The paper has been quite extensively revised and polished following both recommendations from the reviewers and also from further developments and input from the authors. Foremost, we agreed with Reviewer Anthes about the omissions related to GPS RO and we invited Rick to become an author. He accepted and some major revisions resulted, including two new figures and new text. He also had input on other aspects, including the new NRC report (mid-course assessment of the Decadal survey) that is now discussed in section 5. Fairly major revisions were also undertaken to improve the flow and maintain focus, so that one figure was removed and the associated science was also removed. This was a direct response to the reviewer A.

I am pleased to note that all authors have read the revised version end-to-end and all make very favorable comments. They have all resulted in several minor changes in wording throughout, so that the resulting manuscript is now mature and polished. The biggest outstanding issue is the reference to Wiffjels et al. OSC paper???

The extent of the changes are sufficient that a tracked change version is a mess, but I can provide that if requested.

Responses to comments in red.

**Reviewer A**

**Review of “Challenges of a Sustained Climate Observing System”**

Trenberth, Belward, Brown, Haberman, Karl, Running, Ryan, Tanner, and Wielicki

**General Comments:**

This paper provides an overview of the current climate observing system and the challenges involved in maintaining and improving it. It describes how requirements for climate monitoring differ from what is needed for weather prediction, both from the satellite and in situ perspectives. The paper summarizes the recommendations from earlier GCOS reports as well as those from the 2009 Third World Climate Conference, and defines what an “operational” system for monitoring Earth’s climate should look like. A summary of current and planned satellite and in-situ observations is provided, and the challenges associated with satellite accuracy and transitions between observation systems are discussed. Also included is a discussion of the role of assessments and reprocessing both for satellite-derived ECVs and reanalysis. Finally, the paper summarizes the improvements needed in both in situ and satellite observations, as well as data stewardship (documentation and metadata).

The major conclusion of the paper is that while the need for climate information has increased, the effort required to meet the demand has not. The authors note that improvements in data accuracy, independence, continuity, and prioritization within the observation system are needed. It points out that while the challenges are enormous, meeting them is feasible with international cooperation.

The paper provides a nice summary of the current state of the climate system and its shortcomings. However, I got the impression the authors were somewhat tepid in their recommendation of the need to improve the climate observing system. They did not provide a great sense of the urgency with statements like, “much more remains to be done before a fully functional climate observing system exists” in the abstract. I couldn’t tell if we’re 25% or 75% of the way toward a fully functional climate observing system. I would recommend that the authors more forcefully state why it is important to improve the climate observing system and provide a sense of the urgency required. The discussion of $ amounts provides an indication of this. This is now emphasized a bit more but we did not want to expand the text much further.

 Another weakness with the paper is that some of the terms used may be confusing to readers unfamiliar with jargon commonly used in GCOS-style documents. For example, the term “Operational” is often used in the context of climate monitoring. Operational is often thought of as a routine effort with a quick turnaround (e.g., transportation, weather, etc.). To produce climate-quality data products via the model shown in Fig. 3 as an example, it is not obvious that the “operational” model really fits. As problems arise (e.g., in the “instrument calibration” and/or “quality control of data products” steps), it is not acceptable to simply ignore them and release bad data. The problems need to be studied and solutions need to be found before proceeding to the “dissemination and access” step. This is true even if the algorithms used are “mature”. Perhaps what the authors really mean is “sustained” instead of “operational”? Or maybe the authors are defining an “operational” system to serve the needs of the climate applications community that requires data products on a routine basis, at the expense of high accuracy and stability? Such a system does not serve the climate research community very well, as they require the best available accuracy and stability (e.g., to study decadal variability). It would help the reader tremendously if the authors clearly defined such terms and identified what part of the climate user community (e.g., research or applications) they refer to. In the first part “operational” refers to weather and related networks. The term is highlighted in heading 2.3. Here operational climate is introduced and refers to the developing Global Framework for Climate Services and related routine climate services. Operational really means routine and with an institutional commitment as opposed to PI-driven or one-off team instruments. Needs for interannual climate variability (ENSO etc) can tolerate greater noise in the system owing to bigger signal. The text has been revised.

Finally, it is clear that several authors contributed to this paper, as some sections diverge aimlessly into different directions, and are only loosely connected with the rest of the paper. I recommend the authors do some editing to improve the readability of the paper. I’ve pointed out some of these in the Specific Comments section. This seems to relate mainly to the carbon and land surface section, and substantial changes were made by removing one figure and associated text and improving the rationale for the rest and improving the flow.

