High-resolution Cloud Climatology for Global Land Areas

Adam M. Wilson & Walter Jetz

# Introduction

Global change biology depends upon climate data that adequately capture the abiotic environment experienced by organisms. Micro-climatic (100-103m) variability is extremely important to

The vast majority of regional to global analyses use interpolated climate datasets to estimate parameters of interest in locations with no station observations. Clouds are a vitally important yet often overlooked driver of ecological processes. We introduce a new high resolution (30-arcsecond, ≈1km) set of monthly cloud frequency climatologies and related metrics derived from twice daily observations over the full 2000-2014 MODIS archive for global land areas. The climatologies were validated using cloud observations from a global set of over 5000 stations over 1970-2009. These data provide a new lens through which to understand the fine-grain spatial variability of global cloud cover.

# Introduction

Clouds are an extremely important component of the global climate system, affecting energy balance, latent heat flux, radiation flux, and moisture transport1. Cloud cover also affects many ecological processes, such as drought stress2, tree growth2, available photosynthetically active radiation3, eco-physiology4, and animal behavior5,6. Furthermore, cloud frequency can be a better predictor than interpolated precipitation for plant distributions7. Additionally, changes in cloud patterns (including, but beyond changes in precipitation) have multi-fold consequences for biodiversity and ecosystems and thus require fine-grain understanding and monitoring4. Cloud dynamics can vary drastically over small spatial and temporal grains due to atmospheric circulation, topography, and even land cover. Topographic morphology, for example, has profound effects on clouds and precipitation at grains finer than 2km in mountainous areas8. Even soil moisture9 and deforestation10 can affect mesoscale convective events, cloud cover, and precipitation. In the ecological community, there is growing recognition of the importance of fine-grain species-habitat (“microhabitat”) associations in a variety of systems11 and thus a need to increase the spatial resolution of environmental datasets12.

Existing long-term cloud climatologies are typically available and analyzed at relatively coarse grains. The majority of global analyses of cloud climatologies have been conducted at grains coarser than 10km and often over 100km13–15 despite our understanding of important cloud dynamics operating at grains as fine as 1-2km 8,10,16,17. For example, the recent GEWEX systematic review of satellite-derived cloud climatologies13 and all MODIS level three (L3) atmosphere products are summarized at 1**°** (≈110km) resolution. While this resolution is appropriate for study of global cloud dynamics (and necessary for cross-platform comparison), it is far too coarse to capture fine-grain variability important in many ecological questions18.

There are a few examples of finer-grain climatologies based on other sensors, such as HIRS14 (~20km), GridSAT19 (~8km), and AVHRR PATMOS-x20 (~8km), but these are 8-20 times coarser than possible with the MOderate resolution Imaging Spectroradiometer (MODIS). To date there have been two efforts to produce high-resolution (<=1km) cloud climatologies from the MODIS archive. One is based on the MOD35 250m visible cloud mask21, but is spatially bounded to the tropics and incorporates only seven years of data (2000-2006). Additionally, these data were derived from the problematic collection 5 MODIS (MOD35) cloud mask and thus contain significant land-cover and processing-path biases in cloud frequency22. The other MODIS-derived 1km cloud climatology avoids the problematic MOD35 algorithm through a simple cloud masking procedure based on scaled visible wavelength (RGB) images from the MODIS “Rapid Response” system23. Douglas, et. al.24,25, developed an algorithm that applies a user-defined threshold to convert RGB “brightness” to “cloudiness.” However, the product is based on a derivative of surface reflectance data designed to be visually appealing, is strongly dependent on the brightness threshold, and is problematic over high-albedo surfaces (such as urban areas or snow). Furthermore, this approach does not exploit more sophisticated tests used in most cloud detection algorithms such as cloud-top infrared temperature13 and is only available for scattered regions around the globe. As a result, cloud cover is infrequently used in ecological studies, likely due in large part to the complexity of attaining it at sufficiently high resolution for a study region. Goldsmith, Matzke, and Dawson18, for example, needed to process a decade of four daily swath-level (ungridded) satellite images (over 14,000 files) simply to extract mean cloud frequency at two locations in Costa Rica.

