PLOS Climate

Pangeo-Enabled ESM Pattern Scaling (PEEPS): A customizable dataset of emulated Earth System Model output

--Manuscript Draft--

Manuscript Number:	PCLM-D-22-00161
Article Type:	Research Article
Full Title:	Pangeo-Enabled ESM Pattern Scaling (PEEPS): A customizable dataset of emulated Earth System Model output
Short Title:	Pangeo-Enabled ESM Pattern Scaling (PEEPS)
Corresponding Author:	Ben Kravitz Indiana University Bloomington Bloomington, IN UNITED STATES
Order of Authors:	Ben Kravitz
	Abigail C Snyder
Keywords:	Earth System Model; emulator; pattern scaling; Pangeo
Abstract:	Emulation through pattern scaling is a way of rapidly producing climate fields (like temperature or precipitation) from existing Earth System Model (ESM) output that, while inaccurate, is often useful for a variety of downstream purposes. Conducting pattern scaling has historically been a laborious process, in large part due to the increasing volume of ESM output. Here we introduce Pangeo-Enabled ESM Pattern Scaling (PEEPS), a Jupyter notebook-based repository of pattern scaled output from CMIP6. Because all source data for PEEPS is cloud-based, users do not need to download and house the ESM output, meaning that emulated patterns are essentially on demand. Validating PEEPS on the CMIP6 archive for temperature, precipitation, and near-surface relative humidity, pattern scaling performs well over a variety of future scenarios except for regions in which there are strong, potentially nonlinear climate feedbacks. Although pattern scaling is normally conducted on annual mean ESM output, it works equally well on monthly mean ESM output. We identify several downstream applications for PEEPS, including impacts assessment and evaluating climate system uncertainties.
Additional Information:	
Question	**Response**
Financial Disclosure Enter a financial disclosure statement that describes the sources of funding for the work included in this submission and the role the funder(s) played. This includes grants and any commercial funding of the work or authors. This statement will be typeset if the manuscript is accepted for publication. *Please review the submission guidelines and the instructions link below for detailed*	This work was conducted with the support of the U.S. Department of Energy, Office of Science, as part of research in MultiSector Dynamics, Earth and Environmental System Modeling Program. Support for B.K. was provided in part by the Indiana University Environmental Resilience Institute and the Prepared for Environmental Change Grand Challenge initiative. The Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle Memorial Institute under contract DE-AC05-76RL01830. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. Data analysis was performed using the Pangeo platform, which was supported by NSF Earthcube Award 1740648.

requirements and guidance.	
Competing Interests On behalf of all authors, disclose any competing interests that could be perceived to bias this work. This statement will be typeset if the manuscript is accepted for publication. *Please review the instructions link below and PLOS Climate's competing interests policy to determine what information must be disclosed at submission.*	The authors declare no competing interests.
Data Availability Before publication, Authors are required to make fully available and without restriction all data underlying their findings. Please see our PLOS Data Policy page for detailed information on this policy. A **Data Availability Statement**, detailing where the data can be accessed, is required at first submission. Insert your Data Availability Statement in the box below. Please see the data reporting section of our submission guidelines for instructions on what you need to include in your Data Availability Statement. This statement will be typeset if the manuscript is accepted for publication. *PLOS allows rare exemptions to address legal and ethical concerns. If you have legal or ethical restrictions, please use the box below to detail these in full sentences for the Journal team to consider.*	All code and output is available via Zenodo (doi:10.5281/zenodo.7139978) under a Creative Commons license.

Pangeo-Enabled ESM Pattern Scaling (PEEPS): A customizable dataset of emulated Earth System Model output

Ben Kravitz[1,2,*] and Abigail Snyder[3]

1 Department of Earth and Atmospheric Sciences, Indiana University, Bloomington, IN, USA
2 Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, WA, USA
3 Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD, USA

These authors contributed equally to this work.
* bkravitz@iu.edu

Abstract

Emulation through pattern scaling is a way of rapidly producing climate fields (like temperature or precipitation) from existing Earth System Model (ESM) output that, while inaccurate, is often useful for a variety of downstream purposes. Conducting pattern scaling has historically been a laborious process, in large part due to the increasing volume of ESM output. Here we introduce Pangeo-Enabled ESM Pattern Scaling (PEEPS), a Jupyter notebook-based repository of pattern scaled output from CMIP6. Because all source data for PEEPS is cloud-based, users do not need to download and house the ESM output, meaning that emulated patterns are essentially on demand. Validating PEEPS on the CMIP6 archive for temperature, precipitation, and near-surface relative humidity, pattern scaling performs well over a variety of future scenarios except for regions in which there are strong, potentially nonlinear climate feedbacks. Although pattern scaling is normally conducted on annual mean ESM output, it works equally well on monthly mean ESM output. We identify several downstream applications for PEEPS, including impacts assessment and evaluating climate system uncertainties.

Author summary

We produce a dataset that uses pattern scaling, a common method of emulating climate models. Our dataset is built on the Pangeo CMIP6 archive, which has the advantage that we don't need to actually download the climate model output. Here we demonstrate the utility of our dataset, called Pangeo-Enabled ESM Pattern Scaling (PEEPS). Pattern scaling on the CMIP6 archive performs well except for a few areas marked by nonlinearity, where we wouldn't expect pattern scaling to do well. Many pattern scaling approaches look at annual mean temperature and precipitation changes; we look at these as well, but also relative humidity, and we extend pattern scaling to monthly output instead of annual mean output. Monthly mean pattern scaling performs as well as annual mean, and relative humidity performs quite well. The dataset, which is encapsulated in a Jupyter notebook, is flexible and can be extended to multiple scenarios and multiple variables, as long as they are in the Pangeo-accessible archive.

Introduction

Earth System Models (ESMs) are a standard tool for projecting future climate changes in response to forcing. The Coupled Model Intercomparison Project (CMIP) leads this effort and is one of the largest activities in climate science. [1] Nevertheless, ESMs are computationally expensive, which limits the research community's ability to explore various kinds of uncertainties. For example, only a limited number of future scenarios of climate change are feasible to explore, especially in the context of a multi-model intercomparison. [2] Also, exploring internal variability or rare events requires large ensembles of simulations (e.g., [3]) which are similarly computationally prohibitive for many modeling groups and are difficult to explore for a wide variety of scenarios.

Climate model emulators are simple tools that are trained on ESM outputs that produce less precise or less accurate output (as compared to the underlying ESM) but at a tiny fraction of the computational cost. Different emulation approaches often focus on emulating specific aspects of ESM outputs rather than the output as a whole. These approaches often take into account how finescale the emulated outputs (temporally or spatially) need to be and which variables are required for applications. Dorheim et al. [4] calibrate the parameters of a global carbon cycle model to reproduce all four CMIP5 realizations run by a specific ESM from a single parametrization, allowing confident generation of time series of global average temperature anomalies in novel scenarios. Other approaches focus on emulating the internal variability of a particular ESM from a one or two variable subset of outputs to produce numerous ensemble members in seconds or minutes. [5, 6] These areas have recently received an increased amount of attention, in which emulators built on existing scenarios are used to provide low-cost projections for other scenarios [7–10] or exploring sources of error in emulating indices of extreme events. [11] There are several commonly used methods of building emulators. Some include using a simple climate model that reproduces gross Earth system behavior (e.g., Hector; [4]) (https://jgcri.shinyapps.io/HectorUI/), assuming an underlying functional form as in an impulse response model, [12, 13] or training a deep learning model on ESM output. [14, 15]

Many of these methods are effective at emulating global average or large-scale change. For more regional impacts, one needs a way of relating those global changes to finer scales. A simple, commonly used way of doing this is via pattern scaling, in which the spatial pattern of change is assumed to scale linearly with global mean temperature. [16, 17] While technically inaccurate, pattern scaling has been shown to be reasonably effective at reproducing large-scale mean field CMIP5 behavior for temperature and precipitation over a range of scenarios [7, 18] and for emulating indices of extreme events. [11]

Lynch et al. [18] built a pattern scaling archive for available CMIP5 output (https://github.com/JGCRI/CMIP5_patterns) that can be used by a variety of stakeholders, such as impacts modelers [19] and multisectoral dynamics modeling efforts exploring the co-evolution of the integrated human-Earth system. [20] Using the options available to them, Lynch et al. downloaded climate model output from 12 of the 41 available models for the historical period and four future scenarios (the Representative Concentration Pathways), [21] processed the output locally to create a library of patterns, and then hosted that library in a repository. Put simply, this required a lot of time and labor. CMIP6 has over 100 participating models running even more scenarios than in CMIP5, and many models that previously participated in CMIP5 have undergone improvements, including increased spatial resolution. The "climate services" approach undertaken by Lynch et al. is not sustainable.

Here we introduce Pangeo-Enabled ESM Pattern Scaling (PEEPS), a Jupyter notebook to conduct pattern scaling on the CMIP6 archive using python packages produced by the Pangeo community. [22] By interfacing with Pangeo, no downloading of

ESM output is required, which substantially reduces barriers to stakeholders who wish to create a library of pattern scaled output. With PEEPS, one simply selects the scenarios, models, and variables one wishes to analyze, and the Jupyter notebook will produce NetCDF-formatted pattern scaling output for each scenario/model/variable combination directly in the user's storage space. We further host the files examined here in a Zenodo repository [23] for direct download by users uninterested in running Python code themselves. Because of Pangeo, the dataset and the PEEPS Jupyter notebook are effectively the same thing, offering more flexibility to researchers for accessing these patterns in the way that works best for them. It also provides a head start for those interested in customizing the notebook to explore other patterns than those described here.

In this manuscript we document PEEPS and the resulting data set of CMIP6 patterns for the research community. Section 2 describes the pattern scaling method and ESM output that we analyzed. Section 3 produces a validation of PEEPS output (pattern scaling) for the CMIP6 archive, exploring scaling on both annual mean and monthly mean output. Section 4 describes PEEPS itself, including options, formatting of the output files, and performance. Section 5 provides a discussion of the potential uses of this package.

Methods

There are numerous methods of conducting pattern scaling, including calculating a linear change from the underlying data [6, 7, 17] or assuming an underlying functional form [24] or model relating global temperature to local changes. Here we conduct ordinary linear regression in each grid cell of the ESM data to obtain a slope ($m(\mathbf{x})$) and an intercept ($b(\mathbf{x})$) value for each model and each scenario separately, where $\mathbf{x}$ indicates the spatial dimension. The emulated (pattern scaled) output at any grid point ($T(\mathbf{x}, t)$) at any time t is then obtained from the global mean temperature ($\bar{T}(t)$) via

$$T(\mathbf{x}, t) = m(\mathbf{x}) \cdot \bar{T}(t) + b(\mathbf{x}) \tag{1}$$

If there are multiple ensemble members, the regression is computed for each ensemble member individually, and the reported slope and intercept is the ensemble average. Figure 1 shows a sample of pattern scaling for precipitation.

Fig 1. An illustrative example of pattern scaling for annual mean precipitation in a single model (in this case, CESM2-WACCM) over the historical period. Top left shows the slopes (precipitation per degree global warming) of the regression lines at each grid point (mm day^{-1} K^{-1}). Top right shows the intercepts of the regression lines (mm day^{-1}). Bottom left shows the generated precipitation ($P(\mathbf{x})$) time series calculated as $P(\mathbf{x}) = m(\mathbf{x})\overline{T} + b(\mathbf{x})$ where $m(\mathbf{x})$ is slope, $\overline{T}$ is global mean temperature averaged over the last 20 years of simulation (for historical XX-YY), $b(\mathbf{x})$ is intercept, and $\mathbf{x}$ is the spatial dimension. Bottom right shows the residual (mm day^{-1}) of generated minus the actual model output, again averaged over the last 20 years of simulation.

Pattern scaling is often conducted using annual mean climate model output. In addition to that, we explore pattern scaling using monthly mean output. We explore error in both of these temporal resolutions for surface air temperature (tas), precipitation (pr), and relative humidity (hurs). The Pangeo architecture enables analysis of as many other variables and scenarios as are available in the CMIP6 archive (and, likely, archives for future projects). Table 1 describes the number of models we evaluated for each scenario/variable combination. The scenarios we explore are the latest future projections used in CMIP [2] and are the subject of in-depth discussions in the latest assessment report by the Intergovernmental Panel on Climate Change. [25]

Table 1. The number of models available via Pangeo for each scenario/variable combination. 60 models had their output available in Pangeo at the time of our analysis.

	tas	pr	hurs
historical	38	37	30
ssp126	25	26	20
ssp245	24	26	20
ssp370	20	22	18
ssp434	6	6	6
ssp585	28	27	22

Validation

Annual Mean Pattern Scaling

As a metric of validation, Figures 2–4 show a measure of how large the annual mean residual is compared to the underlying data. At each grid point, for each model and scenario, the average residual over the last 20 years of simulation (1995–2014 for historical; 2081–2100 for the others) is computed. This value is then compared to the standard deviation of the model output over that 20-year period for each model and scenario. Values in Figures 2–4 indicate the percent of models for each scenario for which the residual is within one standard deviation of the model output. The standard deviation is computed assuming each year within that 20 year period is independent, which is an erroneous assumption that results in a more conservative estimate (smaller standard deviation).

Fig 2. In each grid box, percentage of models where the surface air temperature (tas) residual for annual mean pattern scaling is within one standard deviation of the model output.

Fig 3. As in Figure 2 but for precipitation (pr).

Fig 4. As in Figure 2 but for surface relative humidity (hurs).

For the most part, pattern scaling appears to be an effective means of replicating annual mean climate model output for a wide variety of models and scenarios, with many regions showing that for 90+% of the models the residuals are within one standard deviation of the baseline. Figure 5 further illustrates this, showing the number of scenarios for which the panels displayed in Figures 2–4 are at least 90%. For surface air temperature, 85.6% of all area (96.3% of land area) has all six scenarios that meet this criterion. For precipitation the analogous figures are 98.5% of all area (98.9% of land area), and for relative humidity 78.4% of all area (78.5% of land area). Relative humidity shows worse performance than the other two models, as well as large inter-scenario differences.

There are regions where, for some scenarios and some models, pattern scaling introduces error. This could be due to a few reasons:

Fig 5. An aggregate of Figures 2–4 illustrating for each grid point the number of scenarios for which the value (percent of models within one standard deviation of baseline) is at least 90%.

1. Pattern scaling doesn't work for that region. This could be due to some response in the climate system that results in a nonlinear relationship between global mean temperature and the local response (for example, feedbacks leading to Arctic Amplification); nonmonotonicity in global mean temperature (for example, ssp126 has an overshoot, which could affect regressions); or a low trend in global mean temperature (again, possible under ssp126) resulting in a poor linear fit.

2. The baseline may have low variability. This would result in a greater probability of exceeding one standard deviation.

3. There is a low number of models (as is the case for ssp434), so having a high residual for even one model can result in a large change in Figures 2–4.

Figures 6 and 7 provide further insight into potential sources of error in pattern scaling for the three variables considered here. For all variables in all scenarios, the inter-quartile range never exceeds 0.5 standard deviations, indicating that any errors tend to be due to a smaller percentage of models rather than general features of pattern scaling. Among those models that exceed the inter-quartile range, relative humidity tends to have greater error than the other two variables, and high latitudes tend to have more error than other regions. Historical and ssp245 tend to have the least error, ssp585 has the most error, and ssp434 has too few models to ascertain a robust comparison with other scenarios. The greatest error tends to vary across variables and scenarios, that is, there is no group of models that performs poorly in all cases. If a model/scenario/variable combination has error in one spatial region, it tends to have high error in the other regions; based on the results in Figures 2–5, this result is likely dominated by local features with high residuals. Nevertheless, there are few values in Figure 7 that exceed one standard deviation, and they are almost entirely found in ssp126, ssp370, and ssp585. Further spatial analysis (not pictured) indicates that indeed the mean residuals on a gridpoint basis are quite small, further reinforcing that on average pattern scaling does well except for a few outliers. A notable exception includes tropical precipitation, which is a known difficulty for pattern scaling due to nonlinear behavior of intense precipitation. [7]

Fig 6. Box plot of root mean square error (RMSE) of pattern scaling, calculated as the number of standard deviations the generated output is from the actual model output (calculated over the last 20 years of simulation), for each scenario (panels) and for temperature (tas), precipitation (pr), and relative humidity (hurs) in a variety of regions. Red lines indicate the median model, blue boxes indicate the inter-quartile range, and whiskers indicate the full model range. Because so few models participated in ssp434, we show the RMSE values for each model.

Fig 7. Heatmap of the number of standard deviations (colors) the generated output is from the actual model output for each model in each scenario (panels; calculated from the last 20 years of each simulation) for temperature (tas), precipitation (pr), and relative humidity (hurs) in a variety of regions. White squares (marked by NaN) indicate that there is no model output available for that model/variable combination on Pangeo.

Figures 2–5 have some areas where pattern scaling performance is consistently worse than others. In addition to the tropics, these areas include the North Atlantic, the high latitudes (predominantly the Arctic), the Southern Ocean, and oceanic areas associated with eastern boundary currents. These are all areas associated with feedback-dominated behavior where pattern scaling might not be expected to perform well: the "warming hole" in the North Atlantic associated with the Atlantic Meridional Overturning Circulation and cloud feedbacks [26]; Arctic amplification associated with strong feedbacks like the ice albedo feedback, lapse rate feedback, and changes in atmospheric and oceanic heat transport [27, 28]; cloud feedbacks in the Southern Ocean [29]; and persistent marine stratocumulus decks off the western coasts of continents. [30] Regarding the Atlantic Meridional Overturning Circulation, the Southern Ocean, and marine stratocumulus decks, these areas are not over land so are not directly relevant for many impacts models, for example agriculture or hydrological models. While these regions are important in general, one would not presume that pattern scaling is an effective tool for studying these sorts of complex processes and feedbacks, so it could be argued that pattern scaling performance in these regions is less important. The high latitudes are important for many impacts studies, notably sea level rise; due to substantial uncertainties in feedback strength at the high latitudes resulting in large model spread, [31] we urge caution in using this package to evaluate impacts of high latitude change. Figures 2-5 do indicate, however, that even at these latitudes, there are many ESMs amenable to pattern scaling in many scenarios for many variables.

The regression approach undertaken here is not well suited to capturing interannual variability (e.g., the El Niño Southern Oscillation or the North Atlantic Oscillation). Regions strongly affected by interannual variability are unlikely to show major sources of error if the oscillation period is substantially smaller than 20 years (the averaging period of our results). If the oscillation changes under climate change such that one mode becomes more dominant than the other, the regression should be able to capture those changes, similarly resulting in low error. A potential caveat is if the oscillation has a longer period than can be captured in the 20-year average (such as the Pacific Decadal Oscillation). [32]

For temperature, pattern scaling on ssp126 is worse-performing than the other scenarios, and for precipitation and relative humidity, pattern scaling on the high forcing scenarios (ssp370 and ssp585) is worse than for the others. ssp126 has little global mean temperature change, what change it does have is nonmonotonic, [8] so it is difficult to obtain a confident regression slope. Nevertheless, because the amount of climate change is so small in this scenario, internal variability is likely a larger source of uncertainty than climate change, so perhaps pattern scaling is not even necessary for this scenario. Under ssp370 and ssp585, because the amount of global warming is greater, there is greater excitation of temperature-related feedbacks, which is more likely to result in behavior that cannot be captured by linear regression.

Monthly Mean Pattern Scaling

Most pattern scaling is conducted using annual mean variables, but climate model output is often available as monthly averages. Monthly averages also often have more utility in many impacts models. This begs the question as to whether pattern scaling on monthly output will work. There is strong reason to suspect that it would, as the seasonal cycle is by far the most dominant source of variability in monthly output, and removing that cycle is standard procedure for creating climatologies on which residuals are calculated.

