

Using machine learning to build temperature-based ozone parameterizations for climate sensitivity simulations

Peer Nowack^{1,2,3*}, Peter Braesicke⁴, Joanna Haigh^{1,2}, Luke Abraham^{5,6}, John Pyle^{5,6}
and Apostolos Voulgarakis²

¹Grantham Institute for Climate Change, Imperial College London, UK.

²Department of Physics, Blackett Laboratory, Imperial College London, UK.

³Data Science Institute, Imperial College London, UK.

⁴IMK-ASF, Karlsruhe Institute of Technology, Germany.

⁵National Centre for Atmospheric Science, UK.

⁶Department of Chemistry, University of Cambridge, UK.

Supplementary Information

Content:

1. Climate model details
2. Tables of regression
3. Supplementary figures

1. Climate model details

The data for this study was produced using the atmosphere-ocean coupled configuration of the Hadley Centre Global Environment Model version 3 (HadGEM3-AO) from the United Kingdom Met Office (Hewitt et al. 2011), coupled to the United Kingdom Chemistry and Aerosols (UKCA) atmospheric chemistry scheme. The atmosphere part is the UK Met Office's Unified Model based on a regular grid with a horizontal resolution of 3.75°longitude by 2.5°latitude and 60 vertical levels up to a height of ca. 84 km. The version used here has a self-contained stratospheric Quasi-Biennial Oscillation. The ocean component is the Ocean Parallélisé part of the Nucleus for European Modelling of the Ocean model version 3.0 (Madec 2008). The ocean configuration here deploys a tripolar, locally anisotropic grid which has 2°resolution in longitude everywhere, but an increased latitudinal resolution in certain regions with up to 0.5°in the tropics. For atmospheric chemistry, we use an updated version of the detailed stratospheric chemistry configuration of UKCA (Morgenstern et al. 2009, Nowack et al. 2016, 2018, 2017, 2014), which is coupled to the atmosphere model. A relatively simple tropospheric chemistry scheme that simulates hydrocarbon oxidation is included. Nitrogen oxide emissions from lightning are parameterized according to (Price & Rind 1992, 1994). Photolysis rates are calculated interactively using the FastJX-photolysis scheme (Bian & Prather 2002, Neu et al. 2007, Telford et al. 2013, Wild et al. 2000). In total 159 chemical reactions involving 41 chemical species are considered.

2. Tables of regressions

Daily mean regressions:

Regression method	No temperature/ozone components	α /Others	Scenario, Timestep
Ridge (*)	5	636500	pi, t
Ridge (*)	100	205000	pi, t
Ridge (*)	250	151500	pi, t
Ridge (*)	500	119500	pi, t
Ridge (*)	1000	89000	pi, t
Ridge (*)	2000	62500	pi, t
Ridge (*)	3000	52000	pi, t
Ridge (*)	5	638000	pi, t-1
Ridge (*)	100	215500	pi, t-1
Ridge (*)	250	166500	pi, t-1
Ridge (*)	500	138000	pi, t-1
Ridge (*)	1000	108000	pi, t-1
Ridge (*)	2000	82500	pi, t-1
Ridge (*)	3000	73500	pi, t-1
Ridge	2000	89000	pi, t-1
Lasso (Tibshirani 1996) (*)	5	0.105	pi, t
Lasso	100	0.0385	pi, t
Lasso	500	0.02075	pi, t
Lasso	100	0.0375	pi, t-1
Lasso	500	0.022	pi, t-1
Ridge (*), 10 years training	1500	53000	4xCO ₂ , t
Ridge (*), 25 years training	1500	58500	4xCO ₂ , t
Ridge (*), 25 years training	2000/2000	40000	4xCO ₂ , t
Ridge (*), 25 years training	2500/2500	58500	4xCO ₂ , t-1
Ridge, 25 years training	2500/3000	68000	4xCO ₂ , t-1

5-day-mean regressions:

Regression method	No temperature/ozone components	α /Others	Scenario, Timestep
Ridge (*), 25 years training	1000/1500	50000	4xCO2, t-1
Ridge, 5 years training	1000/1500	57000	4xCO2, t-1
Ridge, 10 years training	1000/1500	59500	4xCO2, t-1
Ridge, 15 years training	1000/1500	57500	4xCO2, t-1
Ridge, 20 years training	1000/1500	55500	4xCO2, t-1
Ridge, 25 years training	1000/1500	54500	4xCO2, t-1
Ridge, 50 years training	1000/1500	47500	4xCO2, t-1
Ridge (*)	2	105000	pi, t
Ridge (*)	5	148000	pi, t
Ridge (*)	10	178500	pi, t
Ridge (*)	25	93500	pi, t
Ridge (*)	50	69000	pi, t
Ridge (*)	100	54000	pi, t
Ridge (*)	250	41000	pi, t
Ridge (*)	500	33000	pi, t
Ridge (*)	750	28000	pi, t
Ridge (*)	1000	24500	pi, t
Ridge (*)	2000	18500	pi, t
Ridge (*)	2	110000	pi, t-1
Ridge (*)	5	153500	pi, t-1
Ridge (*)	10	195000	pi, t-1
Ridge (*)	25	125000	pi, t-1
Ridge (*)	50	105500	pi, t-1
Ridge (*)	100	97000	pi, t-1
Ridge (*)	250	82000	pi, t-1
Ridge (*)	500	72000	pi, t-1
Ridge (*)	750	67000	pi, t-1
Ridge (*)	1000	63500	pi, t-1
Ridge (*)	2000	57500	pi, t-1
Lasso	2000	64000	pi, t-1
Lasso	2000/2500	61000	pi, t-1
Lasso	1000	0.0285	pi, t
Neural network ¹	1000	2e-4	pi, t

¹Feedforward neural network optimized over a range of learning rates (final 2e-5), using L2 regularization. Two hidden layers each consisting of 1024 units. 10000 epochs maximum iteration, but early stopping after 6191 epochs.

Monthly-mean regressions:

Regression method	No temperature/ozone components	α /Others	Scenario, Timestep
Ridge (*)	5	35500	pi, t
Ridge (*)	50	21500	pi, t
Ridge (*)	100	17500	pi, t
Ridge (*)	250	13000	pi, t
Ridge (*)	500	9000	pi, t
Ridge (*)	600	7500	pi, t
Ridge (*)	5	40500	pi, t-1
Ridge (*)	50	51500	pi, t-1
Ridge (*)	100	49000	pi, t-1
Ridge (*)	250	44000	pi, t-1
Ridge (*)	500	38500	pi, t-1
Ridge (*)	599	38000	pi, t-1
Lasso (*)	500	0.11	pi, t
Lasso (*)	500	0.272	pi, t
Random Forest (*) ²	500	see footnote	pi, t
Ridge (*), 10 years training (*)	119	20500	4xCO ₂ , t-1
Ridge (*), 25 years training (*)	200	30500	4xCO ₂ , t-1

Notes:

- pi=preindustrial, 4xCO₂=abrupt 4xCO₂, t=matching temperature and ozone field at the same time, t-1=prediction regression one timestep in advance.
- For a few mappings, we also applied PCA to the output ozone field (thus mapping temperature components onto ozone components), which reduces the number of regression fitting procedures by orders of magnitude, but also makes an additional decomposition and afterwards re-factoring to ozone concentrations necessary, thus adding computational complexity. We found that the corresponding loss in predictive skill is marginal, making this a valid alternative method depending on whether the user wishes to speed up the α optimization step at the expense of additional large matrix inversion calculations.
- For all regression methods we used only one global regularization parameter, even though a regression is build for every grid cell individually. Certainly, every grid cell regression would have its own best regularization parameter so that significant improvements might be gained by carrying out individual regressions. However, when testing this effect by covering any feasible range of α parameters for each individual regression (a large number due to 22 vertical levels, 73 latitudes and 96 longitudes → 154176 optimizations for each regression in the table), we found that this did not actually lead to significant improvements. On the contrary, we found that a global regularization parameter, if refined sufficiently, typically gave better results. The global regularization parameter was typically refined until the GridSearchCV cross-validation score did not change by more than one digit on the sixth decimal (Pedregosa et al. 2011).
- For regressions marked by (*), we performed normal five-fold cross validation rather than time series cross-validation (Bishop 2006). We found hardly any significant differences due to the different cross-validation methods in terms of the final prediction error.
- The Ridge- α parameter is defined according to the definition of the cost function of Ridge regression in the scikit-learn implementation (Pedregosa et al. 2011) in Python 3.

²Scikit-learn random forest regressor with standard settings apart from min_samples_leaf=3, 300 estimators, bootstrap=True,min_samples_split=2.