In summary, cloud frequency is a vitally important integrative predictor in biogeographical processes, but there are no globally available, high-quality, high-resolution, cloud frequency data sets. In this study we develop and validate a 30-arc-second (≈1km) global monthly cloud frequency climatologies using an alternative MODIS cloud masking algorithm.

# Results

### Spatial distribution

The MODCF captures global patterns of cloud frequency, with mean annual cloud frequencies greater than 80% over much of equatorial South America, the Congo River basin in Africa, and Southeast Asia. Intra-annual variability (standard deviation of mean monthly values) is highest over India, Brazil, and the Savannas of Africa (Figure 4).

There is considerable seasonable variability in some regions; intra-annual standard deviations range from X to Y.

Figure : Cloud frequency seasonality and variability for each terrestrial biome separated by geographic realm. The grey lines represent the seasonality for 1000 randomly selected locations within each region and the blue line is a thin-plate spline representing the overall seasonality.

### Summary by biome

Temporal variance and trends: variance within and between months, perhaps trends through time?.

The MODIS MOD09 cloud detection algorithm is capable of explaining nearly 80% of the variability in cloud cover observed at a global set of validation stations. Furthermore, the MODCF data was nearly as accurate over the full time period (1970-2009) versus the MODIS-era (2000-2009) alone. This suggests that the MODCF, although it only represents 14 years of observations (2000-2013), is a useful metric of multi-decadal fine-grain spatial patterns of cloud frequency.

Figure 2 illustrates the increased fine-grain detail available in MODCF compared with easily available coarse-grain cloud data from the GEWEX Cloud Assessment in a region near Parque Nacional Jaua-Sarisariñama in Southern Venezuela. The 1-km dataset captures the effects of orographic cloud formation due to the complex topography (compare with elevation in the lower right panel). Figure 2 also shows mean annual precipitation from the WorldClim dataset26. WorldClim is available at the same resolution (30-arc seconds) as the product described here, but was developed from interpolated station data using only latitude, longitude, and elevation as covariates. The artifacts near stations and treatment of precipitation as a simple function of elevation are apparent in the interpolated precipitation, while much finer detail of orographic cloud effects is apparent in the Cloud Frequency. While these products (cloud frequency and precipitation) are not directly related, the possibility of incorporating cloud frequency in the interpolation of precipitation is promising.

