

## **Review of Climate and Forecast Metadata Conventions Implementation and Operational Suitability**

NASA's Earth Science Data Systems Standards Process Group (SPG) is considering the Climate and Forecast (CF) Metadata Conventions, for adoption as a community standard. You are invited to review this Requests For Comment (RFC) in the context of your **implementation experience** with this specification and its **suitability for operational use**. You only need to answer questions that are applicable to you. Please send completed review to:

[spg-rfc-021@lists.nasa.gov](mailto:spg-rfc-021@lists.nasa.gov).

*Implementation Experience questions:*

1. *(Your background)* Describe in a sentence or two your overall implementation experience related to the proposed specification. (*e.g., specification implementer, tools developer, data provider, scientific analyst, science user, etc.*) Have you directly implemented the CF metadata conventions? Did you use pre-existing software, and if so, what did you use?

I work at an organization which archives and distributes oceanographic data. The data could either be produced internally within the organization or by external entities. I have converted datasets from other formats, mostly text, to CF compliant netcdf format. I used netcdf C and perl libraries, provided by Unidata for this task.

2. *(Completeness)* Does the specification (the online documents referenced) provide all the detail you need to implement it in software? (*e.g., to read or write a data file; to implement or modify a profile or extension; or develop a tool such as a metadata translator*) If not, describe what is missing in the specification.  
It depends on what it means to be complete. I found a few things missing from the specification. One area which needs to be more complete is Coordinate Systems (using `grid_mapping`) attribute. The convention does not seem to provide datum transformation parameters (did I miss something??). If CF only allows geocentric datums, it should make it clear. Also, it is not always possible to easily transform from one datum to another using limited number of parameters.  
Another area is the attributes for documentation. "comment" attribute is not the best attribute to describe the variables. There should be another attribute consisting of free-form text to describe the variables in detail. This is especially useful for variables which do not have a standard `_name`.  
Also, variables representing data quality information should be explicitly identified.  
The convention is vague when it comes to attributes for climatological data. Limited number of attributes try to represent too much information making is difficult for the implementer.  
ISO- 19115, which is more complete, should be used as a reference. There should be a way to map each ISO- 19115 metadata object with netcdf attributes.  
Also ACDD attributes for data discovery should be considered for adoption into CF.

3. (*Accuracy*) Do any parts of the specification contain inaccuracies, or internal inconsistencies? If so, please provide details.
4. (*Clarity*) Is any part of the specification ambiguous, or poorly explained? If so, please provide details.

I found the attributes for describing climatological data are ambiguous. Again, just giving datum parameters does not tell users how to transform from one datum to another.

It should be made clear what set of attributes are suitable as global attributes and which are more appropriate as variable attributes.
5. (*Balance*) Does the standard describe the right set of concepts and attributes and enable the appropriate operations for its intended users? In particular, have the guiding principles outlined in section 5.2 been followed in the development of standard names?

Standard names are useful but giving some structure to the standard names using well defined ontologies goes a long way in identifying data. This is especially true for variables which do not have a standard name but related to other concepts described in ontology.
6. (*Usefulness*) How well does this specification meet your information sharing needs? (e.g., *Does it properly represent and describe your datasets? What are the pros and cons of these metadata convention attributes?*)

Having data which is machine readable (to a high semantic level) and which GIS software, among other software, can read is important for us. CF allows that. But more needs to be done to make the convention even better.
7. (*Implementation*) What implementation challenges does the proposed standard present? (e.g., *does it conflict with other metadata requirements for your data? Is it compatible with the data formats you use?*)

There aren't any major implementation challenges when using netcdf as the dataformat. But there are some minor duplication of attributes when we use ACDD along with CF.
8. (*Flexibility*) In what software environment(s) have you used the CF metadata conventions (e.g., Solaris, Linux, Windows, Mac OS X)?

We implemented the CF conventions on Linux systems. And served it using Thredds running on Tomcat.
9. (*Standard Names*) In your opinion, does the standard name table provide an adequately comprehensive set of names for the metadata representation?

Standard name table is not comprehensive and I am not sure it will ever be. Of course standard names are good. But linking the standard names with well known ontologies is a good way for users to convey what standard names their data is closely related to in cases when there is no existing standard name

*Operational Suitability questions:*

10. Do you currently use or plan to use CF conventions in a production setting? What types of applications do you use with CF Conventions? Does the metadata model work well with the data types and data manipulations in your application?  
We are using CF in production settings. We are not end users.
11. Why do you choose to use the CF metadata conventions for your applications?  
Repeating earlier point, it is important that user applications are able to read and understand the data to the highest semantic level possible. CF allows reading the axes, coordinate systems, temporal attributes, etc. CF is more comprehensive than other conventions. Also, CF allows different feature types.
12. Have you or your users encountered any difficulty when using some of the data access or visualization tools (e.g., IDL, GrADS, etc.) on files with CF metadata? If you have, please provide a brief description of your experience.  
We used tools like ncBrowse and GrADS with CF metadata. We haven't encountered any major problems. But then aren't heavy users of the data.
13. Does the CF metadata conventions meet your requirements for discovering, accessing, providing interoperability of data and metadata? (e.g., *Can it handle the data types in your applications? Do you provide catalog services that utilize CF conventions?*)  
CF can handle various feature types which is important for us. CF has attributes which allow interoperability. But it is not complete and it doesn't have enough number of attributes for data discovery.
14. What operational challenges or limitations do the CF metadata conventions present? (e.g., *Does it take a long time to learn how to use it? Does it require advanced processing power, large amounts of memory, complex configuration, etc.*)  
Hardly any operational challenges. But there are lack of GUI based tools which can work with CF. User support group is great.
15. What benefits do CF conventions present? Do the benefits of CF conventions outweigh the challenges? (e.g., *Do the conventions offer the flexibility you want to package the data types in your applications? Do they facilitate interdisciplinary studies?*)  
Using a standard/convention, even if it is not that great, usually outweighs the challenges. CF is reasonably good. It is flexible enough to accommodate various feature types. They do facilitate interdisciplinary studies.
16. How much data do/ will you provide using these CF metadata specifications? (*number of distinct data products or data sets, total data volume, number of files.*)  
We provide some data with CF specifications but are on the way to make CF more important in all our future data archiving/distribution activities. No concrete numbers yet.

17. How many users and user-groups do you have or expect to have for data using CF metadata conventions, and what is your expected user community?

Our expected community of users is anyone interested in oceanographic data including oceanographers, marine biologists, coastal managers etc.

18. (*User comments*) Any additional comments, observations or criticisms of CF metadata conventions and the RFC can be provided here.

With climate becoming important for government administrators and policy makers, having interoperable datasets becomes important when integrating data from various fields. GIS is the language most governments understand. CF committee should give special attention to make CF compliant data more easily readable/understandable with popular GIS software.