Response to reviewer comments to manuscript "Reviews and syntheses: Systematic Earth observations for use in terrestrial carbon cycle data assimilation systems"

We thank the two reviewers for their careful inspection of the manuscript. In the following we address their comments point-by-point. We use text in italics to repeat the reviewer comments, normal text for our response, and bold faced text for quotations from the manuscript, with added text marked in colour.

Comments by referee 1, Natasha MacBean:

Firstly, I appreciate the distinction the authors try to make between their review and that of Raupach et al. (2005) and Ciais et al. (2014) by placing an emphasis on EO data versus in situ atmospheric CO2 and eddy covariance data; however, these two sources of data are one of the most widely used in carbon cycle DA studies, and therefore I think it is worth having a separate section that briefly summarizes these data and their uncertainties, while keeping the focus on EO. Otherwise the description of updates to the eddy covariance uncertainty estimates in the general section on observational uncertainties (Section 3.1) could feel out of place. In addition, Section 3.2 discusses operational carbon observing systems which currently include many in situ networks.

We have extended the description of the in situ atmospheric CO_2 and eddy covariance data in the beginning of Section 3 such that the update to the eddy covariance uncertainty estimates does not feel out of place. However, we do not include a whole new section on those data as it has extensively been covered elsewhere. Also, we have removed Section 3.2 and included a shortened version of the text in the Conclusions with reference to other international observation networks.

Section 3.1

As mentioned before, Raupach et al. (2005) have already reflected on the main properties of the data and their error covariances for observations of remotely sensed land surface properties (mainly the normalised differential vegetation index, NDVI), atmospheric CO_2 concentrations, land atmosphere net CO_2 exchange fluxes, and terrestrial carbon stores.
The in-situ measurements of CO_2 concentrations are either based on flask samples or on continuous monitoring stations. The flask sampling network was established in 1961 by Keeling (1961) and has been extended since then to more than 200 sites globally. The continuous in situ network provides measurements at higher precisions and temporal resolution than the flask networks. For both the flask and the continuous stations improvements in precision and in the accuracy have been achieved through propagation of frequent comparisons and international standards (Francey et al., 2001).
The global FluxNet network consists of more than 200 sites globally measuring land-atmosphere fluxes of CO_2, latent and sensible heat and others by the eddy covariance technique at a half-hourly temporal resolution (Baldocchi et al., 2001). Many other (mostly meteorological) variables are measured at these sites as well. In the past years, there has been substantial progress in the homogenisation and availability of these direct CO2 flux measurements. The publically available FLUXNET2015 data set includes more

than 1500 site-years of data covering all major biome types from about 165 sites worldwide spanning a period from1991 (for some sites) up to 2014 (Pastorello et al., 2017). There has also been substantial progress in the specification of uncertainties in eddy-covariance measurements of the land-atmosphere net CO_2 exchange flux (Net Ecosystem Productivity, NEP) and its component fluxes (GPP and ecosystem respiration, Reco). For instance, Lasslop et al. (2008) analysed the error distribution and found that the eddy flux data can almost entirely be represented by a superposition of Gaussian distributions with inhomogeneous variance. Richardson et al. (2008) showed that the measurement errors in NEP are heteroscedastic, i.e. the error variance varies with the magnitude of the flux. In a more recent study Raj et al. (2016) investigated the uncertainty of GPP derived from partitioning the eddy covariance NEP measurements. They used a light-use efficiency model embedded in a Bayesian framework to estimate the uncertainty in the separated GPP from the posterior distribution at half-hourly time steps.

Conclusions

In the context of carbon cycle data assimilation this paper reviews the requirements and summarises the availability and characteristics of some selected observations with a special focus on remotely sensed Earth observation data. Observations are key for understanding the carbon cycle processes and are an important component for any data assimilation system. In this context the provision of systematic and sustained observing systems on an operational basis is becoming more and more important.

An example for such an operational network for in situ data is the Integrated Carbon Observing System (ICOS, see also https://www.icos-ri.eu). ICOS is a pan-European infrastructure for carbon observations, which provides high-quality in situ observations (both fluxes as well as atmospheric concentrations) over Europe and over ocean regions adjacent to Europe with a long-term perspective. ICOS consists of central facilities for co-ordination, calibration and data in conjunction with networks of atmospheric, oceanic and ecosystem observations as well as a data distribution centre, the Carbon Portal, providing discovery of and access to ICOS data products such as derived flux information. Other (quasi-)operational networks measuring atmospheric CO_2 concentrations are maintained, for instance, by the National Oceanic and Atmospheric Administration (NOAA) and the Scripps Institution of Oceanography, both USA, and the CSIRO Global Atmospheric Sampling Laboratory, Australia.

An example for an operational space-based Earth observing programme in Europe is the fleet of so-called Sentinel satellites of the Copernicus programme. Copernicus aims at providing Europe with continuous and independent access to Earth observation data and associated services (transforming the satellite and additional in situ data into value-added information by processing and analysing the data) in support of Earth System Science (Berger et al., 2012). Currently, six different Sentinel missions are planned (and have partly been launched). So far, a dedicated mission for monitoring the carbon cycle, i.e. an instrument measuring the atmospheric CO_2 composition, is not yet included in the Copernicus monitoring programme (see Ciais et al., 2015), however, the series of Sentinel satellites will be extended in the future and likely include a CO_2 mission. Other operational EO programmes are operated by e.g. NOAA and the Japanese Aerospace Exploration Agency.

Secondly, I suggest a slight re-structuring so all the examples of DA studies with these data are incorporated into one specific section, and possibly after the description of the different types of observations. Currently, there are examples in Section 2.3 and the introduction to 3.3. Whilst the examples given in the latter are specifically pertinent to EO data, the use of EO data in an assimilation system has been discussed already in Section 2.3, and therefore the lines are somewhat blurred.

As suggested, we have slightly restructured the DA examples in the manuscript, but we have kept them in two places (Sections 2.3 and 3.2) to distinguish between general examples (Sec 2.3) and examples making use of the EO data on which this manuscript focusses (Sec. 3.2). We have also clarified this approach in the manuscript.

Section 2.3

Recent advances focus on multiple independent data stream assimilation to provide a more rigorous constraint on the multiple components of terrestrial ecosystem models and avoid equifinality, i.e. different parameter solutions providing the same cost function value at the minimum. Examples for such studies on local/regional scale are the assimilation of eddy covariance CO_2 fluxes together with observations of vegetation structural information or carbon stocks (e.g. Richardson et al., 2010; Keenan et al., 2012; Thum et al., 2017). The assimilation of multiple data streams can be performed either in a step-wise (e.g. Peylin et al., 2016) or simultaneous approach (e.g. Kaminski et al., 2012); in the case of non-linear models or non-linear observation operators only the simultaneous assimilation makes optimal use of the observations (MacBean et al., 2016). In Section 3.2 we provide more terrestrial carbon cycle data assimilation examples using some of the remotely sensed products discussed in the following.

Section 3.2

FAPAR has already been demonstrated to provide a strong constraint on terrestrial carbon and water fluxes through its impact on the phenology components of the carbon cycle model either by assimilating only FAPAR data (e.g. Knorr et al., 2010) or in combination with other data streams (e.g. Kaminski et al., 2012; Kato et al., 2013; Forkel et al., 2014).

Finally, it would be good to include websites/references for data access in all data tables (as in Table 1), and, given the important emphasis on observation uncertainty, note if uncertainty estimates come with the data.

Except for Table 3 all other Tables contain websites or references to the data products. Since SIF is a relatively new product data access is distributed among several websites (not official ones in some case), which may not be maintained after some years. Therefore, we did not include URLs but nevertheless added example references to Table 3.

Lines 109-111: worth pointing out that a better fit between the posterior maximum likelihood simulation and the observations does not necessarily mean you have the correct parameters and/or model structure (e.g. MacBean et al., 2016).

Included this point as suggested:

In contrast, data assimilation, in particular when used for parameter optimisation, potentially identifies structural model and/or data deficiencies if the model-data mismatch (or the benchmark test) is still inadequate after optimisation (see also Figure (1)). On the other hand, a better fit between the posterior maximum likelihood simulation (i.e. using the optimised parameters) and the observations is not necessarily an indication for correct parameters and/or model structure as has been pointed out by MacBean et al. (2016).

Section 2.2: The distinction between sequential and variational DA could be slightly confusing for the lay reader. I suggest the following:
Line 133: make it clear that sequential assimilation happens at the point of having an observation – otherwise one may wonder "at which discrete time steps?"

To make the distinction between sequential and variational DA clearer, we changed this sentence to:
We distinguish two basic approaches in data assimilation: sequential assimilation, which assimilates observations subsequently at discrete model time-steps, and variational assimilation,…

Lines 137-139: I think this could read as if J is only evaluated in the variational approach (though that may be helped by changing the caption of Figure 1 – see below). I suggest that instead of just discussing the inner loop you could make a distinction about when J is evaluated and at what point the minimum is found for both approaches. In addition, it might be helpful to the reader to have a sentence that qualitatively describes what the cost function represents and to explicitly say that the aim is to minimize the cost function around lines 132-139.

Changed the wording to:
In the sequential approach the assimilation loop is evaluated sequentially over time following the dynamics of the model. In the case of variational assimilation the assimilation loop is evaluated iteratively (assuming a non-linear model). Both cases evaluate a cost function J, formulated in the Bayesian framework as:

Figure 1: I like this figure, but I cannot see a "Model-data comparison" box as you describe in the caption. I guess you mean "Evaluation of J"?

Indeed, corrected. We have also slightly updated the figure and caption following the suggestion by referee 2.
The loop between the 'Evaluation of J' box to 'Model and observation operator' box) indicates the assimilation process (assimilation loop). Often, the analysis of residuals in model-data comparison leads to either model improvements or adjustment of the measurement strategies ('model improvement' and 'adjusting measurement strategy' arrows).

Section 2.3: Line 195: Maybe add paper by Bloom and Williams (2015) and latest CLM paper by Post et al. (2017)?

Added the Post et al. (2017) reference; the paper by Bloom and Williams (2015), although it is also a model-data fusion study, has a slightly different focus insofar as it uses ecological 'common sense' constraints and may not fit that well to the context here.

Line 200: maybe worth adding ". . .same cost function value at the minimum"?

Added as suggested.

Line 203: Could add Thum et al. (2017) here.

Added as suggested.

Section 3.1: Worth mentioning that observation errors in a DA system should include the models errors, and what could give rise to errors in the model?

We have included the following short paragraph on model errors:
These off-diagonal elements are usually hard to specify, but they are important to quantify in a data assimilation system because they have considerable impact on the solution because of their influence on the weight of the respective observations in the cost function.
In addition to the observational errors, models also have errors, which, in a data assimilation system, are usually included in the observation errors. These errors in dynamical models are mainly caused by process parameterizations (instead of resolving the process), and by the discretization of analytical dynamics into a numerical model. A more detailed description of the different model error sources is given in Scholze et al. (2012).

Section 3.2: This section is very focused on Europe. It would be worthwhile detailing efforts that are underway in other regions, e.g. example the NASA Carbon Monitoring System (http://carbon.nasa.gov). This section also feels a little out of place. I might suggest incorporating it into the introduction to Section 3 or having it as a perspectives section at the end of the article.

We have removed this section and moved a shortened version of the text to the Conclusions to give a perspective on operational monitoring systems at the end of the article as suggested, see also answer to the first comment above.

Section 3.3: Lines 306-308: worth pointing out that using level 2 products may increase the observation uncertainty, particularly given parameters/processes implemented in retrieval algorithm may not be consistent with corresponding equivalent parameters/processes in the underlying model (and that this may be a benefit of using level 1 products – e.g. Quaife et al., 2008). Also perhaps worth explaining that for vegetation activity that VIs are an intermediate step in that they are "lower order" products – i.e. they are raw radiances but also do not require a complex retrieval algorithm; instead they require an atmospheric transport model and limited calculations.

We are not quite sure what the referee is referring to here. The study by Quaife et al. (2008) is based on Level 2 data and not Level 1. Nevertheless, we have included a sentence on the possible inconsistencies when using Level 2 data.
However, there is the risk that when using level 2 or higher products the parameters/processes implemented in the retrieval algorithm may not be consistent with corresponding equivalent parameters/processes in the underlying model, and thus cause additional errors in the assimilation.

Line 316-318: worth including that NDVI has also been used (e.g. MacBean et al., 2015a), and the advantages/disadvantages of using Vis.

We have included the reference and referred to Section 3.3.2 (where we describe the disadvantages of such VIs over a physically based quantity such as FAPAR.) for a discussion on the difference between VIs and FAPAR.
…and recently developed products based on biogeochemical processes, such as sun-induced fluorescence (SIF). Leaf area index (LAI, e.g. Liu et al., 2014), which is in effect closely related to FAPAR, is another geophysical parameter representing vegetation activity. There is also a range of remotely sensed vegetation indices, of which NDVI is an example. Both LAI and NDVI have been used in data assimilation studies: an example for NDVI is given by MacBean et al. (2015) and for LAI by Luke (2011) and Barbu et al. (2014). In Section 3.2.2 we detail the difference between NDVI and FAPAR, and explain that FAPAR is based on physical principles. FAPAR has already been demonstrated to provide a strong constraint…

Line 318: Forkel et al. (2015) is another example of the use of FAPAR with a terrestrial model.

Included the Forkel et al. (2014) reference.

Line 321: and by optimizing parameters related to phenology and photosynthesis (MacBean et al., 2015b).

We prefer not to cite a conference presentation.

Line 322: Saying "Also assimilation of XCO2" comes a bit out of the blue here as you have just been talking about vegetation activity. Please could you say what is meant by XCO2, or refer to section 3.3.1.

Changed the beginning of the sentence and added a reference to section 3.3.1:
Remotely sensed atmospheric CO_2 concentration (XCO_2, see Section 3.3.1) has also been assimilated into a diagnostic terrestrial carbon cycle model to derive net CO_2 fluxes consistent with independent in situ measurements of atmospheric CO_2 and to reduce posterior uncertainties in the inferred net and gross CO_2 fluxes (Kaminski et al., 2016).

Lines 325-332: Other examples of the impact of soil moisture (and LAI and FAPAR) data assimilation on LAI and C fluxes include the work at CNRM with the ISBA-A-gs model, e.g. Barbu et al. (2014).

Included the Barbu et al. (2014) reference:
Remotely sensed atmospheric CO_2 concentration (XCO_2, see Section 3.3.1) has also been assimilated into a diagnostic terrestrial carbon cycle model to derive net CO_2 fluxes consistent with independent in situ measurements of atmospheric CO_2 as well as to reduce posterior uncertainties in the inferred net and gross CO_2 fluxes (Kaminski et al., 2016). Barbu et al. (2014) and Albergel et al. (2017) assimilated both soil moisture and LAI data into a land surface model, but their focus was on improving the hydrological and land surface physical quantities and not the carbon cycle.

Line 334: several studies have demonstrated the added benefit of aboveground biomass, including articles already cited (Richardson et al., 2010; Williams et al., 2005; Keenan et al., 2012). Might be worth listing a few examples, or, combining this section with aforementioned examples of C cycle related DA studies (section 2.3).

We have added some references:
So far, remotely sensed biomass data have not been used in carbon cycle data assimilation studies, but several studies (e.g. Richardson et al., 2010; Keenan et al., 2012; Thum et al., 2017) have demonstrated the added value of in situ above-ground biomass observations in constraining the terrestrial carbon cycle.

Line 340: LAI has been used in C cycle DA (see Barbu et al., 2014). Further to my comment on VIs above, perhaps it would be worth explaining somewhere in the text the differences between using VIs, FAPAR and LAI, why one would use one vs another?

See answer above to comment Line 316-318.

Line 341: Worth mentioning the dataset of Li et al. (2011) that has been used in several studies investigating trends in biomass. In fact, I expect that VOD data will be increasingly widely used for optimizing biomass in terrestrial biosphere models, and therefore I would suggest adding a discussion of what these data actually represent in Section 3.3.5 (i.e. how reliably can you estimate biomass (leaf or total aboveground?) from what is essentially a measure of water content).

See answer below to comment Line 675.

Line 343: I think that LST might be used in a similar manner to soil moisture in DA in the future, and not just as an input/boundary condition. Therefore perhaps it can be included with VOD in this context?

Changed as suggested:
However, these products are rather used as input or boundary conditions for terrestrial carbon cycle models (burned area and land cover) or, in the case of land surface temperature and vegetation optical depth, they have so far not been used in carbon cycle data assimilation studies.

Section 3.3.2 Lines 477: you mean significant difference in the absolute magnitude between the products (as the temporal and spatial patterns are quite consistent, as you state)? This was also a conclusion drawn by D'Odorico et al. (2014) and, to some extent, Tao et al. (2015); therefore, it is worth mentioning that these studies agree on this point. As mentioned above, here or elsewhere I think it would be beneficial to have a discussion of the use of VIs and LAI. Arguably LAI is the variable that is most closely linked to standard terrestrial model state variables, therefore the reader should understand why one might use any of these three options for optimizing vegetation dynamics/activity, and the advantages/disadvantages of each. For example, if a modeler is mostly concerned with optimizing the overall magnitude of vegetation activity, careful choice of which FAPAR product to use is important (likely the same for LAI). If they are more concerned with temporal dynamics, one could argue that using a normalized lower order product (e.g. NDVI) that does not require such a sophisticated retrieval algorithm might be more appropriate. Perhaps you do not agree! But in any case, a discussion would be useful here. At the end of this section there is a particular focus on the JRC-TIP FAPAR product as opposed to one of the others, MODIS for example. It would be good to explain the reason for this choice, or to see more information on some of the other commonly available products.

We have emphasised that the difference between the products lies mainly in the absolute magnitude and added that D'Odorico et al. (2014) and Tao et al. (2015) came to the same conclusion:
Pickett-Heaps et al. (2014) concluded that although all six evaluated products display robust spatial and temporal patterns there is considerable disagreement in the absolute magnitude amongst the products and none of the products outperforms the others. This has also been confirmed by the studies of Dodorico et al. (2014) and Tao et al. (2015).

We extended the discussion on the difference between Vis and FAPAR by adding LAI into it and mentioning the shortcomings of VIs and LAI compared to FAPAR:
These indices generally exhibit some improvement in one respect but at the expense of degradation in another respect. Pinty et al. (2009) demonstrate the limitations of such VIs in representing the complex radiative properties of the canopy-soil system over the visible to NIR albedo range. Satellite-derived LAI products (e.g. Liu et al., 2014) seem to be an alternative to VIs. LAI is, however, model-dependent, i.e. the correct interpretation of this variable depends on the formulation of the model used in the retrieval scheme, and may differ from the interpretation adopted by the land biosphere model used for assimilating the LAI product (Disney et al., 2016).
A rational approach to addressing all these issues together is to design a physically-based quantity which is determined by the state of the canopy-soil system.

The reason why there is a slight focus on the JRC-TIP product is mentioned in this section 3.3.2. We made this clearer now in the manuscript at two places:
The JRC-TIP (Pinty et a., 2007) is an inverse modelling system that was explicitly designed to retrieve a set of land surface variables, including FAPAR, in a form that is compliant with the requirements for assimilation into terrestrial biosphere models, hence we focus in the following on this product.

TIP uses observed broadband albedo in the NIR and visible spectral domains as

input. The prior information used in the retrieval is constant in space and time, i.e. all variability is determined from space (Kaminski et al., 2017). This is in contrast to other retrieval approaches, which are based on prescribed land cover maps (e.g. Liu et al., 2014).

Lines 484-485: Please could you be clearer how the products in this sentence link to Table 2? JRC MGVI is not described in Table 2 for example.

The JRC MGVI product is included as a footnote in Table 2 because it uses the same algorithm as used for the SeaWiFS product. We clarified this in the footnote and explicitly mention now JRC MGVI.

Section 3.3.4 Lines 489-491: Although I appreciate you do not wish to provide an exhaustive description of retrieval algorithms, I think it would be helpful to qualitatively describe the difference between passive and active retrieval algorithms in one or two sentences here, as well as the fact different algorithms may produce either volumetric water content (absolute values) vs relative soil moisture values. I would be interested to see a discussion of GRACE land water content in this section.

We included a short description on the retrieval of soil moisture from active instruments and added a short sentence on GRACE. We did not include a discussion of GRACE land water measurements here because they reflect the amount of ground water. This is different to the plant available soil moisture used in terrestrial ecosystem models and relevant for simulation of the terrestrial carbon cycle:
Both passive radiometer systems, measuring the emitted microwave radiance ('brightness 590 temperatures'), and active radar systems, measuring backscattered microwave radiance, can be used to retrieve soil moisture. Various approaches exist that convert brightness temperatures and backscatter measurements into estimates of soil moisture, including radiative transfer model inversion approaches (e.g. Kerr et al. 2012, Owe et al. 2008), neural networks (e.g., Rodríguez Fernández et al. 2015), linear regressions (e.g., Al-Yaari et al. 2016), and change detection methods (Wagner et al., 1999). The latter is commonly applied to scatterometer measurements and yields, in contrast to the other approaches which provide soil moisture as volumetric water content, soil moisture as a percentage of total saturation.

Only Synthetic Aperture Radar is able to provide much higher spatial resolutions, up to a few tens of meters, yet at the cost of long revisit times. Also observations made by the Gravity Recovery and Climate Experiment (GRACE; Rodell et al. 2009) are sensitive to soil moisture, but the estimation of soil moisture content from these observations is not straightforward because they are also sensitive to changes in snow, surface water, groundwater, and vegetation.

Section 3.3.5 Line 670: do not need to reiterate what an active sensor is here.

Removed the half sentence on what an active sensor is.

Line 675: I see you do refer to the VOD product of Liu et al. here. Still, I think it would be beneficial to detail that this is based on VOD data and describe briefly how VOD are derived (following on from the mention of VOD in the soil moisture section) and how biomass is estimated from VOD and their expected use/value for optimizing biomass (as discussed above), as well as for better understanding discrepancies in other sources of biomass data that you discuss towards the end of Section 3.3.5.

We have added that the Liu et al (2015) AGB product is based on VOD. We did not include a discussion on how VOD is derived; that would be outside our scope.
Furthermore, the emphasis is on the AGB of forests, although a global data set of AGB in all biomes for the period 1993-2012 has been produced based on VOD data from global passive microwave sensors, hence with spatial resolution of 10 km or coarser (Liu et al., 2015). The AGB product is derived from a regression of VOD against observations of AGB from ground-based inventory data.

Lines 676-684: updated reference: Santoro et al. (2015) – update to aforementioned papers providing biomass estimates across a wider range of biomes in the northern hemisphere.

Updated the reference.
Santoro et al. (2015) provide a high resolution dataset (0.01°) over the northern hemisphere with a relative RMSE against National Forest Inventory between 12% and 45%.

Line 711: Could you provide the biomass limit that the P-band BIOMASS mission will be able to resolve (to compare with the NISAR mission)?

We added the following text to provide a biomass limit from P-band:
The ESA BIOMASS mission (European Space Agency, 2012), to be launched in 2021, is a P-band radar that will provide near global measurements of forest biomass and height. Measurements from airborne sensors indicate that even in dense tropical forests affected by topography, the P-band frequency used by BIOMASS will give sensitivity to biomass up to 350-450 t/ha (Minh et al., 2014; Villard and Le Toan, 2015).

Given you mention the international soil moisture network in Section 3.3.4, it may be worth mentioning the international tree ring data bank in this section (https://data.noaa.gov/dataset/international-tree-ring-data-bank-itrdb), as these data represent a promising new direction for optimizing biomass across a range of biomes.

Included as suggested, we added the following sentence at the end of Section 3.3.5:
As well as limitations caused by mission lifetimes, satellite measurements of biomass are unlikely to be sensitive enough to measure biomass increment except in rapidly growing plantations and tropical forests. Hence an important ancillary dataset for studies aiming to relate biomass to climate and environment is tree ring data (https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/tree-ring).

Minor comments and typos Line 252: maybe "between" better than "among"? Line 309-310: sentence could be simplified Line 111: benchmark Line 135: measurement Line 145:

knowledge Line 234 and 241: related Line 255: diagonal Line 311: terrestrial Line 315: biogeochemical Line 420: that than Line 440: reflectance Line 576: complementarily Line 727: satellite Line 1395: Updated Thum et al. (2016) reference – see below.

All corrected.

Comments by referee 2, Thomas Kaminski:

1. Model-data fusion and data assimilation: L 37 states that both terms mean the same. Is this true? If yes, I suggest to keep one of the two for the rest of the manuscript instead of switching between the two. If not – maybe because by model-data fusion we could also understand some blending of observations with pre-computed model output – then be more precise in the definitions here and below use the appropriate term depending on context.

We clarified the terminology by adding the following sentence and used 'data assimilation' throughout the manuscript.

The term model-data fusion is sometimes understood in a more general way, where observational data is blended with (pre-computed) model output, whereas the term 'data assimilation' refers to a robust mathematical framework for improving model predictions with observational data.

2. L 47: "new observation": maybe rather "new data stream" or "new type of observation"

Changed as suggested.

3. L64: "In contrast to Ciais et al. (2014), who focus on carbon-cycle observations, we focus here on any kind of relevant observational data to be (potentially) assimilated in a terrestrial carbon cycle data assimilation system (CCDAS). In a CCDAS the observations are used to constrain the underlying model (i.e. to move model output quantities closer to the observations and reduce their posterior uncertainties) usually by parameter optimisation." This formulation could be improved. "any kind of" could be dropped in the first sentence, and the second sentence could read (for example): In a CCDAS non-carbon observations can be exploited to constrain the simulated carbon cycle indirectly through the relations implemented in the process model. Such observational constraints act by ruling out combinations of the unknowns in a CCDAS (typically a combination of process parameters, initial- or boundary conditions) that are inconsistent with the observations and thereby reduce uncertainties in the simulated output."

Changed as suggested.

4. L73: "Our focus lies on the terrestrial carbon cycle, because of the higher spatial and temporal variability in the net exchange fluxes and their associated higher uncertainties than form the ocean and anthropogenic components." Maybe not true on all relevant scales. Maybe just drop the sentence, no need to justify the terrestrial focus in this context.

Changed as suggested.

