Reduction and Analysis Techniques

Ronald J. Maddalena
National Radio Astronomy Observatory

Abstract

Single-dish observations can be made in myriad of ways with each observing technique almost always requiring different kinds of data analysis. This chapter will cover the basic continuum and spectral line analysis algorithms common to all single-dish data analysis packages. However, I will not cover the very specialized fields of polarimetry, pulsar, or radar data reduction. In the case of continuum observations the student will learn the steps used to derive the flux of a point source as well as the more common data analysis techniques for generating and analyzing maps of extended sources. For spectral line data, I will concentrate on the analysis algorithms usually applied to single spectra (bandpass and velocity calibration, data averaging and smoothing, baseline fitting, component fitting,...) and how to produce and analyze spectral-line data cubes.

Introduction

Reducing and analyzing single-dish data is similar to reducing other kinds of astronomical data in that it is a time consuming process. Typically, one spends more time analyzing data then taking it. And, in the same way that the reduction of X-ray data differs from the reduction of optical spectra, each type of single-dish data requires different reduction algorithms.

A full description of the gamut of single-dish data reduction techniques would require a textbook a few times the size of the manuals that accompany single-dish software packages. In this introductory paper to single-dish analysis algorithms, I will only have space to concentrate on the common algorithms found in most packages. I will provide the basics so that the reader will be able to effectively start using any analysis system. No attempt will be made to describe the use of any individual software package nor will I give any description of pulsar, polarimetry, or radar analysis techniques.

In the following, I will first list some of the more common data analysis systems for single-dish telescopes (§2). I will then discuss the analysis of continuum (§3) and spectral line (§4) observations. Many of the spectral-line algorithms explained in §4 require an understanding of the continuum algorithms presented in §3. Section 5 gives my thoughts on the future of single-dish data analysis. An appendix provides definitions of the symbols used throughout.

Single-Dish Data Analysis Packages

Each telescope or observatory typically has its own single-dish data reduction system. Users of a telescope should be prepared to learn a new analysis system whenever they observe on a different telescope.

The following table is an incomplete list of single-dish analysis packages and telescopes where they are used.

Table 1: Common Single-Dish Analysis Packages

Package	Telescopes	Web Address
Aips	NRAO 12m	www.cv.nrao.edu/aips
Aips++	NRAO 100m	www.aips2.nrao.edu/docs/aips++.html
Analyz	NAIC 305m	www.naic.edu/menuimag/astronomy.htm
CLASS	IRAM 30m and others	www.iram.fr/doc/class/class.html
IDL	NAIC 305m and others	www.idlastro.gsfc.nasa.gov/homepage.html
SPECX	JCMT	www.jach.hawaii.edu/JACpublic/JCMT/User_documentation/SPECX/part6/node2.html
UniPops	NRAO 43m and 12m	www.cv.nrao.edu/unipops

Each of the above systems differs in the powers and algorithms provided. Since the same class of analysis algorithms in various packages will make different simplifying assumptions, the user cannot expect fully rigorous algorithms in all packages. Some packages have a steep learning curve while others will be easy to master. Discussions of the merits and flaws of different analysis systems, like discussions on religion and politics, are usually very fervent.

Common Traits of Different Analysis Packages

Although differences between analysis systems are major, one can still find some commonalities. For example, almost all still use command line interfaces with graphical user interfaces still relegated to future development. Most analysis systems have some scripting ability to package often-used commands for batch processing.

The structure of the data that each telescope produces, although different in its details, consists not only of a vector or array of data values but also information on how the data were taken, the telescope environment (e.g., weather conditions), and a record of the time the data were taken. Some analysis algorithms often will use some of this extra information.

Most analysis systems have rudimentary data base management systems that allow the user to peruse the collected data and specify selection criteria that the program will use to define a data set to be processed. Some analysis systems are tied into telescope control systems so that it is possible to look at and analyze data in near real time.

All of the most recent systems both read and write FITS files. FITS files should make it simple to export and import data between systems; however, to this day, the exchange of data between some systems still remains painful if at all possible.

Packages that reduce continuum or line data but not both are no longer prevalent. Many systems no longer just analyze one-dimensional data vectors but can analyze and visualize two-dimensional matrices and spectral-line data cubes. The palette of tools presented to the user is almost the same from system to system, only the names of the commands and details about the algorithms they use are different.

Continuum Data Reduction Techniques

The analysis of continuum observations depends somewhat on whether the object observed is a point or extended source. Many of the techniques used in the analysis of point sources will come

in handy when dealing with extended sources. Understanding continuum data reduction will facilitate learning how to reduce spectral-line data.

Continuum Observations of Point Sources

The analysis of observations of point sources will provide us with the rudiments of how one should reduce more complicated types of data. Continuum observations of point sources can be made in many ways of which I will discuss only three:

On-Off Continuum Observations:

In On-Off observations, the telescope records data when the telescope is on the source's position, and then a few beam widths off the source. No data are recorded when the telescope is moving between the signal and reference positions. Such observations are useful for measuring the flux of a source when the pointing of the telescope is well known and the frequency low enough that gain or atmospheric variations are of no concern. In essence, it is only successfully used by cm-wave telescope.

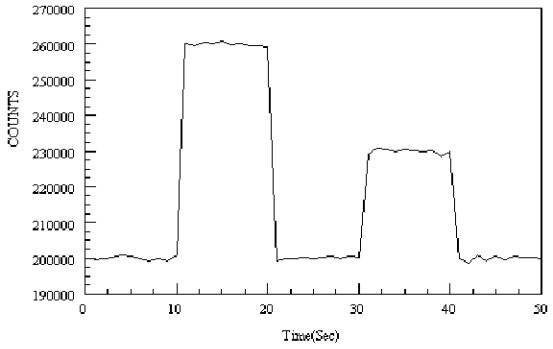

Figure 1: A typical On-Off observation before any data reduction or calibration. The first rise in power is due to a source; the second is due to the turning on of the noise diode or receiver calibration source. From such an observation, one can determine the source flux relative to the calibration source as well as the system temperature.

The difference in power detected by the telescope between the signal and reference position can then be compared to the difference in power between when the noise diode or receiver calibration source is on and off. In many analysis systems, one will find a command to automatically reduce On-Off observations.

The user should be warned that most algorithms make some simplifying assumptions about the observing method and the object observed. If one assumes:

- narrow bandwidths,
- linear power detector,
- source intensity that is a small fraction of the system temperature,
- noise diode temperature that is a small fraction of the system temperature,
- equal time on the reference and source position,
- equal time with the noise diode on and off, and
- insignificant backend blanking time between reference and signal phases and between calibration on and off phases,

then the antenna temperature of the source is given by:

$$T_{src} = \frac{T_{cal}}{P^{reference}_{cal\_On} - P^{reference}_{cal\_Off}} \cdot \frac{\left(P^{signal}_{cal\_On} + P^{signal}_{cal\_Off} - P^{reference}_{cal\_On} - P^{reference}_{cal\_Off}\right)}{2} . \quad (1)$$

The resulting data has a theoretical noise of:

$$\sigma_{T_{src}} = \left(T_{sys} / \sqrt{BW}\right) \cdot \sqrt{\frac{1}{t^{reference}} + \frac{1}{t^{signal}}} \quad (2)$$

where

$$T_{sys} = \frac{T_{cal}}{2} \cdot \left(\frac{\left(P^{reference}_{cal\_Off} + P^{reference}_{cal\_On}\right)}{\left(P^{reference}_{cal\_On} - P^{reference}_{cal\_Off}\right)} - 1 \right) = \frac{T_{cal} P^{reference}_{cal\_Off}}{\left(P^{reference}_{cal\_On} - P^{reference}_{cal\_Off}\right)} . \quad (3)$$

If any of the assumptions is not true, then the above equations need to be modified. A robust formalization of the above for the general case can be found at: http://www.gb.nrao.edu/~rmaddale/GBT/continuumCal.html.

