

**Automatic Enhancement and Detection  
of *IMAGE* Radio Plasma Imager Echoes:**

**1. Derivation and Models**

*Fall 1999, AGU Poster SM12B-05.*

**W. W. L. Taylor, M. L. Rilee, S. A. Boardsen<sup>1</sup>**

**J. L. Green<sup>2</sup>, B. W. Reinisch<sup>3</sup>**

<sup>1</sup>*Raytheon ITSS*, <sup>2</sup>*NASA/GSFC*

<sup>3</sup>*University of Massachusetts-Lowell*

*Contact:* 301-286-4743, Michael.L.Rilee@gsfc.nasa.gov.

*Presented at the*

**1999 Fall Meeting of the AGU, Dec. 13-17, San Francisco, CA.**

**Moscone Center, Hall D**

**1999-12-13**

# Abstract

---

The Radio Plasma Imager (RPI) is a low power radar on board the IMAGE spacecraft to be launched early in year 2000. The principal science objective of RPI is to characterize the plasma in the Earth's magnetosphere by radio frequency imaging. A key product of RPI is the *Plasmagram*, a map of radio signal strength vs. echo delay-time vs. frequency, on which magnetospheric structures appear as curves of varying intensity. Noise and other emissions will also appear on RPI Plasmagrams and when strong enough will obscure the radar echoes. To aid in the analysis of RPI plasmagrams, a computer program is being implemented to automatically detect and enhance the radar echoes. The techniques used are derived within a Bayesian framework and include Maximum Likelihood and Maximum Posterior analyses. A heuristic stochastic global optimization method blending elements of simulated annealing and genetic algorithms is used to determine what model echoes are supported by plasmagram evidence. The application of this work to RPI data will be discussed.

# Outline of this Presentation

This presentation is organized into six vertical display tracks:

1. Introductory track (this track),
2. Background on IMAGE/RPI,
3. High signal-to-noise detection,
4. Low signal-to-noise detection (derivation),
5. Low signal-to-noise detection (results), and
6. Conclusion.

