# The impact of improved satellite retrievals on estimates of biospheric carbon balance

Scot M. Miller[1] and Anna M. Michalak[2]

[1]Department of Environmental Health and Engineering, Johns Hopkins University, Baltimore, MD, USA
[2]Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA

**Correspondence:** Scot M. Miller (smill191@jhu.edu, scot.m.miller@gmail.com)

**Abstract.** The Orbiting Carbon Observatory 2 (OCO-2) is NASA's first satellite dedicated to monitoring $CO_2$ from space and could provide novel insight into $CO_2$ fluxes across the globe. However, one continuing challenge is the development of a robust retrieval algorithm: an estimate of atmospheric $CO_2$ from satellite observations of near infrared radiation. The OCO-2 retrievals have undergone multiple updates since the satellite's launch, and the retrieval algorithm is now on its ninth version. Some of these retrieval updates, particularly version 8, led to marked changes in the $CO_2$ observations, changes of 0.5 ppm or more. In this study, we evaluate the extent to which current OCO-2 observations can constrain monthly $CO_2$ sources and sinks from the biosphere, and we particularly focus on how this constraint has evolved with improvements to the OCO-2 retrieval algorithm. We find that improvements in the $CO_2$ retrieval are having a potentially transformative effect on satellite-based estimates of the global biospheric carbon balance. The version 7 OCO-2 retrievals formed the basis of early inverse modeling studies using OCO-2 data; these observations are best-equipped to constrain the biospheric carbon balance across only continental or hemispheric regions. By contrast, newer versions of the retrieval algorithm yield a far more detailed constraint, and we are able to constrain $CO_2$ budgets for seven global biome-based regions, particularly during the Northern Hemisphere summer when biospheric $CO_2$ uptake is greatest. Improvements to the OCO-2 observations have had the largest impact on glint mode observations, and we also find the largest improvements in the terrestrial $CO_2$ flux constraint when we include both nadir and glint data.

## 1 Introduction

Over the past five years, the field of $CO_2$ remote sensing has evolved rapidly. The sheer number of satellites has increased with the launch of TanSat in 2016 (Yang et al., 2018), GOSAT-2 in 2018 (e.g., Nakajima et al., 2012), and OCO-3 in 2019 (e.g., Eldering et al., 2019). Several additional satellites have also been funded or proposed (e.g., Polonsky et al., 2014; Tollefson, 2016). In addition, the actual $CO_2$ observations or satellite retrievals have also been changing. Roughly once per year, the NASA Atmospheric $CO_2$ Observations from Space (ACOS) science team releases a new version of the OCO-2 and GOSAT observations that incorporates the most recent advances in the retrieval algorithm and addresses observational errors that have been identified by the scientific community (e.g., O'Dell et al., 2012). Early top-down studies of $CO_2$ fluxes using OCO-2

employed version 7 of the observations (e.g., Chatterjee et al., 2017; Crowell et al., 2019; Liu et al., 2017; Nassar et al., 2017), but the ACOS team has subsequently updated the observations through version 9 (at the time of writing).

The OCO-2 observations have changed markedly through this process. One of the largest changes occurred with the release of version 8 of the OCO-2 observations in September 2017 (Fig. 1). This update incorporated a multitude of changes to the quality control prescreening process, the forward spectroscopy model, the retrieval algorithm, and the bias correction (O'Dell et al., 2018b). These changes led to widespread improvements in the observations; version 8 has smaller random errors when compared to ground-based observations, a smaller bias between land nadir and land glint observations, and less bias across many northern high-latitude terrestrial regions (Wunch et al., 2017; O'Dell et al., 2018b). These improvements had a particularly large impact on glint mode observations. For example, a correction to the averaging kernel reduced a 0.3ppm bias in land glint data relative to land nadir (O'Dell et al., 2018b). Previously, inverse modeling studies using version 7 of the OCO-2 retrieval did not assimilate land glint and land nadir observations simultaneously due to this bias (e.g., Crowell et al., 2019). Furthermore, version 7 glint observations had biases greater than 1ppm across the southern ocean that have been remedied in version 8. These errors appeared to be due to high altitude aerosols, so the version 8 algorithm includes a new aerosol layer in the upper troposphere and lower stratosphere that has remedied many of these biases. Overall, the observations rated as good quality in version 8 are very different from those in version 7; 24% of the observations that were marked as high quality in version 7 have been marked as low quality in version 8, and 34% of the observations marked as high quality in version 8 were marked as low quality in version 7.

More recently, version 9 of the OCO-2 observations has been released in October 2018. Improvements in version 9 of the retrieval algorithm yielded smaller changes in the observations (O'Dell et al., 2018a). In particular, this version includes a correction for small-scale biases over land due to topography. Furthermore, the ACOS team relaxed a filter that discards observations collected over dark surfaces, and this change yields more observations over tropical forests (O'Dell et al., 2018a). In spite of these advances, there are still many opportunities for further improving the retrievals. For example, OCO-2 retrievals appear to show biases across most of the northern tropical oceans (O'Dell et al., 2018b).

These improvements to the observations should also improve the reliability or accuracy of $CO_2$ fluxes estimated using the observations. Several studies indicate that errors in the retrieval can have a substantial impact on the strength of the $CO_2$ flux constraint (e.g., Chevallier et al., 2007; Baker et al., 2010; Crowell et al., 2019; Miller et al., 2018). For example, Miller et al. (2018) explored the detectability of biospheric $CO_2$ fluxes using version 7 of the OCO-2 observations. They found that OCO-2 observations can be used to identify variations in biospheric fluxes within continental or hemispheric regions but that the observations have limited ability to constrain biospheric $CO_2$ fluxes across smaller regions. The authors constructed a series of synthetic data experiments to understand the most important factors limiting the $CO_2$ flux constraint; they concluded that atmospheric transport errors and prior flux errors play a role, but retrieval errors are a particularly salient factor. The OCO-2 science team is also developing an ensemble of inverse modeling estimates of $CO_2$ fluxes, and recent comparisons show results that are broadly parallel to Miller et al. (2018): inverse models provide consistent $CO_2$ flux totals for continents or hemispheres but diverge for smaller regions (e.g., Crowell et al., 2019).

