232 maximum tree depth and number of trees from 6 to 22 and from 10 to 150, respectively. For both
233 maximum tree depth and number of trees, NRMSE initially dropped significantly with increasing value,
234 representing sharp improvement in parameterization performance. NRMSE values eventually
235 plateaued, increasing parameterization runtime without noticeably improving performance. A
236 combination of a maximum tree depth of 18 and 100 trees balanced performance with model training
237 and run time.

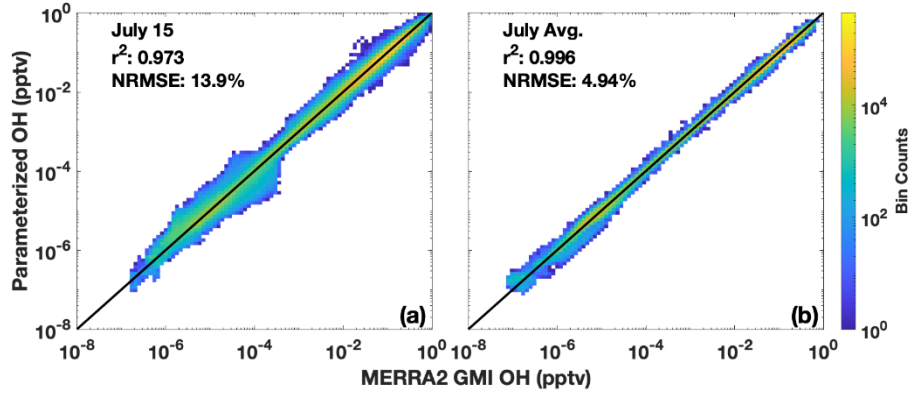
$$NRMSE = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}}{IQR} \tag{1}$$

239   We also evaluated inputs into the parameterization to ensure that each did not lead to decreased
240 performance, finding that no single variable dominates model performance and no variable reduces
241 performance. We retrained the parameterization 27 times for July, removing each input successively, to
242 determine its impact on the NRMSE. When we applied the resultant models to the July 2005 validation
243 dataset, the percentage change in the NRMSE generally increased by less than 1%. The small differences
244 in NRMSE indicate that there are likely variables that provide duplicate information to the
245 parameterization. As will be discussed in Sect. 3.2, however, the relative importance of inputs varies by
246 month, and some variables, though not important on average, have a large influence in specific chemical
247 environments. Because of these factors and a desire to use a consistent set of input variables across all
248 months, we did not remove any inputs from the parameterization as a result of this analysis.
249
250   Finally, we omit NO_X and NO_Y as parameterization inputs because we find that NO_2 is sufficient as an
251 input to capture the impact of reactive nitrogen on OH in the parameterization. Because of the
252 importance of NO_X in OH production (Spivakovsky et al., 2000;Anderson et al., 2021), we tested
253 performance by substituting different reactive nitrogen species for NO_2 as inputs to the
254 parameterization. We trained three additional parameterizations, including ones with NO_X, NO_Y (NO_Y =
255 $NO + NO_2 + PAN + 2N_2O_5 + HNO_3$ + alkyl nitrates), and the individual NO_Y species. Parameterization
256 performance did not improve noticeably with the inclusion of NO_X or the individual NO_Y species.
257 Including NO_Y as a group actually decreased performance.
258
259 **3.0 Evaluation of the parameterization of OH for the ECCOH module**
260   We now evaluate the performance of the parameterization of OH for the ECCOH module created
261 with the machine learning methodology. In Section 3.1, we compare the OH calculated with the
262 parameterization to that from the MERRA2 GMI simulation, showing agreement in both local OH
263 concentrations as well as in global metrics, such as CH_4 lifetime (τ_{CH4}). In Section 3.2, we investigate the
264 parameterization Gain feature importance and SHapley Additive Explanations (SHAP) values to
265 understand the relative contributions of inputs to parameterization performance and to demonstrate
266 that, even though there are no physics- or chemistry-based constraints, parameterization behavior
267 accords with our understanding of OH chemistry. We explore the ability of the parameterization to
268 predict OH in response to strong deviations in its drivers from the climatological mean in Section 3.3, by
269 examining two El Niño events. Finally, we note that we evaluate an offline version of the
270 parameterization of OH and not one integrated within the ECCOH framework. However, the
271 performance will be similar based on preliminary testing and similarities in implementation to the

272 previous parameterization, which has been extensively evaluated (Elshorbany et al., 2016) in the GEOS
273 GCM.
274

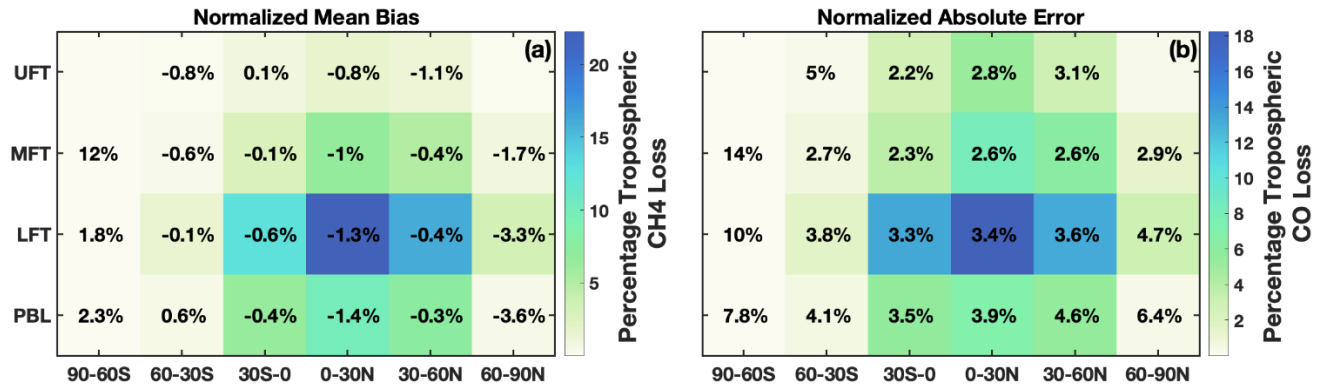
3.1   Ability of the parameterization to reproduce modeled OH and global OH metrics

276   The parameterization is able to reproduce both the spatial distribution and concentration of
277 daily-averaged OH, although with noticeable errors at high latitudes in the winter hemisphere, which is
278 unimportant as OH is seasonally low. Figure 1a shows the fractional difference between OH calculated
279 with the parameterization and OH from the MERRA2 GMI simulation for July 15, 2005, a date omitted
280 from the training dataset. The parameterized and MERRA2 GMI OH fields are shown in Figure S2. The
281 OH in Figure 1 has been averaged over the lower free troposphere (LFT), defined as pressures between
282 the top of the planetary boundary layer (PBL) and 500 hPa. Agreement is similar throughout the
283 troposphere, but we highlight this region because of its importance for CH_4 and CO loss (Spivakovsky et
284 al., 2000). For July 15, there are notable regions of bias, particularly poleward of 30° S where OH is low
285 (Fig. S2). While the source of this error is unknown, it could result from a tendency of regression tree
286 models to have larger bias for lower values (Nowack et al., 2021). This results in a NRMSE for the entire
287 troposphere of 13.9% (Fig. 2a). At higher concentrations, the correlation between the MERRA2 GMI
288 simulation and the parameterized OH is much tighter than at lower concentrations, although the highest
289 density at all concentrations is centered around the 1:1 line. Because the CO and CH_4 lifetimes are much
290 longer than one day, the accuracy of the parameterization on monthly timescales is more relevant to
291 the applications of the parameterization than an individual day.

292
293 *Figure* 1*:* Fractional difference between the parameterized and MERRA2 GMI OH averaged over the LFT (top of the
294 PBL to 500 hPa) for July 15, 2005 (a) and averaged across all days for July 2005 (b). Regions with low OH, defined
295 as a mixing ratio of less than 0.005 pptv, are indicated with stippling.

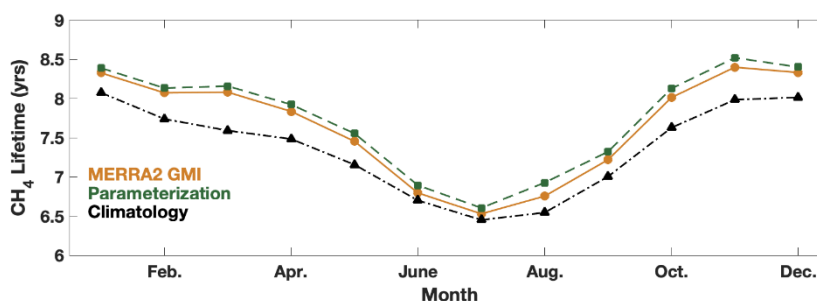
296   When we average the daily output to the monthly scale, the parameterization can reproduce
297 the global OH distribution with little error (Fig. 1-2). For July 2005, the percentage difference between
298 the parameterized OH and output from the MERRA2 GMI simulation in the LFT (Fig. 1b) and throughout
299 the troposphere (Fig. S3) is generally within 10%, outside of the Southern Hemispheric high latitudes,
300 where it is polar night and OH concentrations are negligible. The random errors evident in the daily data
301 in Figure 1a average out on the monthly timescale, resulting in a tight correlation (r^2 = 0.996) and a
302 NRMSE of 4.94% for all tropospheric values (Fig. 2b). Similar results are found for the July model when
303 applied to other years (Table S1) and for parameterizations developed for other months (Fig. S3 and S4).
304 Averaging the daily output over the monthly period yields a better NRMSE by more than a factor of two
305 over climatology (NRMSE = 11%), defined as the mean OH from the MERRA2 GMI simulation averaged
306 over 1980 to 2019.
307

308
309 **Figure 2:** Scatter density plot of tropospheric OH from the MERRA2 GMI simulation plotted against OH calculated by
310 the parameterization for July 15, 2005 (a). Panel (b) shows the 24-hour averaged OH output by the parameterization
311 averaged across all July days for 2005. Colors indicate the number of data points in each bin. The r^2 of a linear least
312 squares regression and the NRMSE are also indicated.

