An improved alternative soil moisture variational analysis algorithm for the Canadian Land Data Assimilation System (CaLDAS)

Pablo J. Grunmann

Presentation layout

1. Nature of the problem:
   - The initialization of soil state variables in numerical weather prediction models. Why we do it and how.
   - Example: operational.
2. CaLDAS:
   - Main features with respect to operational.
   - Short description (G-P Balsamo, December 12, 2005 presentation).
3. CaLDAS Diagnostics:
   1. Early (graphs).
   2. Tools:
      1. Error reduction ratio (estimation of |O-A|/|O-G|).
      2. The distance predicted-observed of screen-level variables (2m temperature and humidity) as a function of initial soil moisture value.
4. CaLDAS_Xsol Prototype.
5. Figures from preliminary comparisons between the original CaLDAS and CaLDAS_Xsol.

1) Nature of the problem:

1. This works intends to improve the analysis of soil moisture.

    - This part contributes to the energy and water fluxes land-atmosphere, influencing from the climate near the ground (temperature and humidity) to the formation of clouds (Ek, M.B., and A.A.M. Holtslag, 2004: Influence of Soil Moisture on Boundary Layer Cloud Development. *J. Hydrometeor.*, **5**, 86–99.)

    - Soil moisture data is currently inferred from readings of other variables, rather than direct, on the field, actual soil moisture measurements.

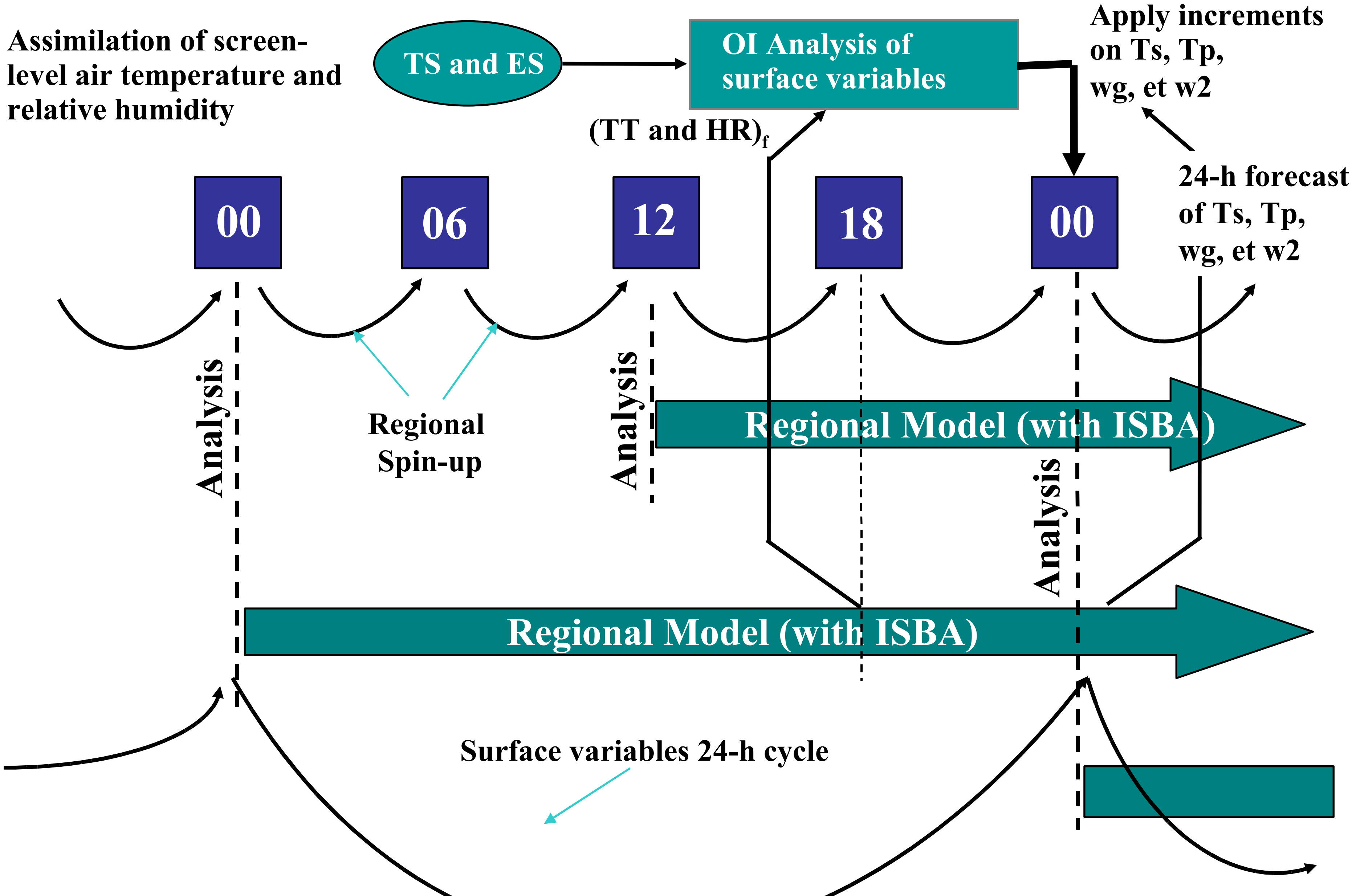
    - The three schemes shown here, Operational (in use), CaLDAS 2.0, CaLDAS_Xsol (or 3.0) derive the initial soil moisture and temperature for the model ISBA from the past discrepancy between predicted and observed screenlevel (2m) temperature and humidity.

    - One of the greatest challenges in techniques like this is how to avoid introducing corrections to the soil moisture that are not related to soil moisture (for example, when other effects on the observed variables become important when compared to the soil condition effect).

2. Example: operational.

Land Surface Surface Assimilation Currently Operational at CMC

Bélair et al. 2003



Meteorological Service of Canada
Environment Canada
RPN

The Canadian land data assimilation system (CaLDAS): an application to soil moisture

CaLDAS
Canadian Land Data Assimilation System

Who's in CaLDAS:

Stéphane BELAIR (CaLDAS co-PI, Hydros PI),
Godelieve DEBLONDE (CaLDAS co-PI, GRIP PM),
Gianpaolo BALSAMO (CaLDAS VF),
Pablo GRUNMANN (CaLDAS VF),
Yingxin Gu (GRIP VF),
Jean-François MAHFOUF (CaLDAS PM, CaPA PM)

Main features of CaLDAS

- Uses 24 hourly OI analyses of surface variables (observed 2m temperature and humidity), from 01z to 24z, of the day before (instead of only one at 18z as in the operational version).
- Simplified 2-DVAR technique to assimilate observations over a time window.
- Capability also for other observations such as satellite brightness temperatures.