Figures 8-10 show a monthly climatology of residuals calculated over six regions of the globe. For most models and scenarios, monthly residuals are small and do not show strong seasonal variations. High latitude temperature is a notable exception, although the mean model/scenario residual is approximately zero. For precipitation and relative humidity, a few models have residuals that separate from the model pack but with few notable seasonal characteristics, indicating that higher residuals are due to the model/scenario combination itself rather than anything intrinsic to monthly pattern scaling. Figure S1 shows a histogram comparing the residual for annual pattern scaling with the averaged residuals for monthly pattern scaling. For all three variables, the differences are several orders of magnitude smaller than the residual fields and are approximately normally distributed. This confirms our hypothesis that pattern scaling on monthly output does indeed work and, in many cases, is indistinguishable in performance from annual mean pattern scaling. Figures S2–S22 show spatial patterns of the residuals, analogous to Figures 2–5. For temperature there are few obvious differences between monthly residuals, and the patterns of success in pattern scaling reflect those of annual mean pattern scaling. There are some seasonal shifts in the residuals for precipitation based on seasonal variations in tropical precipitation; again, this is expected. [7] For relative humidity, differences between months are small and, like temperature, resemble the annual mean residual patterns.

Fig 8. Climatology of monthly climatology surface air temperature (°C) residuals (generated output minus actual model output) calculated over six different regions of the globe. Each line indicates a model/scenario combination (values shown are averages over the last 20 years of simulation), and the thick black line indicates the average over all models/scenarios. x-axis indicates the month of the climatology.

Fig 9. As in Figure 8 but for precipitation (mm day^{-1}).

Fig 10. As in Figure 8 but for surface relative humidity (%).

Data Availability: Pattern Scaling Code and Repository

For each model/scenario/variable combination, the annual mean pattern scaling code produces up to three output files, depending upon the options selected:

1. The patterns for pattern scaling: a NetCDF file containing the slope and intercept from the regression.
2. A NetCDF file containing the timeseries of global average temperature.
3. [optional] A NetCDF file containing the residuals from pattern scaling: a timeseries of ESM output minus the reconstruction from pattern scaling at each grid point.

If the option to produce monthly pattern scaling is selected, items 1 and 3 will produce 12 files each, one per month. If a particular combination has multiple ensemble members, the code will average all ensemble members and then conduct pattern scaling on the average.

Of critical importance is that PEEPS is easy to use and conforms to a standard. The code is written in python and is contained in a fully documented Jupyter notebook. The output uses NetCDF formatting with a structure similar to that found in CMIP6 models, including inherited metadata attributes. In doing so, access to the code and its outputs individually is not a barrier to use for downstream research, such as impacts modeling.

In the notebook itself, the user-editable options are the list of experiments and the list of variables, formatted to CMIP6 standards. The script will then automatically generate a list of all available output (all models) with those specifications and process it for annual or monthly averages, depending upon the block of code the user is running. Within each block there are options regarding whether to fit an intercept (if false, the intercept is assumed to be zero) and whether to save the residuals (item 3 above). The repository includes a tutorial and examples.

It is difficult to benchmark performance because to some degree this depends upon the user's computer. On a laptop, each scenario/variable combination completed within approximately two hours.

Utility

The patterns output by PEEPS for monthly and annual climate data do not, of course, emulate any aspect of the target ESM's internal variability; this can be important for a variety of impacts, including extreme events. [11] While this may be a disadvantage on decadal or shorter timescales, on longer timescales uncertainties tend to be dominated by structural uncertainties (whether models include specific processes) or scenario uncertainties, [33] where mean field emulation is advantageous because it is faster to conduct and easier to calibrate. A ready example of where mean field validation could be useful lies in the flexibility of PEEPS. Although the global mean temperature computation and the slopes/intercepts are by default computed from the same underlying ESM data, this does not have to be the case when using the patterns. One could easily provide a different time series of global mean temperature change, for example, computed by a simple climate model. If this time series is calibrated to replicate global mean temperature behavior of a particular ESM, [4] then one can emulate time series from that ESM under different future scenarios that may not have been run with the full ESM.

Mean field emulation is also valuable for long-term impacts studies, especially if one wishes to explore the consequences of different future scenarios. As an example, an activity in the Agriculture Model Intercomparison Project (AgMIP) [34] involves using global gridded crop models to produce emulators of climatological-mean yield response to uniform perturbations in growing-season mean temperature and precipitation. [35] These yield response emulators presently do not account for variability, meaning the linear patterns presented here for monthly temperature and precipitation can be used to rapidly generate gridded monthly data that can be directly translated into impacts assessments. The areas where pattern scaling performs less well lie over ocean grid cells, which are not relevant for agricultural yields. An additional advantage of PEEPS for impacts modeling is the ability to incorporate numerous ESMs. The spatial patterns for each ESM differ; by sampling this space, one can capture a range of uncertainties in impacts.

There have recently been efforts to create novel scenarios directly from ESM output using time sampling, [36] statistical relationships, [6] or machine learning. [37] By comparing those results with output from PEEPS (or pattern scaling more generally), one can effectively diagnose the effects of linear and nonlinear processes, using emulation to gain deeper understanding of the sources of uncertainties in projections of climate change.

Conclusions and Next Steps

We confirm (unsurprisingly) that, like for CMIP5, pattern scaling is an effective means of obtaining global-to-local relationships in CMIP6. Also, as hypothesized, pattern scaling on monthly output works well, and for temperature and precipitation has nearly identical performance to annual mean pattern scaling. Pattern scaling on relative humidity also works quite well for both annual mean and monthly mean. Model performance can vary across scenarios, variables, and regions (Figure 7), indicating that when using results from pattern scaling one may want to select models that work better for one's purposes. Under the PEEPS framework, exploring how well pattern scaling works for other variables would be nearly trivial.

As discussed in the Introduction, two of the difficulties posed by the large computational expense of ESMs are generating numerous realizations and exploring multiple scenarios. PEEPS could be combined with other efforts to directly address these problems in an accessible way. PEEPS could be used to do the pre-processing for these packages, providing the input from which the multiple realizations could be generated. Similarly, the global mean temperature timeseries used in pattern scaling does not need to be generated from the underlying ESM data. The simple climate model Hector has been calibrated to output a global mean temperature timeseries that matches the output of a variety of ESMs under different scenarios. [4] As such, Hector (or another similarly calibrated simple climate model) could be used to generate a global mean temperature timeseries for a scenario that any particular ESM has not simulated, and PEEPS could then be used to augment that global mean timeseries into a spatially resolved timeseries. (Note that there may be best practices in terms of which scenario should be used to build the patterns. [8, 10]).

It is customary to view a repository as a collection of files. While we have provided PEEPS-generated output calculated from the ESMs, this output is essentially an example. Because the Jupyter notebook is so flexible, any files in the repository can be generated on-demand (after a short waiting period while the patterns are being calculated), so there is less of a need to provide files directly. PEEPS is an example of a new way of viewing climate model output repositories: that the repository and the script to generate the files are the same thing.

A frozen version of the code and output is available for download from Zenodo, doi:10.5281/zenodo.7139978. An institutional repository of the code, which will incorporate future updates, is available at `https://github.com/JGCRI/linear_pattern_scaling`.

Supporting information

S1 Fig. A histogram (count = number of model grid points) of the difference between the residuals (generated minus actual, averaged over the last 20 years of simulation for each scenario) for annual mean pattern scaling and the average of the residuals for all 12 months of monthly pattern scaling. Units are °C, mm/day, and %, respectively.

S2 Fig. In each grid box, the percentage of models where the surface air temperature (tas) residual in the historical simulation for monthly mean pattern scaling is within one standard deviation of the model output. See description in Section 3.1 in the main text for further details.

S3 Fig. As in Figure S2 but for SSP1-2.6.

S4 Fig. As in Figure S2 but for SSP2-4.5.

S5 Fig. As in Figure S2 but for SSP3-7.0.

S6 Fig. As in Figure S2 but for SSP4-3.4.

S7 Fig. As in Figure S2 but for SSP5-8.5.

S8 Fig. As in Figure S2 but for precipitation.

S9 Fig. As in Figure S8 but for SSP1-2.6.

S10 Fig. As in Figure S8 but for SSP2-4.5.

S11 Fig. As in Figure S8 but for SSP3-7.0.

S12 Fig. As in Figure S8 but for SSP4-3.4.

S13 Fig. As in Figure S8 but for SSP5-8.5.

S14 Fig. As in Figure S2 but for surface relative humidity.

S15 Fig. As in Figure S14 but for SSP1-2.6.

S16 Fig. As in Figure S14 but for SSP2-4.5.

S17 Fig. As in Figure S14 but for SSP3-7.0.

S18 Fig. As in Figure S14 but for SSP4-3.4.

S19 Fig. As in Figure S14 but for SSP5-8.5.

S20 Fig. An aggregate of Figures S2-S7 illustrating for each grid point the number of scenarios for which the value (percent of models within one standard deviation of the baseline temperature) is at least 90%.

S21 Fig. An aggregate of Figures S8-S13 illustrating for each grid point the number of scenarios for which the value (percent of models within one standard deviation of the baseline precipitation) is at least 90%.

S22 Fig. An aggregate of Figures S14-S19 illustrating for each grid point the number of scenarios for which the value (percent of models within one standard deviation of the baseline surface air temperature) is at least 90%.

Acknowledgments

This work was conducted with the support of the U.S. Department of Energy, Office of Science, as part of research in MultiSector Dynamics, Earth and Environmental System Modeling Program. Support for B.K. was provided in part by the Indiana University Environmental Resilience Institute and the Prepared for Environmental Change Grand Challenge initiative. The Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle Memorial Institute under contract DE-AC05-76RL01830. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. Data analysis was performed using the Pangeo platform, which was supported by NSF Earthcube Award 1740648.

References

1. Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development. 2016;9(5):1937–1958. doi:10.5194/gmd-9-1937-2016.

2. O'Neill BC, Tebaldi C, van Vuuren DP, Eyring V, Friedlingstein P, Hurtt G, et al. The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. Geoscientific Model Development. 2016;9(9):3461–3482. doi:10.5194/gmd-9-3461-2016.

3. Kay JE, Deser C, Phillips A, Mai A, Hannay C, Strand G, et al. The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability. Bulletin of the American Meteorological Society. 2015;96(8):1333–1349. doi:10.1175/BAMS-D-13-00255.1.

4. Dorheim K, Link R, Hartin C, Kravitz B, Snyder A. Calibrating Simple Climate Models to Individual Earth System Models: Lessons Learned From Calibrating Hector. Earth and Space Science. 2020;7(11). doi:10.1029/2019EA000980.

5. Link R, Snyder A, Lynch C, Hartin C, Kravitz B, Bond-Lamberty B. Fldgen v1.0: an emulator with internal variability and space–time correlation for Earth system models. Geoscientific Model Development. 2019;12(4):1477–1489. doi:10.5194/gmd-12-1477-2019.

6. Beusch L, Gudmundsson L, Seneviratne SI. Emulating Earth system model temperatures with MESMER: from global mean temperature trajectories to grid-point-level realizations on land. Earth System Dynamics. 2020;11(1):139–159. doi:10.5194/esd-11-139-2020.

7. Kravitz B, Lynch C, Hartin C, Bond-Lamberty B. Exploring precipitation pattern scaling methodologies and robustness among CMIP5 models. Geoscientific Model Development. 2017;10(5):1889–1902. doi:10.5194/gmd-10-1889-2017.

8. Wells CD, Jackson LS, Maycock AC, Forster PM. Understanding pattern scaling errors across a range of emissions pathways. Earth system change: climate

prediction; 2022. Available from: https://egusphere.copernicus.org/preprints/2022/egusphere-2022-914/.

9. Alexeeff SE, Nychka D, Sain SR, Tebaldi C. Emulating mean patterns and variability of temperature across and within scenarios in anthropogenic climate change experiments. Climatic Change. 2018;146(3-4):319–333. doi:10.1007/s10584-016-1809-8.

10. Tebaldi C, Knutti R. Evaluating the accuracy of climate change pattern emulation for low warming targets. Environmental Research Letters. 2018;13(5):055006. doi:10.1088/1748-9326/aabef2.

11. Tebaldi C, Armbruster A, Engler HP, Link R. Emulating climate extreme indices. Environmental Research Letters. 2020;15(7):074006. doi:10.1088/1748-9326/ab8332.

12. Schwarber AK, Smith SJ, Hartin CA, Vega-Westhoff BA, Sriver R. Evaluating climate emulation: fundamental impulse testing of simple climate models. Earth System Dynamics. 2019;10(4):729–739. doi:10.5194/esd-10-729-2019.

13. MacMartin DG, Kravitz B. Dynamic climate emulators for solar geoengineering. Atmospheric Chemistry and Physics. 2016;16(24):15789–15799. doi:10.5194/acp-16-15789-2016.

14. Weber T, Corotan A, Hutchinson B, Kravitz B, Link R. Technical note: Deep learning for creating surrogate models of precipitation in Earth system models. Atmospheric Chemistry and Physics. 2020;20(4):2303–2317. doi:10.5194/acp-20-2303-2020.

15. Ayala A, Drazic C, Hutchinson B, Kravitz B, Tebaldi C. Loosely Conditioned Emulation of Global Climate Models With Generative Adversarial Networks. arXiv. 2021;doi:10.48550/ARXIV.2105.06386.

16. Santer B, Wigley TML, Schlesinger ME, Mitchell JFB. Developing climate scenarios from equilibrium GCM results. Max Planck Institute for Meteorology; 1990. 47.

17. Tebaldi C, Arblaster JM. Pattern scaling: Its strengths and limitations, and an update on the latest model simulations. Climatic Change. 2014;122(3):459–471. doi:10.1007/s10584-013-1032-9.

18. Lynch C, Hartin C, Bond-Lamberty B, Kravitz B. An open-access CMIP5 pattern library for temperature and precipitation: description and methodology. Earth System Science Data. 2017;9(1):281–292. doi:10.5194/essd-9-281-2017.

19. Ruane AC, Rosenzweig C, Asseng S, Boote KJ, Elliott J, Ewert F, et al. An AgMIP framework for improved agricultural representation in integrated assessment models. Environmental Research Letters. 2017;12(12):125003. doi:10.1088/1748-9326/aa8da6.

20. Calvin K, Bond-Lamberty B. Integrated human-earth system modeling—state of the science and future directions. Environmental Research Letters. 2018;13(6):063006. doi:10.1088/1748-9326/aac642.

21. Meinshausen M, Smith SJ, Calvin K, Daniel JS, Kainuma MLT, Lamarque JF, et al. The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. Climatic Change. 2011;109:213. doi:10.1007/s10584-011-0156-z.

22. Odaka TE, Banihirwe A, Eynard-Bontemps G, Ponte A, Maze G, Paul K, et al. The Pangeo Ecosystem: Interactive Computing Tools for the Geosciences: Benchmarking on HPC. In: Juckeland G, Chandrasekaran S, editors. Tools and Techniques for High Performance Computing. vol. 1190. Cham: Springer International Publishing; 2020. p. 190–204. Available from: http://link.springer.com/10.1007/978-3-030-44728-1_12.

23. Kravitz B, Snyder AC. Pangeo-Enabled ESM Pattern Scaling (PEEPS): A customizable dataset of emulated Earth System Model output [Dataset]; 2022. Available from: https://doi.org/10.5281/zenodo.7139978.

24. Castruccio S, McInerney DJ, Stein ML, Liu Crouch F, Jacob RL, Moyer EJ. Statistical Emulation of Climate Model Projections Based on Precomputed GCM Runs. Journal of Climate. 2014;27(5):1829–1844. doi:10.1175/JCLI-D-13-00099.1.

25. IPCC. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press; 2021.

26. Keil P, Mauritsen T, Jungclaus J, Hedemann C, Olonscheck D, Ghosh R. Multiple drivers of the North Atlantic warming hole. Nature Climate Change. 2020;10(7):667–671. doi:10.1038/s41558-020-0819-8.

27. Pithan F, Mauritsen T. Arctic amplification dominated by temperature feedbacks in contemporary climate models. Nature Geoscience. 2014;7(3):181–184. doi:10.1038/ngeo2071.

28. Hahn LC, Armour KC, Zelinka MD, Bitz CM, Donohoe A. Contributions to Polar Amplification in CMIP5 and CMIP6 Models. Frontiers in Earth Science. 2021;9:710036. doi:10.3389/feart.2021.710036.

29. Kim H, Kang SM, Kay JE, Xie SP. Subtropical clouds key to Southern Ocean teleconnections to the tropical Pacific. Proceedings of the National Academy of Sciences. 2022;119(34):e2200514119. doi:10.1073/pnas.2200514119.

30. Hill S, Ming Y. Nonlinear climate response to regional brightening of tropical marine stratocumulus: CLIMATE RESPONSE TO CLOUD BRIGHTENING. Geophysical Research Letters. 2012;39(15). doi:10.1029/2012GL052064.

31. Chylek P, Folland C, Klett JD, Wang M, Hengartner N, Lesins G, et al. Annual Mean Arctic Amplification 1970–2020: Observed and Simulated by CMIP6 Climate Models. Geophysical Research Letters. 2022;49(13). doi:10.1029/2022GL099371.

32. Newman M, Alexander MA, Ault TR, Cobb KM, Deser C, Di Lorenzo E, et al. The Pacific Decadal Oscillation, Revisited. Journal of Climate. 2016;29(12):4399–4427. doi:10.1175/JCLI-D-15-0508.1.

33. Hawkins E, Sutton R. The potential to narrow uncertainty in projections of regional precipitation change. Climate Dynamics. 2011;37(1-2):407–418. doi:10.1007/s00382-010-0810-6.

34. Rosenzweig C, Jones JW, Hatfield JL, Ruane AC, Boote KJ, Thorburn P, et al. The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. Agricultural and Forest Meteorology. 2013;170:166–182. doi:10.1016/j.agrformet.2012.09.011.

35. Franke JA, Müller C, Elliott J, Ruane AC, Jägermeyr J, Snyder A, et al. The GGCMI Phase 2 emulators: global gridded crop model responses to changes in CO₂, temperature, water, and nitrogen (version 1.0). Geoscientific Model Development. 2020;13(9):3995–4018. doi:10.5194/gmd-13-3995-2020.

36. Tebaldi C, Snyder A, Dorheim K. STITCHES: creating new scenarios of climate model output by stitching together pieces of existing simulations. Earth system change: climate scenarios; 2022. Available from: https://esd.copernicus.org/preprints/esd-2022-14/.

37. Barnes EA, Hurrell JW, Ebert-Uphoff I, Anderson C, Anderson D. Viewing Forced Climate Patterns Through an AI Lens. Geophysical Research Letters. 2019;46(22):13389–13398. doi:10.1029/2019GL084944.