T_{src} must be corrected for atmospheric attenuation and antenna efficiency, and then converted to flux density. The details on how to apply the corrections are covered in other chapters in this volume.

Beam-Switched Continuum Observations:

On many single-dish telescopes that use Cassegrain or Gregorian optics, the secondary or other small mirror in the system will chop at a rate of a few Hz between two positions on the sky. The positions are usually labeled 'signal' and 'reference'. Or, a telescope receiver might have two feeds that point at slightly separated positions on the sky and a switch that toggles the detected signal between the signal and reference feeds.

The data are collected synchronously in that the data for the signal and reference positions are accumulated separately. Such beam-switched observations are useful for removing fluctuation in power due to the atmosphere at frequencies typically above 10 GHz.

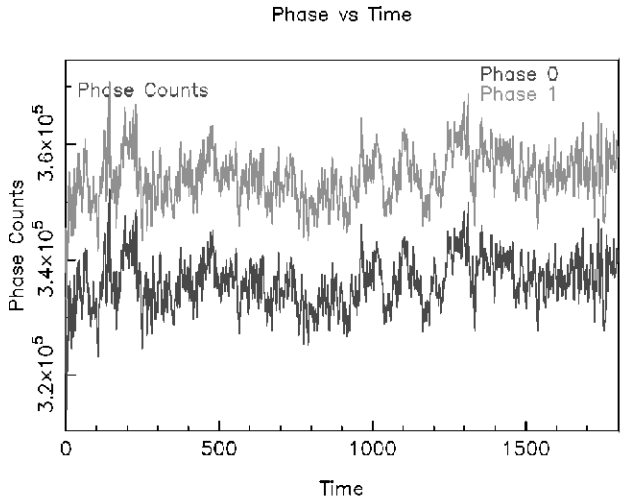

Figure 2: Typical beam-switched observation. Phase 0 data are those with the beam of the telescope on blank sky while Phase 1 are those with the beam on the source. The fluctuations in both phases are extremely similar and typically due to atmospheric fluctuations. The difference between the phases will be proportional to the strength of the source and will have less fluctuation than then either phase's data.

The data are usually represented as a vector of sometimes hundreds of values. The x-axis of the vector is the time at which data were sampled. The y-axis will need to be converted from sampled powers into antenna temperatures (T_{ant}) for both the signal and reference phases of the observations. To derive T_{ant}, the power of each data sample is compared to the difference in power between when the noise diode or receiver calibration source is on and off. If we make similar assumptions as for On-Off observations, then for each data sample i:

$$T_{ant}^{reference}(i) = \left\langle \frac{T_{cal}}{P_{cal\_on}^{reference}(i) - P_{cal\_off}^{reference}(i)} \right\rangle \cdot \frac{\left(P_{cal\_on}^{reference}(i) + P_{cal\_off}^{reference}(i)\right)}{2} \qquad (4)$$

$$T_{ant}^{signal}(i) = \left\langle \frac{T_{cal}}{P_{cal\_on}^{signal}(i) - P_{cal\_off}^{signal}(i)} \right\rangle \cdot \frac{\left(P_{cal\_on}^{signal}(i) + P_{cal\_off}^{signal}(i)\right)}{2} \qquad (5)$$

where the angle brackets represents an average over all data samples. Then,

$$T_{src} = \left\langle T_{ant}^{signal}(i) - T_{ant}^{reference}(i) \right\rangle . \qquad (6)$$

T_{src} must next be corrected for atmospheric attenuation and antenna efficiency, and then converted to flux density.

On-the-Fly Continuum Observations:

In on-the-fly (OTF) observing, the telescope slews across a source and simultaneously records data at a rate of up to a few samples per second. The extent of the slew is typically a few beamwidths. Such an observing technique is useful when one wants to determine the flux of the source or the telescope pointing and beam shape along a certain direction.

The data are usually represented as a vector of sometimes hundreds of points. The x-axis of the vector can be considered either the times or telescope positions at which data were sampled.

The y-axis will need to be converted from sampled powers into antenna temperatures (T_{ant}) before one can derive source temperature, beam parameters, and telescope pointing offset.

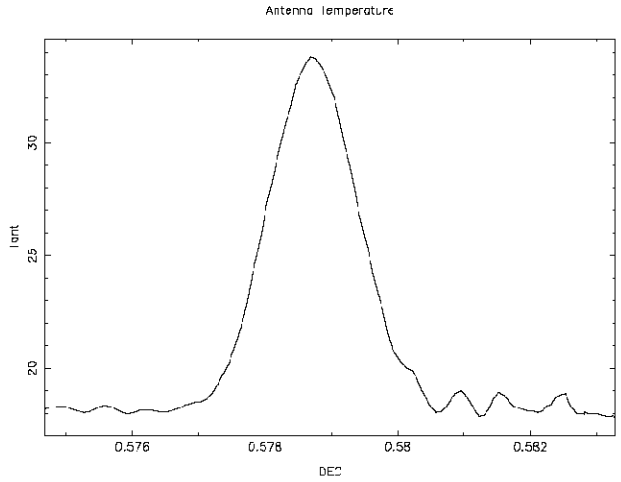

Figure 3: An on-the-fly observation through a source. The data have already been calibrated and converted to T_{ant}. The slew direction was in Declination, as implied by the x-axis of the plot.

To derive T_{ant}, the power of each data sample is compared to the difference in power between when the noise diode or receiver calibration source is on and off. If we make similar assumptions as for On-Off observations, then for each data sample i:

$$T_{ant}(i) = \left\langle \frac{T_{cal}}{P_{cal\_on}(i) - P_{cal\_off}(i)} \right\rangle \cdot \frac{\left(P_{cal\_on}(i) + P_{cal\_off}(i)\right)}{2}. \qquad (7)$$

Some telescopes combine on-the-fly observations with beam switching in which case we would use equations 4 and 5 directly while equation 7 becomes:

$$T_{ant}(i) = T_{ant}^{signal}(i) - T_{ant}^{reference}(i). \qquad (8)$$

The following analysis steps can then performed on $T_{ant}(i)$:

Averaging:

If the same observation is repeated a number of times, you will want to average the data to reduce the noise. In some cases the observations could have been taken under various conditions like different integration times, different bandwidths, and different weather conditions. To produce an optimum average with the highest signal to noise, the data reduction system must perform a weighted average. The weights should be $1/\sigma^2$ where σ is either the theoretical rms (equation 2) or measured rms of the data. The student should examine any textbook on statistics for details on weighted averages (e.g., Mandel 1964; Bevington and Robinson 1992).