Description of the original CaLDAS

(Presented December 12, 2005 by G.P. Balsamo)

- Off-line mode MEC-ISBA (ISBA-soil + MEC-first atmospheric model level) using operational GEMruns as forcing at the interface with the second model level.
- Guess run (24h) uses initial conditions given by previous day (first day uses the operational surface analysis).
- Two perturbations (+and - , centered on guess) are applied to the soil initial conditions and run for 24h to produce finite differences estimate of the Jacobian.
- SVA program calculates the correction to the soil variables.
- 1-step minimization of the cost function (limited to a vicinity of the guess to keep valid the simplifying assumptions).
- Application:
  - Gem-core: 15-km over north America and surrounding waters.

2D Variational surface analysis

Mahfouf (1991), Callies et al. (1998), Rhodin et al. (1999), Bouyssel et al. (2000), Hess (2001), Seuffert et al. (2003), Balsamo et al. (2004), Grunmann (2005).

The Cost function $J(\mathbf{x})$

$$J(\mathbf{x}) = \tfrac{1}{2}\,(\mathbf{x} - \mathbf{x}_b)^T\,\mathbf{B}^{-1}\,(\mathbf{x} - \mathbf{x}_b) + \tfrac{1}{2}(\mathbf{y} - H(\mathbf{x}))^T\,\mathbf{R}^{-1}\,(\mathbf{y} - H(\mathbf{x}))$$

Where:

$\mathbf{x}$	is the control variables vector
$\mathbf{y}$	is the observation vector
H	is the model operator
$\mathbf{B}$	is the background error covariance matrix
$\mathbf{R}$	is the observation error covariance matrix

The analysis is obtained by the minimization of the cost function $J(\mathbf{x})$ with respect to $\mathbf{x}$

Variational surface analysis: minimization

In the original CaLDAS, the minimum of the cost function J_{min} is analytically expressed setting to zero and solving for **x** the gradient expression:

$$\nabla_x J(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{x}))$$

(At the minimum, $\nabla_x J(\mathbf{x}) = 0$)

❑ TL hypothesis : $H(\mathbf{x}) \cong H(\mathbf{x}^b) + \mathbf{H}\cdot(\mathbf{x} - \mathbf{x}^b)$ allows to solve the above expression for **x**, and the solution is indicated as $\mathbf{x}^a$ (analysis)

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{y} - H(\mathbf{x}^b))$$

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y} - H(\mathbf{x}^b))$$

- For high dimensional problems: TL/AD models
- For low dimensional problems: finite differences or Monte Carlo.
- <u>For one dimensional, limited search space: 1-D search techniques.</u>

Limitations

1. ***Simplifying hypotheses (Balsamo, 2005).***
   - ***TL*** (tangent linear approximation).
   - ***2D*** (horizontal decoupling).
   - ***Not 4D**** (reduced control variable space).

*within the assimilation time-window.

- ***Single-step solution relies entirely on the slope of J(x) in a close vicinity of $x=x^b$.***

*Analogous to the scalar problem of finding the minimum of J(**z**).*

Where: $J(z) = \frac{1}{2} [B^{-1}z^2 + R^{-1}(y - H_b + H'.z)^2]; \quad z = x - x^b.$

For B,R, y, H_b *and* H' constants with changes in z (single-step solution), the above function is simply quadratic in z.

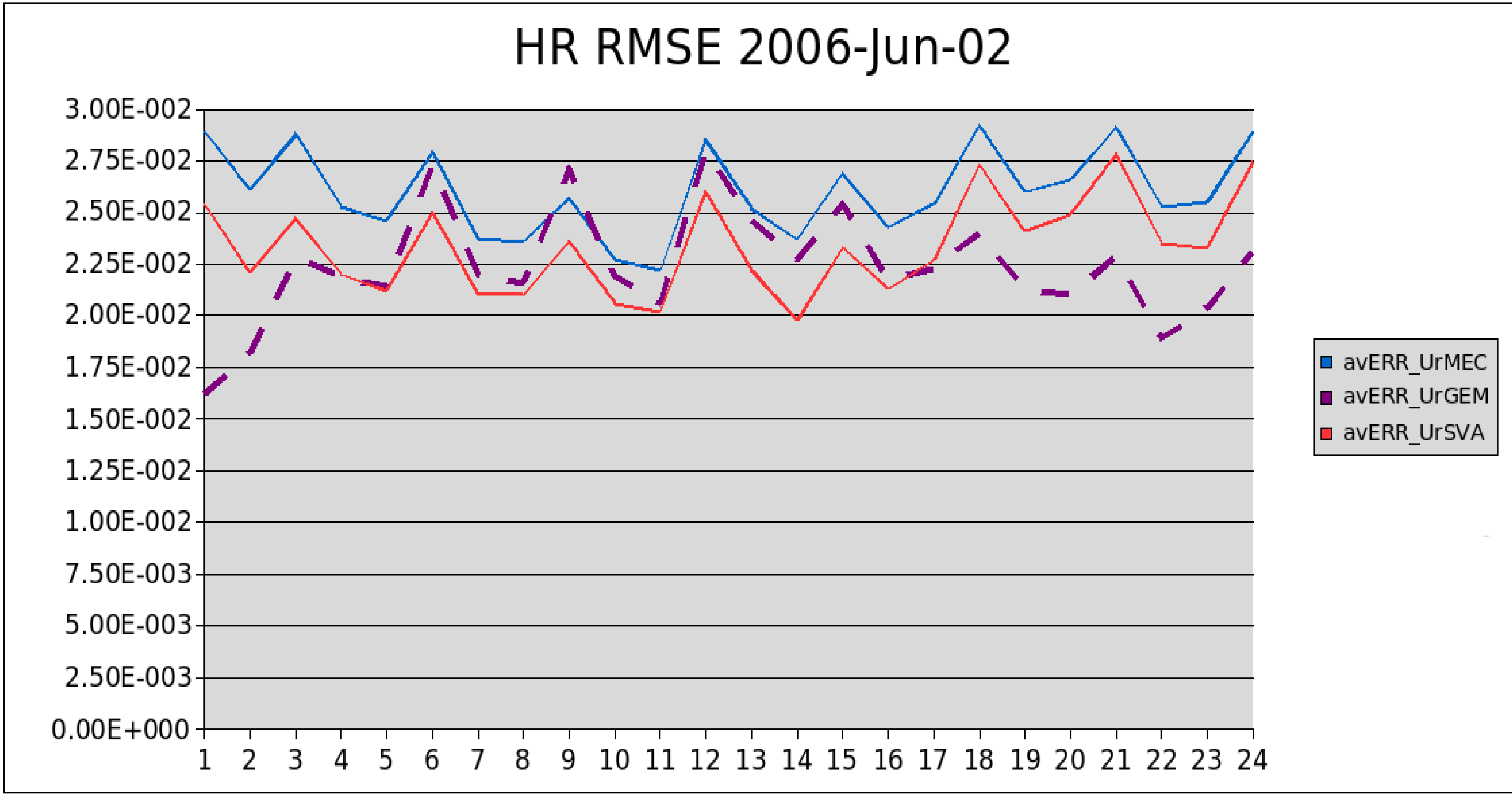
$F(z) = A(2).z^2 + A(1).z + A(0)$; with A(i) constants.

