

Split2Events

From PangaWiki

Split2Events is a software tool to split one file with data from several events into several files, one for each event. The resulting folder with a number of files can automatically be imported with the *Massenimport* routine of 4D. Split2Events may also extract a list of unknown parameters prior to splitting.

The recent version and reference of Split2Events can be found at <http://www.pangaea.de/software/UsefulTools/>

Split2Events is licensed under the GNU General Public License (GPL) (<http://www.gnu.org/licenses/gpl-3.0.txt>) and is freely distributed. (Source-code (http://www.pangaea.de/Software/Source/Split2Events_src.zip) of Split2Events as zip-archive, 265 kB)

Contents

- 1 General
- 2 Just split
- 3 Split with data description
- 4 Split using a metafile
- 5 Split using a metafile and find parameter IDs automatically
- 6 Reference and detailed description

General

- Empty columns and lines which contain GEOCODE only will be removed.
- If data set description is included, 4D will analyze it and format during sequential import.
- When defining data set titles and filename, the placeholder *\$E* may be used to add individual event labels.
- Keywords are not added by Split2Events; they may be set manually in 4D prior to import.

Just split

1. Start Split2Events and drag/drop the file to the program window.
2. Choose the *Split to events...* Tool (F5).
3. The *Mandatory* tab shows the entries of the last session.
4. Click on *New* if a new file collection is processed.
5. Go to the *Options* tab and check *split file to events*.
6. Click *OK*. An import file is written for each event, stored in a new folder.
7. Use 4D (Import/Analytical data/Open folder) for sequential import.

Split with data description

1. Start Split2Events and drag/drop the file to the program window.
2. Choose the *Split to events...* Tool (F5).
3. The *Mandatory* tab shows the entries of the last session.
4. Click on *New* if a new file collection is processed.

5. On the *Mandatory* tab fill out the fields as required.
6. On the *Optional* tab add information as required.
7. On the *Options* tab check *split file to events* and *write data description*.

Split using a metafile

1. Open an import file and choose the option *use metadata file; find parameter by position* on the *Options* tab.
2. Create a template of a metadata file by clicking the button *Create metadata template*. The metadata file will be created with the extension *_metadata.txt*.
3. Open this file with an editor (e.g. drag the file onto an open Excel window) and modify it appropriately. Each line contains the information for one parameter.
 - *Parameter name* as given in the data file. If the ID is given in the data file this entry is empty.
 - *Parameter ID*, in this mode the ID is mostly “unknown”. Only the ID of the GEOCODE is given automatically. You have to fill in the right ID for the parameter.
 - *PI ID* as provided through the PI field of Split2Events. The ID 999999 will be replaced by *@PP@Event label@* (e.g. *@PP@PS2742-5@*)
 - *Method ID*; use ID=43 if *not given*. If left empty, the default method will be used.
 - *Comment* of data series.
 - *Format* as suggested from the precision of the numeric values of the related parameter.
 - *Factor* if a recalculation is required, e.g. may be used to convert units.
 - *Fill empty cells with* add characters, which should be used to fill empty cells. The data entry *@is empty@* will be always replaced by an empty string.
 - *Range min* and *Range max* defines the range of values of parameter (e.g. for temperature in water set this -5 to 50). If a value is outside of the given range, the value will be marked with the flag *not valid* = /.

Split using a metafile and find parameter IDs automatically

1. Prepare a local list of all parameters called *ParameterDB.pdb*. For this step use *File->Create parameter database* or *Tools->Merge new parameters to parameter database*. The name of the parameter database can be given first with *File->Settings....*
2. Browse to the *ParameterDB.pdb* on the *Options* tab.
3. Create a metadata file with the option *use metadata file; find parameter by name*. Split2Events identifies the ID for each parameter by using the parameter database. If a parameter is unknown, the ID is set to *unknown*. If *write parameter import file* on the *Options* tab was checked, a list of the unknown parameters is written to *imp_Parameters.txt*. After completing this file send it to hgrobe@pangaea.de (mailto:hgrobe@pangaea.de) .
4. Continue with step 1.

Reference and detailed description

-- Mandatory tab --

- *Staff ID of principal investigator*. This entry sets the ID for the PI of data in the data description part. The ID 999999 will be replaced by *@GP@Event label@* (e.g. *@GP@PS2742-5@*). With the PanTool function *Search and replace many strings* the PI of data of this event can be set easily.
- *Staff ID of author(s)*. This entry sets a list of IDs for the authors of the datasets. The ID 999999 will be replaced by *@A@Event label@* (e.g. *@A@PS2742-5@*).
- *Institution ID of source*. This entry sets the ID for the source (related to institution) of data. The ID 999999 will be replaced by *@S@Event label@* (e.g. *@S@PS2742-5@*).
- *Reference ID(s)*. This entry sets a list of IDs for the references of the dataset in the data description

part. The ID 999999 will be replaced by @R@Event label@ (e.g. @R@PS2742-5@).

