

PITFALLS IN CLIMATE NETWORK CONSTRUCTION

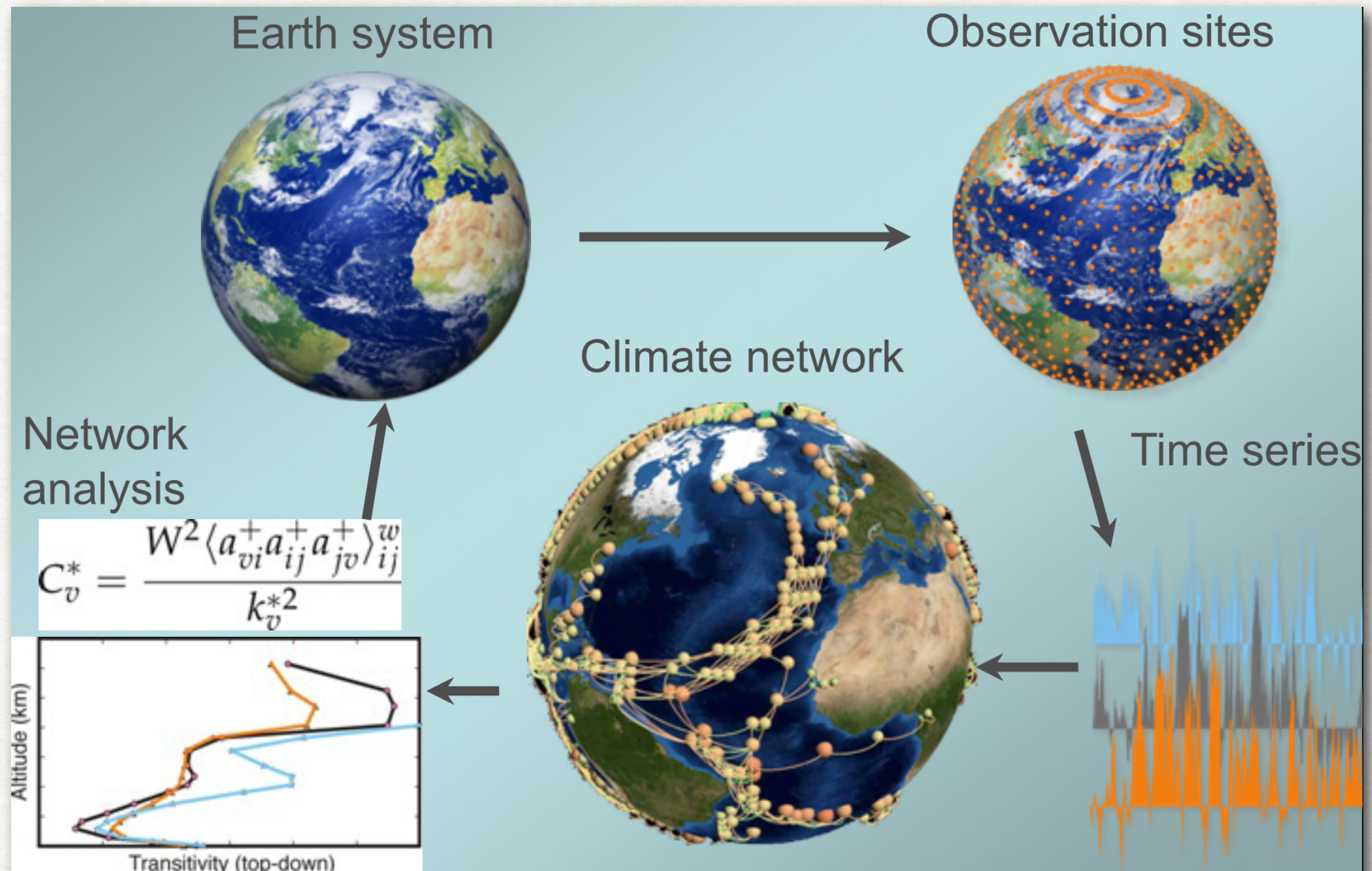
A STATISTICAL PERSPECTIVE

MORITZ HAAS, BEDARTHA GOSWAMI, ULRIKE VON LUXBURG