However, there are problems with using either the theoretical or measured σ. The theoretical σ may be lower than the actual σ (due to such factors as receiver 1/f noise, atmospheric fluctuations) and the measured σ may be higher (due to slow drifts from things like changing ground pickup). It is usually up to the observer to determine which weights to us.

All analysis systems have statistical tools for determining the measured σ. The only thing the user needs to define is the source-free data samples to use in the calculation of σ.

After averaging, what should one then use as the σ of the observation? If you used the measured σ, then you should probably use the same statistical tools to remeasure the σ of the average. If you use the theoretical σ, the σ of the average will be:

$$\sigma_{avrg} = \frac{1}{\sqrt{\sum_j \frac{BW_j \cdot t_j}{Tsys_j^2}}} = \frac{1}{\sqrt{\sum_j \frac{1}{\sigma_j^2}}} \qquad (9)$$

Baseline Fitting:

The derived $T_{ant}(i)$ include not only variations due to sources but any drifts in power due to changes in the atmosphere or ground pickup. The system temperature also introduces a DC offset to the data. Fitting a low-order polynomial to areas of the data thought to be devoid of emission eliminates most of these unwanted factors.

Baseline fitting involves defining in the analysis software the samples which one thinks are devoid of emission and the order of the polynomial to fit. After fitting, the polynomial is subtracted from the data. See Press et al. (1992) or Bevington and Robinson (1992) for discussion on the details of polynomial linear least squares fit algorithms.

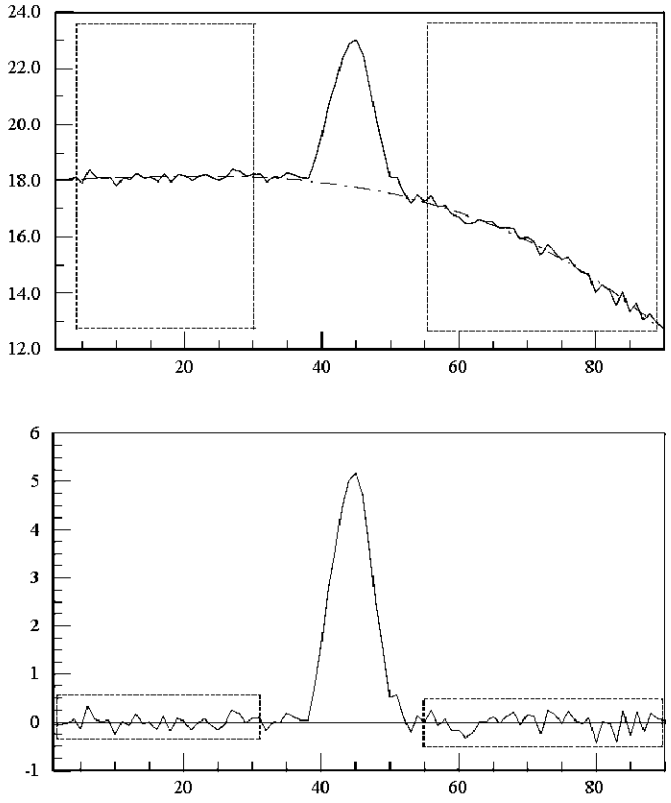

Figure 4: Illustrates baseline fitting to an on-the-fly observation. The top plot is the data before baseline removal and the bottom plot is the same data afterwards. Thin rectangles illustrate the regions thought to be devoid of emission that were used to fit a third order polynomial. The dashed line in the top plot is the fitted polynomial which, when subtracted from the data, produces the bottom plot.

Unless the data have very good signal-to-noise or the baselines start out being exceptionally flat, fitting an arbitrary functional form such as a polynomial to your data will introduce biases in any estimates of source parameters or create false features (Rood, Bania, and Wilson 1984). Two observers fitting baselines to the same data can produce differences in measured source intensities that are larger than the RMS calculated from equation 2. Thus, baseline fitting introduces a random error to an observation, due to the uncertainty of the baseline determination (Lockman 1986).

Gaussian Fitting:

Once a baseline has been removed, one can then try to determine T_{src} by fitting a Gaussian to the residuals of the baseline fit. Note that most telescope beams are approximately Gaussian only between about the beam's half power points. Using data outside of the half power points will introduce systematic errors to the calculated source fluxes and estimated beam widths.

Most analysis systems will allow you to perform the following steps:

- Restrict the data points it will use for Gaussian fitting to those between the half power points,
- Define initial guess for the Gaussian half width, center, and height,
- Set a flag to fit or hold constant the half-width of the Gaussian (or any other Gaussian parameter). This flag should be used if one already knows well the width of the telescope beam,
- Number of iterations to use in the non-linear, least squares fitting routine before the fit will be claimed to be unsuccessful,
- Set the criteria that will be used to tell if the fit has converged. Usually this is the percentage decrease in Chi-square between loops in the non-linear fitting routine.

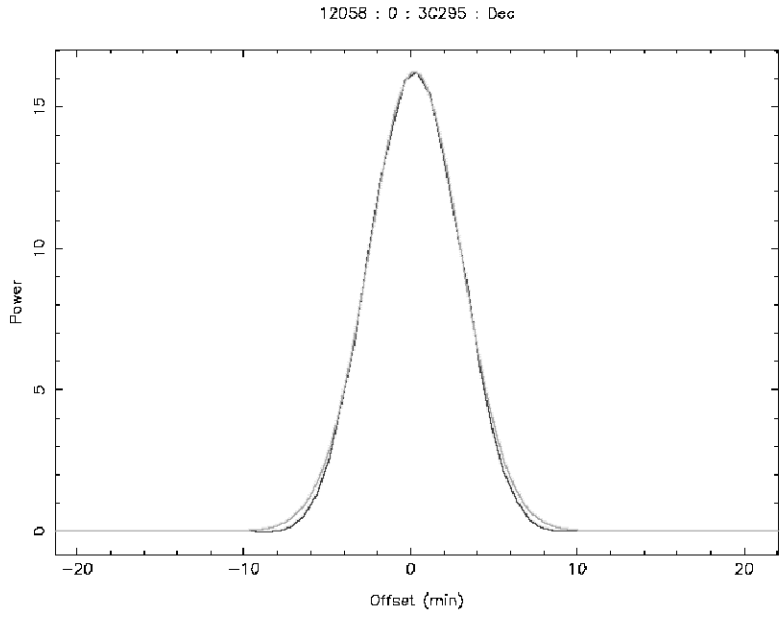

Figure 5: An on-the-fly observation taken with the Green Bank Telescope of the source 3C295. The excellent signal-to-noise of this observation illustrates how the data (dark line) and a Gaussian fitted to the data between the half-power points (thin line) differ in the wings of the Gaussian. Note that a baseline was subtracted from the data before fitting.

In addition to returning the values of the fitted parameters, most algorithms also return the Chi-square of the fit and formal standard deviations of the fitted parameters. Quang-Rieu (1969) provides a discussion of the statistical errors inherent in Gaussian fitting. See Press et al. (1992) or Bevington (1992) for discussion on the details of non-linear least squares fitting algorithms as is typically used for Gaussian fitting.

