Gfdnavi, web-based data and knowledge server software for geophysical fluid sciences

Part I: Rationales, stand-alone features, and supporting knowledge documentation linked to data

Takeshi Horinouchi (Hokkaido U), Seiya Nishizawa (Kobe U), C. Watanabe, A. Tomobayashi, S. Otsuka, T. Koshiro, Y.-Y. Hayashi, and GFD Dennou Club

black: Geophysical scientists | blue: Database scientist
What is Gfdnavi

• = **Geophysical fluid data navigator**

• A suite of software to construct Web-based database of geophysical fluid data

• **Functionality:**
  – Search data
  – Numerical analysis and visualization
  – Documentation of analysis results

• **Available:**
  http://www.gfd-dennou.org/library/davis/gfdnavi/
Introduction
Data we use

- **Observational data** (satellite, station etc etc) / Simulation data (climate prediction etc etc) / other numerical data (assimilation data, idealized data etc)
- Mostly in a few self-descriptive binary formats such as NetCDF, GRIB, HDF-EOS (but not always)
Many organizations/research groups provide data through the web

- They provide data files
- Optionally visualization etc: in many cases custom-made (for each project / organization)
Related work

Live Access Sever (LAS)

• Software to build web server for georeferencing geophysical data
• Data browsing, search and visualization
• Web-page structure: Highly configurable
Problem of current web-based data servers

- Limited visualization / analysis capability
  - Only quick-looks. Need to DL data
  - Service are not available to local data
- Support of non-georeferencing data is weak
Visualization is not the goal

- To others (scientists / society): reports (papers etc)
- While working: memos
- To collaborators: reports / internal documents / discussion

Outputs are documents (not just pieces of images)
Introducing Gfdnavi
Basic requirement

• Support both browser GUI & programability for users

→ Beyond initial quick-look
  • GUI: good to start up / good for novices – interdisciplinary collaboration
  • Programming: infinite degrees of freedom / good to repeat

• Support a wide range of use cases from public data services, group use, to desktop data management
  – Should be easy to install, start up, and manage

• Support documenting & archiving knowledge obtained through data analysis
Two fundamental libraries used to build Gfdnavi (*open-source*)

- **GPhys** – a Ruby library to analyze and visualize geophysical fluid data (*by Horinouchi etc since 2003*)
  - For consolidated access to data in files (*NetCDF, GRIB, GrADS, NuSDAS, HDF5-EOS etc*) or on runtime memory
  - A community infrastructure for data analysis – Key to unite all forms of data access
  - Used by increasing number of scientists

- **Ruby on Rails** – Web application development framework
  - Written in/for Ruby → We can use GPhys directly
  - Equips its products with web server → Easy to deploy
Support programmability in multiple ways

• Web services – Part II (next talk)
• Browser access
  – Download Ruby code and data subset to reproduce visualization
  – Upload Ruby script (for qualified users)
Structure of Gfdnavi

- **Local file system (or OPeNDAP dir)**
- **RDB (metadata etc)**
- **Browser Access**
- **Gfdnavi MVC core**
- **Web server (webrick/Apache)**
- **User**

- **Access numerical data**
- **Meta data scan**
- **Sync with DB**
- **O/R mapping**
- **Web service**

- **Browser Access**
Metadata DB

- **Attributes** (extracted from data files or supplied by additional text files)
- **Directory structure**

```
• Attributes (extracted from data files or supplied by additional text files)
• Directory structure
```

![Diagram of directory tree and data files]

- **directories**
- **variables**
- **attributes**
  - description = “……..”
  - param1 = value1
  - param2 = [val21,val22]
- **spatial_and_temporal_attributes**
- **keywords**
User Interface

**Home**: Independent simple html → replaceable
Functionality

Typical work flow to use Gfdnavi’s browser UI
Functionality

Browse directory tree (Finder)
Search (Explorer)
View docs (Knowledge)
Visualize / analyze (Analysis)
Write knowledge document

Select variables in this file to analyze / visualize

MS Explorer-like tree
Directory contents
Further details (metadata)

Select numerical data
repeat
Functionality

- Browse directory tree (Finder)
- Search (Explorer)
- View docs (Knowledge)
- Visualize / analyze (Analysis)
- Write knowledge document
- Select numerical data
- Select free text
- Select attributes
- Search with Google Maps
- Select a variable to analyze / visualize
Functionality

Browse directory tree (Finder)

Search (Explorer)

View docs (Knowledge)

Visualize / analyze (Analysis)

Write knowledge document

Select numerical data

Save in the DB (login needed)