These were very helpful comments.

**Specific Comments:**

Abstract:

“A major challenge is to adequately deal with the continually changing observing system, especially from satellites and other autonomous platforms such as in the ocean.”

It might not be so obvious to some why a “continually changing observing system” is a challenge. For example, some change, such as instrument improvements, might actually be a positive thing (provided these don’t cause a discontinuity in the record…). I recommend being more explicit about what exactly the challenge is. Detail not appropriate in abstract: small changes made.

p.2: “Mitigation of the human influences, such as by cutting greenhouse gas and aerosol emissions, is a major challenge and the effectiveness of mitigation actions must be documented in order for them to continue.”

This sentence seems odd given how little mitigation is presently occurring. I thought that greenhouse gas emissions during the past few years have exceeded even the worst-case scenario cited in the IPCC Fourth Assessment Report. Consider revising this sentence. reworded

p. 5: “The 2010 update provided cost estimates for fully implementing and operating the climate observing system; around US$2.5 billion each year (in addition to the current annual global expenditure of some US$5-7 billion on global observing systems serving climate and related purposes). Around US$1.4 billion of this additional expenditure is needed for satellites or for *in situ* observation of the open ocean, in both cases for the benefit of all. In addition around US$600 million per year are needed for *in situ* observations in developing countries.”

It appears odd to provide such a detailed breakdown of how much more money is needed to fully implement and operate the climate observing system ($2.5 billion dollars with $1.4 billion for satellites/in situ + $600 million for in situ in developing countries) when the current annual global expenditure is known so poorly ($5-7 billion). The latter makes one wonder how realistic the estimates really are. I would recommend either coming up with a more accurate estimate of the current cost or not show the current expenditure at all, but still provide an estimate of what additional funding is necessary based upon what currently isn’t being done that should.

We disagree with this comment. Here is where the relative investment is addressed. We don’t know the carbon cycle that well either, but the perturbations are what matter. We have made clear (if not already) that these are not our estimates but they come from GCOS. Some small changes made.

p. 7: “Developing operational components”

Please define what is meant by “operational”. One can think of operational weather, operational transportation systems (FAA, TSA, metro/bus system), operational military activities, etc., all of which have a very specific short-term repetitive goal, such as delivering a weather forecast at a certain time each day, ensuring the safety of departing/arriving flights at an airport, making sure metro/buses depart/arrive on time, etc.. As noted in other parts of the paper, measurement accuracy requirements are many times more stringent for climate than weather. Satellite instruments require sophisticated onboard calibration sources that, when combined with other independent information (e.g., satellite intercomparisons, vicarious calibration, reassessment of ground calibration data, etc.), can be used as input to retrieval algorithms that output climate data records of many of the ECVs discussed in this paper. Anyone who’s ever been involved in a satellite program knows that every new instrument is like a new child, with it’s own unique set of issues and problems that require individualized and often creative solutions. This is even true of instruments having the same design. As problems get resolved, there is also a need to periodically reprocess the dataset in order to remove spurious artifacts in the record.

Such a paradigm is hardly an “operational (turn-the-crank)” activity. In this context, therefore, the term “Operational climate” is an oxymoron, which I believe has ultimately led to some very bad decisions (e.g., the notion of “research-to-operations”). Perhaps what is meant by “operational” is really “sustained” or “ongoing”? If so, the authors should clarify this. See above: substantial changes made.

p. 10: Figure 4: Why is Terra missing from this figure? Aqua and Aura can be considered as being part of the A-Train, but Terra is not.

The figure is modified to add Terra: note Cosmic is now there too.

p. 11: “The Earth Observing System (EOS) platforms are currently scheduled to operate through about 2015.”

I wonder where the authors got this? As I understand it, the EOS missions are reviewed every two years through the NASA Senior Review process, which decides whether or not to continue each mission on a case-by-case basis. As long as a mission’s spacecraft and instruments are healthy enough and continue to provide science value, and as long as the science is deemed important enough to continue, a mission typically gets extended for another 2 years. There is no a priori plan to scrap EOS in 2015 that I know of. In fact, there is sufficient fuel onboard Terra to last through 2017 and for Aqua to make it to 2020 (Aura might be similar too, but I’m not sure). Please remove or modify this statement to reflect the actual situation. Small edits made.