THEORY OF MACHINE LEARNING, UNIVERSITÄT TÜBINGEN

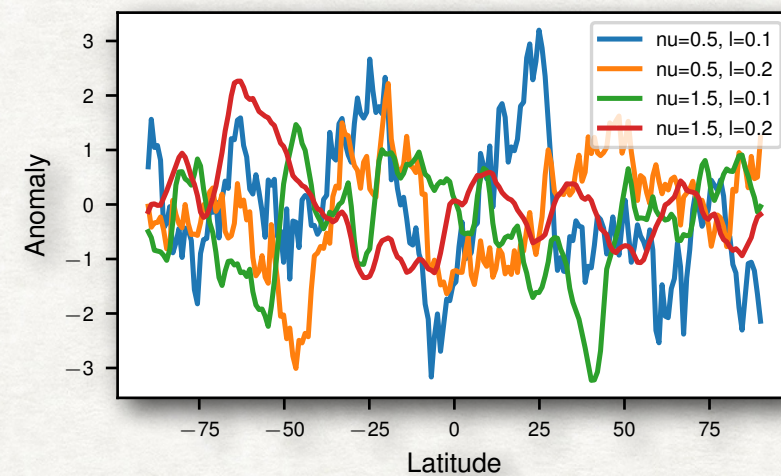
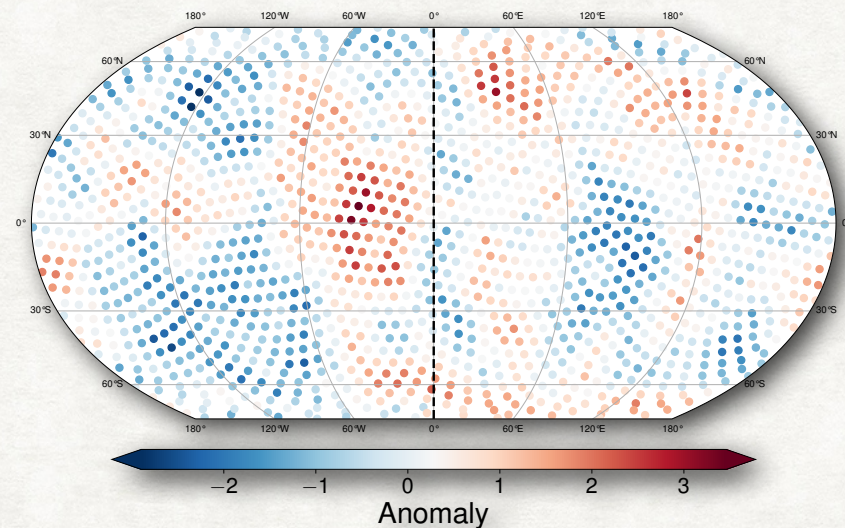
MACHINE LEARNING FOR CLIMATE SCIENCE, UNIVERSITÄT TÜBINGEN



