A Unified Framework for Next-Generation Sea Surface Height Mapping

Jeffrey J. Early
NorthWest Research Associates

Cimarron Wortham
NorthWest Research Associates

Jonathan M. Lilly
Planetary Science Institute

Brian Curtis, Mathias Darr, Daniel Neshyba-Rowe

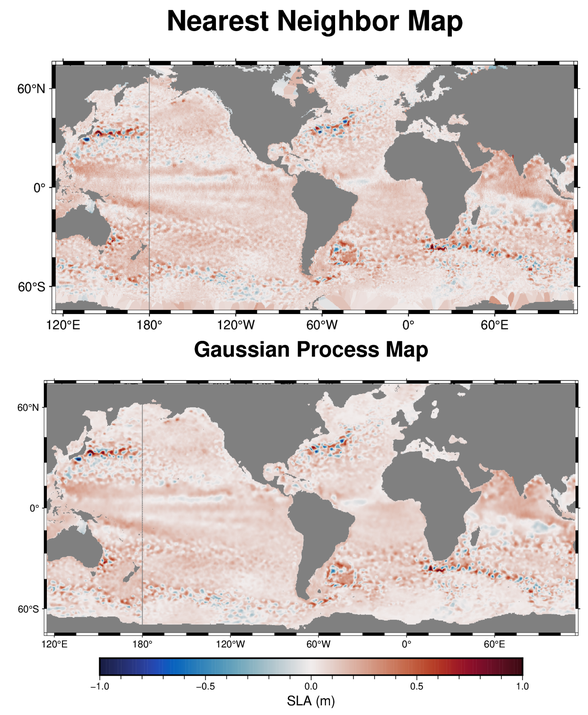
NWRA
Since 1984

Mapping workflow

1. Choose an algorithm

- Nearest neighbor
  - Null hypothesis, 1 parameter model
- Gaussian kernel
  - 2 or 3 parameter, simple model
- Gaussian process regression
  - Includes DUACS covariance function
  - Our own improved “advective” covariance
- Local polynomial fitting
  - Uses generalized weight kernel with few parameters
- Machine learning
  - NeurOST (Martin et al., 2024)
- Your amazing new idea!

2. Show amazing results!



3. Scale-up

- Map 30+ years of data, 2+ billion points
- Dozens of missions, each with unique errors
- Map data in connected basins only
- Map as new data arrives
- Output to standardized file formats
- Compute on the reference ellipsoid
- Compute the error covariances
- Compute the gradients

This step is a significant bottleneck in taking a good mapping idea to full production.

SatMapKit

Lets oceanographers focus on their mapping algorithm and builds standardized scalable output.

New mapping workflow:

1. Set a spatiotemporal region to mapped
2. Implement your **algorithm**
3. SatMapKit uses **OceanDB** to access data and write to a new **file format**
4. Algorithm performance can be **evaluated and uncertainty** assessed.

Algorithm Agnostic Mapping

You create the algorithm, SatMapKit does the mapping

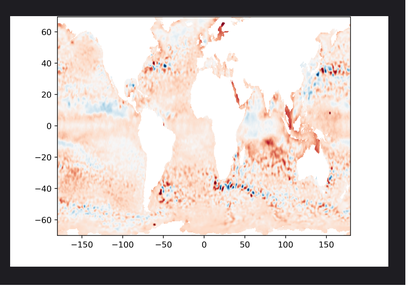
```python
class GeographicGaussianKernelInterpolator( GeographicWindowInterpolator):
    def __init__(self, length_scale: float):
        self.length_scale = length_scale

    def interp(self, data, point):
        r = data.distance
        weights = np.exp(-1/2 * (r/self.length_scale)**2)
        weights /= weights.sum() # normalize
        return np.dot(data.sla_filtered, weights)


date = datetime(year=2013, month=12, day=31, hour=23)
lon_dim = np.arange(-180, 180, 2.5) # resolution of 2.5 deg
lat_dim = np.arange(-70, 70, 2.5)
lon, lat = np.meshgrid(lon_dim, lat_dim)

interp = GeographicGaussianKernelInterpolator(50 * 10**3)
sla = interpolate_using_atdb(lat, lon, date, interp)

plt.figure()
plt.pcolormesh(lon, lat, sla, cmap='RdBu_r')
plt.title('Sea surface anomaly (m)')
```

