Assimilation of Plasma Density Measurements Into the Dynamic Global Core Plasma Model (DGCPM)

A. M. Jorgensen
New Mexico Tech

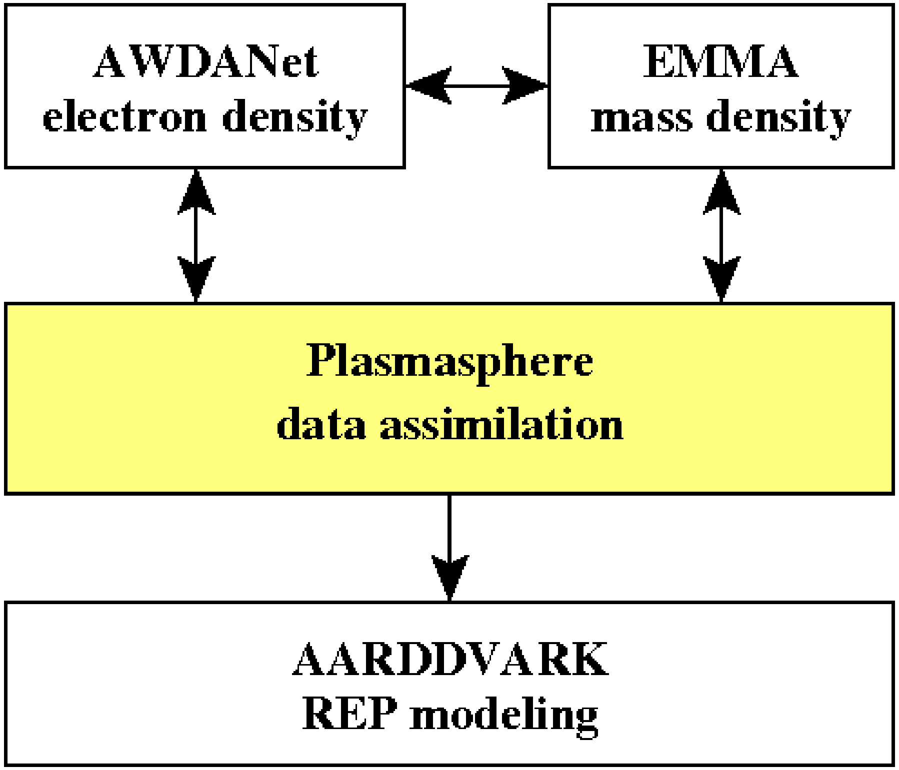
A. J. Ridley, A. M. Dodger, P. J. Chi, J. Lichtenberger, M. B. Moldwin, D. Ober, A. Boudouridis, J. Duffy

2011 AGU Fall Meeting, San Francisco, CA
SM52A-04

Outline

- PLASMON collaboration
- Data assimilation
  - Ensemble Kalman Filter
  - Particle Filter
- Example
- DGCPM + in-situ observations
  - DGCPM
  - Data Model Comparison
  - Assimilation
- Technology
- Conclusion

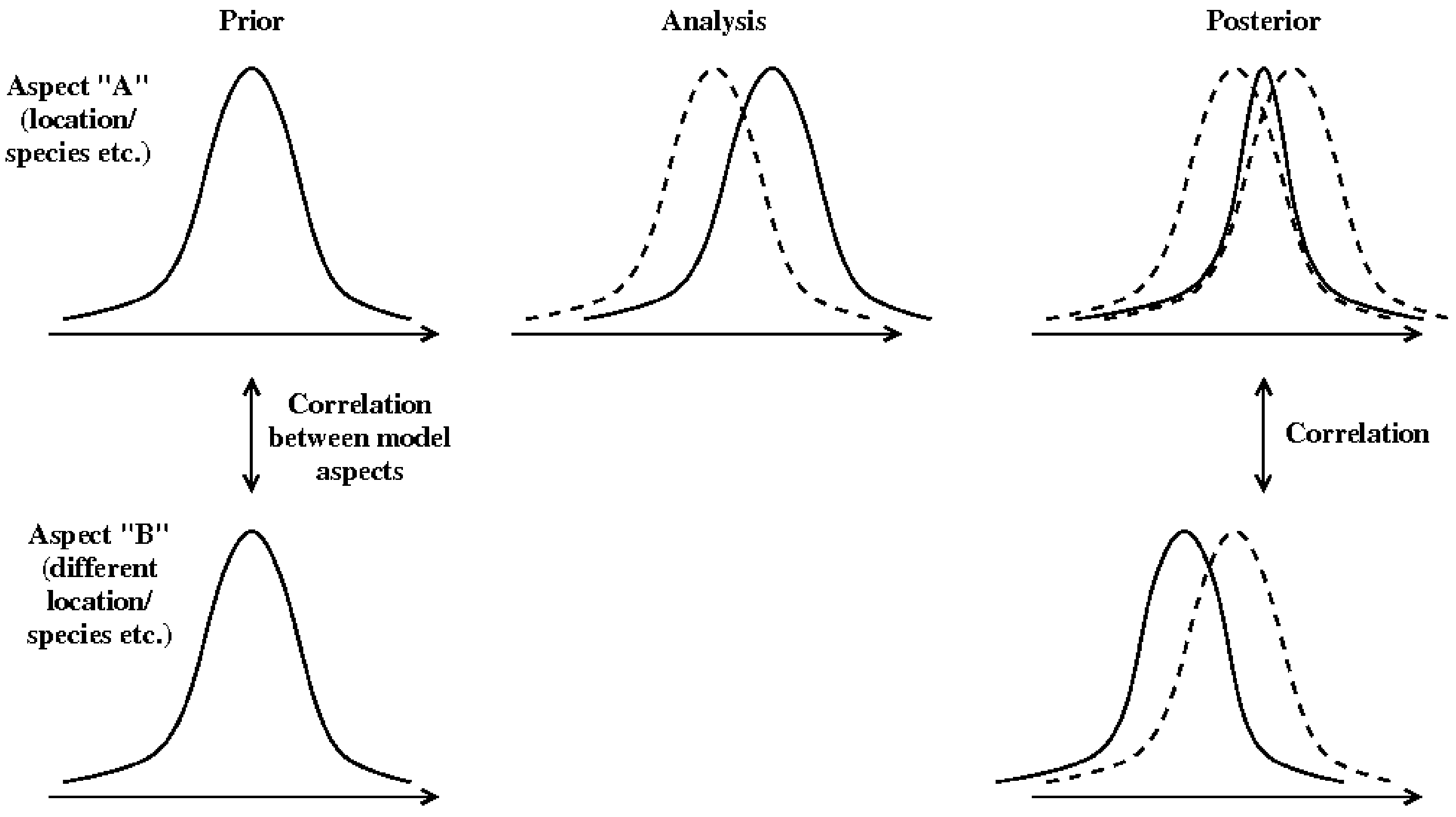
PLASMON collaboration



- Funded by EU FP-7
- 11 institutions
- Expand magnetometer arrays for FLR
- Expand VLF arrays and add automatic whistler analysis.
- Assimilate these into Plasmasphere model to get plasmapause location.
- Expand the AARDDVARK network.
- Use assimilated observations as context for REP studies.