# Overview of the method

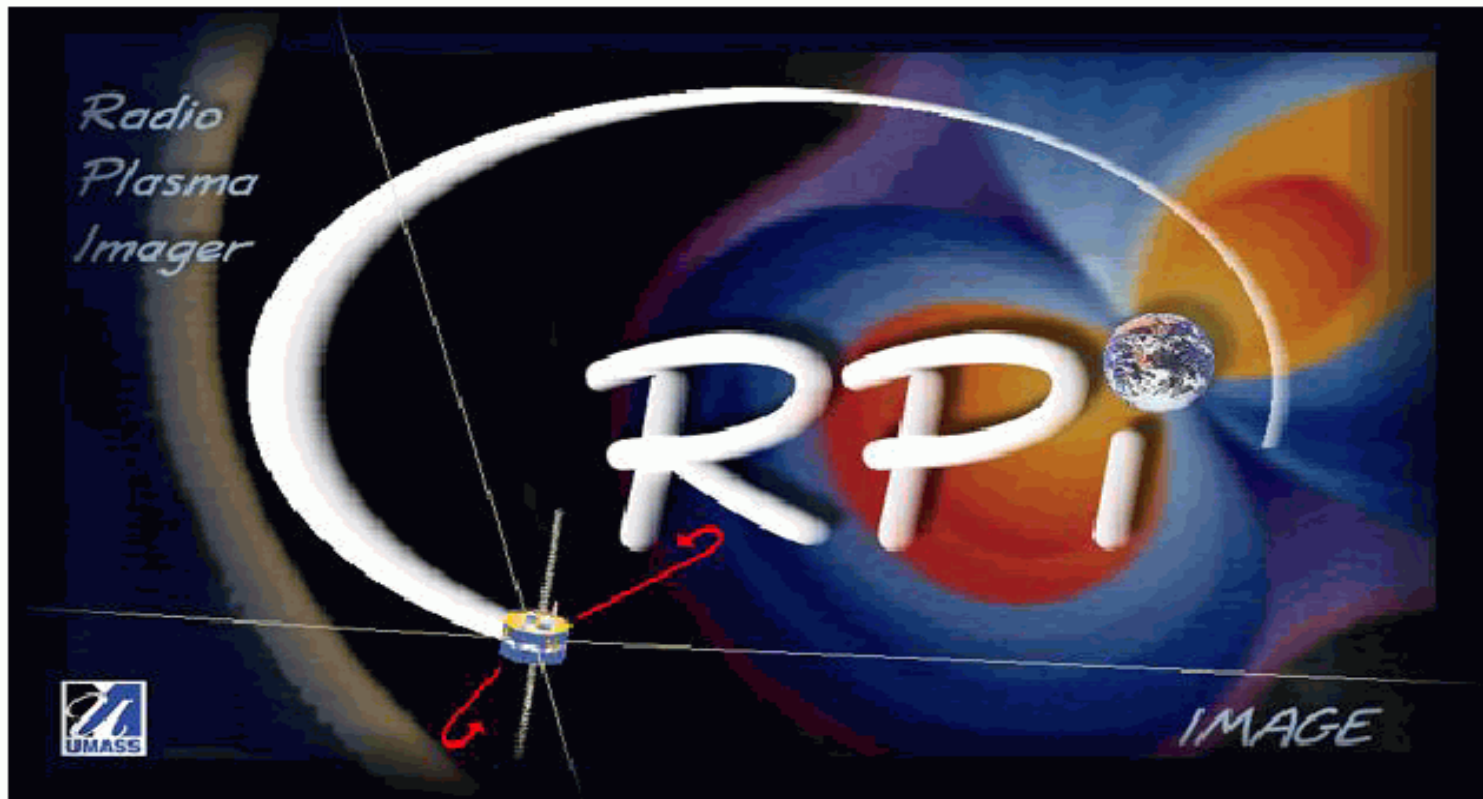
---

Our approach to RPI signal detection and enhancement is to:

1. Use simple thresholding to acquire high signal-to-noise (SNR) echoes;
2. Use high SNR data as the starting point for low SNR modelling;
3. Use noise and signal models built within a Bayesian framework;
4. Identify regions of model parameter space that suggest detections.

The Bayesian framework provides a disciplined way to reason about constraints, models, and data. Questions such as “What are the most likely or probable model-based situations given observations?” can be posed as optimization problems within the framework.

# Radio Plasma Imager



JLG 1999.1201 CUA Presentation.

# IMAGE/Radio Plasma Imager (RPI)

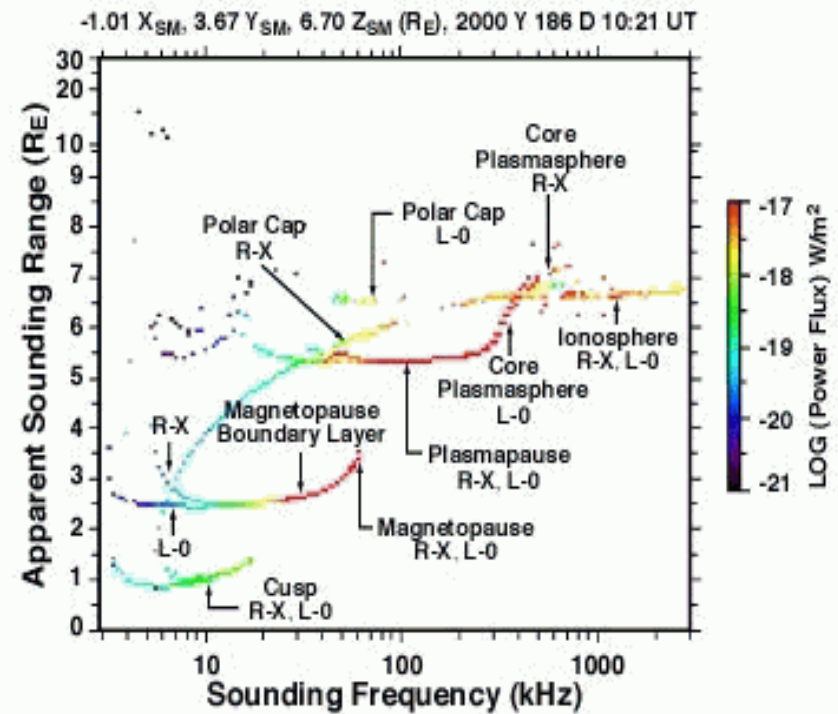
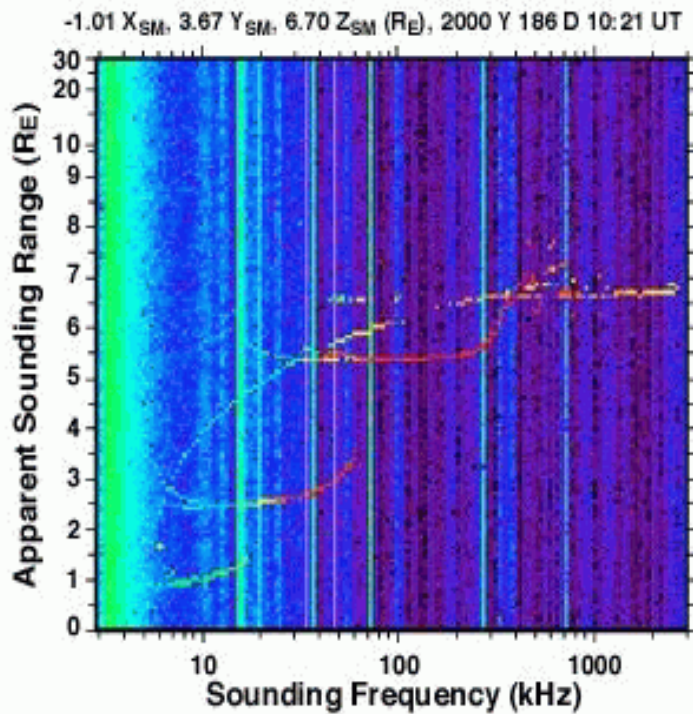
---

IMAGE/RPI is designed to obtain a global profile of the magnetospheric density from the magnetopause to the ionosphere.