- *Project ID(s)*. This entry sets a list of IDs for the projects of the dataset in the data description part. The ID 999999 will be replaced by @Pro@Event label@ (e.g. @Pro@PS2742-5@).
 - *Dataset title*. This entry sets the citation of the dataset in the data description part. The placeholder \$E will be replaced by the event label. The placeholder \$@ will be replaced by the string behind the “@” given in the event label. The ID 999999 will be replaced by @D@Event label@ (e.g. @D@PS2742-5@).
 - *Export filename*. This entry sets the export filename. The placeholder \$E will be replaced by the event label. The placeholder \$@ will be replaced by the string behind the “@” given in the event label.
-

-- Optional tab --

- *Dataset comment* (optional). This entry sets the dataset comment. The placeholder \$E will be replaced by the event label. The placeholder \$@ will be replaced by the string behind the '@' given in the event label. The ID 999999 will be replaced by @C@Event label@ (e.g. @C@PS2742-5@).
 - *URL Further details* (optional). This entry sets the URL of the further details link. The URL has to begin with “doi:”, “hdl:” or “http://”. The placeholder \$E will be replaced by the event label. The placeholder \$@ will be replaced by the string behind the “@” given in the event label. 999999 will be replaced by @F@Event label@ (e.g. @F@PS2742-5@).
 - *URL other version* (optional). This entry sets the URL of the further details link. The URL has to begin with “doi:”, “hdl:” or “http://”. The placeholder \$E will be replaced by the event label. The placeholder \$@ will be replaced by the string behind the “@” given in the event label. 999999 will be replaced by @O@Event label@ (e.g. @O@PS2742-5@).
 - *Topologic type*. Menu to select the topologic type of a dataset.
 - *Status*. Menu to select the status of a dataset.
 - *Login*. Sets the login of a dataset to *on* or *off*.
 - *User(s) of validated datasets*. This entry sets the ID for the *user* of data. The ID 999999 will be replaced by @U@Event label@ (e.g. @U@PS2742-5@).
-

-- Options tab --

- **Split to events options**
 - *split file to events*: Split2Events splits a file to events only if this option is set. Uncheck it for creating surface datasets.
 - *write data description*: The data description part of the import file contains all import options. This is needed for importing a huge number of files. But it can also be useful to configure complex import files outside 4D.
 - *use name of input file for \$E or ID=999999*.
 - *make filename unique*: Needed if you split more than one files with the same events.
 - *marks files with 1, 2 or 3 lines*: Useful to separate data from cores and surfaces are mixed.
 - *override existing datasets*: @I@Event label@ will be written.
- **Metadata file options**
 - *don't use metadata file*: A metadata file will not be used.
 - *use auto metadata file*: Sets the default method and a suitable format for each parameter automatically.
 - *use metadata file; find parameter by position*: If the program splits a file to events, a metadata file is loading. The position of parameters in the metadata file has to be the same order as in the data file. A metadata template can be created with *Create metadata template*.
 - *use metadata file; find parameter by name*: If the program splits a file to events, a metadata file is loading. The program finds the right parameter automatically. A metadata template can be created with *Create metadata template*.
- **Create metadata template options**

- *write parameter import file*: If the option *use metadata file; find parameter by name* is selected and the user creates a metadata template, Split2Events finds missing parameters and writes them into *imp_Parameter.txt*. After completing this file send it to hgrobe@pangaea.de (<mailto:hgrobe@pangaea.de>) .
 - **Parameter database**. To create a metadata template and find missing parameters the program needs a list of all known parameters defined in PANGAEA. Use *Tools->Create parameter database* or *Tools->Merge new parameters to parameter database* to create a parameter database. The name of the parameter database can be given first with *File->Settings....* Browse for it.
 - **Buttons**
 - *Create metadata template*. Pressing this button creates a metadata template file. This depends on the *metadata file options* and the *create metadata template options*. If *use metadata file; find parameter by name* is selected a parameter database is needed.
 - *New*. Resets all settings.
 - *Save*. Saves a project file manually.
 - *Load*. Loads a project file. Overwrites all given settings!
 - *OK*. Starts splitting procedure.
 - *Cancel*. Closes the settings dialog. No settings will be changed.
-

Contact: Dr. Rainer Sieger, Alfred Wegener Institute, Bremerhaven

Retrieved from "<http://wiki.pangaea.de/wiki/Split2Events>"

- This page was last modified on 27 June 2011, at 19:37.