3. Supplementary figures

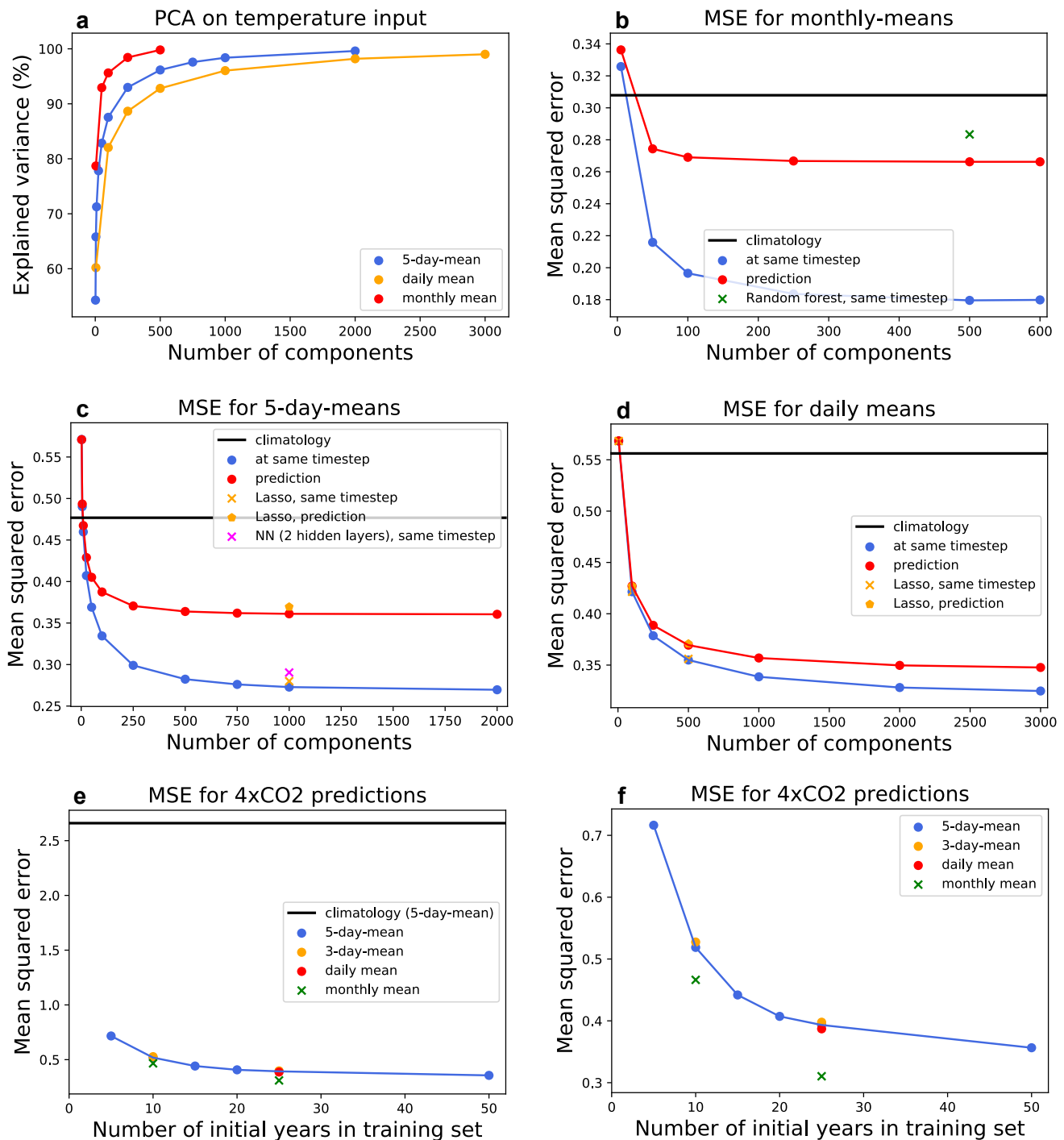


Figure S1. Description, see text below.

Description of Figure S1:

(a) Explained variance of the 50-year long preindustrial training dataset depending on the number of principal components for monthly-mean, 5-day-mean and daily mean data.

(b) Mean squared error (MSE) of various regression methods (as labeled) for monthly-mean data depending on the number of temperature components, calculated on the independent testset. Even a small number of components is better than using a fixed climatology (black line), but including more than 100 components still yields significant improvements in both matching the monthly-mean temperature state to the ozone field over the same period (blue line), or making a one month in advance

prediction for the monthly-mean ozone field (red line). Clearly, the prediction lead time of one month limits how much the MSE can be reduced compared to a climatology, but the prediction still does better than the climatology.

(c) The same for 5-day-mean data. Due to the reduced lead time, the prediction error is in the relative sense closer to the regression carried out for the same time interval. The overall MSE increases though due to increased noise on shorter time scales.

(d) Again the same for daily-mean data. The prediction is now much closer to the best possible case of no lead time, implying an increased predictive capability one day in advance. Most of the error originates from the troposphere (Figure S6), where including even shorter time-scales (ideally instantaneous) could lead to further improvements. However, such short-term fluctuations are not expected to be of particular climatological importance.

(e,f) MSE for the 4xCO₂ predictions only, depending on the number of years taken into account in the training period. (e) includes the MSE of using a fixed climatology on the 175-years 5-day-mean testset (obtained by interpolating a monthly-mean ozone climatology linearly to five-day time resolution), which clearly leads to a large error as the response to 4xCO₂, both in terms of variability and mean changes, is not included. (f) shows the same data, but without the error of the climatology. Clearly, including more initial data in the training set has the potential to reduce the MSE, but most information is gained within the first 10-15 years, which would thus be the recommended minimum time of training for the ML model following a 4xCO₂ forcing.

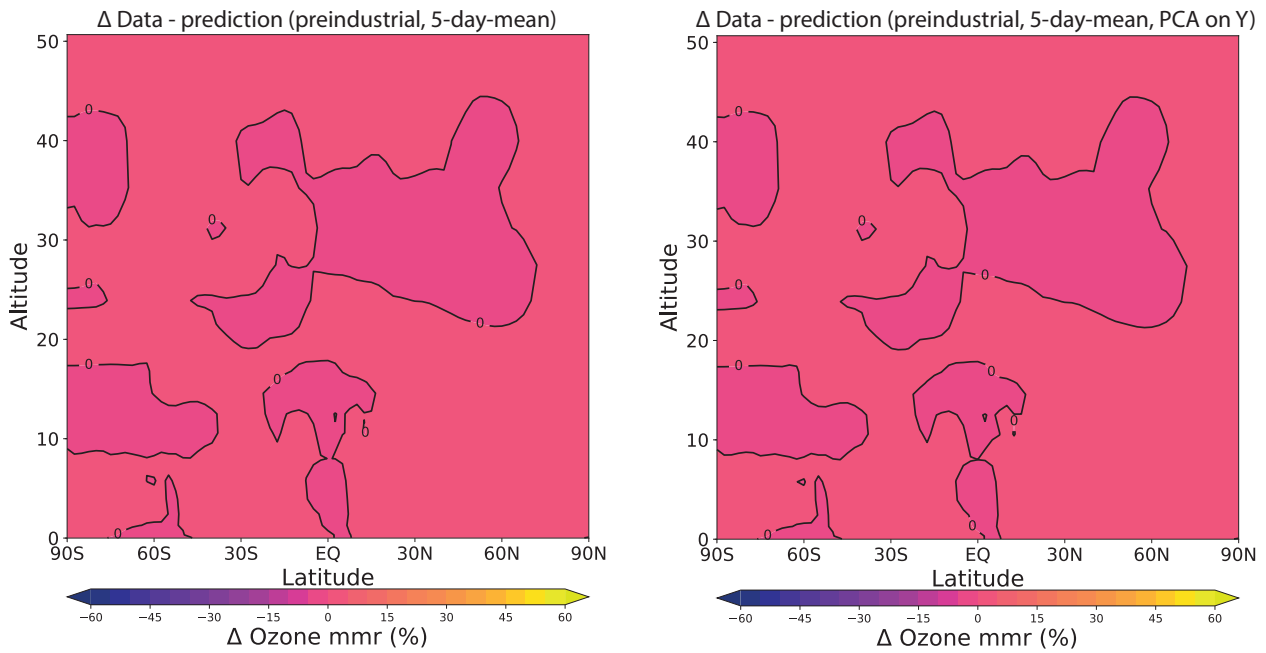


Figure S2. As Figure 2d in the main text but for predictions based on 5-day-average data. While the relative MSE is higher, the resulting average climatology is still well reproduced, in agreement with the intuition that in this steady forcing scenario, the main sources of MSE come from deviations in ozone variability around a well-determined mean state. The plot on the right is equivalent, but was produced by a regression in which the 5-day-mean ozone data was also dimension-reduced prior to the fit. As a result, the fit was a mapping between principal components of temperature and ozone rather than for each ozone grid cell individually. The final ozone results were then obtained by reverse transformation of the predicted ozone components back into the 'concentration space' for each grid cell. Clearly, using this dimension reduction does not lead to significantly different results, see also description of the regressions in Supplementary section 2.

