Land-surface data assimilation systems

Pedro Viterbo
Instituto de Meteorologia & Centro de Geofísica da Universidade de Lisboa

Agradecimentos: Alan Betts, Emanuel Dutra (CGUL), Matthias Drusch (ECMWF), Laura Ferranti (ECMWF)

ECMWF Workshop on atmospheric reanalysis, 19 - 22 June 2006
Layout

- Introduction: Imperfect models and inaccurate data
- Land surface in ERA-40: Strengths and weaknesses
- Soil moisture
- Snow
- Conclusions
Layout

• Introduction: Imperfect models and inaccurate data
  – Model drift
  – State variables
  – Observations
  – Reanalyses practice
• Land surface in ERA-40: Strengths and weaknesses
• Soil moisture
• Snow
• Conclusions
The need for data assimilation

- **Long time scales** in land state variables (deeper soil water and temperature, snow): Forecast drifts are possible. A way of controlling model drift (initialisation of soil variables) is needed.

- Drifts due to errors in forcing (precipitation, radiation, BL humidity) or in the land-surface model/model fluxes

**Mackenzie river basin precipitation: era40 vs. observations**

1958-70

\[ R^2 = 0.733 \]

1987-97

\[ R^2 = 0.648 \]

\[ y = 0.89x \]

\[ y = 1.07x \]

\[ y = 1.025x \] (\[ R^2 = 0.915 \])

\[ y = 1.179x \] (\[ R^2 = 0.919 \])
What to initialise and how?

- There are no **routine direct observations** of soil state variables
- For an effective data assimilation, the state variables with longer timescales need to be initialised:

<table>
<thead>
<tr>
<th>STATE VARIABLES</th>
<th>OBSERVABLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root zone soil moisture</td>
<td>No direct observations globally (see below)</td>
</tr>
<tr>
<td>Snow mass</td>
<td>Snow depth</td>
</tr>
<tr>
<td></td>
<td>Snow cover from remote sensing</td>
</tr>
<tr>
<td></td>
<td>Snow mass from AMSR-E (saturates at higher values; does not work in forest areas)</td>
</tr>
<tr>
<td>Above ground biomass</td>
<td>(Indirectly) from remote sensing: vegetation indices, LAI/fAPAR</td>
</tr>
</tbody>
</table>

- Proxy observations for root zone soil water
  - Screen level T and RH: Linked to Bowen ratio
  - Rainfall rates
  - Window channel brightness temperature: Early morning evolution is linked to evaporation
  - Microwave (L-band, 1.4 GHz) radiances: Top soil moisture
  - C-band passive and active systems: Top soil moisture

- **State variables** are non-linearly related (via the equations of the land-surface scheme) to observations: Complex observation operators
Current practice at global centres

<table>
<thead>
<tr>
<th></th>
<th>Soil water</th>
<th>Snow mass</th>
<th>Biomass</th>
</tr>
</thead>
<tbody>
<tr>
<td>JRA-25</td>
<td>???</td>
<td>Use of snow depth obs</td>
<td>Monthly climate LAI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Snow cover</td>
<td></td>
</tr>
<tr>
<td>ERA-40</td>
<td>OI for root zone, based on 2T/RH</td>
<td>Use of snow depth obs, Cressman</td>
<td>Constant LAI</td>
</tr>
<tr>
<td>NCEP/NCAR R1</td>
<td>Relaxation to climate</td>
<td>Climate</td>
<td>Monthly climate LAI</td>
</tr>
<tr>
<td>NCEP/NCAR R2</td>
<td>(Observed precipitation “ingestion”)</td>
<td>???</td>
<td>Monthly climate LAI</td>
</tr>
</tbody>
</table>

Land surface analysis lies, in terms of methods and data usage, far behind its atmospheric counterpart
G(N/E)LDAS

- Global/North-American/European Land Data Assimilation Systems
  - Running offline a land surface scheme, forced by near-surface meteorology, downwelling radiative fluxes and precipitation
  - Best possible forcing, very often observation-based estimates hybridized with reanalysis
  - Inexpensive, often run with several models and several versions of the forcing. Ideal tool to have a land-surface model climate

- In most cases (GLDAS, NLDAS) LDAS is a misleading name, because there is no data assimilation involved: It assumes that the forcing is correct.

- Global Soil Water Project (GSWP) is one version of “LDAS”, using best available forcing for 1986-1995
Introduction: Imperfect models and inaccurate data

Land surface in ERA-40: Strengths and weaknesses
  - Surface fluxes
  - Soil water
  - Don’t throw the baby with the bath water

Soil moisture

Snow

Conclusions
Surface energy balance
(41°-50°N 2°W-16°E)

- ERA-40: Boxes and whiskers; 2001 e 2002 ☐; 2003 •

- Positive surface solar radiation anomalies since March, associated to anomalously low cloud
- The surface starts responding in June, with an increase in sensible heat flux followed by a decrease in evaporation in July
Soil water index (SWI)

- SWI=1: No restrictions to evapotranspiration due to soil water
- SWI=0: Evapotranspirations shuts