313

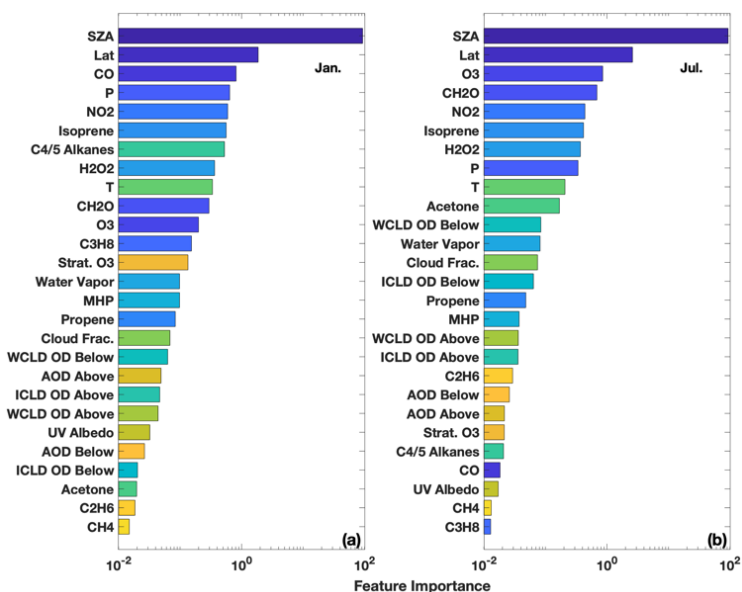
314   In regions where global CO and CH_4 loss are most important, parameterization biases and errors
315 are low. For CO and CH_4, tropospheric loss to OH maximizes in the LFT in the 0 - 30° latitude band of the
316 summer hemisphere with near negligible loss in the winter hemisphere polar region (Fig. 3). The
317 comparatively large over- and underestimates over Antarctica evident in Figure 1 are irrelevant to the
318 OH/CO/CH_4 cycle because of the low loss rate in this region. In contrast, in regions where CO and CH_4
319 loss maximize, the parameterization is biased low by only -0.3 to -1.4%. The normalized absolute error
320 varies between 2.2% and 4.6% in the tropics and Northern Hemispheric mid-latitudes for all
321 tropospheric layers (MFT: pressures between 500 and 300 hPa, UFT: pressures between 300 hPa and the
322 tropopause). Results are similar for other months.

323

324
325 **Figure 3:** (a) Percentage of total tropospheric CH_4 lost to reaction with OH for July 2005 averaged over 30° zonal
326 mean bins and 4 atmospheric layers is shown by the background colors. The percentage loss values account for the
327 mass of each region relative to the total atmospheric mass. Percentages indicate the normalized mean bias of the
328 parameterization with respect to the MERRA2 GMI simulation. Statistics for the polar UFT are omitted because low
329 tropopause heights limit the data amount in these regions. (b) Same as (a) except for tropospheric CO loss and the
330 normalized absolute error.

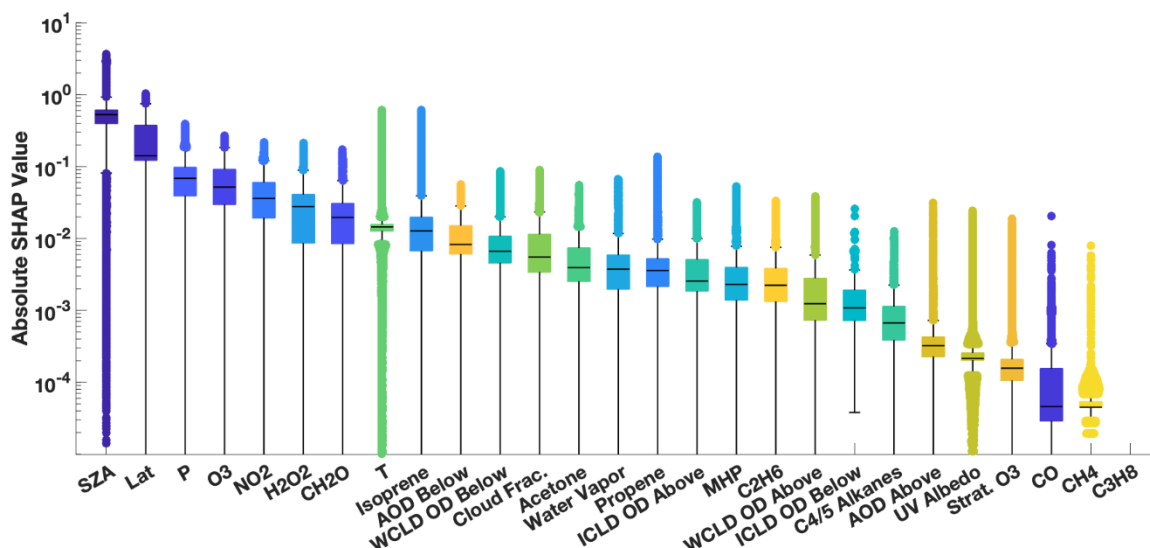
331   The parameterization is also able to reproduce global mean metrics of OH, such as τ_{CH4}, within
332 1.3% on average. For each month of 2005, we calculated the global, mean mass-weighted tropospheric
333 OH as described in Lawrence et al. (2001) and the mean tropospheric τ_{CH4} with respect to OH as
334 described in Nicely et al. (2020) for the MERRA2 GMI simulation, the parameterization, and the
335 climatological mean, defined as the average value from the MERRA2 GMI simulation between 1980 and

336   2019. Results for τ_{CH4} are shown in Figure 4, and for mass-weighted OH in Figure S5. The
337   parameterization captures the seasonality of the τ_{CH4}, with a minimum in boreal summer and a
338   maximum in boreal winter. Agreement varies slightly by month, differing by only 0.8% in January and up
339   to 2.5% in August, although the bias is systematically low for 2005 and the other validation years (Table
340   S1). These values are reasonable and much smaller than the inter-model variability often seen in model
341   intercomparison projects (e.g., Nicely et al., 2020;Voulgarakis et al., 2013). Similar results are found for
342   the global, mean mass-weighted OH. The Northern Hemispheric/Southern Hemispheric OH ratio (Fig.
343   S5) also generally agrees within 0.5% for all months, again with the exception of August. The
344   comparatively weaker performance for August is limited to 2005, however, as performance of the
345   August parameterization in the other validation years (1985, 1995, and 2015) is closer to the 1%
346   difference shown by the parameterizations for the other months. The parameterizations present a
347   significant improvement over the climatological mean, which for 2005, consistently underestimates τ_{CH4}
348   for all months and by up to 6% in March.
349

350
351   ***Figure*** 4*:* Global mean methane lifetime with respect to tropospheric OH from the parameterization (green squares)
352   and MERRA2 GMI (orange circles) for 2005 and the climatological average (black triangles) calculated from MERRA2
353   GMI for 1980 - 2019.

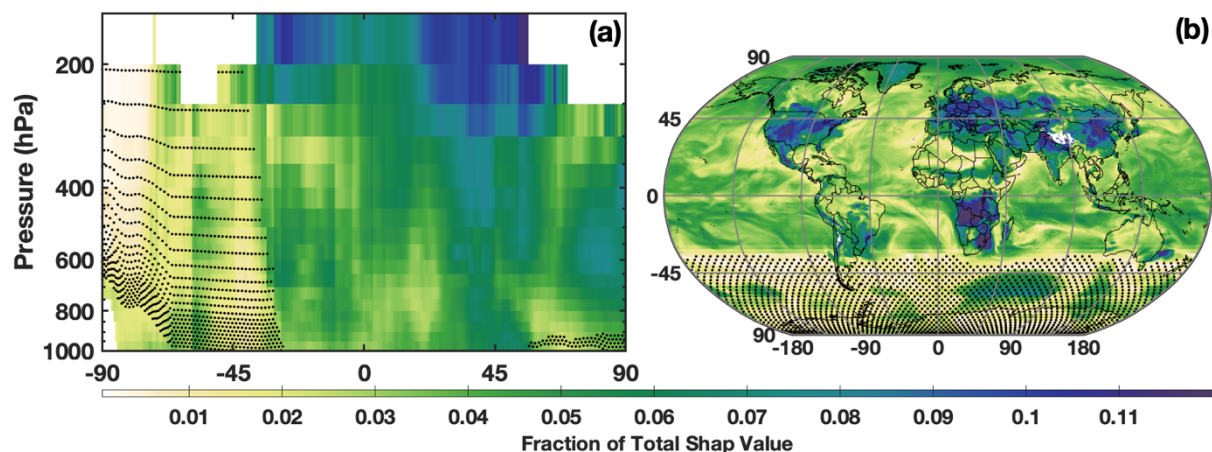
354   **3.2   Understanding the relative importance of input parameters**
355   While we have demonstrated that the parameterization is able to reproduce OH accurately, it is
356   also instructive to understand the relative importance the parameterization places on each of the
357   inputs. Although this parameterization is not process-based, understanding how the parameterization
358   responds to different inputs can help determine if the regression tree is responding in a way consistent
359   with current understanding of OH chemistry, although there are limitations to the information that can
360   be gleaned from these metrics. We evaluate the regression tree parameterization using two metrics,
361   the Gain feature importance as output by the XGBoost package, and SHAP values.
362
363   **3.2.1 Investigating the Gain feature importance**
364   The Gain feature importance (Chen and Guestrin, 2016) is a measure of the improvement in model
365   accuracy achieved from adding branches in the model corresponding to a specific input variable. The
366   Gain value therefore indicates the relative importance of each input for the model as a whole but not
367   for individual datapoints. The Gain values for each input for the January and July models are shown in
368   Figure 5. While there are differences between the two months, several features are similar. Variables
369   that indicate geographic location (e.g., SZA, latitude, and pressure) and chemical species that have
370   previously shown to be dominant drivers of OH variability (e.g., NO_2, O_3, CO) and/or OH proxies (e.g.,
371   HCHO) (Wolfe et al., 2019;Murray et al., 2021) have some of the highest Gain values. As we show
372   below, caution should be used in extrapolating results from the Gain values to a process-based
373   understanding of OH without prior knowledge of its response to chemical and dynamical drivers.
374