The present study is a follow-up to Miller et al. (2018). We re-examine the conclusions of that study in light of recent improvements in OCO-2 observations of $CO_2$. We also identify opportunities for future improvements to the retrievals.

## 2 Methods

### 2.1 Overview

5   Uncertainties in biospheric fluxes are thought to be greater than in other $CO_2$ source types (e.g., National Research Council, 2010; Huntzinger et al., 2012; Le Quéré et al., 2018), and the $CO_2$ signal from biospheric fluxes is often larger than from other source types. Hence, we design a set of top-down experiments to examine whether we can detect variations in biospheric $CO_2$ sources and sinks within different regions of the globe and different months of the year using OCO-2 observations. In the present study, these variations are defined as any spatial or temporal patterns in $CO_2$ fluxes that have been gridded to the
10  resolution of a global atmospheric model – 1° latitude by 1.25° longitude and a 3-hourly time interval.

Detecting variations in $CO_2$ fluxes is a pre-requisite for constraining $CO_2$ budgets or flux totals; we must be able to detect variations in $CO_2$ sources and sinks across a region if we are to constrain budgets across any region of smaller size. We begin with two large hemispheric regions and then decrease the size of those regions to create increasingly challenging tests of the OCO-2 observations (Fig. 2); we use four and seven global regions, respectively, in each of these experiments. All of these
15  regions are based on a map of global biomes presented in Olson et al. (2001). The seven-region map contains broad global biomes aggregated from those in Olson et al. (2001) while the two- and four-region maps have been aggregated from Olson et al. (2001) to form even larger regions. We use a biome-based map because inverse modeling studies often estimate $CO_2$ flux totals for biome-based regions, and these regions have clear ecological significance.

We construct this set of experiments for each of the last three versions of the OCO-2 observations and examine how the
20  results change with the retrieval version. These experiments are identical except for the retrieval version used. Therefore, this setup provides a means to understand how improvements in the observations are improving the constraint on biospheric $CO_2$ fluxes. We examine these questions for each month within the year 2015 – to understand how these results vary by season and by region or biome.

### 2.2 Implementation of the top-down experiments

25  We design a regression framework to determine whether we can detect variations in $CO_2$ fluxes using OCO-2 observations. This section provides an overview of the approach, but Miller et al. (2018) provides full descriptive and mathematical detail. This regression will try to match $CO_2$ observations from OCO-2 using numerous atmospheric model outputs. Each model output estimates the enhancement in total column $CO_2$ ($XCO_2$) from fluxes in a particular region and a particular month. We generate all of these model outputs of $CO_2$ using the Parameterized Chemistry and Transport Model (PCTM) (Kawa et al.,
30  2004). The model setup used here has a spatial resolution of 1° latitude by 1.25° longitude, and we incorporate $CO_2$ fluxes at

a 3-hourly time resolution. The wind fields used to drive PCTM are from the Modern Era Retrospective-Analysis for Research and Applications (MERRA) product (Rienecker et al., 2011). This setup is identical to Miller et al. (2018).

We run many atmospheric model simulations using numerous different biospheric $CO_2$ flux estimates. The regression will try to reproduce OCO-2 observations using a linear combination of these model simulations. For example, in the seven region experiments, we use seven different geographic regions, seven biospheric $CO_2$ flux estimates, and 16 different months (September 2014-December 2015). We discard results from the first four months as model spin-up. These combinations equate to 784 total atmospheric model outputs. We further run atmospheric model simulations using a spatially and temporally constant flux in each region and each month, and we allow the regression to use these model outputs as well. The SI and Miller et al. (2018) describe the $CO_2$ flux estimates and regression in greater detail.

This approach provides a means to evaluate when and where current satellite observations can constrain variations in $CO_2$ fluxes. At least some of the atmospheric model outputs that are driven by biospheric $CO_2$ flux estimates should help reproduce the OCO-2 observations better than the model outputs that are driven by spatially and temporally constant fluxes. If so, a model with spatially and temporally variable fluxes is better able to reproduce OCO-2 observations than a model with constant fluxes. This result would imply that OCO-2 observations can be used to detect variations in biospheric $CO_2$ sources and sinks within a given region for a given month. By contrast, suppose that the atmospheric model outputs driven by biospheric $CO_2$ flux estimates do not reproduce the OCO-2 observations any better than the model outputs with constant $CO_2$ fluxes. This result would imply one of several conclusions. First, the observations may not be sensitive to fluxes from the region or month in question. This outcome may occur if the magnitude of fluxes is small in a given region or if there are no OCO-2 observations near that region. Second, errors in the atmospheric model or in the OCO-2 observations may obscure variations in $XCO_2$ that are due to $CO_2$ fluxes. Lastly, the biospheric $CO_2$ flux estimates used in the atmospheric model may not be skilled and may not reflect real-world biospheric $CO_2$ fluxes. However, in this study, we offer up seven biospheric $CO_2$ flux estimates for each region and each month, and at least one of these estimates should correlate with real-world $CO_2$ fluxes to a reasonable extent. Hence, it is unlikely that this explanation would drive the results. Rather, it is more likely that the observations are not sensitive to fluxes from a given region or that errors in the model–data system are too large.

Note that anthropogenic, biomass burning, ocean, and biospheric fluxes all contribute to $XCO_2$ observed by OCO-2, and we need to account for non-biospheric $CO_2$ fluxes in order to isolate the signal from biospheric fluxes in the regression. We model atmospheric enhancements of $XCO_2$ from anthropogenic emissions using EDGAR v4.2 FT2010 (European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL), 2013; Olivier et al., 2014), climatological ocean fluxes using Takahashi et al. (2016), and biomass burning fluxes using the Global Fire Emissions Database (GFED), version 4.1 (van der Werf et al., 2010; Giglio et al., 2013); each of these model outputs is considered in the regression.

