Ensemble Data Assimilation of GSMaP Precipitation into the Nonhydrostatic Global Atmospheric Model NICAM

**Shunji Kotsuki**\(^1\), Koji Terasaki\(^1\), Guo-Yuan Lien\(^1\), Takemasa Miyoshi\(^1,2\), and Eugenia Kalnay\(^2\)

\(^1\)Data Assimilation Research Team, RIKEN-AICS, Japan
\(^2\)University of Maryland, College Park, Maryland, USA

Data Assimilation Seminar, Dec 25, 2015@ RIKEN-AICS
Motivation

Oki and Kanae (2006)
Outline

• Assimilating GSMaP with NICAM-LETKF
  – Introduction
  – Gaussian Transformation
  – DA-cycle experiments
  – Forecast experiments

• Estimation of global crop calendar with NDVI
  – Current status and challenges
Outline

- Assimilating GSMaP with NICAM-LETKF
  - Introduction
  - Gaussian Transformation
  - DA-cycle experiments
  - Forecast experiments

- Estimation of global crop calendar with NDVI
  - Current status and challenges
Outline

• Assimilating GSMaP with NICAM-LETKF
  – Introduction
  – Gaussian Transformation
  – DA-cycle experiments
  – Forecast experiments

• Estimation of global crop calendar with NDVI
  – Current status and challenges
GPM: Global Precipitation Measurement

Hou et al. 2014
GPM: Global Precipitation Measurement

Hou et al. 2014
Goals

- **To improve NWP** using satellite-derived precipitation data following *Lien et al. (2013, 2015a, 2015b)*
- **To produce a new precipitation product** through data assimilation
Experimental Setting

• **Numerical Model**
    • GL6 (approx. 110 km resolution)

• **Observations**
  - CTL: Radiosondes
  - EXP: Radiosondes + GSMaP/Gauge (Ushio et al. 2009)
    • *with Gaussian transformation*

• **Data assimilation**
  - LETKF (Hunt et al. 2007)
  - NICAM-LETKF (Terasaki et al. 2015) with 36 members
    • 3D-LETKF
    • Localization: 400 km for horizontal & 0.4 log(p) for vertical
    • Relaxation to prior perturbation (Zhang et al. 2004; $\alpha = 0.7$)
NICAM-LETKF (Terasaki et al. 2015)

NICAM Ensemble Forecast

$X^f$

Obs operator

$y^o$ & $HX^f$

LETKF

$X^a$

$X^f$

$H$: obs. operator

$y^o$: observation

$X^{f,a}$: guess, analysis (ensemble)
**Assimilation of GSMaP by NICAM-LETKF**

- NICAM Ensemble Forecast
- Obs operator
- QC & Gaussian Transform.
- y^o & HX^f
- LETKF
- Inverse Transformation
- Precipitation analysis

**Symbols:**
- X^a: guess, analysis (ensemble)
- X^f: forecast
- H: obs. operator
- y^o: observation
- X^{f,a}: guess, analysis (ensemble)
Outline

• Assimilating GSMaP with NICAM-LETKF
  – Introduction
  – **Gaussian Transformation**
  – DA-cycle experiments
  – Forecast experiments

• Estimation of global crop calendar with NDVI
  – Current status and challenges
Gaussion Transformation

\[ F^G(\tilde{y}) = F(y) \quad \iff \quad \tilde{y} = F^{-1}(F(y)) \quad \iff \quad y = F^{-1}(F^G(\tilde{y})) \]

Forward transform (mm/6hr → sigma) Inverse transform (sigma → mm/6hr)

\( y \) : original variable (mm/6hr) \hfill \( \tilde{y} \) : Transformed variable (sigma)
\( F() \) : CDF of original variable \hfill \( F^G() \) : CDF of Gaussian distribution

---

PDF

```
+----------------+----------------+
| 0.8            | 0.6            |
| 0.4            | 0.2            |
| 0.0            | 0.2            |
+----------------+----------------+
| 0 1 2 3 4 5 6 7 8| y (mm/6h)     |
+----------------+----------------+
```

CDF

```
+----------------+----------------+
| 0.8            | 0.6            |
| 0.4            | 0.2            |
| 0.0            | 0.2            |
+----------------+----------------+
| 0 1 2 3 4 5 6 7 8| y (mm/6h)     |
+----------------+----------------+
```

Original variable

---

\(--\) : Model
\(--\) : Obs.