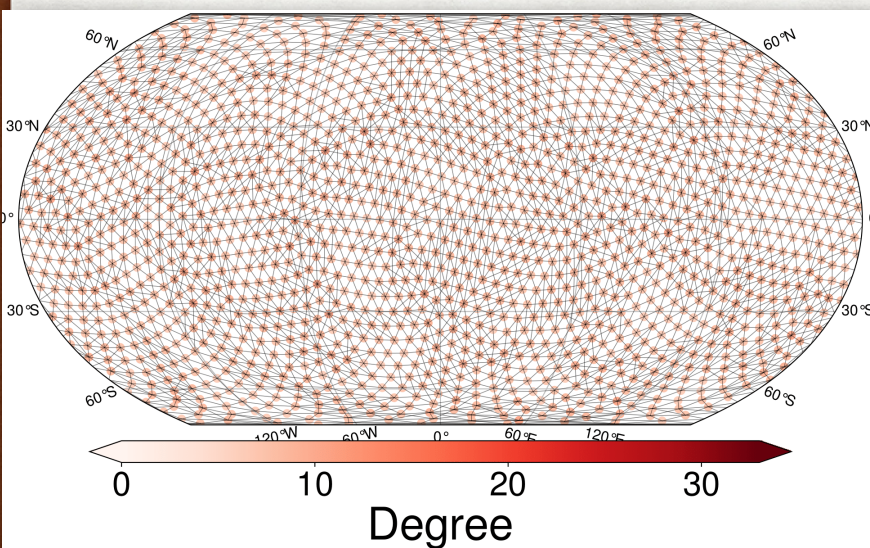


Jürgen Kurths. Climate Networks and Extreme Events.

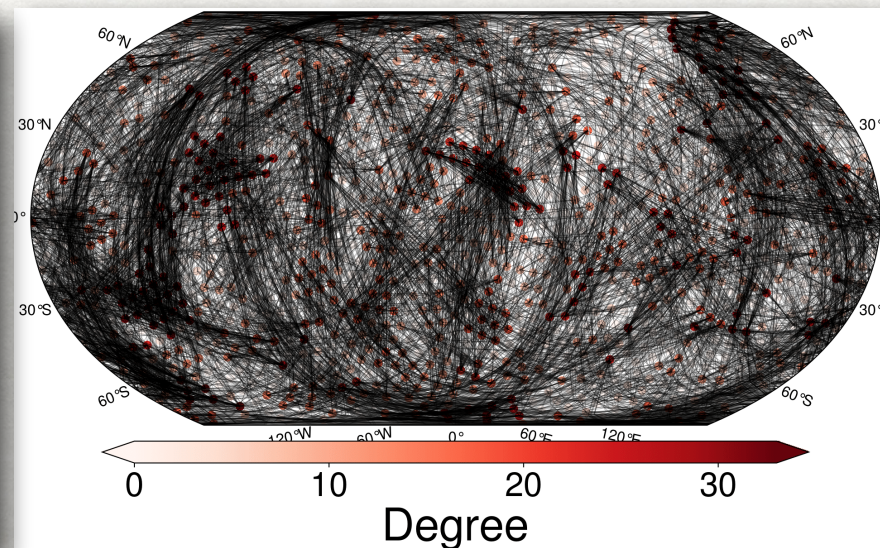
Generate isotropic random fields with various spatio-temporal hyperparameters



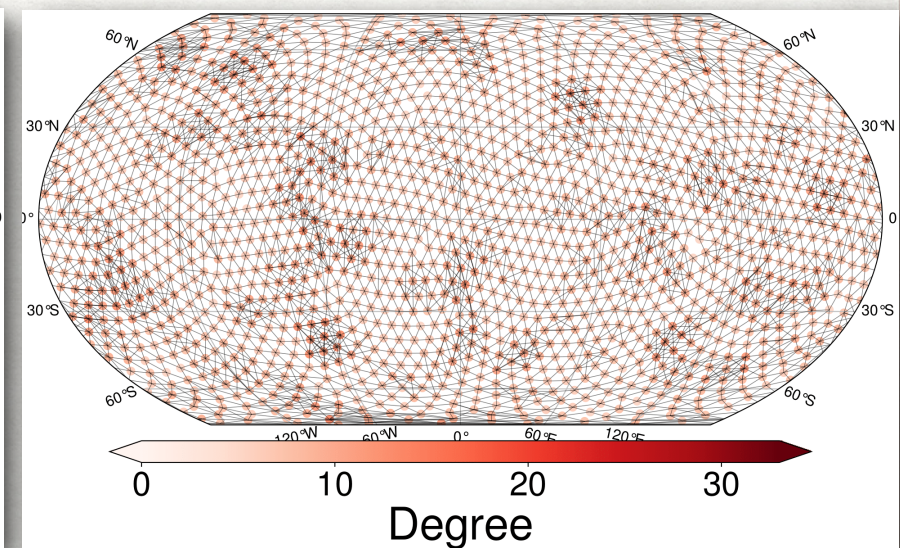
How do estimated networks differ from their ground truth?



