Response to Referee #1

We are grateful to the reviewer for the thorough and thoughtful comments that have led to important improvements of the original manuscript. Our detailed responses to general, major and minor comments are listed below. Reviewer's comments are in italic, and anthers' responses are in dark blue font. Page and line numbers refer to discussion paper *Atmos. Chem. Phys. Discuss., 15, 11763–11797, 2015*.

1. General comments.

Zheng et al. present a study that attempts to link large-scale observations and simulations of isoprene emissions and concentrations. They do so with a regression analysis between observation-derived estimates of GPP and formaldehyde variability and their climatic drivers (as well as regression between the GPP and formaldehyde estimates), which are compared with the respective regressions obtained from an Earth system model, and with the respective regressions to simulated isoprene emissions.

The study is a brave attempt to overcome one of the largest problems in isoprene chemistry modelling: The lack of constraints from large-scale observations on the magnitude and spatial patterns of isoprene emissions (or concentrations). Although the results obtained in the study give little confidence in the possibility to obtain such constraints from observations, the attempt is worth documenting and can provide insight for many researchers in the fields of emission modelling and atmospheric chemistry modelling. As such, I think that this manuscript should be published.

However, I have a major concern that should be addressed prior to publication. The two data sets used (global estimates of gross primary production derived from upscaled eddy covariance measurements of CO2 exchange and variability of column-integrated formaldehyde content obtained from satellite observations) bear large uncertainties in their estimates, and so do the isoprene models, which is illustrated e.g. by the large spread in global isoprene emissions obtained from the three models used. These uncertainties are mentioned in the manuscript, but I think that they should be acknowledged more concretely in the methods used for the analysis. Specifically, the comparison between observations and models is performed very indirectly, by comparing correlations between the drivers and the isoprene proxies for the models and the observations separately, which is probably a result of the abovementioned uncertainties. I think the authors should discuss why they have chosen to do this rather than comparing simulated GPP or formaldehyde concentrations directly with the observations. Maybe such an attempt could be added as an illustration to show the large discrepancies in absolute amounts between models and observations.

All in all, I think that this paper, despite the lack of a strong conclusion with regard to the applicability of large-scale data sets for constraining isoprene emissions or concentrations, should be published after accounting for this, as well as for the remarks below.

Response: We agree that the datasets we use (gross primary productivity dataset derived

from FLUXNET and tropospheric formaldehyde column concentrations retrieved from OMI), as well as the isoprene models bear large uncertainties. We have added uncertainty analysis in the Methods and Results sections (see below). We have also added the direct comparison of absolute amounts of simulated and FLUXNET-derived GPP, and of simulated and satellite-retrieved tropospheric HCHO columns in the supplementary materials (see Supplementary Section S1 and Figure S1, Figure S2). However, we emphasize that direct comparison of simulated and measured absolute magnitude quantities is not a main goal of this particular study. For instance, modelers can easily tune parameters such as basal isoprene emission factors, V_{cmax} etc. to match measurements. The goal of this work is to investigate the large-scale observationally-derived climatic covariance and correlations in the photosynthesis-HCHO system, and to assess the models' ability to reproduce these observationally-derived biosphere-atmosphere system sensitivities and to expose the implications for isoprene emission.

After 11767/16, we added: "The main steps of the upscaling procedure are processing FLUXNET observational data and calculating GPP for each site, training model-tree-ensembles (MTEs) for each GPP using site-level explanatory variables, and applying the established MTEs using global gridded datasets of the same explanatory variables to obtain the global GPP estimates (Jung et al., 2011). Twenty-nine explanatory variables are used to train the MTEs, including the fraction of absorbed photosynthetically active radiation (fAPAR), precipitation, temperature and other climate and land cover data (Jung et al., 2011). The uncertainties are mainly from but not limited to (1) measurement of eddy covariance fluxes (Lasslop et al., 2008; Richardson et al., 2006), (2) the choice of explanatory variables (Jung et al., 2011), (3) gap filling and extrapolation to different environmental domains and temporal periods (Jung et al., 2009), (4) global gridded explanatory variables (Hicke, 2005; Zhao et al., 2006). The derived GPP in tropical and subtropical regions is less well constrained with observations and has larger uncertainties compared to the mid-latitudes (Beer et al., 2010; Jung et al., 2011)."

