Automatic Shape Recognition of Type III Radio Bursts in Solar Wind Dynamical Radio Spectra

S. Vidojević, M. Dražić and M. Maksimovic

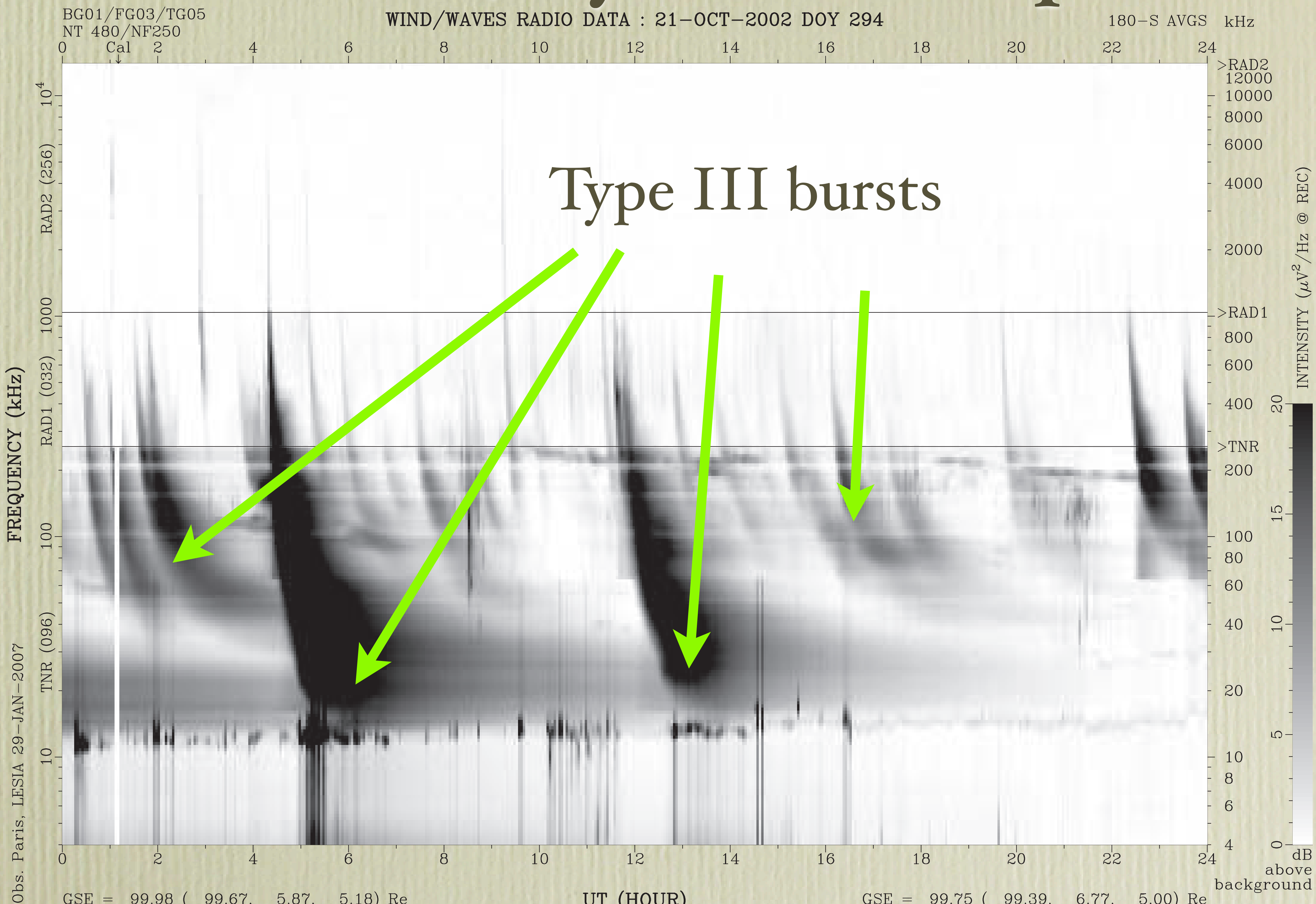
State University of Novi Pazar, V. Karadzica bb, 36300 Novi Pazar , Serbia
Faculty of Mathematics, Studentski trg 16, 11000 Belgrade, Serbia
CNRS and LESIA Observatoire de Paris, 92195 Meudon, France
E-mail: sonja@matf.bg.ac.rs, mdrazic@sezampro.rs, milan.maksimovic@obspm.fr

XI SCSLSA, August 21-25 2017, Šabac, Serbia

Motivation

- Large amounts of data are already recorded and stored - they continue to grow every day.
- People have no time to analyze this data - human attention has become the precious resource.
- So, we must find ways to automatically analyze the data, to automatically classify it, summarize it, to discover and characterize trends in it, to automatically flag anomalies etc.

