

Five reasons why you should know ...



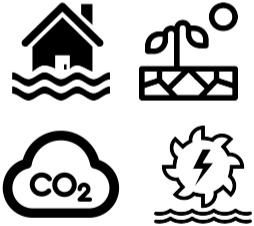
Environment and  
Climate Change Canada  
Environnement et  
Changement climatique Canada



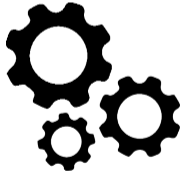
**Juliane Mai, Bryan A Tolson, Kurt C Kornelsen,**  
David Schäfer, Vincent Fortin, Nicolas Gasset, Djamel  
Bouhemhem, Michael Leahy, Paulin Coulibaly, François Anctil, Brent Hall

# Motivation

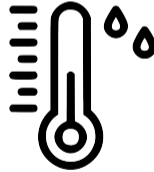
Environmental  
Models



Differences:  
Processes



Similarity:  
Input Data



Forecast Mode:  
Numerical Weather  
Predictions



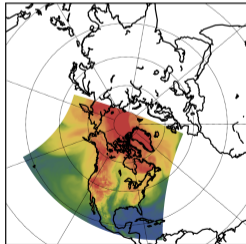
# Aim of CaSPAr

**archive**  
forecasts & analyses  
produced by



Environment and  
Climate Change Canada  
Environnement et  
Changement climatique Canada

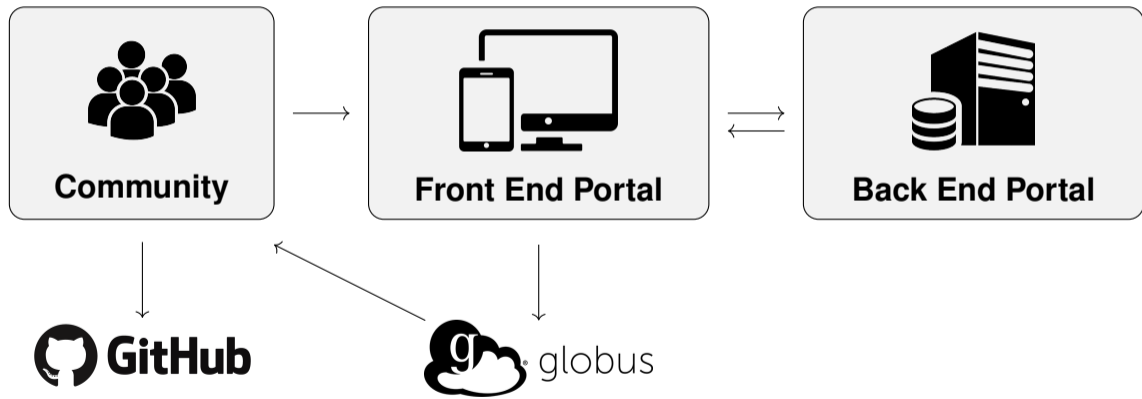
**convert**  
raw FST data to  
standardized NetCDF