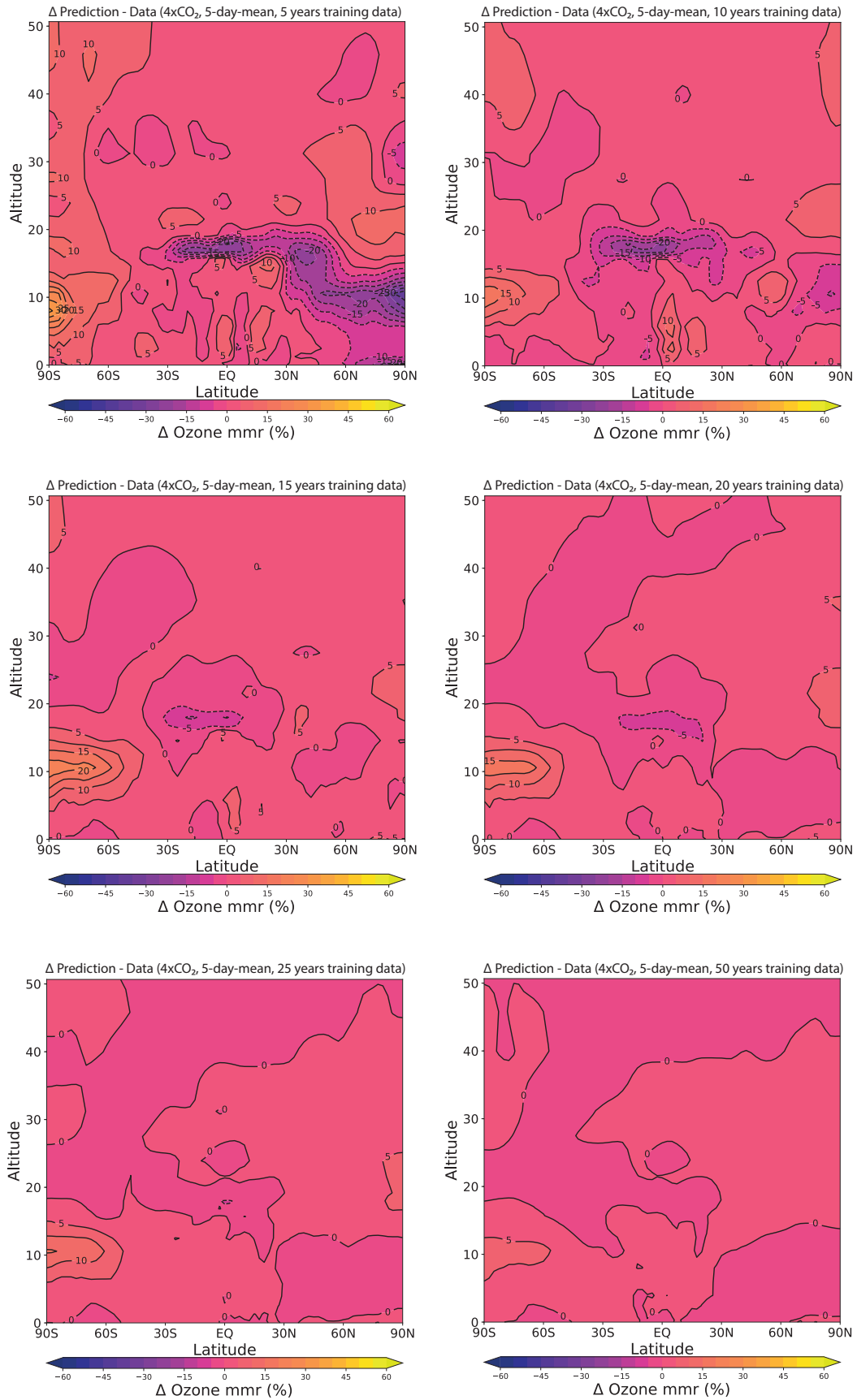


Figure S3. As Figure 2f in the main text, but for different lengths of training periods (5, 10, 15, 20, 25 and 50 years, using 5-day-mean rather than daily-mean data as done in the main text). As can also be deduced from Figure S1f, a good ozone climatology can already be obtained using merely the first 10-15 years of data, although using longer training datasets clearly leads to improvements. All differences are shown for the entire remaining test period, i.e. 195, 190, 185, 180, 175 and 150 years of data.

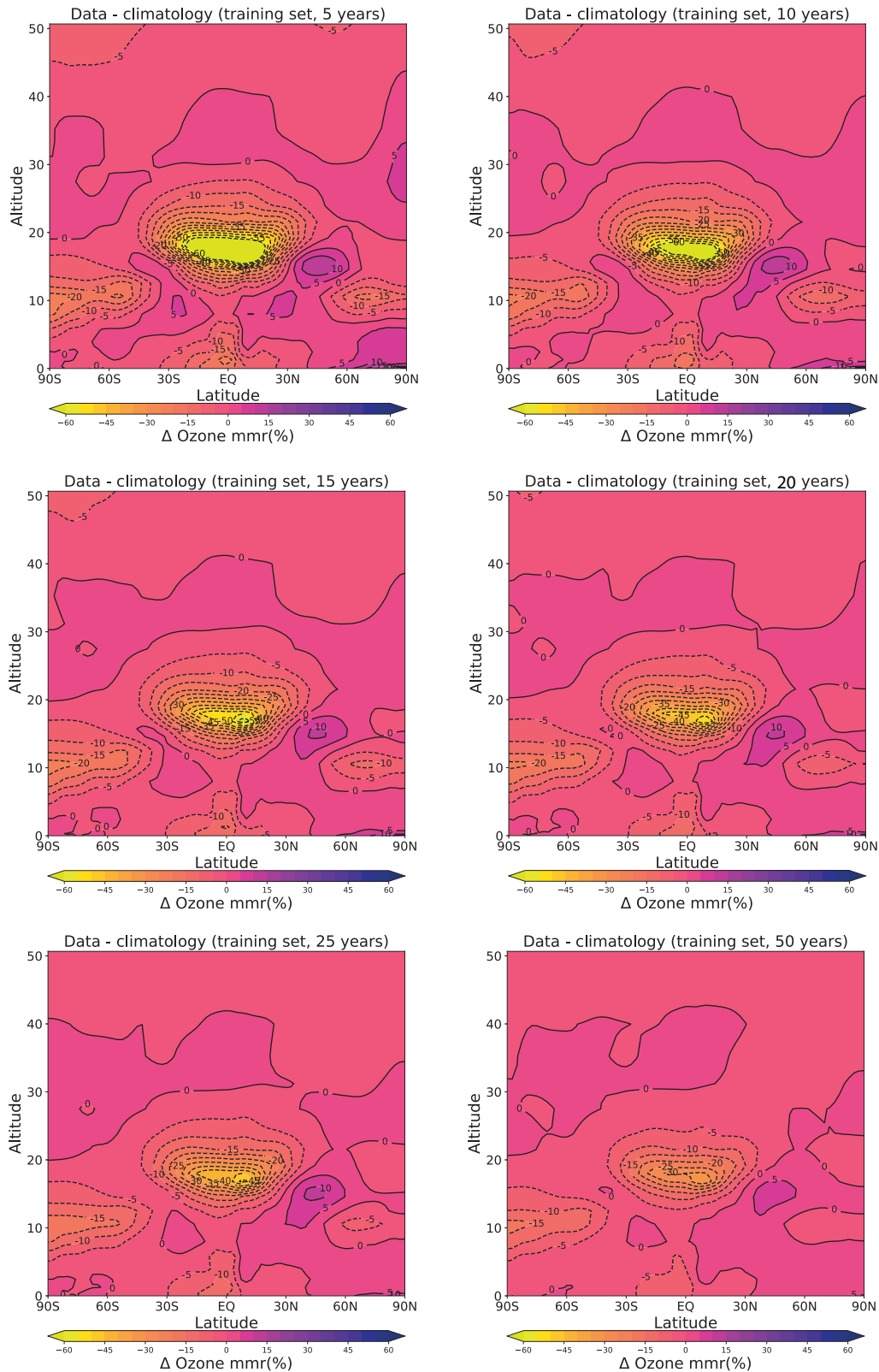


Figure S4. As Figure S3, but now showing the error if the average climatology of the training dataset would be used rather than the preindustrial climatology. While changes in mean upper stratospheric ozone could be fairly well captured by such a climatology due to the fast adaptation of catalytic ozone depletion cycles to the also fast stratospheric cooling following the quadrupling of atmospheric CO₂ (in essence fast stratospheric adjustment), ozone decreases in the lower stratosphere cannot be captured by such a climatology. This is due to the long timescale involved in the lower stratospheric ozone response, which in turn is coupled to tropospheric warming (Chiodo *et al.* 2018, SPARC 2010). Such slow ozone feedbacks are however central to its role in climate sensitivity simulations (Chiodo & Polvani 2016, Dietmüller *et al.* 2014, Nowack *et al.* 2015) and are thus important to include.