Observations, dynamical spectrum

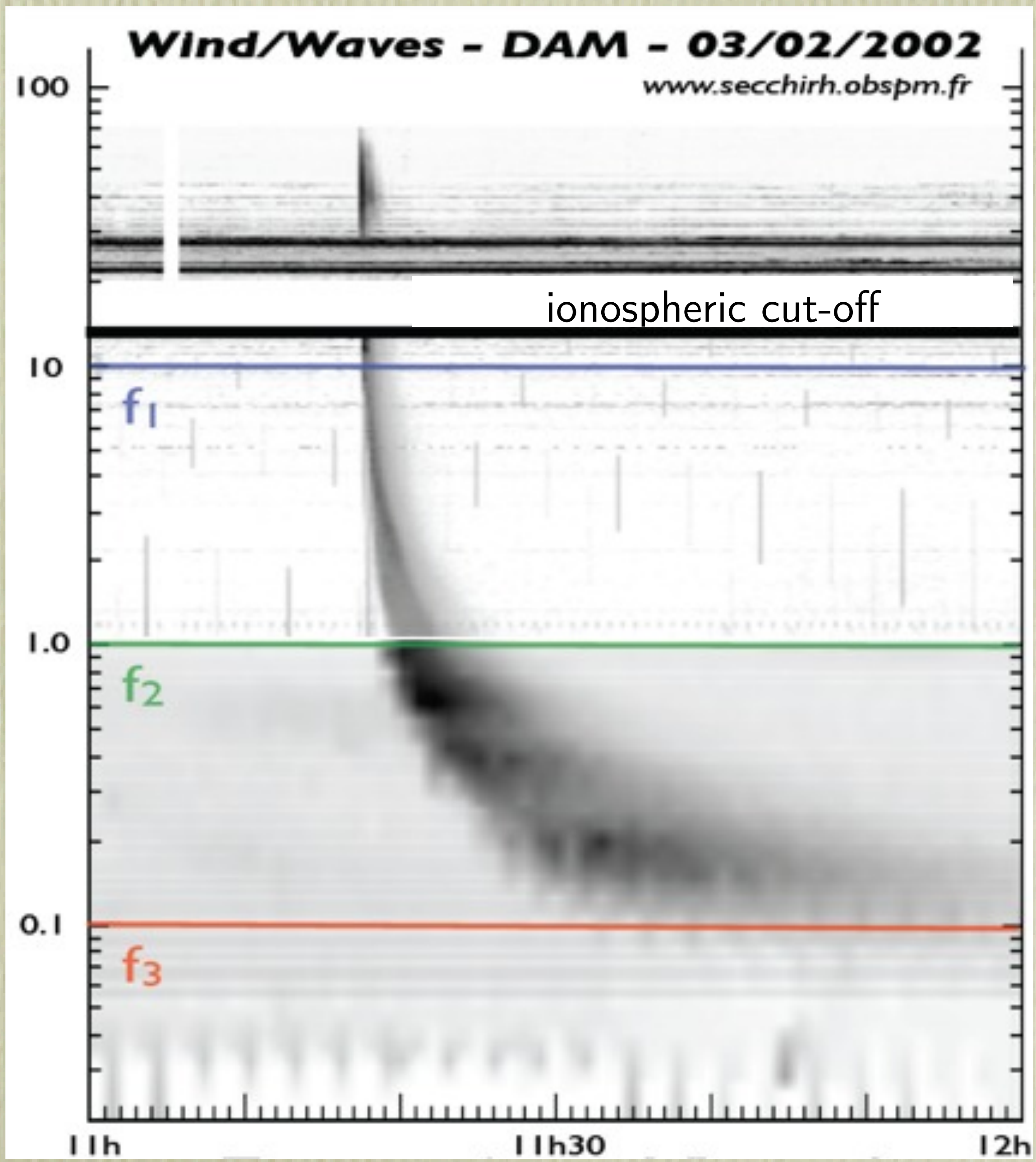


Frequency [4 kHz- 14 MHz]