**provide**  
gis-based web interface  
to make data available



# Architecture

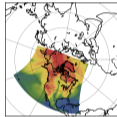


# Frontend

– Interface –



User Login  
and  
Data Management



**ArcGIS Toolbox:**  
Visualization and GUI  
Interactive Request  
Submission



RESTful API



Live  
Documentation

# Frontend

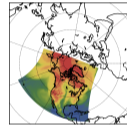
– Control –



**ArcGIS Toolbox:**  
Validate User  
Requests



**PostgreSQL:**  
Maintain Queue  
Request Status  
**submit\_job.py:**  
Send to Backend



**PostgreSQL:**  
Maintain Product  
Records

# Backend

– Conversion and archiving –



02:00 am UTC

Rsynch cron-job  
running on Pegasus  
pushes data to Graham



every hour

Cron-job on Graham  
checking for  
new rsynch report



triggered

**converter.py:**  
jobs converting FST files to NetCDF

**merger.py:**  
jobs reducing files by merging timesteps

**checker.py:**  
job checking results and e-mails report

# Backend

– User requests –

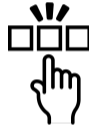


user request

JSON string from  
frontend



create mask  
for chosen  
spatial domain



select forecast  
horizons and  
variables



move to outbox  
JSON string to  
frontend



# Frontend – Transfers –



update Globus ACL &  
email user



Globus handles  
user interaction

A screenshot of the Globus web interface. The top navigation bar includes 'Manage Data', 'Publish', 'Groups', 'Support', and 'Account'. The main content area is titled 'Transfer Files' and shows two panels for 'Endpoint: Globus Tutorial Endpoint 1' and 'Endpoint: Globus Tutorial Endpoint 2'. Each panel has a 'Path' field and a 'Go' button. Below the panels are 'Transfer Settings' and a 'Label This Transfer' field. A red arrow points to the 'select all' button in the second endpoint's file list.

user GUI or  
API transfer

**#1**

# Individual Subsetting of Data

# Frontend

– Submit Your Request –

- Select product
- Credentials
- Domain
- Variables
- Horizons
- Issues
- Time period
- Submitted!

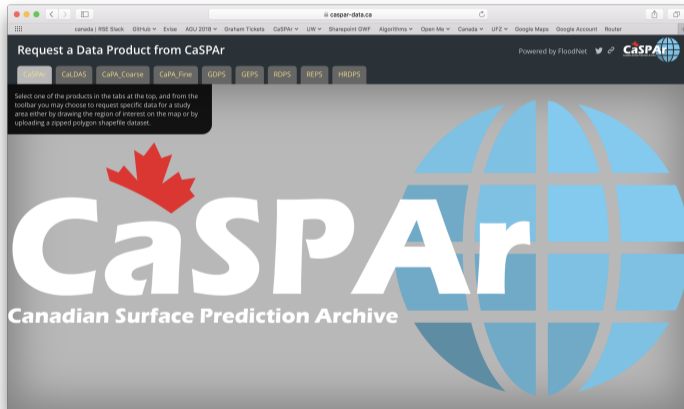


# Frontend

– Submit Your Request –

✓ **Select product**

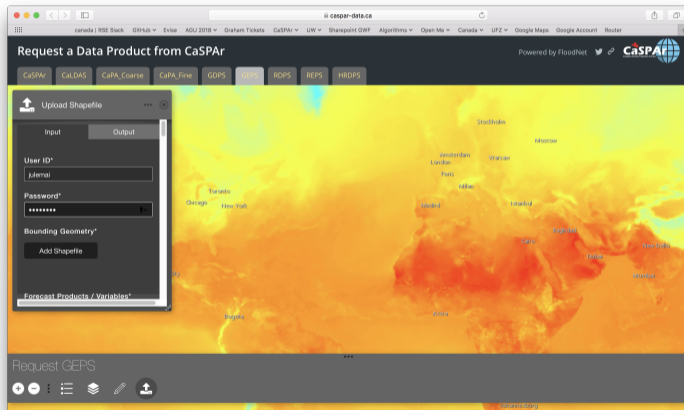
- Credentials
- Domain
- Variables
- Horizons
- Issues
- Time period
- Submitted!



# Frontend

– Submit Your Request –

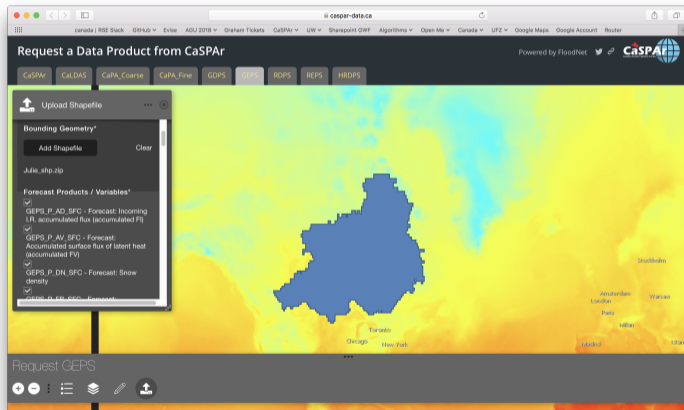
- ✓ Select product
- ✓ Credentials
- Domain
- Variables
- Horizons
- Issues
- Time period
- Submitted!