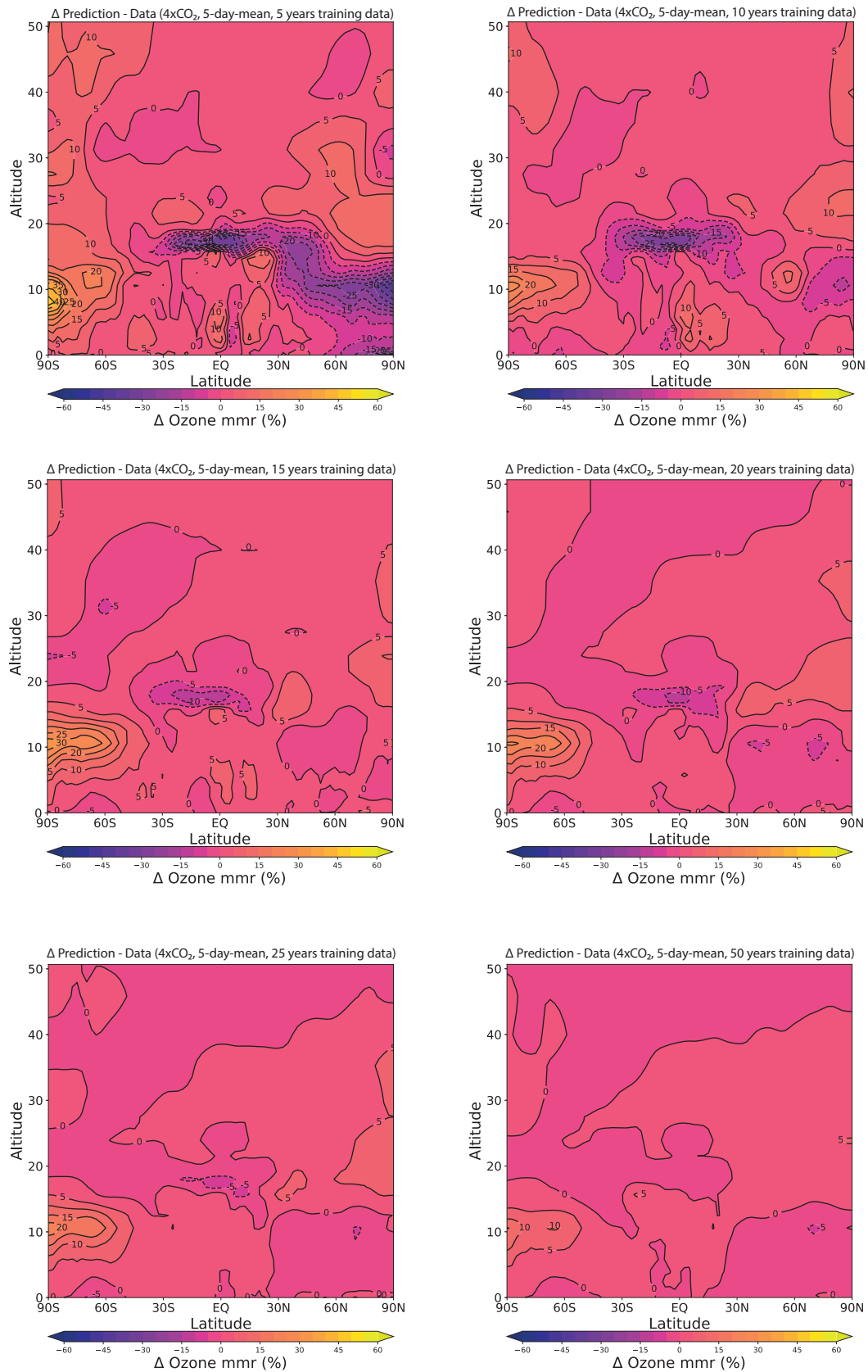
Differences for years 190-200 of the abrupt 4xCO₂ simulation

Figure S5. As Figure S3, but with the differences between the ML predictions taken over years 190-200 after the CO₂ forcing. For training sets longer than about 10 years the prediction error does not increase much towards the end of the simulations (compare to Figure S3 directly), implying that ozone trends are well captured by the linear ML model used here.

Preindustrial, daily mean, 50 years training data

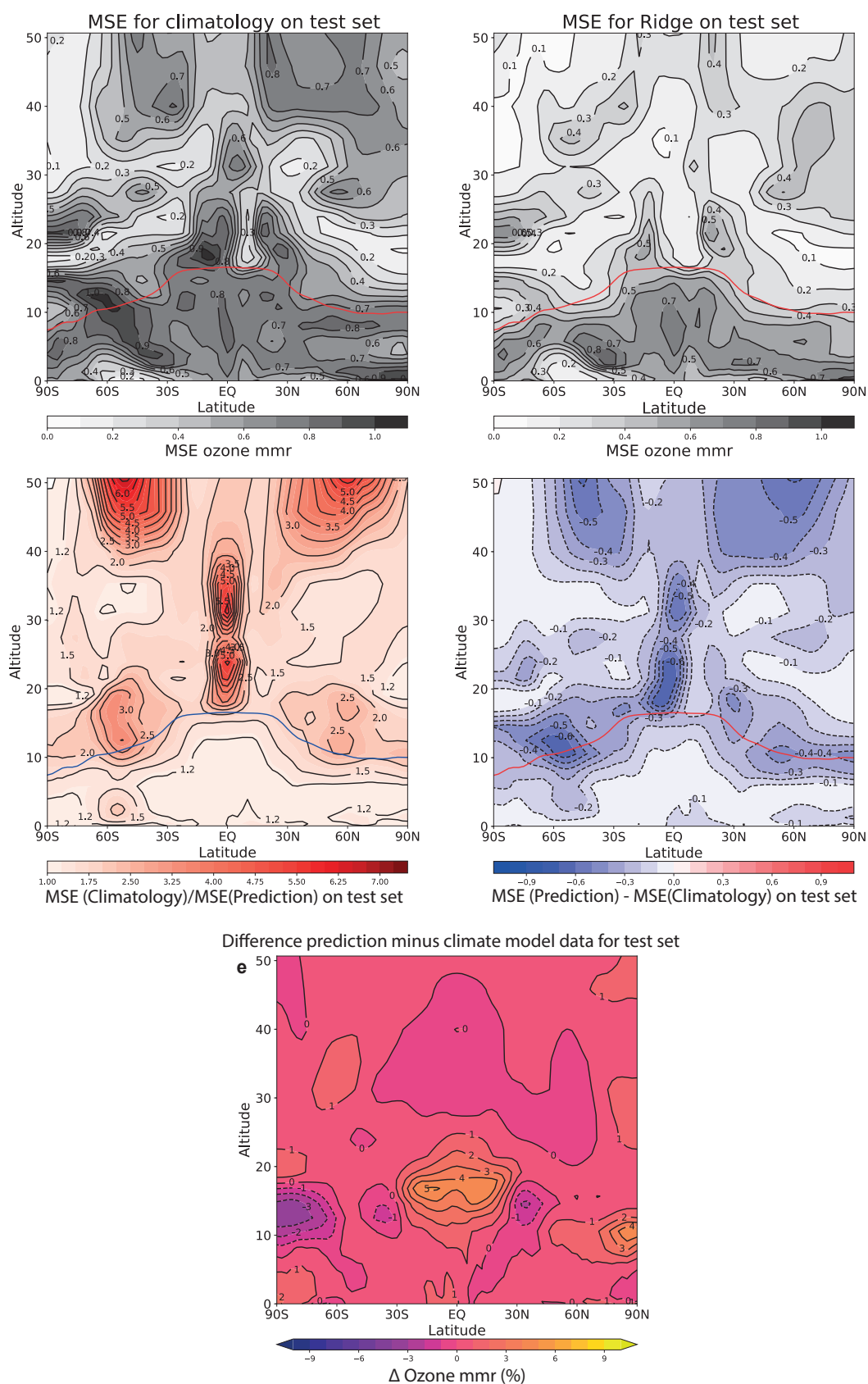


Figure S6. Description, see below.

Description of Figure S6: For preindustrial daily mean data (13-year long testset). Top left: Zonal average mean squared error (MSE) for the fixed ozone climatology relative to the ozone field produced by the fully-interactive climate model simulation. Top right: the same, but for the ML model discussed in the main text. The thin red line in either plot marks the climatological tropopause following the WMO lapse rate definition (WMO 1957). The results are for normalized ozone data, thus making MSE for normally much lower tropospheric ozone mass mixing ratios comparable to those in the ozone-rich stratosphere. Clearly, the MSE is reduced everywhere using the ML model, but most significant gains are found in the slow-moving stratosphere, as expected from the temperature correlation of ozone with the most important mechanisms there and the smaller error introduced by the one-day extrapolation in the predictions for this more slowly changing environment than is the troposphere. Middle row: the left plot shows the MSE ratio (climatology MSE divided by ML MSE), underlining the fractional improvement introduced by using the ML model. As the mean differences in the climatology (shown in the bottom row) are relatively small, the main MSE differences must arise from a better modeling of ozone variability using the ML regression. The most significant improvements are found in the tropical stratosphere (improvement by a factor of up to >7), in the mid-latitude upper stratosphere (also up to factor 7 better) and in the mid-to-high latitude lower stratospheres (up to factor 3). All of these regions are associated with major modes of variability in the form of the Quasi-Biennial Oscillation, the semi-annual oscillation and the mid-latitude jets and polar vortices, suggesting a direct link to the improved modeling of ozone variability there and the MSE reduction. The right plot in the middle row shows the absolute improvements in terms of the MSE.

