Climate Change

EUMETSAT

Copernicus
Europe's eyes on Earth

Validating Atmospheric Motion Vectors with a geospatial database

O. Sus, M. Grant, L. Medici, J. Onderwaater

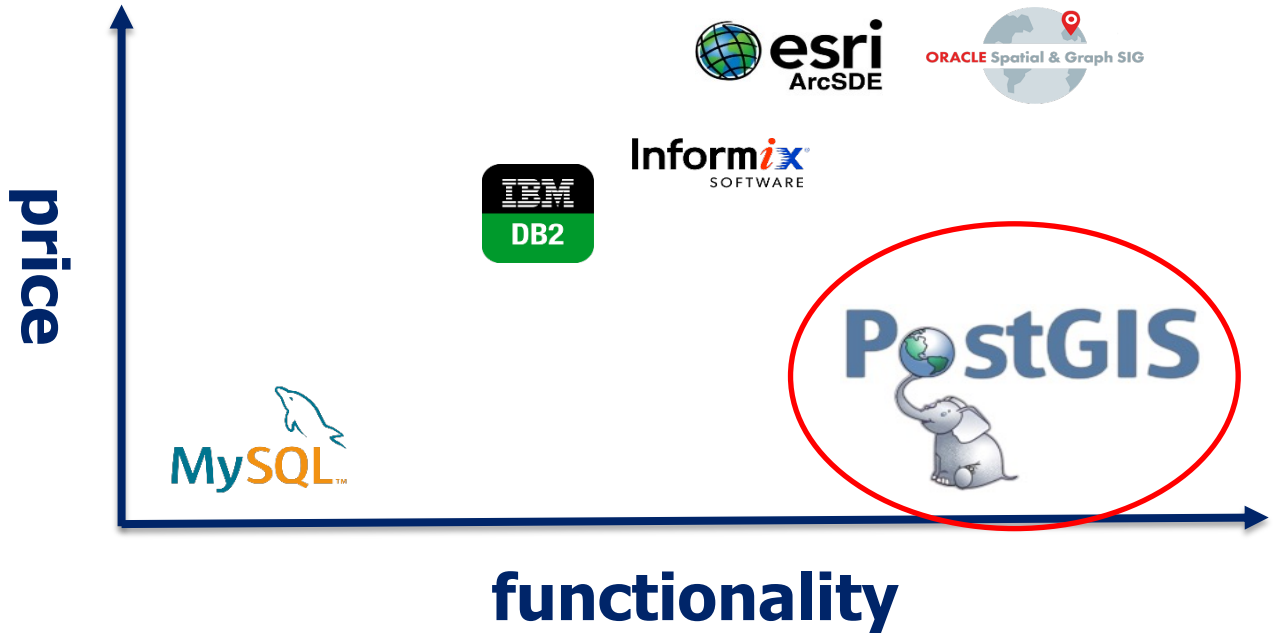
Why a database? Why PostgreSQL?

- PostgreSQL: a relational, open-source database management system (RDBMS)
- Simple, easy to use
- Provides good set of tools for
  - searching and combining data from various sources
  - rearranging data (filtering, grouping, statistical aggregation)
- Automated query optimization + partitioning
- Replication
- Safe simultaneous data access and modification by multiple users
- Standardized SQL methods and software interfaces (e.g. Python GeoPandas)

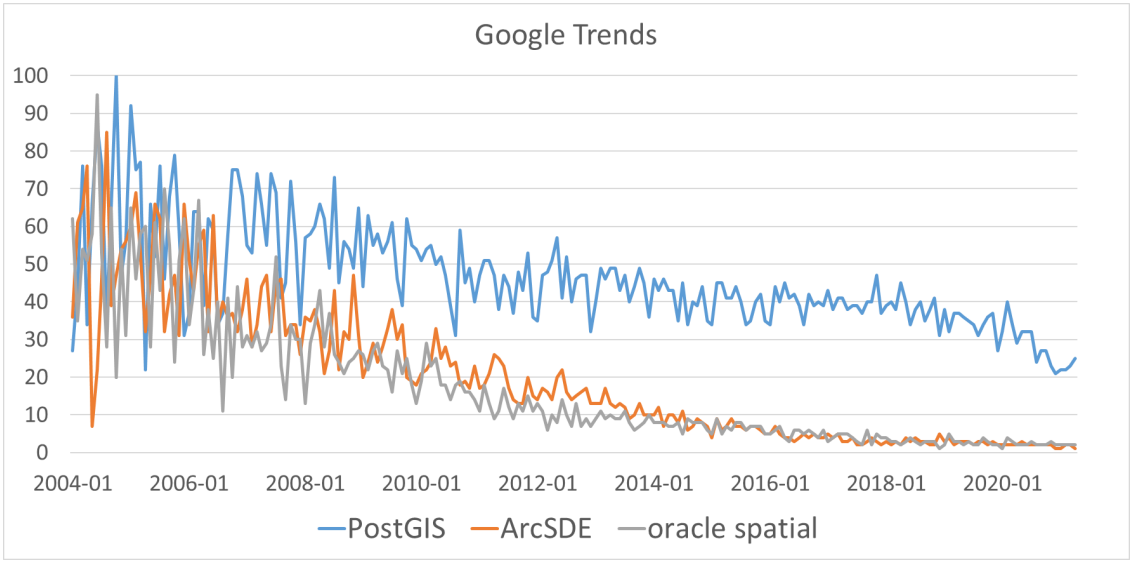
Why PostGIS?