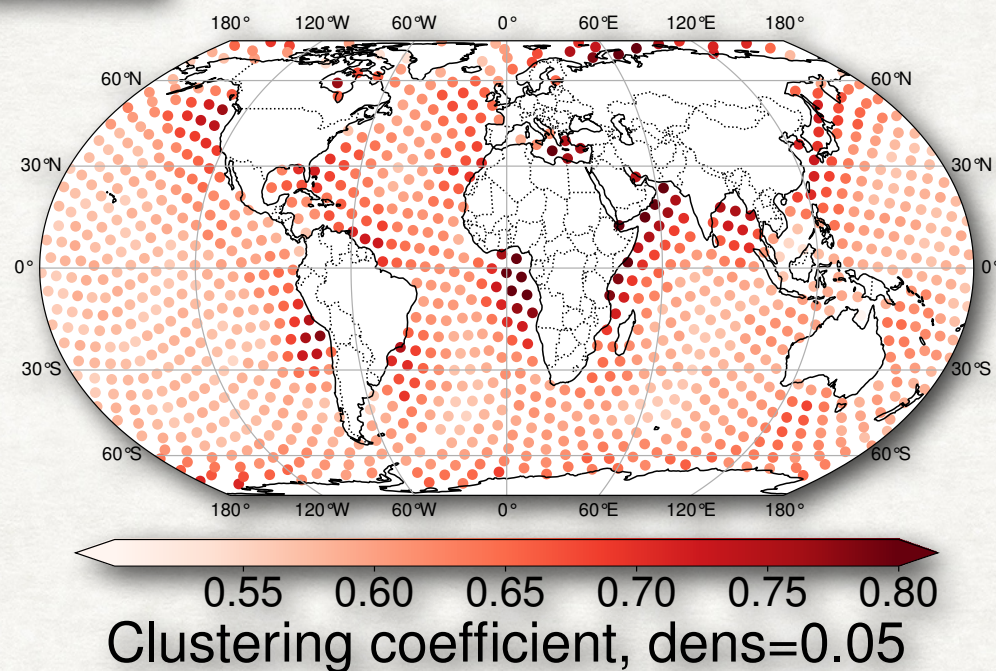
True Network



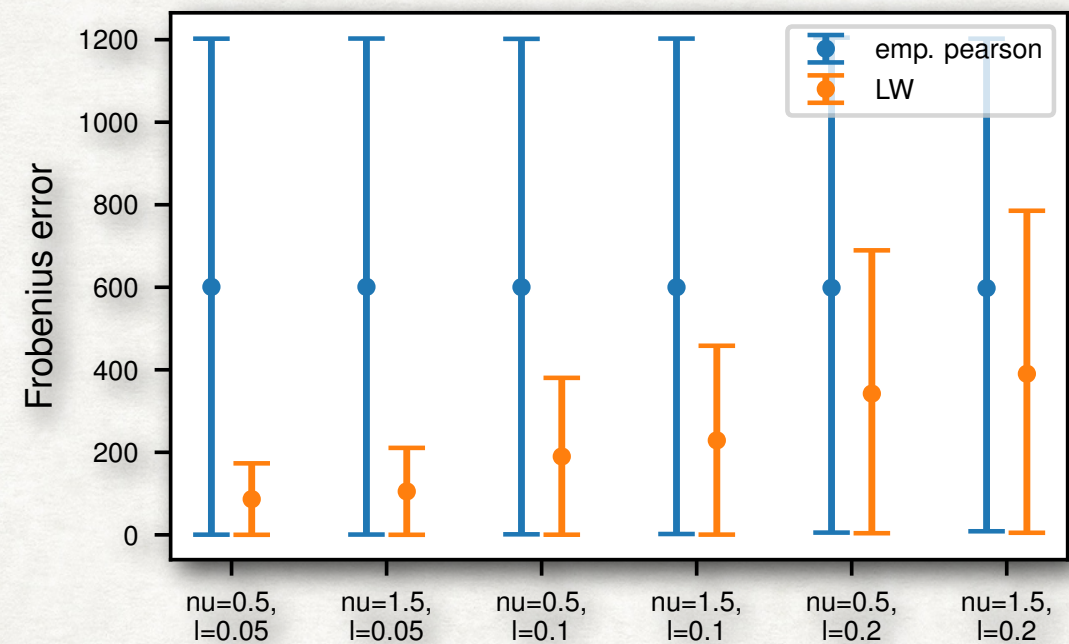
Empirical Pearson
Correlation Network