Get the URL to redraw the img

Select

Repeat

Draw method: You can supply your own Ruby Script & Minimum Subset Data

Download script and data

save diagram

link to this diagram
Functionality

Browse directory tree (Finder)

Search (Explorer)

Select numerical data

Visualize / analyze (Analysis)

View docs (Knowledge)

repeat

Write knowledge document

analyze!
Functionality

Browse directory tree (Finder)
Search (Explorer)
View docs (Knowledge)

Visualize / analyze (Analysis)

Select numerical data
Select
Repeat

Write knowledge document
Functionality

Browse directory tree (Finder)

Search (Explorer)

Select numerical data

Select

Visualize / analyze (Analysis)

View docs (Knowledge)

repeat

Write knowledge document
Functionality

Browse directory tree (Finder)

Search (Explorer)

Select numerical data

Visualize / analyze (Analysis)

View docs (Knowledge)

Write knowledge document

Create Function

Function

name
save directory

FFT[^2] along a specified dimension

description

number of input variables

number of arguments

number of output variables

Function Arguments
Functionality

Gfdnavi

Top Finder Explorer Analysis Knowledge

Browse directory tree (Finder)
Search (Explorer)

Select numerical data
Select

Visualize / analyze (Analysis)
View docs (Knowledge)

Write knowledge document

GFDNAVI

Variables

Axes

Dimensions

longitude 0
latitude 90
levelist 1

Options

Draw Analysis

Temperature

CONTOUR INTERVAL = 5.000E+00

Create a Knowledge Document
with this image(s)
Gfdnavi

Top Finder Explorer Analysis Knowledge

Browse directory tree (Finder)

Search (Explorer)

Select numerical data

Visualize / analyze (Analysis)

View docs (Knowledge)

Select

Write knowledge document

Create a Knowledge Document with these image(s)
Save Images and Create a New Knowledge

Gfdnavi

Top Finder Explorer Analysis Knowledge

Browse directory tree (Finder)
Search (Explorer)
Visualize / analyze (Analysis)
View docs (Knowledge)

Select numerical data
Select
Repeat

Write knowledge document

Figure 1
Caption:
ERA Jan T at 1 hPa

Figure 2
Caption:
ERA Jan T at 100 hPa

Texas

((Figure 1)) shows the climatological temperature at 1 hPa using the ECMWF Reanalysis (ERA40) in the northern hemisphere. It shows that the climatological polar vortex is shifted to the Pacific side.

((Figure 2)) shows the same as Fig. 1 but for 100 hPa. It shows that the westerly jet is strong in the Pacific storm track.
ECMWF Reanalysis January Climatology

Figure 1 shows the climatological temperature at 1 hPa using the ECMWF Reanalysis (ERA40) in the northern hemisphere. It shows that the climatological polar vortex is shifted to the Pacific side.

Figure 2 shows the same as Fig.1 but for 100 hPa. It shows that the westerly jet is strong in the Pacific storm track.
GFDNAVI

Top Finder Explorer Analysis Knowledge User Logout

Browse directory tree (Finder)
Search (Explorer)
View docs (Knowledge)
Visualize / analyze (Analysis)
Write knowledge document

Select numerical data
Select
repeat
draw!
A typhoon occurred at east of the Philippines.

A typhoon occurred on July 2005.
The figure means the amount of rainfall per hour at July 16, 2005. We can see a typhoon east of the Philippines.

Forecast of Course of typhoon

According to the forecast of the Japan Meteorological Agency, typhoon 5 will change direction of movement near Taiwan. The sea around Okinawa is warmed by the intense heat of days, so it seems that the typhoon will move further. There is the possibility of coming off from the expectation and landing on West Japan. Please note the future.

Rainfall rate (mm/hr)

Fig 1. Rainfall

Path: /user/root/knowledge/typhoon8.knmlg
Edit | Back to List

There are no comments on this document.
There are no comment on this document.

Add a Comment on this document.

Title: [Enter] Typhoon Information  
Author: [Enter]  
Textbody:

Choose a default layout: figures under text.  
size of figure: height [100]  
input the number of figures in a row [1]  

Figure 1

Caption:

Figure Path: [image]
A typhoon occurred at east of Philippines

A typhoon occurred on July 2005. The figure means amount of rainfall per hour at July 16, 2005. We can see a typhoon east of Philippines.

Forecast of Course of Typhoon

According to the forecast of Japan Meteorological Agency, typhoon 5 will change direction of movement near Taiwan. The sea around Okinawa is warmed by the intense heat of days, so it seems that the typhoon will further. There is possibility of coming off from the expectation and landing on West Japan. Please keep an eye on the future.