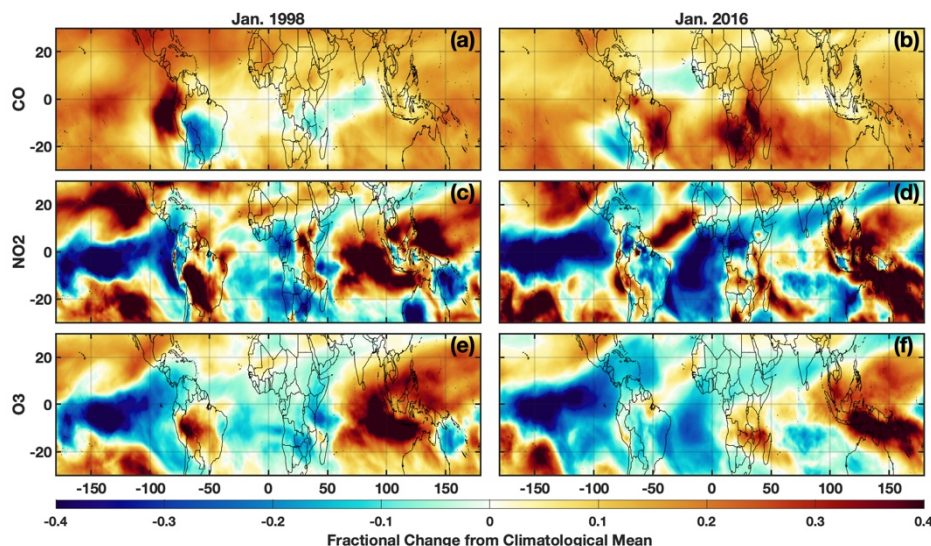
375   The relative importance of variables that indicate location is consistent with OH chemistry and
376 previous parameterization studies. Primary OH production is driven by the photolysis of O_3 followed by
377 the subsequent reaction of the O^1D radical, produced from that photolysis, with water vapor (e.g.,
378 Spivakovsky et al., 2000). Thus, the OH distribution is strongly dependent on SZA, latitude, and
379 pressure. This is consistent with the parameterization, where SZA and latitude have the highest Gain
380 values for both months examined here, as well as with the results of Duncan et al. (2000), who
381 highlighted the importance of latitude in their parameterization.
382
383   Similarly, the chemical species that are most important for controlling OH distribution on the global
384 scale also tend to have higher Gain values. As discussed above, O_3 and NO_X chemistry is instrumental in
385 controlling primary and secondary OH production on global scales (e.g., Spivakovsky et al.,
386 2000;Anderson et al., 2021), consistent with their comparatively high Gain values. HCHO, an oxidation
387 product of the reaction of OH with many VOCs, has been found to be a suitable proxy for OH in the
388 remote atmosphere (Wolfe et al., 2019), consistent with its relative importance in both the July and
389 January models.

390
391 **Figure 5:** The feature importance (gains) of the January (a) and July (b) parameterizations as calculated by XGBoost.
392 Variables are sorted by their relative importance. WCLD = Water cloud; ICLD = Ice Cloud; OD = Optical Depth.
393 "Above" and "below" for the optical depth variables indicate the optical depth above and below a particular model
394 grid box. Colors have no specific meaning but are specific to individual inputs for all panels of Figures 5 and 6.
395
396   The relative importance of global OH sinks in the parameterization, however, demonstrates the
397 limitations of using the Gains values to interpret the regression tree model in a process-based way. CO,
398 the dominant OH sink on a global scale (Spivakovsky et al., 2000), is the most important chemical input
399 for the January parameterization, although it is relatively unimportant in the parameterizations for all
400 other months. While tropical CO variability in MERRA2 GMI and biomass burning emissions (Duncan,
401 2003b) are larger in boreal winter than July, there is no process-based explanation for why CO should be
402 different in January from December or February. Differences in the relative importance of CO between
403 the two months does not imply that OH sensitivity to CO in MERRA2 GMI or the atmosphere varies in
404 the same manner. Instead, the differences simply indicate that the parameterization algorithm finds CO
405 to be more useful in predicting OH in January than July. Similarly, CH_4, the second most important OH
406 sink on the global scale, has low Gain values, suggesting it has little impact on model performance. This

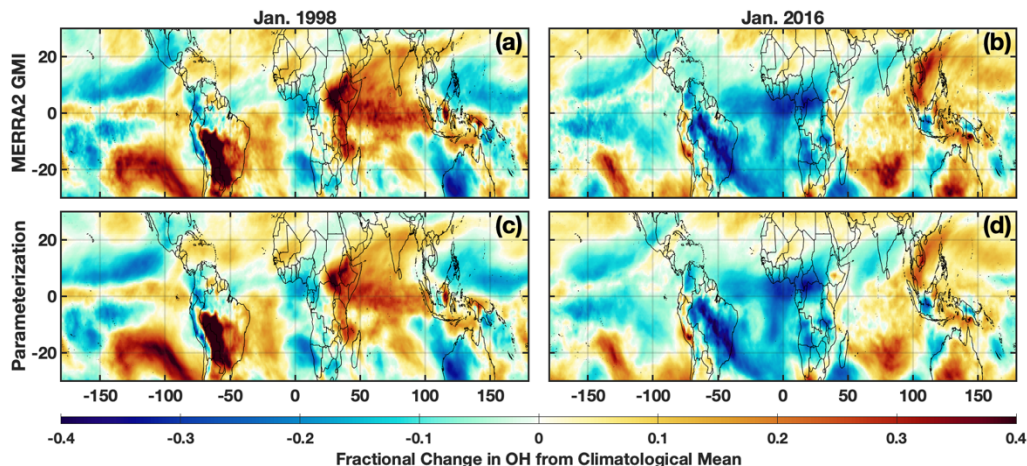
407 is likely because, in the MERRA2 GMI simulation, CH_4 concentrations vary little within a given latitude
408 band due to CH_4 surface concentrations being set as a boundary condition. The methane distribution
409 therefore provides little additional information beyond that already contained in the variables that
410 indicate location.
411
412 **3.2.2. Investigating parameterization SHAP values**
413   While the Gain values indicate the relative importance of species in the parameterization and can
414 provide some information as to whether the parameterization behaves in a manner consistent with our
415 understanding of OH chemistry, the metric only provides information about the dataset as a whole.
416 Gain values can therefore obscure the importance of variables that only strongly impact the
417 parameterization for a small subset of the data. To better understand the relative importance of
418 variables as well as the spatial variability in that importance, we also calculate the SHAP values
419 (Lundberg and Lee, 2017), which provide information on the relative importance of each datapoint input
420 into the model.
421
422   In the context of machine learning, Shapley values, an idea first developed for game theory
423 (Shapley, 1953), indicate the average contribution of an individual model input to all possible
424 combinations of inputs. For example, to calculate the Shapley value of the variable X for a hypothetical
425 machine learning model with three input variables X, Y, and Z, first a model would be trained with all
426 three variables. A new model would then be retrained, omitting X, and the difference between the two
427 models would be calculated to determine the contribution of X. This process would then be repeated
428 with different permutations of input variables (e.g., X and Y, X and Z) to determine the contribution of X
429 in those instances. The final Shapley value is the average of the contribution from these different
430 models. SHAP values use the same concept but in a manner that reduces the computation time
431 (Lundberg and Lee, 2017), as this process could become prohibitive for a model, such as the
432 parameterization of OH, that contains 27 inputs.
433
434   We calculate SHAP values using the TreeExplainer API of the SHAP package available for Python.
435 One limitation of the algorithm used to calculate SHAP values is that it is too computationally expensive
436 to calculate the SHAP values for the tuned regression tree model. Computational time to calculate SHAP
437 values for the troposphere at the native model resolution for one day is several months. Maximizing
438 computational speed by degrading the model resolution and running the SHAP package with GPUs,
439 would take approximately 4 days for one model day. Calculating SHAP values for a model with default
440 model hyperparameters, however, takes minutes. This is due to the large reduction in the number of
441 trees (100 to 10) and the maximum tree depth (18 to 6) in the parameterization, which significantly
442 speeds up the creation of new regression trees needed in the SHAP value calculation.
443
444   We first evaluate the feasibility of using the SHAP values for the untuned model to explain the
445 parameterization behavior. To test this, we created a subset of 5000 OH values from the
446 parameterization training dataset that spanned the full range of OH concentrations. We then calculated
447 the SHAP values for the July parameterization with tuned hyperparameters as well as for a July
448 parameterization using the default XGBoost hyperparameters. For the variables found to be most
449 important for the parameterization (e.g., SZA, NO_2, O_3, isoprene, HCHO, latitude), there are strong
450 correlations (r^2 of 0.97 or higher) for the SHAP values between the tuned and untuned model, resulting
451 in similar spatial distributions, although there are differences in the magnitude.   For other variables,
452 correlation is much weaker, although the relative importance of variables is similar for the tuned and
453 untuned parameterizations. We therefore restrict our analysis primarily to variables that have highly

454 correlated SHAP values between the tuned and untuned models and discussion to the relative
455 importance of the different variables.
456
457   The distribution of SHAP values for the training dataset for July demonstrates the importance of
458 including each of the variables as inputs to the parameterization as well as the large variability in their
459 relative importance. Figure 6 shows the distribution of the SHAP values for each input parameter of the
460 approximately 3.2 million datapoints used to train the July parameterization. The median SHAP values
461 (Fig. 6) show similar ordering as the Gains feature importance (Fig. 5), with variables that indicate
462 location as well as O_3 and NO_2 being the most important in both cases. When looking at the distribution
463 of the SHAP values, however, it becomes apparent that variables that appear to be unimportant for
464 parameterization performance in the mean (e.g., propene and CH_4) can have large importance for
465 individual datapoints. For example, although propene can be locally important for OH chemistry, due to
466 its reactivity, concentrations in the remote atmosphere are low, making the species seem unimportant
467 in the aggregate. Similar results are found for the January parameterization (Fig. S6). As discussed in
468 Section 2.2, the SHAP values provide a rationale for including each of these species in the
469 parameterization.
470

471
472 ***Figure 6:*** Distribution of the absolute SHAP value for each parameterization input for July from an untuned version
473 of the parameterization of OH. Input parameters are sorted by order of relative importance. The median is indicated
474 with the black line, edges of the box represent the interquartile range, and whiskers represent the 5th and 95th
475 percentile. Values outside this range are indicated with circles. Note that the SHAP value for propane is zero,
476 indicating that it is not used by the untuned parameterization.