- Frequencies: 3 kHz to 3 MHz; Plasma Densities: 10 to  $10^5$ /cc
- Two crossed 500m tip-to-tip thin wire dipoles (spin plane)
- One 20m tip-to-tip thin wire dipole (spin axis)
- 10 W peak power, 3 kV max. antenna voltage
- On board processing for enhancement of signal-to-noise
- A relative of *Digisonde Portable Sounder* developed at University of Massachusetts, Lowell (Reinisch et al. 1992).

# Radio Sounding

## Modeled IMAGE RPI Plasmagram



JLG 1999.1201 CUA Presentation.

# Sources of interference

---

- **Solar-Heliospheric**

e.g. Type III bursts and storms

- **Terrestrial Magnetospheric**

Auroral kilometric radiation

Trapped and escaping continuum radiation

- **Astronomical**

- **Instrumental & In situ plasma emission**

Note: The current work is by no means the first line of mitigation for these noise sources. Principal techniques built into RPI's data control system include signal processing, frequency control, and polarization and spatial discrimination (cf. Green et al. 1997).



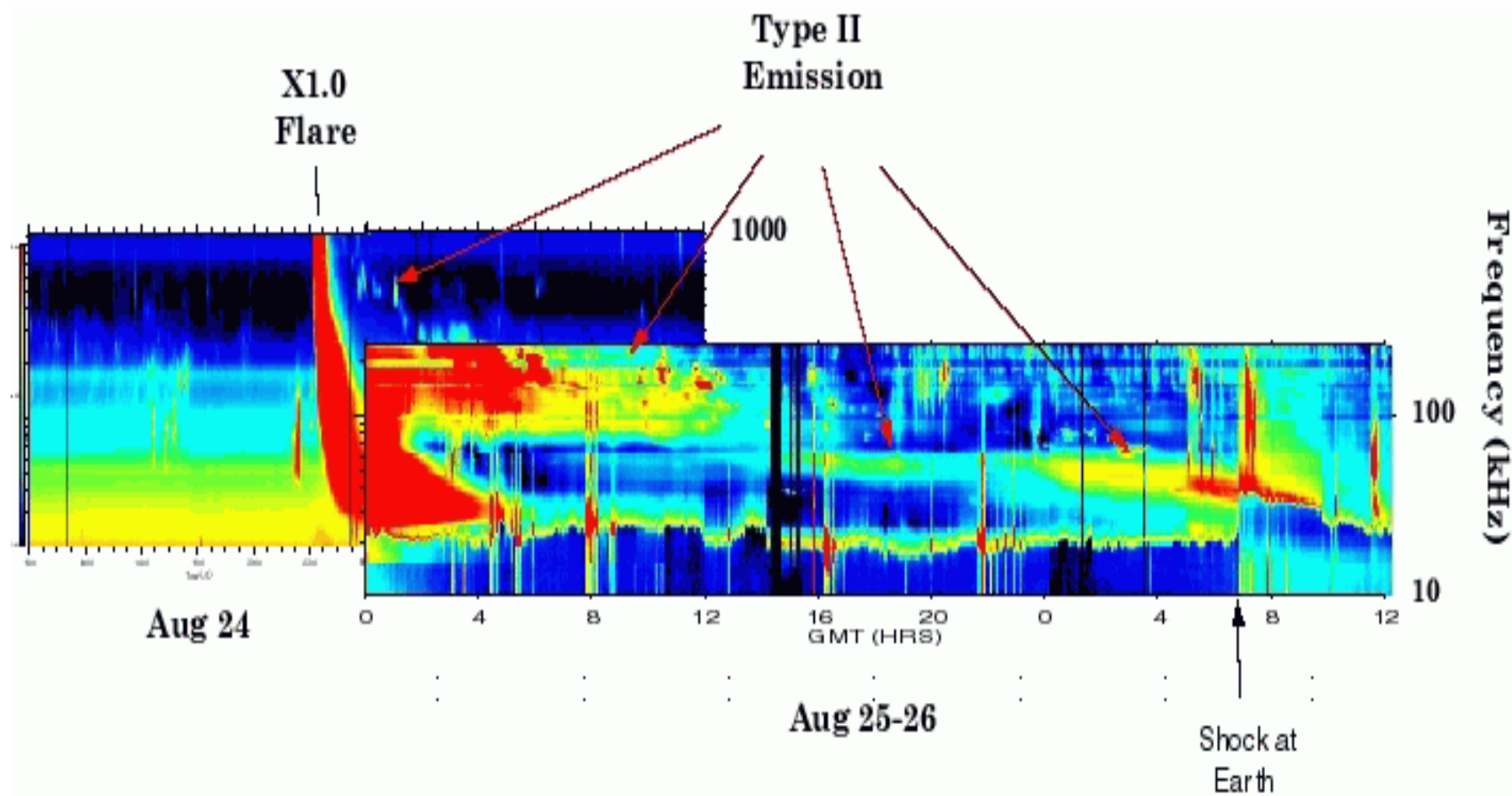


Figure 1: Radio observations by *WAVES* on the *Wind* spacecraft. Emissions such as these will be observed by RPI.