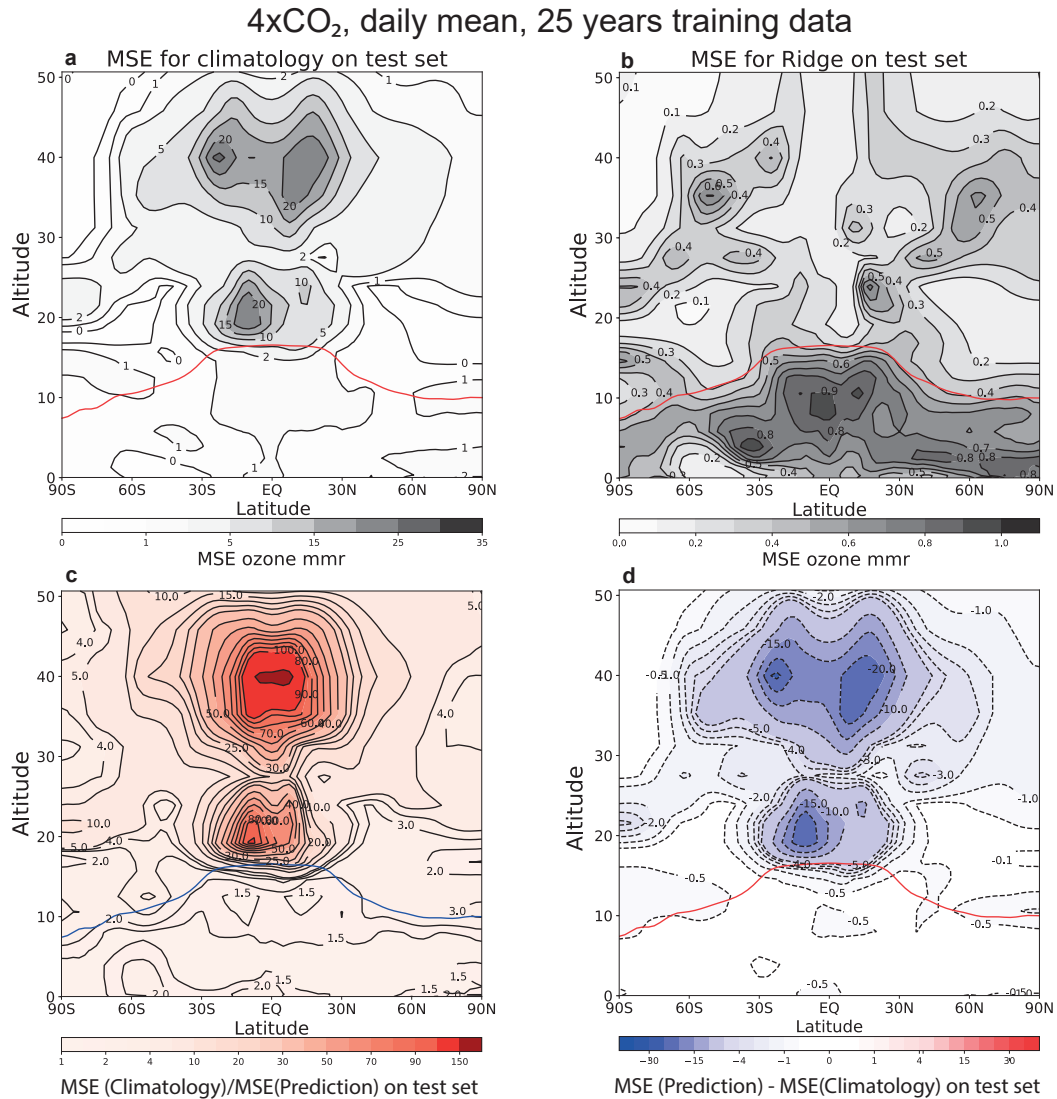


Figure S7. As Figure S6, but for the 4xCO₂ run discussed in the main text. Due to the significant changes in ozone under 4xCO₂ in every model grid cell (Figure 2c), the MSE improvement by using the ML regression model will be caused by both the better representation of ozone variability and trends. The large offsets lead to very large MSE in ozone for the climatology, especially in the upper stratosphere. As a result, the use of the ML model can lead to MSE improvement of factors > 150 as compared to a fixed preindustrial climatology.

Figure S8: Ozone variability in a variety of grid cells as given by the chemistry-climate model and as predicted by the ML model for daily mean preindustrial testdata (as Figure 3 in the main text, an extension for more grid cells). See supplementary file [pi_trends_dm_supp_r.pdf](#) and the figure on the next page.

Figure S9: Ozone variability in a variety of grid cells as given by the chemistry-climate model and as predicted by the ML model for 5-day-mean preindustrial testdata (as Figure 3 in the main text, but different time averaging of the data). See supplementary file [pi_trends_5dm_supp_NOYPCA_r.pdf](#)

Figure S10/S11: Ozone variability in a variety of grid cells as given by the chemistry-climate model and as predicted by the ML model for daily-mean 4xCO₂ testdata (as Figure 4 in the main text, an extension of the figure for more grid cells). See supplementary files [4xco2_trends_dm_supp_r.pdf](#) and [4xco2_trends_dm_supp_2_r.pdf](#)

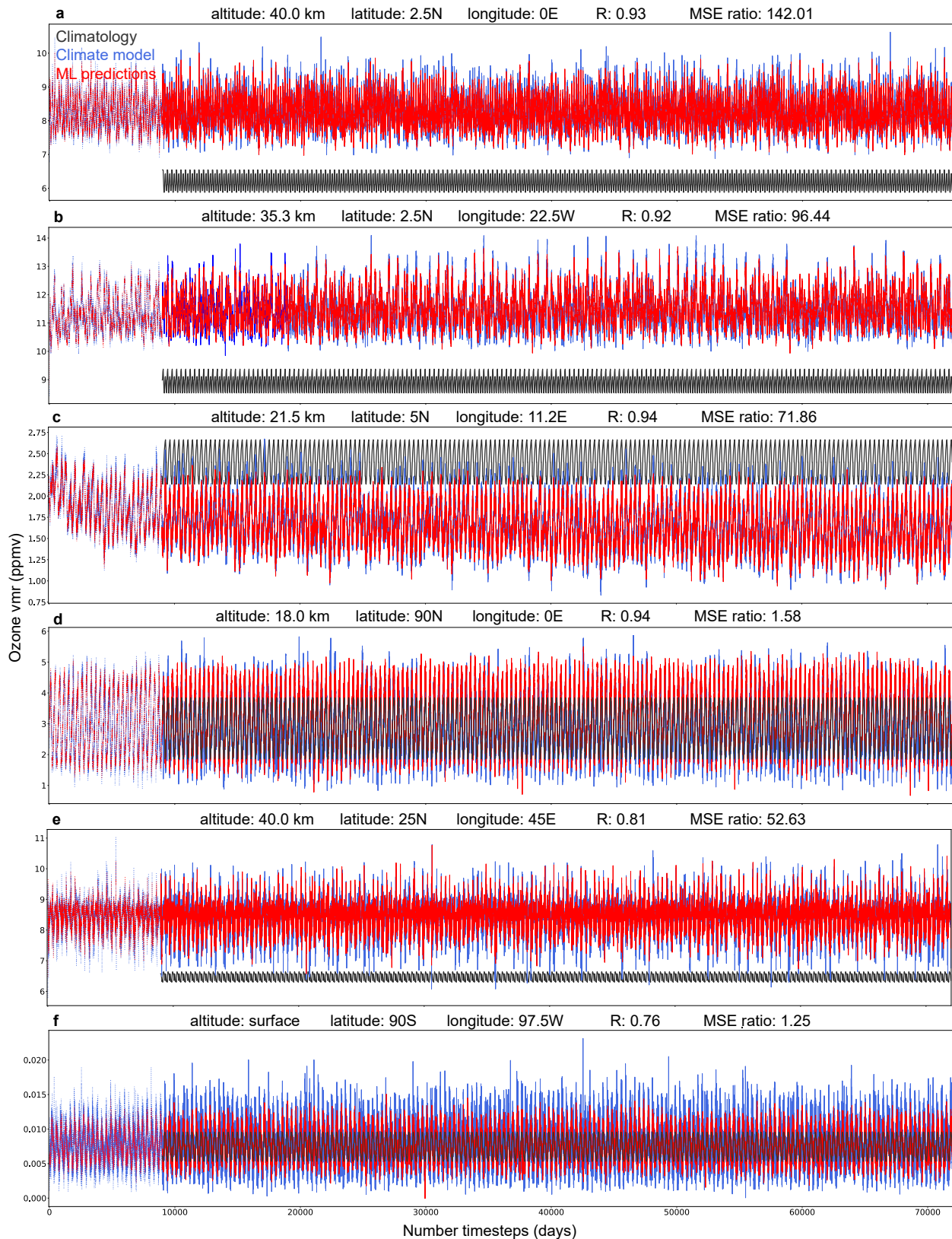


Figure S8 continued. Ozone (ppmv) time series data for various model grid cells following an instantaneous quadrupling of atmospheric CO₂. Color code and labels as in Figure 3 in the main text. The first 25 years are used as training data (indicated by the dotted lines), the subsequent 175 years (timesteps 9001-72000 based on a 360-day calendar) are predictions for the ML model.