$F'(z) = 2A(2).z + A(1) = 0 \Rightarrow z = -A(1) / [2.A(2)]$.

➔ The solution points to the vertex of the F(z) parabola.

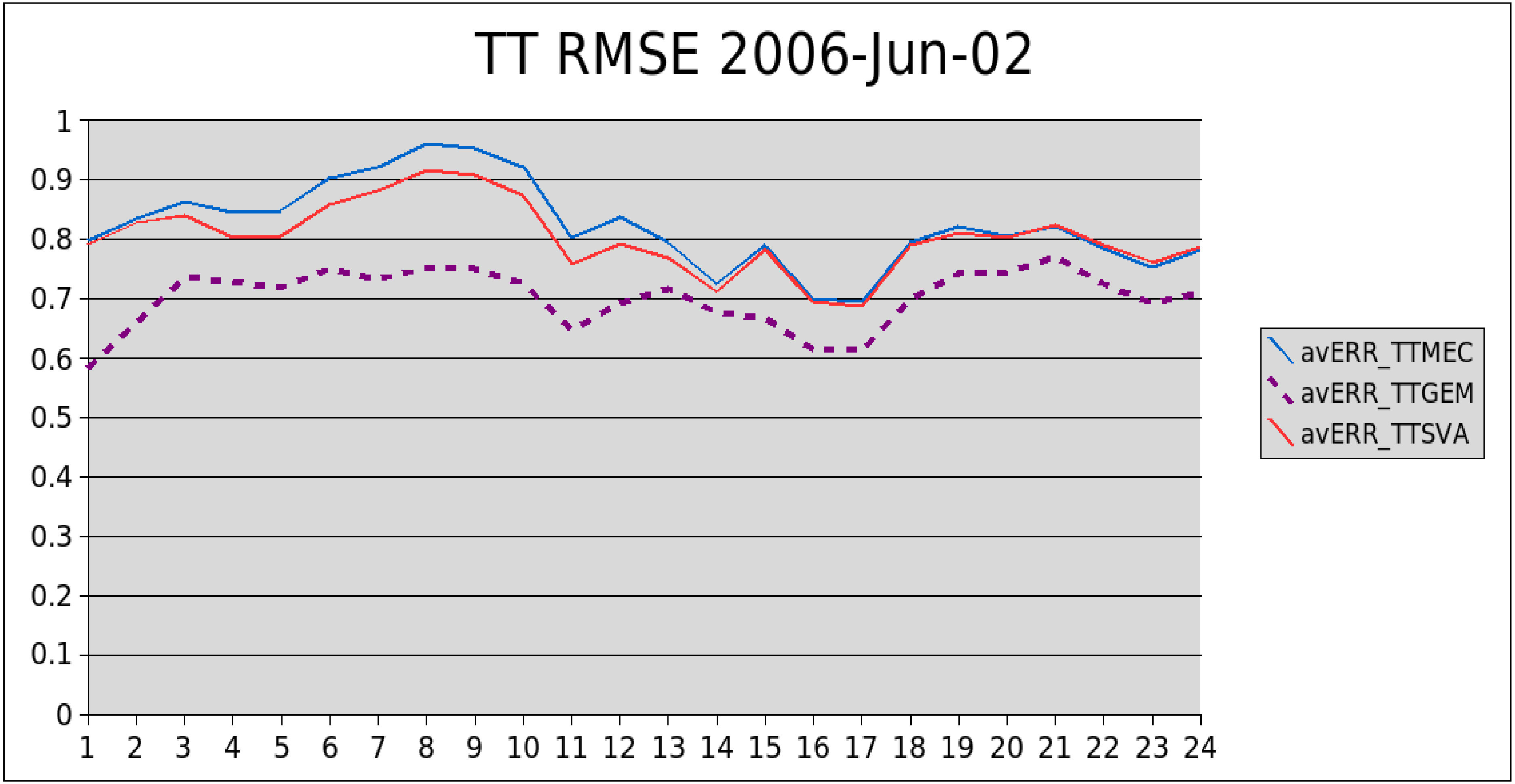
RMSE over N. America

Obs from OI having the guess run as background for MEC and SVA but for GEM, uses the same GEM run as background.



RMSE over N. America

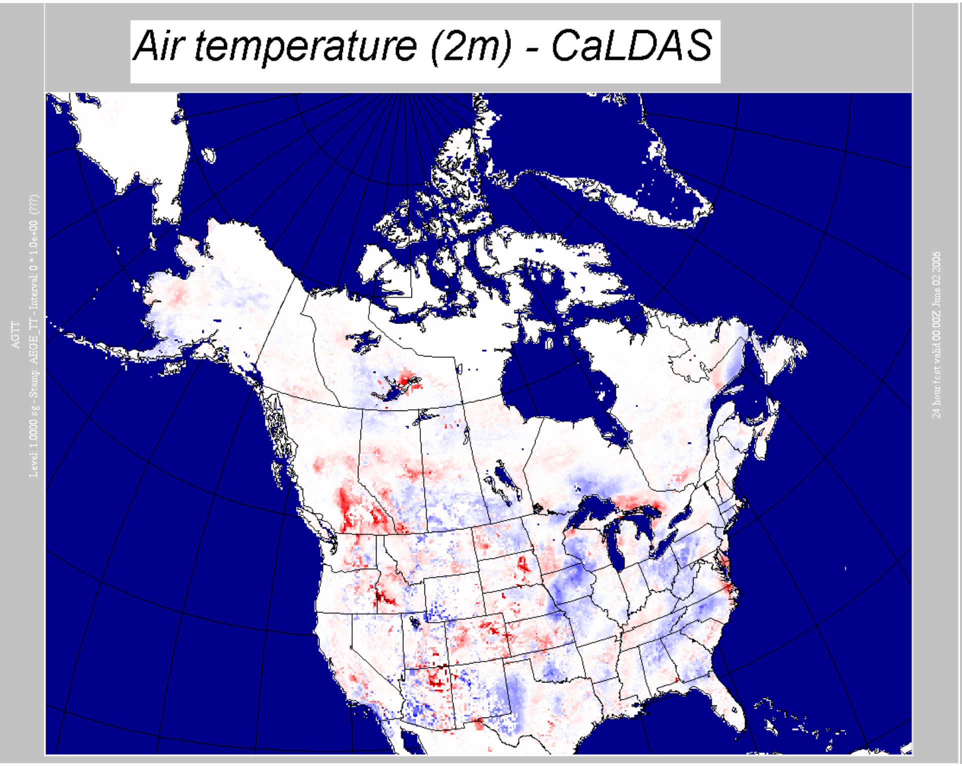
Obs from OI having the guess run as background for MEC and SVA but for GEM, uses the same GEM run as background.