# Frontend

– Submit Your Request –

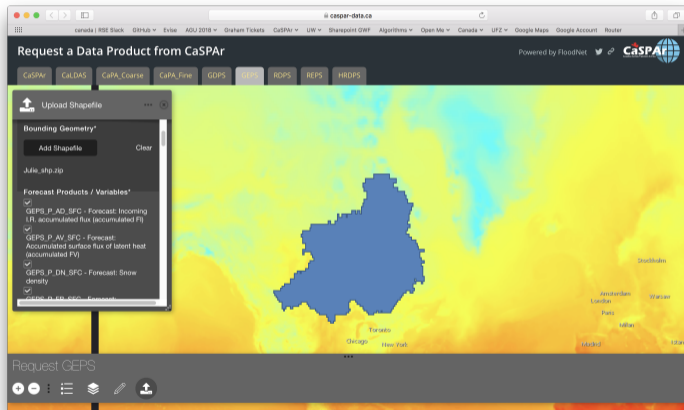
- ✓ Select product
- ✓ Credentials
- ✓ Domain
- Variables
- Horizons
- Issues
- Time period
- Submitted!



# Frontend

– Submit Your Request –

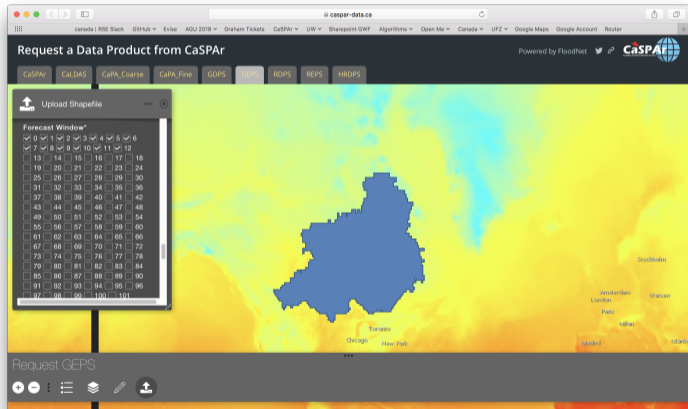
- ✓ Select product
- ✓ Credentials
- ✓ Domain
- ✓ Variables
- Horizons
- Issues
- Time period
- Submitted!



# Frontend

– Submit Your Request –

- ✓ Select product
- ✓ Credentials
- ✓ Domain
- ✓ Variables
- ✓ Horizons
- Issues
- Time period
- Submitted!