- PostGIS is a PostgreSQL extension that allows location queries to be run in SQL
- PostGIS is a full-featured GIS (https://postgis.net/features/):
  - Spatial data types
  - Spatial indexing
  - > 1000 spatial functions
  - Vector and raster data analysis
  - Web mapping capabilities
  - …

Why PostGIS?



PostGIS is powerful, popular, and for free.



Example use case

- Create a time series of wind speed for > 70° lat and UTC between 08:00 and 12:00.
- Flat files approach:
  - Read each file, cut data within spatial and temporal constraints, aggregate statistics
- DB approach:
  - One line query. Direct access to all data, including filtering by time and space.

Setup and performance

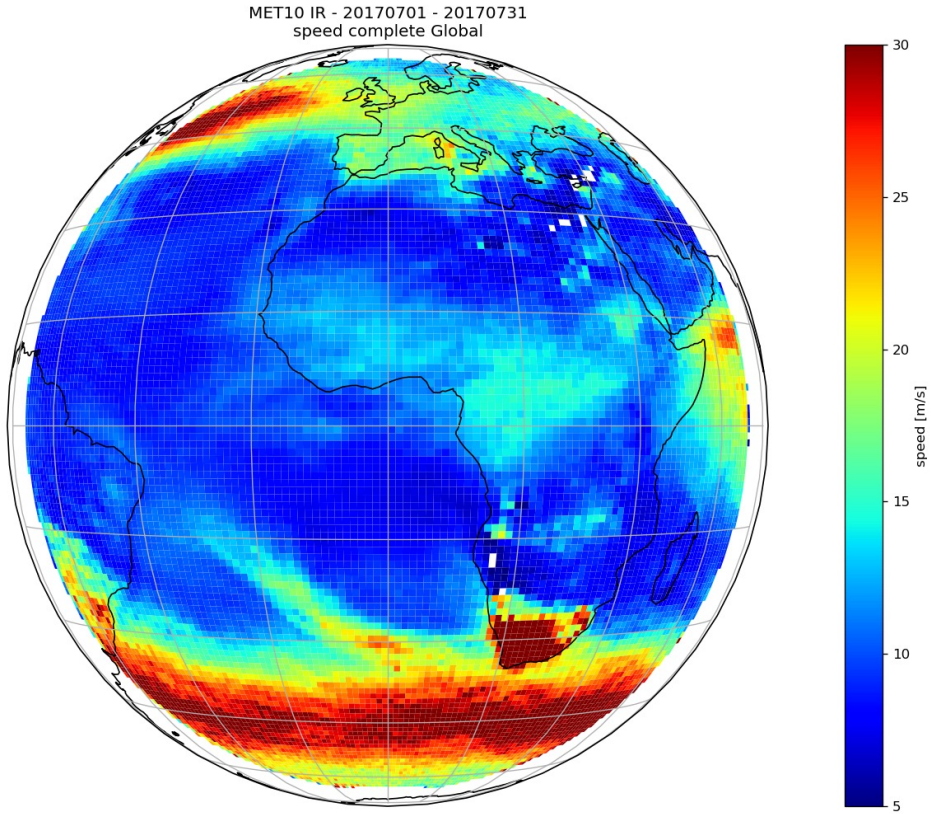
- Entire EUMETSAT archive of AMVs (LEO and GEO) ingested, plus radiosonde and MODIS data
- > 5 TB data, > 20 bn AMVs
- Setup: AMV data partitioned by year, spatially indexed. Both are crucial for a good database performance.