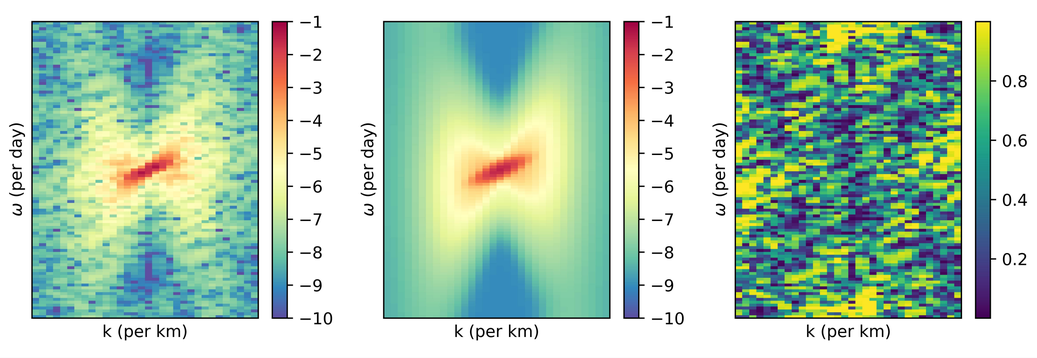


- **Open source, written in python: https://github.com/satmapkit/MapInterp**

An advective covariance function

- Covariance functions based on physically-motivated parametric models not only improve mapping, but also provide physical insight into dynamical processes.
- Using a model that allows for a pdf of velocities, we can create a far more realistic covariance function
- Work with Arthur P. Guillaumin, Queen Mary University of London

Comparison of the spectrum implied by the covariance function to a QG model. The observed periodogram from model output sampled along the altimeter track pattern is shown at left, the expected periodogram with optimized parameters is in the middle, and the magnitude of their difference is shown on the right.



OceanDB: Scalable data access

A database for along-track altimetry and other ocean data

- **Open source, written in python: https://github.com/Nazanne/OceanDB**

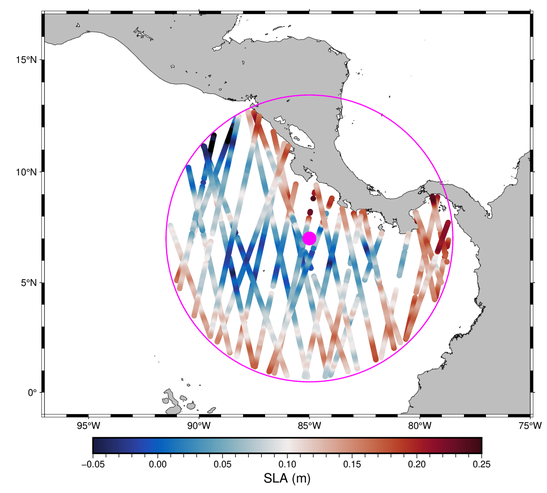
```
$ make shell // Open shell in docker container with Oceandb package installed
$ oceandb init // Creates the database tables
$ oceandb ingest-along-track --start-date 2021-02-01 --end-date 2024-03-01 --missions all
```

- OceanDB indexes location, time, altimetry mission, and ocean basin for optimized queries.

  `geographic_nearest_neighbors_dt`
  `geographic_points_in_r_dt`
  `projected_points_in_dx_dy_dt`
  `projected_points_in_r_dt`

- SatMapKit uses these queries, with the user defined interpolator, to build the maps.
- Basin connection data provided by Josh Willis from S. Fournier et al. (2022): User’s Guide, JPL MEaSUREs Gridded Sea Surface Height Anomalies Version 2205.

Along-track SLA near Panama



Evaluation and Uncertainty

- All linear algorithms can be mapped to the uncertainty estimation, and this can thus be abstracted SatMapKit
- All methods can be assessed using bootstrapping, and thus handled by SatMapKit
- SatMapKit will work directly with the Ocean Data Challenges tests for mapped altimetry products.

A better file format

A new compressed file format based on continuous functions

Users want data at arbitrary locations, e.g.,

```
eta = ssh(lat,lon,date)
```

Traditional equirectangular grids are inefficient, not placing points appropriate to the DOF.