Template Fitting:

Gaussian fitting is flawed in that beams are sometimes not symmetric due to optical aberrations like coma. And, Gaussian fitting can only be performed over the subset of the data collected between the half-power points of the beam. Template fitting is one way to avoid these flaws. It harks back to the old philosophy of fitting a physical model to your data instead of arbitrary functions.

If one has sufficient knowledge of the telescope, the beam of the telescope can be determined a priori. Or, an observer can average a large number of observations to produce a high signal-to-noise representation of the telescope beam. Some analysis systems allow you to fit a template or model of the beam to the collected data.

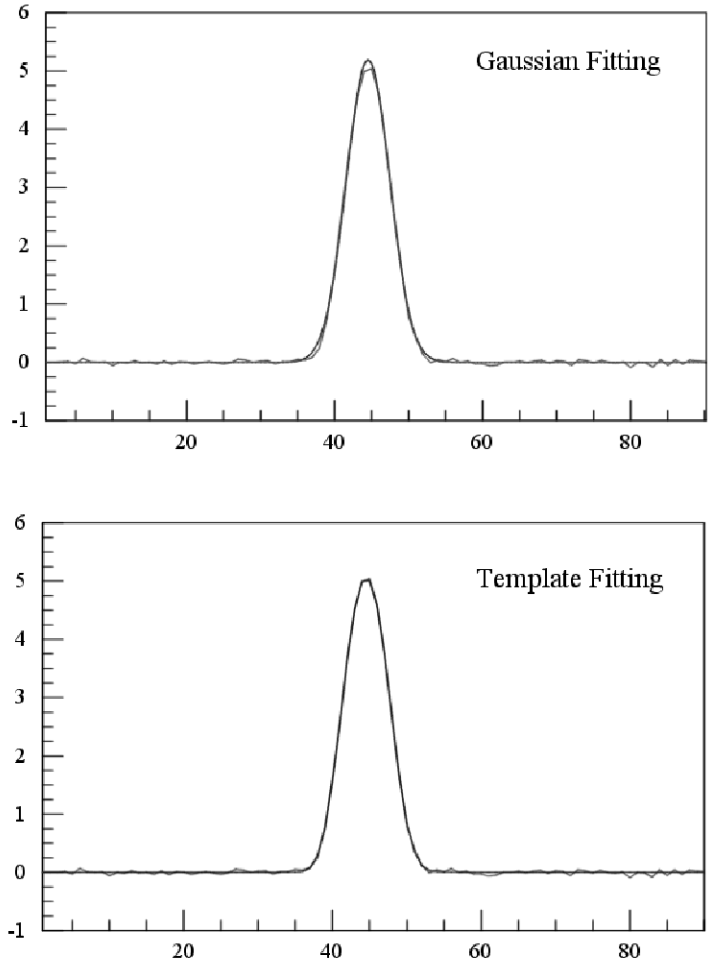

Figure 6: Comparison of Gaussian fitting (top) to template fitting for the same on-the-fly observation. Unlike the fit in Figure 5, the Gaussian fitting was across all of the data, not just between the half-power points. Note how the Gaussian fitting underestimates the intensity of the source and does not fit well the wings of the profile.

The standard algorithm usually convolves the template with the data to determine the x-axis offset of the data. The data are shifted by the offset usually by taking the Fourier transform of the data, adding a phase shift corresponding to the offset, and inverse Fourier transforming the data (Press et al. 1992). Then, the algorithm performs a linear least-squares fit of the template to the data: $T(i) = A \cdot Template(i) + B$ where A corresponds to the fitted scaling of template to the data and B to the fitted DC-offset between the template and the data. One implementation of template fitting can be found at http://aips2.nrao.edu/docs/user/Dish; see Van Zee et al. (1997) for a typical use of template fitting.

If the template is an adequate representation of the data, the results of a template fit will have fewer systematic problems. The results will be more accurate than Gaussian fitting since all of the data will be used in the fit.

Continuum Observations of Extended Sources

The previous section on point sources has provided us with the algorithms and techniques for simple analysis of continuum mapping observations of extended sources. All mapping observations of extended sources can be divided into two major categories:

Continuum Grid Mapping:

In grid mapping, the telescope steps the telescope from pixel to pixel in the map only collecting data when the telescope has settled onto a pixel. While mapping, the observing technique could be beam switching to reduce atmospheric fluctuations.

The reduction of grid mapping requires converting the detected powers of each pixel to T_{ant} using the appropriate analysis described in §3.1. The resulting T_{ant} are then regridded into a matrix. Since grid mapping is going out of vogue, we won't go into any further details.

Continuum On-the-Fly (OTF) Mapping:

OTF mapping has been used for decades by many cm-wave telescopes for strong sources. It is now the preferred technique at mm-waves since at those wavelengths OTF reduces the ill effects of atmospheric fluctuations. In OTF mapping, the telescope slews through a region of the sky and simultaneously records data at a rate of up to a few samples per second. The slewing is usually performed in one cardinal direction to produce a row or column in the map that is most often highly oversampled. Between rows or columns, the telescope is offset in the direction orthogonal to the slew direction by somewhat less than half the beam width and a new row or column is started. While mapping, the observing technique could be beam switching to further reduce atmospheric fluctuations.

As in grid mapping, the first step is to use the appropriate analysis described in §3.1 to convert from detected power to T_{ant}. If warranted, a baseline might be fitted and removed from the data for each row or column in the map separately. The data are then gridded into a matrix using techniques that depend upon the analysis package. Each gridding algorithm has its pluses and minuses. Usually a gridding algorithm performs a weighted average of all data points that fall within some radius of a matrix pixel. The weights usually depend upon the distance a data point lies from the matrix pixel with those further away getting less weight.

For example, a 'top-hat' gridding algorithm uses uniform weighting of all data points that are within a specified radius of a matrix pixel. A top-hat algorithm is useful for when you expect little structure in your map. Other gridding functions use a weight function that is representative of the telescope's beam and are better suited to noisy maps or maps full of small-scale structure. Although Briggs, Schwab, and Sramek (1998) describe gridding algorithms in the context of array imaging, much of the general details of their discussion apply to single-dish gridding algorithms as well.

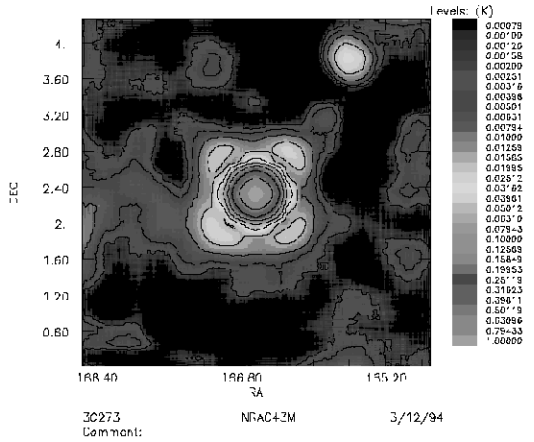

Figure 7: On-the-fly continuum map of the point source 3C273 taken with the NRAO 140-ft telescope at 1400 MHz. There is a second radio source located to the northwest of 3C273. Contours are spaced 1 dB apart and show the beam pattern or response of the telescope to a point source.