Use cases

- Monthly average of GEO AMVs on a regular grid

Gridded using snapped_geo

```
SELECT AVG(speed) AS DATA,
       COUNT(*) AS COUNT,
       snapped_geo::geography AS centre,
       ST_PolygonFromCentroidGeography(snapped_geo::geography, 0.5, 0.5) AS geometry
FROM amv AS a,
     ST_SnapToGrid(a.geolocation::geometry, -179.5, -89.5, 1.0, 1.0) AS snapped_geo
WHERE pressure_final > 0
  AND pressure_final < 1100
  AND a.product_id = 18
  AND a.quality_index_incl_fct >= 60
  AND a.speed >= 5
  AND a.datetime > '2017-07-01 00:00:00'
  AND a.datetime < '2017-07-31 00:00:00'
GROUP BY snapped_geo::geography
```



Runtime: 2:57 minutes

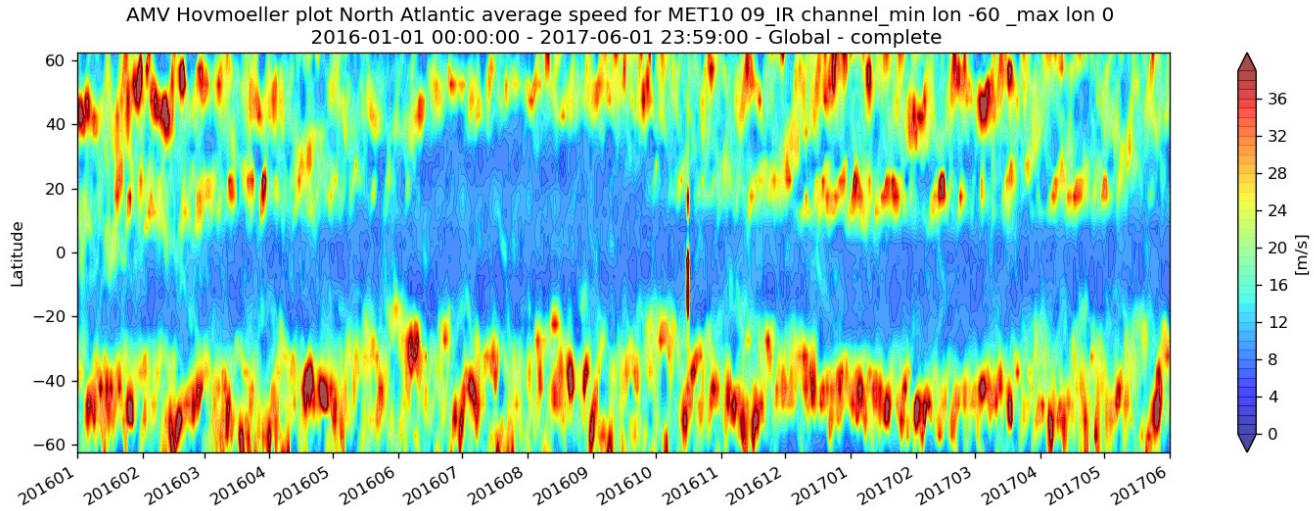
Use cases

- Hovmöller plot for 1.5 years, from -60° to 0° latitude

Runtime: 1:24 minutes

Hovmoeller query

```
SELECT width_bucket(ST_Y(a.geolocation::geometry), -90, 90, 36) AS lats,
       date_trunc('day', a.datetime) AS date,
       AVG(speed) AS value
FROM amv AS a
WHERE pressure_final > 0
  AND pressure_final < 1100
  AND a.datetime > '2016-01-01 00:00:00'
  AND a.datetime < '2017-06-01 23:59:00'
  AND a.product_id = 14
  AND a.quality_index_incl_fct >= 30
  AND a.speed >= 5
  AND ST_X(geolocation::geometry) >= -60
  AND ST_X(geolocation::geometry) <= 0
GROUP BY lats, date
ORDER BY lats
```