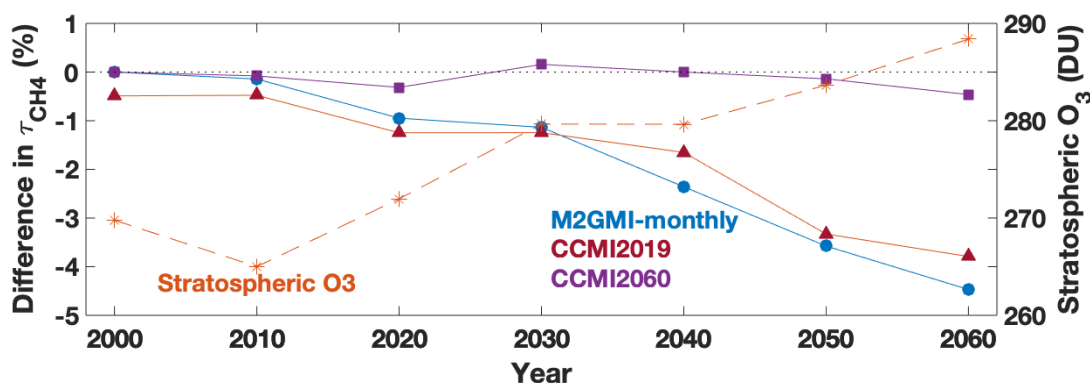
477
478   The SHAP values also demonstrate the spatial distribution of the relative importance of the
479 different chemical OH drivers. Figure 7 shows the relative importance of NO_2, as determined by the
480 SHAP values for the untuned parameterization, for both the zonal mean and the LFT. Note that the
481 untuned parameterization has large errors for low OH concentrations, so SHAP values poleward of 45 °S
482 should be viewed as more uncertain than those elsewhere. In both the horizontal and vertical, the SHAP
483 values demonstrate that the parameterization captures the spatial pattern of the relative importance of
484 NO_X for OH production. The spatial pattern in Figure7a, for example, has the highest contribution of
485 NO_2 to the total SHAP value in the tropical UFT and in the northern hemisphere midlatitudes. This is
486 nearly identical to the spatial pattern of the relative contribution of the $NO + HO_2$ reaction to overall OH

487 production in the MERRA2 GMI simulation (Anderson et al., 2021). Likewise, in the LFT, the contribution
488 from NO$_2$ maximizes over continental regions with high emission and minimizes over the remote oceans.
489 The spatial pattern of SHAP values of isoprene also agree with OH chemistry, maximizing in regions of
490 strong biogenic emissions and minimizing over oceans (Fig. S7). These SHAP values demonstrate that,
491 although the parameterization is not process-based, its behavior at least partially accords with our
492 understanding of OH chemistry.
493

494
495 **Figure 7:** The fraction of the contribution of the NO$_2$ SHAP value to the sum of the absolute SHAP value of all inputs
496 in July is shown for the zonal mean (a) and the LFT (b). Regions where mean OH mixing ratios are below 0.03 pptv,
497 the point below which the untuned parameterization is unable to reasonably predict OH, are indicated by the
498 *stippling*.
499
500 **3.3 Case Study: Testing the parameterization response to the 2016 El Niño Event**
501   While we have demonstrated that the parameterization can satisfactorily reproduce OH during all
502 months of 2005, we now investigate its ability to capture OH accurately during the 2016 El Niño event.
503 which we excluded from the training dataset. Evaluating how the parameterization responds to
504 deviations from the climatological mean of the inputs during a large-scale event on which it was not
505 trained, such as the 2016 El Niño, is a strong test of its ability to predict extremes in OH as well as to
506 respond to deviations from the climatological mean of the parameterization inputs. The response of the
507 parameterized OH to these extremes in inputs will also provide a further test of the ability of the
508 parameterization to behave in a process-based way.
509
510   El Niño events lead to dramatic changes in the concentrations and distributions of many OH
511 drivers, including O$_3$ (Oman et al., 2011;Oman et al., 2013), CO (Duncan, 2003a;Rowlinson et al., 2019),
512 NO$_X$ (Murray et al., 2013;Murray et al., 2014) and water vapor (Shi et al., 2018;Anderson et al., 2021).
513 As such, the El Niño Southern Oscillation (ENSO) is the dominant mode of OH variability throughout
514 much of the troposphere and can result in localized changes in OH on the order of 40 – 50% from the
515 climatological mean (Anderson et al., 2021;Turner et al., 2018). Changes in secondary production from
516 NO$_X$ in the UFT and changes in primary production from O$_3$ in the PBL and LFT drive the ENSO related
517 variability of OH (Anderson et al., 2021).   Methane emissions also vary strongly with the ENSO phase
518 (Zhang et al., 2018;Worden et al., 2013). In order to capture the OH/CH$_4$/CO system correctly, the
519 parameterization must be able to capture ENSO-related OH variability.
520
521   Here, we investigate the ability of the parameterization to capture OH during the El Niño events of
522 1997/98 and 2015/16, two of the largest such events during the period of the MERRA2 GMI simulation

523   according to the Multivariate ENSO Index (Wolter and Timlin, 2011). The 1997/98 event, which was
524   included in the training dataset, was a prototypical example of an Eastern Pacific (EP) El Niño,
525   characterized by sea surface temperature (SST) anomalies extending to the coast of South America. In
526   contrast, the 2015/16 event was a blend of an EP and a Central Pacific (CP) El Niño, also known as El
527   Niño Modoki, where SST anomalies extend only to the international dateline (Paek et al., 2017).   These
528   different "flavors" of El Niño affect atmospheric distributions of OH drivers, such as water vapor (Du et
529   al., 2021), in different ways, suggesting different impacts on OH. While we did include other blended El
530   Niños (1986/87, 1987/88, and 1991/92) (Kug et al., 2009) in the training dataset, each had responses in
531   the atmospheric distribution of OH and its drivers distinct from the 2015/16 event. We focus our
532   investigation on January and the MFT, the time and location of the strongest correlation between ENSO
533   and OH (Anderson et al., 2021) in the MERRA2 GMI simulation. We also restrict the analysis to the OH
534   precursors, NO_2, CO, and O_3, as they have both a strong influence in the variability of ENSO-related OH
535   production/loss and have comparatively large feature importance and SHAP values in the January
536   parameterization.
537
538   For both the 1997/98 and 2015/16 El Niño events, each OH driver examined deviates strongly
539   from the climatological mean, defined as the average value from the MERRA2 GMI simulation over all
540   Januarys from 1980 – 2019. Both O_3 and NO_2 have pronounced positive anomalies over the western
541   Pacific and maritime continent and negative anomalies over the eastern Pacific (Fig. 8) that extend
542   throughout much of the free troposphere (Fig. S8), likely associated with changes in the Walker
543   Circulation as described in Oman et al. (2011).   The positive anomalies over Indonesia show a distinct
544   westward shift during the 1997/98 event as compared to 2015/16, highlighting the variability in the
545   effects of ENSO on emissions and transport. CO has a large positive anomaly over much of the globe,
546   attributable to the increases in biomass burning during El Niño events (e.g., Duncan, 2003a). As with O_3
547   and NO_2, there are large differences in the spatial pattern of the CO anomalies between the two events,
548   particularly over the Indian Ocean, central Africa, and South America.
549

550
551   **Figure 8:** Fractional difference of the indicated variable between January 1998 (left) and the climatological mean
552   (1980 – 2019) calculated from the MERRA2 GMI simulation for the MFT. The same values but for January 2016 are
553   indicated on the right. Species shown are CO (a,b), NO_2 (c,d) and O_3 (e,f).
554
555   The differences in anomalies of the OH drivers between the 1997/98 and 2015/16 El Niño events
556   lead to distinct anomaly patterns in OH itself. During the 1997/98 event, in the MFT, there are

557 noticeable positive anomalies in OH over much of the Indian Ocean basin, the southeastern Pacific,
558 South America, and the western Atlantic Ocean (Fig. 9). During 2015/16, the positive anomalies were
559 more limited and most noticeable in the tropical western Pacific Ocean and southern Indian Ocean.
560 Along the equator, the positive anomalies extended throughout a larger portion of the troposphere
561 during January 1998 than 2016. Both the parameterization inputs and the OH itself respond strongly
562 and in different ways to each El Niño event, providing a strong test to determine the robustness of the
563 parameterization.
564   The parameterization reproduces the ENSO-related OH anomalies for both El Niño events with
565 remarkable fidelity. We ran the parameterization for all Januarys from 1980 to 2016 to calculate a
566 climatology and calculated the deviations for 1998 and 2016 from that value. For both events, the
567 parameterization accurately captures the location and magnitude, as well as the spatial pattern, of the
568 OH anomalies with a few minor exceptions in the horizontal and vertical (Figs. 9 and S8). Correlation
569 between the MERRA2 GMI and parameterized anomalies plotted in Figure 9 has an r^2 of 0.93 or higher
570 for both years. The parameterization is therefore capable of reproducing both the climatological mean
571 of OH as well as large deviations in the mean in response to strong climatological deviations in the
572 model inputs, even for years excluded from the training dataset.
573

574
575 ***Figure 9:*** Fractional difference of the indicated variable between January 1998 (left) and the climatological mean
576 (1980 – 2019) calculated from the MERRA2 GMI simulation averaged over the MFT. The same values but for January
577 2016 are indicated on the right. Species shown are OH from the MERRA2 GMI simulation (a,b) and OH calculated by
578 the parameterization (c,d).