<http://lep694.gsfc.nasa.gov/waves/waves.html>, M. L. Kaiser, author.

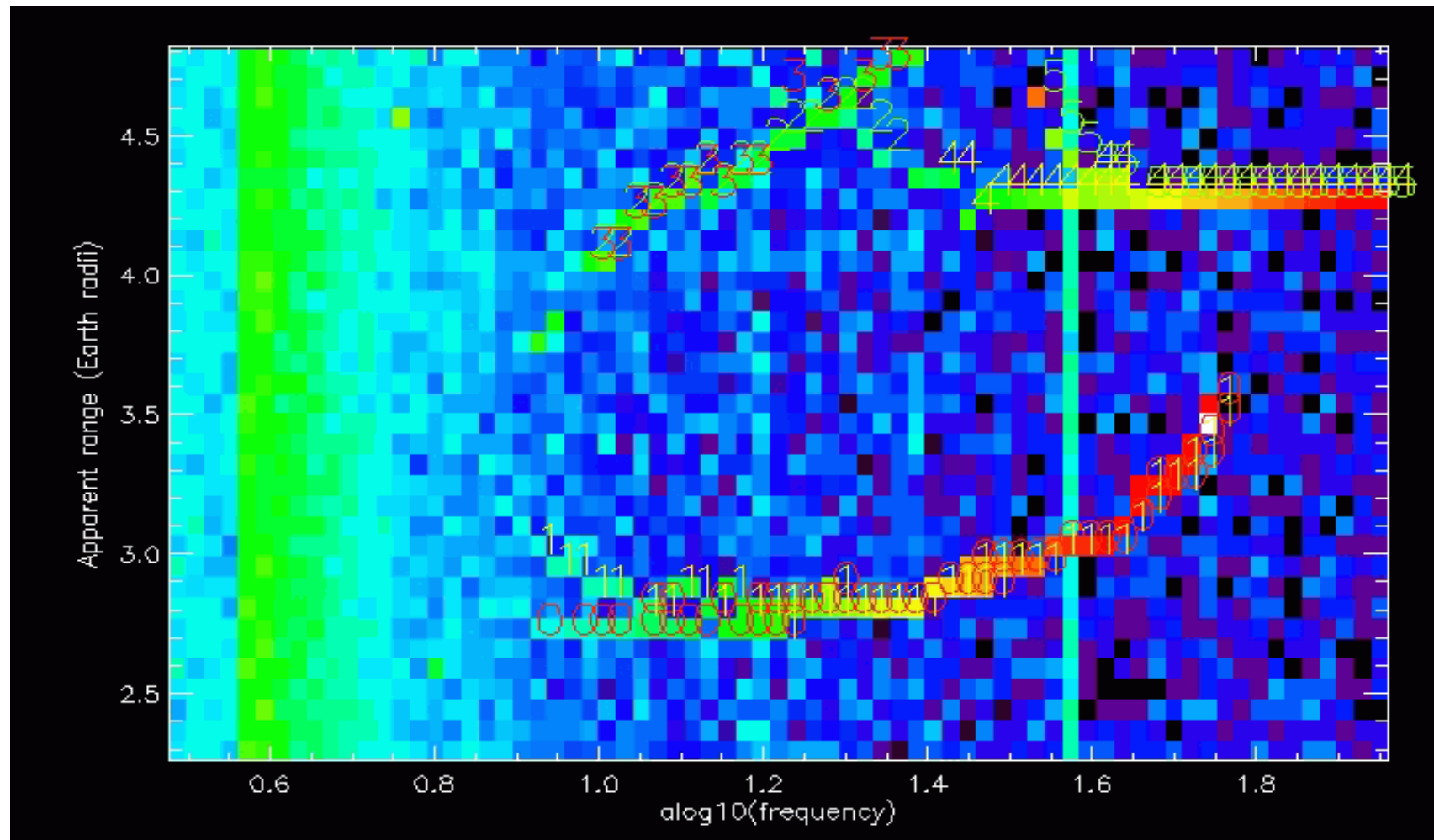
# Analysis at High Signal-to-Noise Ratios

At high signal to noise ratios, we can be confident that a detection has occurred. We have implemented the following simple scheme to automatically trace linear features that occur in RPI plasmagram data.