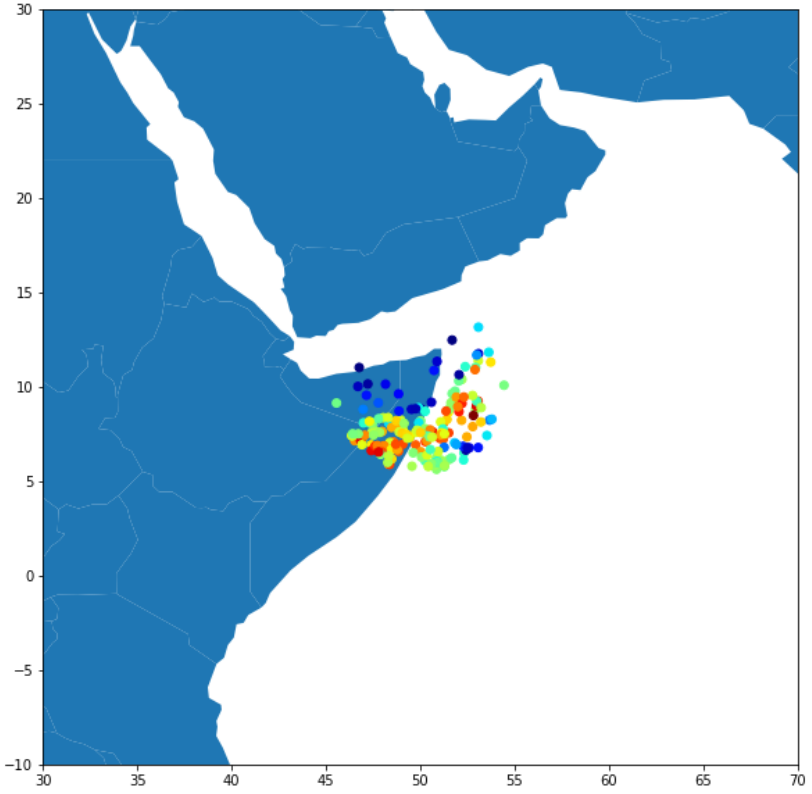
Use cases

- AMVs within 500 km and 3 h of a particular time and location

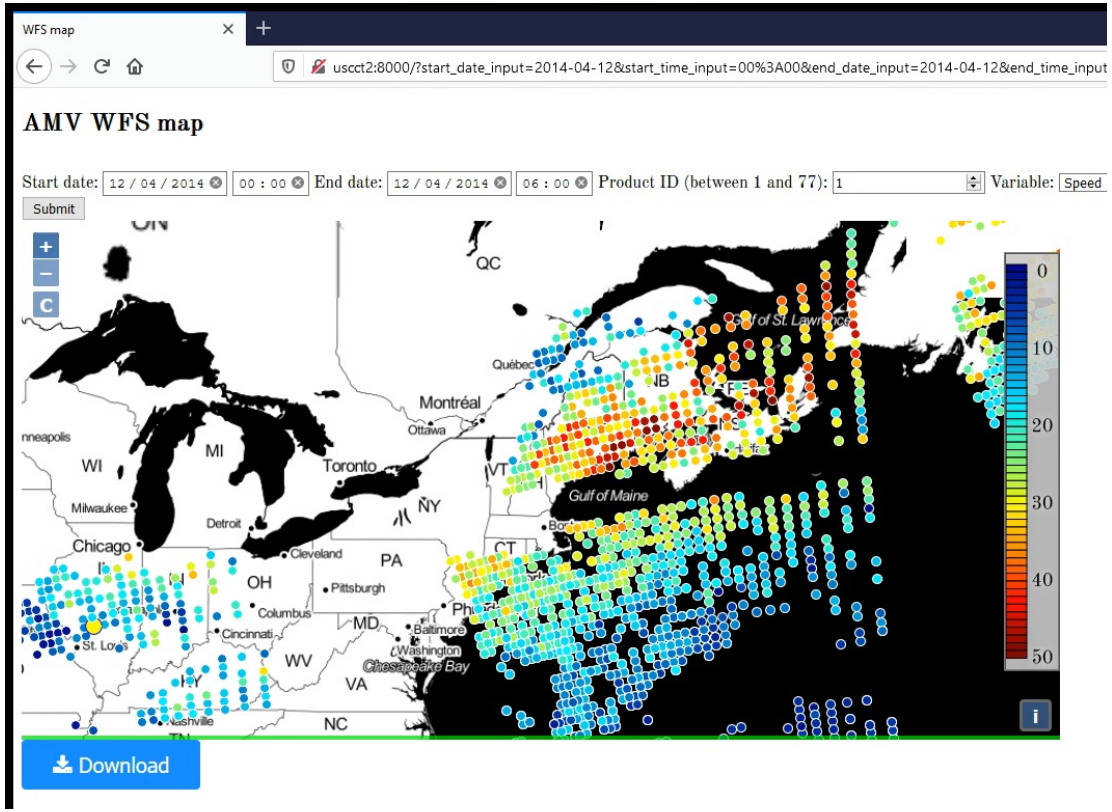
Runtime: 3:19 minutes

Count all AMVs for product ID 18 within 500 km and 3 h of a particular datetime and geolocation in lon/lat.

```
SELECT COUNT(*)
FROM amv
WHERE ABS(EXTRACT(EPOCH
                  FROM (datetime - '2015-03-01 12:00:00'))) < 10800
  AND ST_DWithin(geolocation, ST_GeogFromText('POINT(50.0 10.0)'), 500000)
  AND product_id = 18
  AND datetime BETWEEN '2015-01-01' AND '2016-01-01';
```