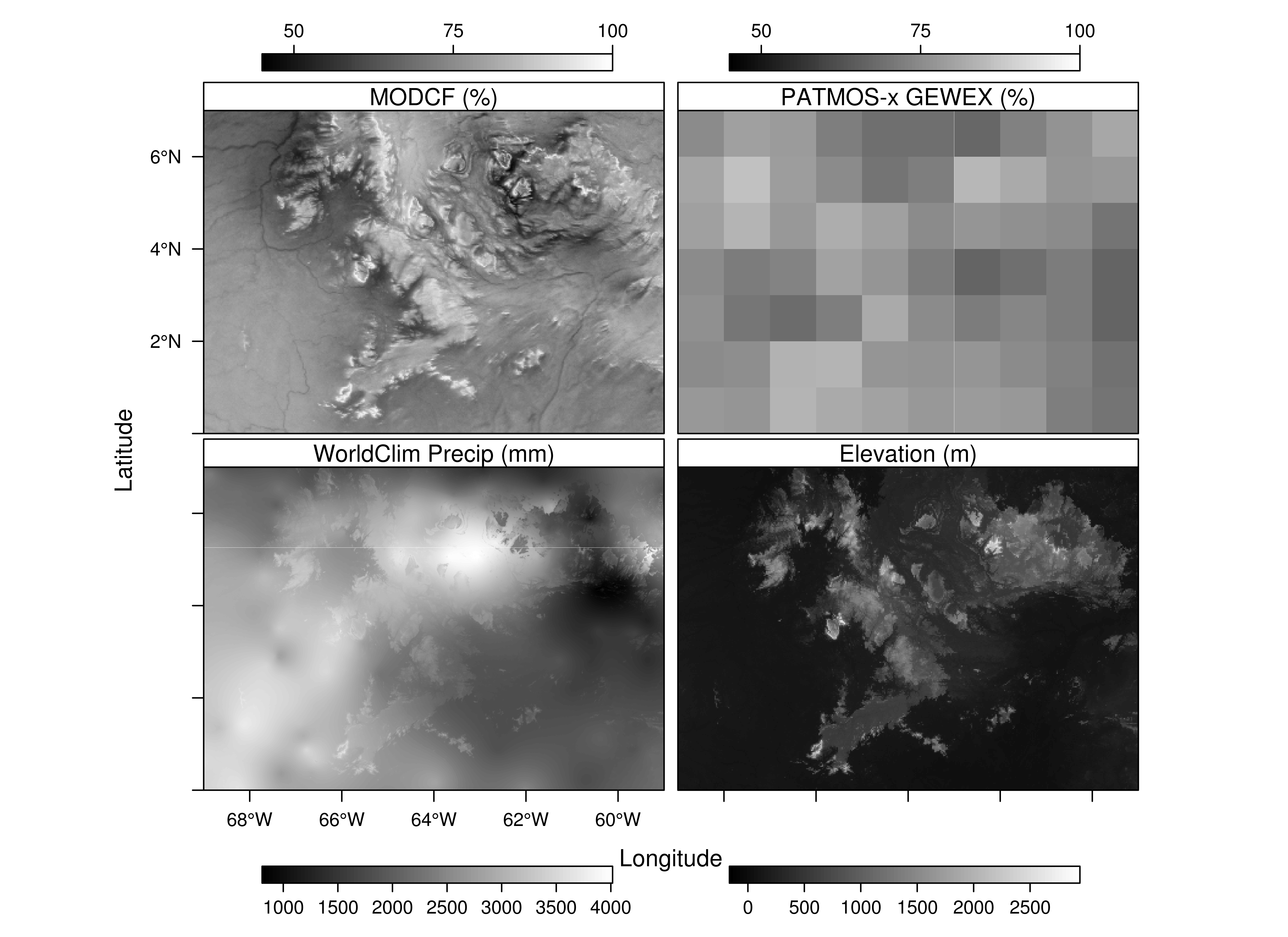


Figure : Comparison of cloud related products for a region near Parque Nacional Jaua-Sarisariñama in Southern Venezuela. Top left: MODCF developed in this paper (~1km resolution). Top right: PATMOS-x AVHRR data formatted for the Global Energy and Water cycle Experiment (GEWEX) Cloud Assessment (1 degree, ~110km). Lower left: WorldClim mean annual precipitation (mm). Lower right: Elevation (m).

### Applications to biodiversity and ecological modeling

Cloud cover is known to be an important factor for many biological processes ranging from ecosystem parameters (such as net primary productivity) to animal behavior. For example, Graham, et. al 3 used data from a micrometeorological station to determine that seasonal availability in light due to cloud cover limited CO2 uptake in a rainforest tree species. Goldsmith, Matzke, and Dawson18 compared MODIS MOD35 cloud frequency to understand how cloud cover and leaf wetting varied across sites in tropical montane and pre-montane cloud forests. They found that the satellite-derived cloud cover was a useful predictor of leaf wetness even between sites only 2km apart. They used the MOD35 cloud data only for the pixels over their forested study locations, so the land-cover biases present in the MOD35 data22 were unlikely to affect the analysis. Fischer, Still, and Williams 2 modeled drought stress in bishop pine (*Pinus* muricata) and found that persistent cloud cover near the coast of Santa Cruz Island (California, USA) reduced annual drought stress by 22-44% compared to less cloudy areas further inland. Hare and Cree 4 experimentally altered available radiation to simulate increased cloud cover and found it lowered maternal pregnancy success and slowed growth rates of female McCann’s skinks (*Oligosoma maccanni*). However, these studies all had very limited spatial extents and required either local cloud observations or extensive processing of satellite observations.

Quintero, et. al, recently evaluated the use of MODCF in a test of the “Asynchrony of Seasons Hypothesis” which suggests that seasonal differences in climate-derived food availability could lead to asynchronous breeding seasons and gene flow barriers. The study was conducted using genetic data from ~1600 individuals of 74 species of bird from locations around North and South America. WorldClim monthly precipitation 26 and the MODCF cloud climatologies were both evaluated as proxies for the seasonality of food availability and MODCF improved model performance.