Advanced Techniques for Continuum Mapping:

Once the data in a continuum map have been gridded, some systematic problems are often evident in the data. The powers available to you will depend on the analysis system you are using and you should consult the program's manual for suggestions.

One common problem is striping along rows or columns in the matrix due to instrumental or atmospheric drifts from one row or column to the next. Fourier transforming the matrix, removing the low-order frequencies, and inverse transforming the matrix, can reduce stripping (Emerson 1995; also Klein and Mack 1995 for using the CLEAN algorithm on single-dish data).

If the data have been beam-switched, the separation of the beams on the sky might be smaller than the size of the mapped source. In the later case, certain algorithms, such as that of Emerson, Klein, and Haslam (1979) can be used to reconstruct the image.

For creating maps with the highest fidelity, one might consider making multiple maps maybe with each map taken with the telescope slewing in a different cardinal direction. Each map should be taken very quickly. The resulting maps should then be averaged using some measure of each maps σ for weights.

After a matrix is created, and any artifacts removed, many analysis programs have tools to calculate source strengths and sizes, take slices through the matrix, and remove drifts from varying ground pickup. Since different analysis programs possess different functionalities, the user must peruse the software's documentation for a list of available algorithms.

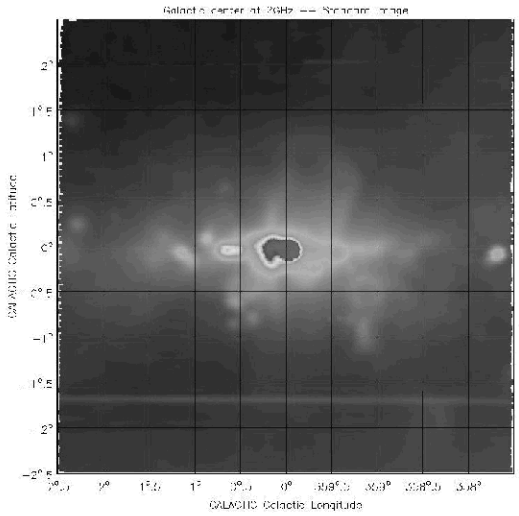

Figure 8: On-the-fly continuum map of the galactic center taken with the Green Bank Telescope at a frequency of 2 GHz. Note the stripping at galactic latitude of –1.6 that will need to be fixed either with more observations or one of the techniques described above.

Spectral-line Data Reduction Techniques

Most of the principles behind continuum data reduction also apply to spectral-line data reduction. Thus, this section will build upon what we have already discussed in the previous section.

Spectral-Line Observations of Point Sources

As with continuum observations, we will start by discussing observations of point sources before tackling spectral-line mapping.

Spectral Line Switching Schemes

There are three different categories of spectral-line observations, each depending upon how a reference spectrum is created. Two of these switching schemes, beam switching and position switching, are very similar to the beam switching and On-Off techniques described above for continuum observations. Frequency switching is commonly used only for spectral line observing.

Position-Switched Observations:

Position-switched observations are the spectral line equivalent of On-Off continuum observations. All analysis programs have some command to create a difference spectrum from the signal and reference observations. The difference spectrum is derived by comparing the difference in power between the signal and reference spectra with the difference in power between when the noise diode or calibrator is on and off. With the same assumptions we made in §3.1, the difference spectrum is:

$$T_{ant}(f) = \frac{T_{cal}(f)}{2} \cdot \left[\frac{\left(P^{signal}_{cal\_on}(f) + P^{signale}_{cal\_off}(f)\right) - \left(P^{reference}_{cal\_on}(f) + P^{reference}_{cal\_off}(f)\right)}{\left(P^{reference}_{Cal\_On}(f) - P^{reference}_{Cal\_Off}(f)\right)} \right] \tag{10}$$

$$= \left(T_{sys}(f) + \frac{T_{cal}(f)}{2} \right) \cdot \left[\frac{\left(P^{signal}_{cal\_on}(f) + P^{signale}_{cal\_off}(f)\right) - \left(P^{reference}_{cal\_on}(f) + P^{reference}_{cal\_off}(f)\right)}{\left(P^{reference}_{cal\_on}(f) + P^{reference}_{cal\_off}(f)\right)} \right]$$

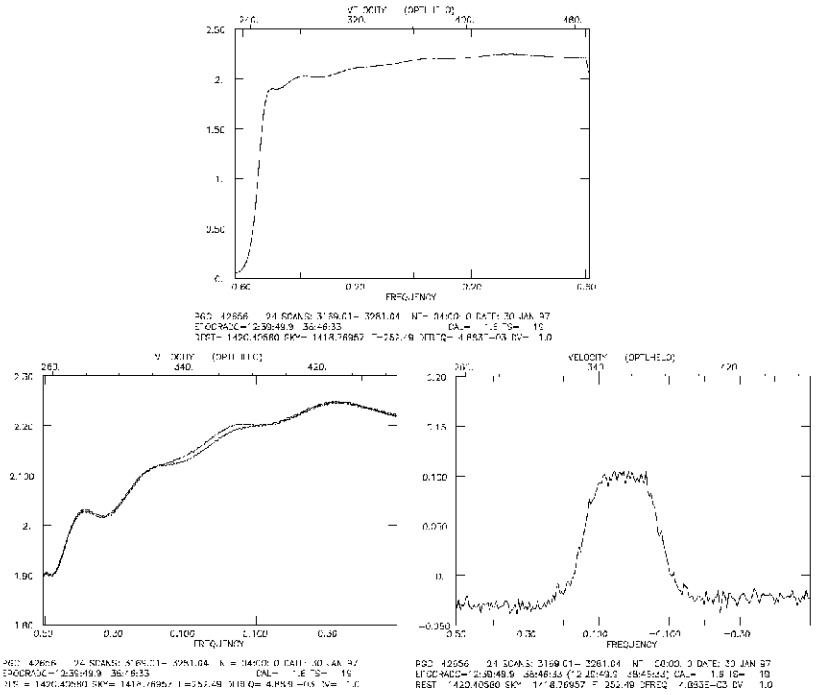

Figure 9: A position switched, atomic hydrogen observation taken with the NRAO 140-ft telescope toward an extragalactic source. At top is the bandpass of the 'off' position observation. At lower left is the superposition of the bandpasses of the 'on' and 'off' observation. Note how the two bandpasses are almost identical except between 320 and 380 km s^{-1}. The difference spectra, as defined by equation 10, is at lower right and clearly shows the hydrogen spectral line profile from the galaxy.

The per-channel noise of the difference spectrum is:

$$\sigma_{T_{ant}} = Constant \cdot \left(T_{sys} \Big/ \sqrt{BW / N} \right) \cdot \sqrt{\frac{1}{t^{reference}} + \frac{1}{t^{signal}}} \qquad (11)$$

where the constant depends upon the details of the spectrometer (e.g., channel spacing and number of sampling levels for an autocorrelation spectrometer). Note that the noise is higher in the difference spectra than in either the signal or reference spectra.

$T_{ant}(f)$ must be corrected for atmospheric attenuation and antenna efficiency, and, if one is observing a point source, converted to flux density.