1. Estimate noise floor from histogram.
2. List datapoints that exceed a desired, fixed SNR.
3. Construct lists of data points that form contiguous traces.
4. Search for multiple peaks in time as a sign of branching.
5. Construct a set of single valued echo traces for further analysis.

At this point individual traces can be treated as individual datapoints.

# High signal-to-noise ratio trace extraction



## Example usage: Automated model fitting

---

Once individual traces have been extracted, we can fit models to them. A two step process has proven efficient and reliable:

1. Fit the strength of the trace as a function of frequency;
2. Then fit the “delay time” behavior of the trace.

We factor the echo trace into a portions responsible for the echo strength  $W_{echo}$  and apparent range  $\mathcal{C}_{echo}$  as functions of frequency  $f$ .

$$\tilde{S}_{echo} = W_{echo} \mathcal{C}_{echo}$$

After examining several different forms, the form of echo strength used here is a log-log-polynomial:

$$\ln W_{echo} \equiv \ln W_{echo}(f) = \sum_k \eta_k (\ln f)^k .$$

The form of the apparent range is

$$\ln \mathcal{C}_{echo} \equiv \ln \mathcal{C}_{echo}(t, f) = -a \sum_j (t - \tilde{t}_j)^2.$$

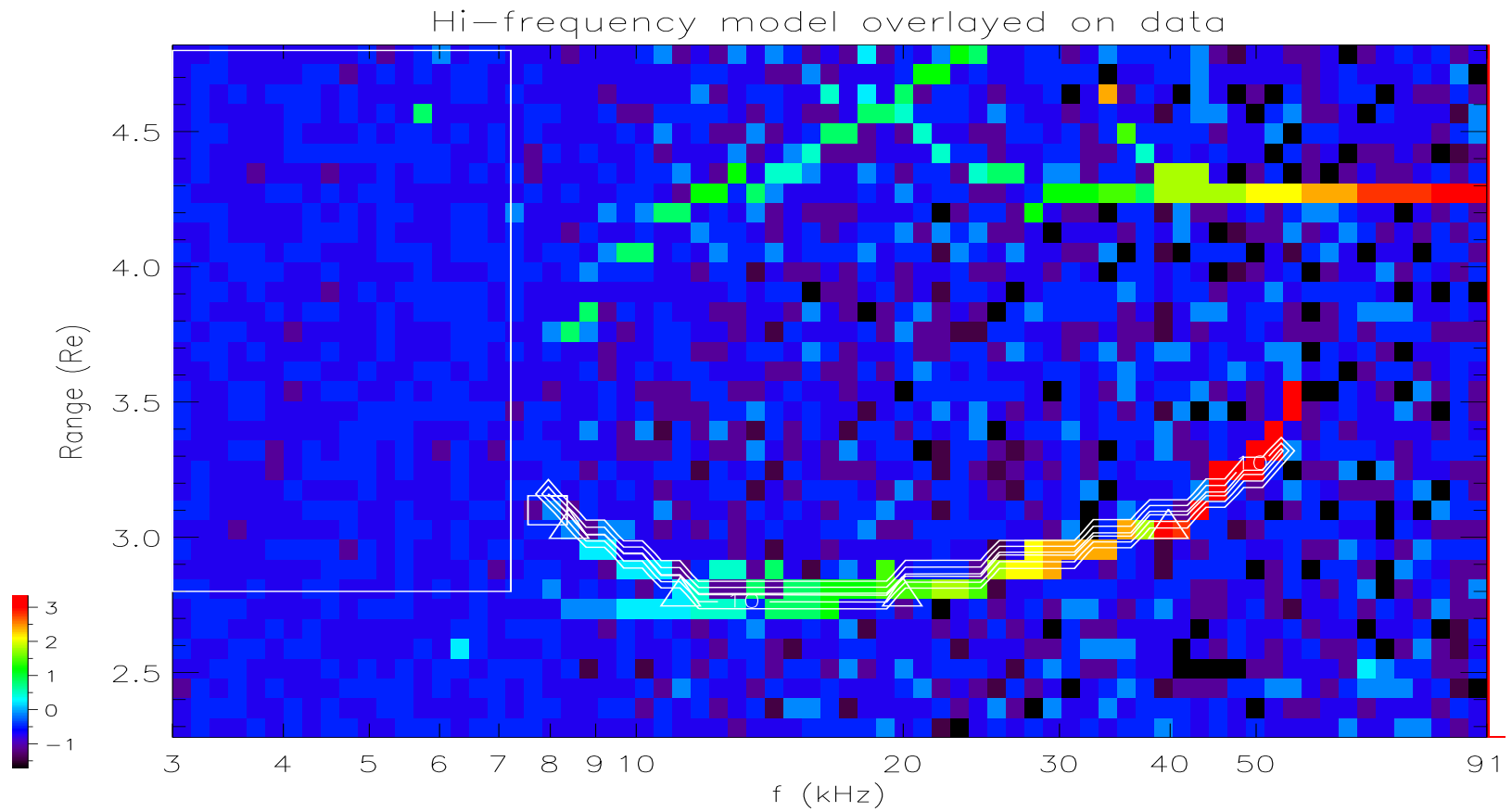
The parameter  $a$  is the reciprocal width of the modeled echo trace, and for the work here  $a$  is set so that the model traces are not resolved in time. Times are presented as apparent ranges measured in Earth radii. The maxima of  $\mathcal{C}_{echo}$  occur when the left hand side of the previous equation is zero. The locations of these zeros are governed by