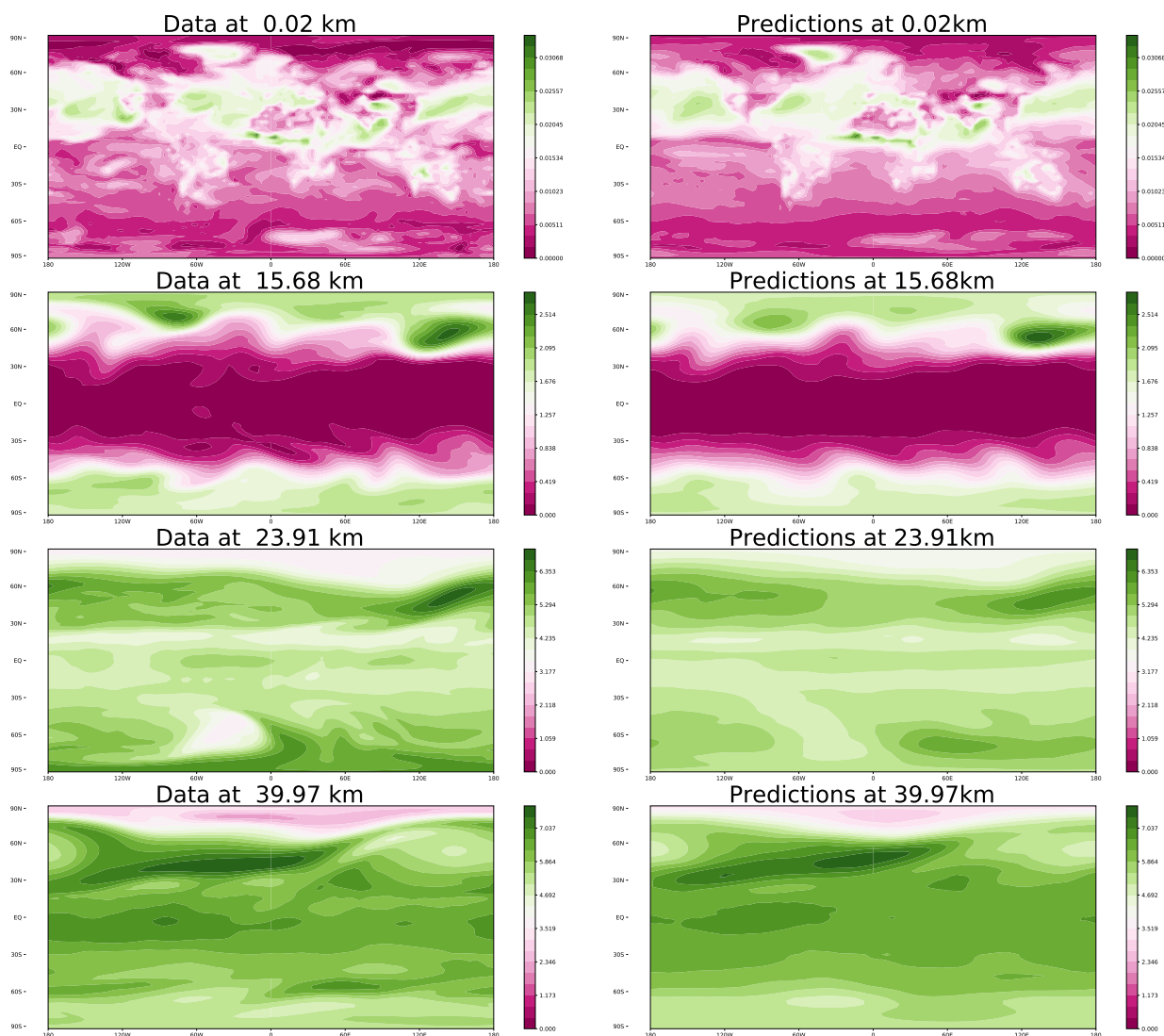


Figure S12.1 Ozone cross sections at randomly chosen timesteps and altitudes for the pre-industrial daily-mean ozone data. Left: Data produced by the chemistry-climate model, right: the corresponding ML prediction.

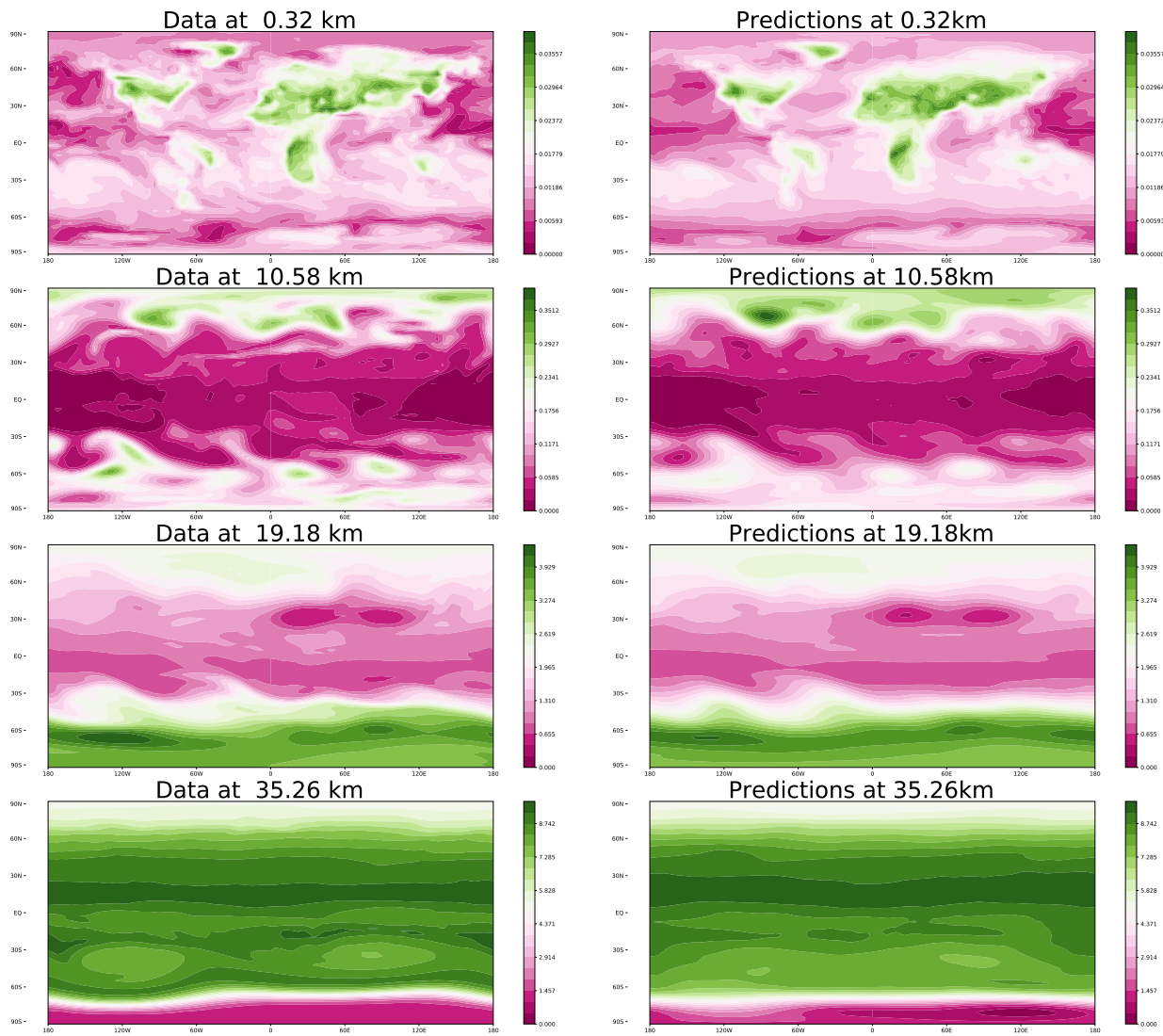


Figure S12.2 Ozone cross sections at randomly chosen timesteps and altitudes for the pre-industrial daily-mean ozone data. Left: Data produced by the chemistry-climate model, right: the corresponding ML prediction.