We further implement model selection to evaluate when and where current satellite observations can constrain variations in biospheric $CO_2$ fluxes. Model selection will determine which combination of atmospheric model outputs to include in the regression based upon which best reproduces the OCO-2 observations. If this combination includes at least one biospheric $CO_2$ flux model for a given region and season, we conclude that the observations likely can be used to constrain variations in

$CO_2$ fluxes. However, if this combination does not include any biospheric $CO_2$ flux model for a given region and season, we conclude that the observations likely cannot be used to constrain flux variations for that region and season.

We specifically employ a form of model selection known as the Bayesian Information Criterion (BIC), an approach commonly used in regression modeling (e.g., Ramsey and Schafer, 2012, chap. 12) and more recently in atmospheric inverse modeling (e.g., Gourdji et al., 2012; Miller et al., 2013; Shiga et al., 2014; Fang et al., 2014; Fang and Michalak, 2015). To this end, we create different combinations of model outputs and use each combination in the regression. We score each combination based upon how well it reproduces the OCO-2 observations; combinations with a lower weighted sum of squares error receive a better score. Each combination is also scored based upon the total number of model outputs in that combination. Specifically, combinations with a greater number of model outputs receive a larger penalty for complexity, and this penalty prevents combinations that overfit the data from receiving an anomalously good score. The best combination of atmospheric model outputs is the one with the lowest score. We subsequently examine this combination and tally whether at least one atmospheric model output using a biospheric flux estimate was selected for each region and each month of the year. Miller et al. (2018) and the SI describe this approach in greater detail, including the specific equations for the BIC.

## 3 Results & discussion

### 3.1 Strength of the biospheric $CO_2$ flux constraint

The constraint on $CO_2$ fluxes using recent versions of the OCO-2 observations is a step-change improvement relative to previous versions. Overall, there was only a limited ability to detect variations in monthly $CO_2$ fluxes across individual biomes using version 7 of the retrievals (Fig. 3a-c, Miller et al., 2018). However, these capabilities have changed using versions 8 and 9 of the observations (Fig. 3d-i). Variations in $CO_2$ fluxes are detectable across tropical biomes much of the year and across temperate biomes in northern hemisphere summer when fluxes from these regions are most variable. These results imply that the updated OCO-2 observations can be used to detect and constrain variations in monthly $CO_2$ fluxes from seven biome-based regions in certain circumstances – in about two thirds of all months in the tropics and during northern hemisphere summer in the extra-tropics.

The improvement in the flux constraint is particularly evident in the four- and seven-region experiments (Figs. 3b-c and 3e-f). In the four-region model selection experiments, the OCO-2 observations can be used to detect variations in tropical fluxes for most months of the year (Fig. 3e). In other words, at least one biosphere flux model is found to explain a sufficiently large fraction of the observed variability in $XCO_2$ as to be selected via the BIC model selection procedure for the tropical regions for most months. This result indicates that spatiotemporal variability in $CO_2$ fluxes from within each of these regions is preserved in the OCO-2 observations. This represents a marked improvement over results when using observations from version 7 of the OCO-2 retrieval algorithm (Figs. 3b and 3e, Miller et al., 2018). The results using the newer versions 8 and 9 also show substantial improvements in other regions, including dryland and dry monsoon regions, temperate regions, and high-latitude regions (Figs. 3e and 3h).

The seven-region model selection experiments are an even more challenging test of current observations. These experiments examine whether we can detect spatiotemporal variations in biospheric fluxes across seven broad, aggregated global biomes. These experiments produce much better results using versions 8 and 9 of the observations. Specifically, biospheric flux models are selected across tropical and subtropical biomes for at least one month of every season. The same is true across all temperate and high-latitude biomes for a minimum of one month during northern hemisphere summer.

These improvements appear greatest across tropical biomes. There is a consistent flux signal from many tropical regions throughout the year, and hence we are able to detect variations in fluxes from tropical regions across different seasons using versions 8 and 9 of the observations. By contrast, the atmospheric signal due to biospheric $CO_2$ fluxes in northern mid- and high-latitudes has the largest absolute magnitude during northern hemisphere summer. As a result, we see a large improvement in the flux constraint in mid-latitudes in northern hemisphere summer but not in other times of year when the absolute magnitude of $CO_2$ fluxes is smaller. Furthermore, there are far fewer land nadir and land glint observations in northern mid- and high-latitudes in northern hemisphere winter relative to summer.

One notable feature of all model selection experiments is the result for dryland and dry monsoon regions (Fig. 2c). At first glance, it may appear surprising that biospheric flux models are selected for so many months in this region, given that some parts of this region are very dry and presumably have small $CO_2$ fluxes. Several semiarid regions within this classification have a very distinct monsoon that can bring over 500mm of precipitation per month (e.g., northeastern Brazil, western India, and Pakistan). As a result, there is a large spatial contrast in $CO_2$ fluxes across these regions during northern hemisphere spring and summer – large $CO_2$ uptake in places with a spring and summer monsoon and little to no fluxes in places like the Sahara or the Arabian Peninsula.

Note that the results using version 9 of the observations are not very different from those using version 8. The change in the observations between versions 8 and 9 is only incremental (e.g., Fig. 1b). Version 9 has a lower quality control threshold for surfaces with low albedo, resulting in more observations across tropical rainforests (O'Dell et al., 2018a), and this version includes a topography correction that mostly manifests at small spatial scales. The latter change could be very important for studies that estimate point sources or urban emissions using OCO-2. However, these changes are unlikely to make a large difference in this study both given the large size of the regions examined and the $1° \times 1°$ spatial resolution of the atmospheric model simulations. The SI includes a detailed discussion of the subtle differences between the model selection results using versions 8 and 9 of the observations.

