### Data Reduction and Analysis Techniques

**Ronald J. Maddalena**

www.nrao.edu/~rmaddale/Education

---

**Continuum - Point Sources**

*On-Off Observing*

- Observe blank sky for 10 sec
- Move telescope to object & observe for 10 sec
- Move to blank sky & observe for 10 sec
- Fire noise diode & observe for 10 sec
- Observe blank sky for 10 sec

---

**Continuum - Point Sources**

*On-Off Observing*

- Known:
  - Equivalent temperature of noise diode or calibrator ($T_{cal}$) = 3 K
  - Bandwidth ($\Delta \nu$) = 10 MHz
  - Gain = 2 K / Jy

- Desired:
  - Antenna temperature of the source ($T_A$)
  - Flux density ($S$) of the source.
  - System Temperature ($T_s$) when OFF the source
  - Accuracy of antenna temperature ($\sigma_{TA}$)

---

**Continuum - Point Sources**

*On-Off Observing - noise estimate*

1. Write down data analysis equation:

$$T_s = \frac{T_{cal} \cdot P_{off,cal} - P_{off,cal} \cdot T_{cal,off}}{P_{off,cal} - P_{off,cal}}$$

2. Use “propagation of errors”:

$$\sigma_{T_A} = \sqrt{\sum (\frac{\partial T_A}{\partial P})^2 \sigma_P^2}$$

3. Use the following substitutions:

$$\sigma_P = \frac{1}{\sqrt{\Delta \nu t}}$$

---

**Continuum - Point Sources**

*On-Off Observing*

![Graph showing data analysis and noise estimation](image-url)
Continuum - Point Sources
On-Off Observing – noise estimate

\[ T_A = \frac{P_{\text{cal, on}}}{P_{\text{cal, off}}} \left( P_{\text{signal, on}} - P_{\text{signal, off}} \right) \]

\[ \sigma_{\text{cal}}^2 = \sum \left( \frac{\partial T_A}{\partial P_{\text{cal}}} \right)^2 + \left( \frac{\partial T_A}{\partial P_{\text{signal}}} \right)^2 \]

\[ \sigma_{\text{signal}}^2 = \left( \frac{T_{\text{cal}}}{P_{\text{cal, on}} - P_{\text{cal, off}}} \right)^2 \left( \sigma_{\text{cal}}^2 + \sigma_{\text{signal}}^2 \right) \]

\[ \text{SNR} = \frac{1}{\sqrt{103 + 30}} \left( 10^2 \right) \sim 900 \quad \text{(Not 3000!)} \]

Continuum - Point Sources
Assumptions:

- “Classical” Radiometer equation assumes:
  - Narrow bandwidths,
  - Linear power detector,
  - \( T_A \ll T_s \),
  - Noise diode temperature \( \ll T_s \),
  - \( t_{\text{ref, on}} = t_{\text{signal}} \)
  - \( t_{\text{cal, on}} = t_{\text{cal, off}} \)
  - Blanking time \( \ll t_{\text{signal}} \)
  - No data reduction!

Phases of an Observation
Total Power

\[ \text{SNR} = \frac{1}{\sqrt{103 + 30}} \left( 10^2 \right) \sim 900 \quad \text{(Not 3000!)} \]

- \( T_{\text{cal}} = 4 \text{ K} \)
- \( T_s = 100 \text{ K} \)

- Shapes very similar
- Excess noise from atmospheric fluctuation
**Phases of an Observation**

**Double Beam Switched Power**

- **Detector Ref Cal On**
- **Detector Ref Cal Off**
- **Detector Sig Cal On**
- **Detector Sig Cal Off**

