Minimizing Reanalysis Jumps Due to New Observing Systems

Weather-Chaos Group Meeting
11/05/2012
by Yan Zhou

Advisors: Eugenia Kalnay
Junye Chen
I. Motivation

Fig 1. Typical global 6-h analysis cycle (Kalnay, 2002)

- Data assimilation is the process by which observations are incorporated into a numerical weather/climate model.
I. Motivation

Data assimilation is the process by which observations are incorporated into a numerical weather/climate model.

In most global operational systems, an “analysis cycle” (Fig 1) is commonly used; typically a 6-h cycle performed four times a day (Kalnay, 2002).

Fig 1. Typical global 6-h analysis cycle (Kalnay, 2012)
I. Motivation

- Although the numerical weather/climate model and data assimilation system are fixed, the observing system greatly change in time.
I. Motivation

• Although the numerical weather/climate model and data assimilation system are fixed, the observing system changes greatly in time.
• New satellites and other remote sensing platforms launch and end at irregular time intervals, and even existing satellites have an expected lifetime of 5-10 years, much shorter than what is needed for climate studies (Bosilovich et al. 2011).
Fig 1. Typical global 6-h analysis cycle (Kalnay, 2012)

- Although the numerical weather/climate model and data assimilation system are fixed, the observing system changes greatly in time.
- New satellites and other remote sensing platforms launch and end at irregular time intervals, and even existing satellites have an expected lifetime of 5-10 years, much shorter than what is needed for climate studies (Bosilovich et al. 2011).
- Such variations in the observing system can lead to systematic changes in the reanalysis time series, especially for the derived fields like precipitation (Zhang et al. 2012).
Fig 2. Global monthly mean precipitation (mm/day) time series for MERRA, NCEP-CFST, ERA-Interim, JRA_25, NCEP2, and ERA-40, compared with GPCP and CMAP (Chen et al., 2012)
The discontinuities in the time evolution associated with observing system changes are common in reanalysis products, especially starting with the ‘satellite period’ in 1979.
I. Motivation

Fig 2. Global monthly mean precipitation (mm/day)

These discontinuities are a very difficult problem. People have documented the “climate jumps” for different observing systems, but nobody has been able to solve this problem.
The overall purpose of this research is to find a solution to minimize the “climate jumps” associated with observing system changes.
II. Reanalysis System and Data

- **MERRA**
  - NASA’s Modern Era Retrospective-analysis for Research and Applications (MERRA, 1979-present) is a recent high-resolution reanalysis.
II. Reanalysis System and Data

• MERRA
  o NASA’s Modern Era Retrospective-analysis for Research and Applications (MERRA, 1979-present) is a recent high-resolution reanalysis.
  o It utilizes the NASA global data assimilation system, the Goddard Earth Observing System Model, Version 5 – Data Assimilation System (GEOS5-DAS), to produce a long-term synthesis, which introduces the historical satellite and conventional data records into a global gridded dataset.
II. Reanalysis System and Data

- MERRA
  - With emphasis on exactly closing the water and energy budgets, MERRA surpasses other latest reanalyses in assimilating time series of global monthly mean precipitation (Rienecker et al, 2011).
II. Reanalysis System and Data

- Special Sensor Microwave/Imager (SSM/I) product
  - SSM/I was generated by Remote Sensing Systems Inc. using a unified, physically based algorithm to simultaneously retrieve ocean wind speed (at 10 meters), atmospheric water vapor, cloud liquid water, and rain rate.
II. Reanalysis System and Data

• Special Sensor Microwave/Imager (SSM/I) product
  - SSM/I was generated by Remote Sensing Systems Inc. using a unified, physically based algorithm to simultaneously retrieve ocean wind speed (at 10 meters), atmospheric water vapor, cloud liquid water, and rain rate.
  - The SSM/I channels are sensitive to water vapor, cloud, precipitation and surface parameters, rather than temperature. Since only clear sky over ocean data is used, and the surface roughness is not involved in assimilation process, the direct impact of SSM/I is mostly on the moisture field.
II. Reanalysis System and Data

GEOS5 incorporated Incremental Analysis Update procedure (IAU, Bloom et al. 1996) to minimize the 6-hourly shock from the observation input (Kennedy et al. 2011).