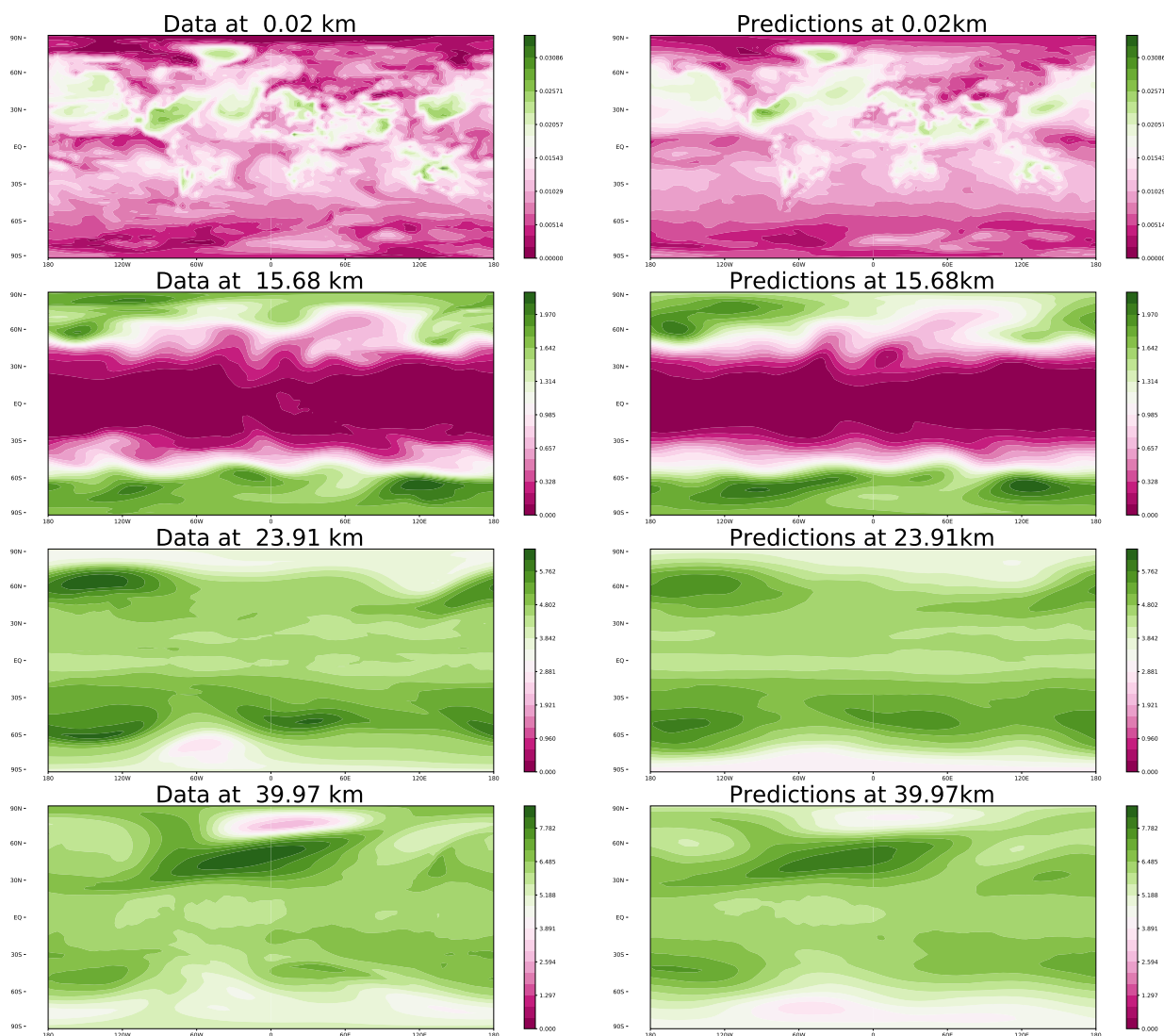


Figure S12.3 Ozone cross sections at randomly chosen timesteps and altitudes for the pre-industrial daily-mean ozone data. Left: Data produced by the chemistry-climate model, right: the corresponding ML prediction.

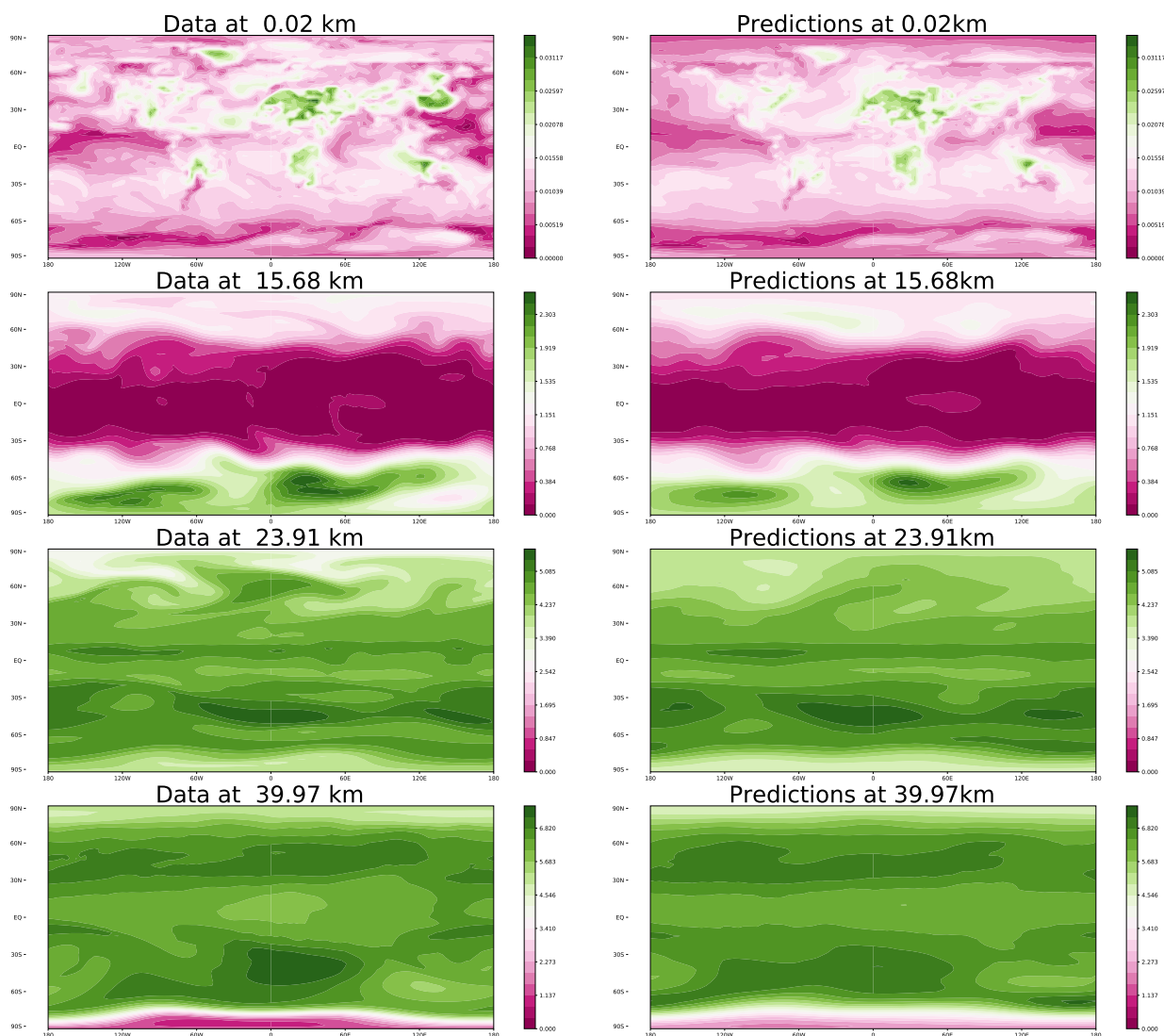


Figure S12.4 Ozone cross sections at randomly chosen timesteps and altitudes for the pre-industrial daily-mean ozone data. Left: Data produced by the chemistry-climate model, right: the corresponding ML prediction.