Data Assimilation

Combining observations and a model provides a more complete and more accurate picture of the system state. Makes it possible to estimate system aspects not measured because they may be coupled through the model to other aspects which are measured.



Ensemble Kalman Filter

- Kalman filter
  - State is a vector, all operations are linear, and all uncertainties are Gaussian, represented by a covariance matrix

$$\psi = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_m \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & \sigma_2^2 & & \sigma_{2m} \\ \vdots & & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_m^2 \end{bmatrix} \quad m \times m$$

  - Advancing: based on data

$$\psi_k \longrightarrow \psi_{k+1} \qquad \Sigma_k \longrightarrow \Sigma_{k+1}$$

  - Huge arrays:

$$m = 200 \times 200 = 40000 \quad \Rightarrow m \times m = 40000^2 = 1.6 \times 10^9$$

  - Assumes linearity: $\psi_c = \alpha\,\psi_a + \beta\,\psi_b$ is also a valid solution

Ensemble Kalman Filter

- Ensemble Kalman filter: sample of the Kalman Filter

$$\boldsymbol{A} = \begin{bmatrix} \psi_{11} & \psi_{12} & \cdots & \psi_{1N} \\ \psi_{21} & \psi_{22} & & \psi_{2N} \\ \vdots & & \ddots & \vdots \\ \psi_{m1} & \psi_{m2} & \cdots & \psi_{mN} \end{bmatrix}$$

Each column of A is a running model. Each model runs with different parameters/drivers/ noise.

- Posterior ensemble is linear combination of prior ensemble

$$\boldsymbol{A}_{\text{posterior}} = \boldsymbol{A}_{\text{prior}} \times \boldsymbol{X}$$

... and we can compute any statistics we want from N values of each cell in each model.

$$\boldsymbol{A} : \boldsymbol{m} \times \boldsymbol{N} = \mathbf{40000} \times \mathbf{100} = \mathbf{4} \times \mathbf{10^6}$$

$$\boldsymbol{X} : \boldsymbol{N} \times \boldsymbol{N} = \mathbf{100^2} = \mathbf{10^4}$$

Much smaller matrices

Computing the Ensemble Transformation

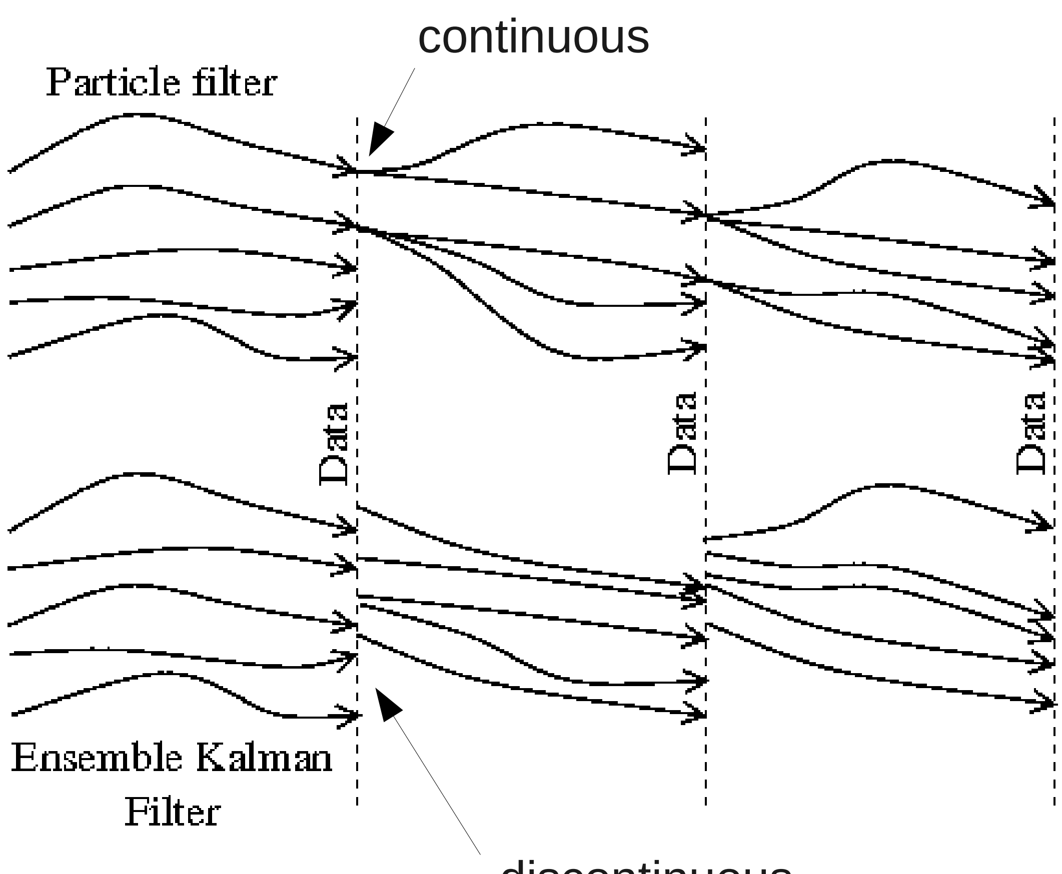
X in

$$\boldsymbol{A}_{\text{posterior}} = \boldsymbol{A}_{\text{prior}} \times \boldsymbol{X}$$

.... it's not important, let's skip it – but it is based on the difference between data and model, and on the assumption of Gaussian statistics.

Particle Filter

EnKF filter assumes linearity (Any linear combination is itself a valid solution). Not generally true. Negative plasma density is – in principle - possible.



$$\boldsymbol{A}_{\text{posterior}} = \boldsymbol{A}_{\text{prior}} \times \boldsymbol{X}$$

For a particle filter X contains
only 1s and 0s, sampling
The models at the time of
Observation based on the
Probability of the models in
The face of observation –
Using any probability distribution
We like.