Time [24 h]

XI SCSLSA, August 21-25 2017, Šabac, Serbia

Type III Bursts



- Short (sec → hrs) & very intense (→10-14 Wm-2Hz-1) radio emissions;
- Emission frequencies decrease rapidly (GHz → kHz);
- Emission at fundamental fp or at harmonic of fp;
- Often associated with solar flares;
- Associated with the propagation of electrons supra-thermal (c/10 → c/3);

TIII's Frequency Drift*) (1/2)

- The frequency, related to the local plasma frequency ($f_{pl} \propto \sqrt{n_e}$, $n_e \propto 1/R^2$, $f_{pl} \propto 1/R$), drifts downward as the emission region rapidly propagates outward.

- Since the radio burst is generated by local plasma emission processes, radio emissions at high frequencies (high plasma densities) occur very near the Sun ~ $2R_\odot$ for 16 MHz, while those at low frequencies (low plasma densities) occur far from the Sun (~ 1 AU) for 20 kHz.

- Type III radio bursts are therefore characterized by a rapid drift to lower frequencies due to the near-relativistic speeds of the burst electrons.

*) Vidojevic S., Maksimovic M.: *Preliminary Analysis of Type III Radio Bursts from STEREO/SWAVES Data*, XV National Conference of Astronomers of Serbia, 2–5 October 2008, Belgrade, Serbia, Publ. Astron. Obs. Belgrade No. **86** (2009), 287 - 291. http://publications.aob.rs/86/pdf/287-291.pdf

TIII's Frequency Drift*) (2/2)

- For about 100 bursts we have computed the frequency drift rates obtained from all the maxima of the power spectral density profiles at each of the covered frequencies. The profiles are fitted by Gram-Charlier type A function.

- Obtained maxima are further aproximated by linear function in log-log scale.

- **$df/dt = -10^a f^\alpha$**. The negative sign denotes that the starting frequency is observed to drift from high to low values. The least square fit of a straight line through all of observed maxima gives:

- **$\alpha = 1.80 \pm 0.05$ and $a = -1.70 \pm 0.03$.**

Spectra Shape Modelling*)

- The choice of the best-suited statistical distribution for data modelling is not a trivial issue;
- Unless a sound theoretical background exists for selecting a particular distribution, one will usually try to test various candidates and select a distribution based on its fit to the observed data;
- It is more efficient to define a sufficiently general family that can be used for this purpose.

*) S. Vidojevic Shape Modelling with Family of Pearson Distributions, 9th SerbianConference on Spectral Line Shapes in Astrophysics, Banja Koviljaca, Serbia, May 13-17, 2013, Book of abstracts, p. 52, http://www.scslsa.matf.bg.ac.rs/Book_of_abstracts_9thSCSLSA.pdf

Pearson system - great diversity of shapes:

- unimodal, bimodal, U-shaped, J-shaped and monotone probability distribution functions,
- ...which may be symmetric and asymmetric, concave and convex,
- ...with smooth, abrupt, truncated, long, medium or short tails.

Pearson system*)

- First derivative of probability density function:

$$\frac{1}{f(x)}\frac{\mathrm{d}f(x)}{\mathrm{d}x} = -\frac{a+x}{c_0+c_1x+c_2x^2}$$

- Asymmetry ($As^2 = \beta_1$)

$$\beta_1 = \frac{\mu_3^2}{\mu_2^2}$$

- Excess (β_2)

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

Using only 2 parameters: Squared Asymmetry (β_1) and Excess (β_2), calculated from observations, Type of Pearson distribution can be retrieved.

*) Pearson, K.: 1895, *Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material.* Philosophical Transactions of the Royal Society of London, **186**, 343 – 414

Method of moments

$$c_0 = (4\beta_2 - 3\beta_1)(10\beta_2 - 12\beta_1 - 18)^{-1}\mu_2$$

$$a = c_1 = \sqrt{\beta_1}(\beta_2 + 3)(10\beta_2 - 12\beta_1 - 18)^{-1}\sqrt{\mu_2}$$

$$c_2 = (2\beta_2 - 3\beta_1 - 6)(10\beta_2 - 12\beta_1 - 18)^{-1}$$

$$\kappa = \frac{1}{4}c_1^2(c_0c_2)^{-1} = \frac{1}{4}\beta_1(\beta_2 + 3)^2(4\beta_2 - 3\beta_1)^{-1}(2\beta_1 - 6)^{-1}$$