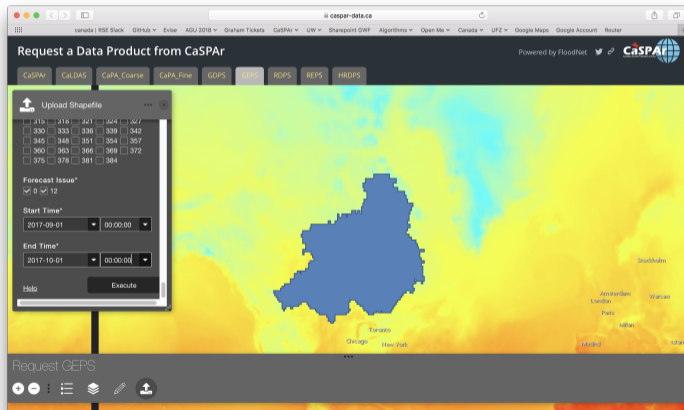




# Frontend

– Submit Your Request –

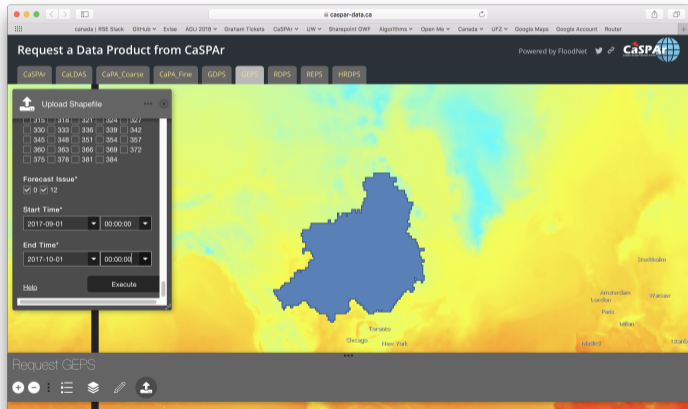
- ✓ Select product
- ✓ Credentials
- ✓ Domain
- ✓ Variables
- ✓ Horizons
- ✓ Issues
- Time period
- Submitted!



# Frontend

– Submit Your Request –

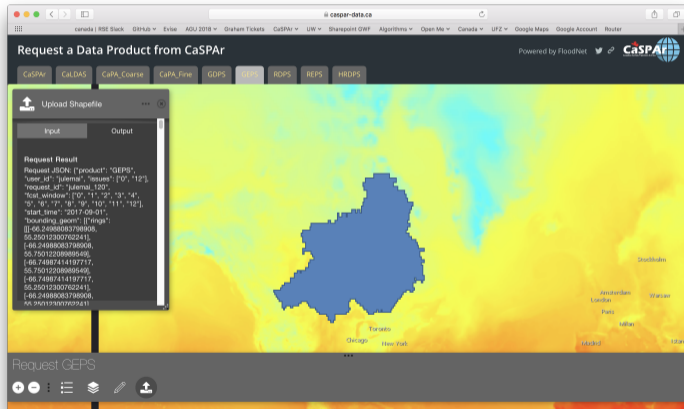
- ✓ Select product
- ✓ Credentials
- ✓ Domain
- ✓ Variables
- ✓ Horizons
- ✓ Issues
- ✓ Time period
- Submitted!



# Frontend

– Submit Your Request –

- ✓ Select product
- ✓ Credentials
- ✓ Domain
- ✓ Variables
- ✓ Horizons
- ✓ Issues
- ✓ Time period
- ✓ Submitted!