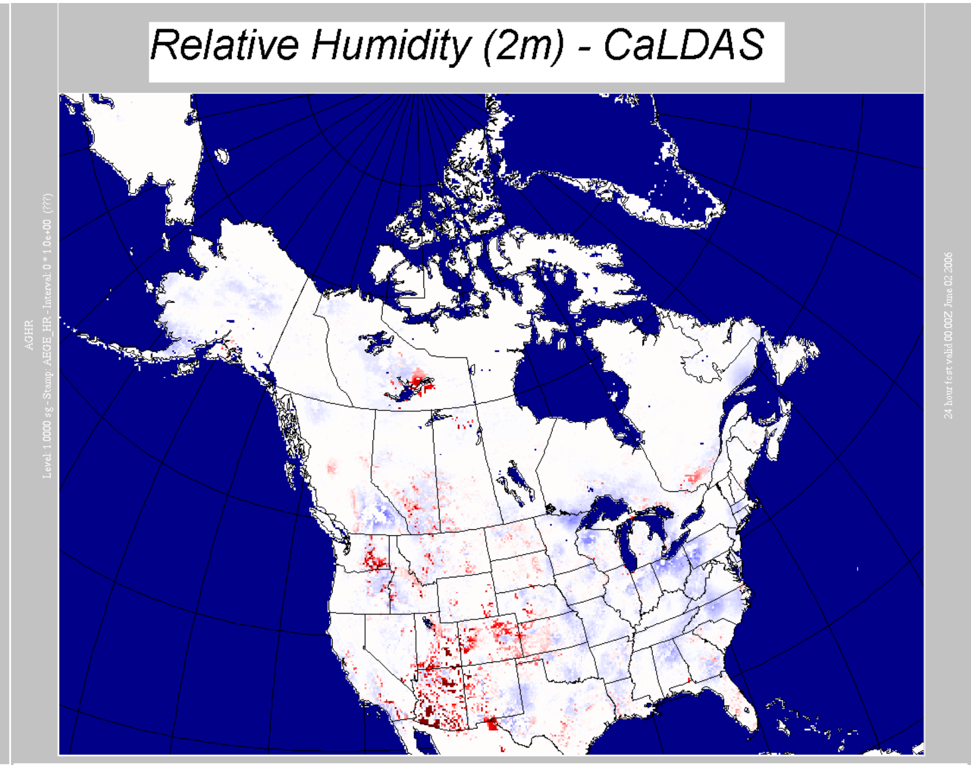


Estimate of |O-A|/|O-G|

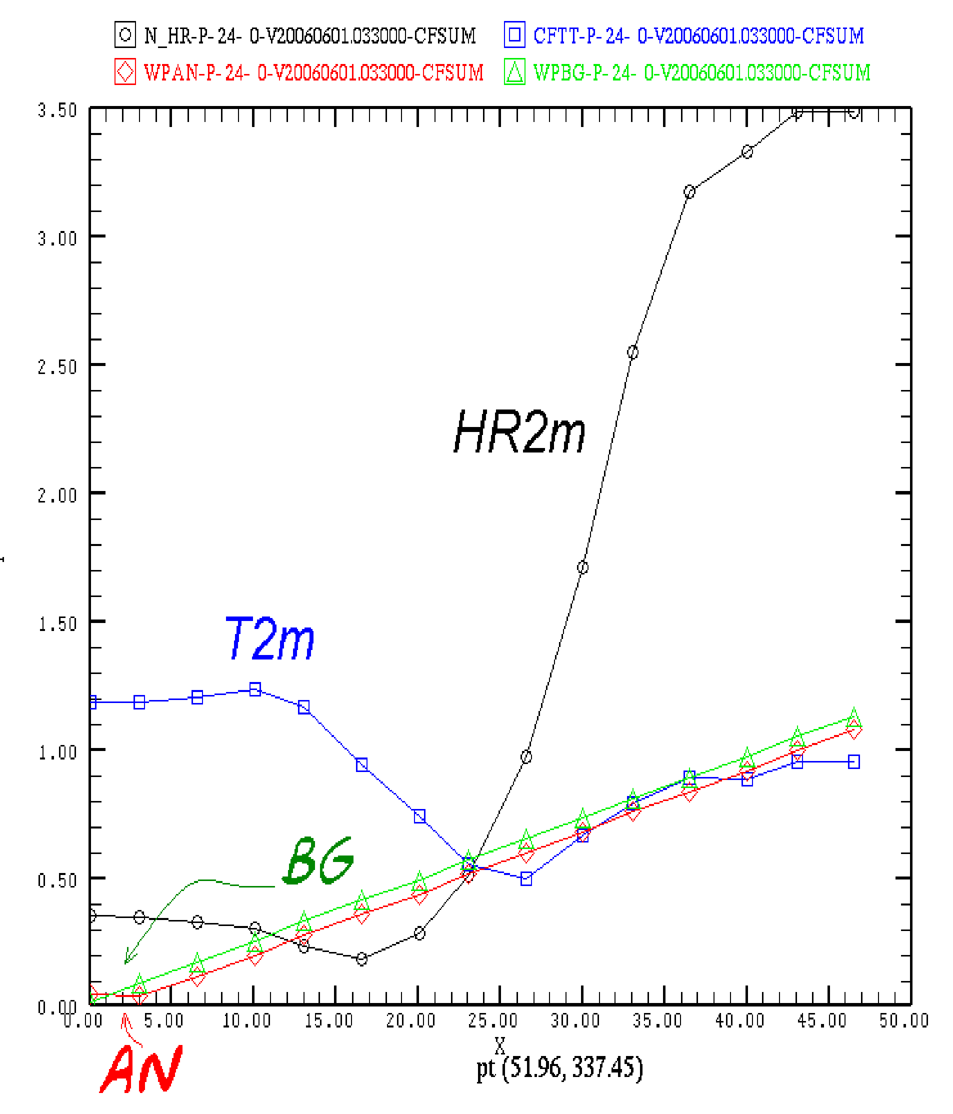
Impact of Analysis on error

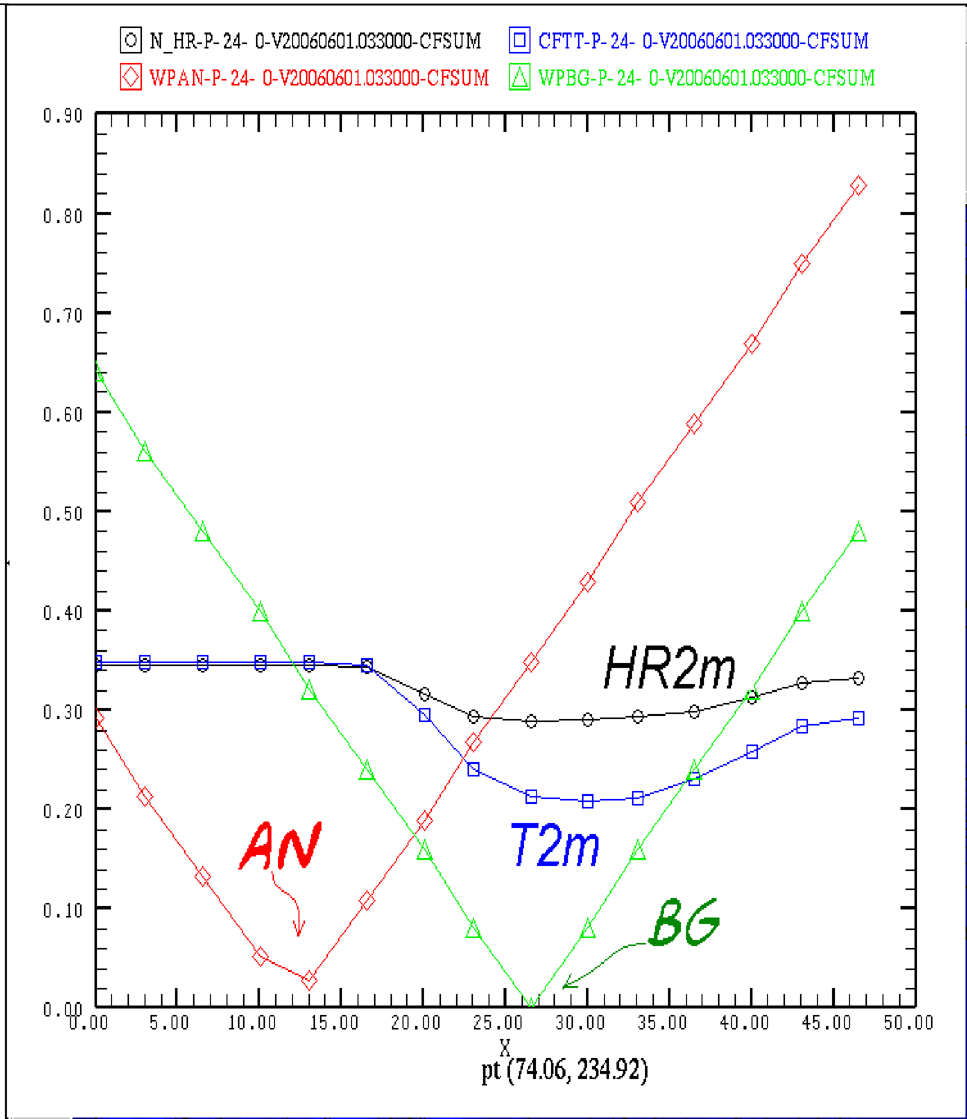
(“After”/“Before” ratio estimate)



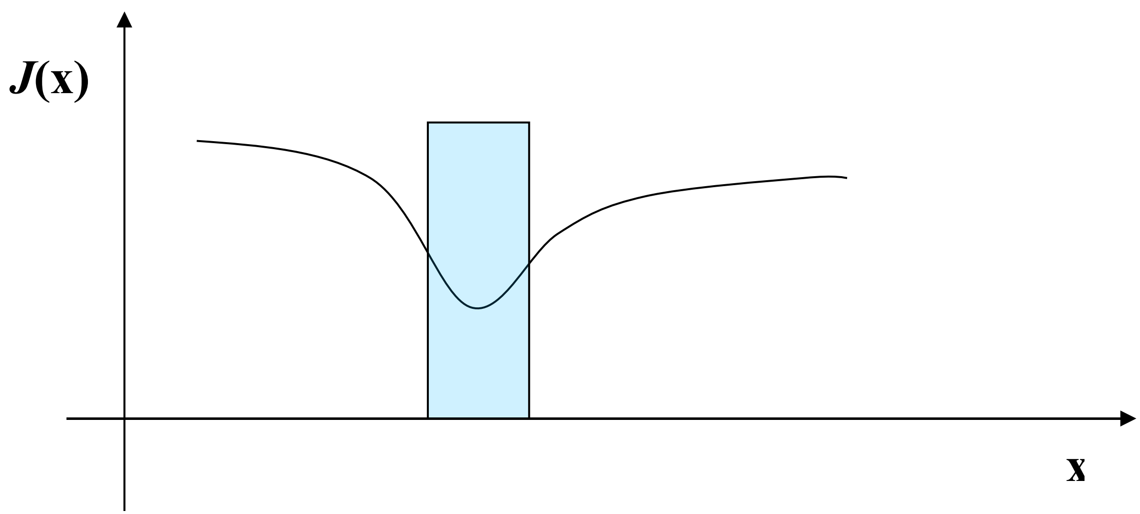


Study of CaLDAS behavior with respect to the cost function





- If the guess value is outside a region that can be approximated as a parabola, then the single step solution may not point to the minimum:



- In the case when x is soil moisture, this is quite frequent.

