ROSINA DFMS Time Series Abundance Files Data Description

The time series files in this archive were derived using calibrated L3 high resolution mass spectra measured by the ROSINA/DFMS instrument archived on the PDS SBN. The L2 to L3 process description can be found in "Rinaldi, M., DFMS PDS L2-to-L3 Data Processing Documentation, Version 1.5, 2018." The main detector of the DFMS is a position-sensitive imaging detector. The ionized neutrals of a selected mass impinge on the front of the Micro Channel Plate (MCP), releases secondary electrons that are then accelerated toward the back of the MCP. The secondary electrons knock out more electrons leading to an electron cascade. This electron cascade is detected by two independent rows of 512 pixels each called the Linear Electron Detection Arrays (LEDA). As described in the ROSINA User Guide in section 3.6, p. 14, row A or row B was normally more sensitive than the other row due to the electric field distribution in the ion source (contaminants). This caused the signal to be higher on one of the rows than on the other at different mission phases.

We performed peak fitting to each species on a given mass of the spectra collected by both rows A and B. The derived abundance values of rows A and B can either be added to enhance the signal, or they can be used separately.

The peak fitting routine to the ROSINA DFMS high resolution mass spectra to derive the time series files was performed via the following steps:

1) Read in file. For each Row,
2) Remove the first and last 25 pixels, to remove unreal spikes at the edges of the Row and create subset
3) Remove pixels with highest count rates (one-by-one) until mean <= median
4) Determine threshold: related to standard deviation of mean (of subset; where mean <= median)
5) Locate blocks of pixels with count rates that exceed threshold (each block is described as a 'peak' but may be composed of multiple species (as many as 7 species within a single peak), to be fit with multiple Gaussians (as many as 7 Gaussians at once))
6) For each block, use smoothing (averaging) and first, second derivatives to identify localized maxima and changes in slopes, identifying potential overlapping species in an automated fashion
7) If multiple potential overlapping species are identified, then sort the indices from smallest abundance to largest abundance. If a successful fit cannot be found and the smallest local maximum (or change in slope) is less than e.g., 0.1 of the next smallest, then it may not be a real species (i.e., noise), and may be removed
8) Determine how many Gaussians to fit to each species and to the overall peak (large local maxima are fit with 2 Gaussians with common center)

9) Multi-Gaussian fit attempts are performed, by providing ranges of initial guesses for each coefficient of each Gaussian, for multi-parameter fit attempts. Using multiple nested 'forloops', thousands of sets of initial guesses may be examined. Three distinct methodologies are used:

a) IDL LMFIT internal routine: a non-linear least squares fit to a function with an arbitrary number of parameters. LMFIT uses the Levenberg-Marquardt algorithm, and searches for a convergent solution. If no convergent solution is found,

b) IDL CURVFIT internal routine: uses a gradient-expansion algorithm to compute a non-linear least squares fit to a user-supplied function with an arbitrary number of parameters. The user-supplied function may be any non-linear function where the partial derivatives are known or can be approximated. If no convergent solution is found,

c) A brute force determination of chi-squared value using the initial guesses of coefficients is performed. This is most likely to be used for an obvious single peak with species that are significantly asymmetric and are unable to be properly fit with the internal routines

10) The solution with the smallest chi-squared value is stored, and the various parameters are printed to output arrays and files (i.e., Gaussian fit coefficients, integrated count rates for each species, overall peak integral (from Gaussians) and summed count rates)

Definitions:

Peak: A large scale increase over 'background'. May include one or more species, and may be fit with multiple Gaussians, depending in part on the number of local maxima

Species: A specific element or molecule. May be fit by one Gaussian, or in hi-res with a peak count rate greater than a prescribed threshold, by two Gaussians with a common center

Gaussian: A single Gaussian function: fn = an(0)*exp(-(pix–an(1))^2/(2*an(2))

Reference Datasets:

The datasets used in the generation of these data are in order of mission phase and as follows:

Altwegg, K., ROSETTA-ORBITER 67P ROSINA 3 PRL V2.0, RO-C-ROSINA-3-PRL-V2.0, ESA Planetary Science Archive, NASA Planetary Data System, 2019. https://doi.org/10.5270/esa-hr4u8ut

Altwegg, K., ROSETTA-ORBITER 67P ROSINA 3 ESC1 V2.0, RO-C-ROSINA-3-ESC1-V2.0, ESA Planetary Science Archive, NASA Planetary Data System, 2019. https://doi.org/10.5270/esa-pui3jpq

Altwegg, K., ROSETTA-ORBITER 67P ROSINA 3 ESC2 V2.0, RO-C-ROSINA-3-ESC2-V2.0, ESA Planetary Science Archive, NASA Planetary Data System, 2019. https://doi.org/10.5270/esa-6o6j4ir

Altwegg, K., ROSETTA-ORBITER 67P ROSINA 3 ESC3 V2.0, RO-C-ROSINA-3-ESC3-V2.0, ESA Planetary Science Archive, NASA Planetary Data System, 2019. https://doi.org/10.5270/esa-jz1r3o6

Altwegg, K., ROSETTA-ORBITER 67P ROSINA 3 ESC4 V2.0, RO-C-ROSINA-3-ESC4-V2.0, ESA Planetary Science Archive, NASA Planetary Data System, 2019. https://doi.org/10.5270/esa-fbb2kn4

Altwegg, K., ROSETTA-ORBITER 67P ROSINA 3 EXT1 V2.0, RO-C-ROSINA-3-EXT1-V2.0, ESA Planetary Science Archive, NASA Planetary Data System, 2019. https://doi.org/10.5270/esa-w27m438

Altwegg, K., ROSETTA-ORBITER 67P ROSINA 3 EXT2 V2.0, RO-C-ROSINA-3-EXT2-V2.0, ESA Planetary Science Archive, NASA Planetary Data System, 2019. https://doi.org/10.5270/esa-krtpsgp

Altwegg, K., ROSETTA-ORBITER 67P ROSINA 3 EXT3 V2.0, RO-C-ROSINA-3-EXT3-V2.0, ESA Planetary Science Archive, NASA Planetary Data System, 2019. https://doi.org/10.5270/esa-2fqg97y