Figure 1

Click here to access/download;Figure;figure1.png

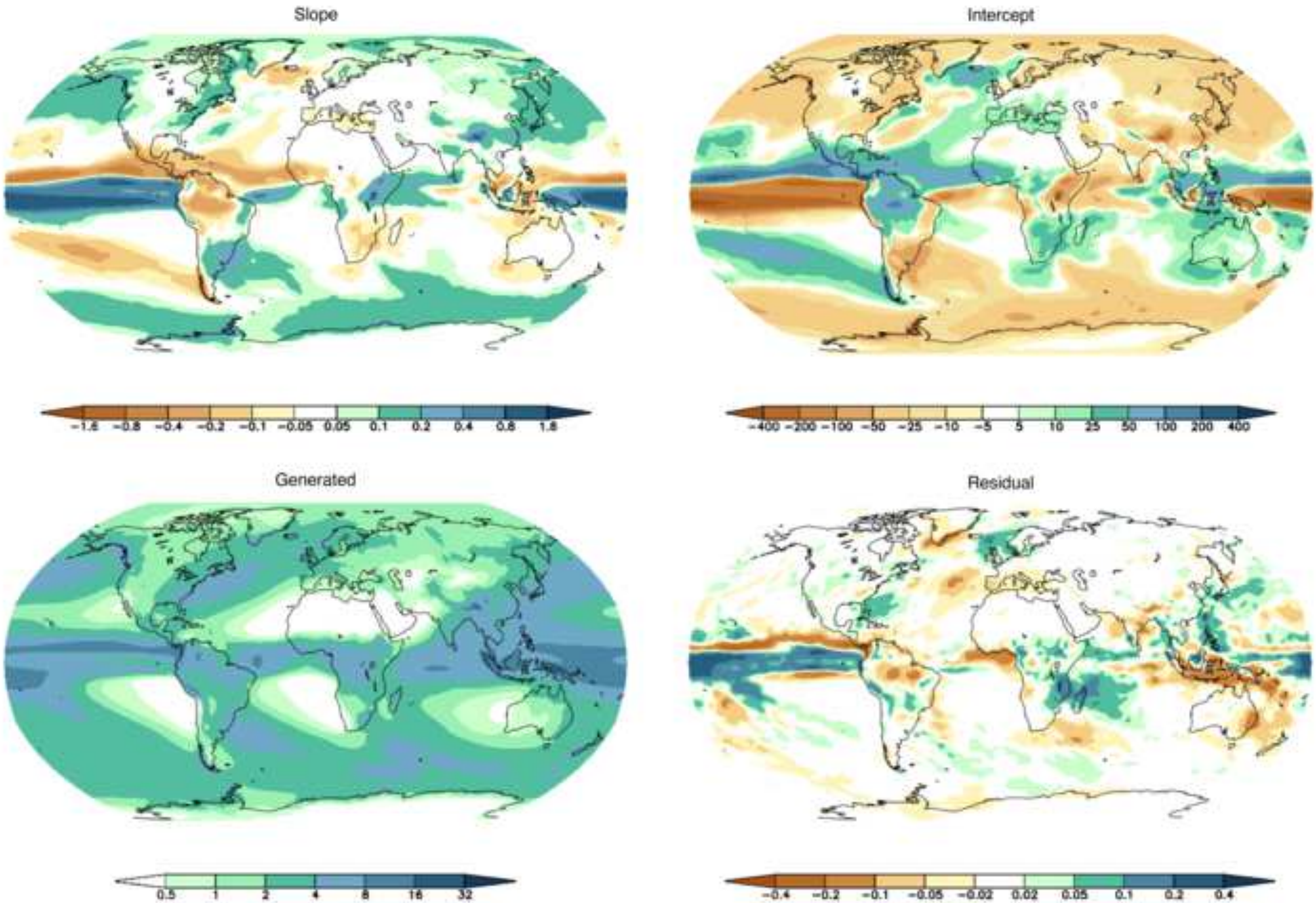


Figure 2

Click here to access/download;Figure;figure2.png

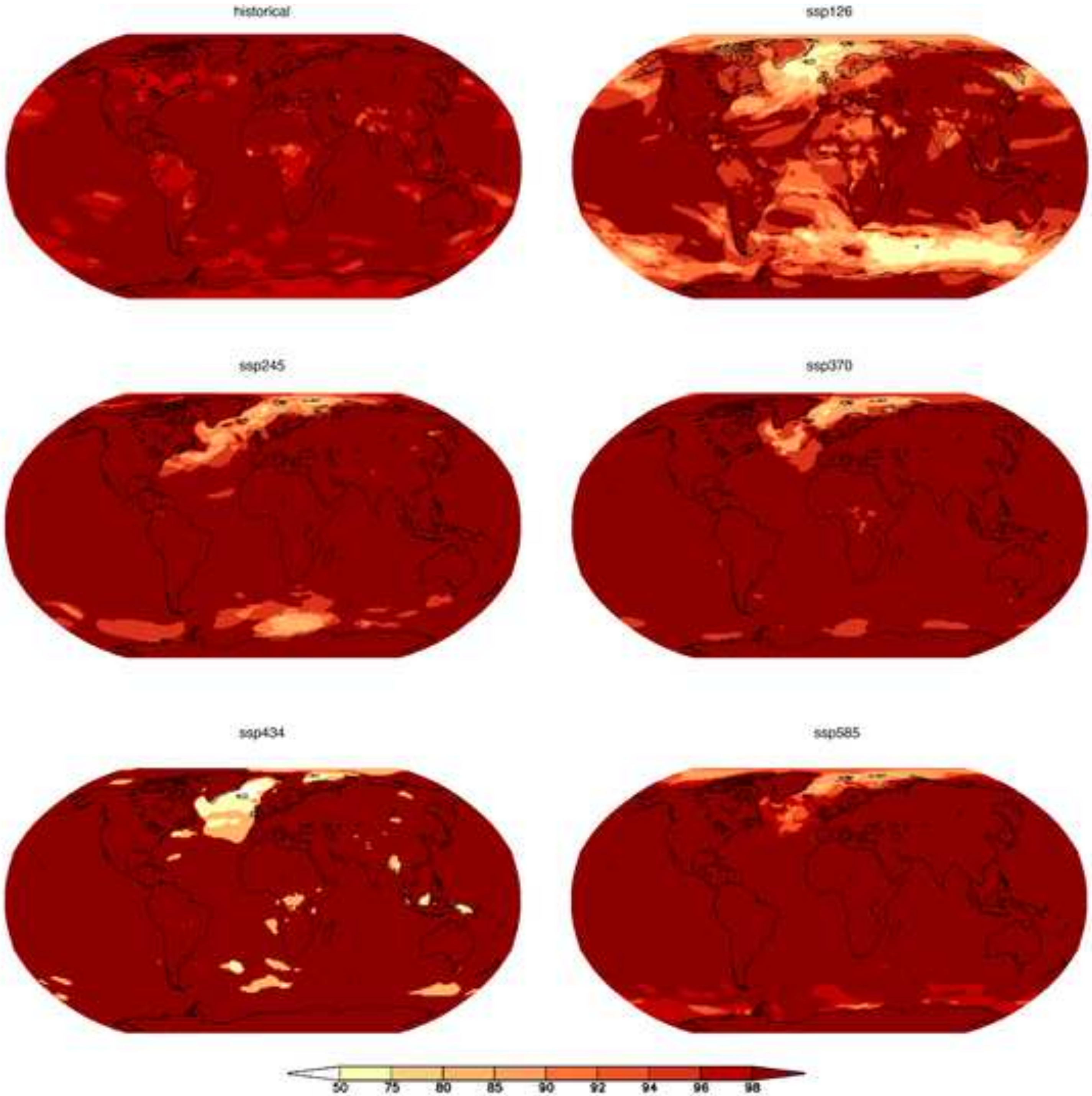


Figure 3

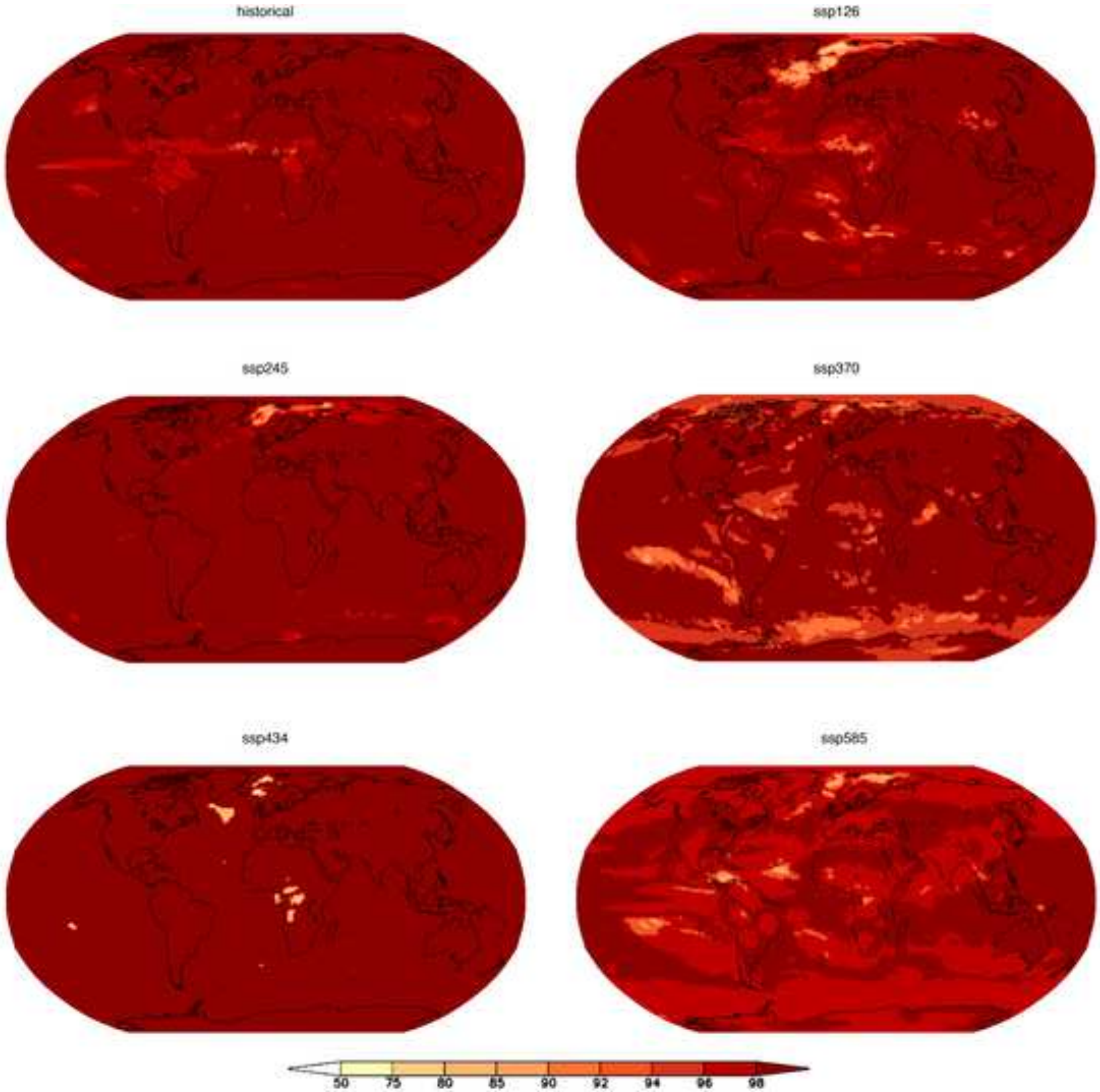

historical
ssp126
ssp245
ssp370
ssp434
ssp585
50
75
80
85
90
92
94
96
98

Figure 4

Click here to access/download;Figure;figure4.png

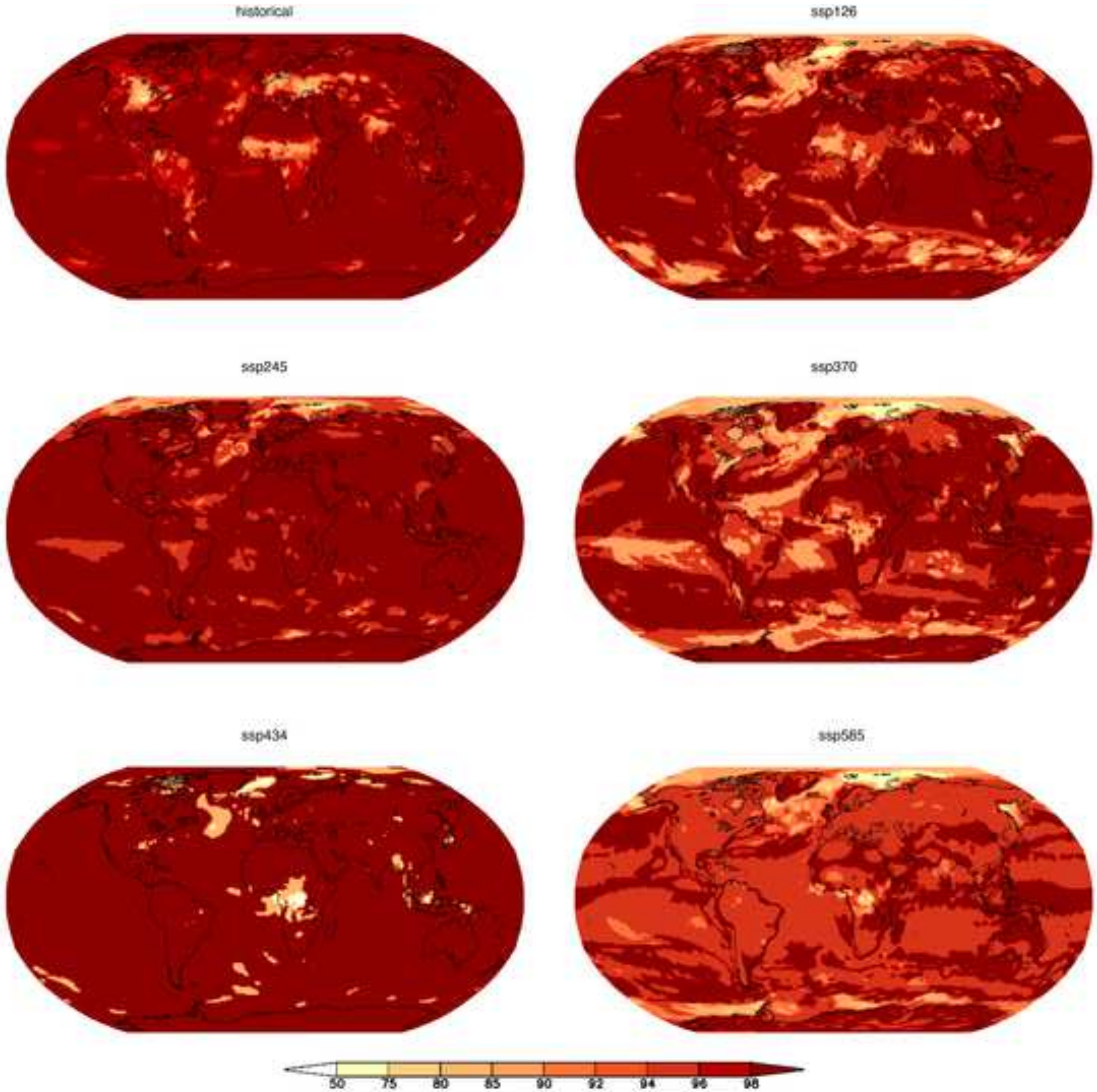


Figure 5

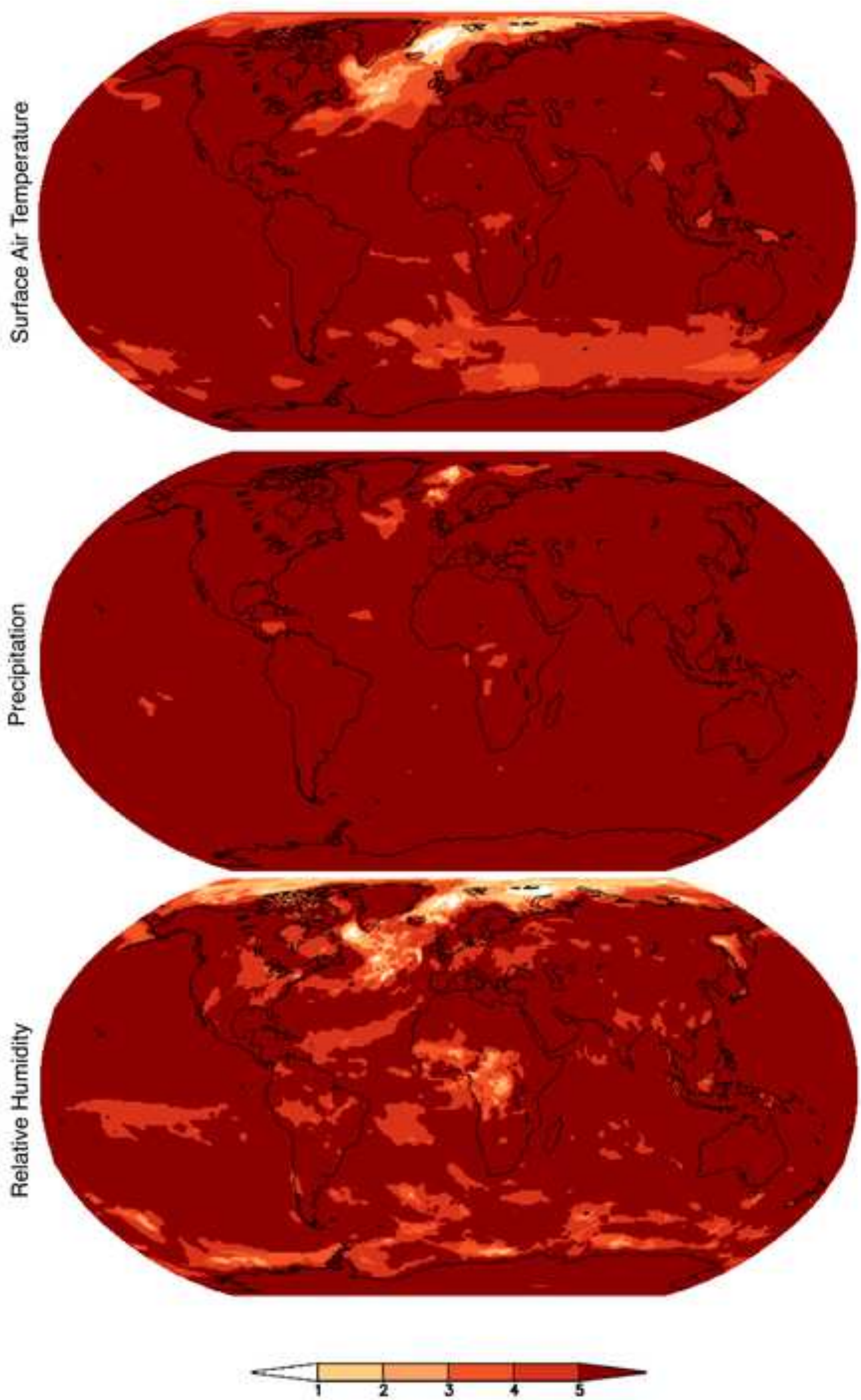
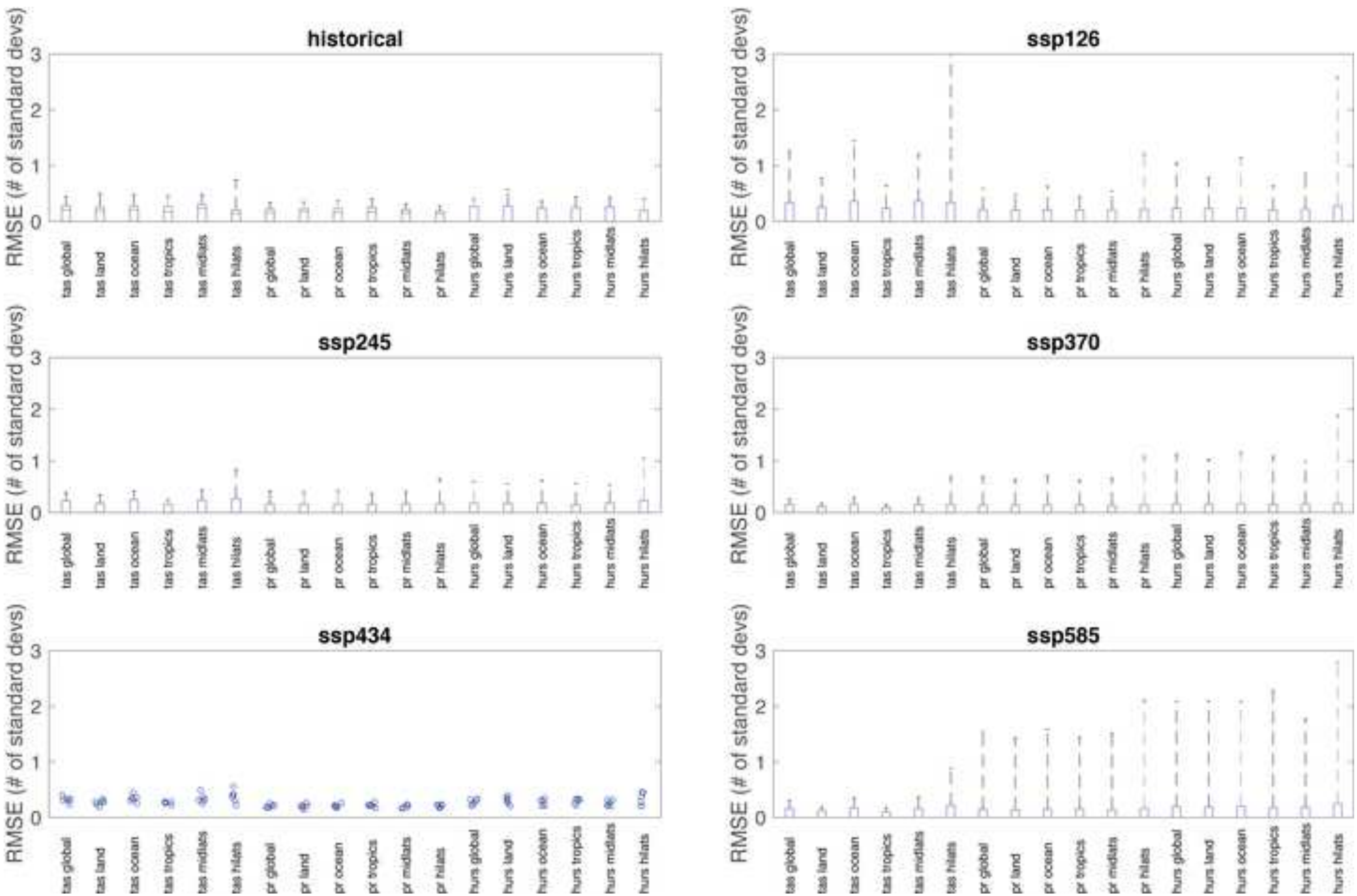