Model Noise

- To get models to diverge to explore model space we add noise. Multiple ways to do it.
- Here we choose “driver parameters” modeled as red noise

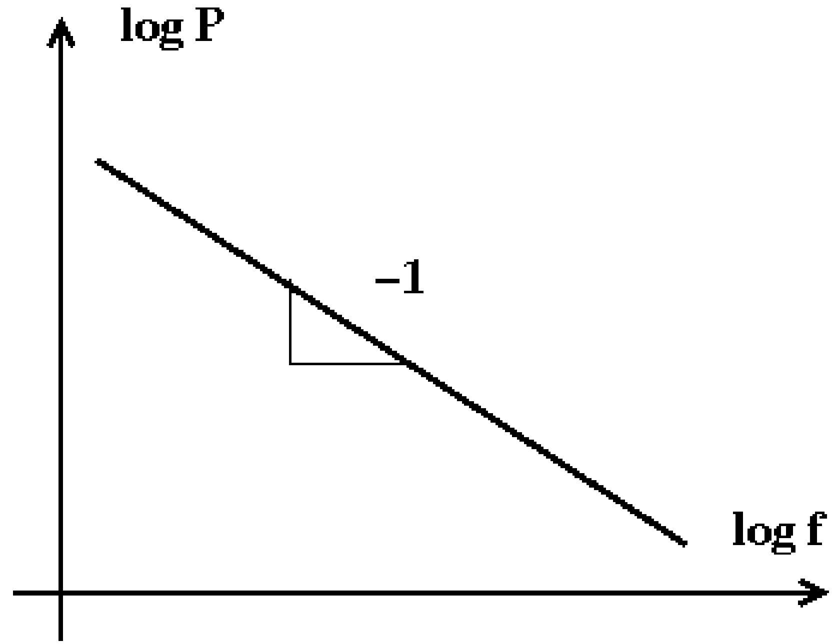
$$q_k = \alpha q_{k-1} + \sqrt{1-\alpha}\, w_k$$

White noise

$$\psi^* = \begin{bmatrix} q \\ \psi_1 \\ \psi_2 \\ \vdots \\ \psi_m \end{bmatrix}$$

Driver parameter (E-field)

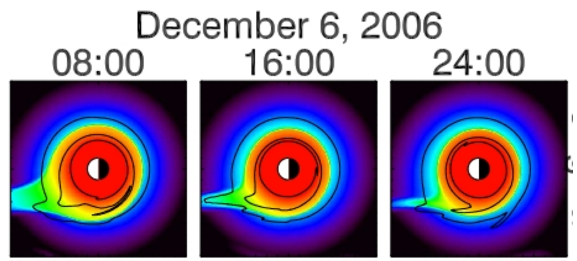
Evolve according to model equations



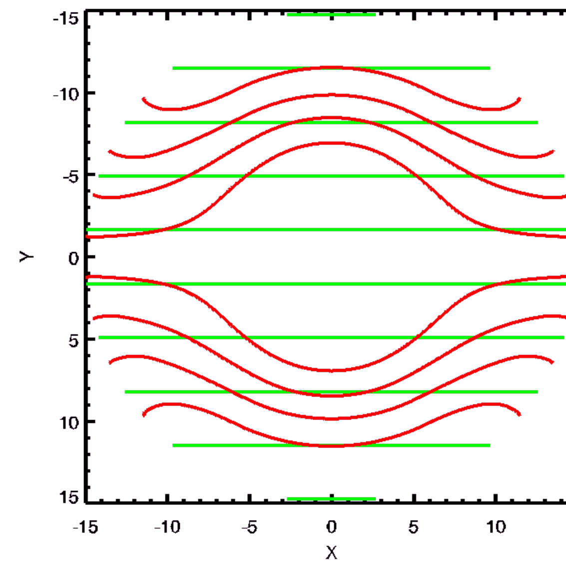
Evensen (2003)

DGCPM + In-situ measurements

- Dynamic Global Core Plasma Model (e.g. Ober [1997])
- Data-model comparison
- Assimilation



Gallagher et al. (1995)



Sojka et al. (1986)

$$F_n = -\frac{NB_i}{\tau}$$

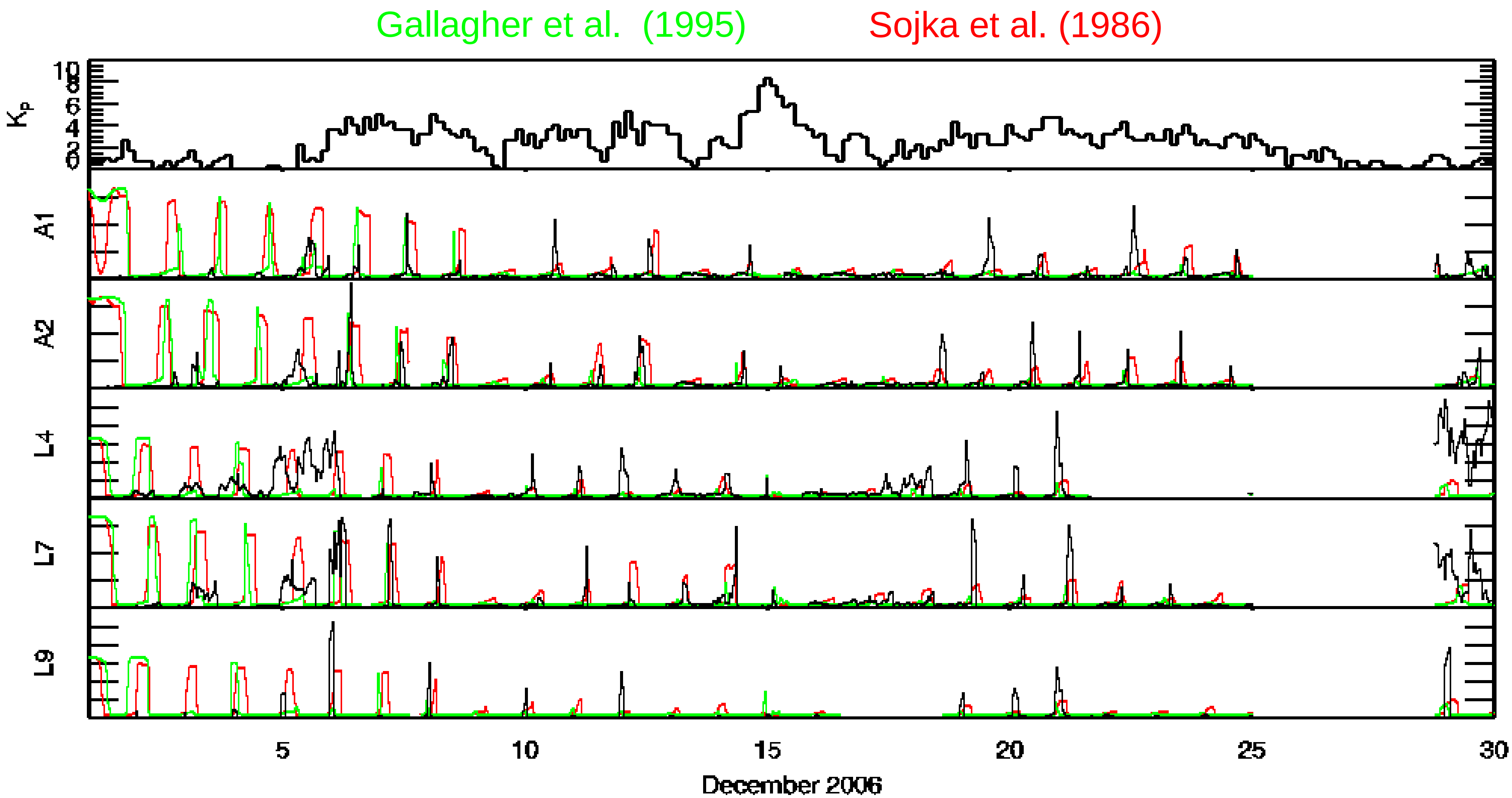
$$F_d = \frac{n_{\mathrm{sat}} - n}{n_{\mathrm{sat}}} F_{\mathrm{max}}$$