579 **4.0 Discussion and Summary: The parameterization of OH as a tool for scientific research**
580   In this manuscript, we present a new methodology to generate a parameterization of OH that, as
581 compared to previous methods (e.g., Spivakovsky et al., 1990; Duncan et al., 2000), is efficient and easy
582 to use, allowing the user to rapidly update the parameterization of OH as necessary. The new method
583 uses GBRTs and a full-chemistry simulation from a CCM as the training data to generate the
584 parameterization of OH with a high degree of accuracy relative to the full-chemistry simulation. We
585 illustrated our methodology with a parameterization designed for the ECCOH module of the GEOS CCM.
586
587   The parameterization of OH accurately reproduces OH from the full-chemistry simulation on
588 which it was trained, but it may not produce the desired accuracy for a given time period or scenario
589 outside the range represented in the training data. Of course, the degree of degradation in accuracy
590 depends on how far inputs exceed the ranges of the training dataset. In addition, the parameterization
591 of OH generated using inputs from one model may not be portable to another model or even a different

592   configuration of the same model as shown below. The simulated relationships between OH and the
593   input parameters may differ because of inter-model variations in the chemical, dynamical, and radiative
594   schemes. Ultimately, it is up to the user to determine an acceptable level of degradation for a specific
595   research topic. In this section, we give an example of the degree of degradation in accuracy for a
596   parameterization of OH that uses 1) a different time period for the same model and 2) input parameters
597   from a different model.
598
599   **4.1: Input parameters from a different time period for the same model setup**
600   Analysis of a separate model simulation, the Chemistry Climate Model Initiative (CCMI) GEOS
601   simulation (Morgenstern et al., 2017), highlights possible limitations in extending the parameterization
602   to years outside of those on which it was trained, particularly if there is a strong trend in one of the
603   inputs. The GEOS CCMI simulation has unconstrained meteorology, spans 1960 – 2100, and has a
604   horizontal resolution of 2.0° latitude × 2.5° longitude. Emissions for precursors of tropospheric O_3 and
605   aerosols are from the RCP6.0 scenario. We trained two new parameterizations on the CCMI dataset,
606   denoted CCMI2019 and CCMI2060, using data from 1980 – 2019 and 1980 – 2061, respectively. We
607   used the same methodology to create the training datasets as for the MERRA2 GMI parameterization.
608   CCMI output are only available at monthly resolution, so we trained the CCMI parameterizations on
609   monthly, instead of daily, averaged values. Every 10th year, staring in 2000, was omitted from the
610   training dataset for validation.
611
612   While the CCMI2019 parameterization performed similarly to the MERRA2 GMI
613   parameterization for years included in the training dataset, performance degraded significantly for years
614   beyond 2019. The CCMI2019 parameterization captured the τ_{CH4} for 2000 and 2010 within 1% (Fig. 10,
615   red line) and the NRMSE within 5% (not shown). When we applied the parameterization to years
616   outside of the training window, however, performance degraded quickly and, by 2060, underestimated
617   τ_{CH4} by about 4%. The CCMI2060 parameterization, on the other hand, captures the τ_{CH4} lifetime within
618   0.5% for all validation years.
619
620   The reason for this performance degradation is likely due to input parameters that extend
621   beyond the range used in the training dataset.   For example, there is a strong positive trend in the
622   stratospheric O_3 column (Fig. 10), which results in chemical environments in 2060 that did not exist in
623   the 1980 – 2019 training dataset. Other variables with strong trends, such as CH_4 and temperature, as
624   well as different large-scale dynamical patterns, could also decrease parameterization performance.
625   These results strongly suggest caution when applying the parameterization to future scenarios outside
626   of the training window. For example, experiments investigating relative changes in a species (e.g.,
627   understanding how a 10% decrease in OH would affect tropospheric CH_4 abundance) could likely be
628   successfully implemented, while those trying to understand absolute differences in CH_4 would require
629   significant care to ensure valid results.
630

631
632 ***Figure 10:*** *Time series showing the percent difference between τ_{CH4} calculated from the CCMI simulation and from*
633 *three separate parameterizations: one trained on MERRA2 GMI output from 1980 – 2019 (blue circle), one trained*
634 *on CCMI output spanning 1980 – 2019 (red triangle), and one trained on CCMI output spanning 1980 – 2060 (purple*
635 *square). The stratospheric O_3 column (orange star) from the CCMI simulation averaged over 30 S to 60 N, the region*
636 *where tropospheric CH_4 loss to OH maximizes (Fig. 3), is also shown. All data are for July.*

637 **4.2 Input parameters from a different model setup**
638   Similar to applying the parameterization outside of the timeframe on which it was trained,
639 applying the parameterization to a different model setup also warrants caution, as the differences can
640 result in new chemical environments on which the parameterization was not trained. We now discuss
641 parameterization performance when output from the CCMI simulation discussed in Section 4.1 is input
642 into the MERRA2 GMI-trained parameterization. Despite both simulations being from the GEOS
643 framework, the CCMI simulation setup differs from the MERRA2 GMI simulation in emissions, time
644 frame, resolution, and meteorology (unconstrained vs specified dynamics), among others. Because
645 CCMI output is only available at a monthly resolution, we created a separate parameterization,
646 hereafter referred to as "M2GMI-monthly", using MERRA2 GMI output with identical parameterization
647 inputs as the daily parameterization but using monthly-averaged values. Performance is similar to that
648 of the parameterization trained on daily data and averaged over monthly timescales, with the NRMSE
649 for the troposphere on the order of 6 -7% depending on the year.
650
651   When output from the CCMI simulation is used as inputs to the M2GMI-monthly
652 parameterization, performance degrades significantly. For July 2000, for example, there are distinct
653 regions of both positive and negative biases (Fig. 11a) in parameterized OH, resulting in a NRMSE of
654 13%, on par with using climatology as an OH estimate. Because the largest discrepancies are centered
655 outside of the tropics and/or in regions with low concentrations, τ_{CH4} for year 2000 is identical between
656 the CCMI and parameterized OH. When applied to 2060 (Fig. 11c), which is far outside the training
657 period of the M2GMI-monthly parameterization, there is a near universal high bias in parameterized OH,
658 resulting in a NRMSE of 16% and a τ_{CH4} biased low by 4.5%. This overestimate results in a negative trend
659 in τ_{CH4} from parameterized OH from 2000 to 2060 (Fig. 10, blue line), despite the trend in the CCMI
660 simulation being positive. Applying the MERRA2 GMI parameterization to a study using the CCMI setup
661 would therefore misrepresent the OH/CH_4 cycle.
662

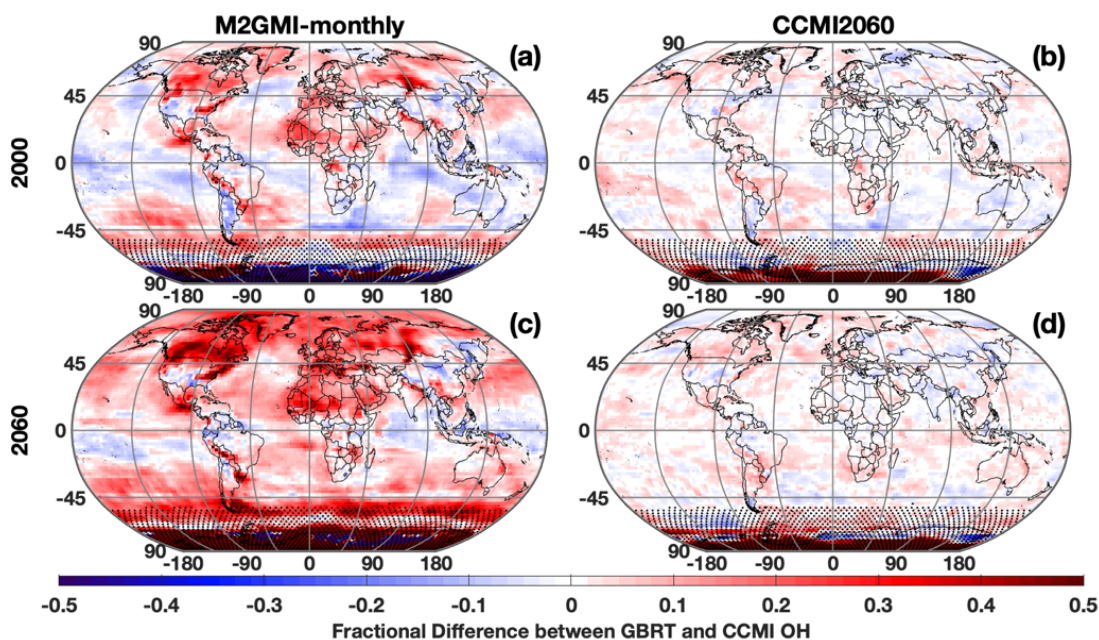



Figure 11: *Fractional difference between OH calculated by the M2GMI-monthly (left) and the CCMI2060 (right) parameterizations and OH output from the CCMI simulation. Data are averaged between 850 and 500 hPa for July 2000 (top) and July 2060 (bottom). Regions with low OH, defined as a mixing ratio of less than 0.005 pptv, are indicated with stippling.*