Surface Air Temperature
Precipitation
Relative Humidity
1
2
3
4
5

Figure 6

Click here to access/download;Figure;figure6.png



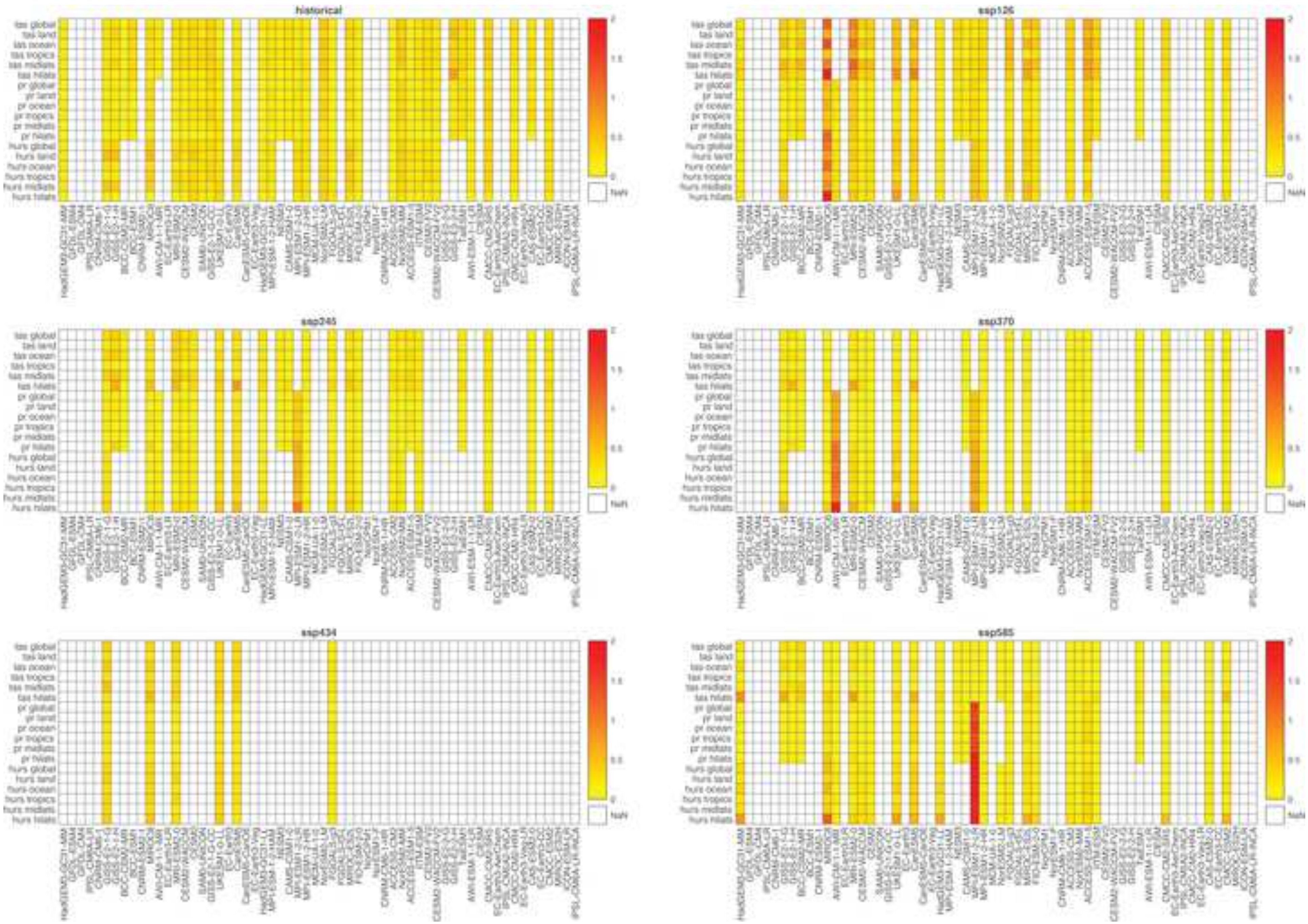

historical
ssp126
ssp245
ssp370
ssp434
ssp585
tas global
tas land
tas ocean
tas tropics
tas midlats
tas hilats
pr global
pr land
pr ocean
pr tropics
pr midlats
pr hilats
hurs global
hurs land
hurs ocean
hurs tropics
hurs midlats
hurs hilats
2
1.5
1
0.5
0
NaN

Figure 8

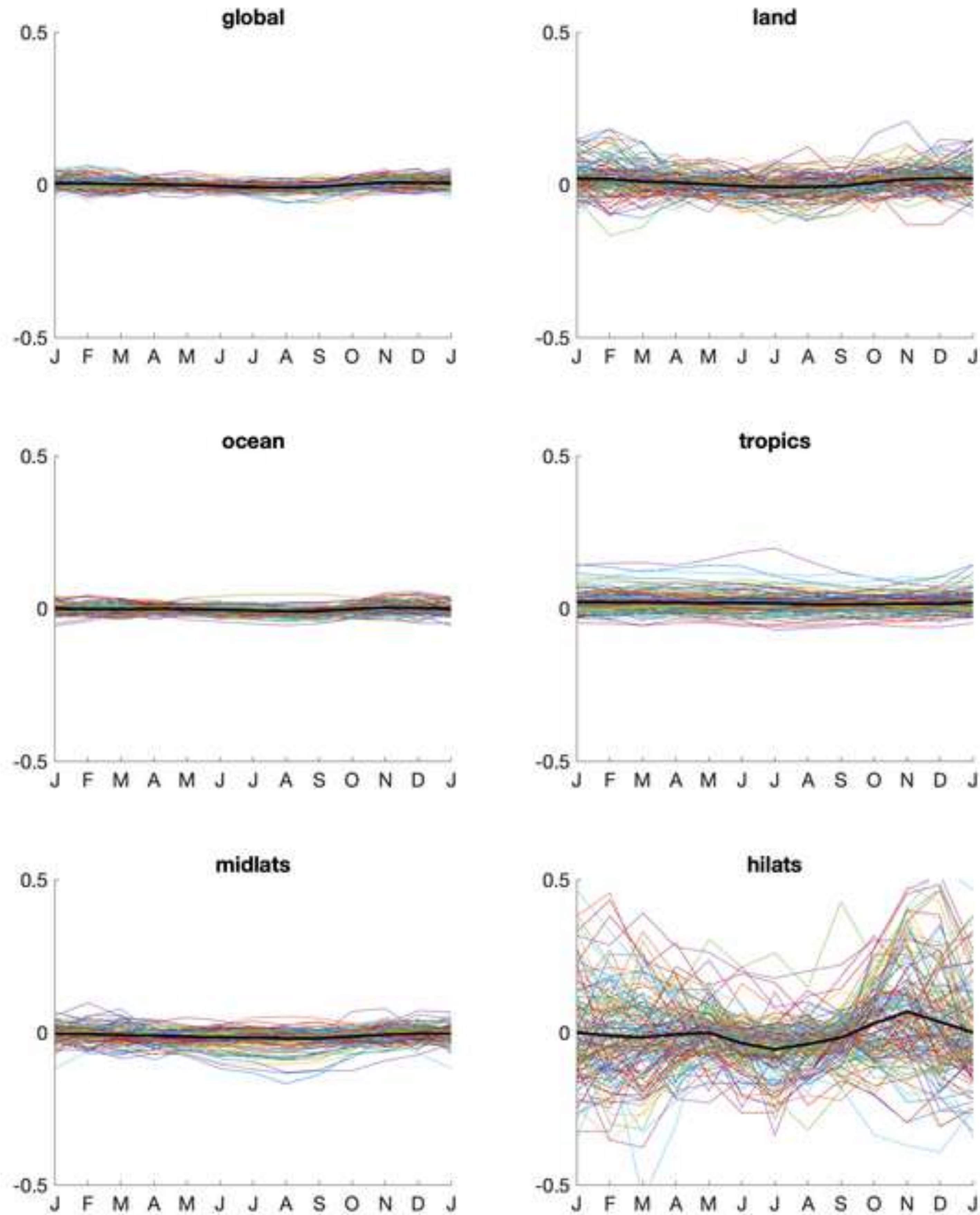

global
land
ocean
tropics
midlats
hilats
0.5
0
-0.5
J F M A M J J A S O N D J

Figure 9

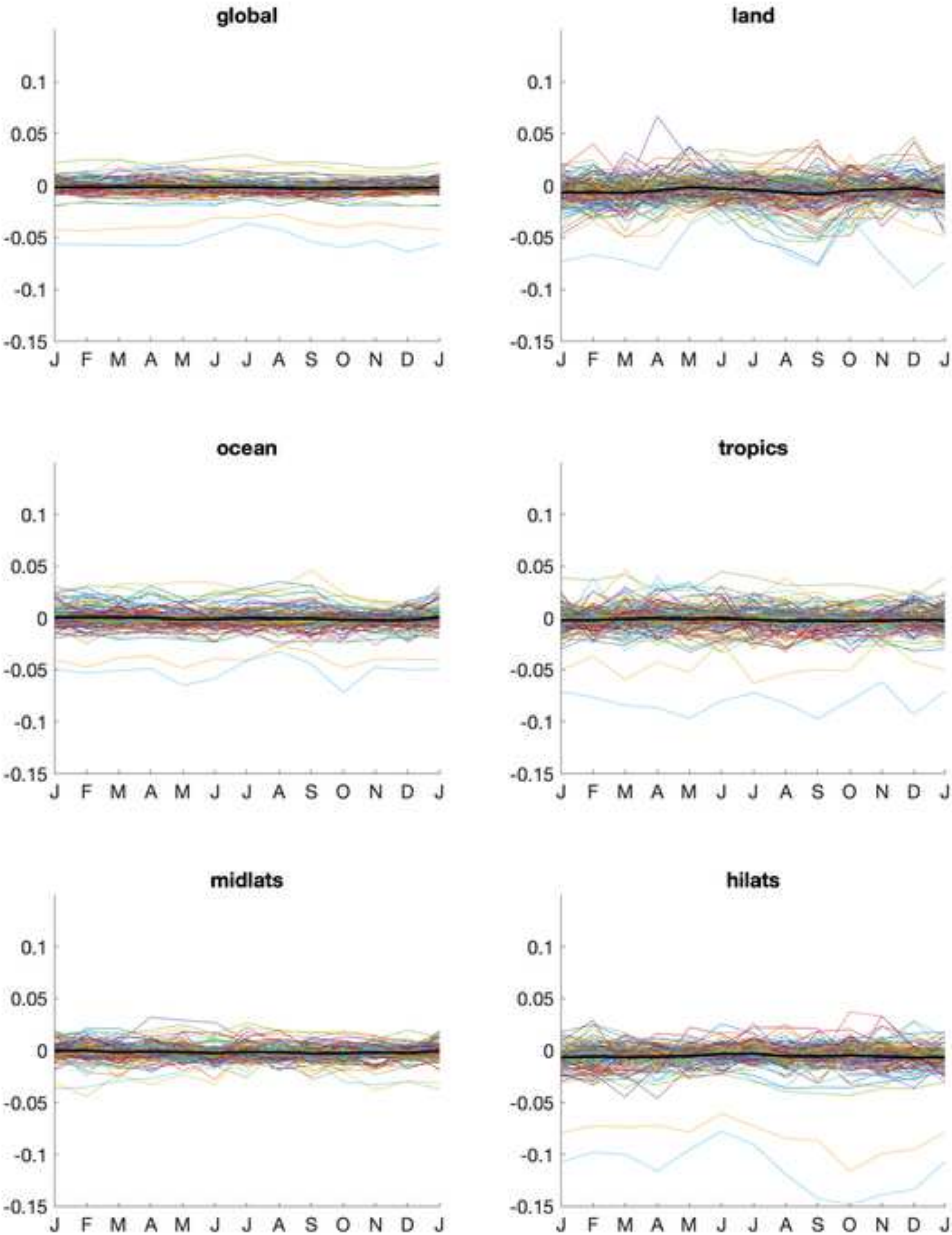

global
land
ocean
tropics
midlats
hilats
0.1
0.05
0
-0.05
-0.1
-0.15
J F M A M J J A S O N D J