$$\vec{B}(\vec{r})$$

$$\vec{E}(\vec{r}) \rightarrow \frac{D_\perp N}{Dt} = \frac{F_N + F_S}{B_i}$$

Global E-field is aspect that we would like to determine but can't measure. Or plasmapause location is what we want to determine. These will be based on other measurement and a parametrized E-field model.

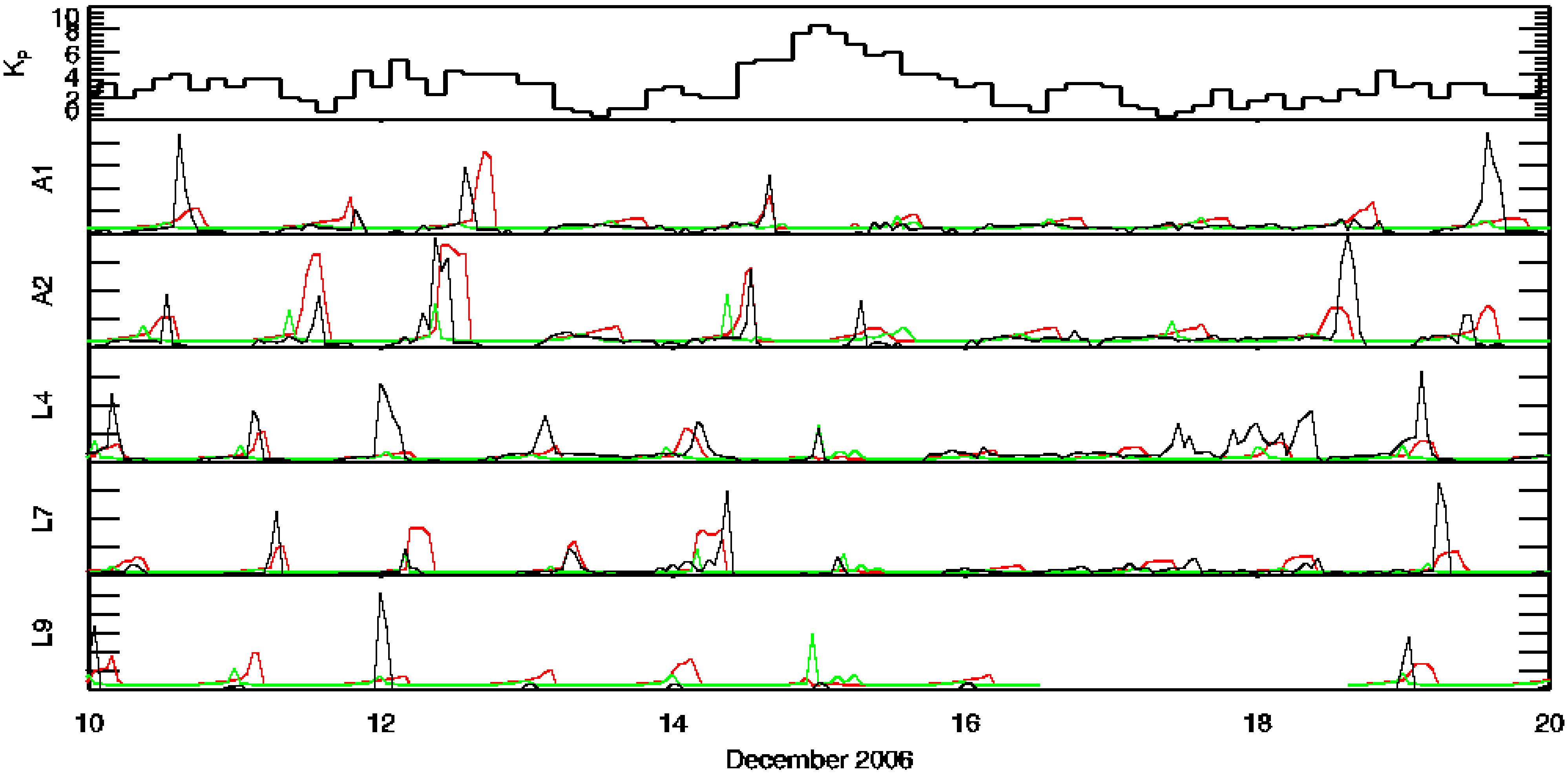
Data/Model Comparison



Enough similarities that there is some hope that data assimilation is possible. Enough differences that data assimilation is likely needed. Data uncertainties and model noise must be chosen carefully.