## 3.2 Drivers of the results

Numerous factors affect the accuracy of $CO_2$ fluxes estimated from satellite data. These factors include the accuracy and precision of the observations, the atmospheric transport model, and the prior flux estimate used in the inverse model. Improvements in any of these inverse modeling inputs could improve the constraint on biospheric $CO_2$ fluxes. We find that recent improvements to the retrieval are having a particularly large impact on the strength of the $CO_2$ flux constraint. Furthermore, these improvements are not restricted to a single satellite like OCO-2. Rather, the ACOS retrievals and bias correction (O'Dell et al.,

2012, 2018b) will be directly applicable to other NASA carbon monitoring missions, including the recently-launched OCO-3 mission (Eldering et al., 2019) and the planned GeoCarb mission (Polonsky et al., 2014).

These improvements to the retrieval algorithm have had an effect on both glint and nadir observations from OCO-2 collected in almost every region of the globe. The sheer number of different changes makes it challenging to pinpoint exactly which have had the largest impact on the $CO_2$ flux constraint; there have been numerous updates to the quality control prescreening, the forward spectroscopy model, the retrieval algorithm, and the bias correction. Furthermore, these updates have had multiple effects on the reported $CO_2$ observations, reducing white noise, reducing bias, and changing which observations do or do not pass quality control. O'Dell et al. (2018b) detail these changes in much greater detail.

With that said, a few of these improvements appear to have a particularly salient impact on the results of this study. For example, the largest improvements have generally been to the glint mode observations. A 0.2 to 0.3 ppm bias between land nadir and land glint observations in version 7 has been remedied in version 8, and version 8 glint observations show smaller biases across many ocean regions. Furthermore, version 8 exhibits less random noise in all types of observations, but that noise reduction is largest in glint observations, both over land and over the oceans (O'Dell et al., 2018b).

Indeed, we also see the largest improvement in the flux experiments conducted in this study when we include glint mode observations. Figure 4 displays the results of the model selection experiments when the glint data are excluded. The figure shows results using version 7, and 8, and 9 of the observations. The improvement between versions 7 and 8 is much smaller when the glint observations are excluded than when they are included (Fig. 3). Even in terrestrial regions, these glint observations may play a key role in the overall flux constraint. For example, the absolute number of nadir and glint observations over land are roughly equal; there are $4.3 \times 10^6$ land nadir observations with a positive quality control flag for 2015 and $4.3 \times 10^6$ land glint observations during the same time period.

Note that this study focuses on detecting variations in $CO_2$ fluxes from terrestrial regions in individual months. To that end, certain types of flux estimation problems are beyond the scope of the current study. For example, there is strong evidence that OCO-2 observations are still biased across northern tropical oceans, and reductions in these biases could improve ocean flux estimates derived from OCO-2 (Baker, 2018; O'Dell et al., 2018b). Furthermore, there is always a possibility that the observations have a bias that is correlated across regions larger than those examined in this study. For example, the observations show a small, time-dependent drift from one year to another (O'Dell et al., 2018b). The approach used in this study would be unlikely to detect the impact of those biases.

## 4   Conclusions

$CO_2$ observations from the OCO-2 satellite have changed enormously with recent improvements to the retrieval algorithm. New observations are more self-consistent (e.g., better agreement between glint and nadir data) and compare better against ground-based observations. In some regions, these changes are comparable in magnitude to the atmospheric $CO_2$ enhancement due to biospheric $CO_2$ sources and sinks.

In this study, we specifically examine how these changes to the retrieval algorithm have improved the constraint on biospheric $CO_2$ fluxes, and we find that the improvement is large. Using observations based on version 7 of the retrieval algorithm, we find that biospheric fluxes can only be constrained across continental or hemisphere-size regions, as these observations can rarely be used to detect or constrain variations in $CO_2$ fluxes across smaller regions. By contrast, we find a step-change improvement

5   in the biospheric $CO_2$ flux constraint using updated versions of the OCO-2 observations, based on versions 8 and 9 of the retrieval algorithm. Specifically, these improvements make it possible to detect variations in $CO_2$ fluxes within seven global biome-based regions during many seasons of the year. This improvement is particularly large when both nadir and glint data are included.

This study indicates that improvements to space-based $CO_2$ observations are yielding large improvements in global mon-
10   itoring of biospheric carbon fluxes. As new $CO_2$ monitoring missions like OCO-3 and GeoCarb launch into orbit, these improvements will have a lasting impact on space-based monitoring of $CO_2$.

*Data availability.*   All OCO-2 observations are available from NASA's Goddard Earth Sciences Data and Information Services Center (GES DISC) at https://disc.gsfc.nasa.gov/OCO-2 (last access: 24 Feb. 2019).

*Author contributions.*   S.M.M and A.M.M. designed and wrote the study.

15   *Competing interests.*   The authors declare that they have no conflicts of interest.

*Acknowledgements.*   We thank Christopher O'Dell, Annmarie Eldering, and David Crisp for their feedback on the research. The OCO-2 data are produced by the OCO-2 project at the Jet Propulsion Laboratory, California Institute of Technology, and obtained from the OCO-2 data archive maintained at the GES DISC. This work is funded by NASA ROSES grant no. 80NSSC18K0976.

# References

Baker, D.: OCO-2 v9 ocean glint bias correction, in: Fall Science Team Meeting, NASA OCO-2 Science Team, Boulder CO, 2018.

Baker, D. F., Bösch, H., Doney, S. C., O'Brien, D., and Schimel, D. S.: Carbon source/sink information provided by column $CO_2$ measurements from the Orbiting Carbon Observatory, Atmos. Chem. Phys., 10, 4145–4165, https://doi.org/10.5194/acp-10-4145-2010, 2010.

5  Chatterjee, A., Gierach, M. M., Sutton, A. J., Feely, R. A., Crisp, D., Eldering, A., Gunson, M. R., O'Dell, C. W., Stephens, B. B., and Schimel, D. S.: Influence of El Niño on atmospheric $CO_2$ over the tropical Pacific Ocean: Findings from NASA's OCO-2 mission, Science, 358, https://doi.org/10.1126/science.aam5776, 2017.