Figure 10

Click here to access/download;Figure;figure10.png

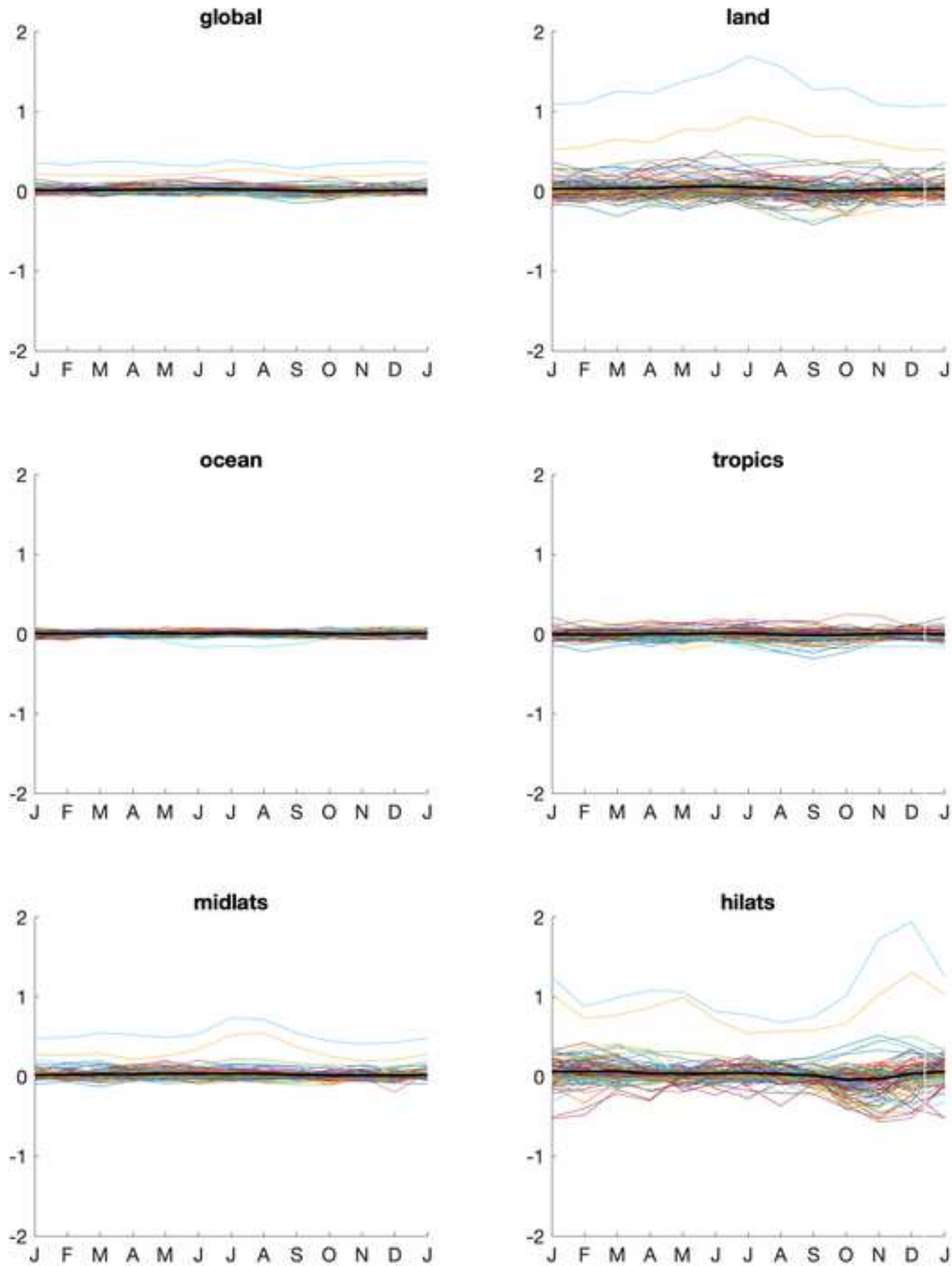


Figure S1

Click here to access/download;Figure;figureS1.png

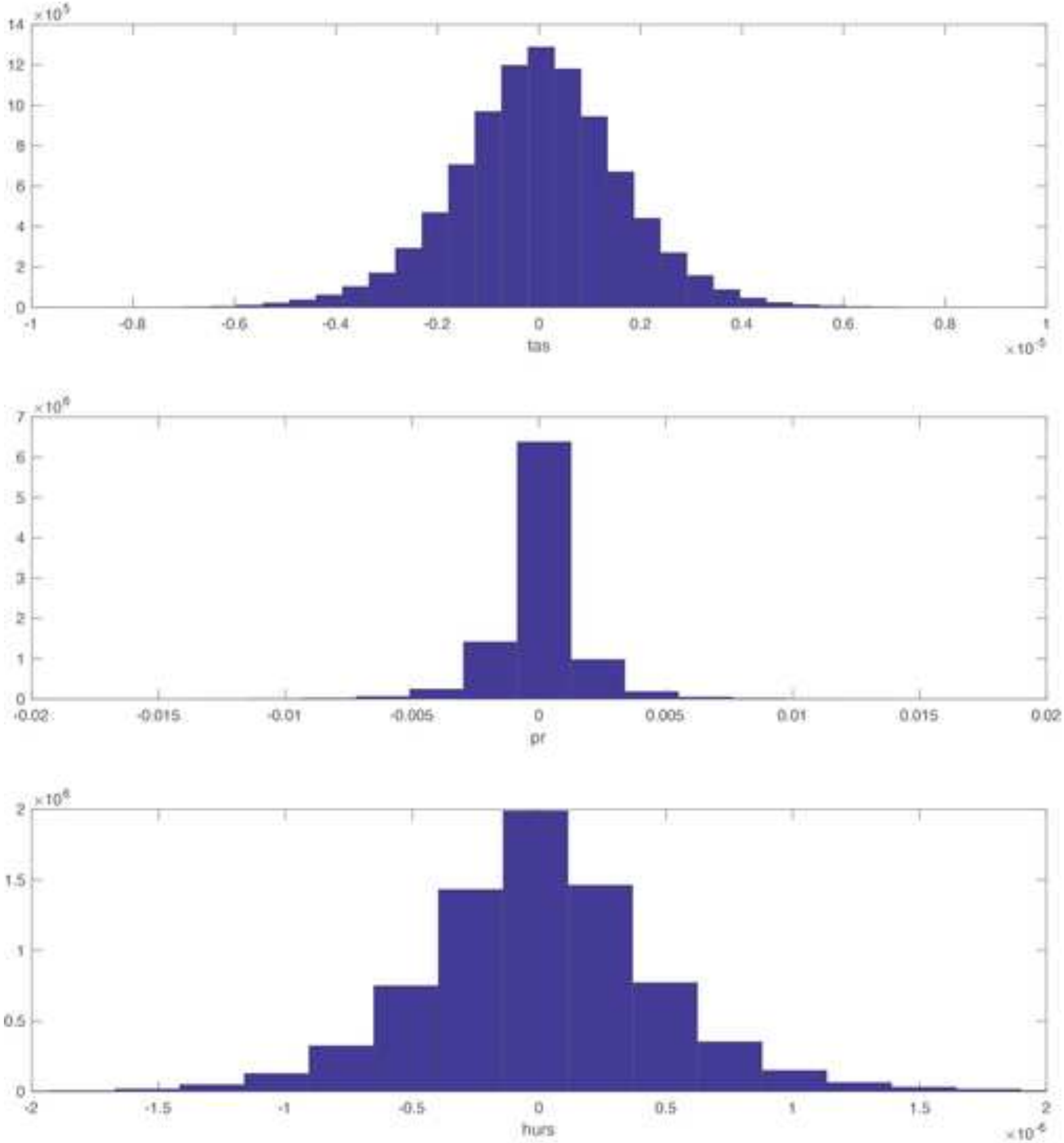


Figure S2

Click here to access/download;Figure;figureS2.png

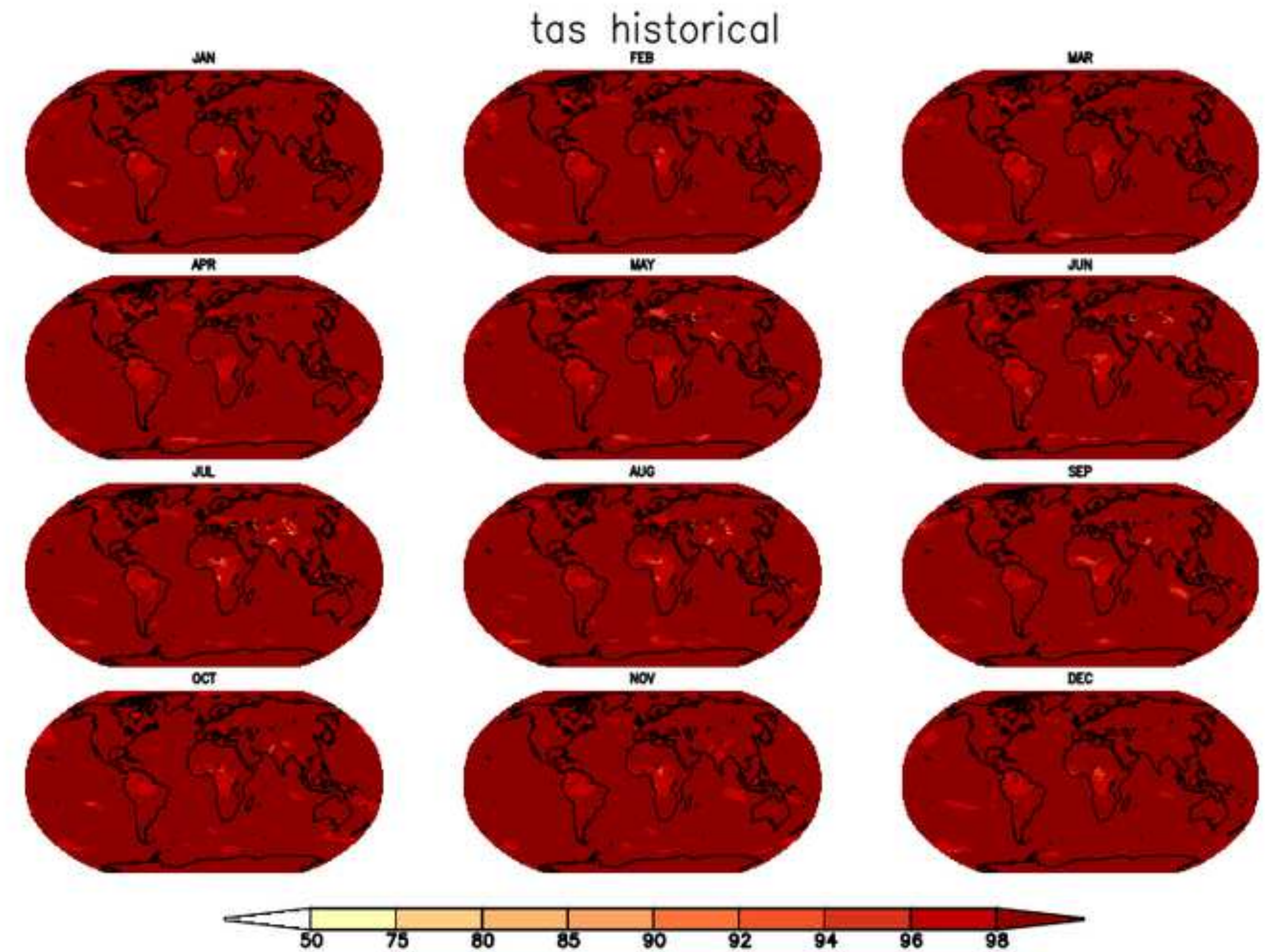


Figure S3

Click here to access/download;Figure;figureS3.png

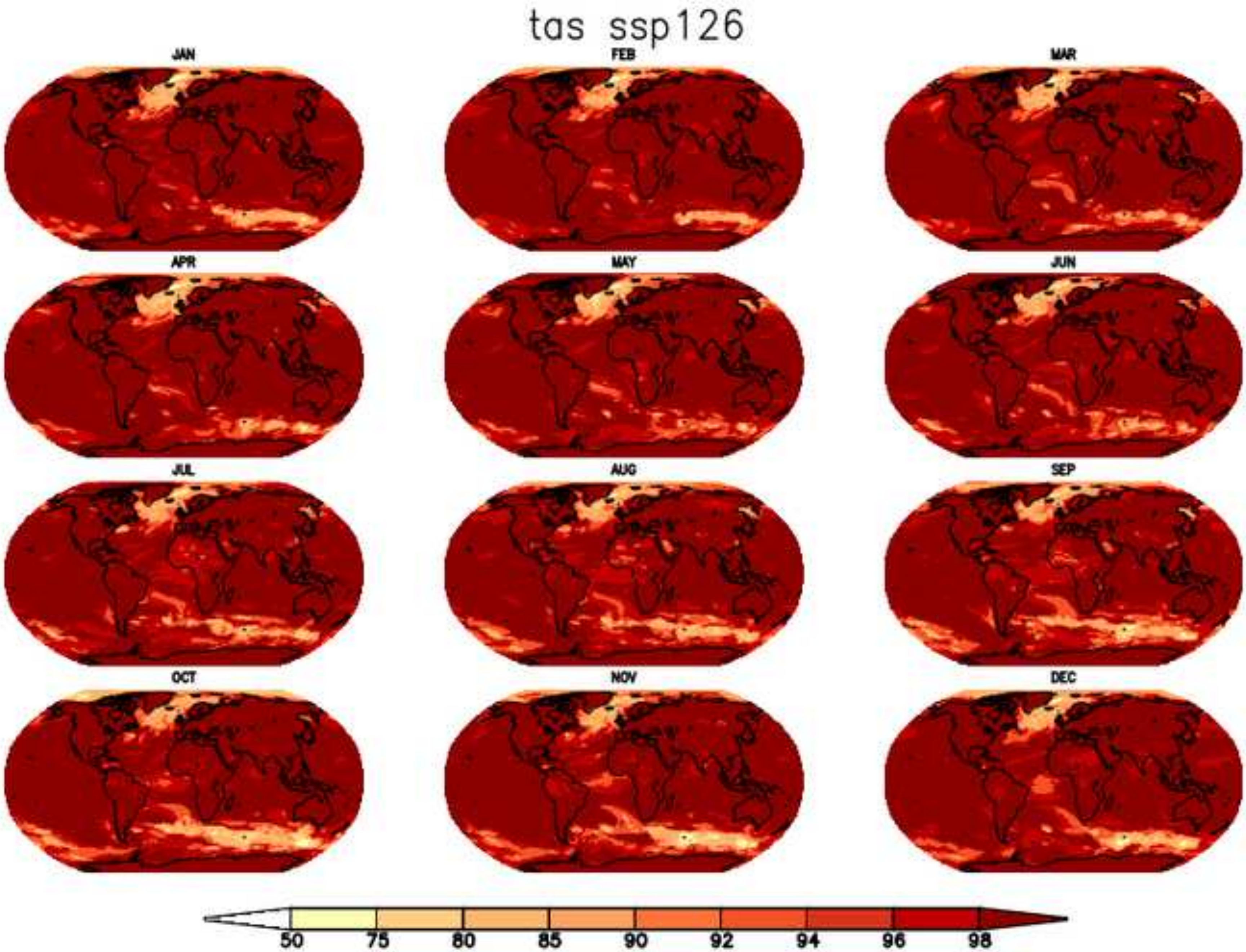


Figure S4

Click here to access/download;Figure;figureS4.png

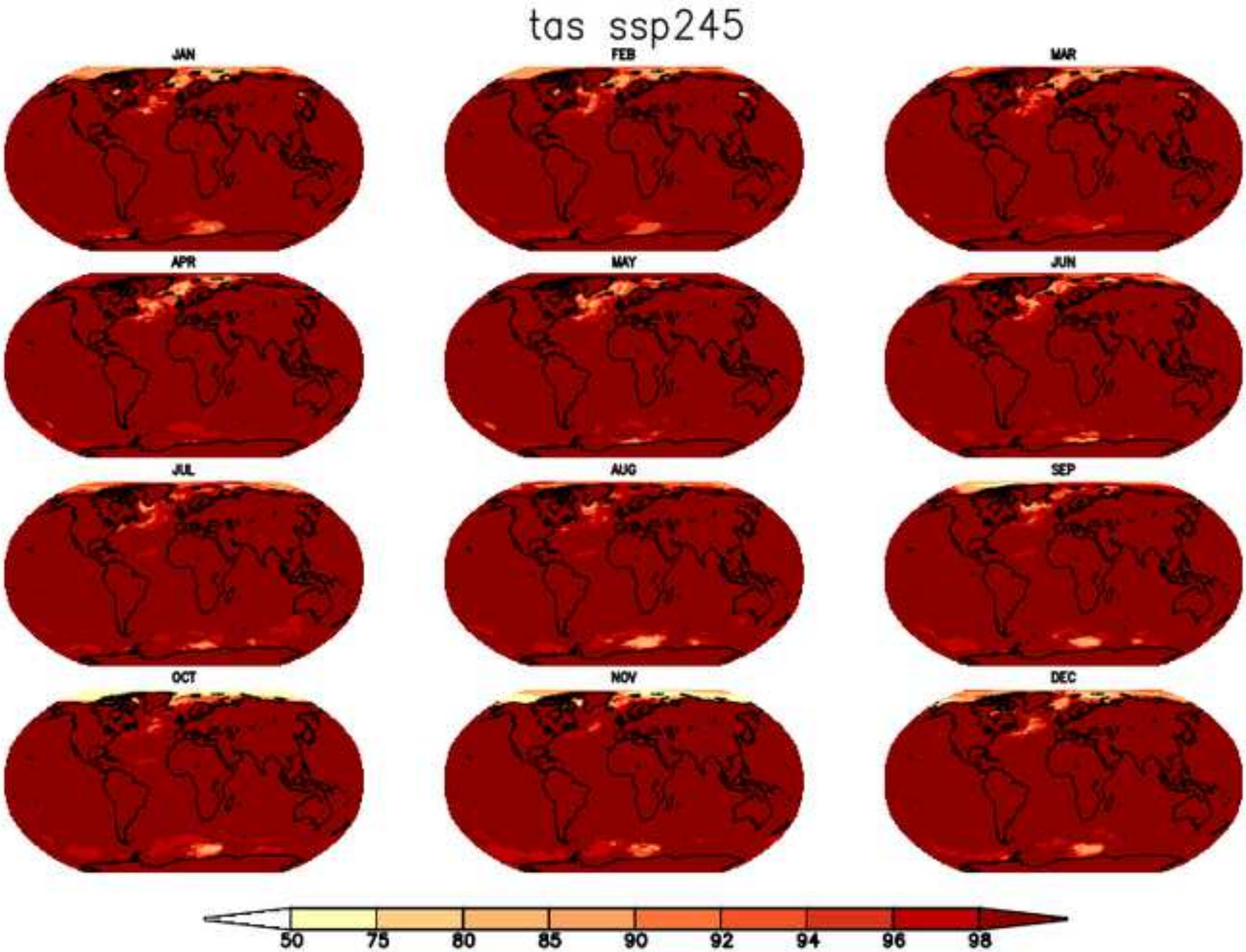


Figure S5

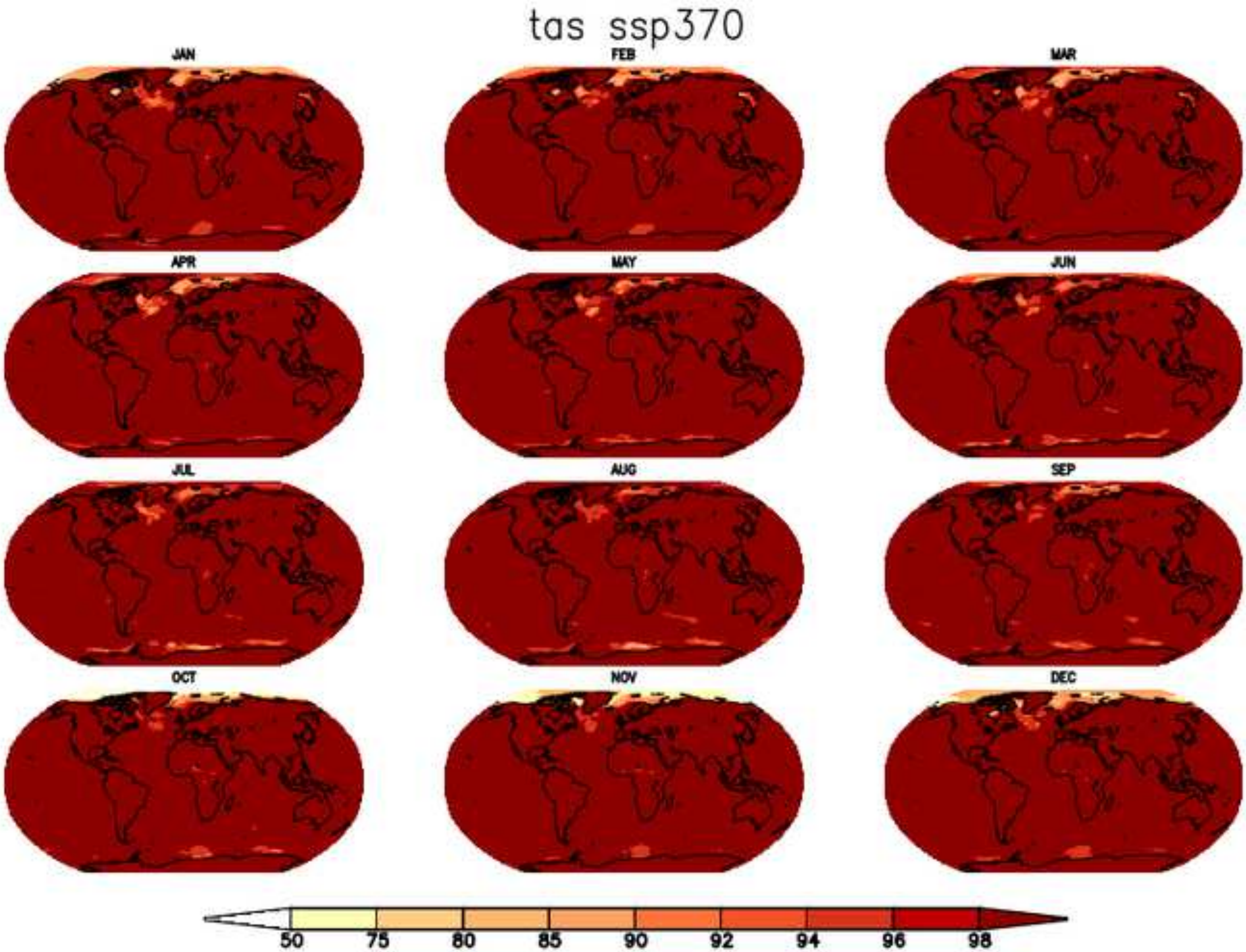

tas ssp370
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
50
75
80
85
90
92
94
96
98