5. L101: in fact the weighting is in inverse proportion of the uncertainty, also appears below where Eq 1 is described

Corrected.

6. L103: either is maybe not appropriate?

Indeed, removed 'either'.

7. L118: Maybe you want to put: "Here, we follow the notation as introduced by Rayner et al. (2016)" at the beginning of the subsection, i.e. before you start using their notation.

We included the reference to Rayner et al. (2016) in the first sentence of the subsection:
The general problem of data assimilation can be formulated (following the notation of Rayner et al. [2016]) as follows:…

8. As we are dealing with assimilation of "multiple data streams" you could mention that usually each data steam requires its own observation operators, and in fact already here refer to Kaminski and Mathieu (2016/7), maybe even their octopus Figure. And for Eq 1 you could say that, for convenience of notation, now you combine all of them into a single H().

Added two sentences to clarify this point. The first one in the second paragraph of this section:
A data assimilation system consists of three main ingredients: a set of observations, a dynamical model including the observation operator, and an assimilation method. When assimilating multiple data streams each data stream usually requires its own observation operator (see e.g. Kaminski and Mathieu, 2016).

And a second one after Equation (1):
When multiple data streams with different observation operators are assimilated there will be several summands of the form of the second term on the right hand side of Equation (1), one for each data stream.

9. L122: Maybe drop this sentence. In fact the state are mixing ratios.

Changed as suggested.

10. L133: "thus evolves"?

Changed the corresponding sentence to:
We distinguish two basic approaches in data assimilation: sequential assimilation, which assimilates observations subsequently at discrete model time-steps, and variational assimilation,…

11. L136: "optimality" maybe you can find a better word? Maybe "adequacy"?

Indeed, changed as suggested.

12. Figure 1 is confusing in several respects (prior info enters cost, inner/outer loop to be confused with NWP terminology, U(o) not necessarily only a model output, cost function at minimum does not imply availability of A, etc ...). Maybe you just want to drop it with the two sentences that describe it?

We keep Figure 1 because referee #1 found it useful, but we updated the figure and changed the wording (inner/outer loop) in the text and caption to not be confused with NWP.

13. L113-150 starting with "From Equation ..." could also be clearer, shortened or dropped (It does not follow from Eq. 1 that uncertainties are to be taken into account, but Eq. 1 follows from combining PDF descriptions of prior, observations, and model with a few simplifications, Mean and variance are not sufficient to characterise a multi-variate Gaussian, ...) Same holds for next paragraph ("assimilation problem is Gaussian" does not make sense; division by the matrix B is not straightforward...) Maybe just explain variables in Eq. 1 and then directly move to the paragraph starting with Rayner et al. (2016).

Dropped the whole paragraph as suggested.

14. L161: The "and" between citations is missing ("citep" would have worked for multiple citations), same problem occurs a few times further down below.

Changed.

15. L205: what about non-linear observation operators?

Included also 'observation operators' in the text:
…; in the case of non-linear models or non-linear observation operators only the simultaneous assimilation…

16. Section 3.1: It is good to introduce the different forms of errors. It would also be instructive to provide definitions of precision and accuracy.

We think a definition is not needed here (and can be looked up in a dictionary if needed), it is more important that the terms are related to random and systematic errors, which we have done at the end of the respective bullet points.

17. L242: Is is worth mentioning that the scale at which we trust the model may be larger than a grid cell?

Changed as suggested:
For instance a quantity simulated by a model is 'representative' for a given spatial and temporal resolution of the model grid. In fact, the scale at which we trust a model may be larger than a grid-cell.

L247: "In the case of satellite-based observations the representation error also includes errors in inferring a biophysical quantity from the photons measured at the sensor. We come back to this issue later." I would think that such errors in the retrieval rather go into the above two categories?

Indeed this may be misleading and we have removed this sentence here.

18. L 255: "they affect the prediction of the optimal solution in the same way as" could maybe be replaced by "they have considerable impact on the solution. This is because of their influence on the weight of the respective observations in the cost function." It is very good you stress this point. In fact you should take it up in the presentation of each data stream. So far it is only addressed in the XCO2 and the biomass sections.

Changed the sentence as suggested. We did not mention this for each data stream explicitly again because here it is mentioned in general and not related to a particular data stream.

19. L 263: What is inhomogeneous variance?

Here, it means that the variance of each of the superposed Gaussian distributions is not the same.

20. Section 3.2: Maybe add reference/web page of ICOS? Is is worth mentioning similar programmes outside Europe? "The measurements are designed": maybe better "the network" or "the observing system"?

Added URL for ICOS web page and changed as suggested.

L282: Paragraph may fit better into the beginning of section 3.3. Where you discuss the relevant observations provided by the sentinels, you are using our current perspecitve, i.e. S1-5. You could make this clear, because in a few years time are reader could wonder why you don't mention observations by S6 ... etc...

This section has been removed and parts of it are now in the Conclusions sections, see also answer to first comment from referee #1.

21. L 305 and 310: On L305 you write EO, then Earth Observation, then EO... something to be checked throughout ...

After introducing EO in the beginning of the manuscript we now consistently use EO throughout the manuscript.

22. L344: For example Luke (2011) assimilates LAI.
We have changed this section (see answers to referee #1) and included the reference to Luke (2011).

23. Section 3.3.1: Is it worth to briefly explain how a total column value can be sensitive or insensitive to surface fluxes?

The following has been added:
In the following we focus the discussion on sensors that have already delivered multi-year XCO_2 and XCH_4 data sets, i.e. SCIAMACHY and TANSO.
These satellite-derived XCO_2 and XCH_4 data products are sensitive to surface fluxes because CO_2 and CH_4 emission and uptake by surface sources and sinks results in the largest changes of the atmospheric CO_2 and CH_4 mixing ratio close to the Earth's surface and therefore modifies the observed vertical columns. This results in local or regional atmospheric enhancements (e.g., Buchwitz et al., 2017, discussing localized methane sources) or large-scale atmospheric gradients (e.g., Reuter et al., 2014, discussing CO_2 uptake by the terrestrial biosphere).

L400: ")" should be "("

Corrected.

L413: You could mention how the aggregation of errors to the 5 degree grid was performed.

The following has been added:
Each 5°×5° monthly grid cell also contains an estimate of the overall uncertainty (also shown in Fig. (3)) which has been computed taking into account random and systematic error components. The grid-cell uncertainty is computed from two terms: (i) using the reported uncertainties as given in the Level 2 (individual ground pixel) product files for each of the used satellite products (using an ensemble of SCIAMACHY and GOSAT Level 2 products) and (ii) using a term accounting for potential regional/temporal biases as obtained from validation using TCCON ground-based data (see above). The first term depends on the number of individual observations added (the error reduces in proportion to the square root of the number of observations added) whereas the latter term is constant and in the range 0.57 – 0.87 ppm depending on satellite XCO_2 product or in the range 6 – 10 ppb for XCH_4.

L435: Maybe update reference to latest version of the CCI CAR.

Done.

L439: Is is worth mentioning planned XCO2 missions?

Note that additional missions, not discussed in detail in our manuscript, are already mentioned. Therefore, we have added here planned missions only shortly by adding the following:
…China's TanSat (launched end of 2016), which will deliver XCO2 with similar characteristics to NASA's OCO-2. It can be expected that future satellites will provide improved measurements, in particular with respect to more localized emission sources (e.g., Bovensmann et al., 2010; Buchwitz et al., 2013; Ciais et al., 2015).

24. Section 3.3.2: L445: remove one ")".

Corrected.

L458: "closely follows the state of the vegetation" could be "is determined by the state of the canopy-soil system"

Changed as suggested.

L475: Disney et al. (2016) also compare two products.

Included the Disney et al (2016) reference here:
McCallum et al. (2010) looked at four FAPAR data sets over Northern Eurasia for the year 2000, Pickett-Heaps et al. (2014) evaluated six products across Australia, D'Odorico et al. (2014) compared three products over Europe, Tao et al. (2015) assessed five products over different land cover types, and Disney et al. (2016) compared two FAPAR products derived from GlobAlbedo and MODIS data.

L 480: To simplify the sentence maybe move the part in brackets up to the definition of L460.

Done as suggested, the text with the definition of FAPAR now reads:
The Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), which is a normalised fraction with values ranging from 0 to 1, provides information on the photosynthetic activity of the land vegetation.

L503: "see 2" should be "see table 2".

Corrected.

Regarding correlation of uncertainty you might add on L506: after "periods.": To reduce disk space, by default, JRC-TIP products are delivered without correlations among the uncertainties between individual variables, even though these correlations are available. An estimate of uncertainty correlation in space and time is not provided. The JRC-TIP products derived from MODIS (collection 5) broadband albedos minimise temporal uncertainty correlation as each collection 5 albedo value is derived as integral exclusively of observations over non-overlapping 16-day periods.

Changed as suggested, except for the last sentence (which seems to be too specific).

25. Section 3.3.3: L 514: "directly related" In the context of data assimilation, is it worth mentioning that there are complex processes which require complex models as observation operators for SIF, in order to extract the maximal benefit from this data steam?

Included a sentence on this at the end of the paragraph:
But in the context of DA and in order to extract the maximal benefit from SIF data, the complex processes responsible for SIF in the plants' photochemical systems (as mentioned above) require complex models as observation operators for SIF.

L 524: "lies" could be "relies"?

Changed as suggested.

L 530: Why is the simplicity of the forward model related to the fact that least squares is applied, which might also work with complex models?

Changed to:
The retrieval forward model is thus simple and can be linearised (e.g. Guanter et al., 2012; Köhler et al., 2015}, which simplifies the inversion.

L 540: Does "a compromise" make sense here? Isn't it rather the definition of the grid size that is determined by a compromise and the number of retrievals then just a function of this choice of grid size (plus the other factors mentioned)? When discussing the spatial and temporal sampling, it might be instructive to mention the variability in time as you do it in space.

Typically, the smallest spatial or temporal grid size is defined as a function of the application and region of interest (if not global studies are considered), which then defines the other dimension (temporal and spatial), and the number of retrievals to be considered as the reviewer is pointed out. A short clarification has been added to the text.
The number of retrievals to be aggregated into a given grid-cell results from a compromise between spatial resolution, temporal resolution and precision of the gridded product, the size of the spatial and temporal bins being exchangeable in terms of their effect on the random uncertainty.

26. Section 3.3.4: L600: "cost of the": maybe better "cost of long"

Changed as suggested.

L652: ",5."

Corrected, should read 'see Figure 5'.

L658: "to improve the model's hydrology" in fact in a CCDAS we are after the indirect constraint on carbon, so this restriction may not be needed here?

Indeed, we have slightly changed the text:
…that allows for a systematic assimilation into land surface models. These products have been used to improve model hydrology by, for example, Martens et al. (2016) who showed that…

27. Conclusions: L724: "observational characteristics of the observational data" maybe you meant error or uncertainty characteristics?

Yes, corrected.

L730: "correlations": "uncertainty correlations" or "error correlations"

Corrected to ‘error correlations’.

L732: “For example, while FAPAR data constrain mainly the phenology component of a terrestrial carbon cycle model, soil moisture data, in contrast, constrain the hydrological component,” see above regarding the indirect constraints. You probably do not need to write this. In fact FAPAR can provide an important constraint on hydrology (see Kaminski et al., 2012)

Removed this sentence as suggested, this also increases the readability of this section.

28. Fig 3: I’d suggest to go for a 6 panel figure, the four maps are tiny; Better use degree symbol in capton.

We have improved the figure along the lines suggested and for the revised version of the manuscript we use the degree symbol in the caption.

29. Table 1: I suggest to replace “parameter” by “variable”

Changed as suggested.

30. Table 4: You could add wave lengths to the bands, for many colleagues the band names don’t mean anything.

We do not think that adding the wave lengths provides important additional information in the context of the paper here, so we kept the table as is.

There are quite a number of typos. Many of them (e.g. “Reflectamce-based” or “assessemt” or “observeing”) can be detected by a spell checker.

All corrected.

Manuscript prepared for Biogeosciences
with version 2015/04/24 7.83 Copernicus papers of the LaTeX class copernicus.cls.
Date: 19 June 2017

Reviews and syntheses: Systematic Earth observations for use in terrestrial carbon cycle data assimilation systems

Marko Scholze[1], Michael Buchwitz[2], Wouter Dorigo[3], Luis Guanter[4], and Shaun Quegan[5]

[1]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden
[2]Institute of Environmental Physics (IUP), University of Bremen, Bremen, Germany
[3]Department of Geodesy and Geoinformation, Vienna University of Technology (TU Wien), Vienna, Austria
[4]Remote Sensing Section, German Research Center for Geosciences (GFZ), 14473 Potsdam, Germany
[5]Centre for Terrestrial Carbon Dynamics, The University of Sheffield, Sheffield S3 7RH, U.K.

Correspondence to: M. Scholze (marko.scholze@nateko.lu.se)

Abstract.

The global carbon cycle is an important component of the Earth system and it interacts with the ~~hydrological~~hydrology, energy and nutrient cycles as well as ecosystem dynamics. A better understanding of the global carbon cycle is required for improved projections of climate change including corresponding changes in water and food resources and for the verification of measures to reduce anthropogenic greenhouse gas emissions. An improved understanding of the carbon cycle can be achieved by ~~model-data fusion or~~ data assimilation systems, which integrate observations relevant to the carbon cycle into coupled carbon, water, energy and nutrient models. Hence, the ingredients for such systems are a carbon cycle model, an algorithm for the assimilation, and systematic and well ~~error-characterized~~ error-characterised observations relevant to the carbon cycle. Relevant observations for assimilation include various ~~in-situ~~ in situ measurements in the atmosphere (e.g. concentrations of CO_2 and other gases) and on land (e.g. fluxes of carbon water and energy, carbon stocks) as well as remote sensing observations (e.g. atmospheric composition, vegetation and surface properties).

We briefly review the different existing data assimilation techniques and contrast them to model benchmarking and evaluation efforts (which also rely on observations). A common requirement for all assimilation techniques is a full description of the observational data properties. Uncertainty estimates of the observations are as important as the observations themselves because they similarly determine the outcome of such assimilation systems. Hence, this article reviews the requirements of data assimilation systems on observations and provides a non-exhaustive overview of current observations and their uncertainties for use in terrestrial carbon cycle data assimilation. We report on

progress since the review of model-data synthesis in terrestrial carbon observations by Raupach et al. (2005) emphasising the rapid advance in relevant space-based observations.

1 Introduction

The anthropogenic ~~pertubation~~ perturbation of the global carbon cycle has led to a global mean increase of 43% in atmospheric CO_2 (from 280 ppm to 398 ppm) in 2014 compared to pre-industrial (before 1750) levels (WMO, 2015), and is the main driver for climate change. The main causes for the increase in CO_2 are burning of fossil fuels and land use change, which amount to emissions of 9.8 ± 0.5 GtC in 2014. However, only about 44% of these emissions stay in the atmosphere, the remainder is currently taken up by the land biosphere ($\approx$ 30%) and the surface ocean ($\approx$ 26%) (Le Quéré et al., 2015). Positive climate-carbon cycle feedbacks, ~~predominatly~~ predominantly acting on land processes, may reduce this sink capacity and thus accelerate global warming (Matthews et al., 2007). Also, the sink strength of the terrestrial biosphere is more variable than that of the ocean (Ciais et al., 2013) and its quantification by process-based terrestrial carbon cycle models exhibit large uncertainties (Le Quéré et al., 2015).

A common way to reduce uncertainties from process-based modelling is by confronting these models with observational data. Raupach et al. (2005) pointed out that the systematic combination of observational data with process modelling, which is commonly ~~refered~~ referred to as 'model-data fusion', is an effective strategy for observing the Earth system. ~~Model-data fusion , or more formally known as data assimilation,~~ The term model-data fusion is sometimes understood in a more general way, where observational data is blended with (pre-computed) model output, whereas the term 'data assimilation' refers to a robust mathematical framework for improving model predictions with observational data. Data assimilation is motivated by several benefits to make best use of observations and models (Mathieu and O'Neill, 2008). These benefits include, among others, (1) forecasting and initialisation (forward predictions in time based on past observations), (2) model and data quality control (regular and systematic confrontation of model output with observations within their uncertainty statistics), (3) combination of various data streams (combined constraints of independent observations can be stronger than the sum of the individual constraints), (4) filling in regions with sparse observations (consistent propagation of information from data-rich regions to data-poor regions), (5) estimating unobservable quantities (through process-based relations in the model observations constrain modelled quantities which are not directly measured) and (6) observing system design (what is the delta of a new ~~observation/instrument~~type of observation).

Systematic observations are a key ingredient for ~~model-data fusion~~ data assimilation studies. Here, we focus on the carbon cycle and the land-atmosphere system. The land-atmosphere components of the carbon cycle are an important part of an integrated Earth observation system because of the close interactions on land between the carbon cycle and the water and energy cycles, and hence its

importance for climate projections and climate change mitigation strategies through the monitoring and management of terrestrial greenhouse gas sources and sinks.

Raupach et al. (2005) provide an analysis of the various elements of a Terrestrial Carbon Observation System (TCOS). The need, design and steps to be taken towards a TCOS were already outlined by others before (Cihlar et al., 2002; Global Carbon Project, 2003) but Raupach et al. (2005) systematically reviewed two major components of a TCOS: the ~~model-data fusion~~ data assimilation methods and the observational data and data uncertainty ~~characterisctics~~ characteristics for some selected, main kinds of relevant data. The requirements for a policy-relevant carbon observing system have been outlined by Ciais et al. (2014). They review the current systematic carbon-cycle observations and illustrate the implementation of such a policy-relevant carbon observing system.

In this paper we provide an update of the observational data and data uncertainty characteristics as assessed by Raupach et al. (2005) with a focus on ~~exisitng~~ existing but also new and upcoming, relevant space-based observations~~(~~, in the following referred to as Earth Observation (EO) data~~)~~. In contrast to Ciais et al. (2014), who focus on carbon-cycle observations, we focus here on ~~any kind of~~ relevant observational data to be (potentially) assimilated in a terrestrial ~~carbon cycle data assimilation system~~ Carbon Cycle Data Assimilation System (CCDAS).

In a CCDAS ~~the observations are used~~ non-carbon observations can be exploited to constrain the ~~underlying model(i. e. to move model output quantities closer to the observations and reduce their posterior uncertainties) usually by parameter optimisation~~simulated carbon cycle indirectly through the relations implemented in the process model. Such observational constraints act by ruling out combinations of the unknowns in a CCDAS (typically a combination of process parameters, initial- or boundary conditions) that are inconsistent with the observations and thereby reduce uncertainties in the simulated output. In that sense we are somewhat broader in terms of observed variables because also the 'non-carbon' observations (such as soil moisture or land surface temperature) are able to constrain the carbon cycle indirectly through process information embedded in the underlying models. At the same time, the focus of our review is narrower than that of Ciais (2014), who also addressed ocean and anthropogenic components. ~~Our focus lies on the terrestrial carbon cycle, because of the higher spatial and temporal variability in the net exchange fluxes and their associated higher uncertainties than form the ocean and anthropogenic components.~~

The paper is ~~organized~~ organised as follows: in the next section we contrast data assimilation to recently established benchmarking activities and give a brief overview of commonly used data assimilation approaches and their applications in terrestrial carbon cycling. We continue with a short overview on data characteristics ~~inlcuding~~ including an update on progress for some of the observations discussed in Raupach et al. (2005). Since there has been much developments in the provision of remotely sensed observations we focus here on the characteristics of the most relevant EO data streams.

2 ~~Model-data fusion~~Data Assimilation

2.1 Data assimilation versus benchmarking

In the recent past the international land surface and terrestrial ecosystem modelling communities have ~~recognized~~ recognised the importance of model benchmarking and evaluation (e.g. Luo et al., 2012; Foley et al., 2013). One of the reasons for this development is the huge range of model results from different models in key diagnostics of the land-atmosphere interface such as gross primary productivity (GPP) and latent heat flux (Prentice et al., 2015).

In general 'benchmarking' is understood as the quantification of performance against a reference using some pre-defined metrics. The reference can either be output from some previous model simulations, other (ensembles of) models or reference ~~datasets~~ data sets based on observations if the model simulates the analogue quantity. Luo et al. (2012) suggest a theoretical framework for benchmarking land models based on based on ~~standardized~~ standardised references and metrics to measure model performance skills. A large variety of such metrics and their characteristics is introduced by Foley et al. (2013). Some examples of benchmarking terrestrial carbon cycle models (either standalone or coupled to climate models) are given by e.g. Randerson et al. (2009), Cadule et al. (2010) and Kelley et al. (2013).

The commonality between benchmarking/evaluation and data assimilation lies in the quantitative assessment of model output. In benchmarking the quantitative assessment is performed by calculating some metrics against either observations or other references, while in data assimilation this is achieved by defining a cost function, which quantifies the mismatch of some simulated model quantity against observations weighted by the inverse of their uncertainties (including a model uncertainty). But data assimilation goes beyond benchmarking as it minimises the quantified mismatch to improve model performance directly by adjusting ~~either~~ initial and boundary conditions, state variables and/or model process parameters.

As pointed out by Prentice et al. (2015) there is a need for both model benchmarking and data assimilation: Benchmarking as a routine application to improve confidence and evaluate the performance (over time) in terrestrial carbon cycle modelling. However, if a benchmark test for a given model fails this could simply imply that the model parameter values have not been specified correctly and ~~optmised~~ optimised against observations. In contrast, data assimilation, in particular when used for parameter optimisation, potentially identifies structural model and/or data deficiencies if the model-data mismatch (or the ~~benchmar~~ benchmark test) is still inadequate after optimisation (see also Figure (1)). On the other hand, a better fit between the posterior maximum likelihood simulation (i.e. using the optimised parameters) and the observations is not necessarily an indication for correct parameters and/or model structure as has been pointed out by MacBean et al. (2016).

2.2 Data assimilation methods

The general problem of ~~model-data fusion, or, more strictly speaking,~~ data assimilation can be formulated ~~like this~~(following the notation of Rayner et al. (2016)) as follows: Given a model M, a set of observations $\mathbf{y}$ of some observables $\mathbf{o} = H(\mathbf{z})$, with $\mathbf{z}$ being the state variables of the model and H the observation operator, and prior information on some target variables $\mathbf{x}$, produce an updated description of $\mathbf{x}$. $\mathbf{x}$ may include elements of $\mathbf{z}$ and $\mathbf{p}$ (parameter, quantities not changed by the model, i.e. process parameters, boundary and initial conditions). ~~Here, we follow the notation as introduced by Rayner et al. (2016).~~ The observation operator maps the model state onto observables. In the case of a CCDAS assimilating atmospheric CO_2 the observation operator is the atmospheric transport model mapping the net CO_2 surface exchange fluxes as calculated by the terrestrial carbon cycle onto simulated atmospheric CO_2 concentrations. ~~In transport inversions the dynamical model, the atmospheric transport model, is also the observation operator.~~

A ~~data-assimilation~~ data assimilation system consists of three main ingredients: a set of observations, a dynamical model including the observation operator, and an assimilation method. When assimilating multiple data streams each data stream usually requires its own observation operator (see e.g. Kaminski and Mathieu, 2016). In the Bayesian formulation of the assimilation problem uncertainties (i.e. the description of quantities by probability density functions, PDFs) are central to the concept of data assimilation. Both observations as well as models have errors arising for various reasons. We will detail the observational errors in the next section. Dynamical models as well as observation operators have errors arising from the parameterizations, and the discretization of analytical dynamics into a numerical model; for a more complete description of uncertainty in Earth System models or components of such we refer to Scholze et al. (2012).

We distinguish two basic approaches in data assimilation: sequential assimilation, which assimilates observations ~~at discrete~~ subsequently at discrete model time-steps~~and thus evolves over time according to the dynamical model; and~~ , and variational assimilation, which assimilates all observations at once at their respective ~~measyurement~~ measurement time over a given period, the so-called assimilation window. They differ in their numerical efficiency and ~~optimality~~ adequacy for their specific use. A general data-assimilation scheme is shown in Figure (1). In the sequential approach the ~~inner~~ assimilation loop is evaluated sequentially over time following the dynamics of the model. In the case of variational assimilation the ~~inner~~ assimilation loop is evaluated iteratively (assuming a non-linear model)~~until~~ . Both cases evaluate a cost function ~~minimum is found. The cost function is formulated as~~ J, formulated in the Bayesian framework as:

$$J = \frac{1}{2}\left[(\mathbf{x}-\mathbf{x}^{\mathrm{b}})^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^{\mathrm{b}}) + (\mathrm{H}(\mathbf{x})-\mathbf{y})^T\mathbf{R}^{-1}(\mathrm{H}(\mathbf{x})-\mathbf{y})\right], \quad (1)$$

where $\mathbf{x}^{\mathrm{b}}$ is the prior information, $\mathbf{B}$ the prior uncertainty covariance, and $\mathbf{R}$ the observational uncertainty covariance. ~~From Equation 1 follows that data and prior knowledge cannot be treated separately from their respective uncertainties (Raupach et al., 2005). In other words, observations~~

~~(or prior knowlegde) for data assimilation are only complete if we know the full probability density function (PDF), which, in the case of a Gaussian, can be characterised by its mean and variance. In practical terms, the observational uncertainty covariances weight the model-data mismatch, while the prior uncertainty covariances weight the deviation of the target variables from their prior values. We note here that in the Gaussian case the model and observation operator errors can be added quadratically to the observation errors.~~

~~An important diagnostic in data assimilation is the posterior uncertainty, which usually, because of its high dimension, is hard to compute. If the assimilation problem is Gaussian the computation of the posterior uncertainty covariance matrix simplifies and it can be approximated by the inverse of the Hessian ($2_n d$ derivative)of the cost function. Typically, gradient-based optimisation approaches approximate the Hessian, alternatively ensembles can be used to derive realisations of the posterior PDF. The uncertainty reduction relative to the prior (i.e. $1 - \mathbf{U}_{\mathbf{xpo}}/\mathbf{B}$ with $\mathbf{U}_{\mathbf{xpo}}$ the posterior uncertainty) then is a measure of the observational constraint on the target variables.~~

When multiple data streams with different observation operators are assimilated there will be several summands of the form of the second term on the right hand side of Equation (1), one for each data stream. Rayner et al. (2016) introduce the theory fundamental to data assimilation and illustrate how the different implementations of data assimilation relate to this theory in a more narrative style A more complete and mathematically precise introduction to the concepts of data assimilation is given in the textbooks by e.g. ~~Daley (1991); Tarantola (2005)~~Daley (1991) and Tarantola (2005).