$$\tilde{t}_j \equiv \tilde{t}_j(f) = \text{Spline}_k(\eta_{jk}, \ln f).$$

We typically used cubic splines to control the model's echo return time. For the implementation used the model parameters  $\eta_{jk}$  are the points through which the spline passes.

# A fit to high SNR data

---



The signal strength depicted above is in counts that only roughly reflect actual RPI values.

# Dealing with interference

---

Our approach is based on Bayes' rule (e.g. Ó Ruanaidh and Fitzgerald 1996):

$$p(\tilde{S}|SI) = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}} = \frac{p(S|\tilde{S}I)p(\tilde{S}|I)}{p(S|I)}$$

which prescribes a way to combine models, data, and one's *a priori* expectations.

The model connecting observed data with physical parameters is the likelihood:

$$l \equiv p(S|\tilde{S}I) = \frac{(\tilde{S}\delta t)^{S\delta t} e^{(-\tilde{S}\delta t)}}{(S\delta t)!} = \frac{(\sum_x S_x \tilde{S}_x \delta t)^{S\delta t} e^{-\sum_x \tilde{S}_x \delta t}}{(S\delta t)!}$$

These forms arise when the signal statistics are adequately modeled by a Poisson distribution, which is the observed case for RPI data.

# Signal+Noise<sup>a</sup>

---

We have assumed the signal and noise to be additive

$$\tilde{S} = \sum_x \tilde{S}_x = \tilde{S}_{echo} + \tilde{S}_{noise}.$$

The noise model is written

$$\tilde{S}_{noise} = W_{noise} \equiv W_{noise}(f),$$

where  $W_{noise}$  is only a function of frequency for this work, but could be a more general function. For most purposes,  $W_{noise}$  can be determined during passive RPI observations. For example, we have set  $W_{noise}$  by time-averaging data that does not contain echoes.

---

<sup>a</sup>The duration of a given observation  $\delta t$  is a function of time and frequency and is absorbed into the signal and noise powers.



# Log-likelihood

---

It is more convenient to work with the log-likelihood:

$$\ln l = \ln p(S|\tilde{S}I) = S \ln \tilde{S} - \tilde{S} - \ln(S)! \approx -S \ln \frac{S}{\tilde{S}} - \tilde{S}.$$

Signals are measured and modeled on a set of discrete times and frequencies, here denoted as a set of elements  $tf$ . If we take the statistics of measurements obtained at different  $tf$  to be independent, we can write the log-likelihood as