Classification

I: $\kappa < 0$

II: $\beta_1 = 0, \ \beta_2 < 3$

III: $2\beta_2 - 3\beta_1 - 6 = 0$

IV: $0 < \kappa < 1$

V: $\kappa = 1$

VI: $\kappa > 1$

VII: $\beta_1 = 0, \ \beta_2 > 3$

Method of Maximum Likelihood

Likelihood function

$$L(\boldsymbol{\theta}|\mathbf{x}) \equiv f(\mathbf{x}|\boldsymbol{\theta}) = \prod_{i=1}^{n} f_i(x_i|\boldsymbol{\theta})$$

applying logarithm, one obtain:

$$\mathcal{L}(\boldsymbol{\theta}|\mathbf{x}) = \ln L(\boldsymbol{\theta}|\mathbf{x}) = \sum_{i=1}^{n} \ln f_i(x_i|\boldsymbol{\theta})$$

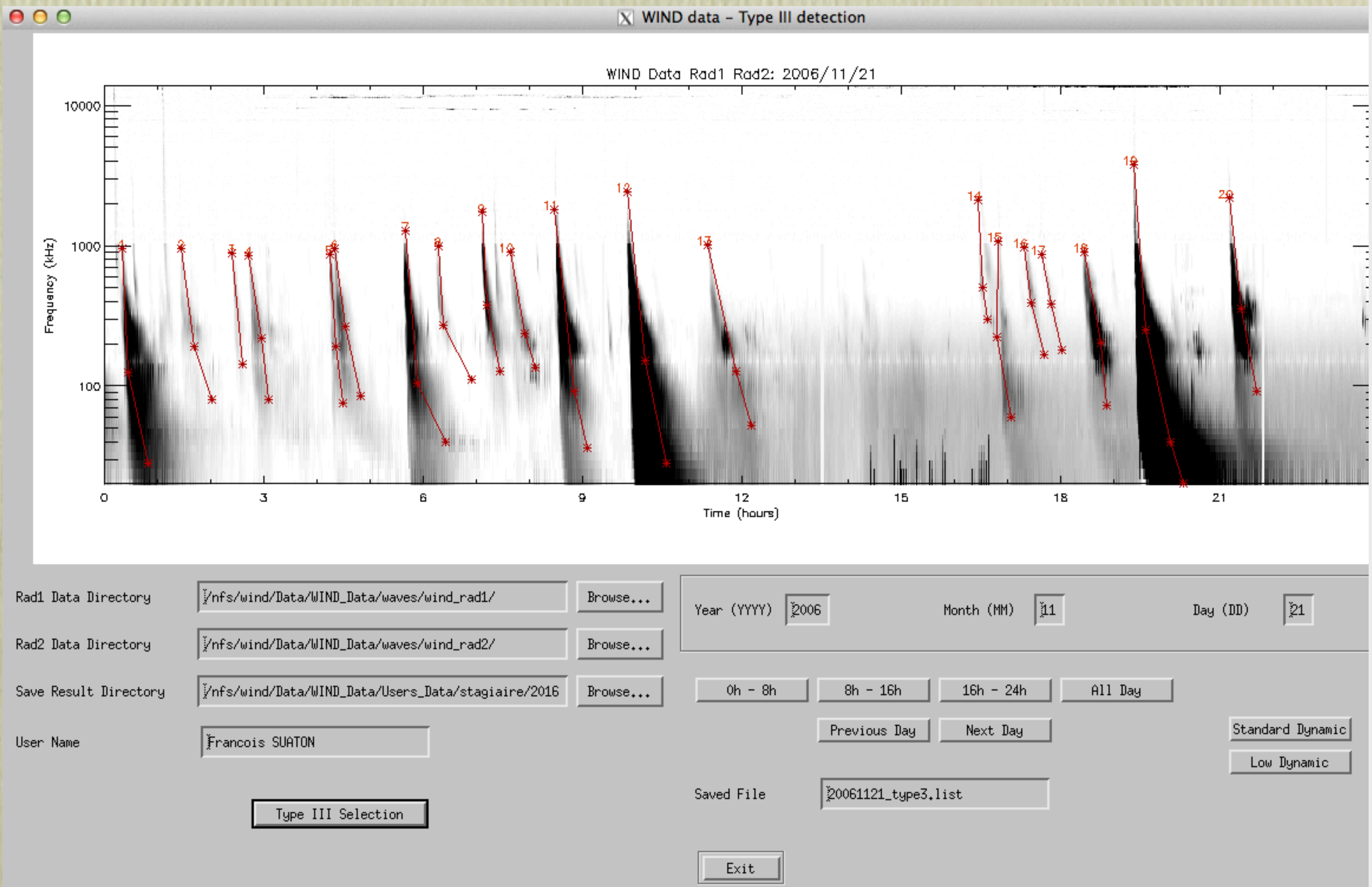
Looking for $\boldsymbol{\theta}^*$

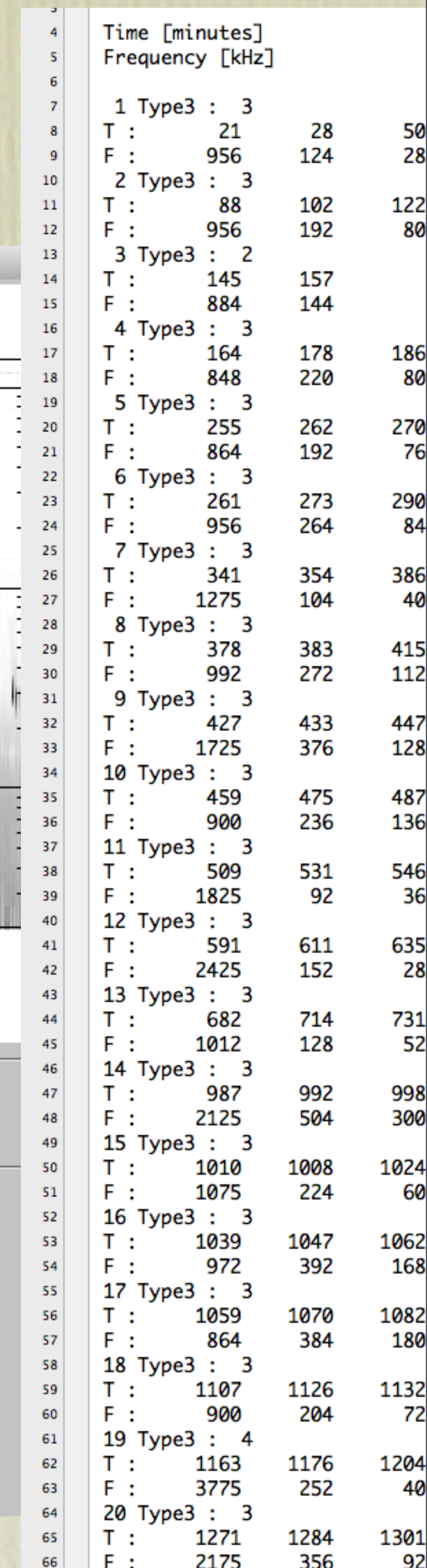
- Looking for $\boldsymbol{\theta}^*$ which maximizes likelihood

$$\mathcal{L}(\boldsymbol{\theta}^*|\mathbf{x}) = \max_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}|\mathbf{x}) = \max_{\boldsymbol{\theta}} \sum_{i=1}^{n} \ln f_i(x_i|\boldsymbol{\theta})$$

- It is not possible to solve this task analytically, thus, we apply numerical methods of optimization.