2.3 Examples of terrestrial carbon cycle data assimilation

A variety of the methods as described by Rayner et al. (2016) have been applied by the carbon cycle community. One example making use of formal assimilation ~~methdologies~~ methodologies for inferring surface-atmosphere CO_2 exchange fluxes is based on atmospheric transport inversions. As mentioned before, in atmospheric inversions the observation model is an atmospheric tracer transport model. In atmospheric inversions both sequential and variational methods have been used together with observations of atmospheric trace gas concentrations such as from the flask sampling network, continuous ~~in-situ~~ in situ and aircraft measurements and more recently also remotely sensed total column measurements. The techniques for atmospheric transport inversions have been detailed in Enting (2002) and a recent comparison of results from different transport inversion is given by Peylin et al. (2013).

A more recent development is the assimilation of observations into terrestrial biosphere models. Here, various methods and observations have been used to optimise model process parameters at different scales. A comparison of a whole suite of these assimilation methods applied to a test case using a simplified model at local-scale is given by Trudinger et al. (2007) and Fox et al. (2009).

Kaminski et al. (2002) were among the first who applied a formal algorithm together with observations of atmospheric CO_2 concentrations to constrain the Simple Diagnostic Biosphere Model at

global scale. This work was continued by the development of the first Carbon Cycle Data Assimilation System (CCDAS) with a process-based model (BETHY) at its core (Rayner et al., 2005). The advantage of using a process-based model at the core of a CCDAS is that once the process parameters have been optimised the the constrained model can also be used for predictions as demonstrated by Scholze et al. (2007). Also, such systems are capable of ingesting multiple independent data streams besides atmospheric CO_2 concentrations. Kaminski et al. (2013) provide an overview on the developments of the CCDAS-BETHY since its first application while Scholze et al. (2016) demonstrate the latest application of CCDAS-BETHY assimilating atmospheric CO_2 and remotely sensed surface soil moisture simultaneously. Since then several global terrestrial ecosystem models have been included in ~~Carbon Cycle Data Assimilation Systems~~ CCDAS employing a variational approach (e.g. Schürmann et al., 2016; Peylin et al., 2016).

Concurrently, there have been several studies at the local/regional scale assimilating various types of observations. For instance, Barrett (2002) used a genetic algorithm to infer soil carbon turnover times in a terrestrial carbon cycle model over Australia from in situ observations of plant production, biomass, litter and soil carbon~~observations~~. Local eddy covariance flux tower measurements of net exchange of CO_2 and latent and sensible heat fluxes have been assimilated to ~~optimize~~ optimise parameter related to photosynthesis, respiration and energy fluxes of terrestrial ecosystem models, using Monte Carlo type methods ~~(e.g. Braswell et al., 2005; Knorr and Kattge, 2005; Moore et al., 2008; Ricciuto et al., 2008)~~ (e.g. Braswell et al., 2005; Knorr and Kattge, 2005; Moore et al., 2008; Ricciuto et al., 2008; Post et al., 2017), sequential methods (Williams et al., 2005), as well as variational approaches (e.g. Wang et al., 2001; Kuppel et al., 2012; Raoult et al., 2016)

Recent advances ~~are focusing~~ focus on multiple independent data stream assimilation to provide a more rigorous constraint on the multiple components of terrestrial ecosystem models and avoid equifinality, i.e. different parameter solutions ~~provide~~ providing the same cost function value at the minimum. Examples for such studies on local/regional scale are the assimilation of eddy covariance CO_2 fluxes together with observations of vegetation structural information or carbon stocks ~~(e.g. Richardson et al., 2010; Keenan et al., 2012) or together with remotely sensed vegetation activity such as the Fraction of Absorbed Photosynthetic Active Radiation (FAPAR) (e.g. Kato et al., 2013; Bacour et al., 2015)~~ (e.g. Richardson et al., 2010; Keenan et al., 2012; Thum et al., 2017). The assimilation of multiple data streams can be performed either in a step-wise (e.g. Peylin et al., 2016) or simultaneous approach (e.g. Kaminski et al., 2012); in the case of non-linear models or non-linear observation operators only the simultaneous assimilation makes optimal use of the observations (MacBean et al., 2016). In Section 3.2 we provide more terrestrial carbon cycle data assimilation examples using some of the remotely sensed products discussed in the following.

3 Data characteristics and provision

Observations are our measurable representation of the 'Truth'. They come with different characteristics in terms of spatial and temporal resolution, coverage of the observed system, and errors. In analogy, models are also some representation of the 'Truth', however, via knowledge embodied in some form of functional relationships (with their own errors as mentioned before). The paper by Raupach et al. (2005) has been instrumental in highlighting the challenges in combining models and observational data for building a TCOS ~~focussing~~ focusing on the observational requirements. Ciais et al. (2014) argue for a globally integrated carbon observation system to improve our understanding of the carbon cycle for predicting future changes and to be able to independently verify the impact of emission reduction measures. Such a system relies on atmospheric carbon observations as a backbone but also concerns observations of the terrestrial and ocean carbon cycle. They focus on a strategy towards a global carbon-cycle monitoring system for achieving the above mentioned objectives.

Figure (2) depicts exemplarily the main observations of a TCOS and their space-time characteristics. In the following we briefly summarise the aspects of uncertainty in the observations and highlight progress on the specification of uncertainty for some of the observations in Fig. (2) as well as on their monitoring since Raupach et al. (2005).

3.1 Observational uncertainty

As mentioned before an important ingredient to any ~~model-data fusion~~ data assimilation system are not only the observations ~~themsleves~~ themselves but also the uncertainties associated to them. We distinguish three main types of observation errors:

- Random: Random errors are always present in measurements and are caused by unpredictable changes in the measurement system (e.g. electronic noise in electrical instrument). They show up as different readings of the same repeated measurement, and thus can be reduced by taking the average of multiple measurements. Random ~~erros~~ errors are usually assumed to be Normal (Gaussian) distributed, however, in some cases the random error distribution is log-normal (e.g. precipitation) or skewed by outliers due to unpredictable corruptions of the measurement system. Random ~~erros are therefore relatad~~ errors are therefore related to the precision of a measurement system.

- Systematic (bias): Systematic errors in observations are usually due to some recurring problems in the overall measurement system. They are caused by instrument miscalibrations or interferences with the measurement system. They can vary in space and time but they affect the measurement system in a predictable way. Biases can be both additive (absolute mean bias) and multiplicative (biases in the dynamic range affecting the amplitude of a signal). If the source for systematic errors is known they can usually be fixed and ~~shoud~~ should be re-

moved. Systematic ~~erros are therefore relatad~~ errors are therefore related to the accuracy of a measurement system.

- Representativeness. The representation error occurs when information is represented at a scale different from the source of the information. For instance a quantity simulated by a model is 'representative' for a given spatial and temporal resolution of the model grid. In fact, the scale at which we trust the model may be larger than a grid-cell. An individual measurement, however, represents information influenced by the local environment not resolved by the model grid (e.g. representation of atmospheric flask data in an atmospheric transport model grid-cell). ~~In the case of satellite-based observations the representation error also includes errors in inferring a biophysical quantity from the photons measured at the sensor. We come back to this issue later.~~

For both random and systematic errors not only the magnitude of the error for a single observation is important, i.e. the diagonal elements in the observational uncertainty covariance matrix $\mathbf{B}$, but also the correlations ~~among~~ between errors for different observations. Hence there is a need to specify the off-diagonal elements in the error covariance matrix $\mathbf{B}$. These off-diagonal elements are usually hard to specify, ~~however, they are~~ but it is important to quantify them in a data assimilation system~~because they affect the prediction of the optimal solution in the same way as the diagonall elements.~~ . They have considerable impact on the solution because of their influence on the weight of the respective observations in the cost function.

In addition to the observational errors, models also have errors, which, in a data assimilation system, are usually included in the observation errors. These errors in dynamical models are mainly caused by process parameterizations (instead of resolving the process), and by the discretization of analytical dynamics into a numerical model. A more detailed description of the different model error sources is given in Scholze et al. (2012).

As mentioned before, Raupach et al. (2005) have already reflected on the main properties of the data and their error covariances for observations of remotely sensed land surface properties (mainly the normalised differential vegetation index, NDVI), atmospheric CO_2 concentrations, land-atmosphere net CO_2 exchange fluxes, and terrestrial carbon stores. The in situ measurements of CO_2 concentrations are either based on flask samples or on continuous monitoring stations. The flask sampling network was established in 1961 by Keeling (1961) and has been extended since then to more than 200 sites globally. The continuous in situ network provide measurements at higher precision and temporal resolution than the flask networks. For both the flask and the continuous stations improvements in precision and in the accuracy have been achieved through propagation of frequent comparisons and international standards (Francey et al., 2001).

The global FluxNet network consists of more than 200 sites globally measuring land-atmosphere fluxes of CO_2, latent and sensible heat and others by the eddy covariance technique at a half-hourly temporal resolution (Baldocchi et al., 2001). Many other (mostly meteorological) variables are measured

at these sites as well. In the past years, there has been substantial progress in the homogenisation and availability of these direct CO_2 flux measurements. The publically available FLUXNET2015 data set includes more than 1500 site-years of data covering all major biome types from about 165 sites worldwide spanning a period from 1991 (for some sites) up to 2014 (Pastorello et al., 2017). There has also been substantial progress in the specification of uncertainties in eddy-covariance measurements of the land-atmosphere net CO_2 exchange flux (Net Ecosystem Productivity, NEP) and its component fluxes (GPP and ecosystem respiration, R_{eco}). For instance, Lasslop et al. (2008) analysed the error distribution and found that the eddy flux data can almost entirely be represented by a superposition of Gaussian distributions with inhomogeneous variance. Richardson et al. (2008) showed that the measurement errors in NEP are heteroscedastic, i.e. the error variance varies with the magnitude of the flux. In a more recent study Raj et al. (2016) investigated the ~~uncertainty~~ uncertainty of GPP derived from partitioning the eddy covariance NEP measurements. They used a light-use efficiency model embedded in a Bayesian framework to estimate the uncertainty in the separated GPP from the posterior distribution at half-hourly time steps. ~~The availability of the eddy covariance data has also been heavily improved; the latest release of the FLUXNET2015 dataset now contains data from about 165 sites worldwide spanning a period from 1991 (for some sites) up to 2014 (FLUXNET2015).~~

3.2 ~~Towards operational carbon observation systems~~

~~In the European framework there have recently been major developments towards systematic in-situ observations for use in terrestrial carbon cycle data assimilation systems. The Integrated Carbon Observing System (ICOS) is a novel pan-European infrastructure for carbon observations, which will provide high-quality in-situ observations (both fluxes as well as atmospheric concentrations) over Europe and over ocean regions adjacent to Europe with a long-term perspective. ICOS consists of central facilities for co-ordination, calibration and data in conjunction with networks of atmospheric, oceanic and ecosystem observations as well as a data distribution centre, the Carbon Portal, providing discovery of and access to ICOS data products such as derived flux information. The ICOS network runs in an operational mode, and greenhouse gas concentrations and fluxes will be determined on a routine basis. The measurements are designed to allow up to daily determination of (mainly natural) sources and sinks at scales down to approximately 50 x 50 km^2 for the European continent.~~

~~An example for an operationalised, space-based Earth observing programme is the fleet of so-called Sentinel satellites of the European Copernicus programme. Copernicus aims at providing Europe with a continuous and independent access to Earth observation data and associated services (transforming the satellite and additional in-situ data into value-added information by processing and analysing the data) in support of Earth System Science (Berger et al., 2012). So far, six different Sentinel missions are planned out of which three are in operation and the remainder is scheduled to be launched during the next years. Each type of the currently foreseen Sentinels has a specific objective and~~

~~will deliver a range of EO products. Some of these products will be suitable for constraining the terrestrial carbon cycle, such as soil moisture (Sentinel 1), FAPAR, leaf chlorophyl and water content and land cover (Sentinel 2 and 3), land surface temperature (Sentinel 3), atmospheric methane and flouresecence (Sentinel 5 and precursor). So far, a dedicated mission for monitoring the carbon cycle, i.e. an instrument measuring the atmospheric CO_2 composition, is not yet included in the Copernicus monitoring programme (see Ciais et al., 2015), however, the series of Sentinel satellites is likely to be extended in the future.~~

3.2 Examples of systematic observations from satellite EO data

There has been a vast extension of EO ~~capabilites~~ capabilities during the past 10 years or so both in terms of product quality (including, for instance, improved accuracy) but also quantity (new products).

In any data assimilation system using satellite EO data one needs to decide in the design phase of the system whether to assimilate observations at the sensor level (i.e. the spectral radiances for optical sensors or brightness temperatures for microwave sensor, referred to as level 1 data) or to assimilate the bio-geophysical variable derived from the radiances through a retrieval algorithm (level 2 data product). When assimilating level 1 data the retrieval algorithm is part of the observation operator linking the model state to the observations in the data assimilation system. A more detailed description of the two alternatives in assimilating EO satellite observations into models of the Earth system is given by Kaminski and Mathieu (2016). In carbon cycle data assimilation systems level 2 data products (or even level 3 data, which are provided on a regular space-time grid) are most commonly used. However, there is the risk that when using level 2 or higher products the parameters/processes implemented in the retrieval algorithm may not be consistent with corresponding equivalent parameters/processes in the underlying model, and thus cause additional errors in the assimilation.

In the next subsections we present some selected ~~, and~~ remotely sensed Earth Observation products, which are relevant for terrestrial carbon cycle data assimilation~~most relevant remotely sensed Earth Observation products~~ , in more detail~~. The EO products described below (~~:

- atmospheric CO_2 ~~, vegetation activity , soil moisture , terrestriaial biomass)~~
- vegetation activity (FAPAR and SIF)
- soil moisture
- terrestrial biomass.

These EO products either have already been used, are in the process of being used, or would potentially be a useful data constraint in a CCDAS. For vegetation activity we distinguish two major types of products: more 'traditional' reflectance- or radiative-based products such as fraction of

absorbed photosynthetically active radiation (FAPAR) and recently developed products based on ~~biogeochamical processes~~biogeochemical processes, such as sun-induced ~~flouresecence~~ fluorescence (SIF). ~~For instance, FAPAR~~ Leaf area index (LAI, e.g. Liu et al., 2014), which is in effect closely related to FAPAR, is another geophysical parameter representing vegetation activity. There is also a range of remotely sensed vegetation indices, of which NDVI is an example. Both LAI and NDVI have been used in data assimilation studies: an example for NDVI is given by MacBean et al. (2015) and for LAI by Luke (2011) and Barbu et al. (2014). In Section 3.2.2 we detail the difference between NDVI and FAPAR, and explain that FAPAR is based on physical principles. FAPAR has already been demonstrated to provide a strong constraint on terrestrial carbon ~~as well as~~ and water fluxes through its impact on the phenology components of the carbon cycle model ~~(e.g. Knorr et al., 2010; Kaminski et al., 2012)~~ either by assimilating only FAPAR data (e.g. Knorr et al., 2010) or in combination with other data streams (e.g. Kaminski et al., 2012; Kato et al., 2013; Forkel et al., 2014). SIF is a promising observation to constrain the gross uptake of CO_2 by plant photosynthesis. First assimilation results using SIF observations in a CCDAS show that the uncertainty in global annual GPP is largely reduced by constraining parameters that describe leaf phenology (Norton et al., 2016). ~~Also assimilation of XCO2~~ Remotely sensed atmospheric CO_2 concentration (XCO_2, see Section 3.2.1) has also been assimilated into a diagnostic ~~terrestial~~ terrestrial carbon cycle model ~~has been shown~~ to derive net CO_2 fluxes consistent with independent ~~in-situ~~ in situ measurements of atmospheric CO_2 ~~as well as~~ and to reduce posterior uncertainties in the ~~infered~~ inferred net and gross CO_2 fluxes (Kaminski et al., 2016). Barbu et al. (2014) and Albergel et al. (2017) assimilated both soil moisture and LAI data into a land surface model, but their focus was on improving the hydrological and land surface physical quantities and not the carbon cycle. van der Molen et al. (2016) assessed the impact of assimilating various remotely sensed soil moisture products into the SiBCASA ecosystem model on simulated carbon fluxes in Boreal Eurasia. Although the impact of assimilating ASCAT surface soil moisture was significant, its skill in this hydrologically complex environment strongly depends on surface water and vegetation dynamics. In contrast, Scholze et al. (2016) showed that when assimilating SMOS soil moisture simultaneously with ~~in-situ~~ in situ atmospheric CO_2 concentrations the reduction of uncertainty in gross and net CO_2 fluxes relative to the prior is considerably higher than for assimilating CO_2 only, which quantifies the added value of SMOS observations as a constraint on the terrestrial carbon cycle. So far, remotely sensed biomass data have not been used in carbon cycle data assimilation studies, ~~however, Thum et al. (2017)~~ but several studies (e.g. Richardson et al., 2010; Keenan et al., 2012; Thum et al., 2017) have demonstrated the added value of ~~in-situ observations of biomass increment in reducing uncertainties in simulated above ground biomass mainly through the calibration of parameters in the carbon allocation scheme of the~~ in situ above-ground biomass observations in constraining the terrestrial carbon cycle~~model~~.

This list of EO products described in this paper is admittedly subjective and there is of course a whole range of additional remotely sensed products available, which are relevant for carbon cycle

studies as well, e.g. burned area (e.g. Giglio et al., 2013), land cover (e.g. Bontemps et al., 2012), land surface temperature (e.g. Li et al., 2013), ~~leaf area index (which is in effect closely related to FAPAR) (e.g. Liu et al., 2014)~~ or vegetation optical depth (VOD) (e.g. Konings et al., 2016). However, these products are rather used as input or boundary conditions for terrestrial carbon cycle models ~~or, for instance~~ (burned area and land cover) or, in the case of ~~vegetation optical depth~~land surface temperature and VOD, they have so far not been used in carbon cycle data assimilation studies.

3.2.1 Atmospheric CO_2 and CH_4

Satellite retrievals of atmospheric carbon dioxide (CO_2) and methane (CH_4) are available from several satellite instruments such as mid-tropospheric CO_2 and CH_4 columns from Infrared Atmospheric Sounding Interferometer (IASI) (e.g. Crevoisier et al., 2009a, b) on EUMETSAT's Metop satellite series, vertical profiles with highest sensitivity in the mid/upper troposphere from AIRS on Aqua (e.g. Xiong et al., 2013), stratospheric profiles from MIPAS on ENVISAT limb observations (e.g. Laeng et al., 2015) and from the solar occultation observations of SCIAMACHY on ENVISAT (Noël et al., 2011, 2016) and ACE-FTS (e.g. Boone et al., 2005; Foucher et al., 2009). These observations have however only little or no sensitivity to CO_2 and CH_4 concentration changes close to the Earth's surface and therefore contain only limited information on regional or local CO_2 and CH_4 sources and sinks. Satellites with high near-surface sensitivity are nadir (~~downlooking~~down-looking) satellites which measure radiance spectra of reflected solar radiation in the relevant spectral bands in the near-infrared/shortwave-infrared (NIR/SWIR) spectral region, which are located around 1.6 μm (CO_2 and CH_4) and around 2.0 μm (CO_2). Satellites instruments which perform (or have performed) these observations are SCIAMACHY onboard ENVISAT (2002–2012) (Burrows et al., 1995; Bovensmann et al., 1999), TANSO-FTS onboard GOSAT (launched in 2009) (Kuze et al., 2009, 2014) and NASA's Orbiting Carbon Observatory 2 (OCO-2) mission (launched in 2014) (Crisp et al., 2004; Boesch et al., 2011).

The main CO_2 and CH_4 data products of these sensors are near-surface-sensitive column-averaged dry-air mole fractions of CO_2 and CH_4, denoted ~~XCO2 and XCH4~~XCO_2 and XCH_4. The quantities ~~XCO2 and XCH4~~ XCO_2 and XCH_4 are both retrieved from SCIAMACHY/ENVISAT (ground pixel size: 30×50 km^2 (along track times across track); swath width 960 km with contiguous ground pixels) and TANSO-FTS/GOSAT (10 km pixel size; several (e.g. 3 or 5) non-contiguous pixels across track). OCO-2 delivers ~~XCO2~~ XCO_2 (8 ground pixels across track, each ≈1.3 km) and ~~in the near future other satellites~~ other satellites have been or will be launched such as Europe's Sentinel-5-Precursor satellite (S5P) (Veefkind et al., 2012), which will deliver (among several other parameters) ~~XCH4~~ XCH_4 (7 km pixel size at nadir, 2600 km swath width with contiguous ground pixels; planned launch: ~~mid~~ autumn 2017) (Butz et al., 2012) and China's TanSat (~~planned launch~~ launched end of 2016), which will deliver ~~XCO2~~ XCO_2 with similar characteristics ~~as~~ to NASA's OCO-2. It can be expected that future satellites will provide improved measurements, in particular with respect to

more localised emission sources (e.g., Bovensmann et al., 2010; Buchwitz et al., 2013; Ciais et al., 2015). In the following we ~~will~~ focus the discussion on sensors ~~who~~ that have already delivered multi-year ~~XCO2 and XCH4 year~~ XCO_2 and XCH_4 data sets, i.e. on SCIAMACHY and TANSO.

~~The XCO2 and XCH4~~ These satellite-derived XCO_2 and XCH_4 data products are sensitive to surface fluxes because CO_2 and CH_4 emission and uptake by surface sources and sinks results in the largest changes of the atmospheric CO_2 and CH_4 mixing ratio close to the Earth's surface and therefore modifies the observed vertical columns. This results in local or regional atmospheric enhancements (e.g. Buchwitz et al. (2017), discussing localised methane sources) or large-scale atmospheric gradients (e.g. Reuter et al. (2014), discussing CO_2 uptake by the terrestrial biosphere).

The XCO_2 and XCH_4 data products retrieved from SCIAMACHY and TANSO are generated from the radiance observations using different approaches. Most approaches are based on 'Optimal Estimation' (OE) (e.g. Rogers, 2000; Reuter et al., 2010), also called Bayesian inference. OE permits to constrain the retrieval using a priori information on, e.g. atmospheric vertical profiles of trace gases and aerosols. In general, the radiances are simulated using a radiative transfer model (RTM) and RTM and other parameters (state vector elements) are adjusted until an 'optimal' match is achieved between observed and simulated radiances. One algorithm (WFM-DOAS (WFMD) (Buchwitz et al., 2000; Schneising et al., 2008, 2009)) is based on least-squares and does not use a priori information to constrain the fit parameters. As a consequence, the resulting ~~XCO2 and XCH4~~ XCO_2 and XCH_4 products are typically somewhat 'noisier' compared to the OE products.

The ~~XCO2 and XCH4~~ XCO_2 and XCH_4 data products from SCIAMACHY are generated within the GHG-CCI project (Buchwitz et al., 2015) of ESA's Climate Change Initiative (CCI, Hollmann et al. (2013)) and these products are available from the GHG-CCI website (http://www.esa-ghg-cci.org/). ~~XCO2~~ XCO_2 and/or ~~XCH4~~ XCH_4 products from GOSAT are generated at several institutions in Japan, Europe and the USA and are available from several sources as shown in Table (1). The quality of these GHG-CCI products and the ~~XCO2 and XCH4~~ XCO_2 and XCH_4 products generated elsewhere has been significantly improved during recent years (e.g. Schneising et al., 2012; Yoshida et al., 2013; Dils et al., 2014; Buchwitz et al., 2015) and has now reached quite high maturity when compared to user requirements as formulated in, e.g. GCOS (2011). This can be concluded, for example, from the quality of the latest version of the GHG-CCI SCIAMACHY and TANSO ~~XCO2 and XCH4~~ XCO_2 and XCH_4 data set ('Climate Research Data Package No. 3', CRDP3) (Buchwitz et al., 2016). Based on comparisons with ground-based observations of the Total Carbon Column Observing Network (TCCON, Wunch et al. (2010, 2011)) it has been found that the GCOS requirements for systematic error (< 1 ppm for ~~XCO2~~XCO_2, < 10 ppb for ~~XCH4~~XCH_4) and long-term stability (< 0.2 ppm/year for ~~XCO2~~XCO_2, < 2 ppb/year for ~~XCH4~~XCH_4) are met for nearly all products. As also shown in Buchwitz et al. (2016), the single observation (ground pixel) retrieval precision ~~)~~ (random error primarily due to instrument noise) is about 2 ppm for ~~XCO2~~ XCO_2 from SCIAMACHY and TANSO and ≈15 ppb for TANSO ~~XCH4~~XCH_4. For SCIAMACHY

~~XCH4~~ XCH_4 the precision depends on time period and retrieval algorithm and is in the range 35 - 80 ppb. For some products it has also been investigated to what extent the uncertainty can be reduced upon averaging (Kulawik et al., 2016) and recommendations are given how to take into account error correlations (Reuter et al., 2016), i.e. which values to use for the non-diagonal elements of the error covariance matrix, as an important ~~contrbution~~ contribution to the full characterisation of the data needs for data ~~assimilaton~~ assimilation studies.