Chevallier, F., Bron, F.-M., and Rayner, P. J.: Contribution of the Orbiting Carbon Observatory to the estimation of $CO_2$ sources and sinks: Theoretical study in a variational data assimilation framework, J. Geophys. Res. – Atmos., 112, https://doi.org/10.1029/2006JD007375, 10  d09307, 2007.

Crowell, S., Baker, D., Schuh, A., Basu, S., Jacobson, A. R., Chevallier, F., Liu, J., Deng, F., Feng, L., McKain, K., Chatterjee, A., Miller, J. B., Stephens, B. B., Eldering, A., Crisp, D., Schimel, D., Nassar, R., O'Dell, C. W., Oda, T., Sweeney, C., Palmer, P. I., and Jones, D. B. A.: The 2015–2016 carbon cycle as seen from OCO-2 and the global in situ network, Atmos. Chem. Phys., 19, 9797–9831, https://doi.org/10.5194/acp-19-9797-2019, 2019.

15  Eldering, A., Taylor, T. E., O'Dell, C. W., and Pavlick, R.: The OCO-3 mission: measurement objectives and expected performance based on 1 year of simulated data, Atmos. Meas. Tech., 12, 2341–2370, https://doi.org/10.5194/amt-12-2341-2019, 2019.

European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL): Emission Database for Global Atmospheric Research (EDGAR), release EDGARv4.2 FT2010, http://edgar.jrc.ec.europa.eu, 2013.

Fang, Y. and Michalak, A. M.: Atmospheric observations inform $CO_2$ flux responses to enviroclimatic drivers, Global Biogeochem. Cy., 29, 20  555–566, https://doi.org/10.1002/2014GB005034, 2014GB005034, 2015.

Fang, Y., Michalak, A. M., Shiga, Y. P., and Yadav, V.: Using atmospheric observations to evaluate the spatiotemporal variability of $CO_2$ fluxes simulated by terrestrial biospheric models, Biogeosciences, 11, 6985–6997, https://doi.org/10.5194/bg-11-6985-2014, 2014.

Giglio, L., Randerson, J. T., and van der Werf, G. R.: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), J. Geophys. Res. – Biogeo., 118, 317–328, https://doi.org/10.1002/jgrg.20042, 2013.

25  Gourdji, S. M., Mueller, K. L., Yadav, V., Huntzinger, D. N., Andrews, A. E., Trudeau, M., Petron, G., Nehrkorn, T., Eluszkiewicz, J., Henderson, J., Wen, D., Lin, J., Fischer, M., Sweeney, C., and Michalak, A. M.: North American $CO_2$ exchange: inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion, Biogeosciences, 9, 457–475, https://doi.org/10.5194/bg-9-457-2012, 2012.

Huntzinger, D., Post, W., Wei, Y., Michalak, A., West, T., Jacobson, A., Baker, I., Chen, J., Davis, K., Hayes, D., Hoffman, F., Jain, A., Liu, 30  S., McGuire, A., Neilson, R., Potter, C., Poulter, B., Price, D., Raczka, B., Tian, H., Thornton, P., Tomelleri, E., Viovy, N., Xiao, J., Yuan, W., Zeng, N., Zhao, M., and Cook, R.: North American Carbon Program (NACP) regional interim synthesis: Terrestrial biospheric model intercomparison, Ecological Modelling, 232, 144 – 157, https://doi.org/10.1016/j.ecolmodel.2012.02.004, 2012.

Kawa, S. R., Erickson, D. J., Pawson, S., and Zhu, Z.: Global $CO_2$ transport simulations using meteorological data from the NASA data assimilation system, J. Geophys. Res. – Atmos., 109, https://doi.org/10.1029/2004JD004554, d18312, 2004.

35  Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneth, A., Arora, V. K., Barbero, L., Bastos, A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney, S. C., Gkritzalis, T., Goll, D. S., Harris, I., Haverd, V., Hoffman, F. M., Hoppema, M., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Johannessen, T., Jones,

C. D., Kato, E., Keeling, R. F., Goldewijk, K. K., Landschützer, P., Lefèvre, N., Lienert, S., Liu, Z., Lombardozzi, D., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Olsen, A., Ono, T., Patra, P., Peregon, A., Peters, W., Peylin, P., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rocher, M., Rödenbeck, C., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Steinhoff, T., Sutton, A., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., Wright, R., Zaehle, S., and Zheng, B.: Global carbon budget 2018, Earth System Science Data, 10, 2141–2194, https://doi.org/10.5194/essd-10-2141-2018, 2018.

Liu, J., Bowman, K. W., Schimel, D. S., Parazoo, N. C., Jiang, Z., Lee, M., Bloom, A. A., Wunch, D., Frankenberg, C., Sun, Y., O'Dell, C. W., Gurney, K. R., Menemenlis, D., Gierach, M., Crisp, D., and Eldering, A.: Contrasting carbon cycle responses of the tropical continents to the 2015–2016 El Niño, Science, 358, https://doi.org/10.1126/science.aam5690, 2017.

Miller, S. M., Wofsy, S. C., Michalak, A. M., Kort, E. A., Andrews, A. E., Biraud, S. C., Dlugokencky, E. J., Eluszkiewicz, J., Fischer, M. L., Janssens-Maenhout, G., Miller, B. R., Miller, J. B., Montzka, S. A., Nehrkorn, T., and Sweeney, C.: Anthropogenic emissions of methane in the United States, P. Natl. Acad. Scie. USA, 110, 20 018–20 022, https://doi.org/10.1073/pnas.1314392110, 2013.

Miller, S. M., Michalak, A. M., Yadav, V., and Tadić, J. M.: Characterizing biospheric carbon balance using $CO_2$ observations from the OCO-2 satellite, Atmos. Chem. Phys., 18, 6785–6799, https://doi.org/10.5194/acp-18-6785-2018, 2018.

Nakajima, M., Kuze, A., and Suto, H.: The current status of GOSAT and the concept of GOSAT-2, Proc. SPIE, 8533, 8533 – 8533 – 10, https://doi.org/10.1117/12.974954, 2012.

Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., and Crisp, D.: Quantifying $CO_2$ emissions from individual power plants from space, Geophys. Res. Lett., 44, 10,045–10,053, https://doi.org/10.1002/2017GL074702, 2017GL074702, 2017.

National Research Council: Verifying greenhouse gas emissions: Methods to support international climate agreements, The National Academies Press, Washington, DC, https://doi.org/10.17226/12883, 2010.

O'Dell, C., Eldering, A., Crisp, D., Fisher, B., and Gunson, M.: Recent improvements and remaining challenges in the retrieval of $XCO_2$ from the Orbiting Carbon Observatory-2, in: Fall Meeting, American Geophysical Union, Washington DC, 2018a.

O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., Mc-Duffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS $CO_2$ retrieval algorithm – Part 1: Description and validation against synthetic observations, Atmos. Meas. Tech., 5, 99–121, https://doi.org/10.5194/amt-5-99-2012, 2012.

O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Mandrake, L., Merrelli, A., Natraj, V., Nelson, R. R., Osterman, G. B., Payne, V. H., Taylor, T. E., Wunch, D., Drouin, B. J., Oyafuso, F., Chang, A., McDuffie, J., Smyth, M., Baker, D. F., Basu, S., Chevallier, F., Crowell, S. M. R., Feng, L., Palmer, P. I., Dubey, M., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Roehl, C. M., Sha, M. K., Strong, K., Sussmann, R., Te, Y., Uchino, O., and Velazco, V. A.: Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, Atmos. Meas. Tech., 11, 6539–6576, https://doi.org/10.5194/amt-11-6539-2018, 2018b.

Olivier, J., Janssens-Maenhout, G., Munteam, M., and Peters, J.: Trends in global $CO_2$ emissions; 2014 Report, PBL Netherlands Environmental Assessment Agency; European Commission, Joint Research Centre, Ispra, Italy, 2014.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., and Kassem, K. R.: Terrestrial ecoregions of the world: A new map of life on Earth, BioScience, 51, 933–938, https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2, 2001.

Polonsky, I. N., O'Brien, D. M., Kumer, J. B., O'Dell, C. W., and the geoCARB Team: Performance of a geostationary mission, geoCARB, to measure $CO_2$, $CH_4$ and CO column-averaged concentrations, Atmos. Meas. Tech., 7, 959–981, https://doi.org/10.5194/amt-7-959-2014, 2014.

Ramsey, F. and Schafer, D.: The statistical sleuth: A course in methods of data analysis, Cengage Learning, Boston, MA, 2012.

5    Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, J. Climate, 24, 3624–3648, https://doi.org/10.1175/JCLI-D-11-00015.1, 2011.

Shiga, Y. P., Michalak, A. M., Gourdji, S. M., Mueller, K. L., and Yadav, V.: Detecting fossil fuel emissions patterns from subcontinental

10    regions using North American in situ $CO_2$ measurements, Geophys. Res. Lett., 41, 4381–4388, https://doi.org/10.1002/2014GL059684, 2014GL059684, 2014.

Takahashi, T., Sutherland, S., and Kozyr, A.: Global ocean surface water partial pressure of $CO_2$ database: Measurements performed during 1957-2015 (Version 2015), Tech. Rep. ORNL/CDIAC-160, NDP-088(V2015), Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee, https://doi.org/10.3334/CDIAC/OTG.NDP088(V2015), 2016.

15    Tollefson, J.: Next generation of carbon-monitoring satellites faces daunting hurdles, Nature News, 533, 446, https://doi.org/10.1038/533446a, 2016.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), Atmos. Chem. Phys., 10, 11 707–11 735, https://doi.org/10.5194/acp-10-11707-2010, 2010.

20    Wunch, D., Wennberg, P. O., Osterman, G., Fisher, B., Naylor, B., Roehl, C. M., O'Dell, C., Mandrake, L., Viatte, C., Kiel, M., Griffith, D. W. T., Deutscher, N. M., Velazco, V. A., Notholt, J., Warneke, T., Petri, C., De Maziere, M., Sha, M. K., Sussmann, R., Rettinger, M., Pollard, D., Robinson, J., Morino, I., Uchino, O., Hase, F., Blumenstock, T., Feist, D. G., Arnold, S. G., Strong, K., Mendonca, J., Kivi, R., Heikkinen, P., Iraci, L., Podolske, J., Hillyard, P. W., Kawakami, S., Dubey, M. K., Parker, H. A., Sepulveda, E., García, O. E., Te, Y., Jeseck, P., Gunson, M. R., Crisp, D., and Eldering, A.: Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) $X_{CO_2}$ measurements

25    with TCCON, Atmos. Meas. Tech., 10, 2209–2238, https://doi.org/10.5194/amt-10-2209-2017, 2017.