Data/Model Comparison

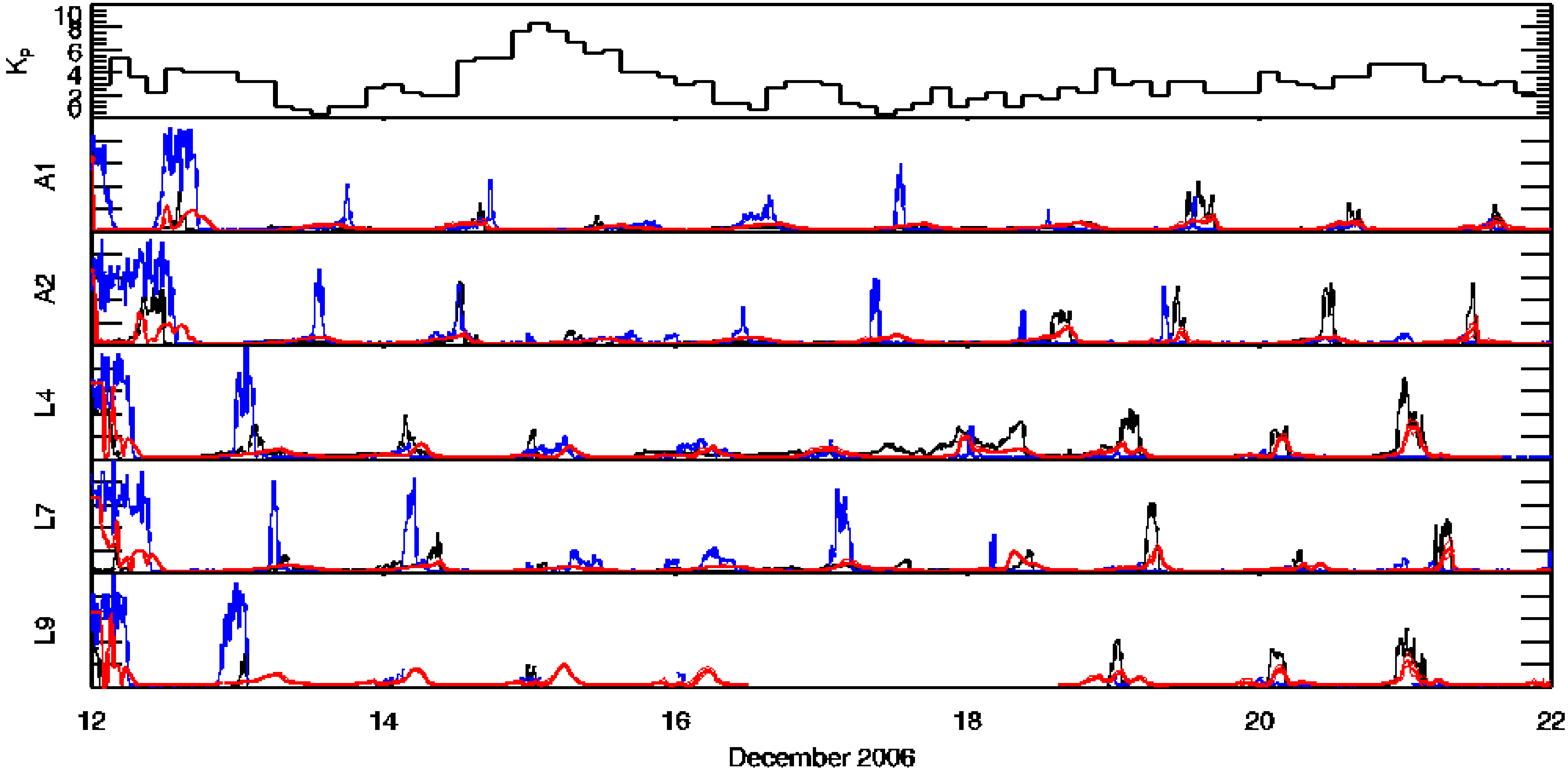
Gallagher et al. (1995) Sojka et al. (1986)



Enough similarities that there is some hope that data assimilation is possible. Enough differences that data assimilation is likely needed. Data uncertainties and model noise must be chosen carefully.