At 11768/20, we added: "The uncertainties on the gridded OMI vertical columns, **mainly due to cloud contamination, the a priori modeled isoprene emissions and the HCHO vertical column retrieval,** are estimated at 5-20%."

At 11773/23, we added: "The goal of this work is to investigate the large-scale observationally-derived climatic covariance and correlations in the photosynthesis-HCHO system, and to assess the models' ability to reproduce these observationally-derived biosphere-atmosphere system sensitivities and to expose the implications for isoprene emission. Therefore, instead of direct comparison between simulated and observed GPP and HCHO columns, we conduct a multiple linear regression (MLR) analysis in Section 3.1 and a correlation analysis in Section 3.2, and use the observationally-derived climatic covariance and correlations to evaluate the NASA ModelE2-YIBs model embedded with three isoprene algorithms. The direct comparison results are included in the supplementary materials for reference: simulated and FLUXNET-derived GPP are of comparable absolute amounts (Figure S1); while simulated tropospheric HCHO columns are considerably higher than that obtained from the OMI retrieval by about a factor of 2 (Figure S2), which is likely due to the large uncertainties in the models as well as the satellite retrieval."

In Supplementary materials, we added Section S1 to compare the direct amounts of observationally-derived and simulated GPP and HCHO columns as follows:

"**S1. Comparison of observationally-derived and simulated GPP and tropospheric HCHO columns**

"In Figure S1, we compare the simulated GPP averaged in each season (second column) with the observational-derived GPP from FLUXNET (first column), and the differences are shown in the third column. The simulated GPP in three model runs Y-PS, Y-MEGAN and Y-MEGAN-SM are identical. The model consistently overestimates GPP in the southeast US. The annual and summertime relative differences are 24% and 16%, respectively. In the Amazon, the simulated GPP is always lower than the FLUXNET-derived GPP: the differences in each season are: -1.5, -2.3, -3.1 and -1.8 g[C] m^{-2} day^{-1} and the relative difference of annual mean GPP is -30%. The model underestimates GPP in most Northern Hemisphere boreal regions in MAM and JJA. Globally, the model simulates a total GPP flux of 116.7 Pg[C] $year^{-1}$, which is 10% lower than the FLUXNET-derived GPP (129.6 Pg[C] $year^{-1}$).

"The simulated HCHO columns are considerably higher than the OMI-retrieved HCHO by about a factor of 2 (Figure S2). This discrepancy is a result of the large uncertainties in both the satellite HCHO retrieval and the HCHO chemistry in the model. In Figure S2, OMI-a, OMI-b and OMI-c represent OMI-retrieved HCHO based on the AMFs of Y-PS, Y-MEGAN and Y-MEGAN-SM, respectively. The models' seasonal averaged spatial patterns are similar. Thus, Figure S2 shows only OMI-a with Y-PS simulated HCHO columns. HCHO columns from OMI and from the models show similar spatial patterns and seasonality in the low- and mid-latitudes. In NH high-latitudes, the satellite-based HCHO columns show an opposite seasonality pattern compared to the models, i.e. higher in winter and lower in summer, which might be unrealistic due to retrieval bias under light-limited conditions (De Smedt et al., 2008; Wittrock et al., 1997)."

2. Major comments.

I consider it misleading to present the GPP product as a "global observation" (11765/22): It is essentially an empirical upscaling from a large, but geographically unevenly distributed, set of CO2 exchange-derived GPP estimates. Although the product is indeed often presented as a global one, there are large regions (notably the tropics and subtropics) for which it is poorly constrained with observations. The nature of this product, which uses amongst others climatic parameters to do the upscaling, differs fundamentally from the satellite-derived HCHO observations. It is no surprise that the product results in good partial regression coefficients (11775/4): the climatic parameters used to determine the regression are an important subset of the parameters used to perform the upscaling. I suggest that the authors highlight this difference when discussing the results of the regression and account for it in their interpretation.

Response: We agree with the referee's important point and made the following modifications.

At 11765/22, “global observations” has been changed to “global observational-based datasets”.

Included in the response to general comments above, after 11767/16, we added “…Twenty-nine explanatory variables are used to train the MTE, including the fraction of absorbed photosynthetically active radiation (fAPAR), precipitation, temperature and other climate and land cover data (Jung et al., 2011)… The derived GPP in tropical and subtropical regions is less well constrained with observations and has larger uncertainties compared to the mid-latitudes (Beer et al., 2010; Jung et al., 2011).”