Step 0: Obtain PDF & CDF

---

Lien et al. (2013, 2015)
Gaussian Transformation

\[ F^G(\tilde{y}) = F(y) \quad \Leftrightarrow \quad \tilde{y} = F^{-1}(F(y)) \quad \Leftrightarrow \quad y = F^{-1}\left[F^G(\tilde{y})\right] \]

- Forward transform (mm/6hr \(\rightarrow\) sigma)
- Inverse transform (sigma \(\rightarrow\) mm/6hr)

\(y\) : original variable (mm/6hr)
\(\tilde{y}\) : Transformed variable (sigma)
\(F()\) : CDF of original variable
\(F^G()\) : CDF of Gaussian distribution

---

Step 0: Obtain PDF & CDF

Step 1: Compute \(F(y)\)

Lien et al. (2013, 2015)
Gaussian Transformation

\[ F^G(\tilde{y}) = F(y) \iff \tilde{y} = F^{G^{-1}}[F(y)] \iff y = F^{G^{-1}}[F^G(\tilde{y})] \]

Forward transform (mm/6hr → sigma)  Inverse transform (sigma → mm/6hr)

\( y \) : original variable (mm/6hr)  \( \tilde{y} \) : Transformed variable (sigma)

\( F() \) : CDF of original variable  \( F^G() \) : CDF of Gaussian distribution

---

Step 0: Obtain PDF & CDF

Step 1: Compute \( F(y) \)

Lien et al. (2013, 2015)
Gaussian Transformation

\[ F^G (\tilde{y}) = F(y) \quad \iff \quad \tilde{y} = F^{-1} \left( F(y) \right) \quad \iff \quad y = F^{-1} \left( F^G (\tilde{y}) \right) \]

Forward transform (mm/6hr\(\rightarrow\)sigma)  Inverse transform (sigma\(\rightarrow\)mm/6hr)

\(y\) : original variable (mm/6hr) \quad \tilde{y} \quad : \text{Transformed variable (sigma)}

\(F()\) : CDF of original variable \quad \(F^G ()\) : CDF of Gaussian distribution

---

Step 0: Obtain PDF & CDF

Step 1: Compute \(F(y)\)

Step 2: Compute

\[ \tilde{y} = F^{-1} \left( F(y) \right) \]

---

Lien et al. (2013, 2015)
Gaussian Transformation

\[ F^G (\tilde{y}) = F (y) \iff \tilde{y} = F^{-1} \left[ F (y) \right] \iff y = F^{-1} \left[ F^G (\tilde{y}) \right] \]

Forward transform (mm/6hr → sigma)  Inverse transform (sigma → mm/6hr)

\( y \) : original variable (mm/6hr)  \( \tilde{y} \) : Transformed variable (sigma)

\( F() \) : CDF of original variable  \( F^G () \) : CDF of Gaussian distribution

---

**PDF**

Original variable  Transformed variable

**CDF**

Step 0: Obtain PDF & CDF

Step 1: Compute \( F(y) \)

Step 2: Compute

\[ \tilde{y} = F^{-1} \left[ F (y) \right] \]

Lien et al. (2013, 2015)
PDF/CDF production

T v.s. ERA Interim (850 hPa, Global [-90:90])

Spin-up period

Experiment

T=1

30 days samples

RMSD

SPRD
PDF/CDF production

Spin-up period

Experiment

30 days samples

T = 1

T v.s. ERA Interim (850 hPa, Global [-90:90])

1 to produce PDF/CDF
PDF/CDF production

Spin-up period

Experiment

① to produce PDF/CDF

T = 1

30 days samples

③ Data Assimilation

② Gaussian-Transformation

T v.s. ERA Interim (850 hPa, Global [-90:90])
PDF/CDF production

T v.s. ERA Interim (850 hPa, Global [-90:90])

Spin-up period

Experiment

① to produce PDF/CDF

SPREAD and RMSD [K]

0 1 2 3 4 5

10/01 10/11 10/21 11/01 11/11

② Gaussian-Transformation

T=1

③ Data Assimilation

T=2

30days samples

④
Gaussian Transformation

NICAM (ques) norm [mm/6hr] 2014110100(UTC)

NICAM(org)

GSMaP (obs) norm [mm/6hr] 2014110100(UTC)

GSMaP(org)
Gaussian Transformation

NICAM (org) norm [mm/6hr] 2014110100(UTC)

GSMaP (org) norm [mm/6hr] 2014110100(UTC)

Transformation (Model CDF)

Transformation (Obs. CDF)

GSMaP (Gauss) sigma

GSMaP (Gauss) sigma
wo Gaussian-Transformation

Innovation Statistic

<table>
<thead>
<tr>
<th>Obs-Guess (org)</th>
<th>average: 0.080</th>
<th>sigma: 2.837</th>
<th>skewness: 6.372</th>
<th>kurtosis: 93.91</th>
</tr>
</thead>
</table>

w Gaussian-Transformation

Innovation Statistic

<table>
<thead>
<tr>
<th>Obs-Guess (GT)</th>
<th>average: -0.015</th>
<th>sigma: 0.729</th>
<th>skewness: 0.418</th>
<th>kurtosis: 0.696</th>
</tr>
</thead>
</table>