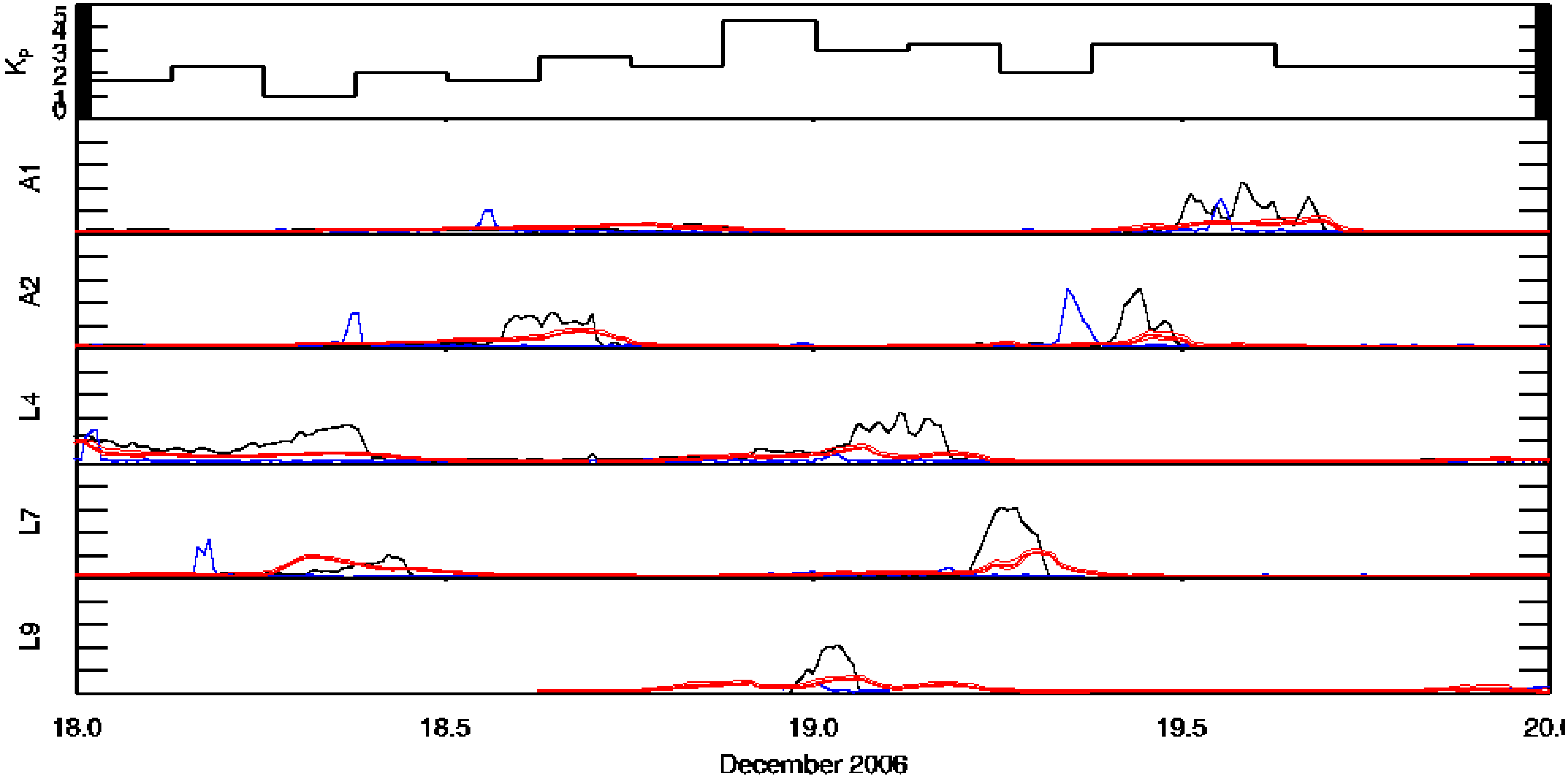
Assimilation

Can we drive the model with the simple parametrized electric field (Sojka, 1986) and improve the agreement with LANL in-situ observations?



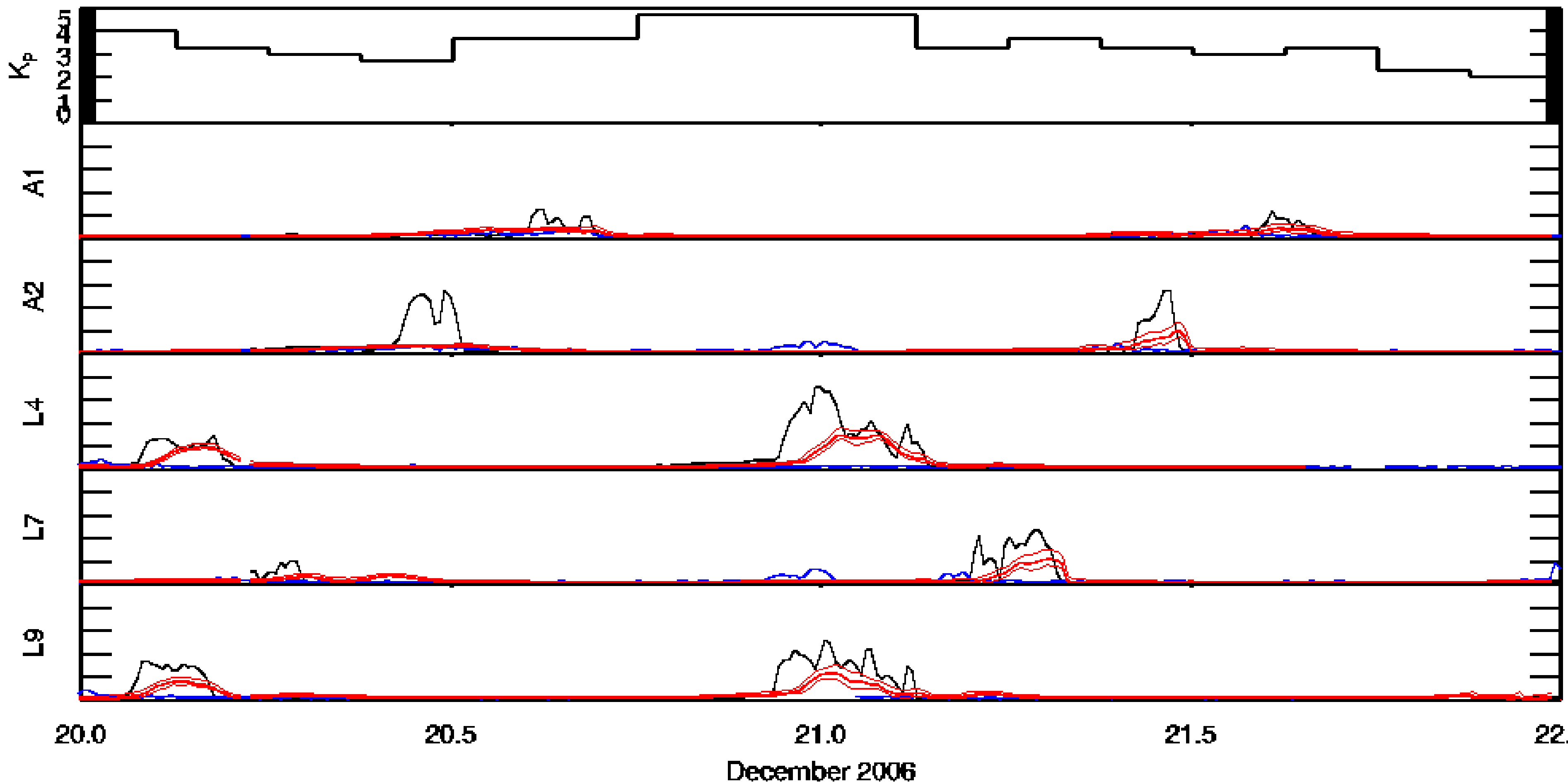
Assimilation

Can we drive the model with the simple parametrized electric field (Sojka, 1986) and improve the agreement with LANL in-situ observations?