At 11775/6, we removed “Several reasons are responsible for the difference in significance…”, and replaced this text with: “It is not surprising that the FLUXNET-derived GPP climatic covariance results have high statistical significance values because this product is an empirically upscaled dataset based on machine learning techniques (see Section 2.1) using a large set of climatic and land cover explanatory variables, and the driving variables T_s, PAR and P used to determine the MLR in this study are an important subset of the original explanatory variables. In contrast, remotely-sensed HCHO columns are relatively noisy due to the satellite retrieval method (Palmer et al., 2001; De Smedt et al., 2008). Other reasons for the differences in statistical significance include: satellite-based HCHO columns have many missing values due to product quality control (e.g. contamination by clouds) and the biomass burning removal (see Section 2.1); the GPP dataset has a longer record (1982-2011) than the HCHO dataset (2005-2013); unlike GPP which has a simple near-parabolic relationship with T_s, HCHO dependence on T_s is more complex. For instance, increasing T_s promotes isoprene emission and oxidation to HCHO, but also accelerates the chemical destruction of HCHO (see supplementary and Fig. S4).”

The regression analysis is consequently shown for three rather than four seasons. Why do you ignore DJF in your analysis?

Response: We ignore DJF results because in the boreal winter, most regions in the Northern Hemisphere mid- and high-latitudes have low GPP and isoprene emission values. Under such conditions, the MLR and correlation coefficients become noisy, nonrealistic and not instructive. Thus, we set a minimum threshold for GPP in all our analyses.

Before 11773/24, we added: “In the following analysis, we apply a minimum threshold for GPP (monthly mean GPP > 0.01 g[C] m^{-2} day^{-1}) to avoid inclusion of meaningless noise, especially in boreal winter (December-January-February) when most NH regions have very low GPP and isoprene emissions.”

In the description of the regression performed, it should be noted that all regressions are temporal, and that spatial regression is not attempted here (if my interpretation is correct). As the regressions are performed on standardized variables (expressed in units of standard deviations), it is the temporal variability that is addressed. I do not understand why the authors express this explicitly by referring to HCHO variability as "HCHOv", but do not use "GPPv" for GPP variability.

Response: We agree that in the regression analysis, the regression coefficients are expressed in units of standard deviations and represent temporal variability. "HCHOv" refers to HCHO zonal anomalies, not the temporal variability. Before we performed the regression, we first calculated "HCHOv", defined as the difference between the total HCHO column in each grid point and the zonal mean HCHO column (see 11768/24). The goal of using "HCHOv" instead of the absolute HCHO columns is to remove the HCHO background, mostly from CH_4 oxidation, and to better represent non-CH_4 VOC emissions, which are dominated by isoprene.

To improve clarity in the definition and its limitations, after 11768/27, we added: "There are two main limitations in using HCHOv as a proxy for isoprene emission: (1) HCHO from CH_4 oxidation is not strictly zonally uniform, thus HCHOv does not purely represent the influence of the non-methane VOCs; (2) HCHOv is dominated by isoprene emission but their relationship is smeared by other VOCs such as biogenic terpenes and anthropogenic VOCs."

The application of a soil moisture dependence in the models is presented as a crucial aspect to get correct large-scale estimates, but the study does not prove that such a dependence is necessary for obtaining the right response - the improved correlation may be the result of other differences between these algorithms (e.g., the soil moisture-independent algorithm Y-MEGAN obtains considerably higher isoprene emissions). The improved correlation does not necessarily point at a causal effect, and interpretation of this should be done with care. Also, the discussion of water stress effects on isoprene is rather concise: There is ample literature on these effects at smaller scales, which could be used here to support the conclusions regarding soil moisture.