Yang, D., Liu, Y., Cai, Z., Chen, X., Yao, L., and Lu, D.: First global carbon dioxide maps produced from TanSat measurements, Advances in Atmospheric Sciences, 35, 621–623, https://doi.org/10.1007/s00376-018-7312-6, 2018.
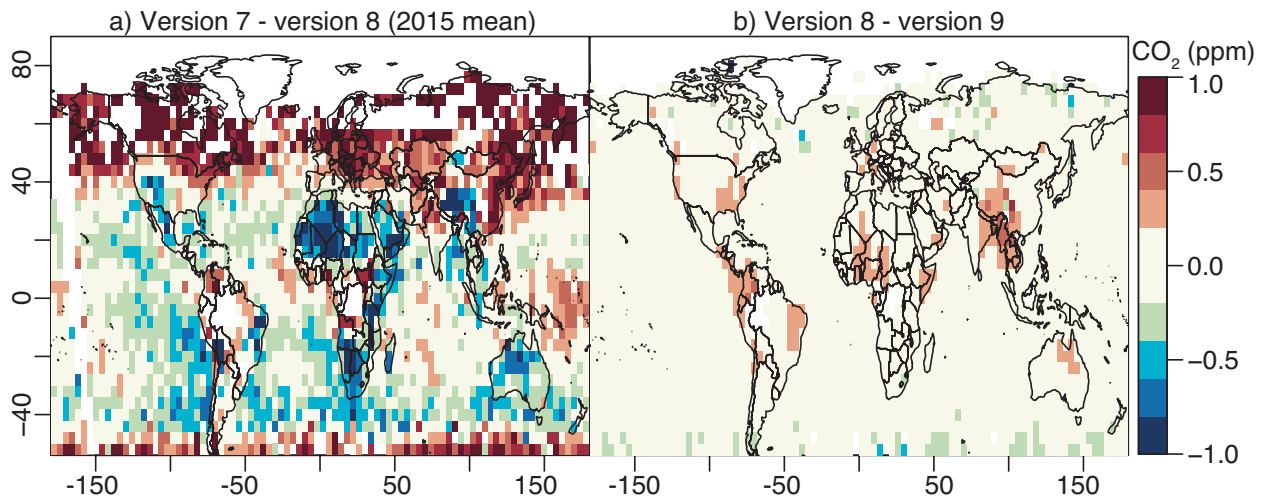
**Figure 1.** Differences between versions 7 and 8 of the OCO-2 observations (a) and between versions 8 and 9 of the observations (b). Version 8 was a much larger update to the observations than version 9. We average all of the differences between observations onto a grid to make the differences more visually apparent. The results shown here are for observations collected in 2015, the time period analyzed in this study. In addition, this map only displays grid boxes with more than 250 total observations in 2015.
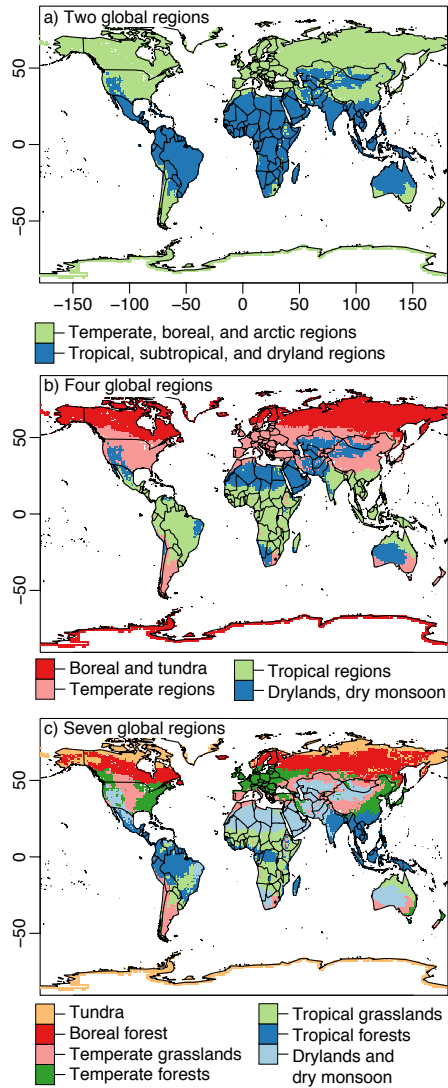
**Figure 2.** The two hemispheric regions (a), four continental regions (b), and seven biome-based regions (c) used in this study. These regions are based upon the world biome map by Olson et al. (2001). The two- and four-region maps are constructed by aggregating individual biomes into larger regions.
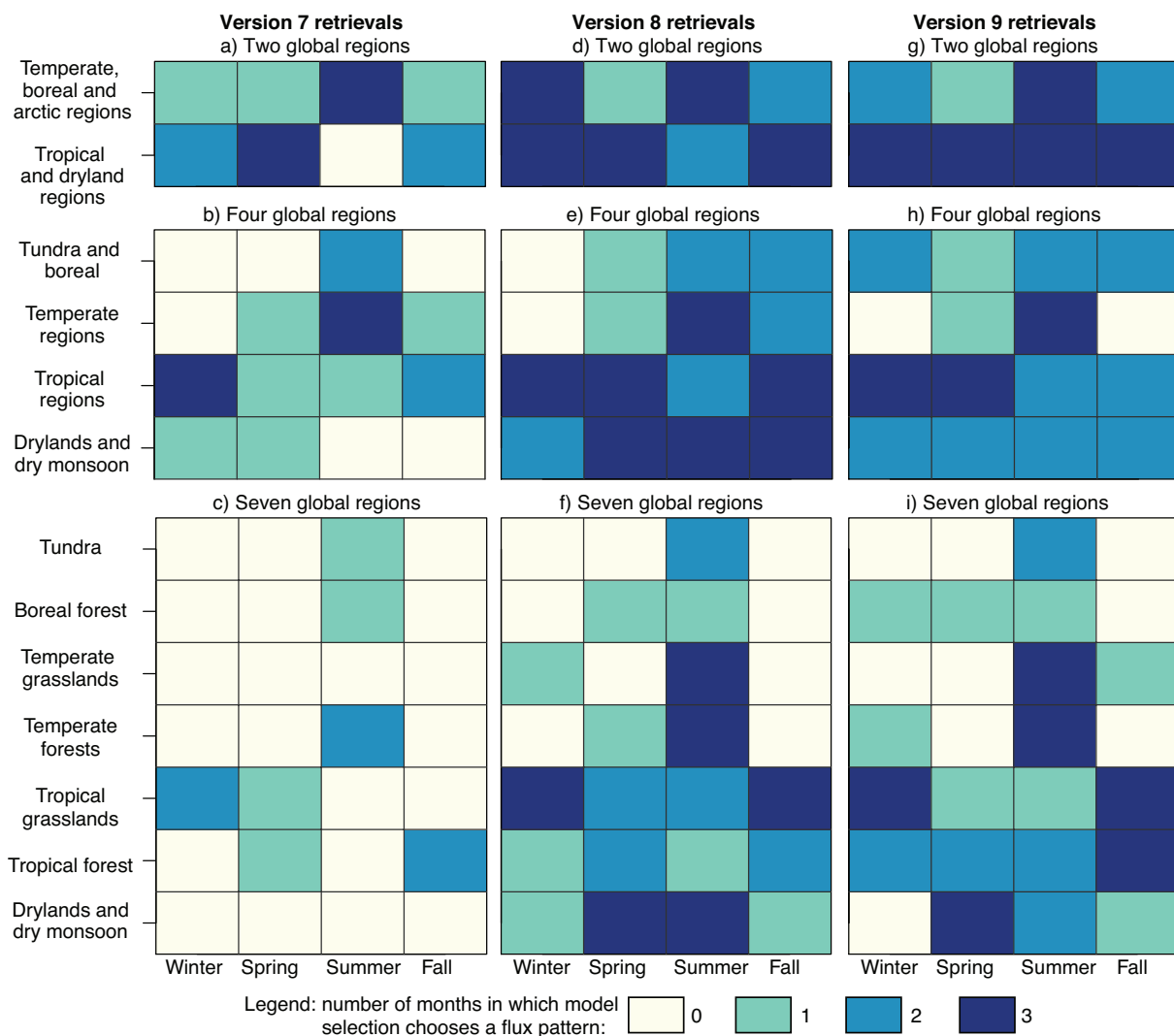