$\left( \frac{dZ}{dt} \right)_{total} = \left( \frac{dZ}{dt} \right)_{dyn} + \left( \frac{dZ}{dt} \right)_{phy} + \left( \frac{dZ}{dt} \right)_{ana} \quad Eq(1)$

$analysis \ tendency \ from \ IAU: \ \left( \frac{dZ}{dt} \right)_{ana} = \frac{analysis \ increment}{6hr}$

Fig 3. A schematic of the IAU implementation in GEOS-5 (Rienecker et al, 2011)
II. Reanalysis System and Data

The model is then run over the six-hour interval again (i.e. corrector segment, Eq(1)) using analysis tendency, defined as AI divided by 6-hour, as an additional forcing term (Cullather and Bosilovich 2011).

\[
\frac{dZ}{dt}_{total} = \frac{dZ}{dt}_{dyn} + \frac{dZ}{dt}_{phy} + \frac{dZ}{dt}_{ana} \quad Eq(1)
\]

\[
\text{analysis tendency from IAU: } \frac{dZ}{dt}_{ana} = \frac{\text{analysis increment}}{6hr}
\]

Fig 3. A schematic of the IAU implementation in GEOS-5 (Rienecker et al, 2011)
In this way, only the tendency of a state can have discontinuities and not the state itself (Rienecker et al., 2011).

\[
\frac{dZ}{dt}_{total} = \frac{dZ}{dt}_{dyn} + \frac{dZ}{dt}_{phy} + \frac{dZ}{dt}_{ana} \quad \text{Eq}(1)
\]

Analysis tendency from IAU: \( \frac{dZ}{dt}_{ana} = \frac{\text{analysis increment}}{6\text{hr}} \)

Fig 3. A schematic of the IAU implementation in GEOS-5 (Rienecker et al., 2011)
Content

I. Motivation

II. Reanalysis System and Data

III. Proposed Hypotheses

IV. Methodology

V. Preliminary Results

VI. Conclusions

VII. Future Research

VIII. Summary
III. Proposed Hypotheses

- By comparing MERRA data and an experiment without SSM/I observations (named noSSMI hereafter), we can obtain the bias between MERRA and noSSMI due to the introduction of this new observing system.

- Based on this bias information, corrections would be made on MERRA data before the introduction of the new observing system, so as to adjust the earlier period data to match with the later period data.

- There are several potential methods to obtain the bias and to make the correction. In this study, we aim to make the correction in the IAU process.

- The correction terms will be derived from the bias of analysis increment between the original MERRA data and the noSSMI.
Content

I.  Motivation

II.  Reanalysis System and Data

III.  Proposed Hypotheses

IV.  Methodology

V.  Preliminary Results

VI.  Conclusions

VII.  Future Research

VIII. Summary
IV. Methodology

- We will test our methodology, inspired by Danforth et al (2007), to solve the SSM/I introduced “climate jump” in MERRA product.

- Danforth et al (2007) compared model data with a reference to generate a forcing term, and then added it to the model tendency equation. The model was run again with new tendency equation, and much reduced the bias.

- Following this method, if MERRA was considered as reference, noSSMI would represent the biased model
IV. Methodology

Fig 4. Schematic on correction definition

- The correction is defined as the bias of analysis increment between MERRA and noSSMI, for any combination of, moisture, temperature, and wind states (Fig 4), which is a 2yr-long 6-hourly dataset.
- By averaging this dataset at the same time within a year, a 365-day 6-hourly dataset is achieved.
This difference contains seasonal variability that we want to keep, and weather variability that we want to filter out because it is sampling noise.
IV. Methodology

- This difference contains seasonal variability that we want to keep, and weather variability that we want to filter out because it is sampling noise.

- The sampling error in the correction is reduced as follows: A Fourier analysis is applied to the 365-day 6-hourly correction time series. The results show that the correction term can be well represented by combining its annual mean, semi-annual (6-month), and seasonal (3-month) oscillations (Fig 6).
This difference contains seasonal variability that we want to keep, and weather variability that we want to filter out because it is sampling noise.

The sampling error in the correction is reduced as follows: A Fourier analysis is applied to the 365-day 6-hourly correction time series. The results show that the correction term can be well represented by combining its annual mean, semi-annual (6-month), and seasonal (3-month) oscillations (Fig 6).

After Fourier filter, the correction data is imported to IAU produced analysis increment every six hour, at corresponding time within a year, to test whether this correction can force noSSMI to approach MERRA. If successful, we can apply this correction to the period before SSM/I was introduced.
IV. Methodology

Challenges: where shall we start?