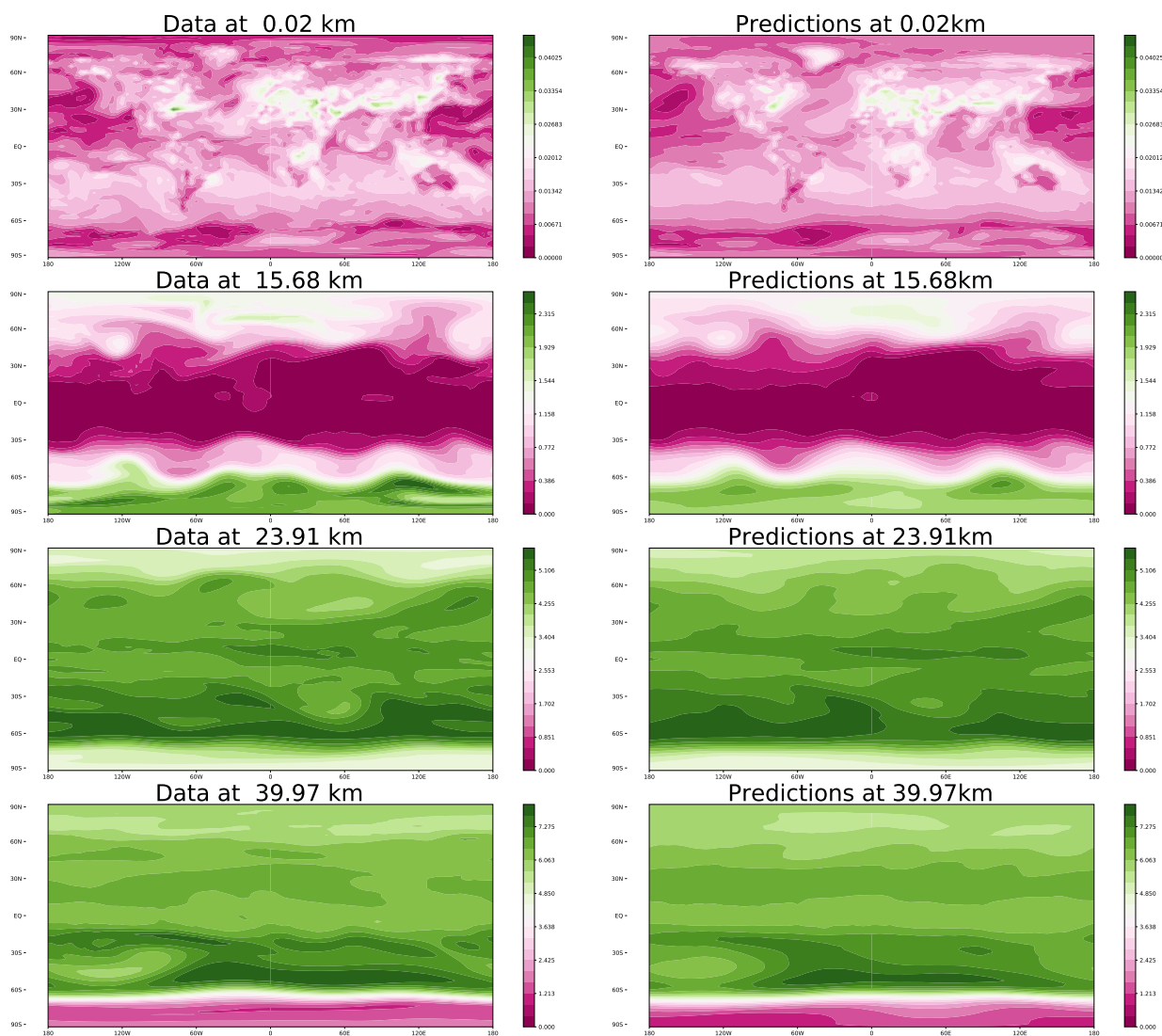


Figure S12.5 Ozone cross sections at randomly chosen timesteps and altitudes for the pre-industrial daily-mean ozone data. Left: Data produced by the chemistry-climate model, right: the corresponding ML prediction.

Bibliography

- Bian, H. & Prather, M. J. (2002), 'Fast-J2: Accurate simulation of stratospheric photolysis in global chemical models', *Journal of Atmospheric Chemistry* **41**(3), 281–296.
- Bishop, C. M. (2006), *Pattern recognition and machine learning*, Springer Science+Business Media, Singapore.
- Chiodo, G. & Polvani, L. M. (2016), 'Reduced Southern Hemispheric circulation response to quadrupled CO₂ due to stratospheric ozone feedback', *Geophysical Research Letters* **43**, 1–10.
URL: <http://doi.wiley.com/10.1002/2016GL071011>
- Chiodo, G., Polvani, L., Marsh, D., Stenke, A., Ball, W., Rozanov, E., Muthers, S. & Tsigaridis, K. (2018), 'The response of the ozone layer to quadrupled CO₂ concentrations', *Journal of Climate* **31**, 3893–3907.
URL: <http://journals.ametsoc.org/doi/10.1175/JCLI-D-17-0492.1>
- Dietmüller, S., Ponater, M. & Sausen, R. (2014), 'Interactive ozone induces a negative feedback in CO₂-driven climate change simulations', *Journal of Geophysical Research: Atmospheres* **119**, 1796–1805.
- Hewitt, H. T., Copsey, D., Culverwell, I. D., Harris, C. M., Hill, R. S. R., Keen, A. B., McLaren, A. J. & Hunke, E. C. (2011), 'Design and implementation of the infrastructure of HadGEM3: The next-generation Met Office climate modelling system', *Geoscientific Model Development* **4**(2), 223–253.
- Madec, G. (2008), NEMO ocean engine, Note du Pole de modelisation, Institut Pierre-Simon Laplace (IPSL), Technical report, France.
- Morgenstern, O., Braesicke, P., O'Connor, F. M., Bushell, A. C., Johnson, C. E., Osprey, S. M. & Pyle, J. A. (2009), 'Evaluation of the new UKCA climate-composition model – Part 1: The stratosphere', *Geoscientific Model Development* **2**(1), 43–57.
- Neu, J. L., Prather, M. J. & Penner, J. E. (2007), 'Global atmospheric chemistry: Integrating over fractional cloud cover', *Journal of Geophysical Research: Atmospheres* **112**, D11306.
- Nowack, L., Grolimund, D., Samson, V., Marone, F. & Wood, V. (2016), 'Rapid Mapping of Lithiation Dynamics in Transition Metal Oxide Particles with Operando X-ray Absorption Spectroscopy', *Scientific Reports* **6**(October 2015), 21479.
URL: <http://www.nature.com/articles/srep21479>
- Nowack, P. J., Abraham, N. L., Braesicke, P. & Pyle, J. A. (2018), 'The impact of stratospheric ozone feedbacks on climate sensitivity estimates', *Journal of Geophysical Research: Atmospheres* **123**.
- Nowack, P. J., Braesicke, P., Abraham, N. L. & Pyle, J. A. (2017), 'On the role of ozone feedback in the ENSO amplitude response under global warming', *Geophysical Research Letters* **44**.
- Nowack, P. J., Luke Abraham, N., Maycock, A. C., Braesicke, P., Gregory, J. M., Joshi, M. M., Osprey, A. & Pyle, J. a. (2014), 'A large ozone-circulation feedback and its implications for global warming assessments', *Nature Climate Change* (December), 2–6.
URL: <http://www.nature.com/doi/10.1038/nclimate2451>

- Nowack, P. J., Luke Abraham, N., Maycock, A. C., Braesicke, P., Gregory, J. M., Joshi, M. M., Osprey, A. & Pyle, J. A. (2015), 'A large ozone-circulation feedback and its implications for global warming assessments', *Nature Climate Change* **5**(1), 41–45.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. (2011), 'Scikit-learn: Machine learning in Python', *Journal of Machine Learning Research* **12**, 2825–2830.
- Price, C. & Rind, D. (1992), 'A simple lightning parameterization for calculating global lightning distributions', *Journal of Geophysical Research: Atmospheres* **97**(D9), 9919–9933.
- Price, C. & Rind, D. (1994), 'Modeling Global Lightning Distributions in a General Circulation Model', *Monthly Weather Review* **122**(8), 1930–1939.
- SPARC (2010), SPARC CCMVal Report on the Evaluation of Chemistry-Climate Models, Technical report, SPARC.
URL: <http://www.sparc-climate.org/publications/sparc-reports/>
- Telford, P. J., Abraham, N. L., Archibald, A. T., Braesicke, P., Dalvi, M., Morgenstern, O., O'Connor, F. M., Richards, N. A. D. & Pyle, J. A. (2013), 'Implementation of the Fast-JX Photolysis scheme (v6.4) into the UKCA component of the MetUM chemistry-climate model (v7.3)', *Geoscientific Model Development* **6**(1), 161–177.
- Tibshirani, R. (1996), 'Regression Shrinkage and Selection via the Lasso', *Journal of the Royal Statistical Society. Series B (Methodological)* **58**(1), 267–288.
- Wild, O., Zhu, X. & Prather, M. J. (2000), 'Fast-J: Accurate simulation of in- and below-cloud photolysis in tropospheric chemical models', *Journal of Atmospheric Chemistry* **37**(3), 245–282.
- WMO (1957), Meteorology - a three dimensional science: second session of the Commission for Aerology, in 'WMO Bull 4', pp. 134–138.