Manual Detection

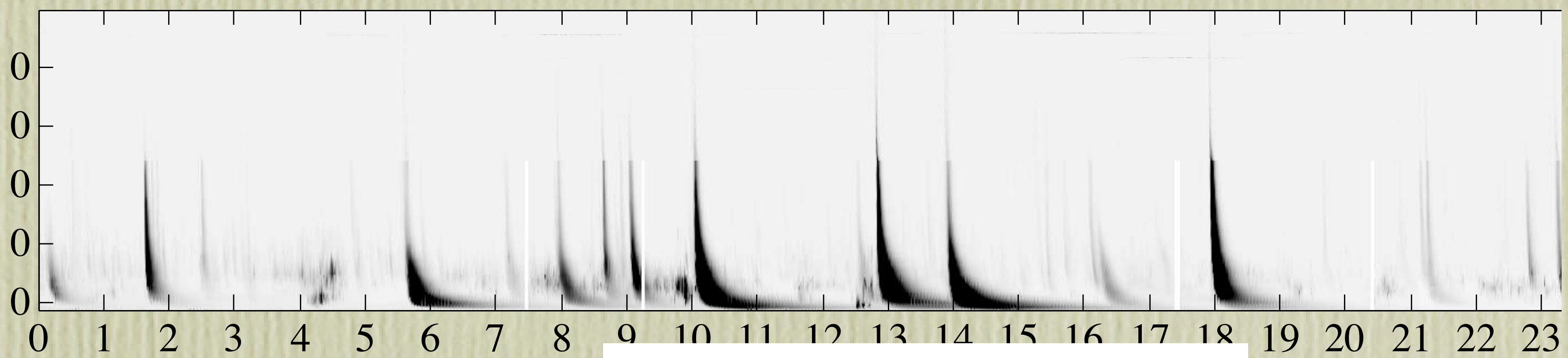




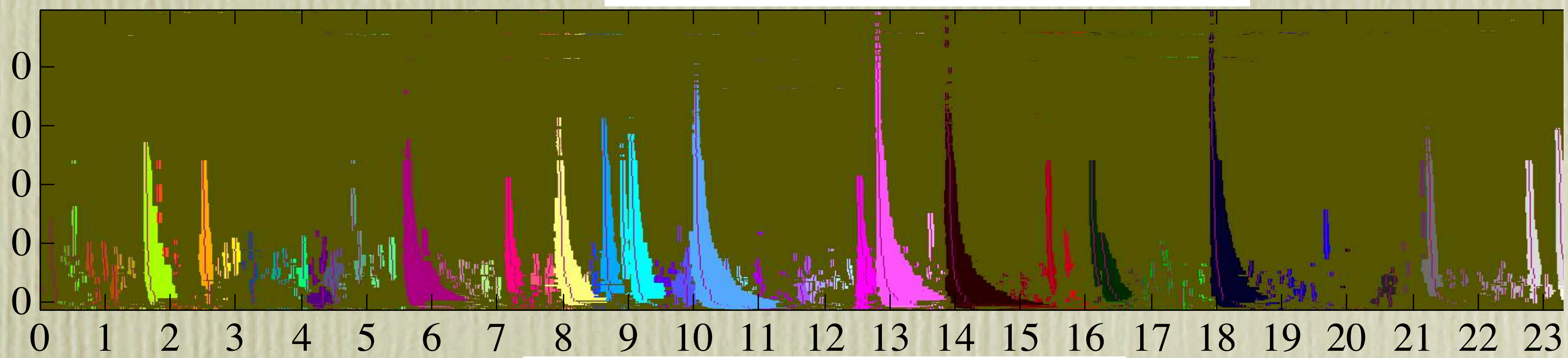
```
Time [minutes]
Frequency [kHz]

 1 Type3 :  3
T :      21     28     50
F :     956    124     28
 2 Type3 :  3
T :      88    102    122
F :     956    192     80
 3 Type3 :  2
T :     145    157
F :     884    144
 4 Type3 :  3
T :     164    178    186
F :     848    220     80
 5 Type3 :  3
T :     255    262    270
F :     864    192     76
 6 Type3 :  3
T :     261    273    290
F :     956    264     84
 7 Type3 :  3
T :     341    354    386
F :    1275    104     40
 8 Type3 :  3
T :     378    383    415
F :     992    272    112
 9 Type3 :  3
T :     427    433    447
F :    1725    376    128
10 Type3 :  3
T :     459    475    487
F :     900    236    136
11 Type3 :  3
T :     509    531    546
F :    1825     92     36
12 Type3 :  3
T :     591    611    635
F :    2425    152     28
13 Type3 :  3
T :     682    714    731
F :    1012    128     52
14 Type3 :  3
T :     987    992    998
F :    2125    504    300
15 Type3 :  3
T :    1010   1008   1024
F :    1075    224     60
16 Type3 :  3
T :    1039   1047   1062
F :     972    392    168
17 Type3 :  3
T :    1059   1070   1082
F :     864    384    180
18 Type3 :  3
T :    1107   1126   1132
F :     900    204     72
19 Type3 :  4
T :    1163   1176   1204
F :    3775    252     40
20 Type3 :  3
T :    1271   1284   1301
F :    2175    356     92
```