$$\ln l = \ln p(S|\tilde{S}I) = \sum_{tf} \ln p(S_{tf}|\tilde{S}_{tf}I).$$

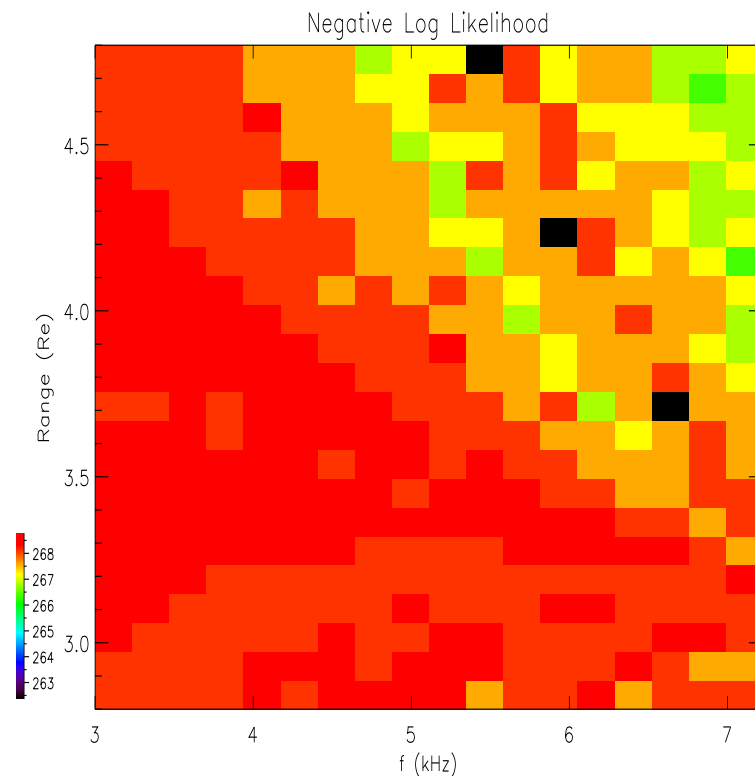
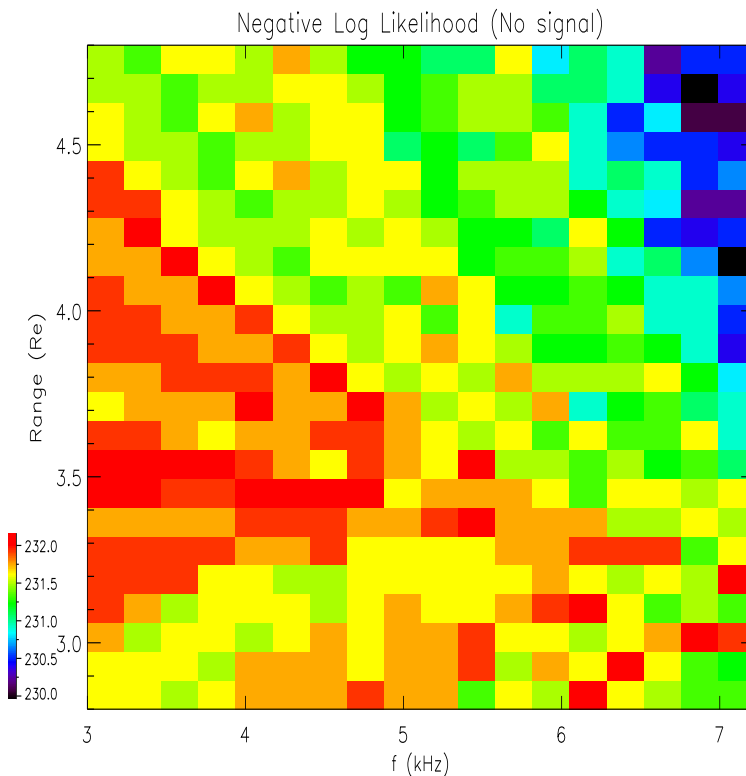
For this work we neglect priors, normalization and deal only with the log-likelihood,  $\ln p(S|\tilde{S}I)$ :

$$\begin{aligned}
\ln l &= \ln p(S|\tilde{S}I) \\
&= \ln p(S|\eta_{echo}, W_{echo}, W_{noise}, I) \\
&\approx - \sum_{tf} \left( S_{tf} \ln \frac{S_{tf}}{\tilde{S}_{tf}} + \tilde{S}_{tf} \right) \quad (1)
\end{aligned}$$

where  $\tilde{S}_{tf} \equiv \tilde{S}_{tf}(\eta_{echo}, W_{echo}, W_{noise})$ . Equation 1 is the foundation of our method. A Maximum Likelihood (ML) approach would entail a search for the maximum of equation 1. A more complete picture of the information contained in the data can be obtained by examining the log-likelihood itself. We have done this by holding those parameters determined during the high SNR analysis fixed while allowing only the delay-frequency position of the low frequency endpoint of the trace to vary.