CONCLUDING PARAGRAPH

solar power development27,28.

# Methods

### MOD09 Cloud Detection Algorithm

The MOD09 surface reflectance product includes an ‘internal cloud mask’ in the PGE11 program which relies on two reflective and one thermal test 29–31. The reflective tests include the shortwave and middle infrared data combined in the ‘middle infrared anomaly’ index (MIRA= ρ20,21-0.82ρ7+0.32ρ6, where ρ indicates MODIS band number). The second test uses reflectance at 1.38 microns (1.38mic=ρ26). The MIRA and the 1.38mic reflectance are designed to be complementary, with MIRA efficiently detecting low or high reflective clouds29, while 1.38mic effectively detects high (and potentially not very reflective) clouds. Additionally, a thermal test is used to identify pixels with high infrared reflectance anomalies (e.g. fires, sun-glint, and high albedo surfaces) with respect to near surface (2m) air temperature computed by the NCEP reanalysis model 32.

The daily 2000-2013 archive (approximately 260TB of data) were processed to calculate the mean and standard deviation of monthly cloud frequency.

Due to the algorithm’s use of tests based on reflectance data, the flag is only available for daytime scenes and thus high latitudes have missing data during winter months. These data are referred to below as the MODIS cloud frequency (MODCF) dataset. To illustrate and contrast the spatial variability in cloud frequency within and between Earth’s ecoregions, we summarized MODCF within each of the up to 14 biomes in each geographic ‘realm’ delineated by the “Terrestrial Ecoregions of the World” dataset33.

# Appendix figures

Figure : Locations of the stations used to validate the cloud product. ADD INDICATION OF STATION ERA (FULL, MODIS)