Example 1

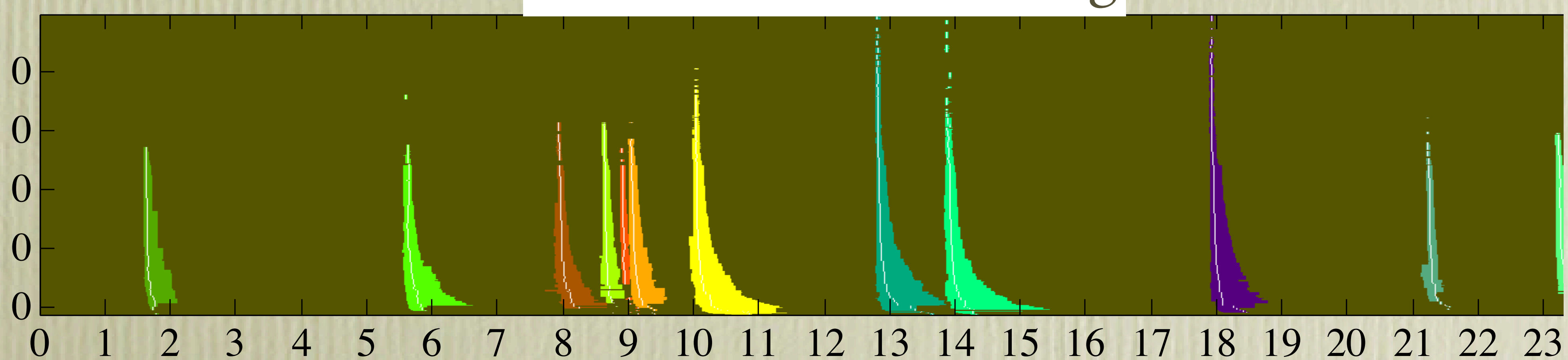
Original data 19971123



Candidates before filtering



Candidates after filtering

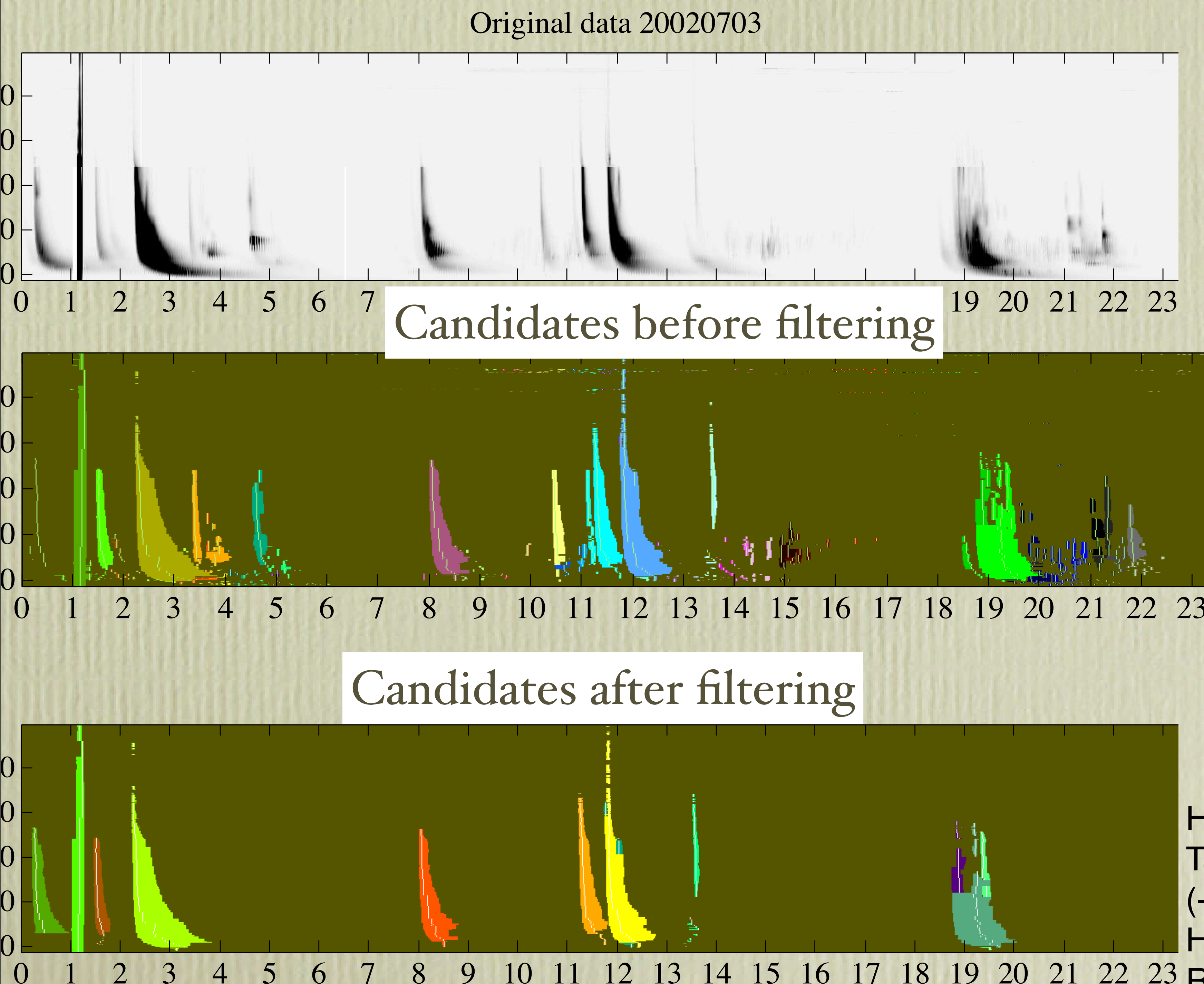


Processing date: 19971123 with parameters:

find_peaks_cutoff: 1.8
find_peaks_slope: 0.2
find_peaks_peakchkdist: 5
find_hills_maxdistforcontnext: 3
find_hills_overlaptol: -4
find_hills_maxdistforcontall: 15
find_hills_peakvalchangetol: [0.5 2]
filter_hills_noisedetect: [4 100 10]
filter_hills_maxdistforcont: 20
filter_hills_minhilllen: 50
filter_hills_notbelowfreq: 1.1
filter_hills_shapecheck: [1 1]
filter_hills_delreport: [0 0 0 0 0 1 1]
image_hills_peakvalue: 50

Hills before filtering: 194
Hills after filtering: 12
Results saved to data/19971123_res.mat

Example 2



- Processing date: 20020703 with parameters:

find_peaks_cutoff: 1.8
find_peaks_slope: 0.2
find_peaks_peakchkdist: 5
find_hills_maxdistforcontnext: 3
find_hills_overlaptol: -4
find_hills_maxdistforcontall: 15
find_hills_peakvalchangetol: [0.5 2]
filter_hills_noisedetect: [4 100 10]
filter_hills_maxdistforcont: 20
filter_hills_minhilllen: 50
filter_hills_notbelowfreq: 1.1
filter_hills_shapecheck: [1 1]
filter_hills_delreport: [0 0 0 0 0 1 1]
image_hills_peakvalue: 50

Hills before filtering: 164
