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 ($T_{\text{cal}}$) = 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 ($\sigma_{T_A}$)
Continuum - Point Sources

On-Off Observing

\[ T_{\text{reference}} = \frac{T_{\text{cal}} \cdot P_{\text{cal, off}}}{P_{\text{cal, on}} - P_{\text{cal, off}}} \]

20 K

\[ T_{\text{signal}} = \frac{T_{\text{cal}} \cdot P_{\text{signal}}}{P_{\text{cal, on}} - P_{\text{cal, off}}} \]

26 K

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

6 K

\[ \sigma_{T_A} = \frac{T_A}{\Delta t} \]

\[ \text{SNR} = 3000 \]
### Continuum - Point Sources

**On-Off Observing – noise estimate**

\[
T_s = \frac{T_{\text{cal on}}}{P^\text{reference}_{\text{cal on}}} - \frac{T_{\text{cal off}}}{P^\text{reference}_{\text{cal off}}} \left( P_{\text{signal}}^{\text{cal off}} - P_{\text{reference}}^{\text{cal off}} \right)
\]

\[
\sigma_{T_s}^2 = \sum \left( \frac{\partial T_s}{\partial P} \right)^2 \sigma_P^2 = \left( \frac{\partial T_s}{\partial P_{\text{signal}}} \right)^2 \sigma_{P_{\text{signal}}}^2 + \left( \frac{\partial T_s}{\partial P_{\text{reference}}} \right)^2 \sigma_{P_{\text{reference}}}^2 + \left( \frac{\partial T_s}{\partial P_{\text{cal on}}} \right)^2 \sigma_{P_{\text{cal on}}}^2
\]

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

\[
\left( \frac{1}{\text{SNR}} \right)^2 = \left( \frac{\sigma_{T_s}}{T_s} \right)^2 = \left[ \left( \frac{P_{\text{reference}}^{\text{cal on}}}{P_{\text{reference}}^{\text{cal off}}} - \frac{P_{\text{cal on}}^{\text{cal off}}}{P_{\text{cal off}}^{\text{cal off}}} \right)^2 + \left( \frac{P_{\text{signal}}^{\text{cal off}}}{P_{\text{cal off}}^{\text{cal off}}} - \frac{P_{\text{cal on}}^{\text{cal off}}}{P_{\text{cal off}}^{\text{cal off}}} \right)^2 \right] \cdot \left( \frac{1}{\Delta V \cdot t} \right)
\]

\[
\text{SNR} = \frac{1}{\sqrt{103+30 \cdot (10^4)}} \approx 900 \quad \text{(Not 3000!)}
\]

### Assumptions:

“Classical” Radiometer equation assumes:

- Narrow bandwidths,
- Linear power detector,
- \( T_A << T_s \),
- Noise diode temperature \( << T_s \),
- \( t_{\text{reference}} = t_{\text{signal}} \),
- \( t_{\text{cal on}} = t_{\text{cal off}} \),
- Blanking time \( << t_{\text{signal}} \),
- No data reduction!
Phases of an Observation

Total Power

• $T_{\text{cal}} = 4 \text{ K}$
• $T_s = 100 \text{ K}$

• $\sigma_{\text{theor}} = 0.1 \text{ K}$
• $\sigma_{\text{meas}} = 1 \text{ K}$

• Shapes very similar
• Excess noise from atmospheric fluctuation
Phases of a Observation

Beam Switched Power
Phases of a Observation

Double Beam Switched Power

Continuum - Point Sources

Beam-Switched Observation

\[ T_{S_{\text{reference}}} (i) = \frac{T_{\text{cal}} (i) - P_{\text{cal}_{-\text{off}}} (i)}{P_{\text{cal}_{-\text{on}}} (i) - P_{\text{cal}_{-\text{off}}} (i)} \left( \frac{P_{\text{cal}_{-\text{on}}} (i) + P_{\text{cal}_{-\text{off}}} (i)}{2} \right) \]

\[ T_{S_{\text{signal}}} (i) = \frac{T_{\text{cal}} (i) - P_{\text{cal}_{-\text{off}}} (i)}{P_{\text{cal}_{-\text{on}}} (i) - P_{\text{cal}_{-\text{off}}} (i)} \left( \frac{P_{\text{signal}_{-\text{on}}} (i) + P_{\text{signal}_{-\text{off}}} (i)}{2} \right) \]

\[ T_{A} = \frac{T_{S_{\text{signal}}} (i) - T_{S_{\text{reference}}} (i)}{2} \]
Continuum - Point Sources
On-The-Fly Observation

If total power:

\[ T_S(i) = \frac{\left( \frac{T \text{\_cal}}{P \text{\_cal\_on}(i) - P \text{\_cal\_off}(i)} \right)}{2} \left( P \text{\_cal\_on}(i) + P \text{\_cal\_off}(i) \right) \]

If beam-switching (switched power):

\[ T_S\text{\_reference}(i) = \frac{\left( \frac{T \text{\_cal}}{P \text{\_cal\_on}(i) - P \text{\_cal\_off}(i)} \right)}{2} \left( P \text{\_cal\_on}(i) + P \text{\_cal\_off}(i) \right) \]

\[ T_S\text{\_signal}(i) = \frac{\left( \frac{T \text{\_cal}}{P \text{\_cal\_on}(i) - P \text{\_cal\_off}(i)} \right)}{2} \left( P \text{\_cal\_on}(i) + P \text{\_cal\_off}(i) \right) \]

*\[ T_A(i) = T_S\text{\_signal}(i) - T_S\text{\_reference}(i) \]*
Baseline Fitting
Polynomials

- Set order of polynomial
- Define areas devoid of emission.