p. 12: “However, it (NPP) will have a ‘quasi-operational’ mandate…”. This will mean nothing to most readers. It’s bad enough that the term “operational” is used, but now you use “quasi-operational”? Do you mean that it will have a weather mandate (because JPSS is mainly about weather, not climate)? Please clarify and reword. Several minor changes made.

p. 12: “The NPP platform contains the VIIRS sensor as the successor…”

This paragraph makes no mention of TSIS, APS or CERES, despite the fact that CERES is the only climate-dedicated instrument on NPP and JPSS-1 (the others are primarily for weather, but obviously VIIRS and CriS are useful for climate also). The failure of Glory means: (i) there is high probability of a gap in the solar irradiance record at the level of accuracy provided by SORCE-TIM; and (ii) there are no plans for a dedicated aerosol instrument like APS, despite the significant uncertainty in climate forcing associated with aerosols. Would any of these instruments even be considered if an additional $2.5 billion were added to the global observing system? This aspect is bolstered and the suggestions are included.

p. 12: “The other relevant land science sensor will be the SMAP (Soil Moisture-Active/Passive) planned for an early 2015 launch, which will monitor surface wetness and freeze/thaw conditions of the land surface.”

Presumably, SMAP is a 5-year mission? If so, then the term “monitor” probably isn’t appropriate. Consider rewording. We added SMOS also, and made small changes.

p. 12: “The overall impact remains to be seen, but it is becoming clear that there is a significant probability of a lack of overlap between the EOS platforms and the next generation operational system (JPSS).”

This paragraph is baffling to me. The previous paragraph correctly points out that NPP launched in October 2011, but this paragraph ignores NPP altogether and claims there is likely to be a gap between the EOS platforms and JPSS. Why is this such a concern given that NPP is now up? Is the concern that a gap will occur between NPP and JPSS-1? What is the actual risk of a gap with EOS–NPP–JPSS-1, assuming JPSS-1 launches in 2017 (the official launch date)? Please revise this paragraph. Some rewrite performed.

p. 12: “Other networks are of vital importance to understanding the physical climate system, including observations of the Earth radiation budget, temperature, greenhouse gases, leaf area index, land cover, albedo, precipitation,…”.

Should “radiation budget” and “albedo” be preceded with “surface”, since this section discusses in-situ observations? minor rewrite made

p. 12-13 General comment about Section 3.2 “Adequacy of in-situ observations”:

With such a provocative section title, I expected some comment on whether or not the authors think the current state of in-situ observations is “adequate”. Instead, the section provides a broad summary of in-situ observations and goes into some detail about “reference observation requirements”. I did not get a sense of the adequacy of in-situ observations. An adequacy comment is now made with cross ref to 2.1 and 5.1

p. 14: “The decadal change forcing and response observations drive the need for very high accuracy and large time/space scales.”

This reads as if there is no need for higher spatial/temporal resolution observations. One can argue that accuracy and stability requirements should be set using large space-time scale metrics (e.g., natural variability), but we shouldn’t stop there. There is some value in resolving variability at finer spatial and temporal scales so that robust spatio-temporal patterns of change can also be resolved. Please revise. Added

p. 15 Section b “Delays and costs”

Starting with the second paragraph, this section appears to drift off into a discussion of the carbon budget and of how Net Primary Production has varied since 1982. It is not clear what point is being made in this section and why the narrow focus on carbon cycle. Surely there are other parts of the observing system that are affected by delays and costs? Please revise this section or remove altogether. We agree and this section was removed. The rest was adjusted as well.

p. 18: “A second key issue is the stability over decades of satellite geophysical retrieval algorithms which all have bias errors larger than decadal climate change signals. Current climate studies assume that these biases remain sufficiently stable to cancel out in observing decadal change anomalies, an assumption that should be verified.”

It is unclear what is meant by “stability over decades of satellite geophysical retrieval algorithms”. If one uses the same algorithm over multiple decades, how can you claim the algorithm isn’t stable? Perhaps the problem is that the retrieval algorithms require ancillary inputs (e.g., reanalysis data), which change with time due to changes in the observing system? Please clarify. Reworded

p. 18: “Another possibility to limit sensitivity to retrieval biases is the use of reflected solar and infrared spectral fingerprinting studies of climate change (Huang et al. 2010; Feldman et al. 2011; Jin et al. 2011). These climate Observing System Simulation Experiment (OSSE) studies have shown that infrared and solar reflected spectral fingerprints are very linear at the large time/space scales relevant to decadal climate change, unlike their highly nonlinear behavior for instantaneous retrievals.”