- Shall we just correct one or all variables? Considering the complex interrelationship between the variables influenced by SSM/I, we need to think carefully before adding the correction to the noSSMI reanalysis.

- The different variables affected by SSM/I interact through physical processes.
IV. Methodology

Challenges: where shall we start?

Fig 5. Specific humidity and temperature zonal mean difference with or without SSM/I observation (Chen. et al, 2012).

- Fig 5 shows the zonal mean differences in JJA between MERRA and the SSM/I withholding run, in Q analysis tendency (upper left), Q (lower left), T analysis tendency (upper right), and T (lower right).
- Similar sign for moisture and its increment but different sigh for T and its increment
IV. Methodology

Challenges: where shall we start?

Fig 5. Specific humidity and temperature zonal mean difference with or without SSM/I observation (Chen. et al, 2012).

- The extra water vapor from SSM/I means more precipitation, thus more latent heat is released to the air. So there is positive changes in T field. The positive T field change is responded by negative T increment from other observation, so to balance the extra latent heat release from precipitation.
IV. Methodology

Challenges: where shall we start?

**Fig 5.** Specific humidity and temperature zonal mean difference with or without SSM/I observation (Chen. et al, 2012).

- This cause-and-effect chain proves that moisture is the “driver” variable in SSM/I observation.
IV. Methodology

Challenges: where shall we start?

Fig 5. Specific humidity and temperature zonal mean difference with or without SSM/I observation (Chen. et al, 2012).

- Since moisture had dominant effect, we will begin with specific humidity correction, followed by specific humidity and temperature correction, and finally a correction in Q, T, and wind together.
## IV. Methodology

- **Experiment scenarios**

<table>
<thead>
<tr>
<th>Experiment name</th>
<th>Correction applied (Fig 6. &amp; Fig 7.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>add_Q</td>
<td>Specific humidity</td>
</tr>
<tr>
<td>add_2Q</td>
<td>2*Specific humidity</td>
</tr>
<tr>
<td>add_Q&amp;T</td>
<td>Specific humidity and temperature</td>
</tr>
<tr>
<td>add_2Q&amp;2T</td>
<td>2<em>Specific humidity and 2</em>temperature</td>
</tr>
<tr>
<td>add_U&amp;V&amp;T&amp;Q</td>
<td>Specific humidity, temperature, U wind, and V wind</td>
</tr>
</tbody>
</table>
Fig 6. Specific humidity (Q) correction and its Fourier transform truncations.

1st column: Q correction July mean;
2nd column: Q correction annual mean;
3rd column: Q correction combining annual mean and semi-annual oscillation (FT01);
4th column: Q correction combining annual mean, semi-annual oscillation (FT01), and seasonal oscillation (FT02).
• The limited Fourier series truncation that we used to reduce sampling errors was sufficient to represent well the observed mean AI difference between MERRA and noSSMI.
• This was also true for the other variables of the analysis, namely T, U, and V (not shown).
Fig 7. Vertical cross-section of zonal specific humidity correction, 1987 07-11

- Fig 7 shows the vertical structure of 5-month mean Q correction (analysis increment difference), July-November 1987.
b) 1987 July - November Total Precipitable Water (TQV)

Fig 8. Daily TQV global mean, July – Nov 1987. Q correction only
Fig 9. Daily TQV global mean, July – August 1987
Fig 10. Daily Total Precipitable Water bias, 1987 Sep-Nov mean

MERRA - noSSMI

Add_Q - noSSMI

Add_2Q - noSSMI
Fig 11. Daily precipitation global mean, July – November 1987
Q correction only
Fig 12. Daily precipitation global mean, July – August 1987
Fig 13. Daily precipitation bias, 1987 Sep-Nov

MERRA - noSSMI

Add_Q - noSSMI

Add_2Q - noSSMI
d) 1987 September – November, Specific Humidity

Fig 14. Qv, 1987 Sep-Nov mean.
Even though the correction was only applied in moisture, all of the four states (Q, T, U, and V) got improvement. It confirms Chen’s argument that moisture is the dominant state in SSM/I data.
d) 1987 September – November, Zonal Wind