- Creates false features
- Introduces a random error to an observation

\[ \sigma_{\text{Peak}}^2 = \sigma_{T_a}^2 + \sigma_{\text{Polynomial}}^2 \]

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?

- $\sigma$ 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.

Where is noise the lowest?

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 => x-offset.
- 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

- $T_A$ changes due to atmosphere emission.
- Use weighted average with weights $= 1/\sigma^2$

$$\langle T_A \rangle = \frac{\sum T_i \frac{1}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2}} \quad \sigma_{avgr} = \frac{1}{\sqrt{\sum \frac{1}{\sigma_i^2}}}$$

- $T_A$ 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

$$T_A^* = T_A \cdot e^{\tau A / \sin(e)} \quad \sigma_{TA}^* = \sigma_{TA} \cdot e^{\tau A / \sin(e)}$$

Gain Correction

$$T_A^* = T_A^*/\eta_A \quad \text{or} \quad T_B^* = T_A^*/\eta_M$$
$$\sigma_{TA}^* = \sigma_{TA}^*/\eta_A \quad \text{or} \quad \sigma_{TB}^* = \sigma_{TA}^*/\eta_M$$
Continuum - Extended Sources
On-The-Fly Mapping

- Telescope slews from row to row. Row spacing: ~0.9 \( \lambda / 2D \)
- A few samples /sec.
- Highly oversampled in direction of slew <0.3\( \lambda / 2D \)
- 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

- Striping (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.
GBT Continuum Images – Rosette

The Rosette at 8.4GHz (GBT Oct 24, 2002)

GBT Continuum Images – M17

Omega Nebula 8.4GHz, Feb 9, 2002
GBT Continuum Images – W3

GBT Continuum Images - Orion
Spectral-line - Point Sources
On-Off Observing

- Observe blank sky for $t_{\text{reference}}$ 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
Spectral-Line - Point Sources
Position-Switched Observing

\[
T_A(f) = T_s^{\text{reference}}(f) \begin{bmatrix} P_{\text{signal}} - P_{\text{reference}} \\ P_{\text{reference}} \end{bmatrix}
\]

- \text{Smoothed/Averaged } T_s \text{ of Denominator}
- \text{Signal (line expected)}
- \text{Reference (No line expected)}

\[
T_s^{\text{reference}}(f) = \left( \frac{T_{\text{cal}}}{2} \right) \left( \frac{P_{\text{reference}}(f) + P_{\text{cal, on}}(f)}{P_{\text{cal, on}}(f) - P_{\text{cal, off}}(f)} \right)_{M \text{ Channels}}
\]

\[
\left( \frac{\sigma_{T_s}}{T_A} \right)^2 \approx \frac{K}{\Delta \nu / N_{\text{channels}}} \left( \frac{1}{T_{\text{cal, on}}} + \frac{1}{T_{\text{cal, off}}} + \frac{\sigma_{T_s}}{T_s} \right)^2
\]

- 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 Frequency
- Local Oscillator

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

\[ T_A = T_s \text{REF}^n \frac{[\text{SIG-REF}]^{\text{REF}}}{\text{REF}} \]

\[ T_A = T_s \text{REF}^n \frac{[\text{SIG-REF}]^{\text{REF}}}{\text{REF}} \]

\[ T_s = T_s \text{REF}^n \frac{[\text{SIG-REF}]^{\text{REF}}}{\text{REF}} \]

\[ T_s = T_s \text{REF}^n \frac{[\text{SIG-REF}]^{\text{REF}}}{\text{REF}} \]

\[ T_s = T_s \text{REF}^n \frac{[\text{SIG-REF}]^{\text{REF}}}{\text{REF}} \]
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, ...
Spectral-Line
RFI Excision

Spectral-Line Mapping
Grid or On-the-Fly
Spectral-Line Mapping
Grid and On-the-Fly

\[ W(\alpha, \delta) = \sum_{V_i = V_{\text{min}}}^{V_{\text{max}}} T(\alpha, \delta, V_i) \cdot \Delta V_i \]
(If \( V_1 = V_2 \Rightarrow \text{Channel Map} \))

For \( \{v=v_{\text{min}}\} \ \{v <= v_{\text{max}}\} \ \{v++\} \)
if \( T(\alpha, \delta, v) > T_{\text{min}} \) then
\[ W(\alpha, \delta) = W(\alpha, \delta) + T(\alpha, \delta, v) \]
endif
endfor

Spectral-Line Mapping
Grid and On-the-Fly

\[ T(\alpha, V) = \sum_{\delta = \delta_{\text{min}}}^{\delta_{\text{max}}} T(\alpha, \delta, V) \]
(Position-velocity map)
Spectral-Line Mapping

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