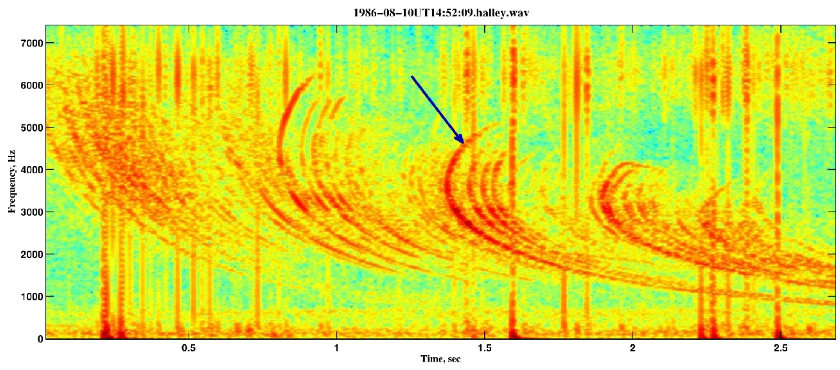


Assimilation

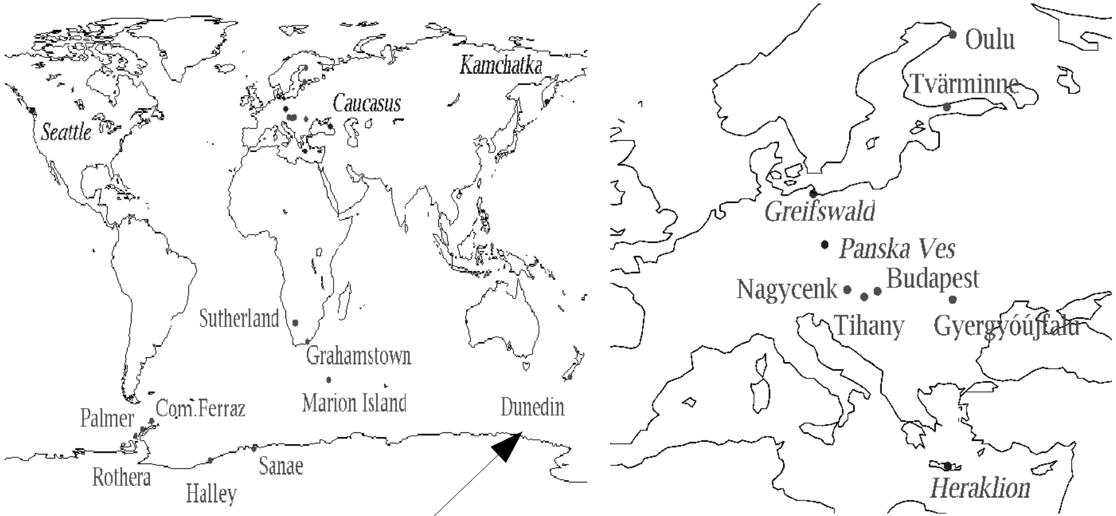
Can we drive the model with the simple parametrized electric field (Sojka, 1986) and improve the agreement with LANL in-situ observations?



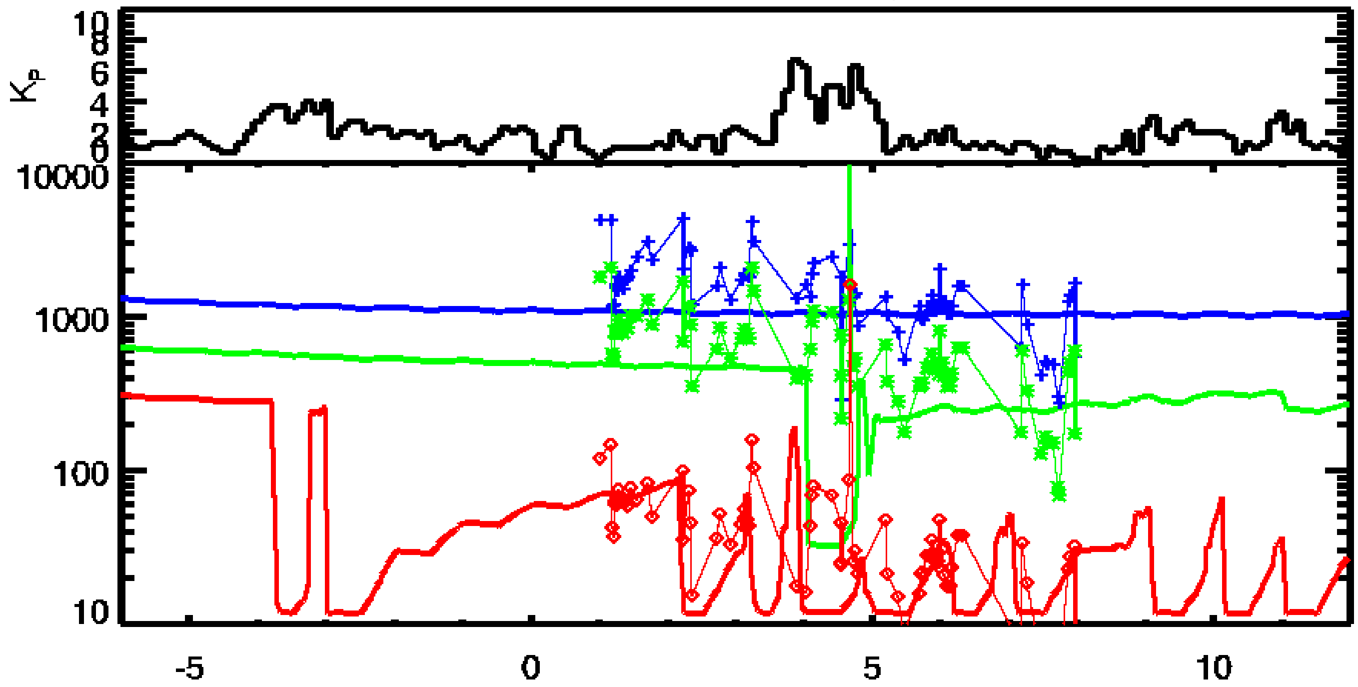
Number density from VLF Whistlers



Lichtenberger (2009)



Dunedin, NZ, Jul/Aug 2010



20 stations

L-shell slice of electron density from each station.