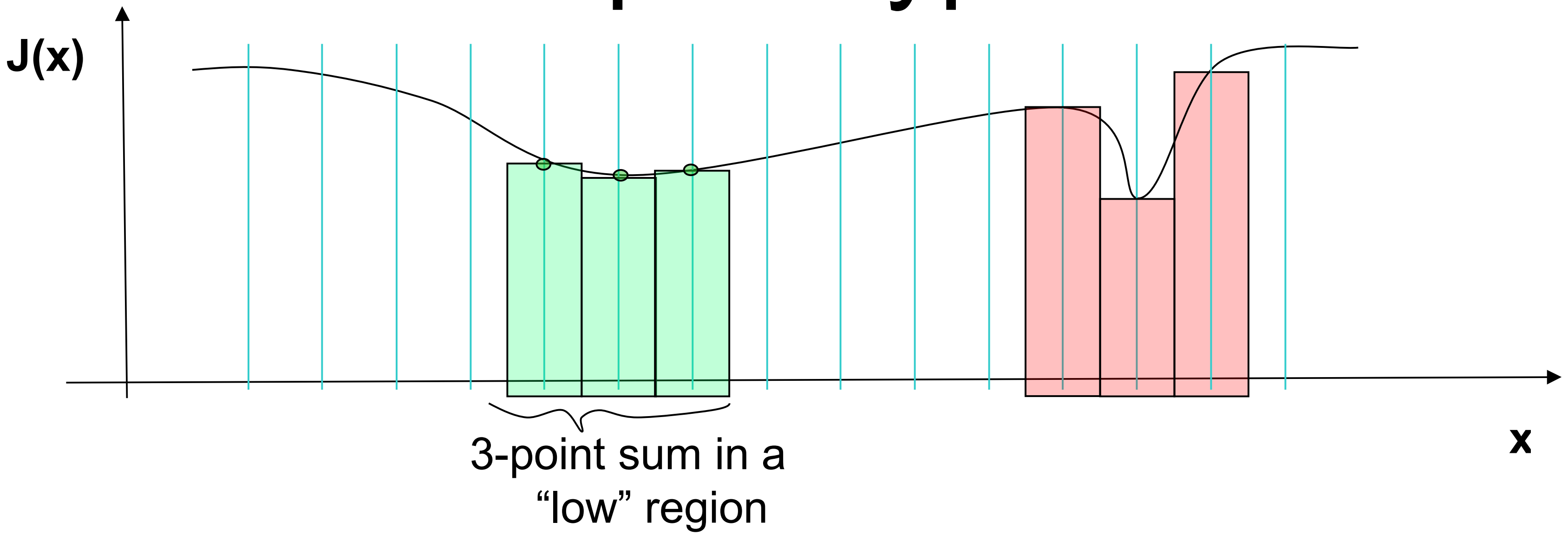
The proposed new method

- In the case of soil moisture, the single step minimization used in CaLDAS would not necessarily lead in the right direction.
- The control variable space in this particular case of soil moisture in MEC-ISBA is quite limited in range (given the physical constrains and precision) and one-dimensional (MEC-ISBA calculations are single-column).
- The above suggest the possibility of using a different method for searching the minimum.

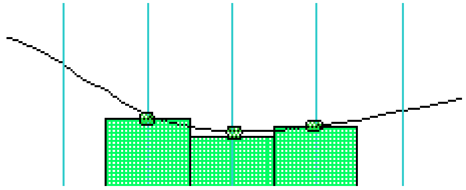
The proposed new algorithm (first prototype)

- The searched value is between dry (no less than 0.0) and saturated (around 0.5).
- Launch 15 runs with initial soil moisture at regular intervals from near dry to saturation.
- This allows to produce a complete exploration of the cost function.
- The search for the minimum starts by selecting the three consecutive points with the lowest sum (a low region)
- If these three points can fit a parabola having a minimum between them, this minimum is the solution. If not, the lowest of the three points is the solution.

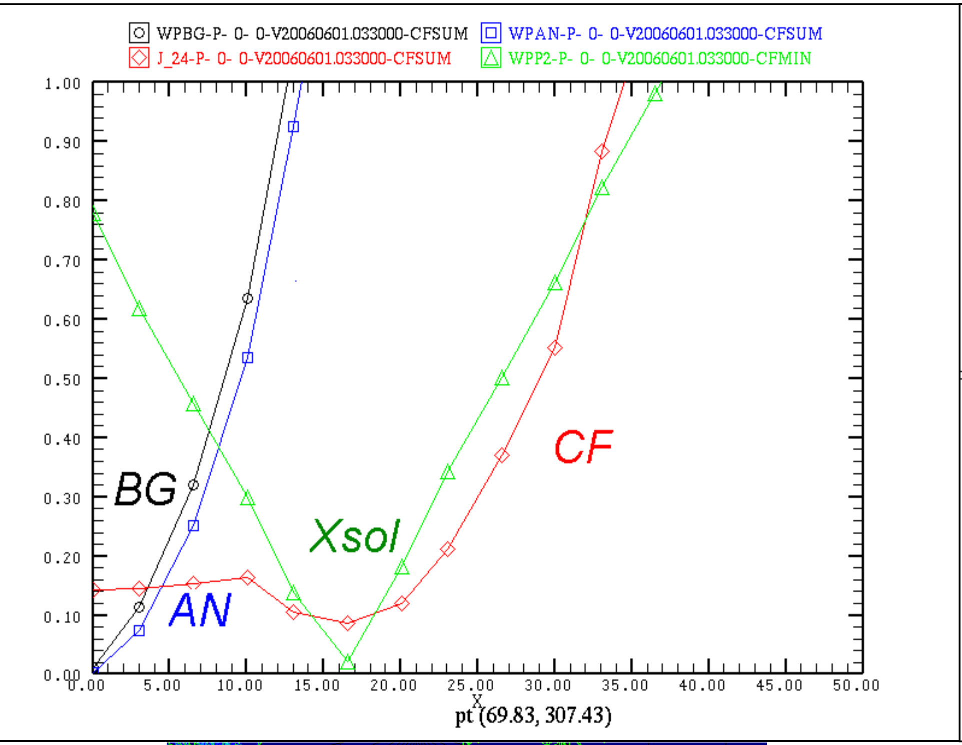
The first prototype solution

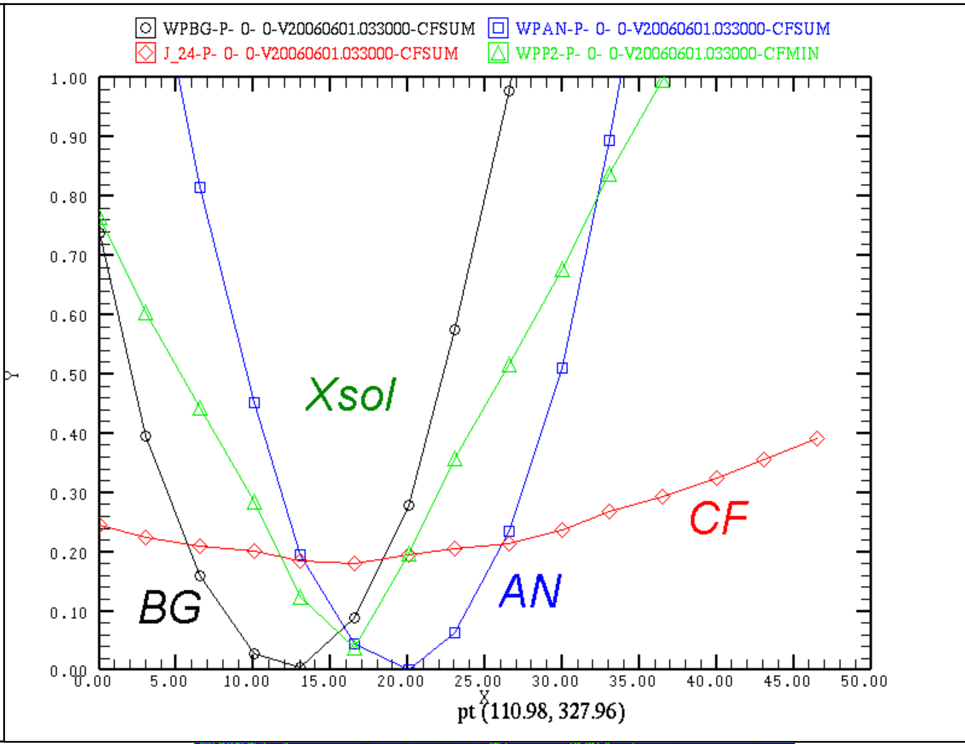


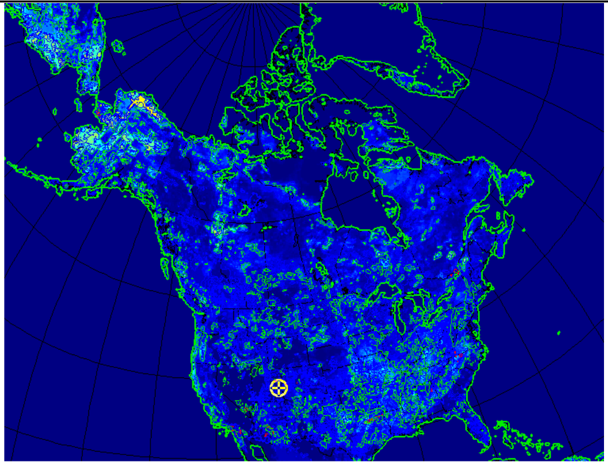
The parabolic minimum is applied to the three consecutive lowest points