Beam-Switched Observations:

Beam switched spectral-line observations are very similar to their continuum cousin and their spectral line position-switched brother. In beam switched observations, both the signal and reference spectra are accumulated simultaneously and synchronously. For some control systems, the difference spectrum (eq. 10) is calculated for the observer and he or she never gets to see the individual signal and reference spectra. With other telescopes, the user will have to use the built-in analysis command to create the difference spectrum.

Frequency-Switched Observations:

Frequency switched observations have some similarities to beam switched observations in that the signal and reference spectra are accumulated simultaneously and synchronously. Frequency-switched observations differ in that it is the center frequency of the observation that is toggled between two values. Depending upon the telescope's control system, the user might be presented with a difference spectrum or the individual signal and reference spectra.

Frequency-switched observations come in two flavors:

Out-of-band frequency switching:

In out-of-band frequency switching, the amount that the frequency is changed is larger than the bandpass of the observation. The spectral line of interest will lie only within the signal spectrum and not the reference. With this type of observing, the user or telescope control system need only apply equation 10 to form a difference spectrum.

In-band frequency switching:

If the frequency change is less than the bandwidth of the observation, it is possible that the spectral line of interest will lie in both the signal and reference spectra. In fact, the frequency change might be less than the width of the line, a specialized observing mode discussed by Liszt (1997). The difference spectrum produced by equation 10 will have a representation of the spectral line and an inverted image of the line separated by the amount of the frequency switch.

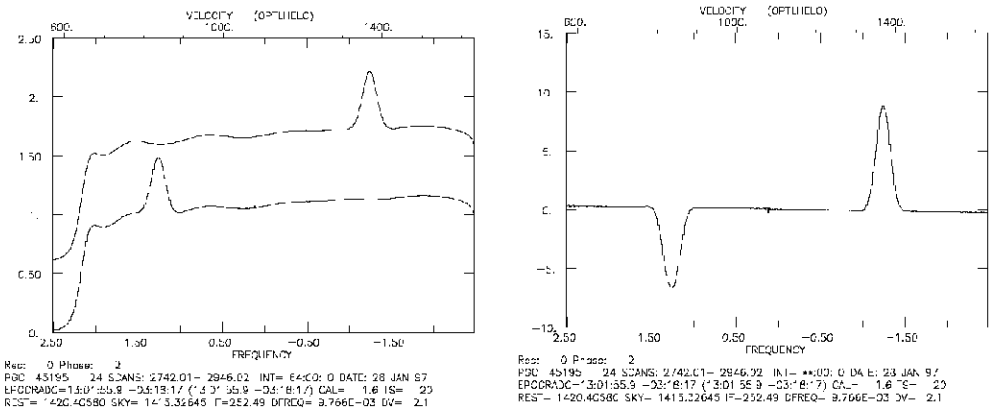

Figure 10: An example of in-band frequency switching. The left panel is the two bandpasses of the signal and reference spectra. The right panel is the difference spectra. Note the spectral line also appears as a negative image of itself but displaced by the amount of the frequency switch, the signature of in-band frequency switched observations.

One is tempted to take the difference spectrum, shift it by the amount of the frequency switch, invert the shifted version, and average the shifted/inverted spectra with the original. However, this simplistic algorithm would produce a systematic lowering of the detected line strength if the line strength is even a small fraction of T_{sys}. The correct algorithm involves creating a second difference spectrum but with the roles of signal and reference reversed. The second difference spectrum is then shifted by the amount of the frequency switch (Δf) and averaged with the original. Luckily, most analysis programs have a command that will 'fold' frequency switched observations.

$$T_{ant}(f) = \frac{T_{cal}(f)}{2} \cdot \left[\frac{\left(P^{signal}_{cal\_on}(f) + P^{signal}_{cal\_off}(f)\right) - \left(P^{reference}_{cal\_on}(f) + P^{reference}_{cal\_off}(f)\right)}{\left(P^{reference}_{Cal\_On}(f) - P^{reference}_{Cal\_Off}(f)\right)} \right] + \tag{12}$$

$$\frac{T_{cal}(f)}{2} \cdot \left[\frac{\left(P^{reference}_{cal\_on}(f+\Delta f) + P^{reference}_{cal\_off}(f+\Delta f)\right) - \left(P^{signal}_{cal\_on}(f+\Delta f) + P^{signal}_{cal\_off}(f+\Delta f)\right)}{\left(P^{signal}_{Cal\_On}(f+\Delta f) - P^{signal}_{Cal\_Off}(f+\Delta f)\right)} \right]$$

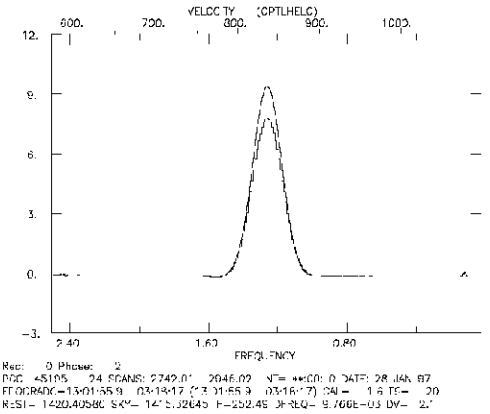

Figure 11: The comparison of the resulting spectra using the correct (upper curve) and incorrect (lower curve) algorithms for 'folding' an in-band frequency-switched observation.

Frequency and Velocity Calibration and Data Shifting:

Once a difference spectrum is formed, it might be necessary to calibrate the frequency and velocity axes of the resulting data. For example, some backends, such as with Acoustical Optical Spectrometers (AOS), the spacing of channels is non-uniform. Each observatory will recommend its own techniques and analysis algorithms for calibrating the frequency axis of their AOS data. Another issue neophyte observers should be aware of is that, since spectrometers sample in frequency, the velocity widths of channels changes across the bandpass.

A subtler problem afflicts all spectral-line observations at some level. Since the Earth rotates and revolves, and the Sun moves relative to the local standard of rest, observations made at two different epochs will need different Doppler shifts since the projected motion of the telescope relative to the direction of the observation will change with time. Usually, the telescope control software will alter the observed frequency so as to align one epoch's data with all other epochs in some chosen rest frame. This observing technique has been sufficient in the days when observing bandwidths were a small fraction of the observing frequency or when only a single spectral line is observed. Today, these hardware limitations are no longer true so the assumptions many analysis systems make are no longer adequate. For wide bandwidths, one can expect the velocity for all but the central channel of a spectrum to shift slowly with time. Additionally, Doppler tracking will only work for a single spectral line; any other simultaneously observed line will be miss-tracked especially if the second line has a rest frequency very different from the first.

All of these issues will require the occasional regridding and shifting of spectra in velocity or frequency. Some analysis programs have automated tools that will align the velocity of two spectra taken at different times. Typically these algorithms take the Fourier transform of the spectrum, apply a phase shift corresponding to the shift in velocity or frequency, and inverse Fourier transform the data. Unfortunately, these algorithms also have the undesired artifact of

slightly altering the noise in your data and require well- or over-sampled data (Brigham 1974; Press et al. 1992).

Averaging Spectra:

Our discussion in §3.1 on averaging data applies to spectral-line data as well. That is, you create an optimum average with the lowest noise only by proper weighting of the individual spectra.