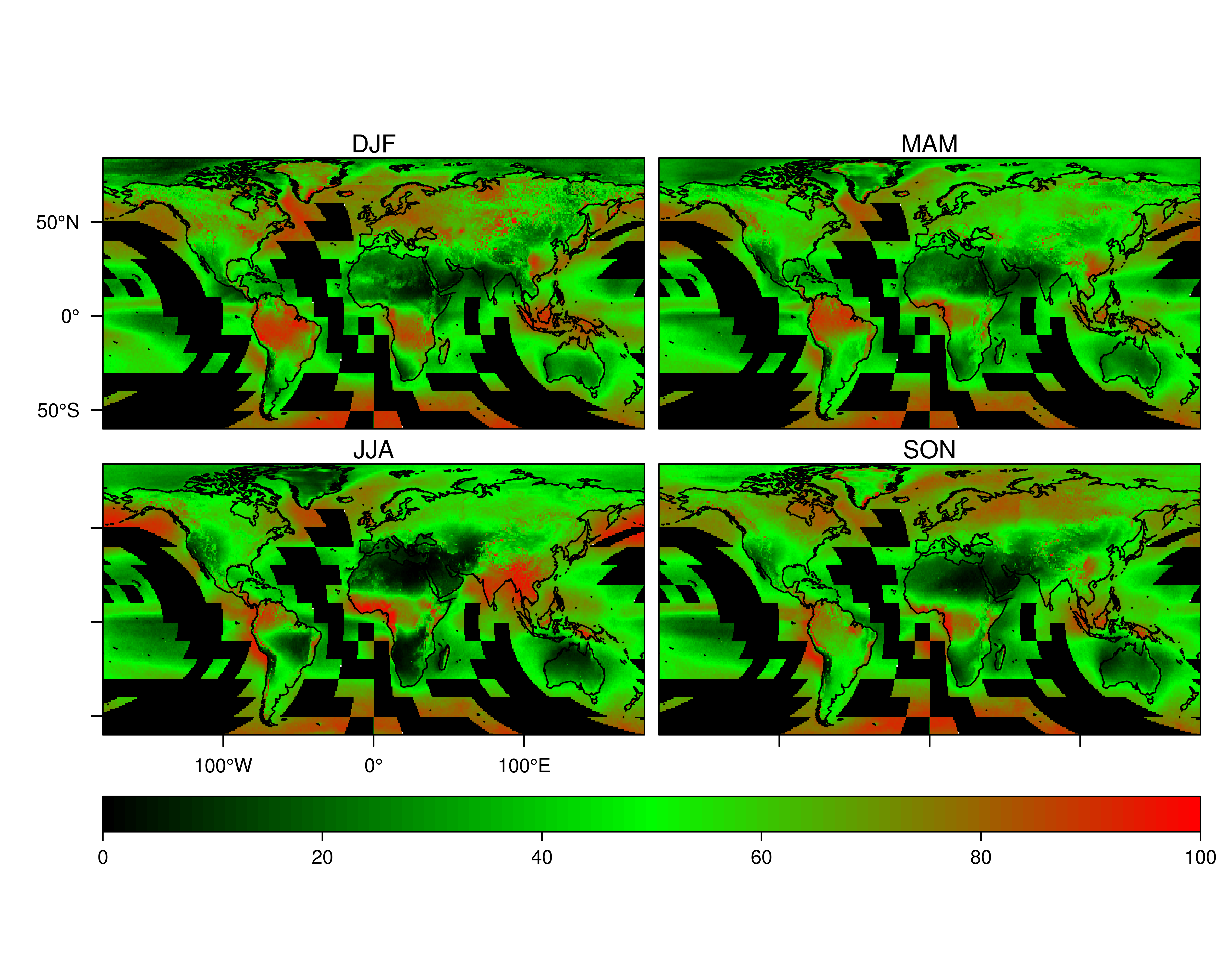


Figure : Near-global 1-km seasonal mean (DJF: December, January, February; MAM: March, April, May; JJA: June, July, August; SON: September, October, November) cloud frequency (proportion of days flagged as cloudy) derived from MODIS MOD09 internal cloud mask algorithm over 2000-2012.

# References

1. Stephens, G. L. & Kummerow, C. D. The remote sensing of clouds and precipitation from space: A review. *Journal of the Atmospheric Sciences* **64,** 3742–3765 (2007).

2. Fischer, D. T., Still, C. J. & Williams, A. P. Significance of summer fog and overcast for drought stress and ecological functioning of coastal California endemic plant species. *Journal of Biogeography* **36,** 783–799 (2009).

3. Graham, E. A., Mulkey, S. S., Kitajima, K., Phillips, N. G. & Wright, S. J. Cloud cover limits net CO2 uptake and growth of a rainforest tree during tropical rainy seasons. *PNAS* **100,** 572–576 (2003).

4. Hare, K. M. & Cree, A. Exploring the consequences of climate-induced changes in cloud cover on offspring of a cool-temperate viviparous lizard. *Biological Journal of the Linnean Society* **101,** 844–851 (2010).

5. Clench, H. K. Behavioral Thermoregulation in Butterflies. *Ecology* **47,** 1021–1034 (1966).

6. Grubb, T. C., Jr. Weather-Dependent Foraging Behavior of Some Birds Wintering in a Deciduous Woodland: Horizontal Adjustments. *The Condor* **79,** 271–274 (1977).

7. Sklenář, P., Bendix, J. & Balslev, H. Cloud frequency correlates to plant species composition in the high Andes of Ecuador. *Basic and Applied Ecology* **9,** 504–513 (2008).

8. Houze, R. A. Orographic effects on precipitating clouds. *Reviews of Geophysics* **50,** (2012).

9. Taylor, C. M., de Jeu, R. A. M., Guichard, F., Harris, P. P. & Dorigo, W. A. Afternoon rain more likely over drier soils. *Nature* **489,** 423–426 (2012).

10. Wang, J. *et al.* Impact of deforestation in the Amazon basin on cloud climatology. *PNAS* **106,** 3670–3674 (2009).

11. Ledo, A., Burslem, D. F. R. P., Condés, S. & Montes, F. Micro-scale habitat associations of woody plants in a neotropical cloud forest. *Journal of Vegetation Science* n/a–n/a (2012). doi:10.1111/jvs.12023

12. Potter, K. A., Arthur Woods, H. & Pincebourde, S. Microclimatic challenges in global change biology. *Global Change Biology* **19,** 2932–2939 (2013).