Through an analysis of the ranges of input parameters from both simulations, we found that the differences in parameterization performance for inputs from MERRA2 GMI and CCMI are likely largely attributable to differences in the stratospheric O_3 distributions between the two simulations. In 2060, for example, CCMI stratospheric O_3 has a much higher frequency of values above 300 DU than the M2GMI-monthly training dataset (Fig. 12). A smaller, but still noticeable, shift in the distribution is also found for the year 2000 (Fig. S9). Likewise, the accuracy of the M2GMI-monthly parameterization decreases from 2000 – 2060 as the stratospheric O_3 burden increases (Fig. 10, red line). Mechanistically, higher stratospheric ozone in CCMI should result in lower tropospheric OH because the reduction in incoming ultraviolet radiation limits tropospheric O_3 photolysis. This could also lead to a higher CO burden, due to the smaller OH sink. Comparisons between the OH and CO distributions from the two simulations are consistent with this hypothesis. Even though the M2GMI-monthly training dataset spanned the full range of stratospheric O_3 values from the CCMI simulation, the frequency of stratospheric O_3 values at higher concentration likely creates chemical environments in the CCMI simulation distinct from those in MERRA2 GMI, forcing the parameterization to extrapolate to a chemical space on which it was not trained. This highlights the need to compare the distribution of any parameterization inputs to that of the training dataset.

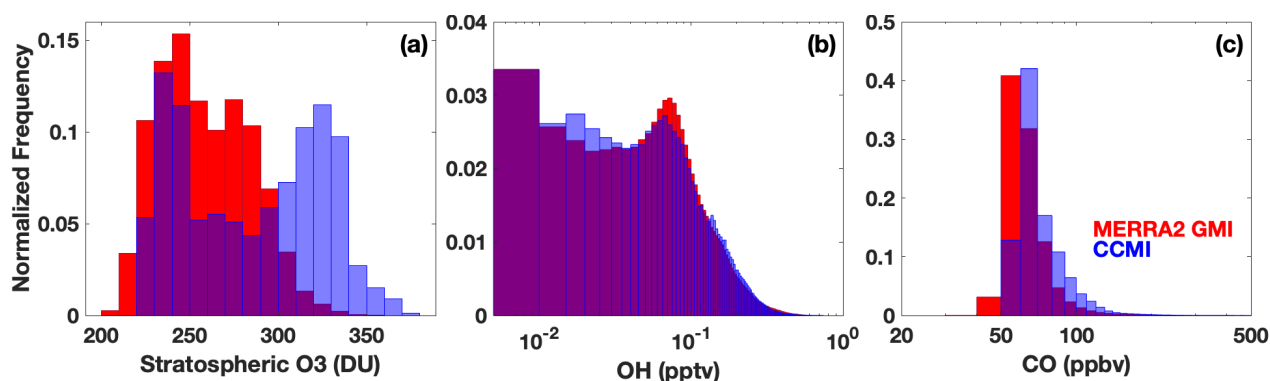



Figure 12: *Histograms showing the distribution of stratospheric column O₃ (a), tropospheric OH (b) and tropospheric CO (c) from the M2GMI-monthly parameterization training dataset (red) and from the CCMI simulation for 2060 (blue). Purple indicates areas of overlap between the two distributions.*

Again, performance improves significantly when we apply output from the CCMI simulation to the CCMI2060 parameterization. The regions of consistent high bias notable when CCMI output was applied to the M2GMI-monthly parameterization are absent for both 2000 and 2060, and NRMSE shows a factor of three improvement over the previously discussed scenario. Likewise, for all validation years, the parameterized OH resulted in a τ_{CH4} that agreed with the CCMI simulation between 0 and -0.46% (purple line, Fig. 10).

We conclude that for best performance, a separate parameterization should be created for each new modeling framework to capture OH variability accurately. This will not create an undue computational expense on an experiment. Because a full chemistry simulation is necessary to create the parameterization inputs of chemical species that are not calculated online, the necessary data to create a training dataset will already be available. The only additional time will be that required to format the regression tree inputs and to train the model, which takes approximately 2 – 3 hours for each month. This process can be performed using previously created python scripts with minimal changes. The flexibility that this modeling framework permits will facilitate its use even if there are major changes to the underlying model chemistry or dynamics.

5.0 Code Availability
The scripts used to create the training datasets and a sample script to create a parameterization have been archived by Zenodo at https://doi.org/10.5281/zenodo.6046037 (Anderson, 2022a). A sample parameterization for the ECCOH module trained on MERRA2-GMI output is available at https://doi.org/10.5281/zenodo.6604130 (Anderson, 2022b).

6.0 Data Availability
Output from the MERRA2 GMI simulation are publicly available at https://acd-ext.gsfc.nasa.gov/Projects/GEOSCCM/MERRA2GMI/. The training dataset and training targets for the July parameterization presented here are available at https://doi.org/10.5281/zenodo.6604130 (Anderson, 2022b). Output from the GEOSCCM simulation for CCMI is available at the Centre for Environmental Data Analysis (CED), the Natural Environment Research Council's Data Repository for Atmospheric Science and Earth Observation, at http://data.ceda.ac.uk/badc/wcrp-ccmi/data/CCMI-1/output.

7.0 Author Contributions

722 DCA wrote the manuscript, performed the data analysis, and created the parameterizations. BND and
723 MBFC developed the idea for the parameterization. SAS performed three-dimensional modeling for the
724 work. JMN and PDI provided advice on machine learning. All authors helped develop ideas for the
725 analysis and contributed to the manuscript,
726

727 **8.0 Competing Interests**
728 The authors declare that they have no conflict of interest.
729
730 **9.0 Acknowledgements**
731 The authors acknowledge funding from the NASA MAP program (grant no. 80NSSC17K0220). In
732 addition, the authors acknowledge funding from the NASA Aura program.

733
734 **10. References**
735

736 Anderson, D. C., Duncan, B. N., Fiore, A. M., Baublitz, C. B., Follette-Cook, M. B., Nicely, J. M., and Wolfe,
737 G. M.: Spatial and temporal variability in the hydroxyl (OH) radical: understanding the role of large-scale
738 climate features and their influence on OH through its dynamical and photochemical drivers,
739 Atmospheric Chemistry and Physics, 21, 6481-6508, 10.5194/acp-21-6481-2021, 2021.

740 Anderson, D.C., Follette-Cook, M.B., Strode, S.A., Nicely, J.M., Liu, J., Ivatt, P.D., Duncan, B.N. (2022a).
741 Code for the development of.a parameterization of OH for CCMs, Zenodo[code],
742 https://doi.org/10.5281/zenodo.6046037.

743 Anderson, D.C., Follette-Cook, M.B., Strode, S.A., Nicely, J.M., Liu, J., Ivatt, P.D., & Duncan, B.N. (2022b)
744 Sample ECCOH OH Parameterization (1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.6604130.

745 Chen, T., and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, KDD '16: Proceedings of the 22nd
746 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, 785-794.

747 Chen, Y.-H., and Prinn, R. G.: Estimation of atmospheric methane emissions between 1996 and 2001
748 using a three-dimensional global chemical transport model, Journal of Geophysical Research:
749 Atmospheres, 111, https://doi.org/10.1029/2005JD006058, 2006.

750 Chin, M., Ginoux, P., Kinne, S., Torres, O., Holben, B. N., Duncan, B. N., Martin, R. V., Logan, J. A.,
751 Higurashi, A., and Nakajima, T.: Tropospheric Aerosol Optical Thickness from the GOCART Model and
752 Comparisons with Satellite and Sun Photometer Measurements, Journal of the Atmospheric Sciences,
753 59, 461-483, 10.1175/1520-0469(2002)059<0461:TAOTFT>2.0.CO;2, 2002.

754 Colarco, P., da Silva, A., Chin, M., and Diehl, T.: Online simulations of global aerosol distributions in the
755 NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth, Journal of
756 Geophysical Research: Atmospheres, 115, https://doi.org/10.1029/2009JD012820, 2010.

757 Du, M., Huang, K., Zhang, S., Huang, C., Gong, Y., and Yi, F.: Water vapor anomaly over the tropical
758 western Pacific in El Niño winters from radiosonde and satellite observations and ERA5 reanalysis data,
759 Atmos. Chem. Phys., 21, 13553-13569, 10.5194/acp-21-13553-2021, 2021.

760 Duncan, B., Portman, D., Bey, I., and Spivakovsky, C.: Parameterization of OH for efficient computation in
761 chemical tracer models, Journal of Geophysical Research: Atmospheres, 105, 12259-12262,
762 10.1029/1999JD901141, 2000.

Duncan, B. N.: Indonesian wildfires of 1997: Impact on tropospheric chemistry, Journal of Geophysical Research, 108, 10.1029/2002jd003195, 2003a.

Duncan, B. N.: Interannual and seasonal variability of biomass burning emissions constrained by satellite observations, Journal of Geophysical Research, 108, 10.1029/2002jd002378, 2003b.

Duncan, B. N., Logan, J. A., Bey, I., Megretskaia, I. A., Yantosca, R. M., Novelli, P. C., Jones, N. B., and Rinsland, C. P.: Global budget of CO, 1988–1997: Source estimates and validation with a global model, Journal of Geophysical Research, 112, 10.1029/2007jd008459, 2007a.

Duncan, B. N., Strahan, S. E., Yoshida, Y., Steenrod, S. D., and Livesey, N.: Model study of the cross-tropopause transport of biomass burning pollution, Atmos. Chem. Phys., 7, 3713-3736, 10.5194/acp-7-3713-2007, 2007b.