Take out hill 126 at 19.08 h : not convex (-3.719713)
Hills after filtering: 12
Results saved to data/20020703_res.mat

More data - satellites:

- WIND spacecraft, launched 1994, still operating.
- STEREO A and B, launched 2006, still operating.
- Solar Probe Plus, to be launched in 2018.
- Solar Orbiter, to be launched in 2019.

Literature

- Vidojevic S., Maksimovic M.: P*reliminary Analysis of Type III Radio Bursts from STEREO/SWAVES Data*, XV National Conference of Astronomers of Serbia, 2–5 October 2008, Belgrade, Serbia, Publ. Astron. Obs. Belgrade No. **86** (2009), 287 - 291. http://publications.aob.rs/86/pdf/287-291.pdf

- S. Vidojevic Shape Modelling with Family of Pearson Distributions, 9th SerbianConference on Spectral Line Shapes in Astrophysics, Banja Koviljaca, Serbia, May 13-17, 2013, Book of abstracts, p. 52, http://www.scslsa.matf.bg.ac.rs/Book_of_abstracts_9thSCSLSA.pdf

- Pearson, K.: 1895, *Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material.* Philosophical Transactions of the Royal Society of London, **186**, 343 – 414.

- Sir Ronald Aylmer Fisher (1890-1962) for the first time presented the idea in 1912 (when he was 22 years old) in the article: On an absolute criterion for fitting frequency curves, Messenger of Mathematics (1912), **41,** 155-160.