Figure (3) presents an overview about GHG-CCI CRDP3 ~~XCO2~~ XCO_2 (left) and ~~XCH4~~ XCH_4 (right) data set in terms of time series and maps. These figures have been generated by gridding the underlying individual ground pixel (Level 2) products to generate a 5°×5° monthly Level 3 'Obs4MIPs' product Buchwitz and Reuter (2016). Each 5°×5° monthly ~~grid cell~~ grid-cell also contains an estimate of the overall uncertainty (also shown in Fig. (3)) which has been computed taking into account random and systematic error components. The grid-cell uncertainty is computed from two terms: (i) using the reported uncertainties as given in the Level 2 (individual ground pixel) product files for each of the used satellite products (using an ensemble of SCIAMACHY and GOSAT Level 2 products) and (ii) using a term accounting for potential regional/temporal biases as obtained from validation using TCCON ground-based data (see above). The first term depends on the number of individual observations added (the error reduces in proportion to the square root of the number of observations added) whereas the latter term is constant and in the range 0.57 – 0.87 ppm depending on satellite XCO_2 product or in the range 6 – 10 ppb for XCH_4. As can be seen from Fig. (3), the uncertainty of the satellite ~~XCO2~~ XCO_2 retrievals for monthly 5°×5° averages is estimated to be typically around 0.5 - 1 ppm (values larger than 1 ppm are typically associated with regions where only few observations per grid cell exist, e.g. due to clouds or higher latitudes corresponding to low sun elevation). For ~~XCH4~~ XCH_4 the uncertainty is on the order of a few ppb (typically 4 - 8 ppb). In Buchwitz and Reuter (2016), also initial TCCON validation results of the Obs4MIPs products are presented. It is shown that the ~~XCO2~~ XCO_2 product agrees with monthly averaged TCCON ~~XCO2~~ XCO_2 within 0.29 ± 1.2 ppm ($1\,\sigma$) and the ~~XCH4~~ XCH_4 product within 2.0 ± 10.7 ppb. This is hardly worse ~~that~~ than the results which have been obtained by careful validation of the individual ground pixel retrievals taking into account the best possible spatio-temporal co-location and considering the averaging kernels, etc. (e.g. Buchwitz et al., 2016). Note that the computed differences of Obs4MIPs monthly 5°×5° satellite products with monthly averaged TCCON include the errors of the satellite data, errors of the TCCON products, errors due to neglecting altitude sensitivity differences (averaging kernels), and representativity error. This indicates that the representativity error is quite small (at least for monthly 5°×5° spatio-temporal sampling and resolution), probably on the order of 0.1 - 0.2 ppm for ~~XCO2~~ XCO_2 and a few ppb for ~~XCH4~~ XCH_4 (it is planned to quantify this error in the future but currently only these rough estimates are available). Note that detailed information on all GHG-CCI products is available on the GHG-CCI website in terms of technical

documents, links to peer-reviewed publications and figures including detailed maps for each month and each individual data product.

The SCIAMACHY and TANSO ~~XCO2 and XCH4~~ XCO_2 and XCH_4 retrievals have been used in a number of scientific studies to address important questions related to the sources and sinks of atmospheric CO_2 and CH_4 by atmospheric inversion studies (e.g. Bergamaschi et al., 2013; Houweling et al., 2015) and more recently also in a data assimilation context for optimising model parameters (Chevallier et al., 2017). Obviously, the longer the time series and the more accurate it is, the larger the information content of a given data set. Therefore, further improvements are desired (Chevallier et al., 2017) and possible (at least in terms of time series extension but likely also in further reduction of remaining biases).

3.2.2 ~~Reflectamce-based~~ Reflectance-based vegetation dynamics/activity

Since the early beginnings of remote sensing the state and evolution of the vegetation has been monitored by satellites. An early attempt to analyse vegetation dynamics from space is to calculate the ~~Normalized~~ Normalised Difference Vegetation Index (NDVI), defined as the ratio between the difference of near-infrared, NIR, and visible red, Red, spectral bands and the sum of NIR and Red: NDVI = (NIR - Red)/(NIR + Red) ~~, Deering et al. (1975))~~(Deering et al., 1975). The advantage of an index such as NDVI lies in its simplicity and applicability to sensors with few spectral bands such as the Advanced Very High Resolution Radiometer (AVHRR). Therefore this index has been applied for numerous purposes over the last 30 years or so. But NDVI is not a geophysical variable and it is sensitive to various perturbing factors such as atmospheric constituents (aerosols, water ~~vapor~~vapour), directional effects (geometry of illumination and observation), changes in soil background ~~color~~ col or changes (depending on soil water content)(e.g. Pinty et al., 1993; Goel and Qin, 1994; Leprieur et al., 1994; Dorigo et al., 2007). There have been many attempts in modifying NDVI and developing additional vegetation indices (VIs) to overcome its limitations, for example: Soil-Adjusted Vegetation Index (Huete, 1988), Atmospherically Resistant Vegetation Index (Kaufman and Tanre, 1992) or Global Environmental Monitoring Index (Pinty and Verstraete, 1992). These indices generally exhibit some improvement in one respect but at the expense of ~~some~~ degradation in another respect. Pinty et al. (2009) demonstrate the limitations of such VIs in representing the complex radiative properties of the canopy-soil system over the visible to NIR albedo range. Satellite-derived LAI products (e.g. Liu et al., 2014) seem to be an alternative to VIs. LAI is, however, model dependent, i.e. the correct interpretation of this variable depends on the formulation of the model used in the retrieval scheme, and may differ from the interpretation adopted by the land biosphere model used for assimilating the LAI product (Disney et al., 2016a).

A rational approach to ~~address~~ addressing all these issues ~~at once~~ together is to design a ~~physically based quantity which closely follows~~ physically-based quantity which is determined by the state of the ~~vegetation~~canopy-soil system. The Fraction of Absorbed Photosynthetically Active Radiation

(FAPAR)~~provides some kind of~~ , which is a normalised fraction with values ranging from 0 to 1, provides information on the photosynthetic activity of the land vegetation. ~~FAPAR~~ It is recognised as an Essential Climate Variable (ECV) (GCOS, 2011) and is based on the land surface radiation budget. It is defined as the fraction of the photosynthetically active radiation (i.e. incoming solar radiation in the spectral region 0.4–0.7μm) that is absorbed by the vegetation canopy (see also Pickett-Heaps et al. (2014) for a mathematical definition). Several FAPAR products are derived from a variety of optical sensors (e.g. ATSR, MERIS, MISR, MODIS, SEVIRI, SeaWiFS, VEGETATION) at different spatial and temporal resolutions. Although there has been substantial efforts to ~~harmonize~~ harmonise products across sensors (Ceccherini et al., 2013) and establish standards and validation practices (e.g. Widlowski, 2010) there are still considerable differences among the products. These differences can mainly be associated to differences in the retrieval methodology as well as to the quality of input variables. A recent overview of various FAPAR products and their specifications, but without an ~~assessemt~~ assessment of product uncertainties, is given by Gobron and Verstraete (2009). Table (2) summarises the characteristics of the most common FAPAR products.

Several studies have compared the performance of different satellite-derived FAPAR products: McCallum et al. (2010) looked at four FAPAR ~~datasets~~ data sets over Northern Eurasia for the year 2000, Pickett-Heaps et al. (2014) evaluated six products across Australia, D'Odorico et al. (2014) compared three products over Europe~~and~~ , Tao et al. (2015) assessed five products over different land cover types, and Disney et al. (2016a) compared two FAPAR products derived from GlobAlbedo and MODIS data. Pickett-Heaps et al. (2014) concluded that although all six evaluated products display robust spatial and temporal patterns there is considerable disagreement in the absolute magnitude amongst the products and ~~non~~ none of the products outperforms the others. This has also been confirmed by the studies of D'Odorico et al. (2014) and Tao et al. (2015). One of the reasons for these differences are different assumptions on the underlying biome types. They also reviewed the consistency of the FAPAR products against ~~in-situ~~ in situ field measurements, the mean difference between the EO products and the ~~in-situ~~ in situ field measurements is around 0.1~~(as FAPAR is a normalised fraction values range from 0 to 1)~~. This estimate is confirmed by the study of Tao et al. (2015) who suggest an average uncertainty of 0.14 from validation against total FAPAR and 0.09 from validation against green FAPAR ~~in-situ~~ in situ measurements. In their comparison of Joint Research Centre–Two-stream Inversion Package (JRC-TIP) MODIS, JRC MGVI (based on MERIS) and Boston University MODIS products (see Tab. 2) D'Odorico et al. (2014) placed special emphasis on the assessment of the product uncertainties by not only comparing the uncertainties (or quality indicators) as proposed by the product teams but also by calculating an independent theoretical uncertainty based on the triple collocation (TC) method (see Sec. 3.2.4). While the uncertainties specified by the product teams differed by up to 0.1 among the products, the TC method suggested more consistent uncertainties among the three products of around 10-20% of the signal.

The JRC-TIP (Pinty et al., 2007) is an inverse modelling system that was ~~deliberately~~ explicitly designed to retrieve a set of land surface variables, including FAPAR, in a form that is compliant with the requirements for assimilation into terrestrial biosphere models, hence we focus in the following on this product. TIP is based on a one-dimensional two-stream representation of the radiative transfer in the canopy-soil system (Pinty et al., 2006) and applies the same inversion approach as CCDAS, which is briefly sketched in 2.2 and detailed in ~~Rayner et al. (2016); Kaminski and Mathieu (2016)~~ Rayner et al. (2016) and Kaminski and Mathieu (2016). In a first step it retrieves a set of model parameters describing the state of the vegetation canopy system including ~~their~~ the full uncertainty covariance of the parameters by combining prior information with observed radiant fluxes. Further, the model is used to propagate this PDF forward onto the simulated fluxes such as FAPAR. TIP uses observed broadband albedo in the NIR and visible spectral domains as input~~from which it retrieves the effective (~~. The prior information used in the retrieval is constant in space and time, i.e. ~~model-dependent) quantities such as FAPAR, leaf area index (LAI) besides other radiatitve quantities~~all variability is determined from space (Kaminski et al., 2017). This is in contrast to other retrieval approaches, which are based on prescribed land cover maps (e.g. Liu et al., 2014). Long-term global records of JRC-TIP products (see Table 2) have been retrieved from broadband albedos provided by MODIS collection 5 (Pinty et al., 2011b, c) and Globalbedo ~~(Disney et al., 2016b)~~(Disney et al., 2016a). Products are provided for each of the respective 16-day (MODIS) and 8-day (Globalbedo) synthesis periods. To reduce disk space, by default, JRC-TIP products are delivered without correlations among the uncertainties between individual variables, even though these correlations are available. An estimate of uncertainty correlation in space and time is not provided. Both JRC-TIP records are provided in the native 1 km resolution of the albedo input products. In order to maintain the above-mentioned compliance with terrestrial models, coarser resolution products are to be derived by applying JRC-TIP to aggregated albedo inputs ~~(as in Disney et al., 2016b)~~(as in Disney et al., 2016a). JRC-TIP products are validated at site (Pinty et al., 2007, 2008, 2011a) and regional scales ~~(Disney et al., 2016b)~~ (Disney et al., 2016a); more details on JRC-TIP are given in Kaminski et al. (2017).

3.2.3 Biogeochemical-based vegetation activity

Sun-induced fluorescence (SIF) is an electromagnetic signal emitted as a two-peak spectrum between 650 and 850 nm by the chlorophyll$-a$ of green plants under solar radiation. SIF can be directly related to photosynthetic electron transport rates and yields a mechanistic link to photosynthesis and the subsequent gross carbon uptake by terrestrial vegetation (GPP) (Porcar-Castell et al., 2014). Recent developments in satellite-based spectroscopy have enabled the first retrievals of SIF from space (Frankenberg et al., 2011c; Joiner et al., 2011), which holds the promise of enabling new approaches to globally monitoring terrestrial photosynthesis. For example, a high linear correlation between data-driven GPP estimates and SIF retrievals at global and annual scales was reported by Frankenberg et al. (2011c); Guanter et al. (2012). The skills of SIF as a proxy for photosynthetic ac-

tivity and GPP were also reported by studies over different ecosystems, like the Amazon ~~rainforest~~ rain forest (Lee et al., 2013; Parazoo et al., 2013), large crop belts (Guanter et al., 2014), and the boreal forests in Eurasia and North America (Walther et al., 2015). But in the context of DA and in order to extract the maximal benefit from SIF data, the complex processes responsible for SIF in the plants' photochemical systems (as mentioned above) require complex models as observation operators for SIF.

The global retrieval of SIF from space ~~lies~~ relies on the principle of *in-filling* of solar Fraunhofer lines by SIF (Frankenberg et al., 2011b). Fraunhofer lines are absorption features in the solar spectrum, caused by elements in the solar atmosphere and sufficiently resolved by atmospheric spectrometers. Because of the additive nature of SIF, the fractional depth of the Fraunhofer lines detected by the satellite instrument decreases with the amount of SIF being emitted at the same wavelength. The retrieval of SIF from space is then based on the evaluation of the depth of the Fraunhofer lines present in red and NIR top-of-atmosphere spectra. The retrieval forward model is thus simple and can be linearised (e.g. Guanter et al., 2012; Köhler et al., 2015b), ~~so the inversioncan be easily solved by least squares optimisation~~which simplifies the inversion.

Fraunhofer line-based SIF retrievals tend to be accurate but not precise: uncertainties are dominated by a random component associated to instrumental noise, which is linearly mapped into SIF retrievals. The amplitude of instrumental noise, and hence 1-σ single-retrieval errors, scale with at-sensor radiance for the most common case of grating-based spectrometers dominated by multiplicative noise. This implies that retrieval errors are mostly driven by surface brightness and sun zenith angles (Guanter et al., 2015). Because of this high contribution of random errors to the total retrieval uncertainty, single SIF retrievals are commonly linearly-aggregated as spatio-temporal composites in which random errors are reduced. The ~~amount~~ number of retrievals to be aggregated into a given ~~gridbox~~ grid-cell results from a compromise between spatial resolution, temporal resolution and precision of the gridded product, the size of the spatial and temporal bins being exchangeable in terms of their effect on the random uncertainty. The random uncertainty of the resulting spatio-temporal composites is then not only driven by surface albedo and illumination, but also by the number of soundings going into a given ~~gridbox~~grid-cell, which is in turn defined by cloudiness and latitude (in the case of overlapping orbits). Detailed analyses of random errors in SIF retrievals for different ~~spaceborne~~ space-borne instruments can be found in Frankenberg et al. (2011b) and Guanter et al. (2015).

Global SIF data sets have been or are being derived from GOSAT, MetOp's Global Ozone Monitoring Experiment-2 (GOME-2), ENVISAT's SCIAMACHY and the OCO-2 mission (Joiner et al., 2011; Frankenberg et al., 2011c; Guanter et al., 2012; Joiner et al., 2012, 2013; Köhler et al., 2015a, b; Wolanin et al., 2015; Joiner et al., 2016; Frankenberg et al., 2014). All four missions except for SCIAMACHY are still operating. Sample SIF maps from GOSAT, GOME-2 and SCIAMACHY for July 2010 are displayed in Fig. 4. ~~All four missions except for SCIAMACHY are still operating.~~

~~The spectal~~The spectral, spatial and temporal sampling of single SIF soundings varies for each instrument, as it is summarised in Table 3. For example, GOME-2 and SCIAMACHY provide SIF retrievals in the red and NIR spectral regions with global coverage and a relatively high temporal resolution. However, this comes at the expense of a coarse spatial resolution, which is $40\times80\,km^2$ for GOME-2 ($40\times40\,km^2$ for GOME-2 on MetOp-A since July 2013) and $30\times240\,km^2$ for SCIAMACHY. On the other hand, GOSAT and OCO-2 do not provide spatially-continuous measurements (i.e. no global coverage), but single soundings in the NIR have a much higher spatial resolution than those of GOME-2 and SCIAMACHY. In particular, OCO-2 soundings correspond to ground areas of about $4\,km^2$, which is substantially finer than that of the other data sets. The number of soundings per day by OCO-2 is also much larger (about 100x) than that by the other instruments (Frankenberg et al., 2014), which makes OCO-2 SIF to be the most suited data set for studies over areas not requiring a continuous spatial sampling but benefiting from a high spatial resolution. This is the case, for example, of tropical and boreal forests: spatial continuity is less critical for those ecosystems because they are relatively homogeneous over large spatial scales, whereas the high spatial resolution is important to maximise the number of clear-sky soundings during the parts of the year with persistent cloudiness.

Concerning near-future perspectives for SIF monitoring, it can be expected that the limitations in spatial resolution and coverage of existing SIF products will be alleviated with the advent of the TROPOspheric Monitoring Instrument (TROPOMI) scheduled for launch onboard the Sentinel-5 Precursor satellite mission by mid 2017 (Table 3). TROPOMI will enable SIF retrievals in the red and NIR regions similar to GOME-2 and SCIAMACHY, but with a 7 km pixel, daily global coverage and a number of clear-sky observations per day ≈200 larger than GOME-2 and ≈600 larger than SCIAMACHY. The SIF product from TROPOMI can therefore be anticipated to have a much higher spatio-temporal resolution and signal-to-noise ratio than those from GOME-2 and SCIAMACHY (Guanter et al., 2015). ~~Complementary~~Complementarily, the FLuorescence EXplorer (FLEX) ~~(Drusch and FLEX Team, 2015)~~ (Drusch et al., 2017) has recently been selected for implementation by ESA, with launch currently expected for 2022. FLEX will provide global measurements of SIF in the red and NIR with at a relatively low temporal resolution, but with the finest spatial resolution of all existing and upcoming ~~spaceborne~~ space-borne instruments.

3.2.4 Soil moisture

Soil moisture is measured ~~in-situ~~ in situ through large-scale soil moisture monitoring networks (Dorigo et al., 2011; Ochsner et al., 2013) or at various FLUXNET sites (Baldocchi et al., 2001). Yet, these point observations have only limited coverage in space time, have spatially very divergent properties (Dorigo et al., 2013), and often contain large representativeness errors at the scale of global ecosystem models (Gruber et al., 2013). Satellite remote sensing in the microwave domain has the potential to overcome many of these issues. Microwave remote sensing uses the contrasting

dielectric properties of water, air, ice, and soil particles to infer the water content in the soil column (Owe et al., 2008). Both passive radiometer systems, measuring the emitted microwave radiance ('brightness temperatures'), and active radar systems, measuring backscattered microwave radiance, can be used to retrieve soil moisture. Various approaches exist that convert brightness temperatures and backscatter measurements into estimates of soil moisture, including radiative transfer model inversion approaches (e.g. Kerr et al., 2012; Owe et al., 2008), neural networks (e.g. Rodriguez-Fernandez et al., 2015), linear regressions (Al-Yaari et al., 2016, e.g.), and change detection methods (Wagner et al., 1999). The latter is commonly applied to scatterometer measurements and yields, in contrast to the other approaches which provide soil moisture as volumetric water content, soil moisture as a percentage of total saturation. Microwave sensors operate in different frequency (wavelength) domains, of which L-band (with a wavelength of ≈23 cm) and C-band (≈5 cm) are most commonly used for retrieving soil moisture (Kerr et al., 2012; Owe et al., 2008; Wagner et al., 1999). Smaller wavelengths are more sensitive to the vegetation canopy covering the soil and increasingly lose their sensitivity to water. Still, frequencies up to 19 GHz (≈1.5 cm) have proven potential for providing robust soil moisture estimates at the global scale for moderately to sparsely vegetated areas (Owe et al., 2008). Due to the relatively low energy levels and the technical challenges in microwave domain, spatial resolutions of the satellite observations are generally coarse (≈25–50 km) but with high revisit frequencies (up to 1 day). Only Synthetic Aperture Radar is able to provide much higher spatial resolutions~~up till a few~~ , up to a few tens of meters, yet at the cost of ~~the~~ long revisit times. Also observations made by the Gravity Recovery and Climate Experiment (GRACE; Rodell et al. (2009)) are sensitive to soil moisture, but the estimation of soil moisture content from these observations is not straightforward because they are also sensitive to changes in snow, surface water, groundwater, and vegetation.

Since the release of the first global soil moisture ~~datasets~~ data sets from microwave sensors in the early 2000s the number of available soil moisture products and missions has rapidly expanded (De Jeu and Dorigo, 2016). Several (pre-)operational products are now available from a wide variety of data providers and space ~~organizations~~ organisations (Table 4). While initially soil moisture products were based on sensors mainly designed for other purposes (such as ASCAT, AMSR2, and Sentinel-1), ESA and NASA launched their own dedicated soil moisture satellite missions SMOS and SMAP (Kerr et al., 2012; Entekhabi et al., 2010). ~~Apart from the Sentinel-1 mission, which primarily targets the provision of high resolution observations over Europe, all currently active missions provide a nearly global coverage at a coarse resolution approximately every 1-2 days.~~ Differences between the various products exist in their technical design, observation bands, and retrieval algorithms, which often result in complementary strengths over different land cover types (Alyaari et al., 2015; Dorigo et al., 2010; Liu et al., 2011). The missions also differ in their degree of operationalization: While SMOS and SMAP are primarily scientific concept demonstrators, AMSR2 continues the legacy of C-band radiometer observations started by JAXA and NASA in 2002 with

the launch of AMSR-E, while ASCAT is embedded in a fully operational program of weather observing satellites with a guaranteed continuation at least until 2023 and a follow-on mission already under development (Wagner et al., 2013). Apart from the target variable surface soil moisture, some products come with estimates of freeze/thaw state and ~~vegetation optical depth~~VOD, which are disentangled from the soil moisture impacts on the measured microwave signal during the retrieval process.

As none of the currently active missions covers a period long enough to study climate change impacts, ESA's Climate Change Initiative (CCI) endorsed the combination of available soil moisture products from active and passive microwave sensors into a consistent multi-decadal record. The ESA CCI soil moisture product currently combines soil moisture products from 11 different sensors into a ~~homogenized~~ homogenised daily product spanning the period 1978-2015 (Liu et al., 2012, 2011; Dorigo et al., 2016). Several studies have demonstrated the value of ESA CCI soil moisture for assessing long-term interactions between soil moisture and vegetation productivity (Barichivich et al., 2014; Chen et al., 2014; Dorigo et al., 2012; Muñoz et al., 2014).

Key to a proper assimilation of remotely sensed soil moisture into carbon models is a correct ~~characterization~~ characterisation of its errors. Apart from instrument errors which are common to all observations, the quality of microwave-based soil moisture retrievals is particularly impacted by vegetation cover, soil frost, snow cover, open water, topography, surface roughness, urban structures, and radio frequency interference (Dorigo et al., 2010; Kerr et al., 2012). Observations where a strong adverse impact of these factors is detected are usually masked during processing, which may lead to data gaps for certain areas or periods of the year (Dorigo et al., 2015). If cases where their impact on the soil moisture retrieval is only moderate, the errors that they introduce are either simulated during the retrieval itself using error propagation methods, or assessed a posteriori against reference data using various statistical methods (Draper et al., 2013).

While the ASCAT and AMSR2 products come with an estimate of the error variance for each observation by error propagation (Naeimi et al., 2009; Parinussa et al., 2011) this is still not common practice for all soil moisture products. Yet, no error propagation model perfectly represents all error sources and interactions (Draper et al., 2013). On the other hand, the use of ~~in-situ~~ in situ soil moisture measurements to estimate random errors is hampered by their heterogeneous nature and large spatial representativeness errors (Gruber et al., 2013). As an alternative, in recent years triple collocation analysis (TCA) has firmly established itself as a robust alternative to estimate random errors in soil moisture ~~datasets~~ data sets without the need of an absolute 'true' reference (Dorigo et al., 2010; Scipal et al., 2008). TCA estimates the error variances of three spatially and temporally collocated soil moisture ~~datasets~~ data sets with independent error structures, e.g. a radiometer-based, a scatterometer-based, and a land surface model soil moisture ~~dataset~~data set. Recently, the TCA has been intensively elaborated, e.g. to solve for collinearities between errors (Gruber et al., 2016b) and non-linear dependencies between ~~datasets~~ data sets (Zwieback et al., 2016). The most remark-

able advancement has been to express TCA-based error estimates as a signal-to-noise ratio, which facilitates a direct intercomparison of the skill of ~~datasets~~ data sets independent of their dynamic ranges (Gruber et al., 2016a), see Figure 5. Although the TCA provides an estimate that is entirely independent of any retrieval model assumptions, it only provides a single average error estimate for the entire period under consideration. Thus, synergistic use of error propagation and triple collocation estimates shall be exploited to better capture the temporal error dynamics needed for an optimal assimilation into carbon models. Due to the recent progress in product quality, error ~~characterization~~characterisation, and operationalization, satellite-based soil moisture products have reached the level of maturity that allows for a systematic assimilation into land surface models~~to improve the models' hydrology . For~~ . These products have been used to improve model hydrology by, for example, Martens et al. (2016) who showed that the assimilation of SMOS and ESA CCI soil moisture generally has a small positive impact on soil water ~~storages~~ storage and evaporative fluxes as simulated by the GLEAM land evaporation model. Surface soil moisture from ASCAT is assimilated operationally in near-real-time into ECMWF Land Data Assimilation System to obtain root-zone soil moisture (Albergel et al., 2012). Global root-zone soil moisture products based on SMOS and SMAP are derived by a slightly different approach, which assimilate the observed brightness temperatures instead of the retrieved surface soil moisture products (Lannoy and Reichle, 2016). The assimilation of satellite-based soil moisture products in terrestrial carbon cycle models has been described above.

3.2.5 Biomass

Continental-scale biomass maps have been produced from space using both radar and lidar~~; these rely on the returns from transmitted power, so are known as active sensors~~. Biomass here refers to above-ground biomass (AGB), since there are no methods to measure the below-ground component, and this is typically inferred from AGB using allometric equations. Furthermore, the emphasis is on the AGB of forests, although a global ~~dataset~~ data set of AGB in all biomes for the period 1993-2012 has been produced based on VOD data from global passive microwave ~~satellite data~~sensors, hence with spatial resolution of 10 km or coarser (Liu et al., 2015). The AGB product is derived from a regression of VOD against observations of AGB from ground-based inventory data.

Using long time series of C-band radar data provided by the ESA Envisat satellite, the growing stock volume of northern hemisphere boreal and temperate forests has been estimated (Santoro et al., 2011). Although available at 0.01° resolution, the accuracy of growing stock volume at this scale is comparatively poor, and spatial averaging provides more reliable results: at 0.5° spacing, estimated growing stock volume has a relative accuracy of 20-30% when tested against inventory data (Santoro et al., 2013). Thurner et al. (2014) used this product to derive the carbon stock (above- and below-ground) in boreal, temperate mixed and broadleaf, and temperate coniferous forests of forests above 30° N (40.7, 24.5 and 14.5 PgC respectively). These values have estimated accuracies of

around 33-39% under a conservative approach to estimate uncertainty. Santoro et al. (2015) provide a high resolution data set (0.01°) over the northern hemisphere with a relative RMSE against National Forest Inventory between 12% and 45%.