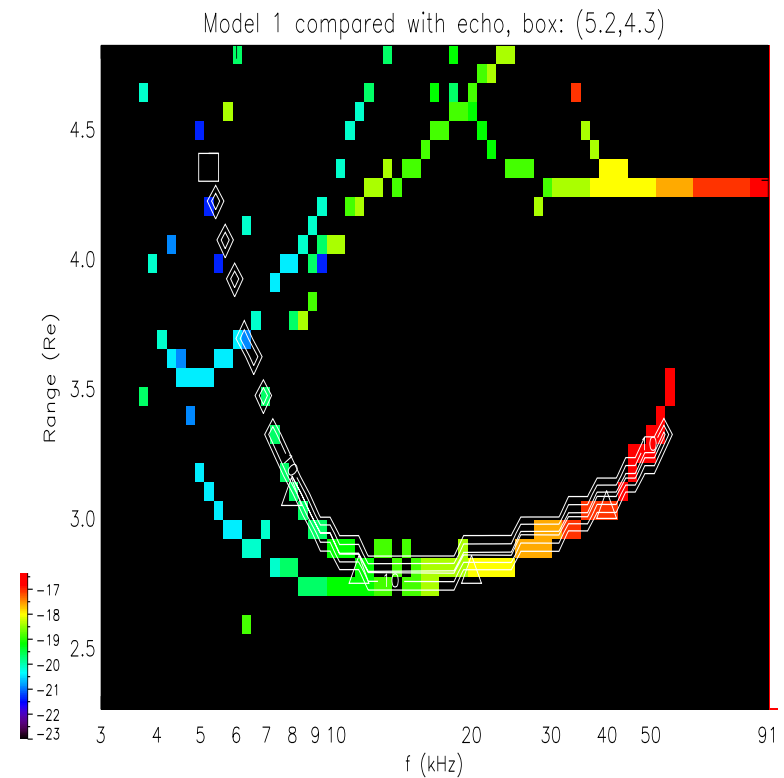
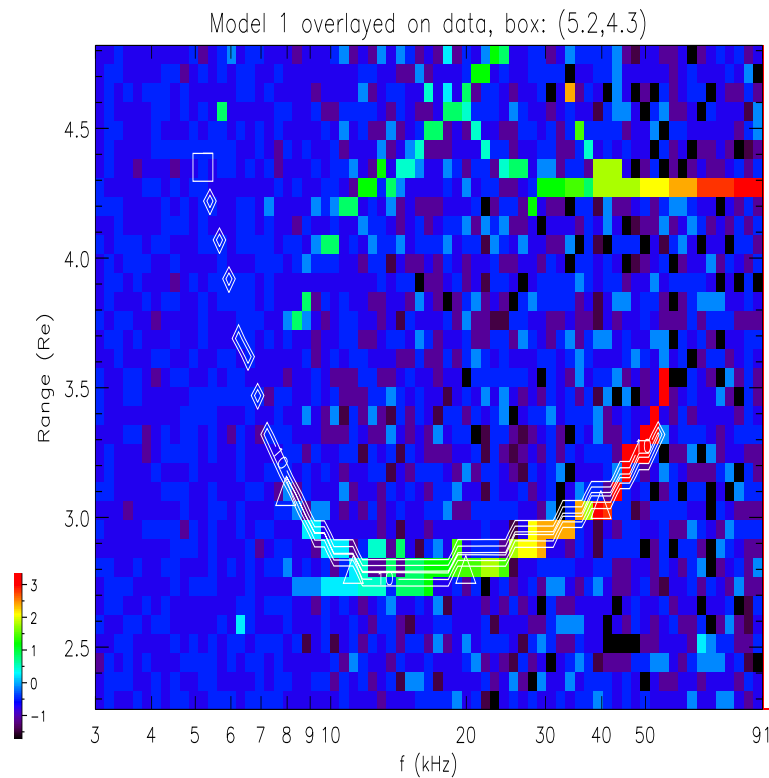
# Likelihood of detection

---



On the left, the signal has been artificially removed. Note how the log-likelihood can change in the presence of a signal.

# Extension of the Magnetopause trace



The box at (5.2,4.3) above shows the endpoint which is also the two-dimensional parameter varied to construct the maps of the log-likelihood.

## Interpretation of the reconstruction

---

This model reconstruction is based on parameters from the light green plain to the immediate left of the black pits visible in the log-likelihood. The black pits correspond to the stronger echoes immediately to the right of the displayed reconstruction.

Within the context of our simple signal+noise model, the data support multiple reconstructions as seen by the multiple local minima of the likelihood.

For operational use, more complicated and perhaps physics-based models that support multiple traces would use more completely the information available in the data.

# Genetic Algorithmic approach to optimization

---

As stated in the abstract, we implemented a program that attempts to minimize the log-likelihood as a function the entire parameter space,  $\eta$ ,  $W_{\text{echo}}$ , and  $W_{\text{noise}}$ .

We found this algorithm excessively good at finding minima of the log-likelihood that had no physical significance. From this we realized that local minima and features appearing in log-likelihood space often have as much or more relevance than global minima. Better behavior will likely be brought out by better models and physically based constraints.

Before more fully exploring these avenues, we have focussed our efforts developing the methods presented here.

# Conclusion and Plans

---

We have developed programs that automatically:

1. Trace, segment, and label high quality RPI echoes;
2. Assess the likelihood for RPI signal detection in extreme noise;
3. Perform 1 and 2 to provide a list of likely RPI echo detections.

**Plans:** The models used in this work are simple and lack features required for operational use. Currently echo reconstructions are based on simple notions of continuity in time and frequency and spline approximation. Feature detection in likelihood space, priors and heuristics based on the magnetospheric phenomenology, and an empirical understanding of the performance of RPI on orbit are elements of a more complete solution.

# Acknowledgements

---

This work has been supported by NASA Contract NASW-97002.

# References

---

- IMAGE Website: <http://image.gsfc.nasa.gov>, E. Bell, NASA Goddard Space Flight Center
- WIND/WAVES Website: <http://lep694.gsfc.nasa.gov/waves/waves.html>, M. Kaiser, NASA Goddard Space Flight Center
- Green, J. L. et al., 1997, in Proceedings of the Chapman Conference on Space Plasma Measurement Techniques, Santa Fe, NM, April 3-7, 1995, esp. the references to Renisch, B. W. et al. (1992, 1995) therein.
- Ó Ruanaidh, J. K. and Fitzgerald, W. J. , 1996, *Numerical Bayesian Methods Applied to Signal Processing*, Springer Verlag.