More Gaussian

Sampling period: 2014110100 - 2014110118
Forward/Inverse Transformations

NICAM (org) norm [mm/6hr] 2014110100 (UTC)

NICAM (org)

Transformation (Model CDF)

NICAM (Gauss)

mm/6hr

sigma
Forward/Inverse Transformations

NICAM (org) norm [mm/6hr] 2014110100 (UTC)

NICAM (Gauss)

Transformation (Model CDF)

Inverse Transformation (Obs. CDF)

GSMaP-like NICAM

mm/6hr

sigma
Precipitation after the first analysis

Noisy field w/ negative values

Precipitation data comparison:
- **No-Transform (NT)**
- **Gaussian-Transform (GT)**

![Map showing precipitation data with and without Gaussian Transform](image)
Precipitation after the first analysis

Improvement in precipitation field

Assimilation

NICAM (guess) - obs[mm/6hr] 2014110100

NICAM-Like NICAM (anal) - obs[mm/6hr] 2014110100(UTC)

GSMaP-Like NICAM (guess)

GSMaP-Like NICAM (anal)

GSMaP (obs)
• Assimilating GSMaP with NICAM-LETKF
  – Introduction
  – Gaussian Transformation
  – DA-cycle experiments
  – Forecast experiments

• Estimation of global crop calendar with NDVI
  – Current status and challenges
RMSDs relative to ERA Interim (in 2014)

U vs. ERA Interim (500 hPa, Global)

RMSD [m/s]

Spin-up period

10/01 10/16 11/01 11/16 12/01 12/16 12/31

GAU-GRD1
GAU-GRD3
GAU-GRD5
RAOBS

—: Raobs: Radiosondes ONLY
—: GRD1: Radiosondes + GSMaP/Gauge (ALL)
—: GRD3: Radiosondes + GSMaP/Gauge (every 3x3 grids)
—: GRD5: Radiosondes + GSMaP/Gauge (every 5x5 grids)
Desrozier’s diagnostics (for precip. obs)

\[
R = \left\langle d^a \left( d^b \right)^T \right\rangle \quad d^{a(b)} = y^o - Hx^{a(b)}
\]

Desrozières et al. (2005)

**NOTE:** Diagnosed with suboptimal experiment GRD1

2014/12/01 – 2014/12/31
Desrozier’s diagnostics (for precip. obs)

\[ R = \left\langle d^a \left( d^b \right)^T \right\rangle \]

\[ d^{a(b)} = y^o - Hx^{a(b)} \]

Desroziers et al. (2005)

Horizontal correlations (ocean)

Horizontal correlations (land)

Strong horizontal correlation !!!

NOTE: Diagnosed with suboptimal experiment GRD1

2014/12/01 – 2014/12/31
Best experiment (RMSD changes)

RMSD (T): DA Precip - Control
NH(Top: 20-60), TR(Mid: -20-20), SH(Bot: -60-20)

RMSD (Qv): DA Precip - Control
NH(Top: 20-60), TR(Mid: -20-20), SH(Bot: -60-20)

RMSD (U): DA Precip - Control
NH(Top: 20-60), TR(Mid: -20-20), SH(Bot: -60-20)

RMSD (V): DA Precip - Control
NH(Top: 20-60), TR(Mid: -20-20), SH(Bot: -60-20)

RMSD (GRD5) — RMSD (Raobs)
Best experiments (MAE changes)

45-days average (2014/11/17-2014/12/31)

(a) $T$ (K) at 500 hPa
(b) $Q_v$ (g/kg) at 700 hPa
(c) $U$ (m/s) at 500 hPa
(d) $V$ (m/s) at 500 hPa

improved $\leftarrow$ RMSD (GRD5) $\rightarrow$ RMSD (Raobs) $\rightarrow$ degraded
Outline

• Assimilating GSMaP with NICAM-LETKF
  – Introduction
  – Gaussian Transformation
  – DA-cycle experiments
  – Forecast experiments

• Estimation of global crop calendar with NDVI
  – Current status and challenges
RMSDs: 120h Forecasts vs. ERA Interim

Validation with mean of ensemble forecasts from different initial dates

---

- **Radiosondes ONLY (RAOBS)**
- **Radiosondes + GSMaP/Gauge (GRD5)**

U vs. ERA Interim (500 hPa, Global)

**Forecast**

**DA-cycle**
RMSDs: 120 hr Forecasts vs. GSMaP/Gauge

**Threat Score (≥ 1 mm/6hr)**

- : Radiosondes ONLY (RAOBS)
- : Radiosondes + GSMaP/Gauge (GRD5)

**Threat Score (≥ 5 mm/6hr)**

Precipitation forecasts are improved !!!

