Runoff Tools

It's all Aidan's fault!

How to map coarse river/runoff fields to the coasts of a fine grid efficiently. DaiTrenberth approx 60 years of monthly data.

I submitted a job with the same specifications as yours, **2400 CPU,** but with a walltime of **2 hours**. Unfortunately it hit the walltime limit and was killed. I am not inclined to try it again with a longer walltime as I think it will just time out again. I originally had a **256 CPU job run for 12 hours and time ou**t. That probably should have been enough time.

Problems identified

Naive searching.
Failure to mask model open ocean and land
No masking of input runoff files.
Huge files.

Not such a problem in the past.
0.1 degree or finer grids common now.

Solution.

Identify coast points of model.

Only use source points with runoff. <1% of points

Use kd-trees to efficiently find nearest neighbours. Both for source to model and spreading large runoffs along coasts.

NetCDF-4 compression.

Results

0.1 GFDL tripolar grid (3600*2700).

1 cpu – 2 minutes!

55GB to 450MB!

Keep intermediate files to allow examination of connections (easy to see in Ferret) and allow hand editing.

Can spread along coast or out to sea.

Intermediate files

runoff_connection_nn.nc
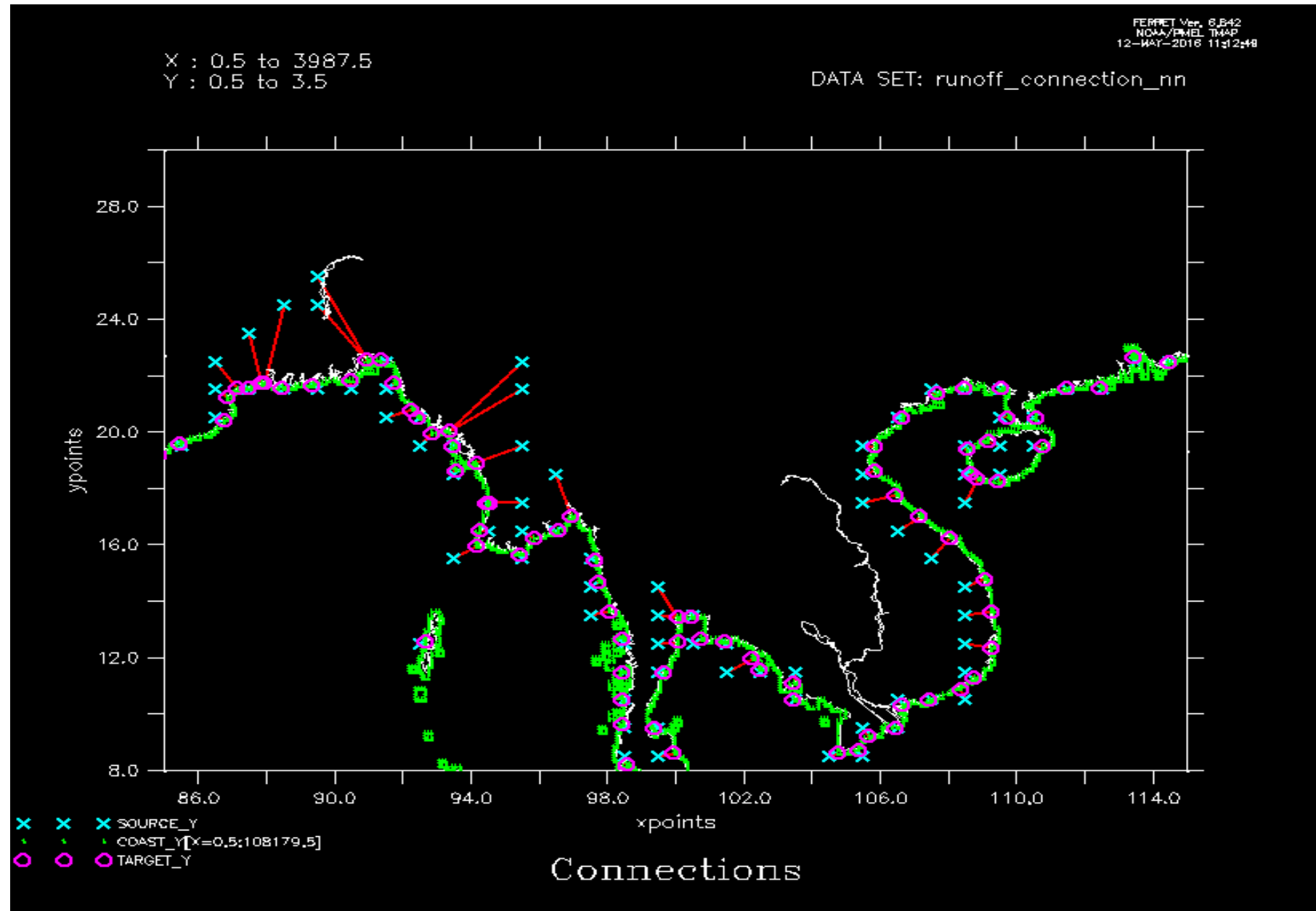- Source and model locations (real and index space)
- Connections between source points and model coast.
- Distances and areas

runoff_weights.nc
- Contains locations and weights of distributed runoff.

NN Connections

Lessons Learned

Brute force doesn't cut it nowadays. We've got to be smarter.

Use efficient data structures.

Only do the work that needs to be done.

Use netCDF4's capabilities! Chunking, compression etc.

Finally

https://github.com/OceansAus/runoff_tools

Zhi Liang looking at incorporating into GFDL suite of tools.
Are there other areas which can benefit from these ideas.