For tropical forests, the key sensor is the Geoscience Laser Altimeter System (GLAS) onboard the Ice, Cloud and land Elevation Satellite (ICESat) which failed in 2009 (Lefsky, 2010). Its archive of forest height estimates was the core ~~dataset~~ data set exploited to produce two pan-tropical biomass maps (Saatchi et al., 2011; Baccini et al., 2012) at grid scales of 1 km and 500 m respectively; Saatchi et al. (2011) also provide a map of the errors associated with the biomass estimates at each pixel. This is produced by combining measurement errors, allometry errors, sampling errors, and prediction errors, which are treated as independent and spatially uncorrelated. Further details are given in the supplementary material to Saatchi et al. (2011). In an attempt to resolve differences between these two maps, Avitabile et al. (2016) used an independent reference ~~dataset~~ data set of field observations to remove the biases in the maps and then combined them to estimate the AGB in the tropical belt (23.4° S to 23.4° N). Testing against a reference ~~dataset~~ data set not used in the fusion process indicated that the fused map had a RMSE 15-21% lower than that of the input maps and nearly unbiased estimates.

However, there are unresolved questions about large-scale biomass patterns across the Amazon inferred from in situ and satellite data. Biomass maps derived from satellite data in Saatchi et al. (2011) and Baccini et al. (2012) differ significantly from each other and from biomass maps derived from in situ plots distributed across Amazonia using kriging (Mitchard et al., 2014). Neither satellite product exhibits the strong increase in biomass from southwestern to northeastern Amazonia inferred from in situ data. Mitchard et al. (2014) attributed this to failure to account for gradients in wood density and regionally varying tree height-diameter relations when estimating biomass from the satellite data. Saatchi et al. (2015) reject this analysis and claim that the trends and patterns in Mitchard et al. (2014) are erroneous and a consequence of inadequate sampling. Resolving this disagreement is of fundamental importance since it raises basic questions about accuracy, uncertainty, and representativeness for both in situ and satellite-derived biomass data.

The next 4-5 years will dramatically improve our global knowledge of biomass, with the launch of three missions aimed at measuring forest structure and biomass. The ESA BIOMASS mission (European Space Agency, 2012), to be launched in 2021, is a P-band radar that will provide near-global measurements of forest biomass and height. Measurements from airborne sensors indicate that even in dense tropical forests affected by topography, the P-band frequency used by BIOMASS will give sensitivity to biomass up to 350-450 t/ha (Minh et al., 2014; Villard and Toan, 2015). Around the same time the NASA-ISRO SAR mission (NISAR) based on an L-band sensor will be deployed, providing measurements of biomass in lower biomass forests (up to 100 t ha^{-1}). These highly complementary missions will be further complemented by the NASA Global Ecosystem Dynamics Investigation vegetation lidar to be placed on the International Space Station around 2019; this aims to

provide the first global, high-resolution observations of the vertical structure of tropical and temperate forests, from which biomass may be estimated.

As well as limitations caused by mission lifetimes, satellite measurements of biomass are unlikely to be sensitive enough to measure biomass increment except in rapidly growing plantations and tropical forests. Hence an important ancillary data set for studies aiming to relate biomass to climate and environment is tree ring data (https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/tree-ring).

4 Conclusions

In the context of carbon cycle data assimilation this paper reviews the requirements and summarises the availability and characteristics of some selected observations with a special focus on remotely sensed Earth observation data. Observations are key for understanding the carbon cycle processes and are an important component for any data assimilation system. In this context the provision of systematic and sustained observing systems on an operational basis is becoming more and more important.

An example for such an operational network for in situ data is the Integrated Carbon Observing System (ICOS, see also https://www.icos-ri.eu). ICOS is a pan-European infrastructure for carbon observations, which provides high-quality in situ observations (both fluxes as well as atmospheric concentrations) over Europe and over ocean regions adjacent to Europe with a long-term perspective. ICOS consists of central facilities for co-ordination, calibration and data in conjunction with networks of atmospheric, oceanic and ecosystem observations as well as a data distribution centre, the Carbon Portal, providing discovery of and access to ICOS data products such as derived flux information. Other (quasi-)operational networks measuring atmospheric CO_2 concentrations are maintained, for instance, by the National Oceanic and Atmospheric Administration (NOAA) Climate Monitoring and Diagnostics Laboratory, and the Scripps Institution of Oceanography, both USA, and the CSIRO Global Atmospheric Sampling Laboratory, Australia.

An example for an operational space-based Earth observing programme in Europe is the fleet of so-called Sentinel satellites of the Copernicus programme. Copernicus aims at providing Europe with continuous and independent access to Earth observation data and associated services (transforming the satellite and additional in situ data into value-added information by processing and analysing the data) in support of Earth System Science (Berger et al., 2012). Currently, six different Sentinel missions are planned (and have partly been launched). So far, a dedicated mission for monitoring the carbon cycle, i.e. an instrument measuring the atmospheric CO_2 composition, is not yet included in the Copernicus monitoring programme (see Ciais et al., 2015), however, the series of Sentinel satellites will be extended in the future and likely include a CO_2 mission. Other operational EO programmes are operated by e.g. NOAA and the Japanese Aerospace Exploration Agency.

The paper also briefly recapitulates the assimilation systems capable of integrating these data, a more comprehensive description of the underlying formalism is given in Rayner et al. (2016) while MacBean et al. (2016) discuss the implementation strategies for a multiple data assimilation system and ~~its~~ their impacts on the results. To take maximum advantage of these data streams in carbon cycle data assimilation studies it is of utmost importance to have the appropriate knowledge of the ~~observational~~ uncertainty characteristics of the observational data, here with a focus on ~~atellite~~ satellite products. This includes an understanding of the observable and its representativeness in order to develop the appropriate observation operator (see also Kaminski and Mathieu, 2016) but also the structure of any biases, random errors and error covariances (that is both the diagonal and off-diagonal elements quantifying the error correlations between different observations).

The benefit of using multiple data streams in a CCDAS lies in the complementarity of the data, and thus in the ability to constrain different components of the underlying process model. ~~For example, while FAPAR data constrain mainly the phenology component of a terrestrial carbon cycle model, soil moisture data, in contrast, constrain the hydrological component, but both components are important elements of the model and determine the simulated carbon fluxes.~~ In fact, because of the model internal interactions and feedbacks among the components the simultaneous assimilation of complementary observations has synergistic effects such that the constraint is larger than the sum of the individual constraints as shown for instance by Kato et al. (2013) ~~assimilating~~ who assimilated observations of FAPAR and latent heat flux.

As a final remark one important aspect of observational data is their continuity since much of the important information is contained in response to climate anomalies. Fortunately, the set up of operational ~~observeing~~ observing systems such as ICOS for ~~in-situ~~ in situ data or Copernicus for satellite data has created the necessary infrastructure to ensure such a long-term perspective in the provision of Earth observations.

Appendix A: List of Acronyms

ACE-FTS	Atmospheric Chemistry Experiment - Fourier Transform Spectrometer
AGB	Above Ground Biomass
AIRS	Atmospheric Infrared Sounder
AMSR2	Advanced Microwave Scanning Radiometer 2
AMSR-E	Advanced Microwave Scanning Radiometer - Earth Observing System
ASCAT	Advanced Scatterometer
ATSR	Along Track Scanning Radiometers
AVHRR	Advanced Very High Resolution Radiometer
CCDAS	Carbon Cycle Data Assimilation System
CCI	Climate Change Initiative
ECMWF	European Centre for Medium-Range Weather Forecasts
ECV	Essential Climate Variable
EO	Earth Observation (in this form generally understood as from space)
ESA	European Space Agency
FAPAR	Fraction of Absorbed Photosynthetically Active Radiation
FLEX	FLuorescence EXplorer
GCOM-W1	Global Change Observation Mission 1st-Water
GLAS	Geoscience Laser Altimeter System
GLEAM	Global Land Evaporation Amsterdam Model
GOME-2	Global Ozone Monitoring Experiment-2
GOSAT	Greenhouse Gases Observing Satellite
GPP	Gross Primary Productivity
IASI	Infrared Atmospheric Sounding Interferometer
ICOS	Integrated Carbon Observing System
ICESat	Ice, Cloud and land Elevation Satellite
ISRO	Indian Space Research Organisation
JAXA	Japan Aerospace Exploration Agency
JRC-MGVI	Joint Research Centre – MERIS Global Vegetation Index
JRC-TIP	Joint Research Centre – Two-stream Inversion Package
LAI	Leaf Area Index
MERIS	Medium Resolution Imaging Spectrometer
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MISR	~~Multiangle~~ Multi-angle Imaging SpectroRadiometer
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NDVI	~~Normalized~~ Normalised Difference Vegetation Index
NIR	Near Infrared
NOAA	National Oceanic and Atmospheric Administration

Obs4Mips	Observations for Model Intercomparisons Project
OCO-2	Orbiting Carbon Observatory 2
OE	Optimal Estimation
PDF	Probability Density Function
SAR	Synthetic Aperture Radar
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric ~~Chartography~~ Cartography
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
SEVIRI	Spinning Enhanced Visible and InfraRed Imager
SIF	Sun-Induced Fluorescence
SMAP	Soil Moisture Active Passive
SMOS	Soil Moisture Ocean Salinity
SWIR	Shortwave Infrared
TANSO-FTS	Thermal And Near infrared Sensor for carbon Observations - Fourier Transform Spectrometer
TCA	Triple Collocation Analysis
TCCON	Total Carbon Column Observing Network
TCOS	Terrestrial Carbon Observation System
TROPOMI	TROPOspheric Monitoring Instrument
VI	Vegetation Index
VOD	Vegetation Optical Depth

Acknowledgements. M.B. has received funding from ESA via the GHG-CCI project. W.D. is supported by the "TU Wien Wissenschaftspreis 2015" a personal grant awarded by the Vienna University of Technology. Fig. 4 was kindly provided by Philipp Köhler, California Institute of Technology. We acknowledge the support from the International Space Science Institute (ISSI). This publication is an outcome of the ISSI's Working Group on "Carbon Cycle Data Assimilation: How to consistently assimilate multiple data streams".

References

Al-Yaari, A., Wigneron, J., Kerr, Y., de Jeu, R., Rodriguez-Fernandez, N., van der Schalie, R., Bitar, A. A., Mialon, A., Richaume, P., Dolman, A., and Ducharne, A.: Testing regression equations to derive long-term global soil moisture datasets from passive microwave observations, Remote Sensing of Environment, 180, 453 – 464, doi:10.1016/j.rse.2015.11.022, 2016.

Albergel, C., de Rosnay, P., Gruhier, C., Muñoz Sabater, J., Hasenauer, S., Isaksen, L., Kerr, Y., and Wagner, W.: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations, Remote Sensing of Environment, 118, 215–226, doi:10.1016/j.rse.2011.11.017, 2012.

Albergel, C., Munier, S., Leroux, D. J., Dewaele, H., Fairbairn, D., Barbu, A. L., Gelati, E., Dorigo, W., Faroux, S., Meurey, C., Le Moigne, P., Decharme, B., Mahfouf, J.-F., and Calvet, J.-C.: Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX_v8.0: LDAS-Monde assessment over the Euro-Mediterranean area, Geoscientific Model Development Discussions, pp. 1–53, doi:10.5194/gmd-2017-121, 2017.

Alyaari, A., Wigneron, J. P., Ducharne, A., Kerr, Y., Wagner, W., De Lannoy, G., Reichle, R., Al Bitar, A., Dorigo, W., Richaume, P., and Mialon, A.: Global-scale comparison of passive (SMOS) and active (ASCAT) satellite-based microwave soil moisture retrievals with soil moisture simulations (MERRA-Land), Remote Sensing of Environment, 152, 614–626, doi:10.1016/j.rse.2014.07.013, 2015.

Avitabile, V., Herold, M., Heuvelink, G. B. M., Lewis, S. L., Phillips, O. L., Asner, G. P., Armston, J., Ashton, P. S., Banin, L., Bayol, N., Berry, N. J., Boeckx, P., de Jong, B. H. J., DeVries, B., Girardin, C. A. J., Kearsley, E., Lindsell, J. A., Lopez-Gonzalez, G., Lucas, R., Malhi, Y., Morel, A., Mitchard, E. T. A., Nagy, L., Qie, L., Quinones, M. J., Ryan, C. M., Ferry, S. J. W., Sunderland, T., Laurin, G. V., Gatti, R. C., Valentini, R., Verbeeck, H., Wijaya, A., and Willcock, S.: An integrated pan-tropical biomass map using multiple reference datasets, Global Change Biology, 22, 1406–1420, doi:10.1111/gcb.13139, 2016.

Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P. S. A., Dubayah, R., Friedl, M. A., Samanta, S., and Houghton, R. A.: Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps, Nature Climate Change, 2012.

Bacour, C., Peylin, P., MacBean, N., Rayner, P. J., Delage, F., Chevallier, F., Weiss, M., Demarty, J., Santaren, D., Baret, F., Berveiller, D., Dufrene, E., and Prunet, P.: Joint assimilation of eddy covariance flux measurements and FAPAR products over temperate forests within a process-oriented biosphere model, Journal of Geophysical Research: Biogeosciences, 120, 1839–1857, doi:10.1002/2015JG002966, 2015.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities, Bulletin of the American Meteorological Society, 82, 2415–2434, doi:10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2, 2001.

Barbu, A. L., Calvet, J.-C., Mahfouf, J.-F., and Lafont, S.: Integrating ASCAT surface soil moisture and GEOV1 leaf area index into the SURFEX modelling platform: a land data assimilation application over France, Hydrology and Earth System Sciences, 18, 173–192, doi:10.5194/hess-18-173-2014, 2014.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Nino, F., Weiss, M., Samain, O., Roujean, J. L., and Leroy, M.: LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm, Remote Sensing of Environment, 110, 275 – 286, doi:10.1016/j.rse.2007.02.018, 2007.

Barichivich, J., Briffa, K. R., Myneni, R., Van der Schrier, G., Dorigo, W., Tucker, C. J., Osborn, T., and Melvin, T.: Temperature and Snow-Mediated Moisture Controls of Summer Photosynthetic Activity in Northern Terrestrial Ecosystems between 1982 and 2011, Remote Sensing, 6, 1390–1431, doi:10.3390/rs6021390, 2014.

Barrett, D. J.: Steady state turnover time of carbon in the Australian terrestrial biosphere, Global Biogeochemical Cycles, 16, 55–1, 2002.

Bergamaschi, P., Houweling, S., Segers, A., Krol, M., Frankenberg, C., Scheepmaker, R. A., Dlugokencky, E., Wofsy, S. C., Kort, E. A., Sweeney, C., Schuck, T., Brenninkmeijer, C., Chen, H., Beck, V., and Gerbig, C.: Atmospheric CH_4 in the first decade of the 21st century: Inverse modeling analysis using SCIAMACHY satellite retrievals and NOAA surface measurements, Journal of Geophysical Research: Atmospheres, 118, 7350–7369, doi:10.1002/jgrd.50480, 2013.

Berger, M., Moreno, J., Johannessen, J. A., Levelt, P. F., and Hanssen, R. F.: ESA's sentinel missions in support of Earth system science, Remote Sensing of Environment, 120, 84 – 90, doi:http://dx.doi.org/10.1016/j.rse.2011.07.023, the Sentinel Missions - New Opportunities for Science, 2012.

Boesch, H., Baker, D., Connor, B., Crisp, D., and Miller, C.: Global Characterization of CO_2 Column Retrievals from Shortwave-Infrared Satellite Observations of the Orbiting Carbon Observatory-2 Mission, Remote Sensing, 3, 270, doi:10.3390/rs3020270, 2011.

Bontemps, S., Herold, M., Kooistra, L., van Groenestijn, A., Hartley, A., Arino, O., Moreau, I., and Defourny, P.: Revisiting land cover observation to address the needs of the climate modeling community, Biogeosciences, 9, 2145–2157, doi:10.5194/bg-9-2145-2012, 2012.

Boone, C. D., Nassar, R., Walker, K. A., Rochon, Y., McLeod, S. D., Rinsland, C. P., and Bernath, P. F.: Retrievals for the atmospheric chemistry experiment Fourier-transform spectrometer, Appl. Opt., 44, 7218–7231, doi:10.1364/AO.44.007218, 2005.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.: SCIAMACHY: Mission Objectives and Measurement Modes, Journal of the Atmospheric Sciences, 56, 127–150, doi:10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2, 1999.

Braswell, B. H., Sacks, W. J., Linder, E., and Schimel, D. S.: Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations, Global Change Biology, 11, 335–355, doi:10.1111/j.1365-2486.2005.00897.x, 2005.

Buchwitz, M. and Reuter, M.: Merged SCIAMACHY/ENVISAT and TANSO-FTS/GOSAT atmospheric column-average dry-air mole fraction of CO_2 (XCO2), Technical Note, Version 1, http://www.esa-ghg-cci.org/?q=webfm_send/319, 2016.

Buchwitz, M., Rozanov, V. V., and Burrows, J. P.: A near-infrared optimized DOAS method for the fast global retrieval of atmospheric CH_4, CO, CO_2, H_2O, and N_2O total column amounts from SCIA-

MACHY Envisat-1 nadir radiances, Journal of Geophysical Research: Atmospheres, 105, 15 231–15 245, doi:10.1029/2000JD900191, 2000.

Buchwitz, M., Reuter, M., Schneising, O., Boesch, H., Guerlet, S., Dils, B., Aben, I., Armante, R., Bergamaschi, P., Blumenstock, T., Bovensmann, H., Brunner, D., Buchmann, B., Burrows, J., Butz, A., Chédin, A., Chevallier, F., Crevoisier, C., Deutscher, N., Frankenberg, C., Hase, F., Hasekamp, O., Heymann, J., Kaminski, T., Laeng, A., Lichtenberg, G., Mazière, M. D., Noël, S., Notholt, J., Orphal, J., Popp, C., Parker, R., Scholze, M., Sussmann, R., Stiller, G., Warneke, T., Zehner, C., Bril, A., Crisp, D., Griffith, D., Kuze, A., O'Dell, C., Oshchepkov, S., Sherlock, V., Suto, H., Wennberg, P., Wunch, D., Yokota, T., and Yoshida, Y.: The Greenhouse Gas Climate Change Initiative (GHG-CCI): Comparison and quality assessment of near-surface-sensitive satellite-derived CO_2 and CH_4 global data sets, Remote Sensing of Environment, 162, 344 – 362, doi:10.1016/j.rse.2013.04.024, 2015.

Buchwitz, M., Dils, B., Boesch, H., Crevoisier, C., Detmers, D., Frankenberg, C., Hasekamp, O., Hewson, W., Laeng, A., Noël, S., Notholt, J., Parker, R., Reuter, M., and Schneising, O.: ESA Climate Change Initiative (CCI) Product Validation and Intercomparison Report (PVIR) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG) for data set Climate Research Data Package No. 3 (CRDP No. 3), Version 4.0, http://www.esa-ghg-cci.org/?q=webfm_send/300, 2016.

Buchwitz, M., Schneising, O., Reuter, M., Heymann, J., Krautwurst, S., Bovensmann, H., Burrows, J. P., Boesch, H., Parker, R. J., Somkuti, P., Detmers, R. G., Hasekamp, O. P., Aben, I., Butz, A., Frankenberg, C., and Turner, A. J.: Satellite-derived methane hotspot emission estimates using a fast data-driven method, Atmospheric Chemistry and Physics, 17, 5751–5774, doi:10.5194/acp-17-5751-2017, 2017.

Burrows, J. P., Hölzle, E., Goede, A. P. H., Visser, H., and Fricke, W.: SCIAMACHY – scanning imaging absorption spectrometer for atmospheric chartography, Acta Astronautica, 35, 445 – 451, doi:10.1016/0094-5765(94)00278-T, 1995.

Butz, A., Hasekamp, O. P., Frankenberg, C., Vidot, J., and Aben, I.: CH_4 retrievals from space-based solar backscatter measurements: Performance evaluation against simulated aerosol and cirrus loaded scenes, Journal of Geophysical Research: Atmospheres, 115, doi:10.1029/2010JD014514, 2010.

Butz, A., Guerlet, S., Hasekamp, O., Schepers, D., Galli, A., Aben, I., Frankenberg, C., Hartmann, J.-M., Tran, H., Kuze, A., Keppel-Aleks, G., Toon, G., Wunch, D., Wennberg, P., Deutscher, N., Griffith, D., Macatangay, R., Messerschmidt, J., Notholt, J., and Warneke, T.: Toward accurate CO_2 and CH_4 observations from GOSAT, Geophysical Research Letters, 38, doi:10.1029/2011GL047888, 2011.

Butz, A., Galli, A., Hasekamp, O., Landgraf, J., Tol, P., and Aben, I.: TROPOMI aboard Sentinel-5 Precursor: Prospective performance of CH_4 retrievals for aerosol and cirrus loaded atmospheres, Remote Sensing of Environment, 120, 267 – 276, doi:10.1016/j.rse.2011.05.030, 2012.

Cadule, P., Friedlingstein, P., Bopp, L., Sitch, S., Jones, C. D., Ciais, P., Piao, S. L., and Peylin, P.: Benchmarking coupled climate-carbon models against long-term atmospheric CO_2 measurements, Global Biogeochemical Cycles, 24, doi:10.1029/2009GB003556, 2010.

Ceccherini, G., Gobron, N., and Robustelli, M.: Harmonization of Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) from Sea-ViewingWide Field-of-View Sensor (SeaWiFS) and Medium Resolution Imaging Spectrometer Instrument (MERIS), Remote Sensing, 5, 3357, doi:10.3390/rs5073357, 2013.

Chen, T., de Jeu, R. A. M., Liu, Y. Y., van der Werf, G. R., and Dolman, A. J.: Using satellite based soil moisture to quantify the water driven variability in NDVI: A case study over mainland Australia, Remote Sensing of Environment, 140, 330–338, 2014.

Chevallier, F., Alexe, M., Bergamaschi, P., Brunner, D., Feng, L., Houweling, S., Kaminski, T., Knorr, W., van Leeuwen, T. T., Marshall, J., Palmer, P. I., Scholze, M., Sundström, A.-M., and Vossbeck, M.: ESA Climate Change Initiative (CCI) Climate Assessment Report (CAR) for Climate Research Data Package No. 4 (CRDP No. 4) of the Essential Climate Variable (ECV) Greenhouse Gases (GHG), Version 4, http://www.esa-ghg-cci.org/?q=webfm_send/385, 2017.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R., Piao, S., and Thornton, P.: Carbon and Other Biogeochemical Cycles, book section 6, p. 465–570, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.015, 2013.

Ciais, P., Dolman, A. J., Bombelli, A., Duren, R., Peregon, A., Rayner, P. J., Miller, C., Gobron, N., Kinderman, G., Marland, G., Gruber, N., Chevallier, F., Andres, R. J., Balsamo, G., Bopp, L., Bréon, F.-M., Broquet, G., Dargaville, R., Battin, T. J., Borges, A., Bovensmann, H., Buchwitz, M., Butler, J., Canadell, J. G., Cook, R. B., DeFries, R., Engelen, R., Gurney, K. R., Heinze, C., Heimann, M., Held, A., Henry, M., Law, B., Luyssaert, S., Miller, J., Moriyama, T., Moulin, C., Myneni, R. B., Nussli, C., Obersteiner, M., Ojima, D., Pan, Y., Paris, J.-D., Piao, S. L., Poulter, B., Plummer, S., Quegan, S., Raymond, P., Reichstein, M., Rivier, L., Sabine, C., Schimel, D., Tarasova, O., Valentini, R., Wang, R., van der Werf, G., Wickland, D., Williams, M., and Zehner, C.: Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system, Biogeosciences, 11, 3547–3602, doi:10.5194/bg-11-3547-2014, 2014.

Ciais, P., Crisp, D., Denier van der Gon, H., Engelen, R., Heimann, M., Janssens-Maenhout, G., Rayner, P., and Scholze, M.: Towards a European Operational Observing System to Monitor Fossil CO_2 emissions, Final Report from the expert group, European Commission, B-1049 Brussels, Belgium, http://www.copernicus.eu/sites/default/files/library/CO2_Report_22Oct2015.pdf, 2015.

Cihlar, J., Denning, S., Ahem, F., Arino, O., Belward, A., Bretherton, F., Cramer, W., Dedieu, G., Field, C., Francey, R., Gommes, R., Gosz, J., Hibbard, K., Igarashi, T., Kabat, P., Olson, D., Plummer, S., Rasool, I., Raupach, M., Scholes, R., Townshend, J., Valentini, R., and Wickland, D.: Initiative to quantify terrestrial carbon sources and sinks, Eos, Transactions American Geophysical Union, 83, 1–7, doi:10.1029/2002EO000002, 2002.

Cogan, A. J., Boesch, H., Parker, R. J., Feng, L., Palmer, P. I., Blavier, J.-F. L., Deutscher, N. M., Macatangay, R., Notholt, J., Roehl, C., Warneke, T., and Wunch, D.: Atmospheric carbon dioxide retrieved from the Greenhouse gases Observing SATellite (GOSAT): Comparison with ground-based TCCON observations and GEOS-Chem model calculations, Journal of Geophysical Research: Atmospheres, 117, n/a–n/a, doi:10.1029/2012JD018087, d21301, 2012.

Crevoisier, C., Chédin, A., Matsueda, H., Machida, T., Armante, R., and Scott, N. A.: First year of upper tropospheric integrated content of CO_2 from IASI hyperspectral infrared observations, Atmospheric Chemistry and Physics, 9, 4797–4810, doi:10.5194/acp-9-4797-2009, 2009a.

Crevoisier, C., Nobileau, D., Fiore, A. M., Armante, R., Chédin, A., and Scott, N. A.: Tropospheric methane in the tropics – first year from IASI hyperspectral infrared observations, Atmospheric Chemistry and Physics, 9, 6337–6350, doi:10.5194/acp-9-6337-2009, 2009b.