Response: We agree that comparing Y-PS and Y-MEGAN solely does not prove the necessary role of soil moisture, because the soil moisture dependence is not the only difference between Y-PS and Y-MEGAN. Therefore, we conducted the third experiment Y-MEGAN-SM, whose only difference with Y-MEGAN is the soil moisture dependence of isoprene emission. Compared to Y-MEGAN, Y-MEGAN-SM demonstrates better skill at reproducing the GPP-HCHOv correlations. Thus the Y-MEGAN-SM model experiment provides supporting evidence for the importance of soil moisture. For instance, Y-MEGAN-SM behaves more like Y-PS, i.e. ISOPe is correlated with GPP rather than HCHOv. This result illustrates the importance of water availability in modeling the photosynthesis-isoprene-HCHO system. Of course, we certainly agree with the referee that correlation does not imply causality. Indeed, the Y-MEGAN-SM experiment still does not perform as well as Y-PS in reproducing the observationally-derived correlations.

We have already discussed possible other effects e.g. at 11777/8:
"Generally, overestimates of GPP-HCHOv r values in the models may be due to over-simplified parameterizations of biogeochemical processes and photochemical oxidation mechanisms, missing, possibly important but highly uncertain, processes in the models, for instance nutrient availability, and the use of generic PFT-specific isoprene emission potentials."

To emphasize the use of the Y-MEGAN-SM experiment, at 11779/1, we added: "**Since the only difference between Y-MEGAN-SM and Y-MEGAN is the soil moisture dependence of isoprene emission,** this result suggests the importance of water availability…"

After 11779/19, we added: "These results suggest water availability could be an important driver of isoprene emission on intraseasonal to interannual time scales."

Before 11779/20, we added the following paragraph about observational evidence of the water stress impacts on isoprene emission: "Multiple field experiments have studied the isoprene response to water deficit conditions on different time scales. Short-time mild drought stress on time scales of a few days affects stomatal conductance and thus the rate of photosynthesis, while does not essentially diminish isoprene emission because photosynthetic electron transport is not inhibited (Fall and Monson, 1992; Niinemets, 2010). Several studies found increases in isoprene emission during the initial stages of mild drought conditions (e.g. Brilli et al., 2007; Pegoraro et al., 2004; Sharkey and Loreto, 1993). Severe drought or prolonged moderate drought conditions on time scales of weeks do result in significant reductions in isoprene emission that are presumably due to decreased leaf carbon availability following sustained reductions in photosynthetic rate (e.g. Brüggemann and Schnitzler, 2002; Funk et al., 2005; Sharkey and Loreto, 1993). Therefore, on the short time scales of a few days, there is a lag between isoprene emission and photosynthetic rate in response to water stress. On longer time scales (weeks to months), isoprene emission is tightly coupled with photosynthesis, both of which are limited by soil moisture deficit. Recent studies have shown the importance of water availability on photosynthesis on interannual scales: Jung et al. (2011) suggest the interannual variability of GPP in semi-arid to semi-humid regions is more sensitive to precipitation rather than temperature; Beer et al. (2010) find that GPP over 40% of the vegetated land is associated with precipitation. Therefore, despite the current lack of direct observations to constrain the soil moisture impact on the interannual variability of isoprene emission, we argue that water availability is likely to be a critical factor regulating isoprene emission on longer times."

3. Minor comments

11765/15: Please define "short-term" and "long-term"

Response: We modified: "While short-term **(hours to days)** weather-related fluctuations in isoprene emission in the temperate zone are well understood (Guenther et al., 1995, 1991), many open questions remain as to the long-term **(months to years)** factors controlling isoprene emission."

11769/27: Is the PFT-specific LAI a value that is globally applicable, or do you deter mine it for each grid cell separately?

Response: PFT-specific LAI is determined for each grid cell separately. We emphasize by adding: "**Gridded spatially-varying** PFT-specific leaf area index (LAI) is derived from…."

11770/3: It seems somewhat crude to refer to the Farquhar model as a Michaelis-Menten model: It bears important elements from Michaelis-Menten kinetics to determine e.g. temperature dependences of the involved enzymes, but the model contains many important elements that are not related to Michaelis-Menten kinetics.

Response: Corrected. Now it's referred as the "Farquhar leaf model".

11771/17: The semicolon at the end of the line is not appropriate here, as there is a new part of the equation starting here. Please add a few words to introduce the empirical parameters C_T and C_L.

Response: We added: "where E is the PFT-specific isoprene emission potential in units of μmolC m^{-2} s^{-1}; **C_T and C_L are defined as follows:**".

References

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rodenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I. and Papale, D.: Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate, Science (80-.)., 329(5993), 834–838, doi:10.1126/science.1184984, 2010.