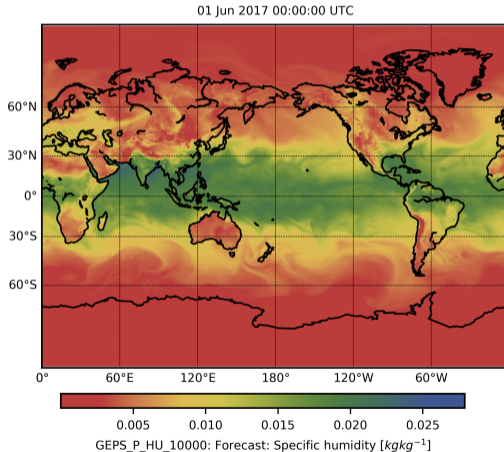
**#2**

## Easy Comparison of NWP Products

# Products Available

– Status Quo –

## 1. GEPS

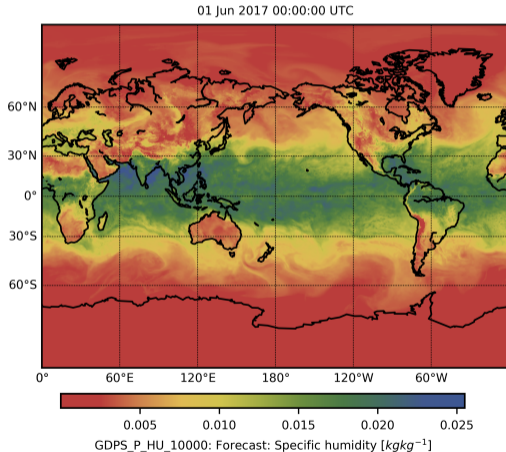


Forecasted time steps [h]:	1,2,..., 72, 96,..., 384
Issues:	2/day
Ensemble size:	21
Number of variables:	55
Resolution:	50km
Memory [GB/month]:	4662

# Products Available

– Status Quo –

1. GEPS
2. GDPS

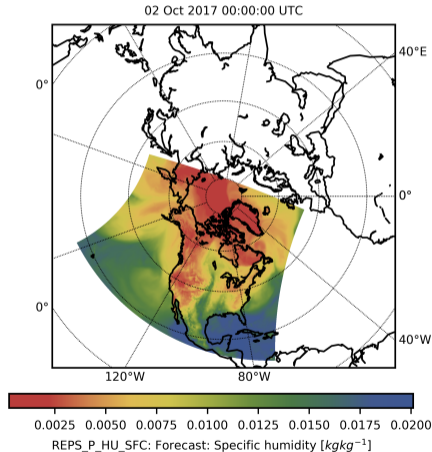


Forecasted time steps [h]:	1,2,..., 144, 147,..., 240
Issues:	2/day
Ensemble size:	1
Number of variables:	55
Resolution:	25km
Memory [GB/month]:	426

# Products Available

– Status Quo –

1. GEPS
2. GDPS
3. REPS

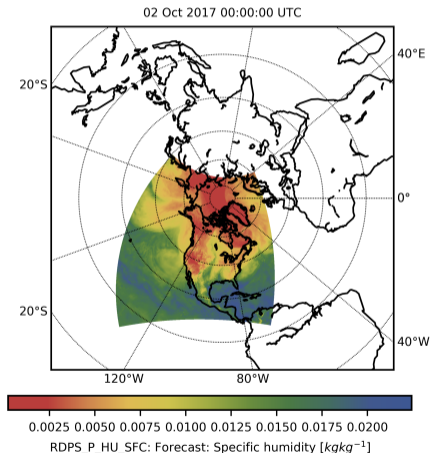


Forecasted time steps [h]:	1,2,..., 72
Issues:	2/day
Ensemble size:	21
Number of variables:	52
Resolution:	15km
Memory [GB/month]:	1638

# Products Available

– Status Quo –

1. GEPS
2. GDPS
3. REPS
4. RDPS



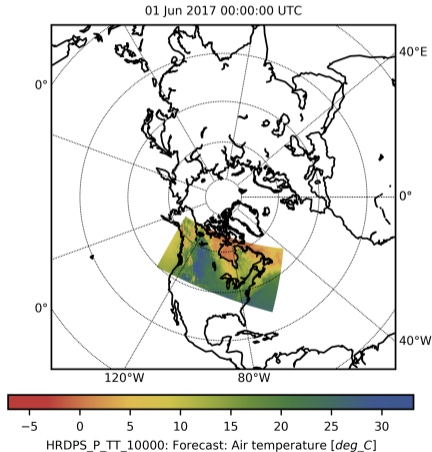
Forecasted time steps [h]:	1,2,..., 84
Issues:	4/day
Ensemble size:	1
Number of variables:	55
Resolution:	10km
Memory [GB/month]:	684