The only caveat of concern for spectra is that one should velocity shift or regrid spectra before averaging so that all spectra will have the same velocity sampling. Failure to perform proper shifting will smear the line and make it have a wider width and lower intensity than would otherwise happen.

Baseline Fitting:

Fitting a polynomial baseline to a spectrum is identical to the fitting of a baseline to a time sequence of continuum data as discussed in §3.1. All of the statistical issues associated with continuum baseline fitting apply to spectral line fitting as well. Most single-dish analysis packages also offer an algorithm that will fit a sinusoid to emission-free regions of a spectrum. A sinusoidal baseline shape in spectral-line data will often occur when the telescope optics produce a standing wave with a frequency that is lower than a few times the spectrum's bandwidths.

Gaussian and Component Fitting:

The steps to fit a Gaussian to spectra are identical to those for continuum observations. It is up to the science at hand that will dictate whether or not a Gaussian is a reasonable model for the expected line profile.

In the case where a spectrum has one or more features, you might want to fit multiple Gaussians. If the features do not overlap, then you can fit each feature separately. However, if the features are blended, or if you have extra information about the features (e.g., the spectrum contains multiple hyperfine transition for which you know the frequency separation), then you will want to simultaneously fit multiple Gaussians. Many but not all packages allow for multiple Gaussian fits. Fewer systems allow you to use already-known information (like our example of known hyperfine separations, or transitions with known ratios of line strengths) to better constrain the Gaussian fit. Thus, if you plan on fitting multiple Gaussians, you should familiarize yourself with the software's documentation.

Spectral-Line moments:

In many cases you will be interested in the integrated intensity or the temperature-weighted velocity (centroid) of a spectral-line feature. All of the standard packages have some command that will perform these calculations of moments for you. The input to the algorithms usually involves defining the region of the spectrum over which spectral-line moments should be taken.

Either the analysis system or the user of a system should be able to generate the statistical uncertainty in the results of a moment calculation. For example, the uncertainty in the integrated intensity is $\sigma\sqrt{n}$ where σ is the per-channel noise and n is the number of channels across which the integrated intensity is taken.

Smoothing:

There are a few reasons why one would want to smooth spectral-line data. If you need to average spectra that were taken at two different velocity or frequency resolutions, first you will need to smooth the data so that all will have the same velocity or frequency resolution as that possessed by the spectrum with the coarsest resolution. Smoothing also affectively removes "ringing" inherent with cross-correlation/FFT-style correlation spectrometer when observing a narrow, strong spectral feature like an RFI spike. Brigham (1974) discusses how leakage, inherent in the

discrete Fourier transform these spectrometers use, manifests itself as ringing. Brigham also discusses how the correct solution for reducing ringing requires specifying a windowing function that the spectrometer should apply before it takes the Fourier transform.

Smoothing algorithms in single-dish analysis packages are of two types: those that decimate data (e.g., the smoothed data have less spectrometer channels than the unsmoothed), and those that don't (smooth and unsmoothed data have the same number of spectrometer channels).

Algorithms that don't decimate data, like the prolific Hanning or Boxcar smoothing functions, are overused and give a statistically false impression of the noise level of your data. Individual data points are no longer independent of each other. Thus, we suggest you limit your use of these functions for those cases you want a quick look at smoothed data.

Better and more statistically correct algorithms are those that decimate data. These algorithms usually convolve the data with some function, usually Gaussian of user-specified half width, before decimating the data. Even with decimating functions, the data points in the resulting spectra are still no longer independent. However, the results from these algorithms, in comparison to algorithms that don't decimate, will produce results that better match what the observations would have been like if it were made at the coarser resolution. Many general texts (Press et al. 1992; Bevington 1992) describe the details of smoothing functions.

RFI Excision:

At some point during your career of using radio telescopes you will encounter radio interference. In some cases the observations can be altered to mitigate the problems, in others you must deal with the consequences in the data analysis stage.

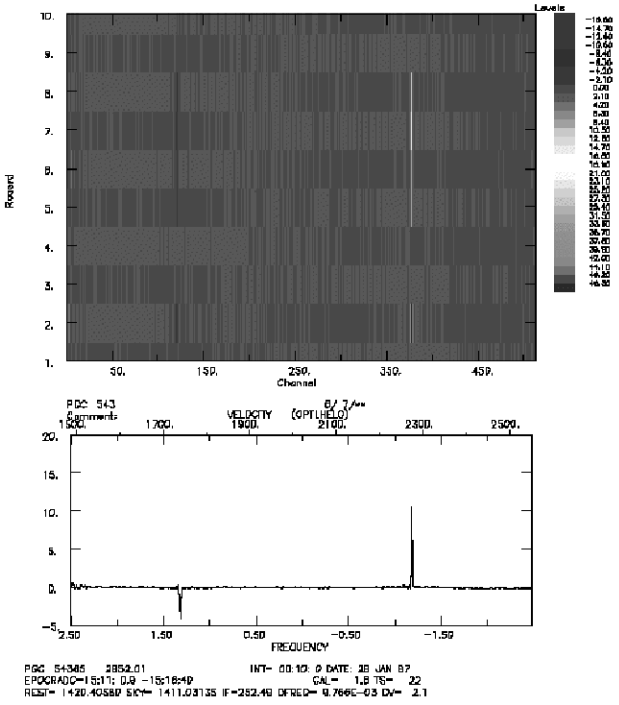

Figure 12: At top, a waterfall display showing ten sequential frequency-switched spectra in gray scale. Note the RFI in the four of the spectra numbered 5-8 near spectral channel 370 (and a weaker RFI frequency-switched into the bandpass near channel 120). The

spectrum at bottom is the weighted average of all ten spectra. By flagging the few channels or spectra with RFI, one could produce a cleaner average.

We have discussed under the previous section on smoothing how a single RFI spike in a spectrum can cause ringing across a bandpass. In other cases, if the RFI comes in bursts, it is sometime possible to excise that fraction of the data that are affected. In such a case, the observer should set up the backend for the fastest sampling times available in the hope that only a small fraction of the resulting spectra will experience RFI. Some data analysis programs have the ability to display the incoming spectra as a 2-dimensional waterfall image of intensity as a function of frequency and time. The user would use the tool to flag which spectra, and maybe which channels, are to be excised. The software will then average all unflagged data to produce a spectra of only those data samples deemed to be free of RFI.

Spectral-Line Observations of Extended Sources

We now have under our belts the necessary background to discuss spectral line mapping of extended sources. Like continuum maps, spectral-line maps can be made by grid mapping or on-the-fly mapping.

Spectral-Line Grid Mapping:

In grid mapping, the telescope steps the telescope from pixel to pixel in a rectangular region of the sky taking spectra only when the telescope has settled onto a pixel. While mapping, the observing technique could use beam or frequency switching. If the observations are position-switched, then the observing technique frequently should pause to take reference spectra.

The reduction of grid mapping requires converting the raw data into difference spectra as described in §4.1. The resulting $T_{ant}(f)$ are then regridded into an appropriate 3-dimensional image commonly called a data cube. The axes of the cube are the two orthogonal sky coordinates in which the map was made and frequency or velocity.