Figure S6

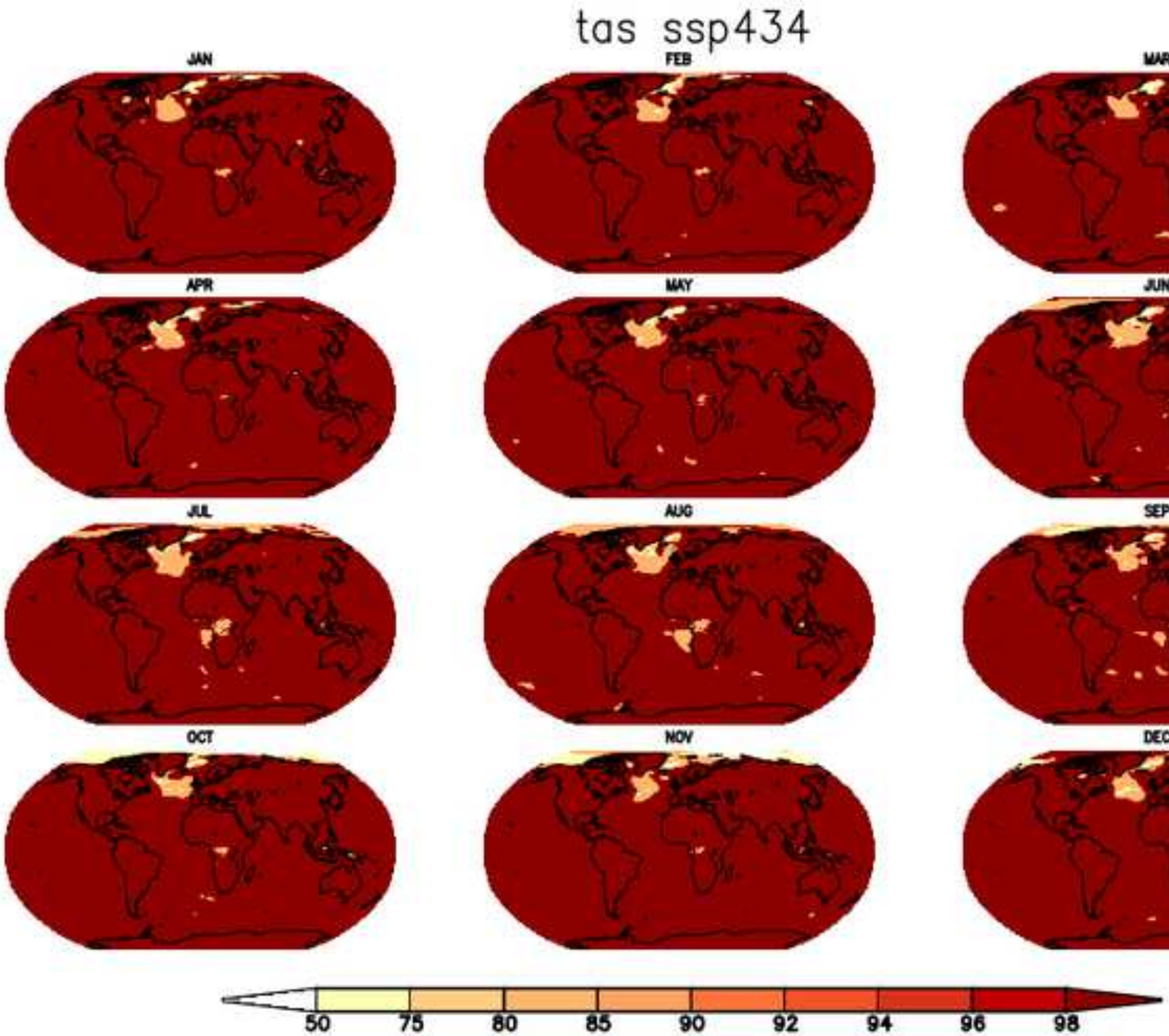

tas ssp434
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
50
75
80
85
90
92
94
96
98

Figure S7

Click here to access/download;Figure;figureS7.png

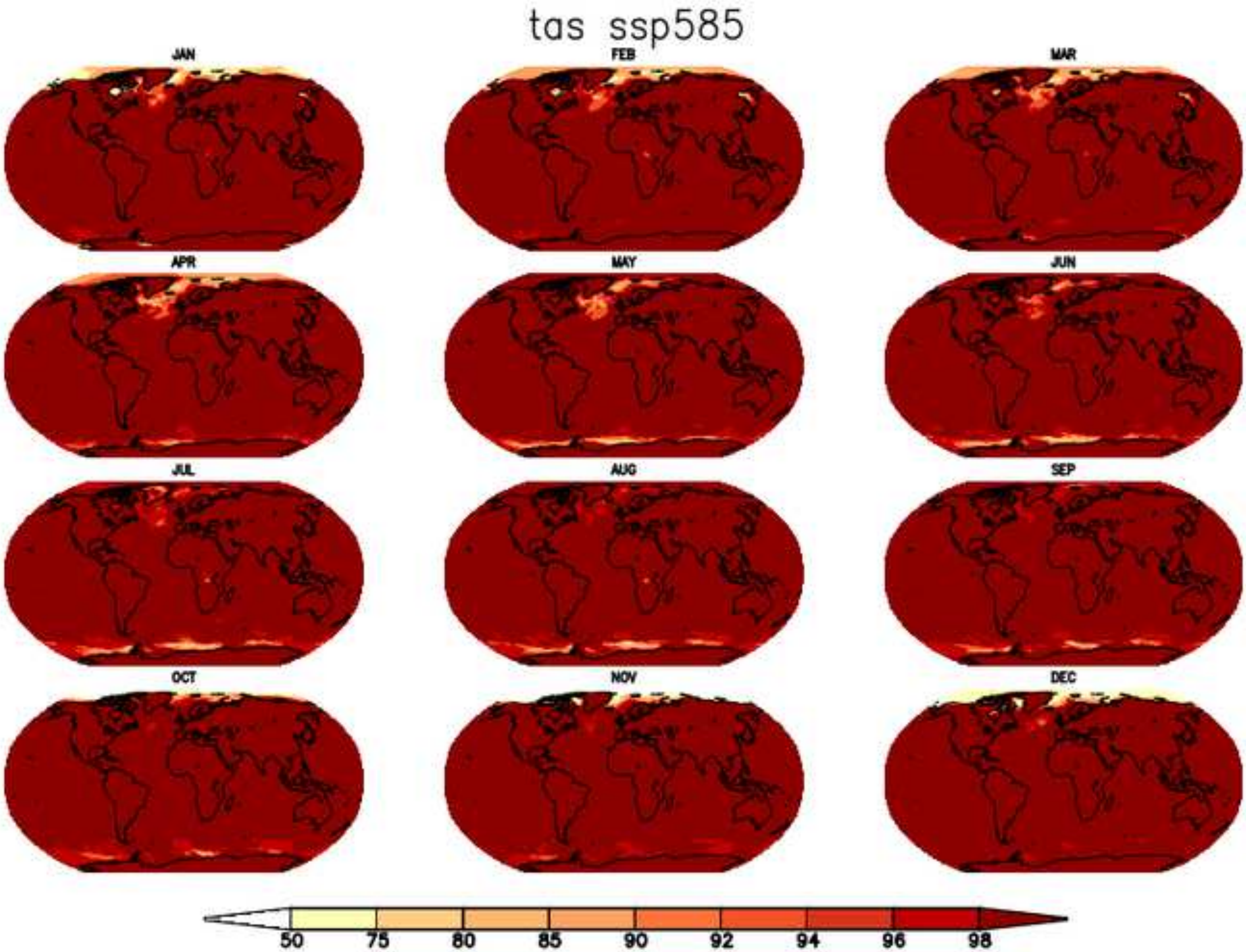


Figure S8

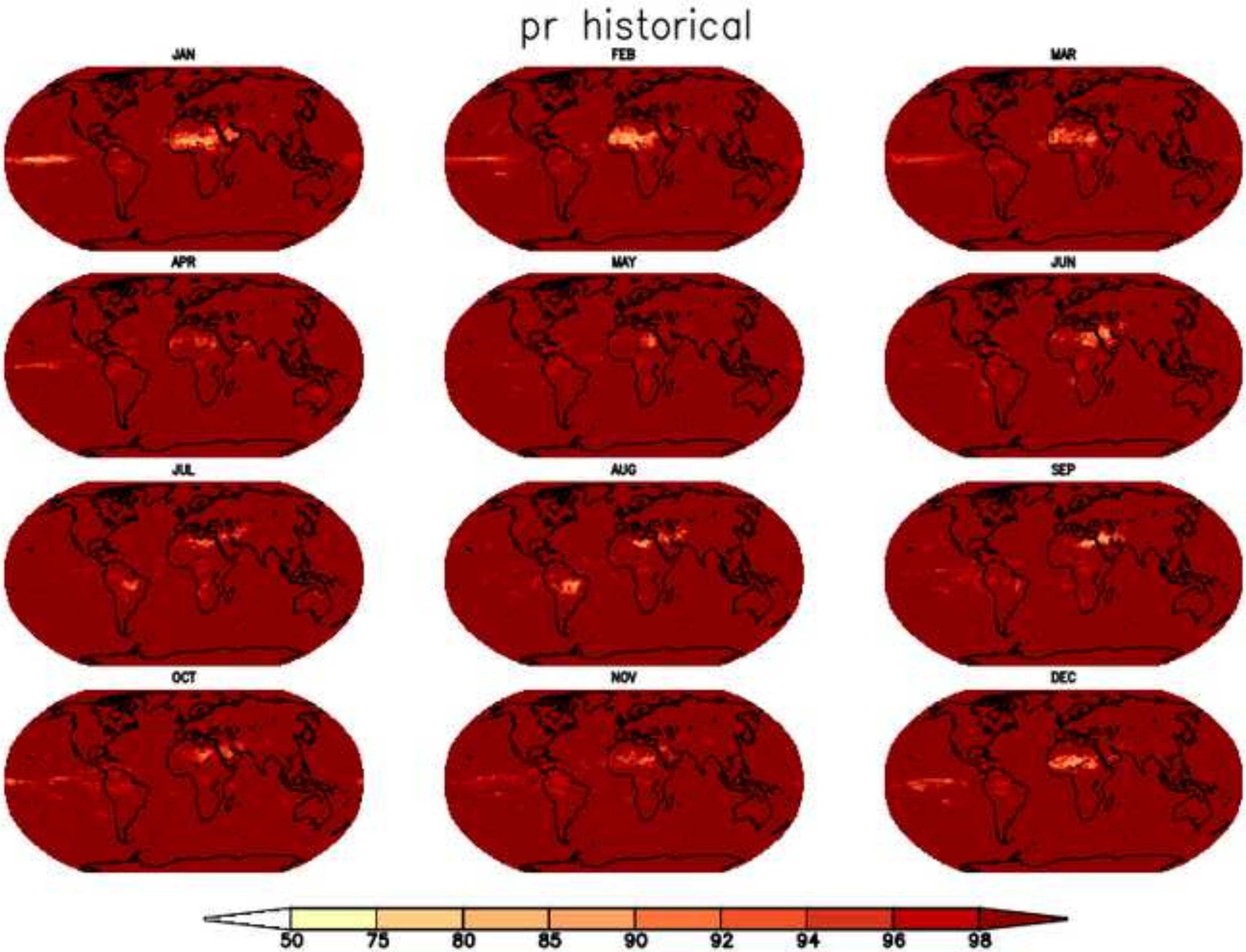
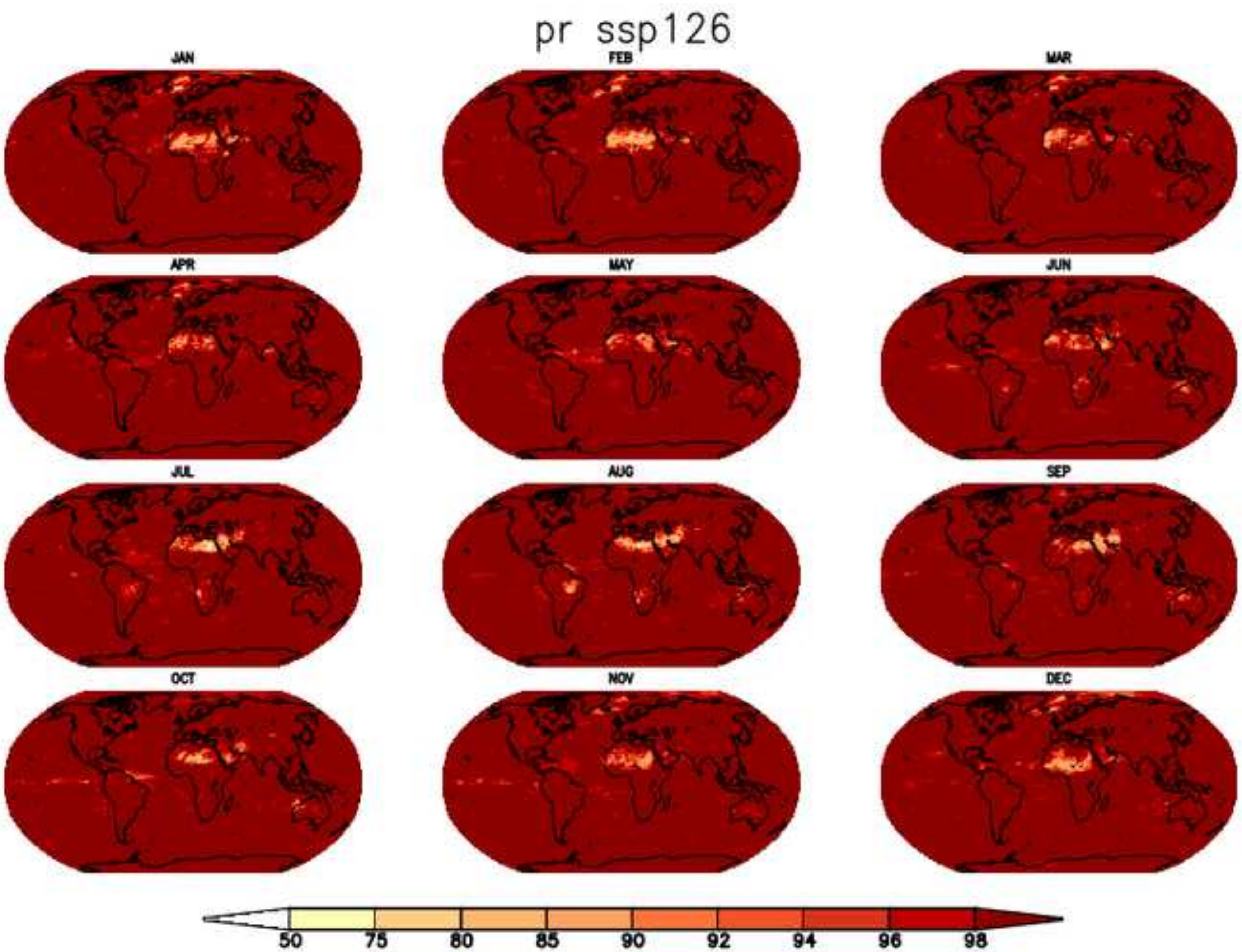

pr historical
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
50
75
80
85
90
92
94
96
98

Figure S9


pr ssp126
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
50
75
80
85
90
92
94
96
98

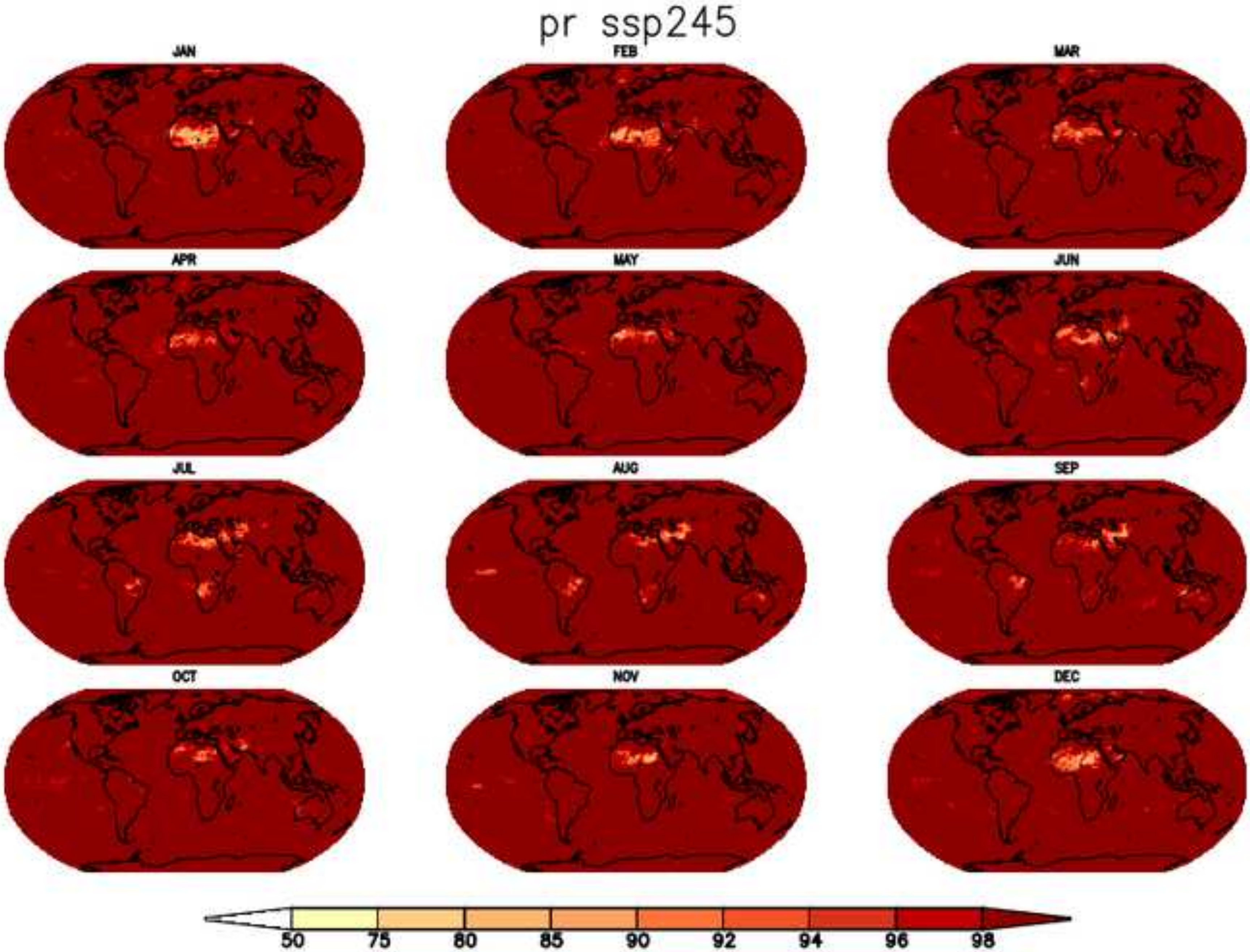

pr ssp245
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
50
75
80
85
90
92
94
96
98

Figure S11

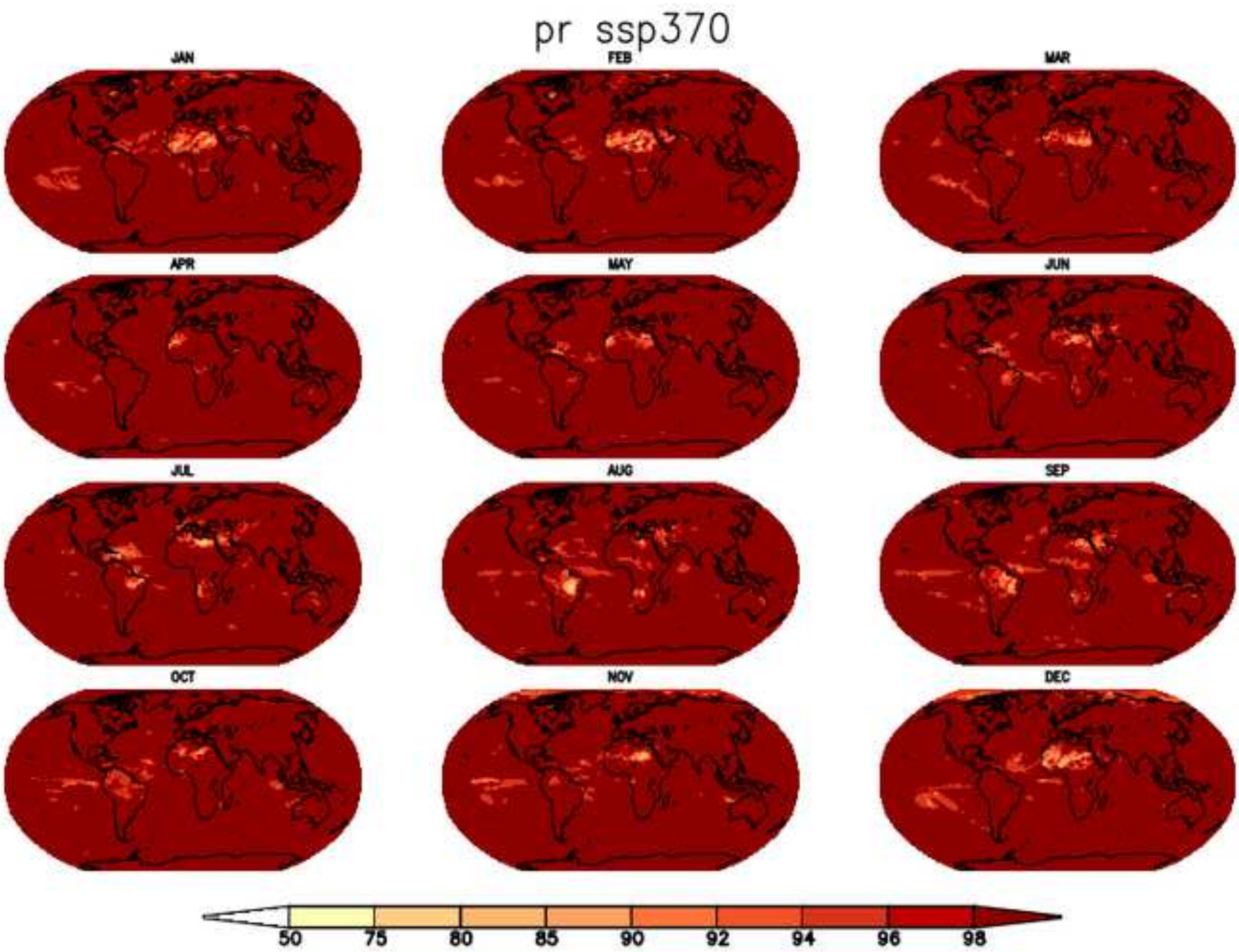

pr ssp370
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
50
75
80
85
90
92
94
96
98