Crisp, D., Atlas, R., Breon, F.-M., Brown, L., Burrows, J., Ciais, P., Connor, B., Doney, S., Fung, I., Jacob, D., Miller, C., O'Brien, D., Pawson, S., Randerson, J., Rayner, P., Salawitch, R., Sander, S., Sen, B., Stephens, G., Tans, P., Toon, G., Wennberg, P., Wofsy, S., Yung, Y., Kuang, Z., Chudasama, B., Sprague, G., Weiss, B., Pollock, R., Kenyon, D., and Schroll, S.: The Orbiting Carbon Observatory (OCO) mission, Advances in Space Research, 34, 700 – 709, doi:http://dx.doi.org/10.1016/j.asr.2003.08.062, trace Constituents in the Troposphere and Lower Stratosphere, 2004.

Crisp, D., Fisher, B. M., O'Dell, C., Frankenberg, C., Basilio, R., Bösch, H., Brown, L. R., Castano, R., Connor, B., Deutscher, N. M., Eldering, A., Griffith, D., Gunson, M., Kuze, A., Mandrake, L., McDuffie, J., Messerschmidt, J., Miller, C. E., Morino, I., Natraj, V., Notholt, J., O'Brien, D. M., Oyafuso, F., Polonsky, I., Robinson, J., Salawitch, R., Sherlock, V., Smyth, M., Suto, H., Taylor, T. E., Thompson, D. R., Wennberg, P. O., Wunch, D., and Yung, Y. L.: The ACOS CO_2 retrieval algorithm – Part II: Global XCO2 data characterization, Atmospheric Measurement Techniques, 5, 687–707, doi:10.5194/amt-5-687-2012, 2012.

Daley, R.: Atmospheric data analysis, Cambridge University Press, Cambridge, UK, 1991.

De Jeu, R. and Dorigo, W.: On the importance of satellite observed soil moisture, International Journal of Applied Earth Observation and Geoinformation, 45, Part B, 107–109, doi:10.1016/j.jag.2015.10.007, 2016.

Deering, D., Rouse, J., Haas, R., and Schell, J.: Measuring forage production of grazing units from Landsat MSS data, Proc. 10th Int. Symp. Remote Sensing Environ., University of Michigan, Ann Arbor, U.S.A., 1975.

Dils, B., Buchwitz, M., Reuter, M., Schneising, O., Boesch, H., Parker, R., Guerlet, S., Aben, I., Blumenstock, T., Burrows, J. P., Butz, A., Deutscher, N. M., Frankenberg, C., Hase, F., Hasekamp, O. P., Heymann, J., De Mazière, M., Notholt, J., Sussmann, R., Warneke, T., Griffith, D., Sherlock, V., and Wunch, D.: The Greenhouse Gas Climate Change Initiative (GHG-CCI): comparative validation of GHG-CCI SCIAMACHY/ENVISAT and TANSO-FTS/GOSAT CO_2 and CH_4 retrieval algorithm products with measurements from the TCCON, Atmospheric Measurement Techniques, 7, 1723–1744, doi:10.5194/amt-7-1723-2014, 2014.

Disney, M., Muller, J.-P., Kharbouche, S., Kaminski, T., Vossbeck, M., Lewis, P., and Pinty, B.: A New Global fAPAR and LAI Dataset Derived from Optimal Albedo Estimates: Comparison with MODIS Products, Remote Sensing, 8, doi:10.3390/rs8040275, 2016a.

Disney, M., Muller, J.-P., Kharbouche, S., Kaminski, T., Voßbeck, M., Lewis, P., and Pinty, B.: A New Global fAPAR and LAI Dataset Derived from Optimal Albedo Estimates: Comparison with MODIS Products, Remote Sensing, 8, 275, doi:10.3390/rs8040275, http://www.mdpi.com/2072-4292/8/4/275, 2016b.

D'Odorico, P., Gonsamo, A., Pinty, B., Gobron, N., Coops, N., Mendez, E., and Schaepman, M. E.: Intercomparison of fraction of absorbed photosynthetically active radiation products derived from satellite data over Europe, Remote Sensing of Environment, 142, 141 – 154, doi:10.1016/j.rse.2013.12.005, 2014.

Dorigo, W., Zurita-Milla, R., de Wit, A., Brazile, J., Singh, R., and Schaepman, M.: A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling, International Journal of Applied Earth Observation and Geoinformation, 9, 165 – 193, doi:10.1016/j.jag.2006.05.003, 2007.

Dorigo, W., De Jeu, R., Chung, D., Parinussa, R., Liu, Y., Wagner, W., and Fernandez-Prieto, D.: Evaluating global trends (1988-2010) in homogenized remotely sensed surface soil moisture, Geophysical Research Letters, 39, L18 405, doi:10.1029/2012gl052988, 2012.

Dorigo, W., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A., Wagner, W., and Drusch, M.: Global automated quality control of in-situ soil moisture data from the International Soil Moisture Network, Vadose Zone Journal, 12, doi:10.2136/vzj2012.0097, 2013.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P., Hirschi, M., Ikonen, J., Jeu, R., Kidd, R., Lahoz, W., Liu, Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., Schalie, R., Seneviratne, S., Smolander, T., and Lecomte, P.: ESA CCI Soil Moisture for improved Earth system understanding: state-of-the art and future directions, Remote Sensing of Environment, under review, 2016.

Dorigo, W. A., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., de Jeu, R. A. M., and Naeimi, V.: Error characterisation of global active and passive microwave soil moisture data sets, Hydrology and Earth System Sciences, 14, 2605–2616, doi:10.5194/hess-14-2605-2010, 2010.

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A., and Jackson, T.: The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, Hydrology and Earth System Sciences, 15, 1675–1698, doi:10.5194/hess-15-1675-2011, 2011.

Dorigo, W. A., Gruber, A., De Jeu, R. A. M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R. M., and Kidd, R.: Evaluation of the ESA CCI soil moisture product using ground-based observations, Remote Sensing of Environment, 162, 380–395, doi:10.1016/j.rse.2014.07.023, 2015.

Draper, C., Reichle, R., de Jeu, R., Naeimi, V., Parinussa, R., and Wagner, W.: Estimating root mean square errors in remotely sensed soil moisture over continental scale domains, Remote Sensing of Environment, 137, 288–298, doi:10.1016/j.rse.2013.06.013, 2013.

Drusch, M. and FLEX Team: FLEX Report for Assessment, ESA SP-1330/2, ESA–ESTEC, Noordwijk (The Netherlands), 2015.

Drusch, M., Moreno, J., Bello, U. D., Franco, R., Goulas, Y., Huth, A., Kraft, S., Middleton, E. M., Miglietta, F., Mohammed, G., Nedbal, L., Rascher, U., Schüttemeyer, D., and Verhoef, W.: The FLuorescence EXplorer Mission Concept - ESA's Earth Explorer 8, IEEE Transactions on Geoscience and Remote Sensing, 55, 1273–1284, doi:10.1109/TGRS.2016.2621820, 2017.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., Kimball, J., Piepmeier, J. R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle, R., Shi, J. C., Spencer, M. W., Thurman, S. W., Tsang, L., and Van Zyl, J.: The soil moisture active passive (SMAP) mission, Proceedings of the IEEE, 98, 704–716, 2010.

Enting, I. G.: Inverse Problems in Atmospheric Constituent Transport, Cambridge University Press, Cambridge, 2002.

European Space Agency: Report for Mission Selection: Biomass, Science authors: Quegan, S., Le Toan T., Chave, J., Dall, J., Perrera, A. Papathanassiou, K., Rocca, F., Saatchi, S., Scipal, K., Shugart, H., Ulander, L. and Williams, M.ESA SP 1324/1, European Space Agency, Noordwijk, the Netherlands., http://esamultimedia.esa.int/docs/EarthObservation/SP1324-1_BIOMASSr.pdf, 2012.

FLUXNET2015: FLUXNET2015 Dataset, available at: http://fluxnet.fluxdata.org/data/fluxnet2015-dataset/, 2015.

Foley, A. M., Dalmonech, D., Friend, A. D., Aires, F., Archibald, A. T., Bartlein, P., Bopp, L., Chappellaz, J., Cox, P., Edwards, N. R., Feulner, G., Friedlingstein, P., Harrison, S. P., Hopcroft, P. O., Jones, C. D., Kolassa, J., Levine, J. G., Prentice, I. C., Pyle, J., Vázquez Riveiros, N., Wolff, E. W., and Zaehle, S.: Evaluation of biospheric components in Earth system models using modern and palaeo-observations: the state-of-the-art, Biogeosciences, 10, 8305–8328, doi:10.5194/bg-10-8305-2013, 2013.

Forkel, M., Carvalhais, N., Schaphoff, S., v. Bloh, W., Migliavacca, M., Thurner, M., and Thonicke, K.: Identifying environmental controls on vegetation greenness phenology through model-data integration, Biogeosciences, 11, 7025–7050, doi:10.5194/bg-11-7025-2014, 2014.

Foucher, P. Y., Chédin, A., Dufour, G., Capelle, V., Boone, C. D., and Bernath, P.: Technical Note: Feasibility of CO_2 profile retrieval from limb viewing solar occultation made by the ACE-FTS instrument, Atmospheric Chemistry and Physics, 9, 2873–2890, doi:10.5194/acp-9-2873-2009, 2009.

Fox, A., Williams, M., Richardson, A. D., Cameron, D., Gove, J. H., Quaife, T., Ricciuto, D., Reichstein, M., Tomelleri, E., Trudinger, C. M., and Wijk, M. T. V.: The {REFLEX} project: Comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data, Agricultural and Forest Meteorology, 149, 1597 – 1615, doi:10.1016/j.agrformet.2009.05.002, 2009.

Francey, R. J., Rayner, P. J., and Allison, C. E.: Global Biogeochemical Cycles in the Climate System, chap. Constraining the global carbon budget from global to regional scales - the measurement challenge, pp. 245 – 252, Academic Press, San Diego, USA, 2001.

Frankenberg, C., Aben, I., Bergamaschi, P., Dlugokencky, E. J., van Hees, R., Houweling, S., van der Meer, P., Snel, R., and Tol, P.: Global column-averaged methane mixing ratios from 2003 to 2009 as derived from SCIAMACHY: Trends and variability, Journal of Geophysical Research: Atmospheres, 116, doi:10.1029/2010JD014849, d04302, 2011a.

Frankenberg, C., Butz, A., and Toon, G. C.: Disentangling chlorophyll fluorescence from atmospheric scattering effects in O2A-band spectra of reflected sun-light, Geophysical Research Letters, 38, doi:10.1029/2010GL045896, 2011b.

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J.-E., Toon, G. C., Butz, A., Jung, M., Kuze, A., and Yokota, T.: New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity, Geophysical Research Letters, 38, doi:10.1029/2011GL048738, 2011c.

Frankenberg, C., O'Dell, C., Berry, J., Guanter, L., Joiner, J., Köhler, P., Pollock, R., and Taylor, T. E.: Prospects for chlorophyll fluorescence remote sensing from the Orbiting Carbon Observatory-2, Remote Sensing of Environment, 147, 1–12, 2014.

GCOS: Global Climate Observing System: Systematic Observation Requirements for Satellite-based Products for Climate, GCOS - 154, https://www.wmo.int/pages/prog/gcos/Publications/gcos-154.pdf, 2011.

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), Journal of Geophysical Research: Biogeosciences, 118, 317–328, doi:10.1002/jgrg.20042, 2013.

Global Carbon Project: Science Framework and Implementation. Earth System Science Partnership (IGBP, IHDP, WCRP, DIVERSITAS), Report No. 1; Global Carbon Project Report No. 1, 69pp, Canberra, 2003.

Gobron, N. and Verstraete, M. M.: FAPAR, fraction of absorbed photosynthetically active radiation – Assessment of the status of the development of the standards for the terrestrial essential climate variables, Version 8, GTOS Secretariat, FAO, Italy, http://www.fao.org/gtos/doc/ECVs/T10/T10.pdf, 2009.

Gobron, N., Pinty, B., Aussedat, O., Chen, J. M., Cohen, W. B., Fensholt, R., Gond, V., Huemmrich, K. F., Lavergne, T., Mélin, F., Privette, J. L., Sandholt, I., Taberner, M., Turner, D. P., Verstraete, M. M., and Widlowski, J.-L.: Evaluation of fraction of absorbed photosynthetically active radiation products for different canopy radiation transfer regimes: Methodology and results using Joint Research Center products derived from SeaWiFS against ground-based estimations, Journal of Geophysical Research: Atmospheres, 111, doi:10.1029/2005JD006511, 2006.

Gobron, N., Pinty, B., Aussedat, O., Taberner, M., Faber, O., Melin, F., Lavergne, T., Robustelli, M., and Snoeij, P.: Uncertainty estimates for the FAPAR operational products derived from MERIS: Impact of top-of-atmosphere radiance uncertainties and validation with field data, Remote Sensing of Environment, 112, 1871 – 1883, doi:10.1016/j.rse.2007.09.011, remote Sensing Data Assimilation Special Issue, 2008.

Goel, N. S. and Qin, W.: Influences of canopy architecture on relationships between various vegetation indices and LAI and Fpar: A computer simulation, Remote Sensing Reviews, 10, 309–347, doi:10.1080/02757259409532252, 1994.

Gruber, A., Dorigo, W., Zwieback, S., Xaver, A., and Wagner, W.: Characterizing coarse-scale representativeness of in-situ soil moisture measurements from the International Soil Moisture Network, Vadose Zone Journal, 12, doi:10.2136/vzj2012.0170, 2013.

Gruber, A., Su, C., Zwieback, S., Crow, W. T., Wagner, W., and Dorigo, W.: Recent advances in (soil moisture) triple collocation analysis, International Journal of Applied Earth Observation and Geoinformation, 45, Part B, 200–211, 2016a.

Gruber, A., Su, C. H., Crow, W. T., Zwieback, S., Dorigo, W. A., and Wagner, W.: Estimating error cross-correlations in soil moisture data sets using extended collocation analysis, Journal of Geophysical Research: Atmospheres, 121, 1208–1219, doi:10.1002/2015JD024027, 2016b.

Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P. E., Gómez-Dans, J., Kuze, A., Suto, H., and Grainger, R. G.: Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements, Remote Sensing of Environment, 121, 236–251, 2012.

Guanter, L., Aben, I., Tol, P., Krijger, J. M., Hollstein, A., Köhler, P., Damm, A., Joiner, J., Frankenberg, C., and Landgraf, J.: Potential of the TROPOspheric Monitoring Instrument (TROPOMI) onboard the Sentinel-5 Precursor for the monitoring of terrestrial chlorophyll fluorescence, Atmospheric Measurement Techniques, 8, 1337–1352, doi:10.5194/amt-8-1337-2015, 2015.

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. A., Frankenberg, C., Huete, A. R., Zarco-Tejada, P., Lee, J.-E., Moran, M. S., Ponce-Campos, G., Beer, C., Camps-Valls, G., Buchmann, N., Gianelle, D., Klumpp, K., Cescatti, A., Baker, J. M., and Griffis, T. J.: Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence, Proceedings of the National Academy of Sciences, 111, E1327–E1333, 2014.

Heymann, J., Reuter, M., Hilker, M., Buchwitz, M., Schneising, O., Bovensmann, H., Burrows, J. P., Kuze, A., Suto, H., Deutscher, N. M., Dubey, M. K., Griffith, D. W. T., Hase, F., Kawakami, S., Kivi, R., Morino, I., Petri, C., Roehl, C., Schneider, M., Sherlock, V., Sussmann, R., Velazco, V. A., Warneke, T., and Wunch, D.: Consistent satellite XCO_2 retrievals from SCIAMACHY and GOSAT using the BESD algorithm, Atmospheric Measurement Techniques, 8, 2961–2980, doi:10.5194/amt-8-2961-2015, 2015.

Hollmann, R., Merchant, C. J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., de Leeuw, G., Forsberg, R., Holzer-Popp, T., Paul, F., Sandven, S., Sathyendranath, S., van Roozendael, M., and Wagner, W.: The ESA Climate Change Initiative: Satellite Data Records for Essential Climate Variables, Bulletin of the American Meteorological Society, 94, 1541–1552, doi:10.1175/BAMS-D-11-00254.1, 2013.

Houweling, S., Baker, D., Basu, S., Boesch, H., Butz, A., Chevallier, F., Deng, F., Dlugokencky, E. J., Feng, L., Ganshin, A., Hasekamp, O., Jones, D., Maksyutov, S., Marshall, J., Oda, T., O'Dell, C. W., Oshchepkov, S., Palmer, P. I., Peylin, P., Poussi, Z., Reum, F., Takagi, H., Yoshida, Y., and Zhuravlev, R.: An intercomparison of inverse models for estimating sources and sinks of CO_2 using GOSAT measurements, Journal of Geophysical Research: Atmospheres, 120, 5253–5266, doi:10.1002/2014JD022962, 2014JD022962, 2015.

Huete, A.: A soil-adjusted vegetation index (SAVI), Remote Sensing of Environment, 25, 295 – 309, doi:10.1016/0034-4257(88)90106-X, 1988.

Jackson, T.: Measuring surface soil moisture using passive microwave remote sensing., Hydrological Processes, 7, 139–152, 1993.

Joiner, J., Yoshida, Y., Vasilkov, A. P., Yoshida, Y., Corp, L. A., and Middleton, E. M.: First observations of global and seasonal terrestrial chlorophyll fluorescence from space, Biogeosciences, 8, 637–651, doi:10.5194/bg-8-637-2011, 2011.

Joiner, J., Yoshida, Y., Vasilkov, A. P., Middleton, E. M., Campbell, P. K. E., Yoshida, Y., Kuze, A., and Corp, L. A.: Filling-in of near-infrared solar lines by terrestrial fluorescence and other geophysical effects: simulations and space-based observations from SCIAMACHY and GOSAT, Atmospheric Measurement Techniques, 5, 809–829, doi:10.5194/amt-5-809-2012, 2012.

Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A. P., Middleton, E. M., Huemmrich, K. F., Yoshida, Y., and Frankenberg, C.: Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2, Atmospheric Measurement Techniques, 6, 2803–2823, 2013.

Joiner, J., Yoshida, Y., Guanter, L., and Middleton, E. M.: New methods for retrieval of chlorophyll red fluorescence from hyper-spectral satellite instruments: simulations and application to GOME-2 and SCIAMACHY, Atmospheric Measurement Techniques Discussions, 2016, 1–41, doi:10.5194/amt-2015-387, 2016.

Kaminski, T. and Mathieu, P.-P.: Reviews and Syntheses: Flying the Satellite into Your Model, Biogeosciences Discussions, 2016, 1–25, doi:10.5194/bg-2016-237, 2016.

Kaminski, T., Knorr, W., Rayner, P., and Heimann, M.: Assimilating Atmospheric data into a Terrestrial Biosphere Model: A case study of the seasonal cycle, Global Biogeochemical Cycles, 16, 14–1–14–16, http://www.agu.org/pubs/crossref/2002/2001GB001463.shtml, 2002.

Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P.: Consistent assimilation of MERIS FAPAR and atmospheric CO_2 into a terrestrial vegetation model and interactive mission benefit analysis, Biogeosciences, 9, 3173–3184, doi:10.5194/bg-9-3173-2012, 2012.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P. J., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J. P., Heimann, M., Hooker-Stroud, A., Houweling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E. N., Köstler, C., Mathieu, P.-P., Pinty, B., Reick, C. H., Rödenbeck, C., Schnur, R., Scipal, K., Sebald, C., Stacke, T., van Scheltinga, A. T., Vossbeck, M., Widmann, H., and Ziehn, T.: The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges, Journal of Geophysical Research: Biogeosciences, 118, 1414–1426, doi:10.1002/jgrg.20118, 2013.

Kaminski, T., Scholze, M., Vossbeck, M., Knorr, W., Buchwitz, M., and Reuter, M.: Constraining a terrestrial biosphere model with remotely sensed atmospheric carbon dioxide, "Remote Sensing of Environment", submitted, 2016.

Kaminski, T., Pinty, B., Voßbeck, M., Gobron, N., and Robustelli, M.: Consistent EO Land Surface Products including Uncertainty Estimates, Biogeosciences, 14, 2527–2541, doi:10.5194/bg-14-2527-2017, 2017.

Kato, T., Knorr, W., Scholze, M., Veenendaal, E., Kaminski, T., Kattge, J., and Gobron, N.: Simultaneous assimilation of satellite and eddy covariance data for improving terrestrial water and carbon simulations at a semi-arid woodland site in Botswana, Biogeosciences, 10, 789–802, doi:10.5194/bg-10-789-2013, 2013.

Kaufman, Y. J. and Tanre, D.: Atmospherically resistant vegetation index (ARVI) for EOS-MODIS, IEEE Transactions on Geoscience and Remote Sensing, 30, 261–270, doi:10.1109/36.134076, 1992.

Keeling, C. D.: The concentration and isotopic abundance of carbon dioxide in rural and marine air, Geochimica et Cosmochimica Acta, 24, 277–298, 1961.

Keenan, T. F., Davidson, E., Moffat, A. M., Munger, W., and Richardson, A. D.: Using model-data fusion to interpret past trends, and quantify uncertainties in future projections, of terrestrial ecosystem carbon cycling, Global Change Biology, 18, 2555–2569, doi:10.1111/j.1365-2486.2012.02684.x, 2012.

Kelley, D. I., Prentice, I. C., Harrison, S. P., Wang, H., Simard, M., Fisher, J. B., and Willis, K. O.: A comprehensive benchmarking system for evaluating global vegetation models, Biogeosciences, 10, 3313–3340, doi:10.5194/bg-10-3313-2013, 2013.

Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M. J., Font, J., Reul, N., Gruhier, C., Juglea, S. E., Drinkwater, M. R., Hahne, A., Martin-Neira, M., and Mecklenburg, S.: The SMOS Mission: New Tool for Monitoring Key Elements ofthe Global Water Cycle, Proceedings of the IEEE, 98, 666–687, doi:10.1109/JPROC.2010.2043032, 2010.

Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., Ferrazzoli, P., Mahmoodi, A., Bitar, A. A., Cabot, F., Gruhier, C., Juglea, S. E., Leroux, D., Mialon, A., and Delwart, S.: The SMOS Soil Moisture Retrieval Algorithm, IEEE Transactions on Geoscience and Remote Sensing, 50, 1384–1403, doi:10.1109/TGRS.2012.2184548, 2012.

Knorr, W. and Kattge, J.: Inversion of terrestrial biosphere model parameter values against eddy covariance measurements using Monte Carlo sampling, Global Change Biology, 11, 1333–1351, 2005.

Knorr, W., Kaminski, T., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P.: Carbon cycle data assimilation with a generic phenology model, Journal of Geophysical Research: Biogeosciences, 115, doi:10.1029/2009JG001119, g04017, 2010.

Köhler, P., Guanter, L., and Frankenberg, C.: Simplified Physically Based Retrieval of Sun-Induced Chlorophyll Fluorescence From GOSAT Data, Geoscience and Remote Sensing Letters, IEEE, 12, 1446–1450, doi:10.1109/LGRS.2015.2407051, 2015a.

Köhler, P., Guanter, L., and Joiner, J.: A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data, Atmospheric Measurement Techniques, 8, 2589–2608, doi:10.5194/amt-8-2589-2015, 2015b.

Konings, A. G., Piles, M., Rötzer, K., McColl, K. A., Chan, S. K., and Entekhabi, D.: Vegetation optical depth and scattering albedo retrieval using time series of dual-polarized L-band radiometer observations, Remote Sensing of Environment, 172, 178 – 189, doi:10.1016/j.rse.2015.11.009, 2016.

Kulawik, S., Wunch, D., O'Dell, C., Frankenberg, C., Reuter, M., Oda, T., Chevallier, F., Sherlock, V., Buchwitz, M., Osterman, G., Miller, C. E., Wennberg, P. O., Griffith, D., Morino, I., Dubey, M. K., Deutscher, N. M., Notholt, J., Hase, F., Warneke, T., Sussmann, R., Robinson, J., Strong, K., Schneider, M., DeÂ Mazière, M., Shiomi, K., Feist, D. G., Iraci, L. T., and Wolf, J.: Consistent evaluation of ACOS-GOSAT, BESD-SCIAMACHY, CarbonTracker, and MACC through comparisons to TCCON, Atmospheric Measurement Techniques, 9, 683–709, doi:10.5194/amt-9-683-2016, 2016.

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., and Richardson, A. D.: Constraining a global ecosystem model with multi-site eddy-covariance data, Biogeosciences, 9, 3757–3776, doi:10.5194/bg-9-3757-2012, 2012.

Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T.: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, Appl. Opt., 48, 6716–6733, doi:10.1364/AO.48.006716, 2009.

Kuze, A., Taylor, T. E., Kataoka, F., Bruegge, C. J., Crisp, D., Harada, M., Helmlinger, M., Inoue, M., Kawakami, S., Kikuchi, N., Mitomi, Y., Murooka, J., Naitoh, M., O'Brien, D. M., O'Dell, C. W., Ohyama, H., Pollock, H., Schwandner, F. M., Shiomi, K., Suto, H., Takeda, T., Tanaka, T., Urabe, T., Yokota, T., and Yoshida, Y.: Long-Term Vicarious Calibration of GOSAT Short-Wave Sensors: Techniques for Error Reduction and New Estimates of Radiometric Degradation Factors, IEEE Transactions on Geoscience and Remote Sensing, 52, 3991–4004, doi:10.1109/TGRS.2013.2278696, 2014.

Laeng, A., Plieninger, J., von Clarmann, T., Grabowski, U., Stiller, G., Eckert, E., Glatthor, N., Haenel, F., Kellmann, S., Kiefer, M., Linden, A., Lossow, S., Deaver, L., Engel, A., Hervig, M., Levin, I., McHugh, M., Noël, S., Toon, G., and Walker, K.: Validation of MIPAS IMK/IAA methane profiles, Atmospheric Measurement Techniques, 8, 5251–5261, doi:10.5194/amt-8-5251-2015, 2015.

Lannoy, G. J. M. D. and Reichle, R. H.: Global Assimilation of Multiangle and Multipolarization SMOS Brightness Temperature Observations into the GEOS-5 Catchment Land Surface Model for Soil Moisture Estimation, Journal of Hydrometeorology, 17, 669–691, doi:10.1175/JHM-D-15-0037.1, 2016.