13. Stubenrauch, C. . J. *et al.* Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX Radiation Panel. *Bulletin of the American Meteorological Society* 130117123745009 (2013). doi:10.1175/BAMS-D-12-00117

14. Wylie, D., Jackson, D. L., Menzel, W. P. & Bates, J. J. Trends in global cloud cover in two decades of HIRS observations. *Journal of climate* **18,** 3021–3031 (2005).

15. Pavolonis, M. J., Heidinger, A. K. & Uttal, T. Daytime global cloud typing from AVHRR and VIIRS: Algorithm description, validation, and comparisons. *Journal of Applied Meteorology* **44,** 804–826 (2005).

16. Allard, J. & Carleton, A. Mesoscale Associations Between Midwest Land Surface Properties and Convective Cloud Development in the Warm Season. *Physical Geography* **31,** 107–136 (2010).

17. Durieux, L., Machado, L. A. T. & Laurent, H. The impact of deforestation on cloud cover over the Amazon arc of deforestation. *Remote Sensing of Environment* **86,** 132–140 (2003).

18. Goldsmith, G. R., Matzke, N. J. & Dawson, T. E. The incidence and implications of clouds for cloud forest plant water relations. *Ecology Letters* **16,** 307–314 (2013).

19. Knapp, K. R. *et al.* Globally Gridded Satellite Observations for Climate Studies. *Bulletin of the American Meteorological Society* **92,** 893 (2011).

20. Heidinger, A. K., Evan, A. T., Foster, M. J. & Walther, A. A naive Bayesian cloud-detection scheme derived from CALIPSO and applied within PATMOS-x. *Journal of Applied Meteorology and Climatology* **51,** 1129–1144 (2012).

21. Mulligan, M. MODIS MOD35 pan-tropical cloud climatology. *MODIS cloud climatology, Version 1* (2006). at <http://www.ambiotek.com/clouds/>

22. Wilson, A. M., Parmentier, B. & Jetz, W. Systematic landcover bias in Collection 5 MODIS cloud mask and derived products – a global overview. *Remote Sensing of Environment* **in press,** (2013).

23. Descloitres, J. *et al.* The MODIS rapid response project. in *Geoscience and Remote Sensing Symposium, 2002. IGARSS ’02. 2002 IEEE International* **2,** 1191–1192 vol.2 (2002).

24. Douglas, M., Beida, R. & Dominguez, A. Developing high spatial resolution daytime cloud climatologies for Africa. in *Preprints, 29th Conf. on Hurricanes and Tropical Meteorology, Tucson, AZ, Amer. Meteor. Soc. P* **2,** (2010).

25. Douglas, M. A high spatial resolution satellite-based cloud climatology for biogeographical applications. in *6th International Conference* (International Biogeography Society, 2013). at <http://www.nssl.noaa.gov/projects/pacs/web/MODIS/>

26. Hijmans, R. J. *et al.* Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25,** 1965–1978 (2005).

27. Ramachandra, T. V., Jain, R. & Krishnadas, G. Hotspots of solar potential in India. *Renewable and Sustainable Energy Reviews* **15,** 3178–3186 (2011).

28. Tapiador, F. J. Assessment of renewable energy potential through satellite data and numerical models. *Energy Environ. Sci.* **2,** 1142–1161 (2009).

29. Petitcolin, F. & Vermote, E. Land surface reflectance, emissivity and temperature from MODIS middle and thermal infrared data. *Remote Sensing of Environment* **83,** 112–134 (2002).

30. Roger, J. C. & Vermote, E. F. A Method to Retrieve the Reflectivity Signature at 3.75 μm from AVHRR Data. *Remote Sensing of Environment* **64,** 103–114 (1998).

31. Vermote, E. F. *et al.* *Land surface reflectance code development, testing and delivery*. 29 (University of Maryland, 2001).

32. Kalnay, E. *et al.* The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society* **77,** 437–471 (1996).

33. Olson, D. M. *et al.* Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience* **51,** 933–938 (2001).