Brilli, F., Barta, C., Fortunati, A., Lerdau, M., Loreto, F. and Centritto, M.: Response of isoprene emission and carbon metabolism to drought in white poplar (Populus alba) saplings, New Phytol., 175(2), 244–254, doi:10.1111/j.1469-8137.2007.02094.x, 2007.

Brüggemann, N. and Schnitzler, J. P.: Comparison of isoprene emission, intercellular isoprene concentration and photosynthetic performance in water-limited oak (Quercus pubescens Willd. and Quercus robur L.) saplings, Plant Biol., 4(4), 456–463, doi:10.1055/s-2002-34128, 2002.

Fall, R. and Monson, R. K.: Isoprene emission rate and intercellular isoprene concentration as influenced by stomatal distribution and conductance., Plant Physiol., 100(2), 987–992, doi:10.1104/pp.100.2.987, 1992.

Funk, J. L., Jones, C. G., Gray, D. W., Throop, H. L., Hyatt, L. a. and Lerdau, M. T.: Variation in isoprene emission from Quercus rubra: Sources, causes, and consequences for estimating fluxes, J. Geophys. Res. D Atmos., 110(4), 1–10, doi:10.1029/2004JD005229, 2005.

Hicke, J. A.: NCEP and GISS solar radiation data sets available for ecosystem modeling: Description, differences, and impacts on net primary production, Global Biogeochem. Cycles, 19(2), 1–18, doi:10.1029/2004GB002391, 2005.

Jung, M., Reichstein, M. and Bondeau, A.: Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, Biogeosciences Discuss., 6(3), 5271–5304, doi:10.5194/bgd-6-5271-2009, 2009.

Jung, M., Reichstein, M., Margolis, H. a., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F. and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, J. Geophys. Res. Biogeosciences, 116, 1–16, doi:10.1029/2010JG001566, 2011.

Lasslop, G., Reichstein, M., Kattge, J. and Papale, D.: Influences of observation errors in eddy flux data on inverse model parameter estimation, Biogeosciences Discuss., 5(1), 751–785, doi:10.5194/bgd-5-751-2008, 2008.

Niinemets, Ü.: Mild versus severe stress and BVOCs: thresholds, priming and consequences, Trends Plant Sci., 15(December), 145–153, doi:10.1016/j.tplants.2009.11.008, 2010.

Palmer, P. I., Jacob, D. J., Chance, K., Martin, R. V., Spurr, R. J. D., Kurosu, T. P., Bey, I., Yantosca, R., Fiore, A. M. and Li, Q.: Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, J. Geophys. Res., 106, 14,539–14,550, doi:10.1029/2003JD003652, 2001.

Pegoraro, E., Rey, A., Greenberg, J., Harley, P., Grace, J., Malhi, Y. and Guenther, A.: Effect of drought on isoprene emission rates from leaves of Quercus virginiana Mill., Atmos. Environ., 38(36), 6149–6156, doi:10.1016/j.atmosenv.2004.07.028, 2004.

Richardson, A. D., Hollinger, D. Y., Burba, G. G., Davis, K. J., Flanagan, L. B., Katul, G. G., William Munger, J., Ricciuto, D. M., Stoy, P. C., Suyker, A. E., Verma, S. B. and Wofsy, S. C.: A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes, Agric. For. Meteorol., 136(1-2), 1–18, doi:10.1016/j.agrformet.2006.01.007, 2006.

Sharkey, T. D. and Loreto, F.: Water stress, temperature, and light effects on the capacity for isoprene emission and photosynthesis of kudzu leaves, Oecologia, 95(3), 328–333, doi:10.1007/BF00320984, 1993.

De Smedt, I., Müller, J.-F., Stavrakou, T., van der A, R., Eskes, H. and Van Roozendael, M.: Twelve years of global observation of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, Atmos. Chem. Phys. Discuss., 8, 7555–7608, doi:10.5194/acpd-8-7555-2008, 2008.

Wittrock, F., Richter, A., Ladstätter-weißenmayer, A. and Burrows, J. P.: Global Observations of Formaldehyde, Science (80-.)., (2), 9–9, 1997.

Zhao, M., Running, S. W. and Nemani, R. R.: Sensitivity of Moderate Resolution Imaging Spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological reanalyses, J. Geophys. Res. Biogeosciences, 111(1), doi:10.1029/2004JG000004, 2006.