The fingerprinting approach needs to be placed in proper context. First, it appears to do away with the notion of ECVs, discussed earlier in the paper. Second, it actually does require “retrieval algorithms” that are more complex than traditional retrieval algorithms in order to determine the physical reason for changes in reflected solar and infrared spectra. Thus, while it is worth pursuing, it too has challenges. For some reason, the text appears to emphasize the shortcomings of traditional retrieval methods, but glosses over the challenges associated with the fingerprinting approach. A more balanced description would be more effective for readers unfamiliar with this area. We disagree: no change.

p. 26-27 Section 5.2 Data documentation and adequacy of metadata.

This section needs some work. The bullets after the first paragraph are especially weak (see below). I am surprised there is no mention of a need to have peer-reviewed documentation in internationally recognized journals. Is this not desirable? Reworded and polished.

p. 26: “Rapid evolution of the global climate fundamentally changes requirements for understanding temporal variations in observed properties”.

This sentence is unclear. Why would the evolution of the global climate change our observing requirements? Perhaps it is the evolution of our understanding about global climate change that leads to changes in observing requirements? Please clarify. reworded

p. 26: “The changing environment increases the importance of older observations…”.

What does “the changing environment” mean? Is this paragraph simply trying to make the case that in order to ensure observational datasets remain useful long after the scientists responsible for them are no longer around (e.g., retired), documentation is needed? Please clarify. reworded

p. 26: Please correct the grammatical mistakes. Also, isn’t this bullet saying basically the same thing as some of the earlier bullets. Two parts were combined and consolidated.

p. 27 Section 5.3 “Tracking climate observing performance”. This section argues for the need to assess independent observations and analyses of related quantities, and presents Fig. 10 as an example. While such comparisons are important, they are insufficient. Too often, assessments of data products simply document differences amongst datasets, and omit to provide any insight as to why there are differences, whether the differences are within or exceed the stated uncertainties of the individual data products, or recommend areas where further study is needed. Modelers will then take the differences amongst the various datasets as the “observational uncertainty”.

Another recent trend is the comparison of “ensembles” of observations (i.e., from different datasets) with “ensembles” of model output (e.g., from multiple runs of the same model or from the output of different models). While this can be quite useful, one problem is that the uncertainty of some of the individual observational datasets might in fact be much smaller than the spread of the observational ensemble, providing a much weaker constraint on the model output. This section would be an appropriate place to comment on what constitutes a useful assessment of observations, and could also provide some insight on the appropriateness of comparing observational ensembles with model output. We have a major activity on this at NCAR, a short para added.

**Reviewer B**

Sorry for the late (and sketchy) comments on the manuscript. It is a well-written and thorough analysis of a long-standing problem with climate observing systems, especially those based on satellite remote sensing. And although the issues are long-standing, there is little that is new. However, my primary concern is that goals of a climate observing system continue to be driven largely by the needs of large-scale climate science, rather than the needs of decision makers who will need to adapt to climate change. The paper's point of view leads to an emphasis on:

-- global-scale observations on multi-decadal time scales

-- detection and attribution of climate change on these time/space scales

-- collections of data to support climate science, which is then delivered to grateful users (see Fig. 3)

-- predominant focus on satellite missions

-- tendency to focus on measurements that increase science understanding, rather than monitoring and continuity (e.g., CLARREO).

There is nothing wrong with this directional emphasis, as clearly we do need to improve our understanding of the climate system. However, it is increasingly apparent that we need a more balanced approach that focuses on time scales of 5-7 years and regional space scales. This is where the adaptive planning decisions will be made. Unfortunately, these scales fall between the weather forecasting and climate observing scales, as we presently define them. I think if we were to do such an analysis, both the variables and the techniques would change. And we would likely have more emphasis on in situ observations, especially in the hydrologic and biological realms. It would also start to bring social science data, such as measures of urbanization and water use. These elements are called out in the recent draft strategic plan for US GCRP. I think we would also explore notions of calibration and validation from a different perspective.

My suggestion would be to retain the paper as is, but it would be useful to add a section on what a climate observing system would look like from this regional-scale, 5-7 year decision support perspective. Also, the NRC Mid-term Assessment of the NASA Decadal Survey should be out soon. I think the authors should consult the report and update their material about NASA's mission plans. Hope these thoughts are helpful!

We can not add much. This comment relates to the Rev 1 comment on operational and climate services, and also interannual as opposed to decadal and longer term climate. The new mid-term assessment is now discussed in section 5. Reworded.