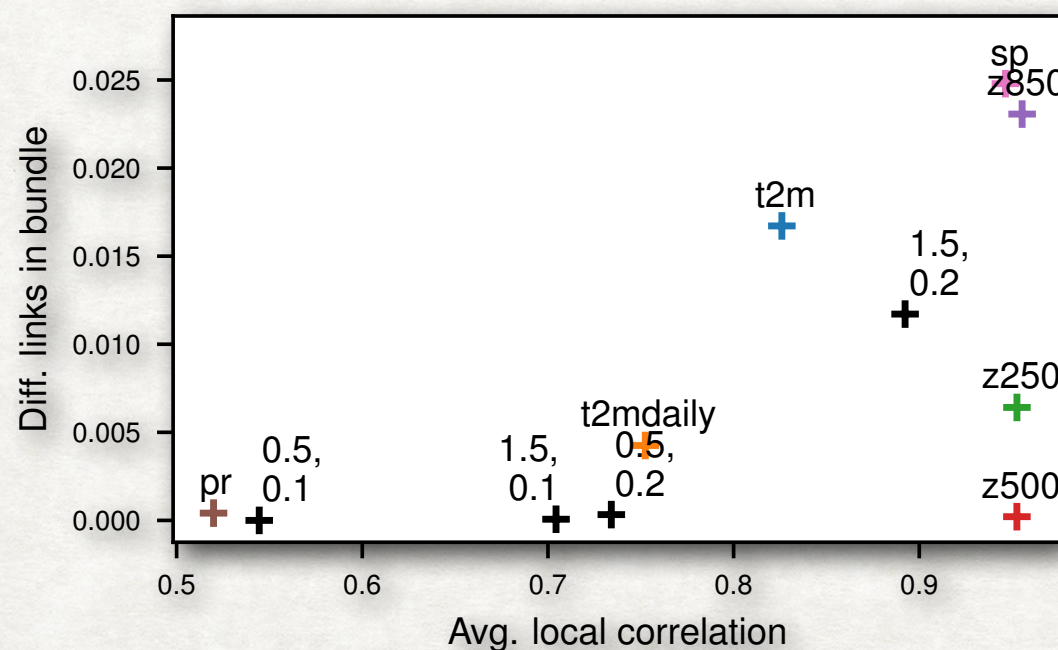
Empirical Spearman
Correlation Network



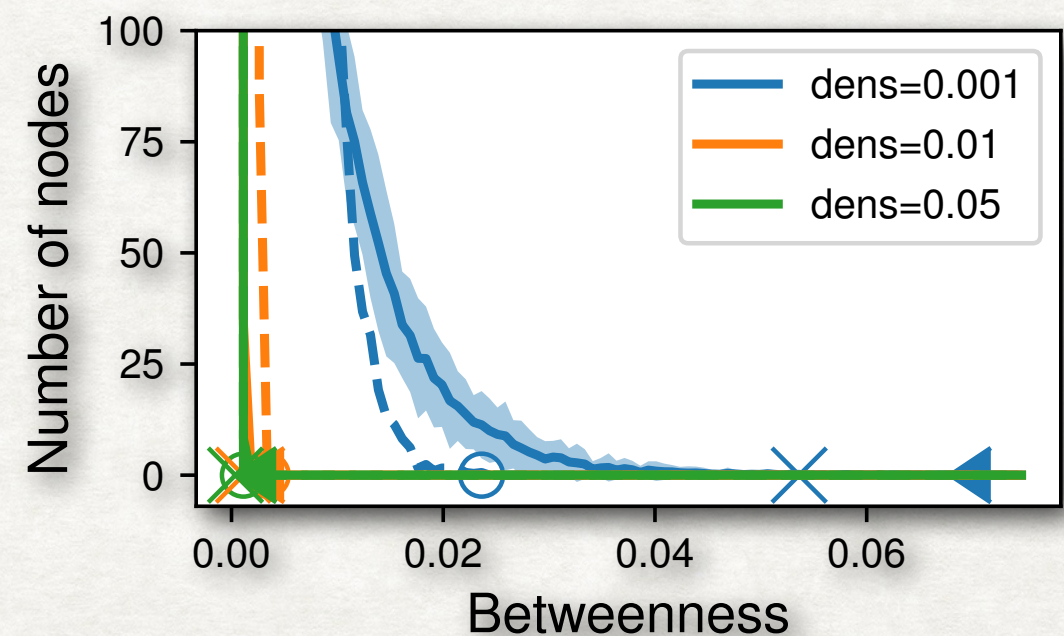