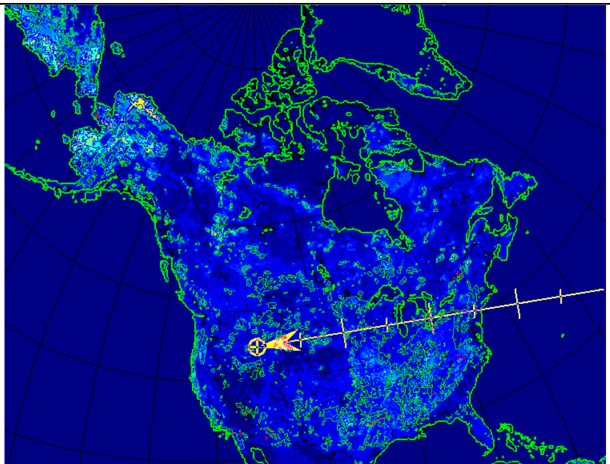


Preliminary tests

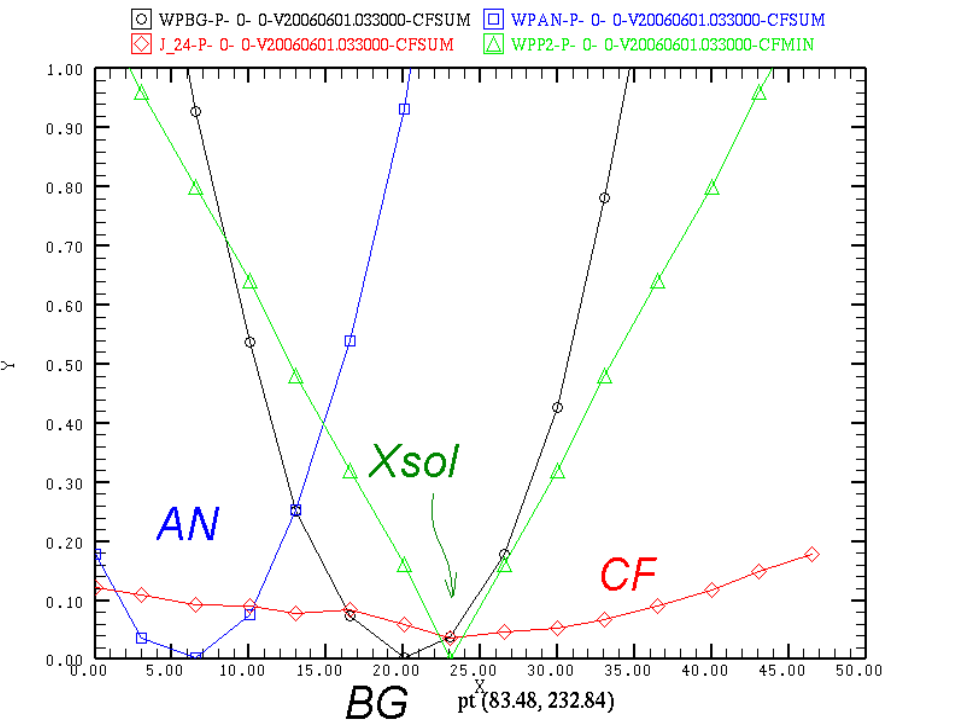


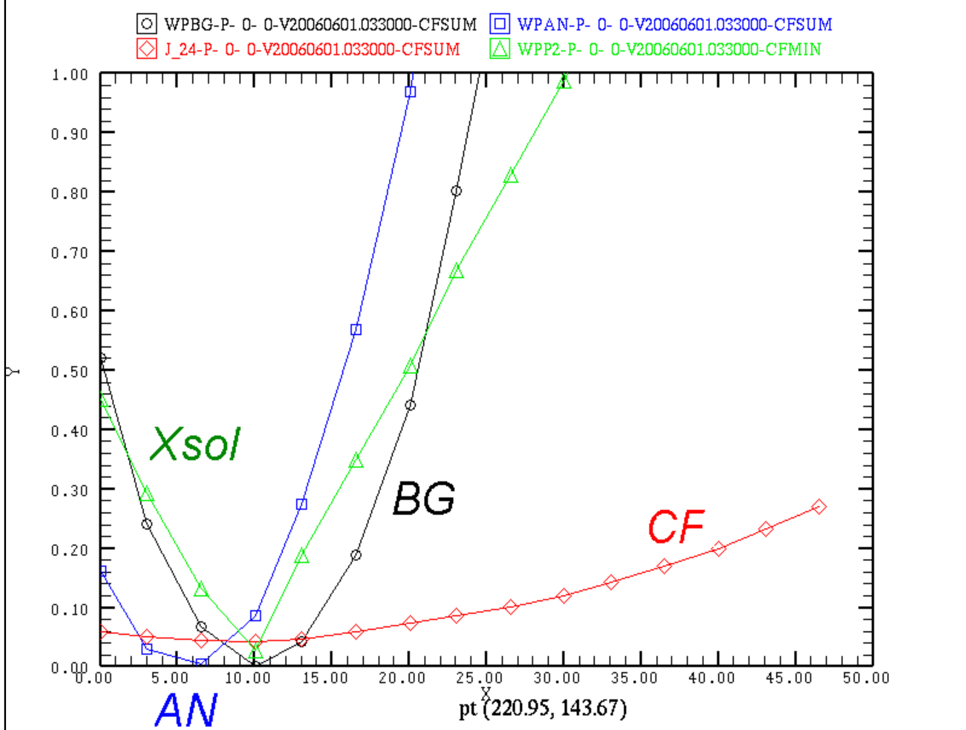


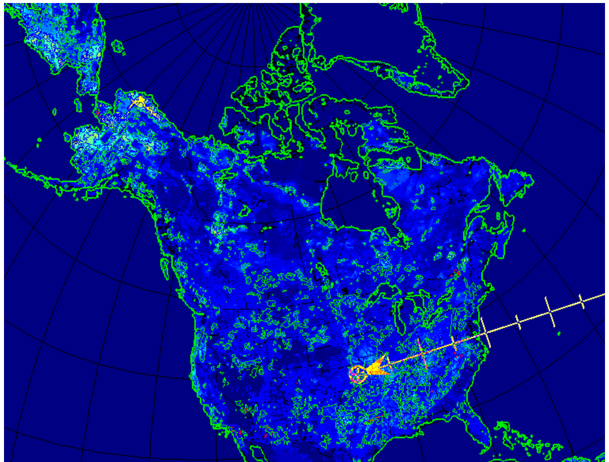