Duncan, B. N., and Logan, J. A.: Model analysis of the factors regulating the trends and variability of carbon monoxide between 1988 and 1997, Atmospheric Chemistry and Physics, 8, 2008.

Elith, J., Leathwick, J. R., and Hastie, T.: A working guide to boosted regression trees, J Anim Ecol, 77, 802-813, 10.1111/j.1365-2656.2008.01390.x, 2008.

Elshorbany, Y. F., Duncan, B. N., Strode, S. A., Wang, J. S., and Kouatchou, J.: The description and validation of the computationally Efficient CH4–CO–OH (ECCOHv1.01) chemistry module for 3-D model applications, Geoscientific Model Development, 9, 799-822, 10.5194/gmd-9-799-2016, 2016.

Fiore, A. M., Horowitz, L. W., Dlugokencky, E. J., and West, J. J.: Impact of meteorology and emissions on methane trends, 1990–2004, Geophysical Research Letters, 33, 10.1029/2006gl026199, 2006.

Gaubert, B., Worden, H. M., Arellano, A. F. J., Emmons, L. K., Tilmes, S., Barré, J., Martinez Alonso, S., Vitt, F., Anderson, J. L., Alkemade, F., Houweling, S., and Edwards, D. P.: Chemical Feedback From Decreasing Carbon Monoxide Emissions, Geophysical Research Letters, 44, 9985-9995, https://doi.org/10.1002/2017GL074987, 2017.

Gelaro, R., McCarty, W., Suarez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A., Gu, W., Kim, G. K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), J Clim, Volume 30, 5419-5454, 10.1175/JCLI-D-16-0758.1, 2017.

Holmes, C. D.: Methane Feedback on Atmospheric Chemistry: Methods, Models, and Mechanisms, Journal of Advances in Modeling Earth Systems, 10, 1087-1099, https://doi.org/10.1002/2017MS001196, 2018.

Ivatt, P. D., and Evans, M. J.: Improving the prediction of an atmospheric chemistry transport model using gradient-boosted regression trees, Atmospheric Chemistry and Physics, 20, 8063-8082, 10.5194/acp-20-8063-2020, 2020.

798   Keller, C. A., and Evans, M. J.: Application of random forest regression to the calculation of gas-phase
799   chemistry within the GEOS-Chem chemistry model v10, Geoscientific Model Development, 12, 1209-
800   1225, 10.5194/gmd-12-1209-2019, 2019.

801   Kelp, M. M., Jacob, D. J., Kutz, J. N., Marshall, J. D., and Tessum, C. W.: Toward Stable, General Machine-
802   Learned Models of the Atmospheric Chemical System, Journal of Geophysical Research: Atmospheres,
803   125, e2020JD032759, https://doi.org/10.1029/2020JD032759, 2020.

804   Kleipool, Q. L., Dobber, M. R., de Haan, J. F., and Levelt, P. F.: Earth surface reflectance climatology from
805   3 years of OMI data, Journal of Geophysical Research: Atmospheres, 113,
806   https://doi.org/10.1029/2008JD010290, 2008.

807   Kug, J.-S., Jin, F.-F., and An, S.-I.: Two Types of El Niño Events: Cold Tongue El Niño and Warm Pool El
808   Niño, Journal of Climate, 22, 1499-1515, 10.1175/2008JCLI2624.1, 2009.

809   Laughner, J. L., Neu, J. L., Schimel, D., Wennberg, P. O., Barsanti, K., Bowman, K. W., Chatterjee, A.,
810   Croes, B. E., Fitzmaurice, H. L., Henze, D. K., Kim, J., Kort, E. A., Liu, Z., Miyazaki, K., Turner, A. J.,
811   Anenberg, S., Avise, J., Cao, H., Crisp, D., de Gouw, J., Eldering, A., Fyfe, J. C., Goldberg, D. L., Gurney, K.
812   R., Hasheminassab, S., Hopkins, F., Ivey, C. E., Jones, D. B. A., Liu, J., Lovenduski, N. S., Martin, R. V.,
813   McKinley, G. A., Ott, L., Poulter, B., Ru, M., Sander, S. P., Swart, N., Yung, Y. L., and Zeng, Z. C.: Societal
814   shifts due to COVID-19 reveal large-scale complexities and feedbacks between atmospheric chemistry
815   and climate change, Proc Natl Acad Sci U S A, 118, 10.1073/pnas.2109481118, 2021.

816   Lawrence, M. G., Jöckel, P., and von Kuhlmann, R.: What does the global mean OH concentration tell
817   us?, Atmos. Chem. Phys., 1, 37-49, 10.5194/acp-1-37-2001, 2001.

818   Lundberg, S. M., and Lee, S.-I.: A Unified Approach to Interpreting Model Predictions, 31st Conference
819   on Neural Information Processing Systems, Long Beach, CA, USA, 2017.

820   Morgenstern, O., Hegglin, M. I., Rozanov, E., Connor, F. M., Abraham, N. L., Akiyoshi, H., Archibald, A. T.,
821   Bekki, S., Butchart, N., Chipperfield, M. P., Deushi, M., Dhomse, S. S., Garcia, R. R., Hardiman, S. C.,
822   Horowitz, L. W., Jöckel, P., Josse, B., Kinnison, D., Lin, M., Mancini, E., Manyin, M. E., Marchand, M.,
823   Marécal, V., Michou, M., Oman, L. D., Pitari, G., Plummer, D. A., Revell, L. E., Saint-Martin, D., Schofield,
824   R., Stenke, A., Stone, K., Sudo, K., Tanaka, T. Y., Tilmes, S., Yamashita, Y., Yoshida, K., and Zeng, G.:
825   Review of the global models used within phase 1 of the Chemistry–Climate Model Initiative (CCMI),
826   Geoscientific Model Development, 10, 639-671, 10.5194/gmd-10-639-2017, 2017.

827   Murray, L. T., Logan, J. A., and Jacob, D. J.: Interannual variability in tropical tropospheric ozone and OH:
828   The role of lightning, Journal of Geophysical Research: Atmospheres, 118, 11,468-411,480,
829   10.1002/jgrd.50857, 2013.

830   Murray, L. T., Mickley, L. J., Kaplan, J. O., Sofen, E. D., Pfeiffer, M., and Alexander, B.: Factors controlling
831   variability in the oxidative capacity of the troposphere since the Last Glacial Maximum, Atmospheric
832   Chemistry and Physics, 14, 3589-3622, 10.5194/acp-14-3589-2014, 2014.

833   Murray, L. T., Fiore, A. M., Shindell, D. T., Naik, V., and Horowitz, L. W.: Large uncertainties in global
834   hydroxyl projections tied to fate of reactive nitrogen and carbon, Proceedings of the National Academy
835   of Sciences, 118, e2115204118, 10.1073/pnas.2115204118, 2021.

836 Nicely, J. M., Salawitch, R. J., Canty, T., Anderson, D. C., Arnold, S. R., Chipperfield, M. P., Emmons, L. K.,
837 Flemming, J., Huijnen, V., Kinnison, D. E., Lamarque, J.-F., Mao, J., Monks, S. A., Steenrod, S. D., Tilmes,
838 S., and Turquety, S.: Quantifying the causes of differences in tropospheric OH within global models,
839 Journal of Geophysical Research: Atmospheres, n/a-n/a, 10.1002/2016JD026239, 2017.

840 Nicely, J. M., Duncan, B. N., Hanisco, T. F., Wolfe, G. M., Salawitch, R. J., Deushi, M., Haslerud, A. S.,
841 Jöckel, P., Josse, B., Kinnison, D. E., Klekociuk, A., Manyin, M. E., Marécal, V., Morgenstern, O., Murray, L.
842 T., Myhre, G., Oman, L. D., Pitari, G., Pozzer, A., Quaglia, I., Revell, L. E., Rozanov, E., Stenke, A., Stone, K.,
843 Strahan, S., Tilmes, S., Tost, H., Westervelt, D. M., and Zeng, G.: A machine learning examination of
844 hydroxyl radical differences among model simulations for CCMI-1, Atmospheric Chemistry and Physics,
845 20, 1341-1361, 10.5194/acp-20-1341-2020, 2020.

846 Nowack, P., Braesicke, P., Haigh, J., Abraham, N. L., Pyle, J., and Voulgarakis, A.: Using machine learning
847 to build temperature-based ozone parameterizations for climate sensitivity simulations, Environmental
848 Research Letters, 13, 104016, 10.1088/1748-9326/aae2be, 2018.

849 Nowack, P., Konstantinovskiy, L., Gardiner, H., and Cant, J.: Machine learning calibration of low-cost NO2
850 and PM10 sensors: non-linear algorithms and their impact on site transferability, Atmos. Meas. Tech.,
851 14, 5637-5655, 10.5194/amt-14-5637-2021, 2021.

852 Oman, L. D., Ziemke, J. R., Douglass, A. R., Waugh, D. W., Lang, C., Rodriguez, J. M., and Nielsen, J. E.: The
853 response of tropical tropospheric ozone to ENSO, Geophysical Research Letters, 38, n/a-n/a,
854 10.1029/2011gl047865, 2011.

855 Oman, L. D., Douglass, A. R., Ziemke, J. R., Rodriguez, J. M., Waugh, D. W., and Nielsen, J. E.: The ozone
856 response to ENSO in Aura satellite measurements and a chemistry-climate simulation, Journal of
857 Geophysical Research-Atmospheres, 118, 965-976, 10.1029/2012jd018546, 2013.

858 Orbe, C., Oman, L. D., Strahan, S. E., Waugh, D. W., Pawson, S., Takacs, L. L., and Molod, A. M.: Large-
859 Scale Atmospheric Transport in GEOS Replay Simulations, Journal of Advances in Modeling Earth
860 Systems, 9, 2545-2560, 10.1002/2017ms001053, 2017.