- **Signal Switching**

**Continuum - Point Sources**

**Beam-Switched Observation**

\[
T_s^{\text{reference}}(i) = \left( \frac{T_{\text{ref}} - T_{\text{cal}}}{T_{\text{ref}} - T_{\text{cal}}} \right)^{1/2} \left( T_{\text{cal}}(i) + T_{\text{cal}}(i) \right) \]

\[
T_s^{\text{signal}}(i) = \left( \frac{T_{\text{ref}} - T_{\text{cal}}}{T_{\text{ref}} - T_{\text{cal}}} \right)^{1/2} \left( T_{\text{cal}}(i) + T_{\text{cal}}(i) \right) \]

\[
T_s(i) = T_s^{\text{signal}}(i) - T_s^{\text{reference}}(i) \]

**Continuum - Point Sources**

**On-The-Fly Observation**

\[
T_s(i) = \left( T_{\text{cal}} - T_{\text{ref}} \right) \left( T_{\text{cal}}(i) + T_{\text{cal}}(i) \right) \]

If beam-switching (switched power):

\[
T_s^{\text{reference}}(i) = \left( \frac{T_{\text{ref}} - T_{\text{cal}}}{T_{\text{ref}} - T_{\text{cal}}} \right)^{1/2} \left( T_{\text{cal}}(i) + T_{\text{cal}}(i) \right) \]

\[
T_s^{\text{signal}}(i) = \left( \frac{T_{\text{ref}} - T_{\text{cal}}}{T_{\text{ref}} - T_{\text{cal}}} \right)^{1/2} \left( T_{\text{cal}}(i) + T_{\text{cal}}(i) \right) \]

\[
T_s(i) = T_s^{\text{signal}}(i) - T_s^{\text{reference}}(i) \]

**Baseline Fitting**

**Polynomials**

- Set order of polynomial
- Define areas devoid of emission.
- Creates false features
- Introduces a random error to an observation

\[
\sigma_{\text{peak}}^2 = \sigma_s^2 + \sigma_{\text{polynomial}}^2
\]

Why Polynomials?

**Continuum - Point Sources**

**Gaussian Fitting**

- Define initial guesses
- Set flags to fit or hold constant each parameter
- Set number of iterations
- Set convergence criteria

- Fitted parameters
- Chi-square of the fit
- Parameter standard deviations.

- Restrict data to between the half power points for fitting to a telescope’s beam
- Multi-component fits should be done simultaneously
Continuum - Point Sources

**Gaussian Fitting**

- Where is noise the highest?
- Where is noise the lowest?

- $\sigma$ changes across the observation.
- Weights ($1/\sigma^2$) for least-square-fit changes across the observation.
- For strong sources, should worry about using proper weights in data analysis.

---

**Template Fitting**

- Create a template:
  - Sufficient knowledge of the telescope beam, or
  - Average of a large number of observations.
- Convolve the template with the data, shift by the x-offset.
- Perform a linear least-square fit of the template to the data.

Always try to fit physically-meaningful functions.

---

**Averaging Data / Atmosphere**

- $T_s$ changes due to atmosphere emission.
- Use weighted average with weights = $1/\sigma^2$:
  \[
  (T_s) = \frac{\sum_j T_j}{\sum_j 1/\sigma_j^2} = \frac{1}{N} \sum_j \frac{T_j}{\sigma_j^2}
  \]
- $T_s$ changes due to atmosphere opacity.
- Opacity from the literature or theory, from a tipping radiometer, from atmospheric vertical water vapor profiles, or by "tipping" the antenna:
  \[
  T_A = T_A e^{\tau_s/\sin(\theta)} \quad \sigma_A = \sigma_A e^{\tau_s/\sin(\theta)}
  \]

---

**Gain Correction**

- Convert Power into $T_S$.
- Fit baseline to each row?
- Grid into a matrix.

- Stripping (Emerson 1995; Klein and Mack 1995).
- If beam-switched, Emerson, Klein, and Haslam (1979) to reconstruct the image.
- Make multiple maps with the slew in different direction.