Figure S12

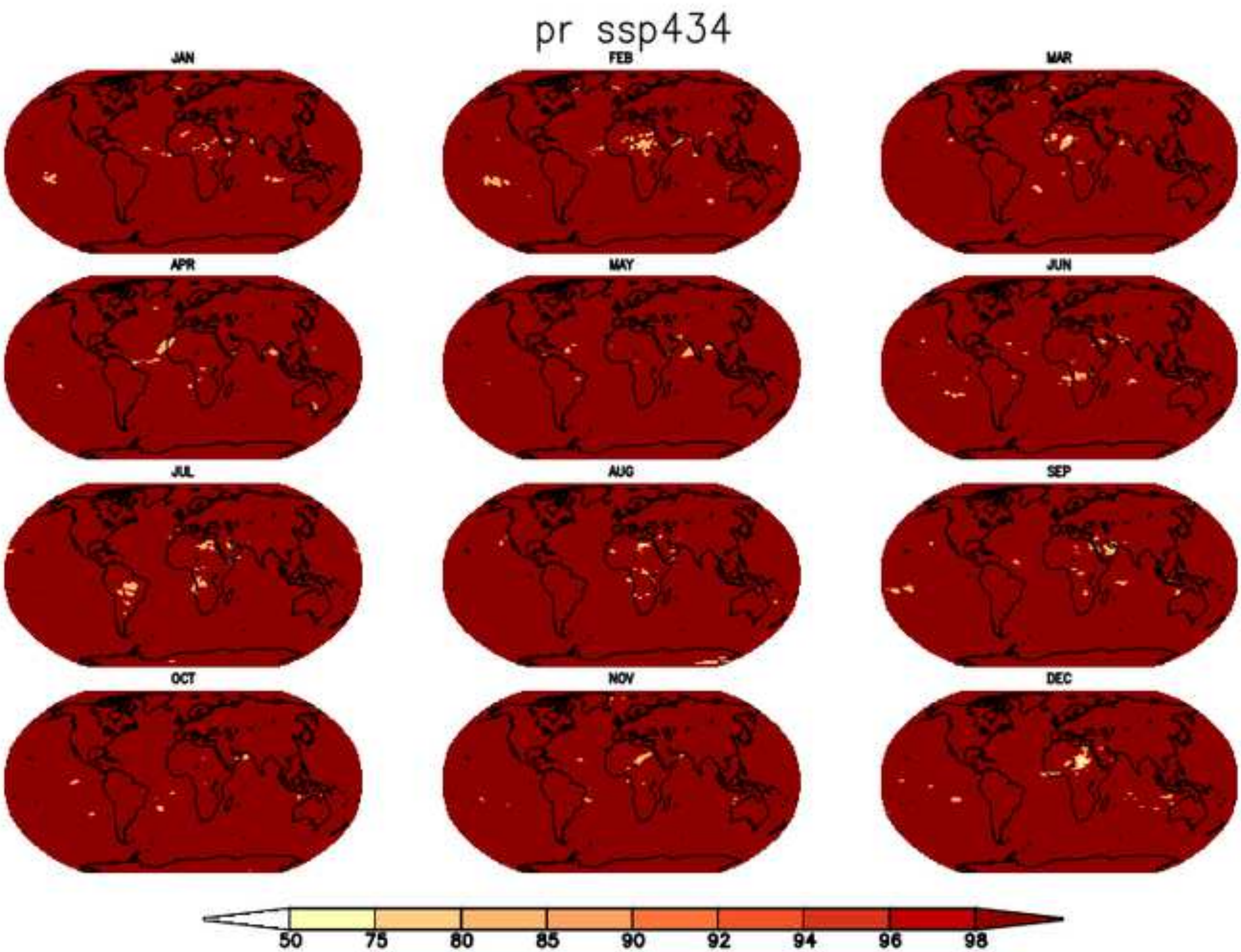

pr ssp434
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
50
75
80
85
90
92
94
96
98

Figure S13

Click here to access/download;Figure;figureS13.png

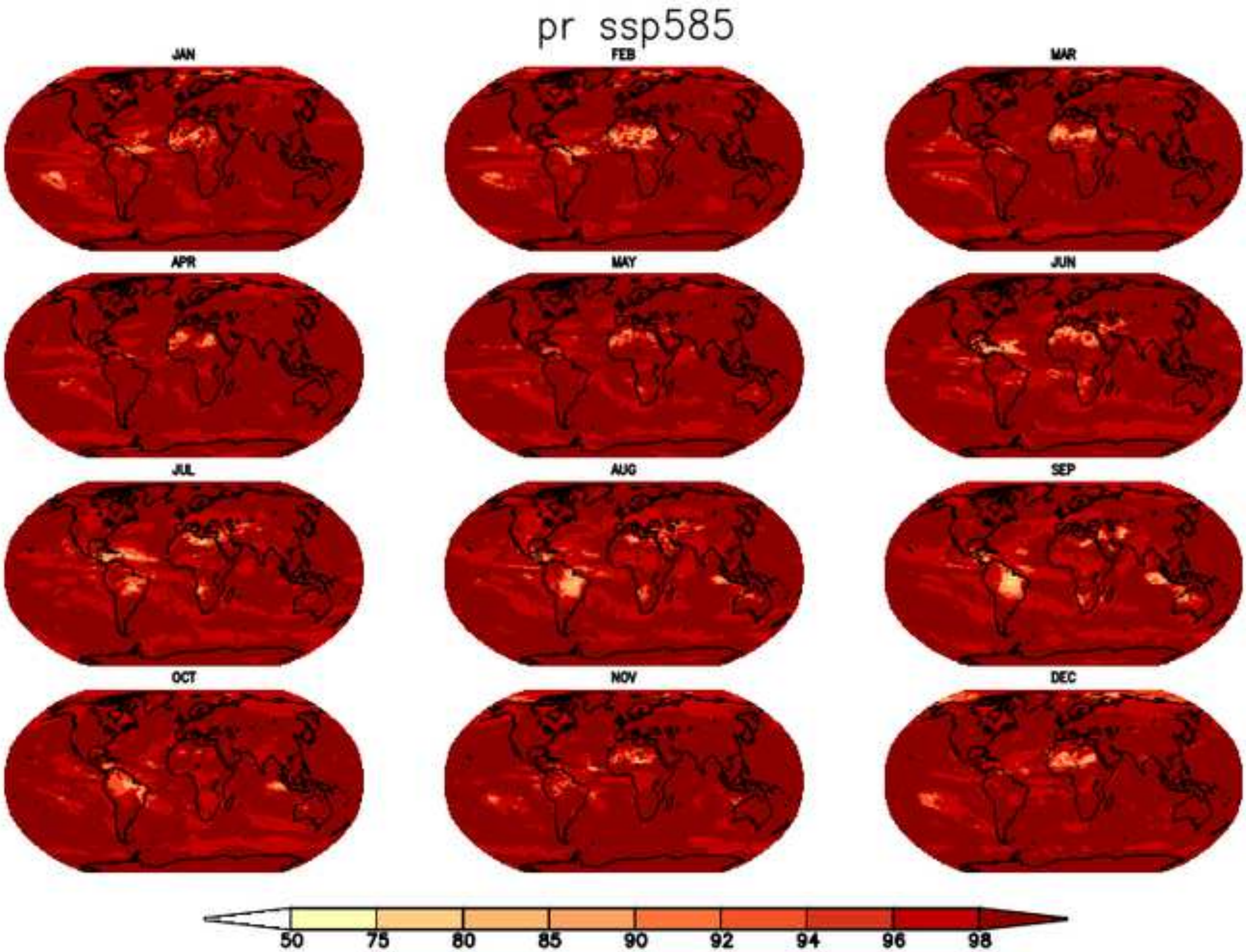


Figure S14

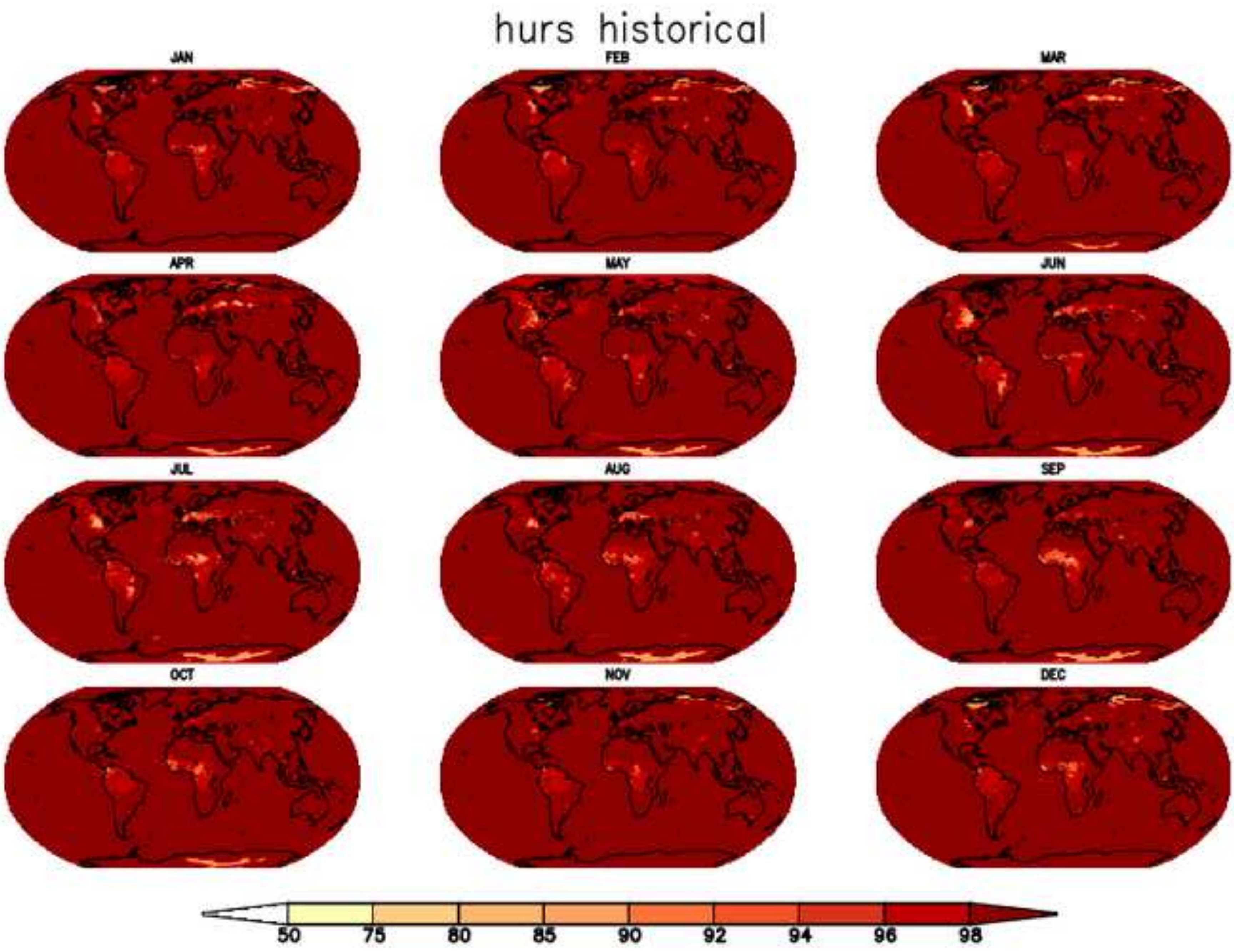

hurs historical
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
50
75
80
85
90
92
94
96
98

Figure S15

Click here to access/download;Figure;figureS15.png

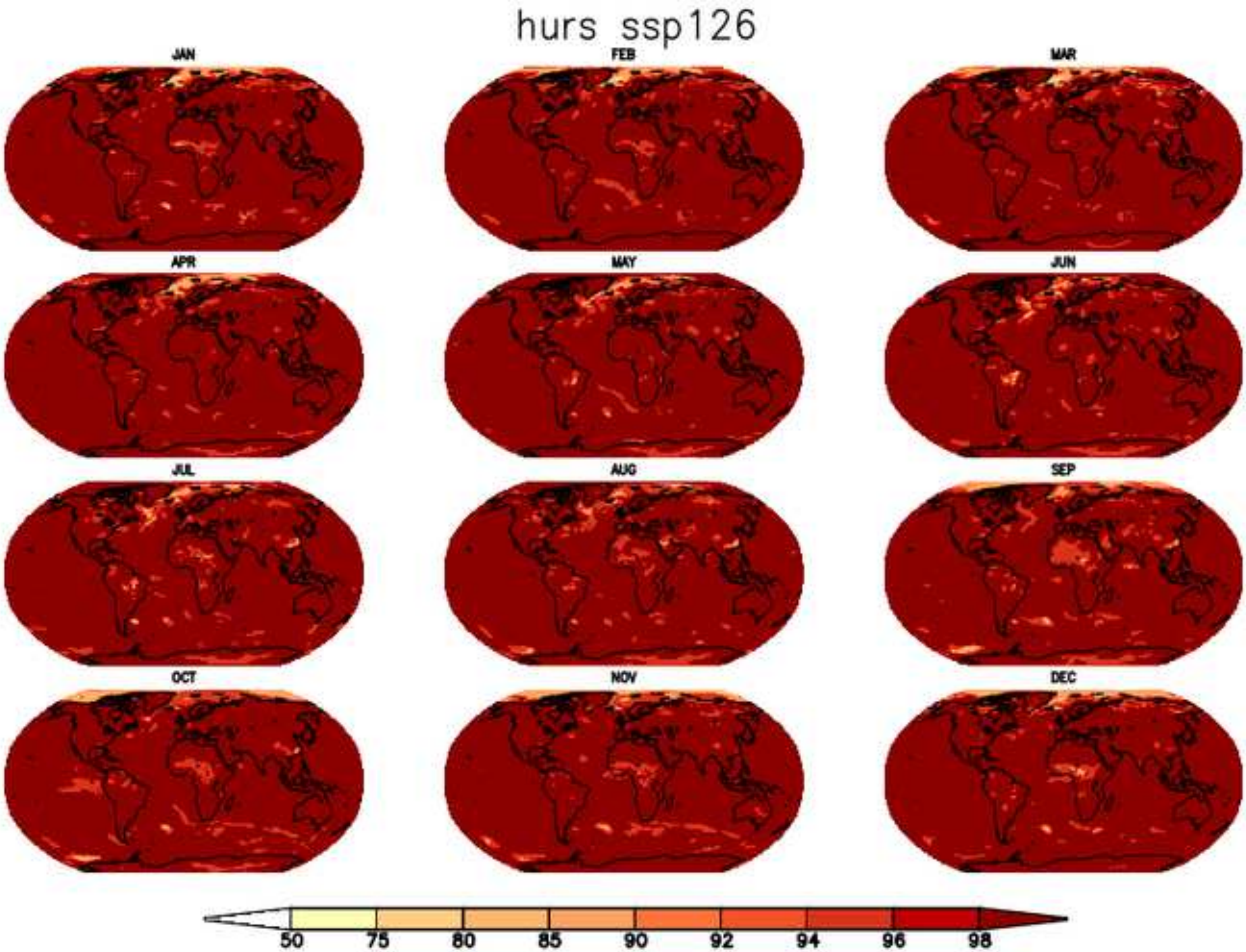


Figure S16

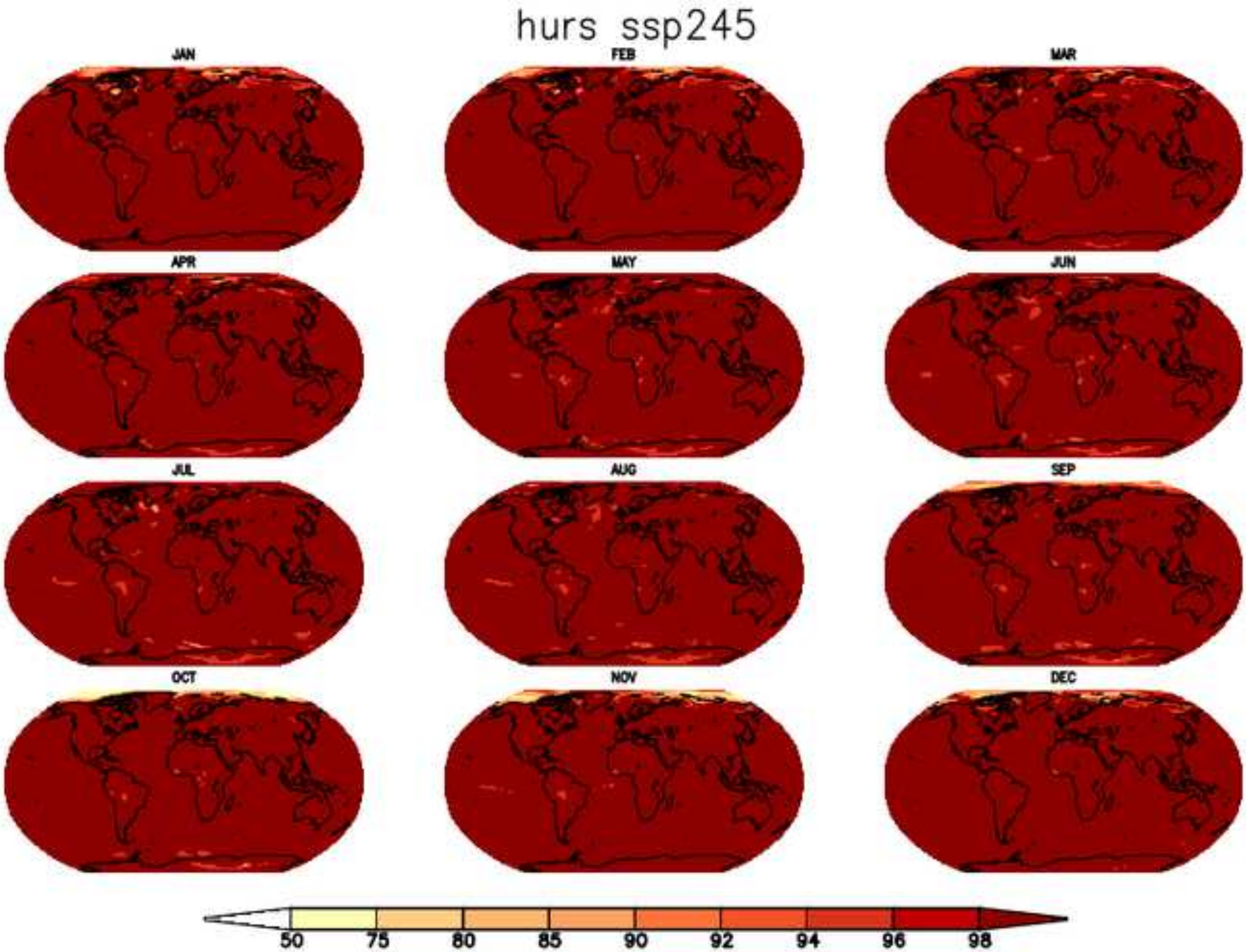

hurs ssp245
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
50
75
80
85
90
92
94
96
98