Isotropic process
But anisotropic grid



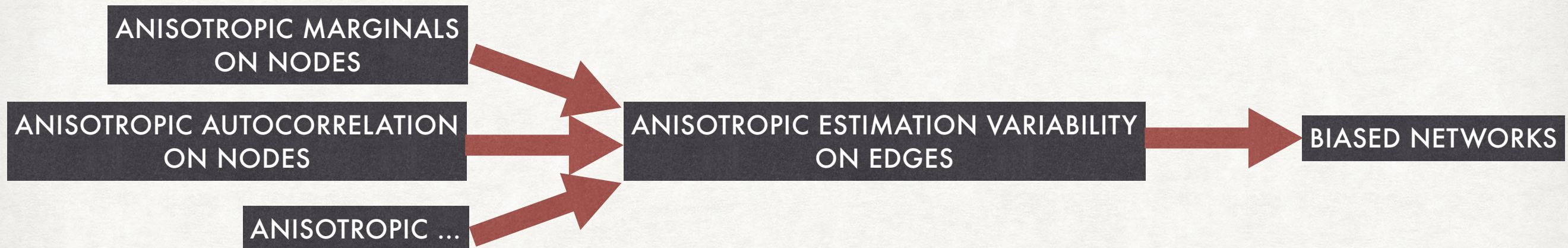
Basic vs Ledoit-Wolf
correlation estimator