Lasslop, G., Reichstein, M., Kattge, J., and Papale, D.: Influences of observation errors in eddy flux data on inverse model parameter estimation, Biogeosciences, 5, 1311–1324, doi:10.5194/bg-5-1311-2008, 2008.

Le Quéré, C., Moriarty, R., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., Friedlingstein, P., Peters, G. P., Andres, R. J., Boden, T. A., Houghton, R. A., House, J. I., Keeling, R. F., Tans, P., Arneth, A., Bakker, D. C. E., Barbero, L., Bopp, L., Chang, J., Chevallier, F., Chini, L. P., Ciais, P., Fader, M., Feely, R. A., Gkritzalis, T., Harris, I., Hauck, J., Ilyina, T., Jain, A. K., Kato, E., Kitidis, V., Klein Goldewijk, K., Koven,

C., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lima, I. D., Metzl, N., Millero, F., Munro, D. R., Murata, A., Nabel, J. E. M. S., Nakaoka, S., Nojiri, Y., O'Brien, K., Olsen, A., Ono, T., Pérez, F. F., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Rödenbeck, C., Saito, S., Schuster, U., Schwinger, J., Séférian, R., Steinhoff, T., Stocker, B. D., Sutton, A. J., Takahashi, T., Tilbrook, B., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Vandemark, D., Viovy, N., Wiltshire, A., Zaehle, S., and Zeng, N.: Global Carbon Budget 2015, Earth System Science Data, 7, 349–396, doi:10.5194/essd-7-349-2015, 2015.

Lee, J.-E., Frankenberg, C., van der Tol, C., Berry, J. A., Guanter, L., Boyce, C. K., Fisher, J. B., Morrow, E., Worden, J. R., Asefi, S., Badgley, G., and Saatchi, S.: Forest productivity and water stress in Amazonia: observations from GOSAT chlorophyll fluorescence, Proceedings of the Royal Society B: Biological Sciences, 280, doi:10.1098/rspb.2013.0171, 2013.

Lefsky, M. A.: A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System, Geophysical Research Letters, 37, n/a–n/a, doi:10.1029/2010GL043622, 2010.

Leprieur, C., Verstraete, M. M., and Pinty, B.: Evaluation of the performance of various vegetation indices to retrieve vegetation cover from AVHRR data, Remote Sensing Reviews, 10, 265–284, doi:10.1080/02757259409532250, 1994.

Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., and Sobrino, J. A.: Satellite-derived land surface temperature: Current status and perspectives, Remote Sensing of Environment, 131, 14 – 37, doi:10.1016/j.rse.2012.12.008, 2013.

Liu, Q., Liang, S., Xiao, Z., and Fang, H.: Retrieval of leaf area index using temporal, spectral, and angular information from multiple satellite data, Remote Sensing of Environment, 145, 25 – 37, doi:10.1016/j.rse.2014.01.021, 2014.

Liu, Y., Parinussa, R., Dorigo, W., De Jeu, R., Wagner, W., Van Dijk, A. I. J. M., McCabe, M., and Evans, J.: Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, Hydrology and Earth System Sciences, 15, 425–436, doi:10.5194/hess-15-425-2011, 2011.

Liu, Y., Dorigo, W., Parinussa, R., De Jeu, R., Wagner, W., McCabe, M., Evans, J., and Van Dijk, A. I. J. M.: Trend-preserving blending of passive and active microwave soil moisture retrievals, Remote Sensing of Environment, 123, 280–297, doi:10.1016/j.rse.2012.03.014, 2012.

Liu, Y. Y., van Dijk, A. I. J. M., de Jeu, R. A. M., Canadell, J. G., McCabe, M. F., Evans, J. P., and Wang, G.: Recent reversal in loss of global terrestrial biomass, Nature Climate Change, doi:10.1038/nclimate2581, 2015.

Luke, C. M.: Modelling aspects of land-atmosphere interation: Thermal instability in peatland soils and land parameter through data assimilation, Ph.D. thesis, University of Exeter, UK, 2011.

Luo, Y. Q., Randerson, J. T., Abramowitz, G., Bacour, C., Blyth, E., Carvalhais, N., Ciais, P., Dalmonech, D., Fisher, J. B., Fisher, R., Friedlingstein, P., Hibbard, K., Hoffman, F., Huntzinger, D., Jones, C. D., Koven, C., Lawrence, D., Li, D. J., Mahecha, M., Niu, S. L., Norby, R., Piao, S. L., Qi, X., Peylin, P., Prentice, I. C., Riley, W., Reichstein, M., Schwalm, C., Wang, Y. P., Xia, J. Y., Zaehle, S., and Zhou, X. H.: A framework for benchmarking land models, Biogeosciences, 9, 3857–3874, doi:10.5194/bg-9-3857-2012, 2012.

MacBean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F.-M., and Ciais, P.: Using satellite data to improve the leaf phenology of a global terrestrial biosphere model, Biogeosciences, 12, 7185–7208, doi:10.5194/bg-12-7185-2015, 2015.

MacBean, N., Peylin, P., Chevallier, F., Scholze, M., and Schürmann, G.: Consistent assimilation of multiple data streams in a carbon cycle data assimilation system, Geoscientific Model Development, 9, 3569–3588, doi:10.5194/gmd-9-3569-2016, 2016.

Martens, B., Miralles, D., Lievens, H., Van der Schalie, R., De Jeu, R., Fernandez-Prieto, D., Beck, H. E., Dorigo, W., and Verhoest, N. E. C.: GLEAM v3.0: satellite-based land evaporation and root-zone soil moisture, Geoscientific Model Development Discussion, 2016.

Mathieu, P.-P. and O'Neill, A.: Data assimilation: From photon counts to Earth System forecasts, Remote Sensing of Environment, 112, 1258 – 1267, doi:10.1016/j.rse.2007.02.040, 2008.

Matthews, H. D., Eby, M., Ewen, T., Friedlingstein, P., and Hawkins, B. J.: What determines the magnitude of carbon cycle-climate feedbacks?, Global Biogeochemical Cycles, 21, 12 PP, doi:10.1029/2006GB002733, gB2012, 2007.

McCallum, I., Wagner, W., Schmullius, C., Shvidenko, A., Obersteiner, M., Fritz, S., and Nilsson, S.: Comparison of four global FAPAR datasets over Northern Eurasia for the year 2000, Remote Sensing of Environment, 114, 941 – 949, doi:10.1016/j.rse.2009.12.009, 2010.

Minh, D. H. T., Toan, T. L., Rocca, F., Tebaldini, S., d'Alessandro, M. M., and Villard, L.: Relating P-Band Synthetic Aperture Radar Tomography to Tropical Forest Biomass, IEEE Transactions on Geoscience and Remote Sensing, 52, doi:10.1109/TGRS.2013.2246170, 2014.

Mitchard, E. T. A., Feldpausch, T. R., Brienen, R. J. W., Lopez-Gonzalez, G., Monteagudo, A., Baker, T. R., Lewis, S. L., Lloyd, J., Quesada, C. A., Gloor, M., ter Steege, H., Meir, P., Alvarez, E., Araujo-Murakami, A., Aragao, L. E. O. C., Arroyo, L., Aymard, G., Banki, O., Bonal, D., Brown, S., Brown, F. I., Ceron, C. E., Chama Moscoso, V., Chave, J., Comiskey, J. A., Cornejo, F., Corrales Medina, M., Da Costa, L., Costa, F. R. C., Di Fiore, A., Domingues, T. F., Erwin, T. L., Frederickson, T., Higuchi, N., Honorio Coronado, E. N., Killeen, T. J., Laurance, W. F., Levis, C., Magnusson, W. E., Marimon, B. S., Marimon Junior, B. H., Mendoza Polo, I., Mishra, P., Nascimento, M. T., Neill, D., Nunez Vargas, M. P., Palacios, W. A., Parada, A., Pardo Molina, G., Peña-Claros, M., Pitman, N., Peres, C. A., Poorter, L., Prieto, A., Ramirez-Angulo, H., Restrepo Correa, Z., Roopsind, A., Roucoux, K. H., Rudas, A., Salomao, R. P., Schietti, J., Silveira, M., de Souza, P. F., Steininger, M. K., Stropp, J., Terborgh, J., Thomas, R., Toledo, M., Torres-Lezama, A., van Andel, T. R., van der Heijden, G. M. F., Vieira, I. C. G., Vieira, S., Vilanova-Torre, E., Vos, V. A., Wang, O., Zartman, C. E., Malhi, Y., and Phillips, O. L.: Markedly divergent estimates of Amazon forest carbon density from ground plots and satellites, Global Ecology and Biogeography, 23, 935–946, doi:10.1111/geb.12168, 2014.

Moore, D. J., Hu, J., Sacks, W. J., Schimel, D. S., and Monson, R. K.: Estimating transpiration and the sensitivity of carbon uptake to water availability in a subalpine forest using a simple ecosystem process model informed by measured net CO_2 and H_2O fluxes, Agricultural and Forest Meteorology, 148, 1467 – 1477, doi:10.1016/j.agrformet.2008.04.013, 2008.

Muñoz, A. A., Barichivich, J., Christie, D. A., Dorigo, W., Sauchyn, D., González-Reyes, A., Villalba, R., Lara, A., Riquelme, N., and González, M. E.: Patterns and drivers of Araucaria araucana forest growth along a

biophysical gradient in the northern Patagonian Andes: Linking tree rings with satellite observations of soil moisture, Austral Ecology, 39, 158–169, doi:10.1111/aec.12054, 2014.

Myneni, R., Hoffman, S., Knyazikhin, Y., Privette, J., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G., Lotsch, A., Friedl, M., Morisette, J., Votava, P., Nemani, R., and Running, S.: Global products of vegetation leaf area and fraction absorbed {PAR} from year one of {MODIS} data, Remote Sensing of Environment, 83, 214 – 231, doi:10.1016/S0034-4257(02)00074-3, the Moderate Resolution Imaging Spectroradiometer (MODIS): a new generation of Land Surface Monitoring, 2002.

Naeimi, V., Scipal, K., Bartalis, Z., Hasenauer, S., and Wagner, W.: An Improved Soil Moisture Retrieval Algorithm for ERS and METOP Scatterometer Observations, IEEE Transactions on Geoscience and Remote Sensing, 47, 1999–2013, doi:10.1109/Tgrs.2009.2011617, 2009.

Noël, S., Bramstedt, K., Rozanov, A., Bovensmann, H., and Burrows, J. P.: Stratospheric methane profiles from SCIAMACHY solar occultation measurements derived with onion peeling DOAS, Atmospheric Measurement Techniques, 4, 2567–2577, doi:10.5194/amt-4-2567-2011, 2011.

Noël, S., Bramstedt, K., Hilker, M., Liebing, P., Plieninger, J., Reuter, M., Rozanov, A., Sioris, C. E., Bovensmann, H., and Burrows, J. P.: Stratospheric CH_4 and CO_2 profiles derived from SCIAMACHY solar occultation measurements, Atmospheric Measurement Techniques, 9, 1485–1503, doi:10.5194/amt-9-1485-2016, 2016.

Norton, A., Rayner, P. J., Scholze, M., and Koffi, E.: Global Gross Primary Productivity for 2015 inferred from OCO-2 SIF and a Carbon-Cycle Data Assimilation System, Abstract B53L-01 presented at 2016, Fall Meeting, AGU, San Francisco, CA, 12-16 December, 2016.

Ochsner, T., Cosh, M., Cuenca, R., Dorigo, W., Draper, C., Hagimoto, Y., Kerr, Y., Larson, K., Njoku, E., Small, E., and Zreda, M.: State of the art in large-scale soil moisture monitoring, Soil Science Society of America Journal, 77, 2013.

Owe, M., de Jeu, R., and Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, Journal of Geophysical Research-Earth Surface, 113, doi:10.1029/2007jf000769, 2008.

Parazoo, N. C., Bowman, K., Frankenberg, C., Lee, J.-E., Fisher, J. B., Worden, J., Jones, D. B. A., Berry, J., Collatz, G. J., Baker, I. T., Jung, M., Liu, J., Osterman, G., O'Dell, C., Sparks, A., Butz, A., Guerlet, S., Yoshida, Y., Chen, H., and Gerbig, C.: Interpreting seasonal changes in the carbon balance of southern Amazonia using measurements of XCO2 and chlorophyll fluorescence from GOSAT, Geophysical Research Letters, 40, 2829–2833, doi:10.1002/grl.50452, 2013.

Parinussa, R., Meesters, A., Liu, Y., Dorigo, W., Wagner, W., and De Jeu, R.: An analytical solution to estimate the error structure of a global soil moisture data set, IEEE Geoscience and Remote Sensing Letters, 8, 779–783, 2011.

Parker, R., Boesch, H., Cogan, A., Fraser, A., Feng, L., Palmer, P. I., Messerschmidt, J., Deutscher, N., Griffith, D. W. T., Notholt, J., Wennberg, P. O., and Wunch, D.: Methane observations from the Greenhouse Gases Observing SATellite: Comparison to ground-based TCCON data and model calculations, Geophysical Research Letters, 38, n/a–n/a, doi:10.1029/2011GL047871, l15807, 2011.

Pastorello, G. Z., Papale, D., Chu, H., Trotta, C., Agarwal, D. A., Canfora, E., Baldocchi, D. D., and Torn, M. S.: A new data set to keep a sharper eye on land-air exchanges, EOS, 98, doi:10.1029/2017EO071597, 2017.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO_2 inversions, Biogeosciences, 10, 6699–6720, doi:10.5194/bg-10-6699-2013, 2013.

Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P. J., Kuppel, S., Koffi, E. N., Kane, A., Maignan, F., Chevallier, F., Ciais, P., and Prunet, P.: A new step-wise Carbon Cycle Data Assimilation System using multiple data streams to constrain the simulated land surface carbon cycle, Geoscientific Model Development Discussions, 2016, 1–52, doi:10.5194/gmd-2016-13, 2016.

Pickett-Heaps, C. A., Canadell, J. G., Briggs, P. R., Gobron, N., Haverd, V., Paget, M. J., Pinty, B., and Raupach, M. R.: Evaluation of six satellite-derived Fraction of Absorbed Photosynthetic Active Radiation (FAPAR) products across the Australian continent, Remote Sensing of Environment, 140, 241 – 256, doi:10.1016/j.rse.2013.08.037, 2014.

Pinty, B. and Verstraete, M.: GEMI: A non-linear index to monitor global vegetation from satellites, Vegetatio, 101, 1335–1372, 1992.

Pinty, B., Leprieur, C., and Verstraete, M. M.: Towards a quantitative interpretation of vegetation indices Part 1: Biophysical canopy properties and classical indices, Remote Sensing Reviews, 7, 127–150, doi:10.1080/02757259309532171, 1993.

Pinty, B., Lavergne, T., Dickinson, R. E., Widlowski, J.-L., Gobron, N., and Verstraete, M. M.: Simplifying the Interaction of Land Surfaces with Radiation for Relating Remote Sensing Products to Climate Models, Journal of Geophysical Research – Atmospheres, 111, doi:10.1029/2005JD005952, 2006.

Pinty, B., Lavergne, T., Voßbeck, M., Kaminski, T., Aussedat, O., Giering, R., Gobron, N., Taberner, M., Verstraete, M. M., and Widlowski, J.-L.: Retrieving Surface Parameters for Climate Models from MODIS-MISR Albedo Products, Journal of Geophysical Research – Atmospheres, 112, 23– PP, doi:10.1029/2006JD008105, 2007.

Pinty, B., Lavergne, T., Kaminski, T., Aussedat, O., Giering, R., Gobron, N., Taberner, M., Verstraete, M. M., Voßbeck, M., and Widlowski, J.-L.: Partitioning the solar radiant fluxes in forest canopies in the presence of snow, Journal of Geophysical Research – Atmospheres, 113, 13– PP, doi:10.1029/2007JD009096, 2008.

Pinty, B., Lavergne, T., Widlowski, J.-L., Gobron, N., and Verstraete, M.: On the need to observe vegetation canopies in the near-infrared to estimate visible light absorption, Remote Sensing of Environment, 113, 10 – 23, doi:10.1016/j.rse.2008.08.017, 2009.

Pinty, B., Andredakis, I., Clerici, M., Kaminski, T., Taberner, M., Verstraete, M. M., Gobron, N., Plummer, S., and Widlowski, J.-L.: Exploiting the MODIS albedos with the Two-stream Inversion Package (JRC-TIP): 1. Effective leaf area index, vegetation, and soil properties, Journal of Geophysical Research – Atmospheres, 116, D09 105, doi:10.1029/2010JD015372, 2011a.

Pinty, B., Clerici, M., Andredakis, I., Kaminski, T., Taberner, M., Verstraete, M. M., Gobron, N., Plummer, S., and Widlowski, J.-L.: Exploiting the MODIS albedos with the Two-stream Inversion Package (JRC-TIP): 2. Fractions of transmitted and absorbed fluxes in the vegetation and soil layers, Journal of Geophysical Research – Atmospheres, 116, D09 106, doi:10.1029/2010JD015373, 2011b.

Pinty, B., Clerici, M., Andredakis, I., Kaminski, T., Taberner, M., Verstraete, M. M., Gobron, N., Plummer, S., and Widlowski, J.-L.: Exploiting the MODIS albedos with the Two-stream Inversion Package (JRC-TIP):

2. Fractions of transmitted and absorbed fluxes in the vegetation and soil layers, Journal of Geophysical Research – Atmospheres, 116, doi:10.1029/2010JD015373, 2011c.

Porcar-Castell, A., Tyystjärvi, E., Atherton, J., van der Tol, C., Flexas, J., Pfündel, E. E., Moreno, J., Frankenberg, C., and Berry, J. A.: Linking chlorophyll$-a$ fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges, Journal of Experimental Botany, doi:10.1093/jxb/eru191, 2014.

Post, H., Vrugt, J. A., Fox, A., Vereecken, H., and Hendricks Franssen, H.-J.: Estimation of Community Land Model parameters for an improved assessment of net carbon fluxes at European sites, Journal of Geophysical Research: Biogeosciences, 122, 661–689, doi:10.1002/2015JG003297, 2015JG003297, 2017.

Prentice, I. C., Liang, X., Medlyn, B. E., and Wang, Y.-P.: Reliable, robust and realistic: the three R's of next-generation land-surface modelling, Atmospheric Chemistry and Physics, 15, 5987–6005, doi:10.5194/acp-15-5987-2015, 2015.

Raj, R., Hamm, N. A. S., Tol, C. V. D., and Stein, A.: Uncertainty analysis of gross primary production partitioned from net ecosystem exchange measurements, Biogeosciences, 13, 1409–1422, doi:10.5194/bg-13-1409-2016, 2016.

Randerson, J. T., Hoffman, F. M., Thornton, P. E., Mahowlad, N. M., Lindsay, K., Lee, Y.-H., Nevison, C. D., Doney, S. C., Bonan, G., Stockli, R., Covey, C., Running, S. W., and Fung, I. Y.: Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models, Global Change Biology, 15, 2462–2484, doi:10.1111/j.1365-2486.2009.01912.x, 2009.

Raoult, N. M., Jupp, T. E., Cox, P. M., and Luke, C. M.: Land surface parameter optimisation through data assimilation: the adJULES system, Geoscientific Model Development Discussions, 2016, 1–26, doi:10.5194/gmd-2015-281, 2016.

Raupach, M. R., Rayner, P. J., Barrett, D. J., DeFries, R. S., Heimann, M., Ojima, D. S., Quegan, S., and Schmullius, C. C.: Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications, Global Change Biology, 11, 378–397, doi:10.1111/j.1365-2486.2005.00917.x, 2005.

Rayner, P., Scholze, M., Knorr, W., Kaminski, T., Giering, R., and Widmann, H.: Two decades of terrestrial Carbon fluxes from a Carbon Cycle Data Assimilation System (CCDAS), Global Biogeochemical Cycles, 19, 20 PP, doi:10.1029/2004GB002254, 2005.

Rayner, P., Michalak, A. M., and Chevallier, F.: Fundamentals of Data Assimilation, Geoscientific Model Development Discussions, 2016, 1–21, doi:10.5194/gmd-2016-148, 2016.

Reuter, M., Buchwitz, M., Schneising, O., Heymann, J., Bovensmann, H., and Burrows, J. P.: A method for improved SCIAMACHY CO_2 retrieval in the presence of optically thin clouds, Atmospheric Measurement Techniques, 3, 209–232, doi:10.5194/amt-3-209-2010, 2010.

Reuter, M., Bovensmann, H., Buchwitz, M., Burrows, J. P., Connor, B. J., Deutscher, N. M., Griffith, D. W. T., Heymann, J., Keppel-Aleks, G., Messerschmidt, J., Notholt, J., Petri, C., Robinson, J., Schneising, O., Sherlock, V., Velazco, V., Warneke, T., Wennberg, P. O., and Wunch, D.: Retrieval of atmospheric CO_2 with enhanced accuracy and precision from SCIAMACHY: Validation with FTS measurements and comparison with model results, Journal of Geophysical Research: Atmospheres, 116, n/a–n/a, doi:10.1029/2010JD015047, d04301, 2011.

Reuter, M., Bösch, H., Bovensmann, H., Bril, A., Buchwitz, M., Butz, A., Burrows, J. P., O'Dell, C. W., Guerlet, S., Hasekamp, O., Heymann, J., Kikuchi, N., Oshchepkov, S., Parker, R., Pfeifer, S., Schneising, O., Yokota,

T., and Yoshida, Y.: A joint effort to deliver satellite retrieved atmospheric CO_2 concentrations for surface flux inversions: the ensemble median algorithm EMMA, Atmospheric Chemistry and Physics, 13, 1771–1780, doi:10.5194/acp-13-1771-2013, 2013.

Reuter, M., Buchwitz, M., Hilker, M., Heymann, J., Schneising, O., Pillai, D., Bovensmann, H., Burrows, J. P., Bösch, H., Parker, R., Butz, A., Hasekamp, O., O'Dell, C. W., Yoshida, Y., Gerbig, C., Nehrkorn, T., Deutscher, N. M., Warneke, T., Notholt, J., Hase, F., Kivi, R., Sussmann, R., Machida, T., Matsueda, H., and Sawa, Y.: Satellite-inferred European carbon sink larger than expected, Atmospheric Chemistry and Physics, 14, 13 739–13 753, doi:10.5194/acp-14-13739-2014, 2014.

Reuter, M., Hilker, M., Schneising, O., Buchwitz, M., and Heymann, J.: ESA Climate Change Initiative (CCI) Comprehensive Error Characterisation Report: BESD full-physics retrieval algorithm for XCO2 for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), Version 2.0, http://www.esa-ghg-cci.org/webfm_send/284, 2016.

Ricciuto, D. M., Davis, K. J., and Keller, K.: A Bayesian calibration of a simple carbon cycle model: The role of observations in estimating and reducing uncertainty, Global Biogeochemical Cycles, 22, doi:10.1029/2006GB002908, gB2030, 2008.

Richardson, A. D., Mahecha, M. D., Falge, E., Kattge, J., Moffat, A. M., Papale, D., Reichstein, M., Stauch, V. J., Braswell, B. H., Churkina, G., Kruijt, B., and Hollinger, D. Y.: Statistical properties of random CO_2 flux measurement uncertainty inferred from model residuals, Agricultural and Forest Meteorology, 148, 38 – 50, doi:10.1016/j.agrformet.2007.09.001, 2008.

Richardson, A. D., Williams, M., Hollinger, D. Y., Moore, D. J. P., Dail, D. B., Davidson, E. A., Scott, N. A., Evans, R. S., Hughes, H., Lee, J. T., Rodrigues, C., and Savage, K.: Estimating parameters of a forest ecosystem C model with measurements of stocks and fluxes as joint constraints, Oecologia, 164, 25–40, doi:10.1007/s00442-010-1628-y, 2010.

Rodell, M., Velicogna, I., and Famiglietti, J. S.: Satellite-based estimates of groundwater depletion in India, Nature, 460, doi:10.1038/nature08238, 2009.

Rodriguez-Fernandez, N. J., Aires, F., Richaume, P., Kerr, Y. H., Prigent, C., Kolassa, J., Cabot, F., Jimenez, C., Mahmoodi, A., and Drusch, M.: Soil Moisture Retrieval Using Neural Networks: Application to SMOS, IEEE Transactions on Geoscience and Remote Sensing, 53, 5991–6007, doi:10.1109/TGRS.2015.2430845, 2015.

Rogers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Publishing, 2000.

Saatchi, S., Mascaro, J., Xu, L., Keller, M., Yang, Y., Duffy, P., Espirito-Santo, F., Baccini, A., Chambers, J., and Schimel, D.: Seeing the forest beyond the trees, Global Ecology and Biogeography, 24, 606–610, doi:10.1111/geb.12256, 2015.

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M., and Morel, A.: Benchmark map of forest carbon stocks in tropical regions across three continents, Proceedings of the National Academy of Sciences, 108, 9899–9904, doi:10.1073/pnas.1019576108, 2011.

Santoro, M., Beer, C., Cartus, O., Schmullius, C., Shvidenko, A., McCallum, I., Wegmüller, U., and Wiesmann, A.: Retrieval of growing stock volume in boreal forest using hyper-temporal series of Envisat ASA ScanSAR

backscatter measurements, Remote Sensing of Environment, 115, 490 – 507, doi:10.1016/j.rse.2010.09.018, 2011.

Santoro, M., Cartus, O., Fransson, J. E., Shvidenko, A., McCallum, I., Hall, R. J., Beaudoin, A., Beer, C., and Schmullius, C.: Estimates of Forest Growing Stock Volume for Sweden, Central Siberia, and Quebec using Envisat Advanced Synthetic Aperture Radar Backscatter Data, Remote Sensing, 5, 4503, doi:10.3390/rs5094503, 2013.

Santoro, M., Beaudoin, A., Beer, C., Cartus, O., Fransson, J. E., Hall, R. J., Pathe, C., Schmullius, C., Shepaschenko, D., Shvidenko, A., Thurner, M., and Wegmueller, U.: Forest growing stock volume of the northern hemisphere: Spatially explicit estimates for 2010 derived from Envisat ASAR, Remote Sensing of Environment, 168, 316 – 334, doi:10.1016/j.rse.2015.07.005, 2015.