Average over 8 ensemble forecasts from different initial dates
• Lien et al. (2015) approach works well with NICAM-LETKF & GSMaP
  – Observation data thinning was essential
    • Horizontal obs error correlation of precipitation
Summary

RMSD improvements relative to CTRL (PREPBUFR) experiment

Thanks to Dr. Terasaki

negative=improved
• Assimilating GSMaP with NICAM-LETKF
  – Introduction
  – Gaussian Transformation
  – DA-cycle experiments
  – Forecast experiments

• Estimation of global crop calendar with NDVI
  – Current status and challenges
Research article

**SACRA – a method for the estimation of global high-resolution crop calendars from a satellite-sensed NDVI**

S. Kotsuki\(^1\) and K. Tanaka\(^2\)

\(^1\)RIKEN Advanced Institute for Computational Science, Kobe, Japan
\(^2\)Disaster Prevention Research Institute, Kyoto University, Kyoto, Japan

Received: 22 Dec 2014 – Published in Hydrol. Earth Syst. Sci. Discuss.: 29 Jan 2015
Reviewed: 05 Oct 2015 – Accepted: 15 Oct 2015 – Published: 05 Nov 2015

**Abstract.** To date, many studies have performed numerical estimations of biomass production and agricultural water demand to understand the present and future supply-demand relationship. A crop calendar (CC), which defines the date or month when farmers sow and harvest crops, is an essential input for the numerical estimations. This study aims to present a new global data set, the Satellite-derived Crop Calendar for Agricultural Simulations (SACRA), and to discuss advantages and disadvantages compared to existing census-based and model-derived products.

We estimate global CC at a spatial resolution of 5 arcmin using satellite-sensed normalized difference vegetation index (NDVI) data, which corresponds to vegetation vitality and senescence on the land surface. Using the time series of the NDVI averaged from three consecutive years (2004–2006), sowing/harvesting dates are estimated for six crops (temperate-wheat, snow-wheat, maize, rice, soybean and cotton). We assume time series of the NDVI represent the phenology of one dominant crop and estimate CCS of the dominant crop in each grid. The dominant crops are determined using harvested areas based on census-based data. The cultivation period of SACRA is identified from the time series of the NDVI; therefore, SACRA considers current effects of human decisions and natural disasters. The difference between the estimated sowing dates and other existing products is less than 2 months (<62 days) in most of the areas. A major disadvantage of our method is that the mixture of several crops in a grid is not considered in SACRA. The assumption of one dominant crop in each grid is a major source of discrepancy in crop calendars between SACRA and other products. The disadvantages of our approach may be reduced with future improvements based on finer satellite sensors and crop-type classification studies to consider several dominant crops in each grid. The comparison of the CC also demonstrates that identification of wheat type (sowing in spring or fall) is a major source of error in global CC estimations.

<table>
<thead>
<tr>
<th></th>
<th>Census-based</th>
<th>Model-based</th>
<th>Satellite-based (this study)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main inputs</strong></td>
<td>Census Data</td>
<td>Atmos. Forcing</td>
<td>Satellite Obs.</td>
</tr>
<tr>
<td><strong>Resolution</strong></td>
<td>Country/State scale</td>
<td>Equal to inputs</td>
<td>5 arc-min</td>
</tr>
<tr>
<td><strong>Detection of cultivation</strong></td>
<td>Hard</td>
<td>Hard</td>
<td>Easy</td>
</tr>
<tr>
<td><strong>Mixture of crops (phenology)</strong></td>
<td>Possible</td>
<td>Possible</td>
<td>Impossible</td>
</tr>
<tr>
<td><strong>Future projection</strong></td>
<td>Impossible</td>
<td>Possible</td>
<td>Impossible</td>
</tr>
</tbody>
</table>
EX) Double cropping grid in India

1. Remove cloud effect
2. Aggregation (time & space)
3. Normalization
EX) Double cropping grid in India

1. Remove cloud effect
2. Aggregation (time & space)
3. Normalization
4. Detect sowing/harvesting date

(a) Cropping map (5 min)
Estimated crop calendar (major crop)

Satellite

Census

Model
Difference btw products

Satellite – Census

Satellite – Model

Earlier signal

Later signal

smaller difference
Summary and Challenges

- Major sources of uncertainty
  - Spring or winter wheat
  - Census-based: no data region (U.N.’s survey)
  - Model-based: sowing date (human’s decision)
  - Satellite-based: land cover (grass or cropland)
Thank you for your attention