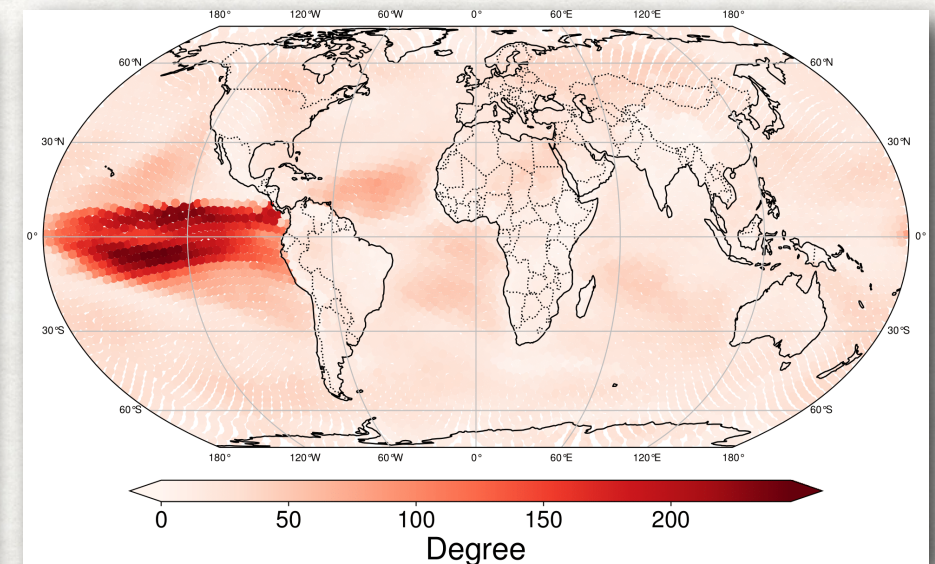
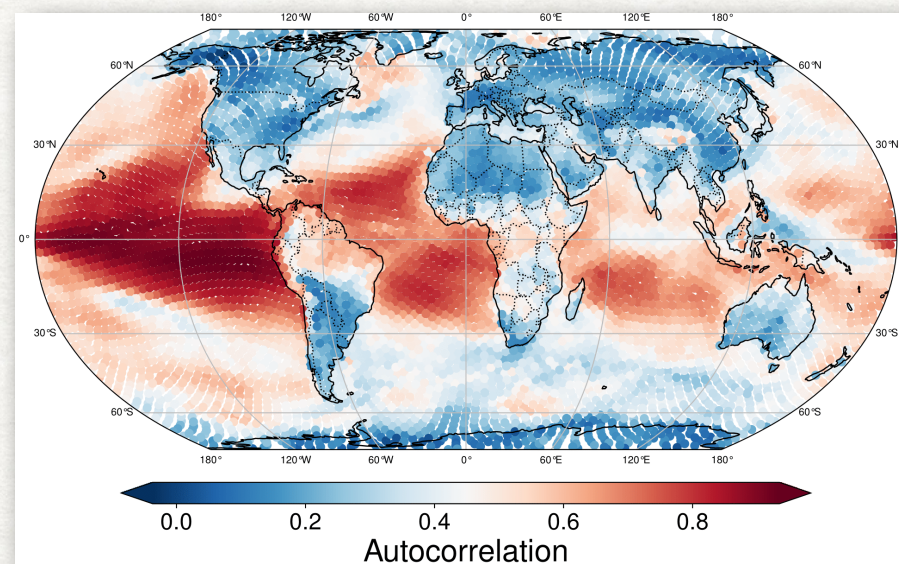
Bundling behaviour of climatic variables