Fig. 1. Rainfall

Path: /usr/root/knowledge/typhoon3.knowledge
Edit | Back to List

1 comment exists.

Show Summary of Comments | Hide Comments | Show full text of Comments

Ref [1]: Typhoon Information  author: Akinori Tomobayashi  by root  last update: Sun Sep 20 After all, typhoon 5 went for Taiwan and landed China.

Write a Comment on this document  Back to List
Functionality

Browse directory tree (Finder)
Search (Explorer)
View docs (Knowledge)
Write knowledge document
Visualize / analyze (Analysis)

Select numerical data
Select
repeat

Fig 1. Rainfall

Path: /usr/root/knowledge/typhoon3.knls
Edit | Back to List

1 comment exists.

Show Summary of Comments Hide Comments Show full text of Comments

Rel[1]: Typhoon Information  author: Akinori Tomobayashi  by root  last update: Sun Sep

After all, typhoon 5 went for Taiwan and landed China.

Fig 1. course of typhoon 5 in 2005.

Path: /usr/root/knowledge/typhoon3_comment1.khls

Write a Comment on this document
Back to List
By linking a document with data and visualization/analysis methods

• One can confirm and extend it
  – falsifiability

• Annotate data with the document
  – scientific metadata

• Applications
  – interactive publication / science collaboration / educational material (incl. interdisciplinary collaboration)
Application Examples

- **JEM/SMILES** data server (JAXA):
  - Satellite obs for ozone etc: Science team (incl. restricted access) + General data service

- International collaboration project to improve **weather forecast in Asia**

International Research for Prevention and Mitigation of Meteorological Disasters in Southeast Asia
NEXT Special Coordination Funds for Promoting Science and Technology for FY 2007 - 2009
Asia S&T Strategic Cooperation Program

[Home]
[Motivations]
[Major Research]
6. Decision support tools for ensemble numerical weather prediction: I. Basic diagrams

6.1 1D line plot

Data
/Nargis/NHM/POM/h.nc (lon, lat, t, member)

Settings
- Axes
  - h member(t)
    - lon = 95.07 degE
    - lat = 16.10 degN
    - (X) t = [0 h, ..., 71 h]
    - (Ens) member = 0, ..., 20
- General Settings
  - Draw method
    - ensemble_1D
- Specific settings
  - style: lines

This diagram is called "Plume diagram".

Result
Time series of surface elevation at Irrawaddy point (95.07 degE, 16.10 degN) for 21 members Fig 1. Some members show storm surge of more than 3 m in height.
Summary

• We have developed Gfdnavi: software to build data and knowledge servers
  – Wide coverage from desktop use to public data service (by having custom web server)
  – Programmability (on browser & by web service)
  – Documentation of analysis results (dynamically reproducible/extendible) (memos / reports / PR / Blog for scientific collaboration)

Next talk by Seiya Nishizawa:
✓ More on programmability
✓ Web services
✓ Network of Gfdnavi
fin
GPhys (Gridded Physical quantity)

VArray (Virtual Array) – Abstracts Data Storage
(Can be in file(s) or multi-D Array on memory; can also be a subset or aggregation of (an)other VArray(s))
Example of GPhys’s associated coordinates

```
(GPhys) "temperature" (4D VArray)
    (Grid)
    (Axis) "x" (1D VArray)
    (Axis) "y" (1D VArray)
    (Axis) "z" (1D VArray)
    (Axis) "t" (1D VArray)

(Array) 0..*
```

- "coordinate variables", but can be simple indices
- coordinate names must be unique to support subsetting by names

Supports transformed grids, scattered data points, etc etc
Supports “coordinates” in CF convention
What is Ruby on Rails
http://www.rubyonrails.org/

- Web development framework in Ruby
- With RDBMS (Mysql, Postgres, SQL Server, SQLite etc)
- Strong prototyping (e.g. Model-View-Controller (MVC) structure)
- Comprehensive library (covering Ajax and Web service)
- Ruby-embedded html
  - suitable to use our Ruby library
- Has a private web server (Webrick); also runs on Apache, lighttpd etc
  - One can personally run a web server anywhere with arbitrary port
From “Understanding Rails MVC”:
http://wiki.rubyonrails.org/rails/pages/UnderstandingRailsMVC
Sister-server method

(a) Basic case: available in LAS. User can’t choose peers
(b) Gfdnavi: one can register any peer by running a Gfdnavi
P2P with directory server

- Direct Use
- Indirect Use
- Query Server
- List
Currently tested by C. Watanabe by using Overlay Weaver (Java-based p2p library) and Rails’ Action Web Service
– Decentralized p2p with distributed hash tables (DHT)