---

**Continuum - Extended Sources**

**On-The-Fly Mapping**

- Telescope slews from row to row. Row spacing: ~0.9 /2D
- A few samples /sec.
- Highly oversampled in direction of slew <0.3A/2D
- Could be beam switching.
- Convert Power into $T_S$.
- Fit baseline to each row?
- Grid into a matrix.

---

**Continuum - Extended Sources**

**On-The-Fly Mapping - Common Problems**

- Stripping (Emerson 1995; Klein and Mack 1995).
- If beam-switched, Emerson, Klein, and Haslam (1979) to reconstruct the image.
- Make multiple maps with the slew in different direction.
Spectral-line - Point Sources
On-Off Observing

- Observe blank sky for \( t_{\text{blank}} \) sec
- Fire noise diode to determine \( T_s \)
- Move telescope to object & observe for \( t_{\text{signal}} \) sec
- Can observe an extended source using this technique -- ‘signal’ observations arranged in a ‘grid’ map.
**Spectral-Line - Point Sources**

**Position-Switched Observing**

\[
T_i(f) = T_i^{\text{position switched}}(f)
\]

- Smoothed/averaged \( T_i \) of Denominator
- Signal expected
- Reference (No line expected)

\[
T_i^{\text{continuum}}(f) = \left( T_i^{\text{position switched}}(f) \right) \frac{P_{\text{continuum}}^{\text{off}}(f) + P_{\text{continuum}}^{\text{off}}(f + \Delta f)}{P_{\text{continuum}}^{\text{off}}(f) + P_{\text{continuum}}^{\text{off}}(f + \Delta f)} \quad \text{(Channels)}
\]

\[
\sigma_{T_i} = \frac{K}{\Delta V / N_{\text{channels}}} \left( \frac{1}{\sigma_{T_i}^2} + \frac{1}{\Delta V} \right) \left( \sigma_{T_i}^2 \right)
\]

- But only for weak lines and no strong continuum!
- Constant depends upon details of the detecting backend

**Phases of a Observation**

**Switched Power – Frequency Switching**

- Signal
- Frequency
- Reference

- Local Oscillator

- Detector
- Calibration

- On

- Off

**Spectral-Line - Point Sources**

**Frequency-Switched Observing - In band**

Line appears twice – should be able to ‘fold’ the spectra to increase SNR

**Spectral-Line - Point Sources**

**Frequency-Switched – “Folding” In Band**

\[
T_i(f) = T_i^{\text{frequency switched}}(f) \left( \frac{P_{\text{continuum}}^{\text{on}}(f) + P_{\text{continuum}}^{\text{on}}(f + \Delta f)}{P_{\text{continuum}}^{\text{off}}(f) + P_{\text{continuum}}^{\text{off}}(f + \Delta f)} + \frac{P_{\text{continuum}}^{\text{off}}(f) + P_{\text{continuum}}^{\text{off}}(f + \Delta f)}{P_{\text{continuum}}^{\text{on}}(f) + P_{\text{continuum}}^{\text{on}}(f + \Delta f)} \right)
\]

**Spectral-Line**

**Baseline Fitting**

- Polynomial: same as before
- Sinusoid

**Spectral-Line**

**Other Algorithms**

- Velocity Calibration
- Velocity/Frequency Shifting & Regridding
  - Doppler tracking limitations
  - Smoothing – Hanning, Boxcar, Gaussian
    - Decimating vs. non-decimating routines
    - For "Optimal Filtering", match smoothing to expected line width
- Filtering – low pass, high pass, median, ...
- Moments for Integrated Intensities; Velocity centroids, ...
The Future of Single-Dish Data Analysis

- Increase in the use of RDBMS.
- Support the analysis of archived data.
- Sophisticated visualization tools.
- Sophisticated, robust algorithms (mapping).
- Data pipelining for the general user.
- Automatic data calibration using models of the telescope.
- Algorithms that deal with data sets.
- Analysis systems supported by cross-observatory groups.
- More will be done with commercial software packages.