861 Paek, H., Yu, J.-Y., and Qian, C.: Why were the 2015/2016 and 1997/1998 extreme El Niños different?,
862 Geophysical Research Letters, 44, 1848-1856, https://doi.org/10.1002/2016GL071515, 2017.

863 Prather, M. J.: Time scales in atmospheric chemistry: Theory, GWPs for CH4 and CO, and runaway
864 growth, Geophysical Research Letters, 23, 2597-2600, https://doi.org/10.1029/96GL02371, 1996.

865 Rigby, M., Montzka, S. A., Prinn, R. G., White, J. W. C., Young, D., O'Doherty, S., Lunt, M. F., Ganesan, A.
866 L., Manning, A. J., Simmonds, P. G., Salameh, P. K., Harth, C. M., Muhle, J., Weiss, R. F., Fraser, P. J.,
867 Steele, L. P., Krummel, P. B., McCulloch, A., and Park, S.: Role of atmospheric oxidation in recent
868 methane growth, Proc Natl Acad Sci U S A, 114, 5373-5377, 10.1073/pnas.1616426114, 2017.

869 Rowlinson, M. J., Rap, A., Arnold, S. R., Pope, R. J., Chipperfield, M. P., McNorton, J., Forster, P., Gordon,
870 H., Pringle, K. J., Feng, W., Kerridge, B. J., Latter, B. L., and Siddans, R.: Impact of El Niño–Southern
871 Oscillation on the interannual variability of methane and tropospheric ozone, Atmospheric Chemistry
872 and Physics, 19, 8669-8686, 10.5194/acp-19-8669-2019, 2019.

873   Saito, R., Patra, P. K., Sweeney, C., Machida, T., Krol, M., Houweling, S., Bousquet, P., Agusti-Panareda,
874   A., Belikov, D., Bergmann, D., Bian, H., Cameron-Smith, P., Chipperfield, M. P., Fortems-Cheiney, A.,
875   Fraser, A., Gatti, L. V., Gloor, E., Hess, P., Kawa, S. R., Law, R. M., Locatelli, R., Loh, Z., Maksyutov, S.,
876   Meng, L., Miller, J. B., Palmer, P. I., Prinn, R. G., Rigby, M., and Wilson, C.: TransCom model simulations
877   of methane: Comparison of vertical profiles with aircraft measurements, Journal of Geophysical
878   Research: Atmospheres, 118, 3891-3904, https://doi.org/10.1002/jgrd.50380, 2013.

879   Shapley, L. S.: A Value for N-Person Games, in: Contributions to the Theory of Games, edited by: Kuhn,
880   H. W., and Tucker, A. W., Ann. Math. Studies, 28, Princeton University Press, 307-317, 1953.

881   Sherwen, T., Chance, R. J., Tinel, L., Ellis, D., Evans, M. J., and Carpenter, L. J.: A machine-learning-based
882   global sea-surface iodide distribution, Earth Syst. Sci. Data, 11, 1239-1262, 10.5194/essd-11-1239-2019,
883   2019.

884   Shi, L., Schreck, C., and Schröder, M.: Assessing the Pattern Differences between Satellite-Observed
885   Upper Tropospheric Humidity and Total Column Water Vapor during Major El Niño Events, Remote
886   Sensing, 10, 10.3390/rs10081188, 2018.

887   Spivakovsky, C. M., Wofsy, S. C., and Prather, M. J.: A numerical method for parameterization of
888   atmospheric chemistry: Computation of tropospheric OH, Journal of Geophysical Research:
889   Atmospheres, 95, 18433-18439, https://doi.org/10.1029/JD095iD11p18433, 1990.

890   Spivakovsky, C. M., Logan, J. A., Montzka, S. A., Balkanski, Y. J., Foreman-Fowler, M., Jones, D. B. A.,
891   Horowitz, L. W., Fusco, A. C., Brenninkmeijer, C. A. M., Prather, M. J., Wofsy, S. C., and McElroy, M. B.:
892   Three-dimensional climatological distribution of tropospheric OH: Update and evaluation, Journal of
893   Geophysical Research: Atmospheres, 105, 8931-8980, 10.1029/1999jd901006, 2000.

894   Stirnberg, R., Cermak, J., Fuchs, J., and Andersen, H.: Mapping and Understanding Patterns of Air Quality
895   Using Satellite Data and Machine Learning, Journal of Geophysical Research: Atmospheres, 125,
896   e2019JD031380, 10.1029/2019JD031380, 2020.

897   Strahan, S. E., Duncan, B. N., and Hoor, P.: Observationally derived transport diagnostics for the
898   lowermost stratosphere and their application to the GMI chemistry and transport model, Atmos. Chem.
899   Phys., 7, 2435-2445, 10.5194/acp-7-2435-2007, 2007.

900   Strode, S. A., Duncan, B. N., Yegorova, E. A., Kouatchou, J., Ziemke, J. R., and Douglass, A. R.: Implications
901   of carbon monoxide bias for methane lifetime and atmospheric composition in chemistry climate
902   models, Atmospheric Chemistry and Physics, 15, 11789-11805, 10.5194/acp-15-11789-2015, 2015.

903   Strode, S. A., Ziemke, J. R., Oman, L. D., Lamsal, L. N., Olsen, M. A., and Liu, J.: Global changes in the
904   diurnal cycle of surface ozone, Atmospheric Environment, 199, 323-333,
905   10.1016/j.atmosenv.2018.11.028, 2019.

906   Turner, A. J., Frankenberg, C., Wennberg, P. O., and Jacob, D. J.: Ambiguity in the causes for decadal
907   trends in atmospheric methane and hydroxyl, Proceedings of the National Academy of Sciences, 114,
908   5367, 10.1073/pnas.1616020114, 2017.

909 Turner, A. J., Fung, I., Naik, V., Horowitz, L. W., and Cohen, R. C.: Modulation of hydroxyl variability by
910 ENSO in the absence of external forcing, Proc Natl Acad Sci U S A, 115, 8931-8936,
911 10.1073/pnas.1807532115, 2018.

912 Voulgarakis, A., Naik, V., Lamarque, J. F., Shindell, D. T., Young, P. J., Prather, M. J., Wild, O., Field, R. D.,
913 Bergmann, D., Cameron-Smith, P., Cionni, I., Collins, W. J., Dalsøren, S. B., Doherty, R. M., Eyring, V.,
914 Faluvegi, G., Folberth, G. A., Horowitz, L. W., Josse, B., MacKenzie, I. A., Nagashima, T., Plummer, D. A.,
915 Righi, M., Rumbold, S. T., Stevenson, D. S., Strode, S. A., Sudo, K., Szopa, S., and Zeng, G.: Analysis of
916 present day and future OH and methane lifetime in the ACCMIP simulations, Atmospheric Chemistry and
917 Physics, 13, 2563-2587, 10.5194/acp-13-2563-2013, 2013.

918 Voulgarakis, A., Marlier, M. E., Faluvegi, G., Shindell, D. T., Tsigaridis, K., and Mangeon, S.: Interannual
919 variability of tropospheric trace gases and aerosols: The role of biomass burning emissions, Journal of
920 Geophysical Research: Atmospheres, 120, 7157-7173, 10.1002/2014jd022926, 2015.

921 Wang, J. S., Logan, J. A., McElroy, M. B., Duncan, B. N., Megretskaia, I. A., and Yantosca, R. M.: A 3-D
922 model analysis of the slowdown and interannual variability in the methane growth rate from 1988 to
923 1997, Global Biogeochemical Cycles, 18, https://doi.org/10.1029/2003GB002180, 2004.

924 Wolfe, G. M., Nicely, J. M., St Clair, J. M., Hanisco, T. F., Liao, J., Oman, L. D., Brune, W. B., Miller, D.,
925 Thames, A., Gonzalez Abad, G., Ryerson, T. B., Thompson, C. R., Peischl, J., McCain, K., Sweeney, C.,
926 Wennberg, P. O., Kim, M., Crounse, J. D., Hall, S. R., Ullmann, K., Diskin, G., Bui, P., Chang, C., and Dean-
927 Day, J.: Mapping hydroxyl variability throughout the global remote troposphere via synthesis of airborne
928 and satellite formaldehyde observations, Proc Natl Acad Sci U S A, 116, 11171-11180,
929 10.1073/pnas.1821661116, 2019.

930 Wolter, K., and Timlin, M. S.: El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an
931 extended multivariate ENSO index (MEI.ext), International Journal of Climatology, 31, 1074-1087,
932 10.1002/joc.2336, 2011.

933 Worden, J., Jiang, Z., Jones, D. B. A., Alvarado, M., Bowman, K., Frankenberg, C., Kort, E. A., Kulawik, S.
934 S., Lee, M., Liu, J., Payne, V., Wecht, K., and Worden, H.: El Niño, the 2006 Indonesian peat fires, and the
935 distribution of atmospheric methane, Geophysical Research Letters, 40, 4938-4943,
936 https://doi.org/10.1002/grl.50937, 2013.

937 Yan, R., Ma, Z., Zhao, Y., and Kokogiannakis, G.: A decision tree based data-driven diagnostic strategy for
938 air handling units, Energy and Buildings, 133, 37-45, https://doi.org/10.1016/j.enbuild.2016.09.039,
939 2016.

940 Zhang, Z., Zimmermann, N. E., Calle, L., Hurtt, G., Chatterjee, A., and Poulter, B.: Enhanced response of
941 global wetland methane emissions to the 2015–2016 El Niño-Southern Oscillation event, Environmental
942 Research Letters, 13, 10.1088/1748-9326/aac939, 2018.

943 Zhu, Q., Laughner, J. L., and Cohen, R. C.: Combining Machine Learning and Satellite Observations to
944 Predict Spatial and Temporal Variation of near Surface OH in North American Cities, Environmental
945 Science & Technology, 10.1021/acs.est.1c05636, 2022.
946