- Soil water anomalies from March onwards, peaking in August
- Summer anomalies not intense enough
Soil water analysis increments
Soil water: 2003 and climate

2001-2003 and ERA-40

Model climate: 40 year AMIP run

Model climate: 6 months ocean coupled Hindcasts+2003

Soil water assimilation reduces seasonal soil water amplitude

ECMWF, Reading, Jul 2006
Mean annual range of soil water
(Dirmeyer et al 1994, JHM)

- ERA-40 has the smallest mean seasonal amplitude
Reanalyses soil water (shaded) and GSWP: ranked validation against obs

Guo et al., 2006: QJRMS, accepted

- ERA-40 is the best reanalysis product, and better than many GSWP models
Layout

• Introduction: Imperfect models and inaccurate data
• Land surface in ERA-40; Strengths and weaknesses
• Soil moisture
  – What we do now: Assimilation based on two-metre temperature/humidity
  – What we should do: Sample the entire physical space, to avoid overfitting or aliasing
  – What we can do: Use LDAS as a weak constraint
• Snow
• Conclusions
Soil water increments: June-July 2002

Accumulated increments

Variance of daily increments
Impact of analysis increments on surface fluxes

Latent heat flux

Sensible heat flux

Control

Open loop - Control

ECMWF, Reading, Jul 2006
Root zone soil moisture: observables and caveats

- BL T/RH

• Fair weather spring/summer conditions
• Low wind speed

What we do now
Synergy of observations

- **Screen level temperature and humidity** are indirect linked to soil moisture through evaporative cooling.
- **Microwave brightness temperature** contains more direct information of near surface soil moisture and is less dependent on atmospheric conditions.
  - Penetration depth of $\mu$w $T_b$ depends on:
    - Soil texture
    - Soil temperature profile
    - Vegetation fraction
    - Vegetation water content
    - Surface roughness
    - LSMEM (Land Surface Microwave Emissivity Model) for model equivalent of $T_b$
- **Rate of change of thermal infrared brightness temperature** contains information on soil moisture, but
  - Clear sky data only;
  - **Model Tskin is very sensitive to aerodynamical resistance (surface roughness)**

ECMWF, Reading, Jul 2006
Root zone soil moisture: observables and caveats

What we should do

- BL T/RH
- Vegetation state (NDVI)

Root zone Soil moisture

- \( \mu w \) Tb
- L- and C- band (d=1-5 cm)

- Fair weather spring/summer conditions
- Low wind speed

- Clear-sky data
- Saturation of LAI=f(NDVI)

- Low water on vegetation
- C-band limited to non-forest areas
ECMWF experimental soil moisture analysis system

- Extended Kalman filter
- Assimilation of two-metre temperature and relative humidity and microwave Tb
- Updated forecast errors
- System is forced with observed estimates of precipitation and downward SW and LW surface radiation, where available
Assimilation of mw Tb: Performance of 2T/RH

- The control simulation (indeed, all simulations) are too warm and too dry. Model day-to-day variability of humidity exceeds observations.
- Assimilation of screen-level parameters decrease the warm/dry bias by 30-40%.
- Assimilation of mw Tb, on top of screen-level parameters, slightly deteriorates the fit to screen-level observations.
Surface soil moisture and Tb (SGP97)

- All assimilation configurations improve the fit to top soil moisture, in particular those using Tb.
On its own, Tb gives a very good simulation for the root layer.
The use of screen-level parameters, w/ or w/o Tb, brings the simulation away from observations.
Evaporative fraction [EF = LE/(H+LE)], the relevant quantity for the surface impact on the atmosphere, is underestimated by the control simulation (cf. dry/warm bias).

- $\mu w$ Tb, on its own, is not effective enough to change EF.
- EF is clearly improved when screen-level parameters are used.
- The synergy of all 3 observations is again visible.
Problem 1: Matching vertical resolution of in-situ, remote sensing, and model soil moisture

Matthias Drusch
Joint and marginal pdf

Matthias Drusch
Problem 2: Bias corrections of remote sensing data

Oklahoma data sets 2002

Matthias Drusch
CDF matching

Cumulative Distribution Function

CDF\textsubscript{ECM}(x') = CDF\textsubscript{TMI}(x)

(Reichle and Koster, 2004)
TMI soil moisture transformation

transfer function
03/2002-10/2002

- CDF matching reduces systematic errors:
The bias has been removed and the dynamic range has been adjusted.

\[ r^2 = 0.18 \rightarrow r^2 = 0.69 \]

Matthias Drusch

ECMWF, Reading, Jul 2006
Transferred TMI images

Volumetric soil moisture 01/10/2002

TMI Pathfinder Data Set

Transferred TMI data (2002 regional transfer functions)