# Products Available

– Status Quo –

1. GEPS
2. GDPS
3. REPS
4. RDPS
5. HRDPS

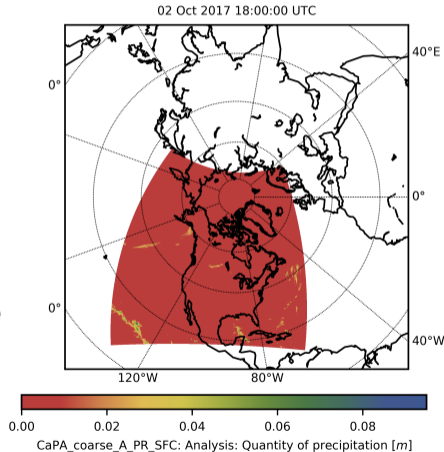


Forecasted time steps [h]:	1,2,..., 48
Issues:	4/day
Ensemble size:	1
Number of variables:	50
Resolution:	2.5km
Memory [GB/month]:	840

# Products Available

– Status Quo –

1. GEPS
2. GDPS
3. REPS
4. RDPS
5. HRDPS
6. CaPA\_coarse

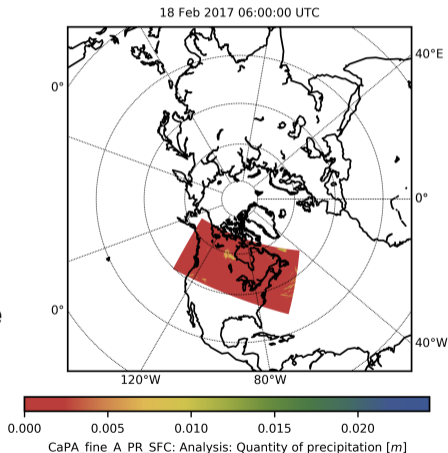


Forecasted time steps [h]:	N/A
Issues:	4/day
Ensemble size:	1
Number of variables:	2
Resolution:	10.0km
Memory [GB/month]:	1

# Products Available

– Status Quo –

1. GEPS
2. GDPS
3. REPS
4. RDPS
5. HRDPS
6. CaPA\_coarse
7. CaPA\_fine

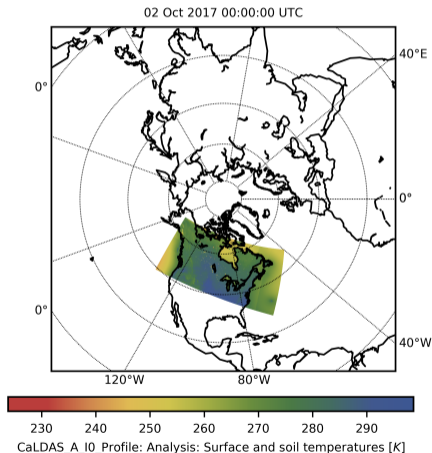


Forecasted time steps [h]:	N/A
Issues:	4/day
Ensemble size:	1
Number of variables:	2
Resolution:	2.5km
Memory [GB/month]:	2

# Products Available

– Status Quo –

1. GEPS
2. GDPS
3. REPS
4. RDPS
5. HRDPS
6. CaPA\_coarse
7. CaPA\_fine
8. CaLDAS



Forecasted time steps [h]:	N/A
Issues:	8/day
Ensemble size:	25
Number of variables:	14
Resolution:	2.5km
Memory [GB/month]:	281

# Products Available

– Status Quo –

11+ TB  
new data  
per month

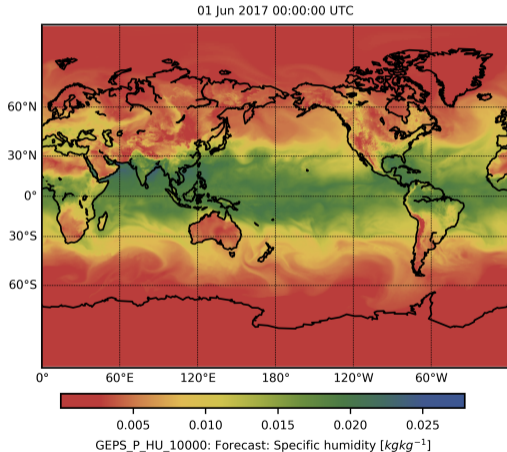
(archived since summer 2017)

