RFC: nagg -l option to specify characteristics of output files

Larry Knox

# Purpose

This document describes gathering characteristics for nagg output files from an example NPP data product file using the example file option of the nagg tool.

# Introduction

Each JPSS file contains groups and attributes that provide enough information to determine several characteristics of the aggregation in the file:

Products to be included

Number of granules of each product in each full aggregation

Packaging

Geolocation yes/no

The proposed –l option will direct nagg to create its output files with the same products and aggregation number as those in an NPP data product example file. The products to be included will be determined from the subgroups of /Data\_Products in the example file, which are named with the same names as the products. These can be converted to DPIDs for the –t list.

The number of granules of each product is stored in an attribute of each <product name>\_Aggr dataset named “AggregateNumberGranules”. The products found will be overridden by –t in the nagg command as will the aggregation number by –n or –A in the nagg command.

The geolocation information can be present as a subgroup of /Data\_Products or in another file specified by the /N\_GEO\_Ref attribute. If a geolocation subgroup is present in addition to another product subgroup, or if the /N\_GEO\_Ref attribute exists geolocation should be “yes”. If a geolocation subgroup is the only product subgroup present, geolocation should be the corresponding DPID. Otherwise geolocation should be “no”.

Files with multiple subgroups of /Data\_Products are packaged. The only unpackaged arrangement that can be detected from an example file is a single data product with separate geolocation.

The origin can be found in the “/N\_Dataset\_Source” attribute. The domain can be found in the “N\_Processing\_Domain” attribute of each product subgroup of /Data\_Products. The values of these attributes are copied from the input files to the output files. There is one “N\_Dataset\_Source” attribute per file and one “N\_Processing\_Domain” attribute in each product group. If multiple files are aggregated from different origins only one of the original “N\_Dataset\_Source” attribute values will be preserved. Likewise if granules from different processing domains are aggregated in a product group, only one of the original “N\_Processing\_Domain” values for the product will be preserved. The origin and domain in the example file for the –l option will be ignored. The origin and domain fields in the filename are taken from the –O and –D command options with defaults of XXXX and XXX.

# Requirement Specification

## Option Syntax

-l <file>, --like=<file>

If -l <filepath> is found in the nagg command, open <filepath> and determine the first 4 characteristics listed above for the ouput files.

If any of the command line options –A, -n, -g or -t are also found, override any of the 4 characteristics according to the command line option.

Return an appropriate error message if the file cannot be opened or does not contain a required characteristic.

## Table of example and input file elements, conditions and expected nagg

|  |  |  |
| --- | --- | --- |
| **Element** | **Example and input file conditions for successful outcome** | **nagg command option equivalent for successful outcome or error** |
| Product groups | All products in example file found in input files | -t <match list>  aggregate matching products |
|  | Example file has one or more products not found in input files | ERROR |
| Aggregation number | 0 < aggr\_num < nagg granule\_limit in example file | -n aggr\_num  create aggregations of size aggr\_num |
|  | O >= aggr\_num or aggr\_num > granule\_limit in example file | ERROR |
| Geolocation | GEO group in example file and GEO available for input files | Aggregate GEO (Default)  Package GEO  (Default) |
|  | GEO group in example file and GEO not available for input files | ERROR |
|  | N\_GEO\_Ref attribute in example file and GEO available for input files | Aggregate GEO (Default) –S  Aggregate product and geo in separate files |
|  | N\_GEO\_Ref attribute in example file and GEO not available for input files | ERROR |
|  | Neither GEO group nor N\_GEO\_Ref attribute in example file | -g no  Aggregate matching products but not GEO |
|  | GEO group only in example file and GEO available for input files | -g GEO  Aggregate GEO product alone |
|  | GEO group only in example file and GEO not available for input files | ERROR |

# Update for Reference Manual

These are the current and proposed reference manual entries for the –l option. The idea of using the first file encountered as the example file has been eliminated.

## Current entry

-l *file*

*(To be supported in future implementation.)*

Package like the example *file* in number or type list.  Options on the command line override the example.  If both -l and -t are omitted, then the first NPP data product file encountered will be used as the example file.

## Proposed entry

-l *file*

Package like the example *file* in number or type list.  Options on the command line override the example.

# User Examples document additions

These examples are intended to demonstrate the common behavior of the –l <example\_file> option and the behavior with possible command options that override a part of the common behavior.

## Example file option with no overriding command options:

nagg –l <example\_file> <input\_files>

## Example file option with –A and –t command option overrides:

nagg –l <example\_file> –A 300,–t <type list> <input\_files>

1. Example file option with –n and –g command option overrides:

nagg –l <example\_file> -n4, -g no <input\_files>

# Test Specification

The nagg tool creates output files that match the example file in terms of products included, aggregation number, packaged or separated, presence or absence of geolocation, origin and domain attributes.

Tests 1 – 7 are for scenarios with the nagg –l option in this command:

nagg –l <example\_file> <input\_files>

Tests 8 – 9 are for the –l option with other command options that override the effects of the example file.

1. Example file has:

Products also found in input files

Any reasonable number of granules

Packaged geolocation product

Output:

Packaged files produced with products and aggregation number matching example file

1. Example file has:

One or more products not found in input files

Output:

nagg: \*\*\*ERROR\*\*\* nagg\_get\_granules(): The number of product types for which

granules were found was less than the number of products requested.

1. Example file has:

A GEO product group in the example file, but input file are without GEO.

Output:

nagg: \*\*\*ERROR\*\*\* nagg\_get\_granules(): no granules found for geoproduct.

1. Example file has:

1 data product and N\_GEO\_Ref attribute in example file

Matching and available GEO product for input files

Output:

Unpackaged data product and GEO files

1. Example file has:

1 data product and N\_GEO\_Ref attribute in example file

No GEO product available for input files

Output:

nagg: \*\*\*ERROR\*\*\* nagg\_get\_granules(): no granules found for geoproduct.

1. Example file has:

Neither GEO product group nor N\_GEO\_Ref attribute in the file

Output:

Files with aggregated data product(s) as in the example file.

1. Example file has:

GEO product group with no data product.

Output:

Files with aggregated GEO product only.

1. Example file characteristics are overridden by –A 300 and –t <type list>

(Either option could be used alone)

Command: nagg –l <example\_file> –A 300,–t <type list> <input\_files>

Output:

Files with aggregation size according to –A 300, products according to <type list> and other characteristics as determined by the example file.

1. Example file characteristics are overriddenby –n4 and –g no

(Either option could be used alone)

Command: nagg –l <example\_file> -n4, -g no <input\_files>

Output:

Files with aggregation size 4, no GEO granules, and other characteristics as determined by the example file.