ECMWF
Model development at ECMWF

• The use of LAI information requires model developments
  – In order to perform data assimilation the model has to look like the observations
• At present, LAI is time-invariant in the ECMWF
• Current development (monthly LAI climatology)
• In preparation (carbon/biomass) (w/ KNMI)
  – Photosynthesis-based evaporation formulation (from ISBA-Ags), linking water and carbon cycles
  – Spin-off: (natural) Land carbon fluxes
  – Vegetation: Biomass evolves according to parameterized growth and mortality functions
• Data assimilation of biomass using LAI retrievals (at Meteo-France)
  – Biomass to LAI observation operator
  – Variational approach minimizing the difference to background and observations
  – State variables: Biomass and soil moisture
• **Analysis of Biomass using LAI observations** (10 days analysis period)

• **Good LAI correction**

• **Overall good Biomass analysis** (particularly in 2002)

• **Strong (negative) impact of root zone soil moisture, w2,** (different LAI → different root water extraction and transpiration rates)

Lionel Jarlan

ECMWF, Reading, Jul 2006
• **Analysis of Biomass and w2 using LAI observations** (10 day analysis period) + non stationary covariance matrix

• Good LAI correction

• Overall good Biomass analysis

• … but w2 better in agreement with observations during high water stress period

Lionel Jarlan

ECMWF, Reading, Jul 2006
Iberian LDAS

- Running TESSEL (land surface model of ERA-40) offline, forced by ERA-40
  - TESSEL results without land assimilation
Standardized Precipitation Index (SPI): ERA-40 vs observations

Emanuel Dutra
Drought indicators

- AAS – Normalized soil water anomaly
- SPI: Standardized precipitation index
- PDSI: Palmer Drought Stress Index (Official drought indicator in Portugal)
SPI and normalized soil water anomaly

**Offline TESSEL**

**ERA-40 soil water**

Emanuel Dutra

ECMWF, Reading, Jul 2006
Correlation of SPI and normalized soil water anomaly

Emanuel Dutra

ECMWF, Reading, Jul 2006
Root zone soil moisture: observables and caveats

What we can do

- Run the surface model offline, forced by 24-48 hour precipitation
- Use output soil water as a weak constraint to OI soil water analysis

• Fair weather spring/summer conditions
• Low wind speed

BL T/RH
Introduction: Imperfect models and inaccurate data

Land surface in ERA-40; Strengths and weaknesses

Soil moisture

Snow
  - Large analysis increments
  - Snow cover vs. snow depth
  - Model problems: Density and melting

Conclusions
Mackenzie river basin era40:
Surface water and snow budget

\[ \eta_n^{a+1} = \eta_n^a + \sum \left[ P + M - E_l - Y + \Delta\eta \right] \]
\[ S_n^{a+1} = S_n^a + \sum \left[ F - E_s - M + \Delta S \right] \]

• Surface analysis increments are of the same order of the seasonal evolution of the soil water and snow mass budget
ECMWF snow depth and snow cover analysis

- Background (modified short term forecast)
- Observation operator uses model snow density to go from model snow mass to snow depth
- Snow depth (conventional observations)
- NOAA/NESDIS snow cover, ScNESDIS product is used
  - ScNESDIS=0 is unambiguous information
    - This information is presented as an observation
  - ScNESDIS > 0 and Scbackground=0 is ambiguous information
    - This information is used to modify the short term forecast, assigning a small value to these points
- In the following, we will compare analysed snow mass with MODIS snow extent and a high-resolution (1 km), high-quality US snow analysis product (SNODAS); both fields are upscaled to model resolution
Comparison with MODIS

Frequency of days where $\text{Sc}_{\text{MODIS}}=0$ and $\text{snow_mass}_{\text{analysis}} > 0$

March 2002

May 2002

Analysis w/o $\text{Sc}_{\text{NESDIS}}$

Full analysis

Matthias Drusch
Comparison with SNODAS: snow mass

SNODAS 30/11/04

Analysis w/o Sc_{NESDIS} 30/11/04

Full analysis 30/11/04

Matthias Drusch
• Introduction: Imperfect models and inaccurate data
• Land surface in ERA-40; Strengths and weaknesses
• Soil moisture
• Snow
• Conclusions
Conclusions

- Land surface data assimilation is necessary to correct drifts in slow components of the land system, caused by deficiencies in the forcing or inaccurate (surface) model physics.
- Surface analysis still lags behind its atmospheric counterpart.
- New methods allow the use of more observations:
  - A more complete sampling of soil water in physical space: Evaporative feedback to the atmosphere (two-metre temperature and humidity), hydrology (1.4 and 6.4 GHz microwave Tb), and vegetation state (vegetation indices, lead area index, ...)
  - Sinergy of 3 observation types reduces the risk of overfitting and/or aliasing.
- Non-linear transfer functions to match model and observation space:
  - Bias correction will always be necessary.
- In case of contradictory information between screen-level parameters and mw Tb on soil moisture, NWP centres will tend to tune the assimilation to fit the evaporative fraction, since that is the quantity impacting on the atmosphere.
- In practice, the output of LDAS can be used to provide a penalty term to soil assimilation.
- Snow:
  - Check the observations beforehand (snow)
  - Need snow cover for whole reanalysis period (it exists since 1966)
  - Optimal ingestion of special observations (Russia and Canada)?