Spectral-Line On-the-Fly (OTF) Mapping:

OTF spectral-line mapping is no longer just the venue of cm-wave telescopes but has also become the preferred technique at mm-wave telescopes. In OTF mapping, the telescope slews through a rectangular region of the sky and simultaneously records data sometimes as fast as a few spectra per second. As in continuum OTF maps, the slewing is usually performed in one cardinal direction to produce a row or column in the map. After finishing a row or column, the telescope is offset in the direction orthogonal to the slew direction by somewhat less than half the beam width and a new row or column is started. While mapping, the observing technique could be beam or frequency switching. If the observations are position-switched, then the observing technique frequently will pause to take reference spectra.

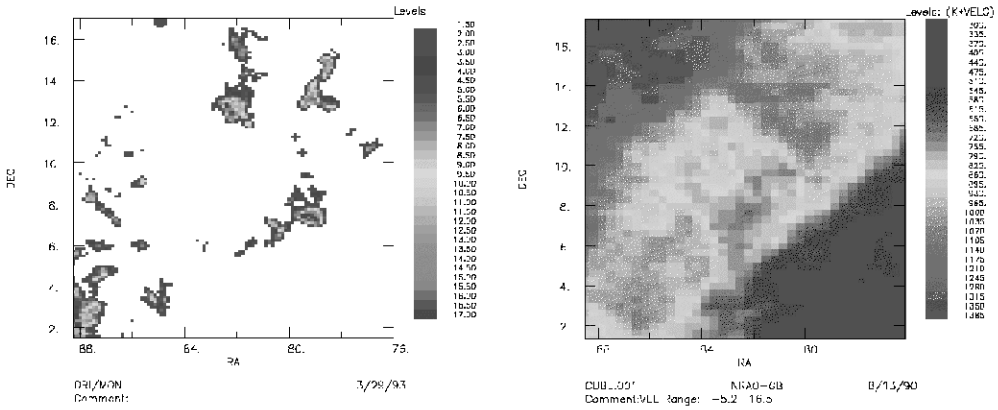

Figure 13: Examples of a grid-mapped (left) and on-the-fly (right) spectral-line observations. At left is the CO emission surrounding the λ Ori HII region as observed with the Columbia 1.2-m mm-wave telescope. At right, the HI emission from the same region as observed with the NRAO 140-ft telescope. Both images were created by slicing a 3-dimensional data cube.

The first step in the data analysis is to use the appropriate analysis described in §4.1 to create difference spectra. The data may then be smoothed, baselined, … before they are gridded into a data cube (see previous section and §3.2).

Rudimentary Analysis of Spectral-Line Maps:

The simpler single-dish packages allow one to produce a two-dimensional image (position-position or position-velocity maps) by taking a slice out of a data cube or integrating along a face of a data cube. Then, one uses the packages two-dimensional analysis tools to further analyze the map. In the single-dish packages, visualization tools usually are no more sophisticated then contour plots and the occasional grayscale or color-scale image. For more sophisticated visualization and analysis one usually needs to turn to systems like Aips, Aips++, and IRAF.

The Future of Single-Dish Analysis

Single-dish observing is becoming more and more sophisticated. Receivers are getting better, astronomers are asking for more sensitive observations or for wide-field maps. Backends are increasing in complexity and the number of channels they produce. Many telescopes are paving the way with multi-feed receivers. New spectrometers will have dozens of inputs, extremely wide bandwidths, and many hundreds of thousands of channels. As with other branches of astronomy, we will continue to see an escalating increase in the data rates coming out of single-dish radio telescopes.

We are also seeing users of single-dish telescopes whose previous observing experience has not been at radio wavelengths. Astronomers are asking for more robust and sophisticated analysis algorithms and better visualization tools. We are seeing a decrease in the amount of real-time, interactive data analysis as many observatories now allow for remote, queue, and service observing.

Single-dish observing hasn't stood still nor should single-dish analysis techniques. I predict that we will see in future single-dish analysis systems:

- An increase in the use of relational data base management systems to help control and organize the huge data sets we will be collecting.

- A unilateral archive of data that, after a reasonable time has passed since observations are completed is perusable by any astronomer, not just the observer who took the data.
- A rise in the sophistication of the visualization tools single-dish astronomers will use.
- More sophisticated, robust algorithms, especially those dealing with large-scale mapping.
- Further exploration of data pipelining for the general user.
- Automatic data calibration as sophisticated models of the telescope are incorporated and used by analysis programs.
- Development of algorithms that deal with data sets instead of individual spectra or continuum slices.

Individual observatories will no longer have the resources needed to provide the sophisticated packages users will request. Instead, within a decade only a few major systems maintained by cross-observatory groups will survive. Some observers and class of observers (e.g., pulsar or radar observers) might still have their specialized packages, but this will be the exception while just ten years ago it was the rule. Only time will tell whether these changes will benefit or hurt the science we love and cherish.

Acknowledgements

I thank Harvey Liszt for his very helpful comments on many points in this paper.

Appendix: Symbols and Definitions

BW	Bandwidth (Hz)
f	Frequency
i	Sample number
N	Number of spectral-line channels
P	Detected power in A/D counts or Volts
σ_T	Standard deviation of detected source temperature (K)
t	Integration time (s)
T_{ant}	Antenna temperature due to all sources: sky, atmosphere, receiver,... (K)
T_{cal}	Temperature of noise diode or calibration source (K)
T_{src}	Source temperature (K)
T_{sys}	System temperature due to all sources but the object of interest (K)

References

Bevington, P.R. & Robinson, D.K. 1992, Data Reduction and Error Analysis for the Physical Sciences (Boston: McGraw-Hill)

Briggs, D.S., Schwab, F.R., & Sramek, R.A. 1998, in Synthesis Imaging in Radio Astronomy II, ed. G.B. Taylor, C.L. Carilli, & R.A. Perley (San Francisco: ASP), 134-145

Brigham, E.B 1974, The Fast Fourier Transform, (Englewood Cliffs: Prentice Hall)

Emerson, D.T 1995, in Multi-Feed systems for Radio Telescopes, ed. D.T. Emerson & J.M. Payne (San Francisco: ASP), 309-317

Emerson, D.T., Klein, U., & Haslam, C.G.T 1979, A&A, 76, 92

Klein, U. & Mack, K.H. 1995, in Multi-Feed systems for Radio Telescopes, ed. D.T. Emerson & J.M. Payne (San Francisco: ASP), 318-326

Liszt, H. 1997, A&A Suppl., 124, 183.

Lockman, F.J., Jahoda, K., & McCammon, D.1986, ApJ, 302, 432

Mandel, J. 1964, The Statistical Analysis of Experimental Data (New York: Dover).

Press W.H., Teukolsky, S.A., Vetterling, W.T., & Flannery, B.P. 1992, Numerical Recipes in Fortran (New York: Cambridge)

Rood, R.T., Bania, T.M., & Wilson, T.L. 1984, ApJ, 280, 629.

Quang-Rieu, N.1969, A&A, 1, 128.

Van Zee, L., Maddalena, R.J., Haynes, M.P., Hogg, D.H., & Roberts, M.S.1997, AJ, 113, 1638.