Schneising, O., Buchwitz, M., Burrows, J. P., Bovensmann, H., Reuter, M., Notholt, J., Macatangay, R., and Warneke, T.: Three years of greenhouse gas column-averaged dry air mole fractions retrieved from satellite â€“ Part 1: Carbon dioxide, Atmospheric Chemistry and Physics, 8, 3827–3853, doi:10.5194/acp-8-3827-2008, 2008.

Schneising, O., Buchwitz, M., Burrows, J. P., Bovensmann, H., Bergamaschi, P., and Peters, W.: Three years of greenhouse gas column-averaged dry air mole fractions retrieved from satellite â€“ Part 2: Methane, Atmospheric Chemistry and Physics, 9, 443–465, doi:10.5194/acp-9-443-2009, 2009.

Schneising, O., Buchwitz, M., Reuter, M., Heymann, J., Bovensmann, H., and Burrows, J. P.: Long-term analysis of carbon dioxide and methane column-averaged mole fractions retrieved from SCIAMACHY, Atmospheric Chemistry and Physics, 11, 2863–2880, doi:10.5194/acp-11-2863-2011, 2011.

Schneising, O., Bergamaschi, P., Bovensmann, H., Buchwitz, M., Burrows, J. P., Deutscher, N. M., Griffith, D. W. T., Heymann, J., Macatangay, R., Messerschmidt, J., Notholt, J., Rettinger, M., Reuter, M., Sussmann, R., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D.: Atmospheric greenhouse gases retrieved from SCIAMACHY: comparison to ground-based FTS measurements and model results, Atmospheric Chemistry and Physics, 12, 1527–1540, doi:10.5194/acp-12-1527-2012, 2012.

Scholze, M., Kaminski, T., Rayner, P., Knorr, W., and Giering, R.: Propagating uncertainty through prognostic CCDAS simulations, Journal of Geophysical Research, 112, doi:10.1029/2007JD008642, 2007.

Scholze, M., Allen, I., Bill Collins, B., Cornell, S., Huntingford, C., Joshi, M., Lowe, J., Smith, R., Ridgwell, A., and Wild, O.: Understanding the Earth System - Global Change Science for Application, chap. 5 Earth System Models: a tool to understand changes in the Earth System, Cambridge University Press, Cambridge, UK, 2012.

Scholze, M., Kaminski, T., Knorr, W., Blessing, S., Vossbeck, M., Grant, J., and Scipal, K.: Simultaneous assimilation of {SMOS} soil moisture and atmospheric {CO2} in-situ observations to constrain the global terrestrial carbon cycle, Remote Sensing of Environment, 180, 334 – 345, doi:10.1016/j.rse.2016.02.058, special Issue: ESA's Soil Moisture and Ocean Salinity Mission - Achievements and Applications, 2016.

Schürmann, G. J., Kaminski, T., Köstler, C., Carvalhais, N., Voßbeck, M., Kattge, J., Giering, R., Rödenbeck, C., Heimann, M., and Zaehle, S.: Constraining a land surface model with multiple observations by application of the MPI-Carbon Cycle Data Assimilation System, Geoscientific Model Development Discussions, 2016, 1–24, doi:10.5194/gmd-2015-263, 2016.

Scipal, K., Holmes, T., de Jeu, R., Naeimi, V., and Wagner, W.: A possible solution for the problem of estimating the error structure of global soil moisture data sets, Geophysical Research Letters, 35, –, doi:10.1029/2008gl035599, 2008.

Tao, X., Liang, S., and Wang, D.: Assessment of five global satellite products of fraction of absorbed photosynthetically active radiation: Intercomparison and direct validation against ground-based data, Remote Sensing of Environment, 163, 270 – 285, doi:10.1016/j.rse.2015.03.025, 2015.

Tarantola, A.: Inverse Problem Theory and methods for model parameter estimation, SIAM, Philadelphia, 2005.

Thum, T., MacBean, N., Peylin, P., Bacour, C., Santaren, D., Longdoz, B., Loustau, D., and Ciais, P.: The potential benefit of using forest biomass data in addition to carbon and water fluxes measurements to constrain ecosystem model parameters: case studies at two temperate forest sites, Agricultural and Forest Meteorology, 234, 48 – 65, accepted, 2017.

Thurner, M., Beer, C., Santoro, M., Carvalhais, N., Wutzler, T., Schepaschenko, D., Shvidenko, A., Kompter, E., Ahrens, B., Levick, S. R., and Schmullius, C.: Carbon stock and density of northern boreal and temperate forests, Global Ecology and Biogeography, 23, 297–310, doi:10.1111/geb.12125, 2014.

Trudinger, C. M., Raupach, M. R., Rayner, P. J., Kattge, J., Liu, Q., Pak, B., Reichstein, M., Renzullo, L., Richardson, A. D., Roxburgh, S. H., Styles, J., Wang, Y. P., Briggs, P., Barrett, D., and Nikolova, S.: OptIC project: An intercomparison of optimization techniques for parameter estimation in terrestrial biogeochemical models, Journal of Geophysical Research: Biogeosciences, 112, n/a–n/a, doi:10.1029/2006JG000367, g02027, 2007.

van der Molen, M. K., de Jeu, R. A. M., Wagner, W., van der Velde, I. R., Kolari, P., Kurbatova, J., Varlagin, A., Maximov, T. C., Kononov, A. V., Ohta, T., Kotani, A., Krol, M. C., and Peters, W.: The effect of assimilating satellite-derived soil moisture data in SiBCASA on simulated carbon fluxes in Boreal Eurasia, Hydrol. Earth Syst. Sci., 20, 605–624, doi:10.5194/hess-20-605-2016, 2016.

Veefkind, J., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H., de Haan, J., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P.: {TROPOMI} on the {ESA} Sentinel-5 Precursor: A {GMES} mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sensing of Environment, 120, 70 – 83, doi:http://dx.doi.org/10.1016/j.rse.2011.09.027, the Sentinel Missions - New Opportunities for Science, 2012.

Villard, L. and Toan, T. L.: Relating P-Band SAR Intensity to Biomass for Tropical Dense Forests in Hilly Terrain: γ^0 or t^0?, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8, 214–223, doi:10.1109/JSTARS.2014.2359231, 2015.

Wagner, W., Lemoine, G., and Rott, H.: A method for estimating soil moisture from ERS scatterometer and soil data, Remote Sensing of Environment, 70, 191–207, 1999.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Z$'$uger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Bl$'$oschl, G., Eitzinger, J., and Steinnocher, K.: The ASCAT Soil Moisture Product: A Review of its Specifications, Validation Results, and Emerging Applications, Meteorologische Zeitschrift, 22, 5–33, doi:10.1127/0941-2948/2013/0399, 2013.

Walther, S., Voigt, M., Thum, T., Gonsamo, A., Zhang, Y., Koehler, P., Jung, M., Varlagin, A., and Guanter, L.: Satellite chlorophyll fluorescence measurements reveal large-scale decoupling of photosynthesis and greenness dynamics in boreal evergreen forests, Global Change Biology, doi:10.1111/gcb.13200, 2015.

Wang, Y. P., Leuning, R., Cleugh, H., and Coppin, P. A.: Parameter estimation in surface exchange models using non-linear inversion: How many parameters can w e estimate and which measurements are most useful?, Glob. Change Biol., 7, 495–510, 2001.

Widlowski, J.-L.: On the bias of instantaneous {FAPAR} estimates in open-canopy forests, Agricultural and Forest Meteorology, 150, 1501 – 1522, doi:j.agrformet.2010.07.011, 2010.

Williams, M., Schwarz, P. A., Law, B. E., Irvine, J., and Kurpius, M. R.: An improved analysis of forest carbon dynamics using data assimilation, Global Change Biology, 11, 89–105, doi:10.1111/j.1365-2486.2004.00891.x, 2005.

WMO: Greenhouse Gas Bulletin. The State of Greenhouse Gases in the Atmosphere Based on Global Observations through 2014., World Meteorological Organization, No. 11, 9 November, 2015.

Wolanin, A., Rozanov, V., Dinter, T., Noël, S., Vountas, M., Burrows, J., and Bracher, A.: Global retrieval of marine and terrestrial chlorophyll fluorescence at its red peak using hyperspectral top of atmosphere radiance measurements: Feasibility study and first results, Remote Sensing of Environment, 166, 243 – 261, doi:10.1016/j.rse.2015.05.018, 2015.

Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jiménez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M. A.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmospheric Measurement Techniques, 3, 1351–1362, doi:10.5194/amt-3-1351-2010, 2010.

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The Total Carbon Column Observing Network, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 369, 2087–2112, doi:10.1098/rsta.2010.0240, 2011.

Xiong, X., Barnet, C., Maddy, E., Wofsy, S., Chen, L., Karion, A., and Sweeney, C.: Detection of methane depletion associated with stratospheric intrusion by atmospheric infrared sounder (AIRS), Geophysical Research Letters, 40, 2455–2459, doi:10.1002/grl.50476, 2013.

Yoshida, Y., Kikuchi, N., Morino, I., Uchino, O., Oshchepkov, S., Bril, A., Saeki, T., Schutgens, N., Toon, G. C., Wunch, D., Roehl, C. M., Wennberg, P. O., Griffith, D. W. T., Deutscher, N. M., Warneke, T., Notholt, J., Robinson, J., Sherlock, V., Connor, B., Rettinger, M., Sussmann, R., Ahonen, P., Heikkinen, P., Kyrö, E., Mendonca, J., Strong, K., Hase, F., Dohe, S., and Yokota, T.: Improvement of the retrieval algorithm for GOSAT SWIR XCO_2 and XCH_4 and their validation using TCCON data, Atmospheric Measurement Techniques, 6, 1533–1547, doi:10.5194/amt-6-1533-2013, 2013.

Zwieback, S., Su, C.-H., Gruber, A., Dorigo, W. A., and Wagner, W.: The Impact of Quadratic Nonlinear Relations between Soil Moisture Products on Uncertainty Estimates from Triple Collocation Analysis and Two

Quadratic Extensions, Journal of Hydrometeorology, 17, 1725–1743, doi:doi:10.1175/JHM-D-15-0213.1, 2016.

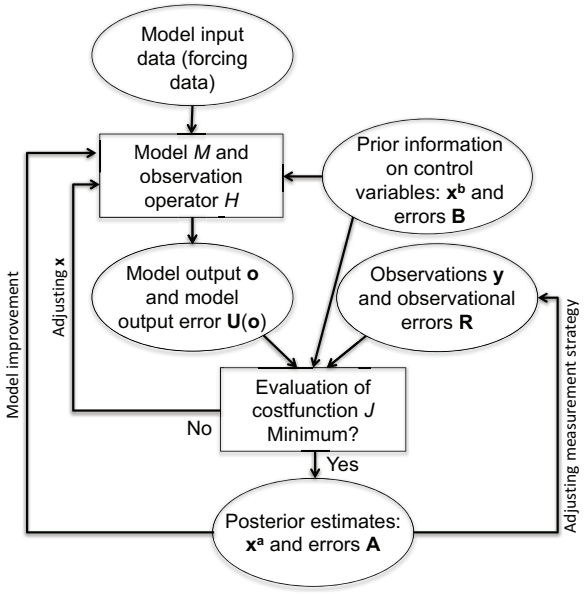

Figure 1. Schematic of a data assimilation system with **x** being the control vector containing the quantities to be updated by the assimilation. The ~~inner~~ loop ~~(~~between the ’~~Model-data comparison~~Evaluation of *J*’ box to ’Model and observation operator’ box ~~)~~ indicates the assimilation process (assimilation loop). Often, the analysis of residuals in ~~model data~~ model-data comparison ~~lead~~ leads to either model improvements or adjustment of the measurement strategies (~~outer loops~~’model improvement’ and ’adjusting measurement strategy’ arrows).

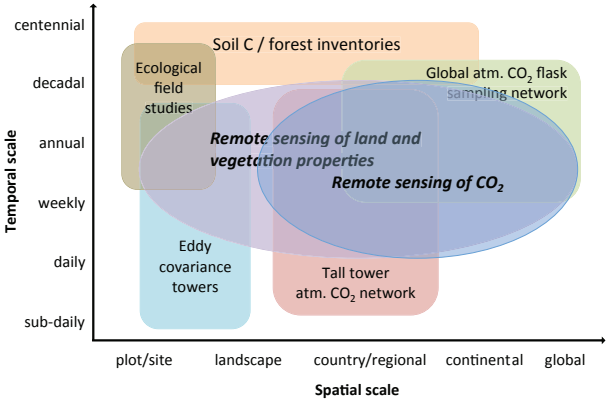

Figure 2. Space-time ~~diagramme~~ diagram for a range of observations relevant for a Terrestrial Carbon Observation System.

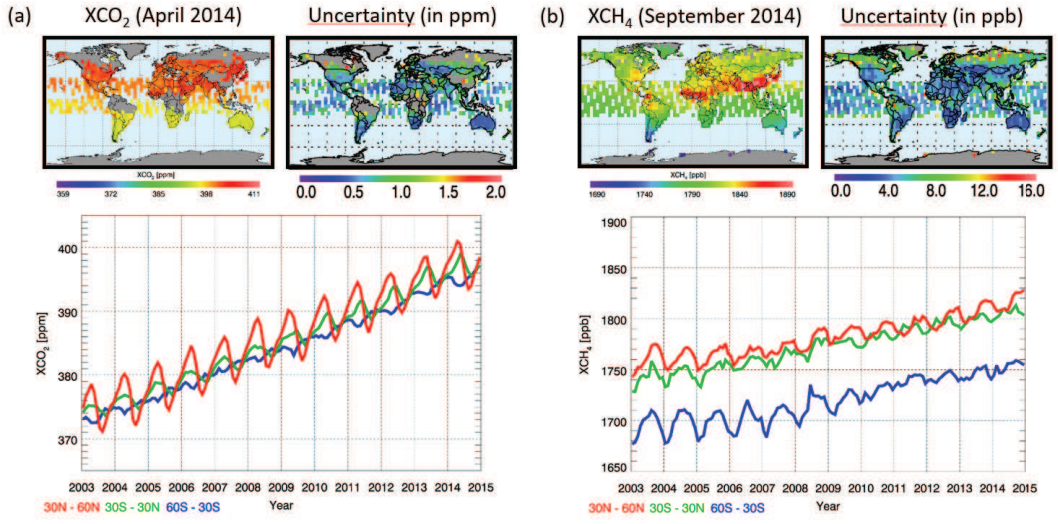

Figure 3. ~~Timeseries~~ Time series of satellite-derived ~~XCO2~~ XCO_2 in 3 latitude bands (see annotation bottom left, e.g. red line: ~~30o-60oN~~30 degree to 60 degree N) and maps showing the spatial distribution of ~~XCO2~~ XCO_2 for April 2014 (top left~~map~~) and corresponding ~~XCO2~~ XCO_2 uncertainty (~~bottom~~top right). (b) As (a) but for ~~XCH4~~ XCH_4 (maps: September 2014).

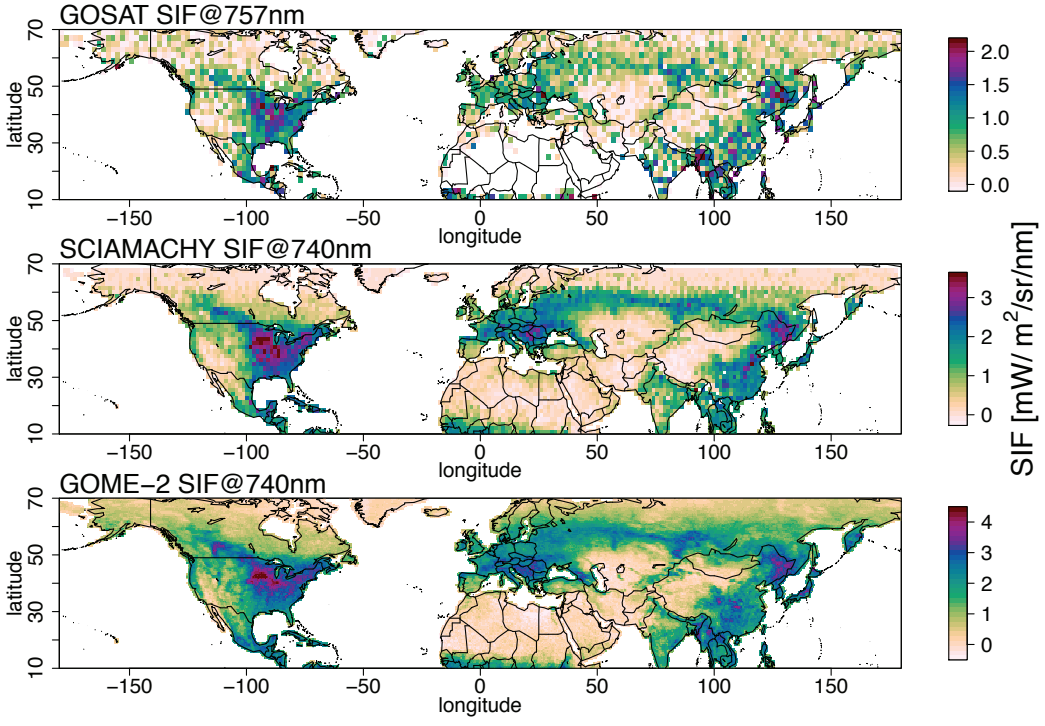

Figure 4. Maps of sun-induced fluorescence (SIF) for July 2010 derived from GOSAT, GOME-2 and SCIAMACHY satellite data.

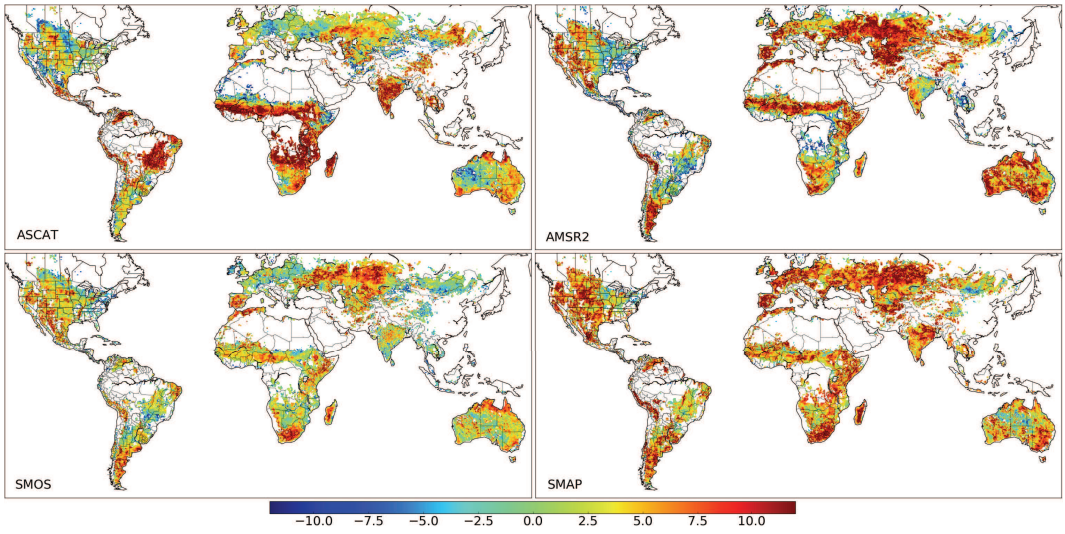

Figure 5. Signal-to-noise ratio (in dB), estimated with the Triple Collocation Analysis for four different satellite-based soil moisture products and a Land Surface Model. a) MetOp-A ASCAT based on the TU Wien method (Wagner et al., 1999); b) AMSR2 based on the LPRM model (Owe et al., 2008); c) SMOS L3 (Kerr et al., 2010); d) SMAP (Jackson, 1993). An SNR of -3 indicates a signal variance that is half of the noise variance, an SNR of 0 a signal variance equal to the noise variance, an SNR of 3 a signal variance that is twice the noise variance, and so on. In areas without data the TC could not be computed, e.g. because of too few observations in one of the ~~datasets~~data sets. For details see Gruber et al. (2016b).

Table 1. Overview SCIAMACHY/ENVISAT and TANSO-FTS/GOSAT ~~XCO2~~ XCO_2 and ~~XCH4~~ XCH_4 Level 2 data products (individual ground-pixel retrievals). For some products also Level 3, i.e. gridded data products are available (e.g. for CO2_SCI_WFMD and CH4_SCI_WFMD from http://www.iup.unibremen.de/sciamachy/NIR_NADIR_WFM_DOAS/ and merged SCIAMACHY and TANSO-FTS ~~XCO2~~ XCO_2 and ~~XCH4~~ XCH_4 products in Obs4MIPs format from http://www.esa-ghg-cci.org/)

~~Parameter~~ Variable	Sensor	Available at: Product (Reference)
~~XCO2~~ XCO_2	SCIAMACHY	http://www.esa-ghg-cci.org/
		CO2_SCI_BESD (Reuter et al., 2011)
		CH4_SCI_WFMD (Schneising et al., 2011)
	TANSO	http://www.gosat.nies.go.jp/en/
		NIES operational GOSAT (Yoshida et al., 2013)
		http://www.esa-ghg-cci.org/
		CO2_GOS_OCFP (Cogan et al., 2012)
		CO2_GOS_SRFP/RemoTeC (Butz et al., 2011)
		http://www.iup.uni-bremen.de/ heymann/besd_gosat.php
		GOSAT BESD (Heymann et al., 2015)
		http://disc.sci.gsfc.nasa.gov/acdisc/documentation/ACOS.html
		NASA ACOS (Crisp et al., 2012)
	SCIAMACHY &	http://www.esa-ghg-cci.org/
	TANSO merged	CO2_EMMA (Reuter et al., 2013)
	OCO-2	http://disc.sci.gsfc.nasa.gov/OCO-2
		NASA OCO-2 (Boesch et al., 2011)
~~XCH4~~ XCH_4	SCIAMACHY	http://www.esa-ghg-cci.org/
		CH4_SCI_WFMD (Schneising et al., 2011)
		CH4_SCI_IMAP (Frankenberg et al., 2011a)
	TANSO	http://www.gosat.nies.go.jp/en/
		NIES operational GOSAT (Yoshida et al., 2013)
		http://www.esa-ghg-cci.org/
		CH4_GOS_OCPR (Parker et al., 2011)
		CH4_GOS_SRPR/RemoTeC (Butz et al., 2010)
		CH4_GOS_OCFP (Parker et al., 2011)
		CH4_GOS_SRFP/RemoTeC (Butz et al., 2011)
	SCIAMCHY &	http://www.esa-ghg-cci.org/
	TANSO merged	CH4_EMMA (Reuter et al., 2013)

Table 2. Characteristics of a variety of FAPAR products, more details and products are provided by D'Odorico et al. (e.g. 2014); Pickett-Heaps et al. (e.g. 2014).

Name	Time period	Temporal resolution	Definition	Reference
MODIS	2000-present	8 days	Green canopy, direct radiation	Myneni et al. (2002)
SeaWiFS[1]	1997-2006	10 days	Green canopy, diffuse radiation	Gobron et al. (2006)
TIP-MODIS	2000-present	16 days	FAPAR/Green canopy, diffuse radiation	Pinty et al. (2011b)
TIP-GlobAlbedo	2002-2011	8 days	FAPAR/Green canopy, diffuse radiation	~~Disney et al. (2016b)~~ Disney et al. (2016a)
Vegetation	1999-present	10 days	FAPAR, direct radiation	Baret et al. (2007)

1 The same algorithm is also used for MERIS (called JRC MGVI, 2002-2011) and SPOT-Vegetation (2012-present) with a 1.2 km, 10 day resolution (Gobron et al., 2008).

Table 3. Selected characteristics of operating and planned spaceborne instruments able to deliver SIF data. Names of upcoming instruments are highlighted in italics. NIR stands for near-infrared. It must be noted that GOME-2 on MetOp-A has been operating with a reduced pixel size of 40×40 km^2 since July 2013. References are examples, the full list is given in the text.

Name	Time period	Overpass time	Spectral sampling	~~Global~~ Spatial ~~coverage~~ resolution	Temporal resolution	Reference
GOSAT	2009–today	Midday	NIR	~~No~~ 10 km diam.	3 days	e.g. Frankenberg et al. (2011c)
GOME-2	2007–today	Morning	red & NIR	~~Yes~~ 40×80 km^2	$<$2 days	e.g. Joiner et al. (2013), Köhler et al. (2015b)
SCIAMACHY	2003–2012	Morning	red & NIR	~~Yes~~ 30×240 km^2	$<$3 days	e.g. Joiner et al. (2013), Köhler et al. (2015b)
OCO-2	2014–today	Midday	NIR	~~No~~ 1.3×2.3 km^2	16 days	Frankenberg et al. (2014)
TROPOMI	~2017	Midday	red & NIR	~~Yes~~ 7×7 km^2	$<$1 day	Guanter et al. (2015)
FLEX	~2022	Morning	red & NIR	~~Yes~~ 0.3×0.3 km^2	$<$27 days	Drusch et al. (2017)

Table 4. Current (pre-)operational global soil moisture missions and products (for abbreviations / acronyms see List of Acronyms

Mission	Organisation	Measurement concept	Band	Mission start	Data access
MetOp - ASCAT	EUMETSAT	Real aperture radar (scatterometer)	C-band	Jan. 2007	http://hsaf.meteoam.it/soil-moisture.php http://land.copernicus.eu/global/products/swi
SMOS	ESA	Interferometric radiometer	L-band	Nov. 2009	http://www.catds.fr/
GCOM-W1 AMSR2	JAXA	Radiometer	C-band	May 2012	http://www.vandersat.com/ http://suzaku.eorc.jaxa.jp/GCOM_W/
SMAP	NASA	Radiometer & radar[1]	L-band	Jan. 2015	http://smap.jpl.nasa.gov/
Sentinel-1	ESA/ Copernicus	Synthetic aperture radar	C-band	Apr. 2014	https://www.eodc.eu/
CCI	ESA	Combined scatterometer and radiometer	L-, C-, X-Ku-band	Nov. 1978	http://www.esa-soilmoisture-cci.org

1 SMAP's radar failed in July 2015