**Figure 3.** Results of the model selection experiments using versions 7, 8, and 9 of the OCO-2 observations. Versions 8 and 9 provide a much stronger constraint on biospheric $CO_2$ fluxes than version 7. The top row displays the results of the experiments with two global regions, the second row with four global regions, and the third row with seven global regions. Each box is color-coded based upon the number of months in which at least one biospheric flux model is chosen using model selection. Dark colors indicate a strong constraint on monthly $CO_2$ fluxes while light colors indicate a weak constraint. Note that these experiments include nadir, target, glint mode observations. In addition, version 7 results are the same as those in Miller et al. (2018).
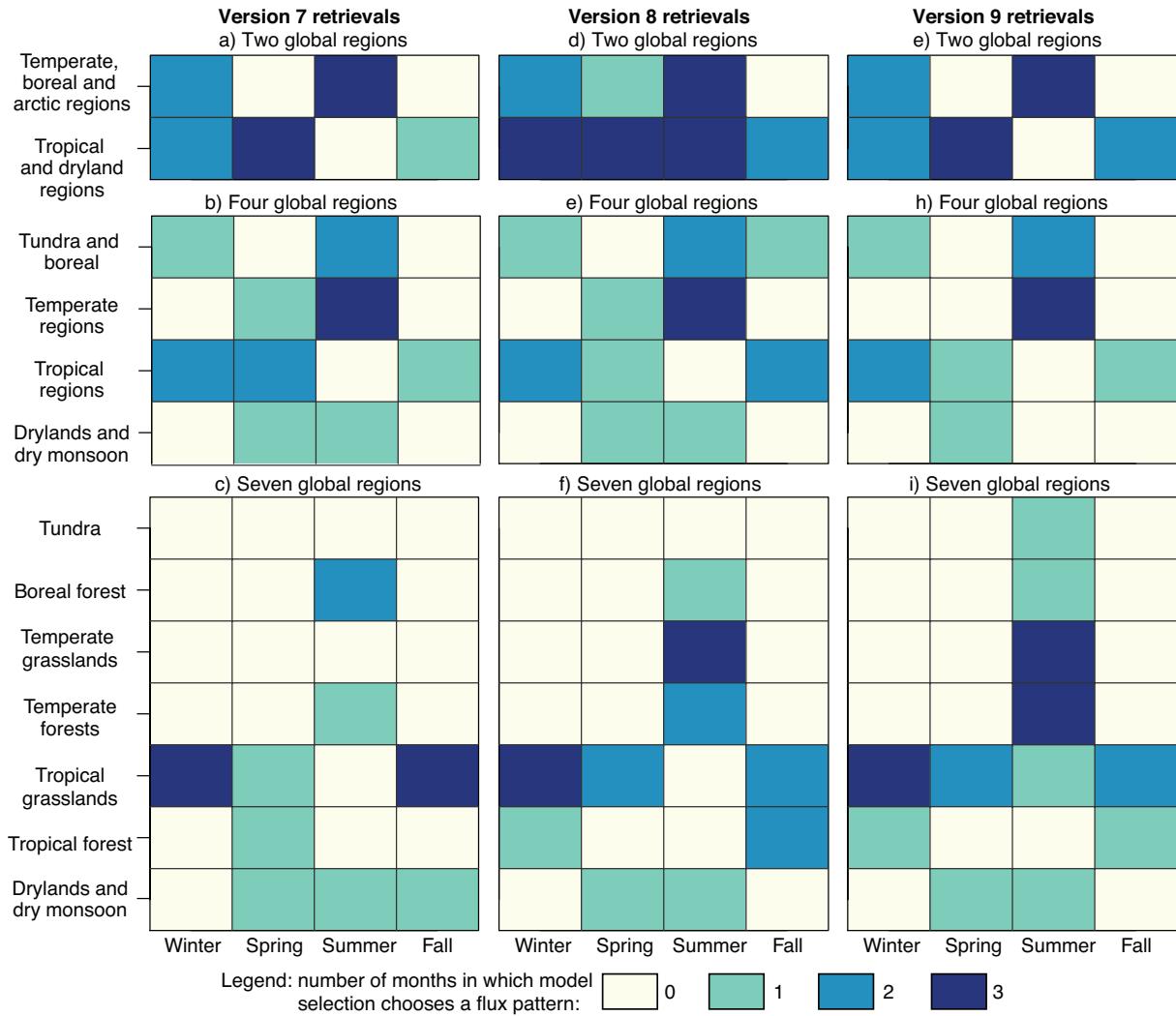
**Figure 4.** Results of the model selection experiments using only nadir and target mode observations. The improvement between versions 7 and 8 is less pronounced when we exclude glint observations and include only nadir and target mode data. Version 7 results here are the same as those in Miller et al. (2018).