**#3**

Drastic Reduction of Data Amount to Download

# User Requests

– GEPS –

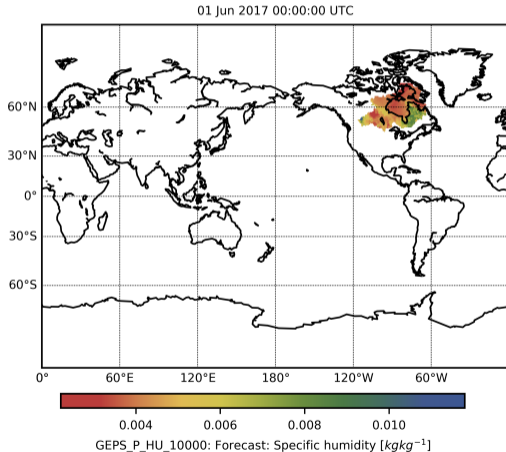


## Request:

- Hudson Bay
- 1 month forecasts
- 13 of 85 forecast horizons
- 4 of 55 variables
- 21 ensemble members

# User Requests

– GEPS –



## Result:

- 330 MB of 4662 GB (0.007%)
- 22.5 min processing time
- 30 (days) x 2 (issues/day) x 21 (ens. members) NetCDF files

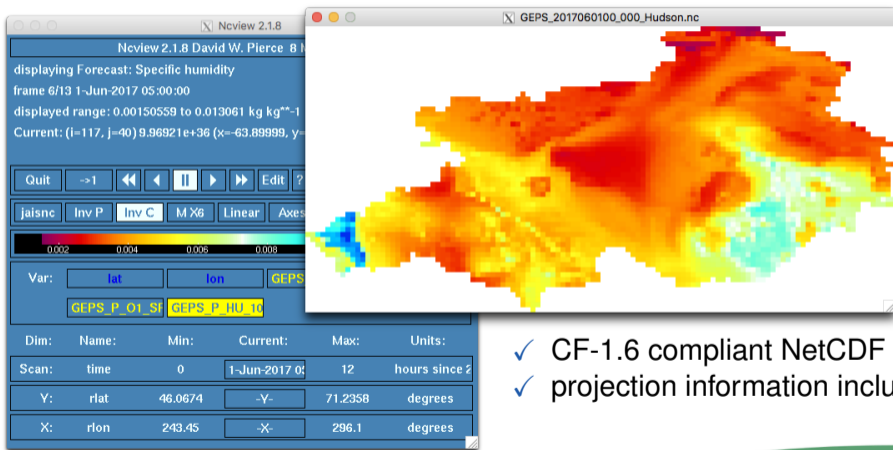


**#4**

## Standardized File Format

# User Requests

– GEPS –



- ✓ CF-1.6 compliant NetCDF files
- ✓ projection information included

**#5**

## Seamless Utilization of Input Data

# Models running with NetCDF inputs directly

- Hydrologic Modeling Framework RAVEN
- Hydrologic Modeling Framework SUMMA
- Variable Infiltration Capacity Model VIC v5.0
- Mesoscale Hydrologic Model – mHM
- MEC-Surface & Hydrology model – MESH
- Weather Research and Forecasting Model – WRF-Hydro
- Community Land Model – CLM
- Multi-parametrization Land Surface Model – Noah-MP

# Summary of CaSPAr highlights

1. **Individual subsetting** of data via web-interface and (soon) API
2. **Easy comparison** of NWP products
3. **Drastic reduction** of data amount to download
4. **Standardized** NetCDF file format (CF-1.6 compliant)
5. **Seamless utilization** of input data for model runs
6. **Fast** data download via Globus
7. **GitHub** library of data processing scripts
8. Modular and decoupled **Frontend and Backend**