Figure S17

Click here to access/download;Figure;figureS17.png

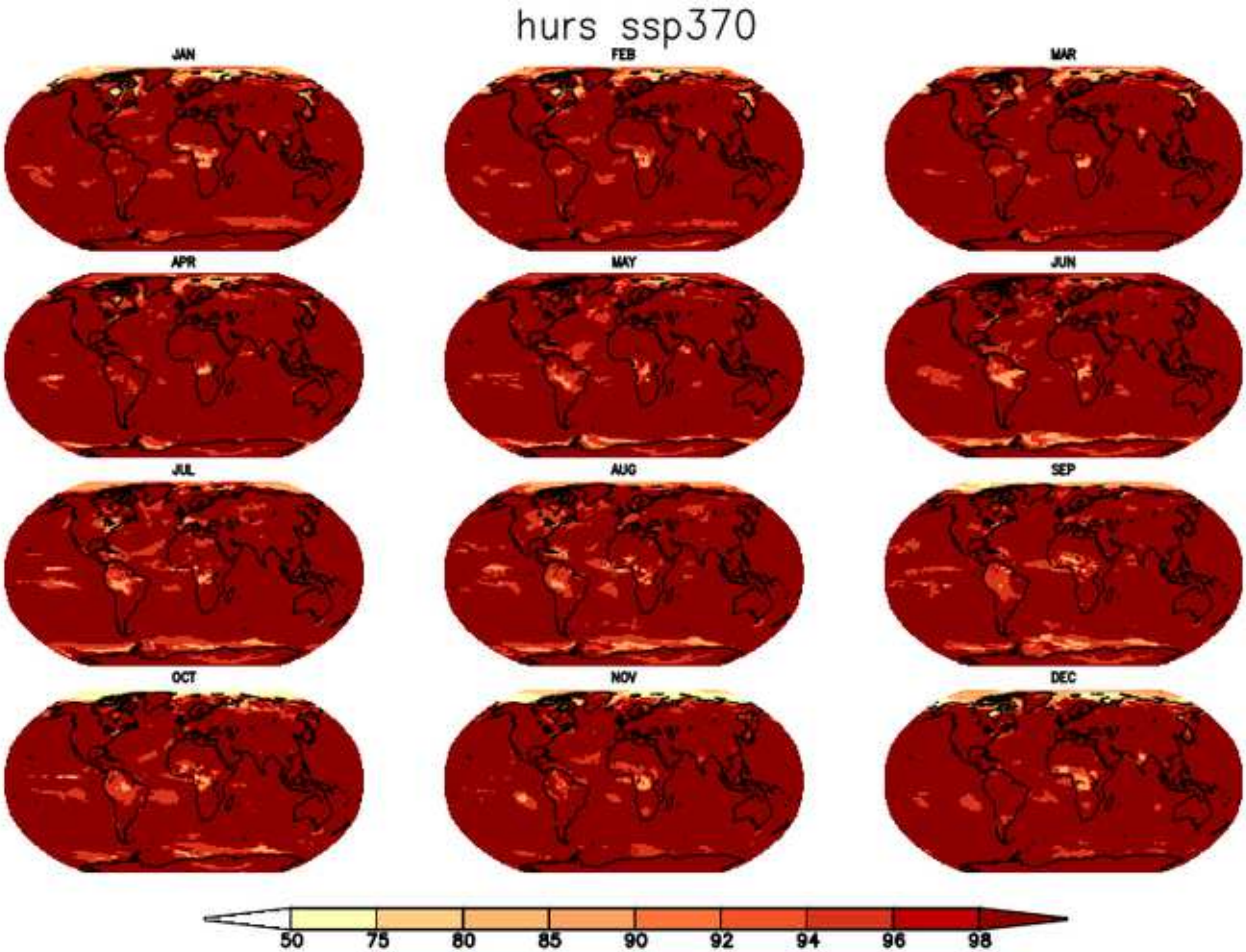


Figure S18

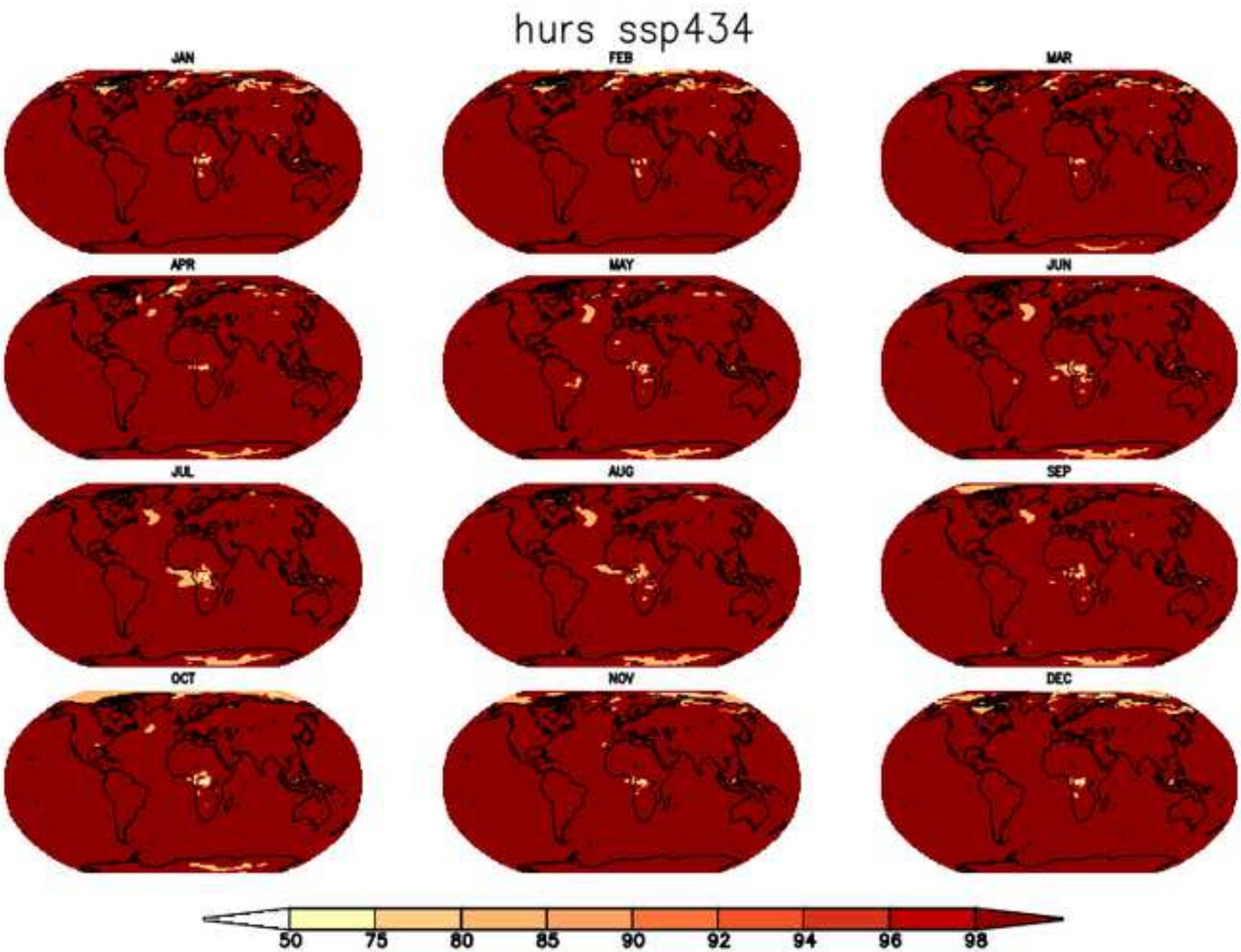

hurs ssp434
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
50
75
80
85
90
92
94
96
98

Figure S19

Click here to access/download;Figure;figureS19.png

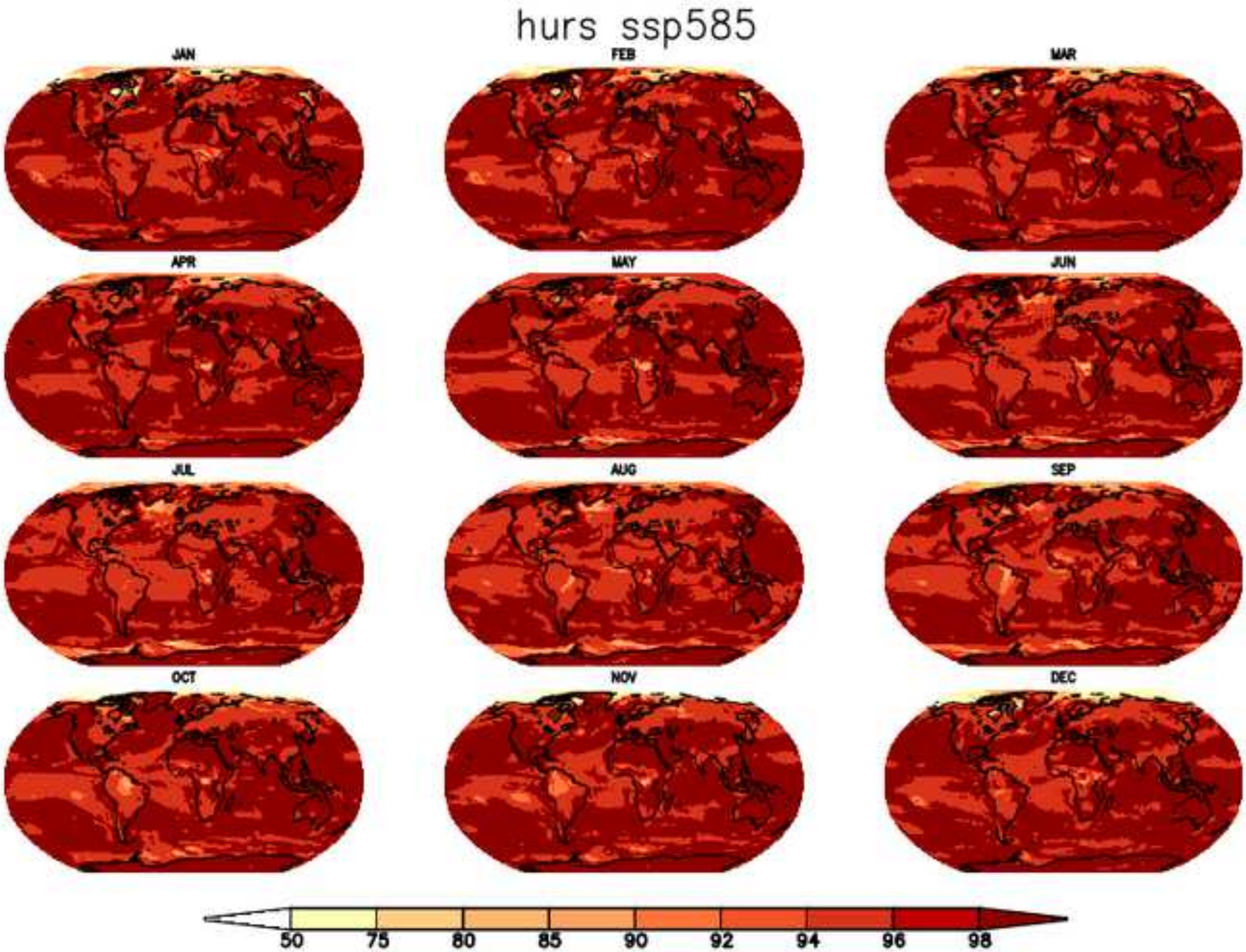


Figure S20

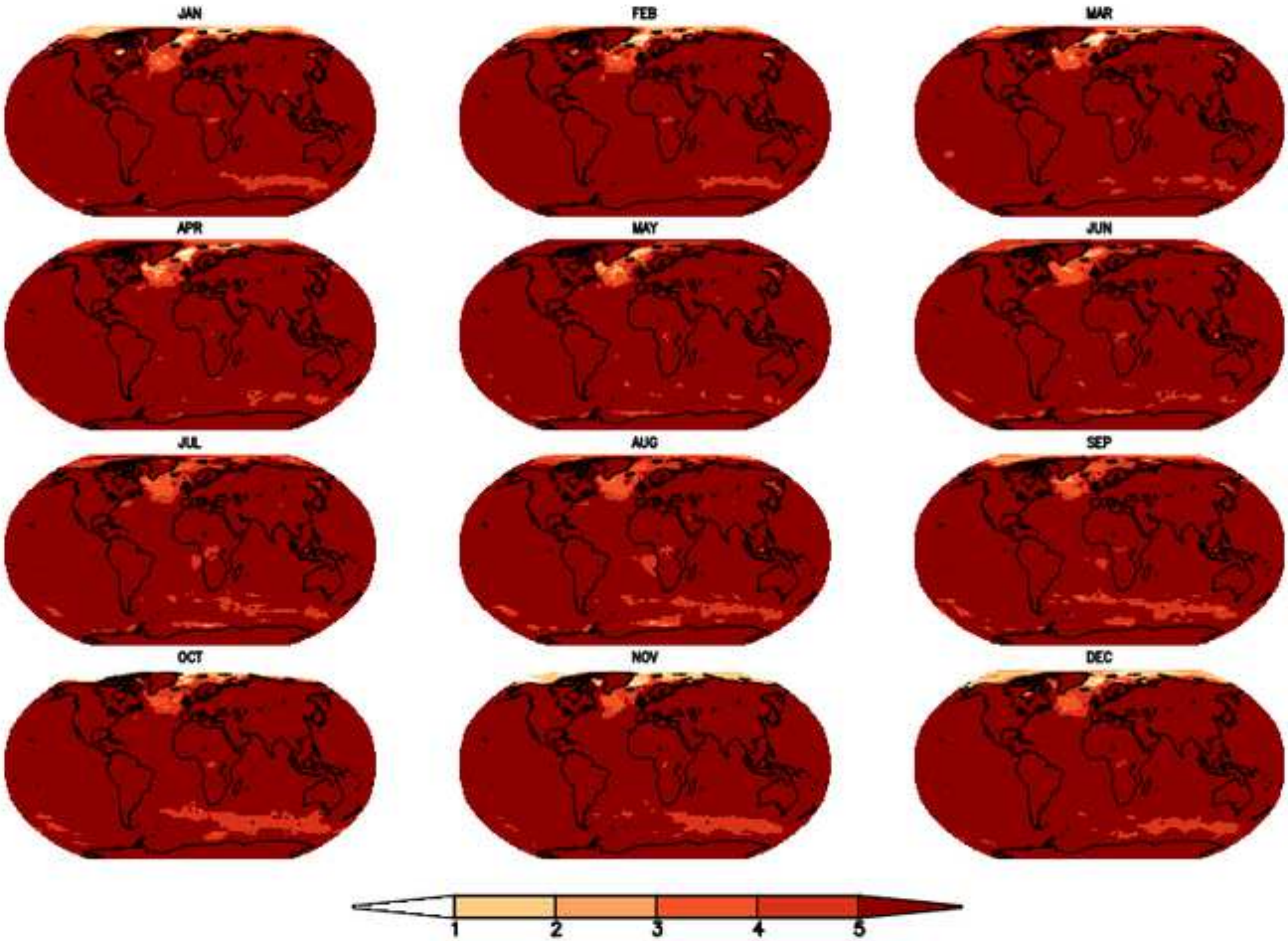


Figure S21

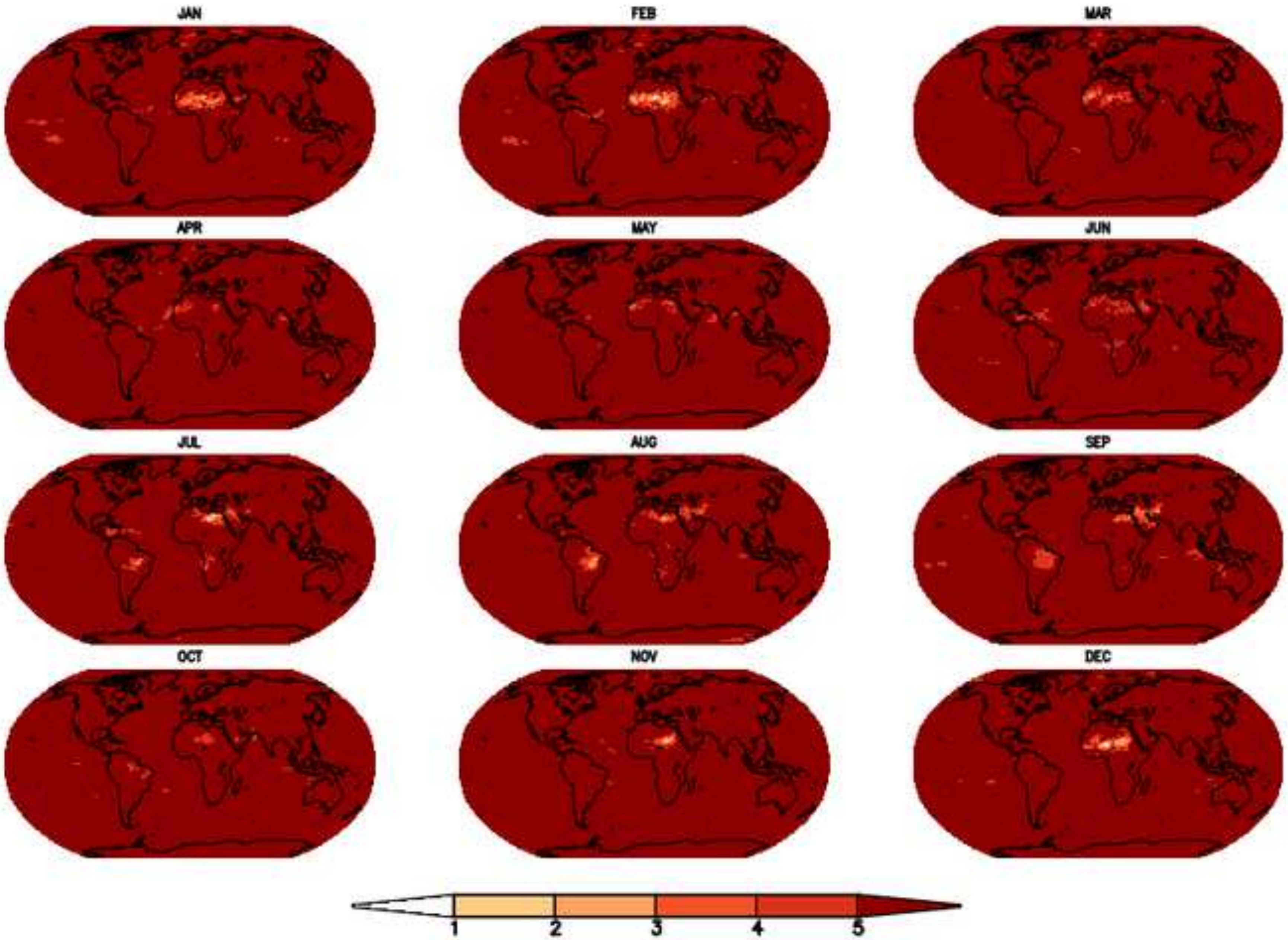


Figure S22

Click here to access/download;Figure;figureS22.png