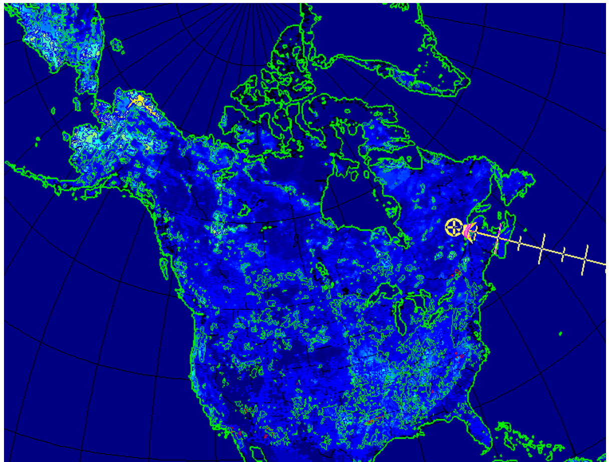


Preliminary tests

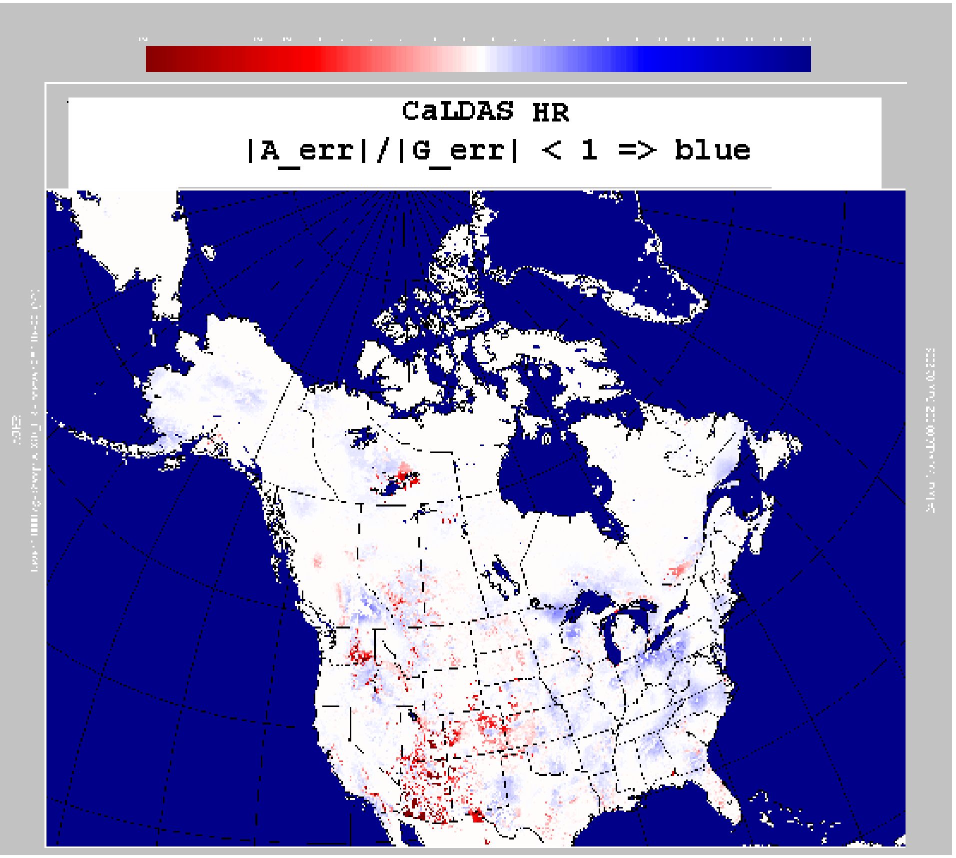


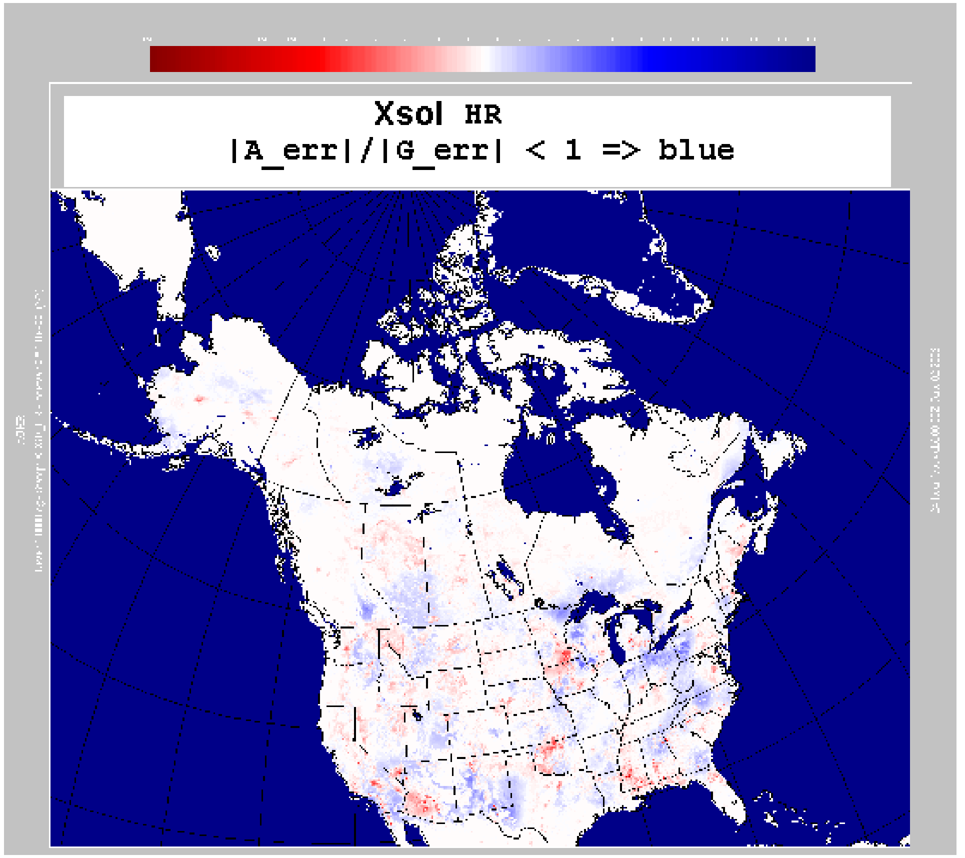






Preliminary tests and comparisons





Preliminary tests and comparisons