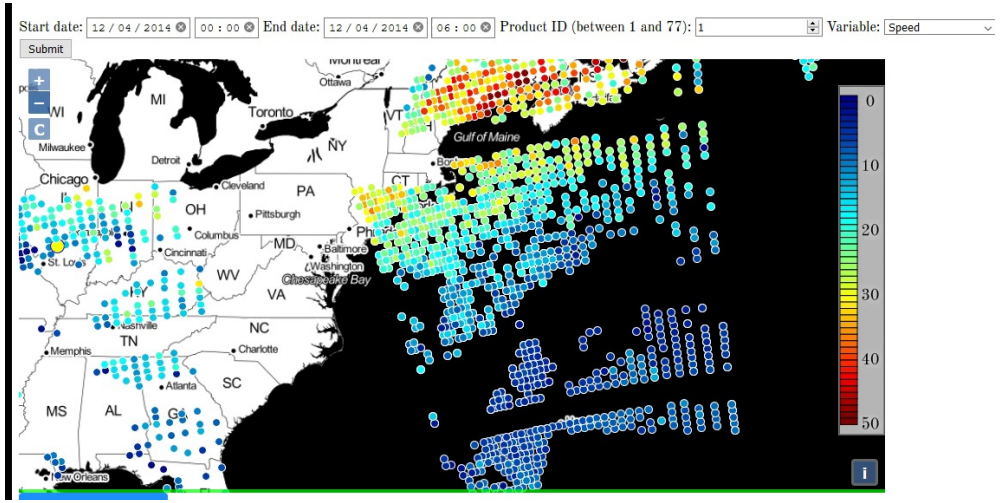


A simple webmapping prototype



A WFS webmap application. Data are directly queried from the PostGIS database via Geoserver. Filtering possible by dataset, variable, height level, and quality level.
Performance improvements to be expected by raster representation of AMVs and pre-tiling.

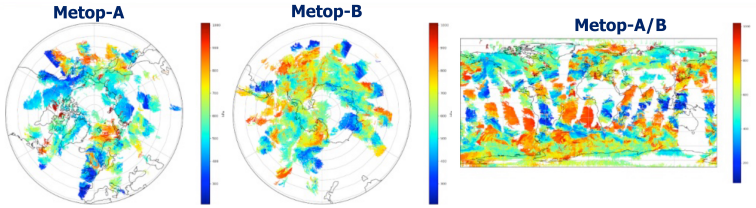
Quick analysis tool of AMVs, including data querying and export.
Overlay of multiple datasets for collocation studies.



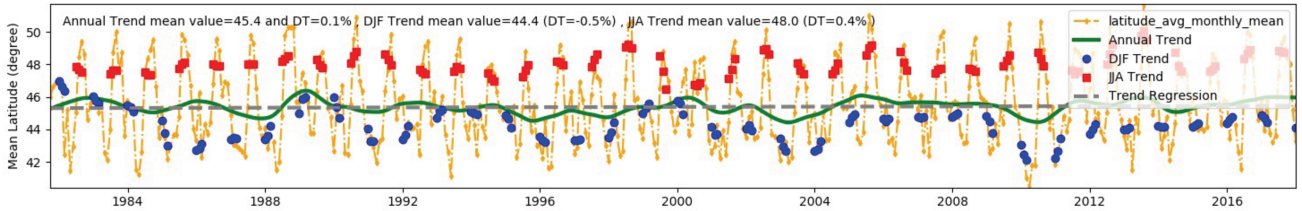
Talks based on data stored in the DB.

Session 6: REPROCESSING AND CLIMATE APPLICATIONS

- Friday 13:50 – 14:00, Marie Doutriaux-Boucher: Climate Data Record of Atmospheric Motion Vectors at EUMETSAT: Status and Perspective



- Friday 14:00 – 14:10, Alessio Lattanzio: Analysis of the Polar Jet with the EUMETSAT Geostationary Atmospheric Motion Vectors Climate Data Record



- Friday 14:20 – 14:30, Roger Huckle: Climate Data Record AMVs from LEO-Satellites