Fig 16. U wind, 1987 Sep-Nov mean.
d) 1987 September – November, Meridional Wind

Fig 17. V wind, 1987 Sep-Nov mean.
I. Motivation
II. Reanalysis System and Data
III. Proposed Hypotheses
IV. Methodology
V. Preliminary Results
VI. Conclusions
VII. Future Research
VIII. Summary
VI. Conclusions

- We have tested the impact of withholding a new observation, i.e., SSM/I, that produced a substantial increase in precipitation.
- So far, we focused on moisture correction only because it is the “driver” of all variables in SSM/I observation.
- The main result we have obtained is that correcting the noSSMI reanalysis with the difference between the MERRA and noSSMI analysis increments is only partially successful because it is too weak.
- In fact, doubling the Q correction gives results that are much closer to the original MERRA.
- We pose the hypothesis that this is because the difference between the analysis increments is not just due to the assimilation of SSMI, but to the accumulation of these increments from the past, which evolve with nonlinear interactions between different variables during the training period.
VI. Conclusions (cont.)

- Alternatively, the proper correction should be obtained by comparing analysis increments between 1) MERRA (exp_MERRA), and 2) an analysis started from MERRA every 6 hours but withholding the SSMI observation (exp_MnoSSMI). (Fig 16.)

We hypothesize that the 6-hr AI difference caused by withholding SSM/I doesn’t contain accumulated error from the past, and there is no nonlinear interactions between different variables within SSM/I.

![Diagram showing time steps during training period, typically 6 hour window.](image)

\[
\text{correction} = \frac{\text{AI difference}}{\text{AI}_{MERRA} - \text{AI}_{MnoSSMI}}
\]

Fig 16. Schematic on 6-hourly correction definition in MERRA system.
More experiments:

A) Correct all variables (U, V, T,Q, ongoing), not just Q. We expect that the results will still be suboptimal, since we are still accumulating analysis increments from the previous correction.

B) Test the proper correction as in Danforth et al (2007) with the MERRA system as described before. However, this is computationally unfeasible (3 hours for one day).

C) Instead we will test the technique with the SPEEDY-LETKF, using the system developed by Dr. Kang. This system is simpler and computationally efficient (5 hours for one year). I will first do a reanalysis with RAOBs and AIRS observations (exp_R&A). Starting from exp_R&A, another reanalysis without AIRS will be conducted every 6 hour (exp_R). (Fig 17.)
More experiments:

The correction will be added to exp_R model tendency equation. We expect this correction could represent the impact of taking off AIRS, and the corrected exp_R experiment can simulate exp_R&A better.

Fig 17. Schematic on 6-hourly correction definition in SPEEDY-LETKF system
Content

I. Motivation

II. Reanalysis System and Data

III. Proposed Hypotheses

IV. Methodology

V. Preliminary Results

VI. Conclusions

VII. Future Research

VIII. Summary
VIII. Summary

◊ The present research studies the discontinuity in the time evolution of MERRA associated with observing system, SSM/I, change. The overall purpose is to find a solution to minimize such discontinuities.

◊ MERRA utilizes the NASA GEOS5-DAS, which is incorporated with Incremental Analysis Update (IAU) procedure. IAU distributes AIs evenly over assimilation cycle as an additional forcing term in model’s tendency equation. So, making correction in IAU process is the same as changing model tendency.

◊ During training period (July 1987 – June 1989), the correction was obtained by comparing the AIs between MERRA and noSSMI experiments. The correction was then applied to IAU process.
This method is only partially successful because it is too weak. In fact, doubling the Q correction gives results that are much closer to the exp_MERRA.

One explanation is that, this method doesn’t take nonlinear interactions between different variables of SSM/I into account. The nonlinear interaction impacts would accumulate during the whole training period and introduce errors not just due to assimilation of SSM/I.

Considering computational feasibility, we will use SPEEDY-LETKF system, developed by Dr. Kang, to test our hypothesis that, the proper correction should be obtained by comparing AIs between one analysis, and another experiment restarting from this analysis while withholding certain observation systems every 6 hour.
**VI. Conclusions (cont.)**

Correlation = $\frac{AI_{\text{MERRA}}}{AI_{\text{noSSMI}}}$

Time steps during training period, typically 6 hour window

$\text{correction} = \frac{AI_{\text{difference}}}{AI_{\text{MERRA}}} = \frac{AI_{\text{MERRA}}}{AI_{\text{MnoSSMI}}}$

Exp_MnoSSMI, starting from MERRA but withholding SSMI