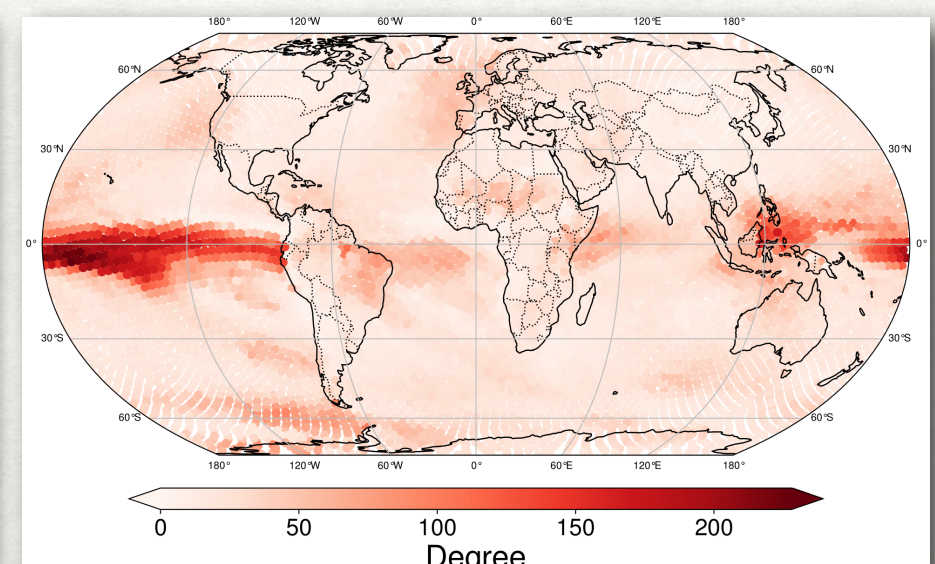
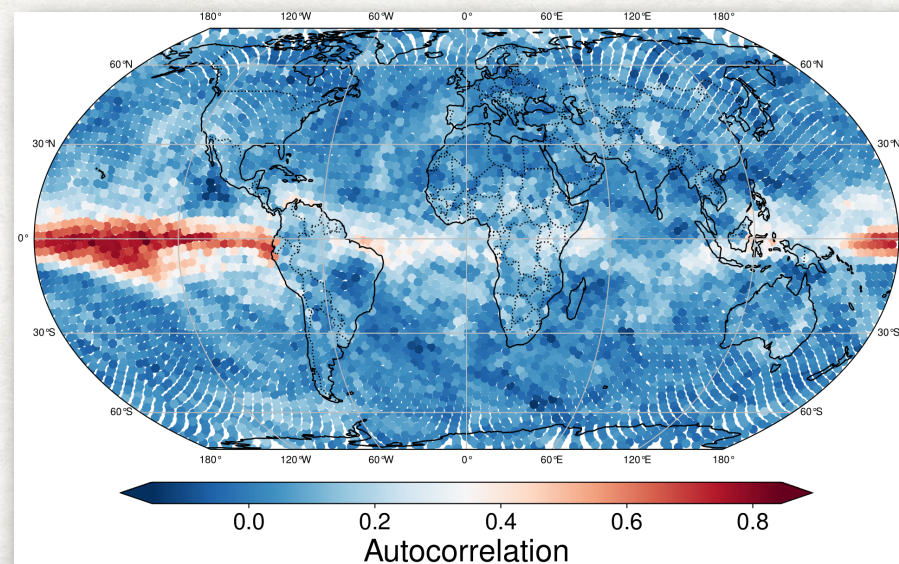
Empirical distortion of network measures



ERA5 monthly t2m



ERA5 monthly precipitation



HOW MANY SPURIOUS LINKS DOES MY NETWORK CONTAIN?

Not adequately addressed so far!

Future work:

Estimate estimation variability in each edge with multiple correlation estimates
(bootstrapping, subsampling, ...)



Get a statistically meaningful sense of how much you can trust your network.



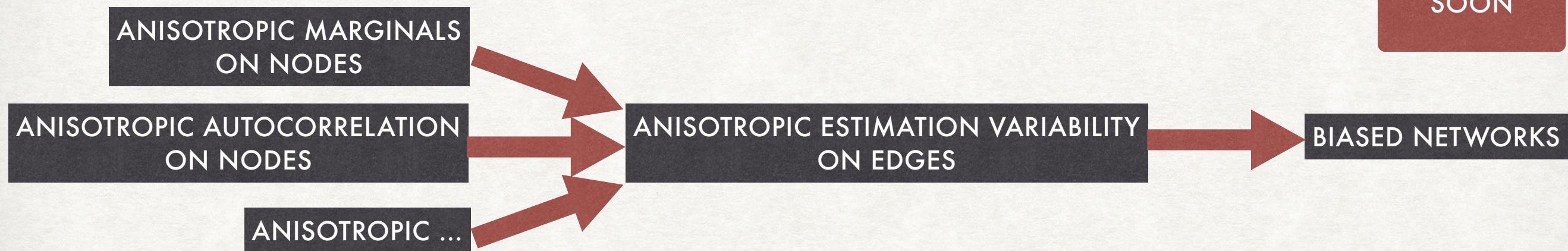
GAINED INSIGHTS



Given non-negligible estimation errors, observe:

- Suitable estimator can prevent many false links!
- Spurious link-bundles and spuriously dense/sparse regions
- Extreme distortion of network measures
(degree, clustering coefficient, betweenness, ...)

PAPER ON
HOMEPAGE
SOON



Trade-off:

Sparse networks

- Low fraction of false links,
- Single false links have high impact,
- ...



Dense networks

- Larger fraction of false links,
- More stable network characteristics,
- More informative for community detection,
- ...