Appendix F

Data Weights

F.1 Introduction

Visibility weight initialization and calibration has undergone several improvements in CASA 4.2.2 and (pending) CASA 4.3. This appendix briefly describes the formal weight definitions, and the changes occurring in these CASA versions. If data sets shall be combined that were reduced with different CASA versions, the weights may need to be adjusted accordingly. This can be achieved, e.g. by running the same version of `statwt` (§ 4.7.9) on all datasets before combination. The best option, however, is to use a single CASA version for all reductions, preferrably 4.2.2 or later.

Note that post-calibration weights, e.g. imaging weights or tapers are not covered by this appendix.

F.2 `SIGMA` and `WEIGHT` columns

Formally, in CASA 4.2.2 and later, the `SIGMA` column in the measurement set will reflect the per-channel noise of the `DATA` as it depends on the channel bandwidth $\Delta\nu$ and the length of an integration Δt:

$$\texttt{SIGMA} = \frac{1}{\sqrt{2\Delta\nu\Delta t}}. \tag{F.1}$$

The factor of $\sqrt{2}$ is for cross-correlation only and auto-correlation data follows $\texttt{SIGMA} = 1/\sqrt{\Delta\nu\Delta t}$.

`SIGMA` will only be updated if the time and channel widths are modified along with any `DATA` column manipulation, e.g. through averaging, binning, smoothing, etc. (tasks like `mstransform`, `cvel`, `split`, `exportuvfits`,...).

The `WEIGHT` column reflects how much weight each `CORRECTED_DATA` sample should receive when data are combined (e.g., in averaging). To start with, `WEIGHT` is initialized from the `SIGMA` column via:

	≤CASA 4.2.1	CASA 4.2.2	≥CASA 4.3 (WEIGHT_SPECTRUM)
Initialization	1	$2\Delta\nu\Delta t$	$2\Delta\nu\Delta t$
System Temperature	$\frac{1}{<\sqrt{T_{\rm sys,k}}>_k^2}$	$\frac{1}{<T_{\rm sys,k}>_k}$	$\frac{1}{T_{\rm sys,k}}$
Gains	$\|\|G\|\|^2$	$\|\|G\|\|^2$	$\|\|G\|\|^2$
Bandpass	$\frac{1}{<\|\|B\|\|^{-1}>_k^2}$	$< \|\|B\|\|^2 >_k$	$< \|\|B\|\|^2 >_k$

Table F.1: Antenna-based WEIGHT calibration factor definitions for different CASA versions. For System Temperator and Bandpass, k is the channel index. ALMA has channelized $T_{\rm sys}$; EVLA does not.

$$\texttt{WEIGHT} = \frac{1}{\texttt{SIGMA}^2} = 2\Delta\nu\Delta t \tag{F.2}$$

Data calibration by `applycal` (§ 4.6.1) with `calwt=T` will calculate and modify the `WEIGHT` values but not `SIGMA`. Calibration applies multiplicative factors and the `WEIGHT` of a visibility on a baseline between antennas i and j is calculated via

$$\texttt{WEIGHT}_{ij} = \frac{\omega_i\omega_j}{\texttt{SIGMA}^2_{ij}} \tag{F.3}$$

where ω_i and ω_j are the antenna-based calibration factors derived by `applycal` ($\omega_i = \omega_j$ for auto-correlation data). In Table F.1 we list the definitions of *antenna-based* ω for different calibration procedures and CASA versions. When more than one calibration is appied, the product of the relevant weight factors is used.

F.2.1 Weights in CASA 4.2.1 and Earlier

The `SIGMA` and `WEIGHT` columns are initialized with values of "1". Traditionally, this convention was adequate for datasets with uniform sampling in time in frequency; a global weight scale factor would not affect calibration and imaging fidelity. In data manipulation operations (e.g., `split`, etc.), SIGMA was treated as a per-channel value and WEIGHT as a per-spw (all channels) weight. Combined with unit initialization, this difference in definition could lead to incongruent weight scales for different spectral windows, in particular if bandwidth and channel count varied. CASA 4.2.1 is *not* recommended for datasets which have variety in spectral window bandwidth and channelization and for which spectral windows are to be combined in imaging.

F.2.2 Weights in CASA 4.2.2

In CASA 4.2.2 the `SIGMA` and `WEIGHT` columns are properly initialized via the definition in Eqs. F.1 and F.2. Both are defined as per-channel values. Also, the weight calibration factors have been subtly updated to improve robustness, as indicated in Table F.1.

F.2.3 Weights in CASA 4.3 and Later

In CASA 4.3 frequency variations of the `WEIGHT` and `SIGMA` values are (optionally) captured in additional `WEIGHT_SPECTRUM` and `SIGMA_SPECTRUM` columns. This allows accommodation of variations of effective sensitivity on a channel by channel basis (e.g. band edges, atmospheric lines, spectral $T_{\rm sys}$ variation etc.). Also, `statwt` will be enhanced to permit calculating weights on a channel-dependent basis.