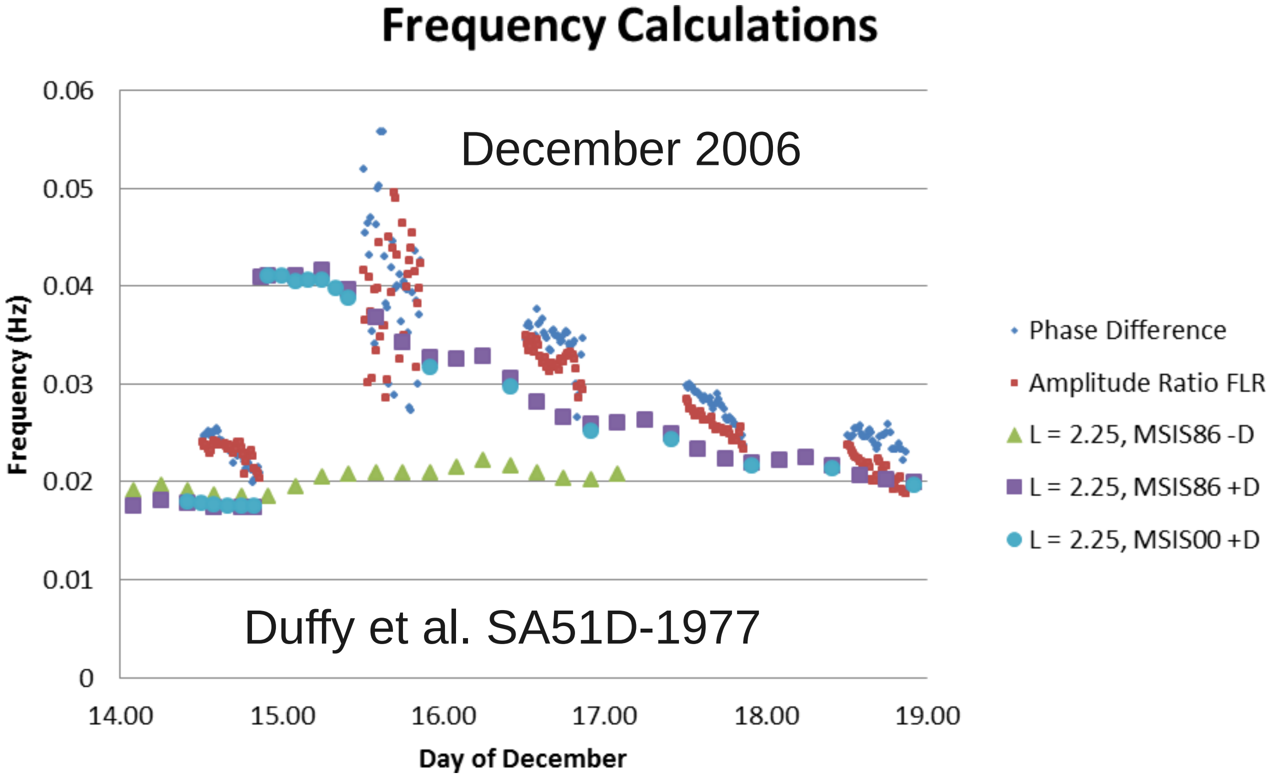
Automatic whistler analysis at every station.

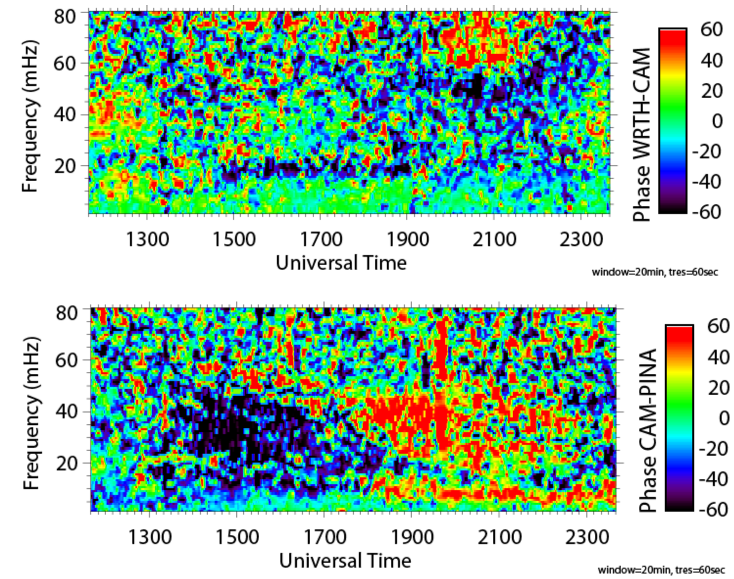
L=2.5

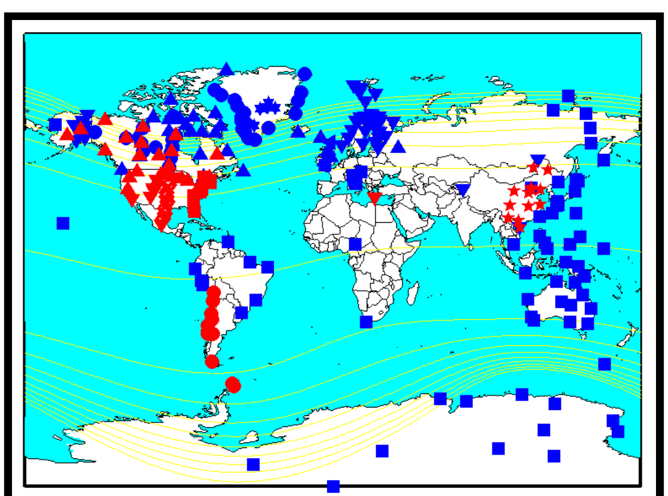
L=3.5

L=4.5

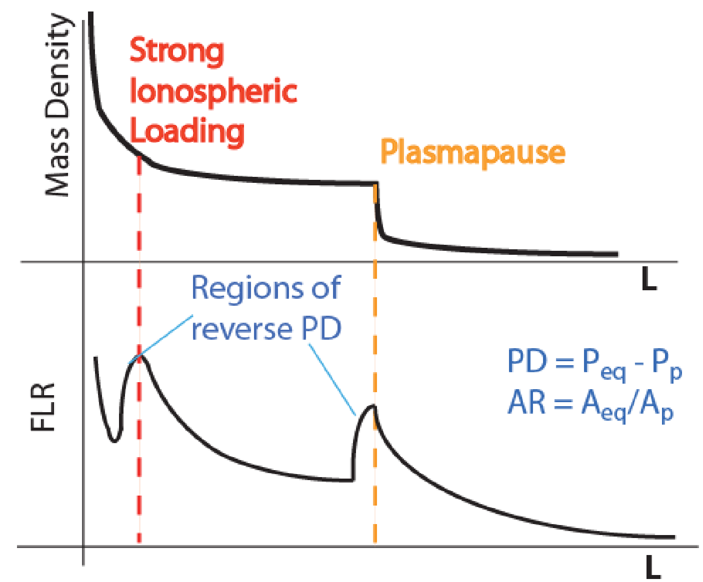
Mass Density from FLR







ULTIMA



Zesta et al. SA43B-06

Summary/Future

- Ensemble Kalman Filter + LANL In-situ works – somewhat
- Need better field model and more data
  - FLR mass density
  - Whistler number density
- GPGPU computing – 100x speedup