# Optimizing the Antarctic Observing Network: Motivation and Theory

Gregory Hakim, Natalia Hryniw, Guillaume Mauger,
Karin Bumbaco, and Eric Steig

*University of Washington*

NCAR Partners: Jordan Powers and Kevin Manning

Sponsor: National Science Foundation through Award # ANT-1043090

# Motivation: Why Objective Network Design?

**Measurements are often sited subjectively**

- Specific needs dictate location
  - E.g. support transportation (airports; bases)
- Often serve other, unintended, purposes
  - E.g. weather forecasting & climate monitoring
  - And we want redundancy and fault tolerance

**Works fine when resources are abundant**

- Distributed costs across agencies (networks of networks)
- Installation, maintenance, and data distribution costs low

# Motivation: Why Objective Network Design?

**Objective network design informs on resource allocation**

- How do we make the best weather forecasts?
- How do we best monitor climate?
- Can use along with other info, but good to know

**Elements of an objective approach**

- Performance measures ("metrics")
  - need to know how to quantify what you want from a network
- Constraints (access, funds, existing stations, etc.)
- A robust algorithm
  - Theory and interpretation in this talk; main results in next talk

# Why Antarctica?

U Wisconsin AMRC/SSEC

1. Surface measurements are sparse
   – access is difficult
   – expensive to install & maintain

2. Impact is potentially large
   – maximum benefit for fixed investment
   – better weather forecasts & climate monitoring

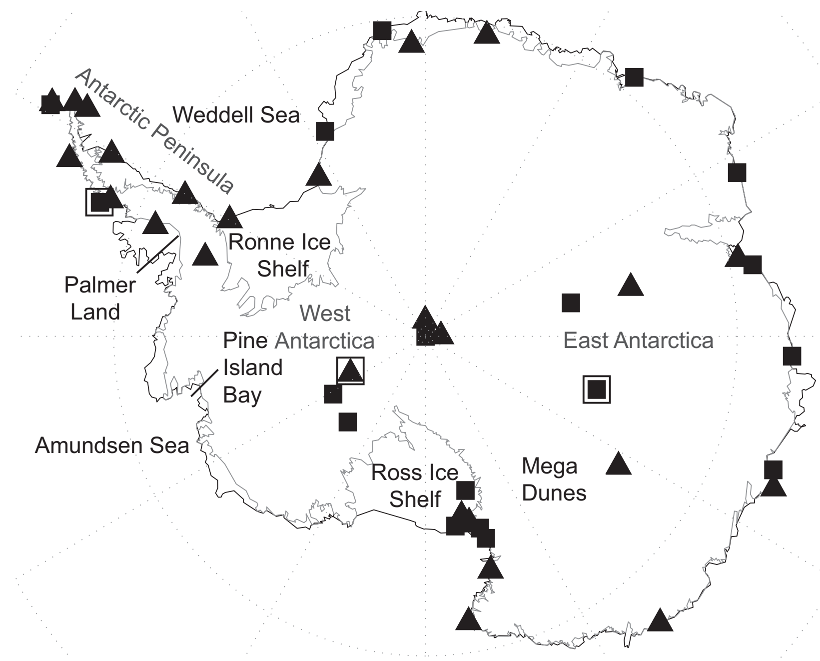
Project goals

– develop techniques to optimize network performance
– test methods on real (OSE) and simulated (OSSE) data
– Recommendations if performance metrics are supplied

# Performance of the Current Network

Bumbaco et al. (2014)

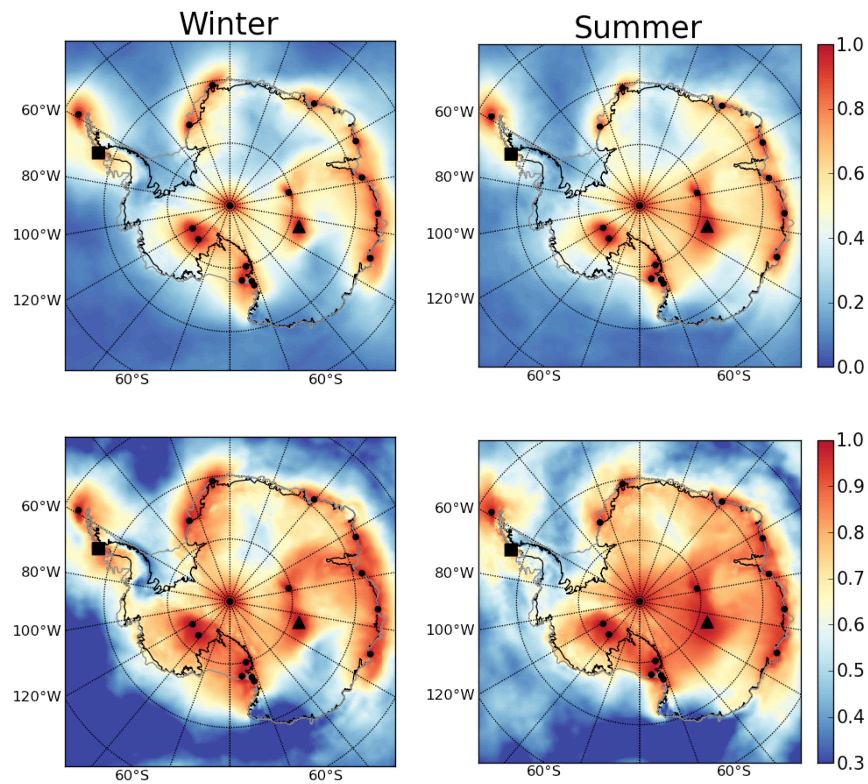
- AWS stations that report regularly
  - 75% and 90% (squares)
  - 2008-2012
- AMPS gridded data
- Two main goals
  - Variance explained
  - Spatial correlations



# Performance of the Current Network

daily

weekly

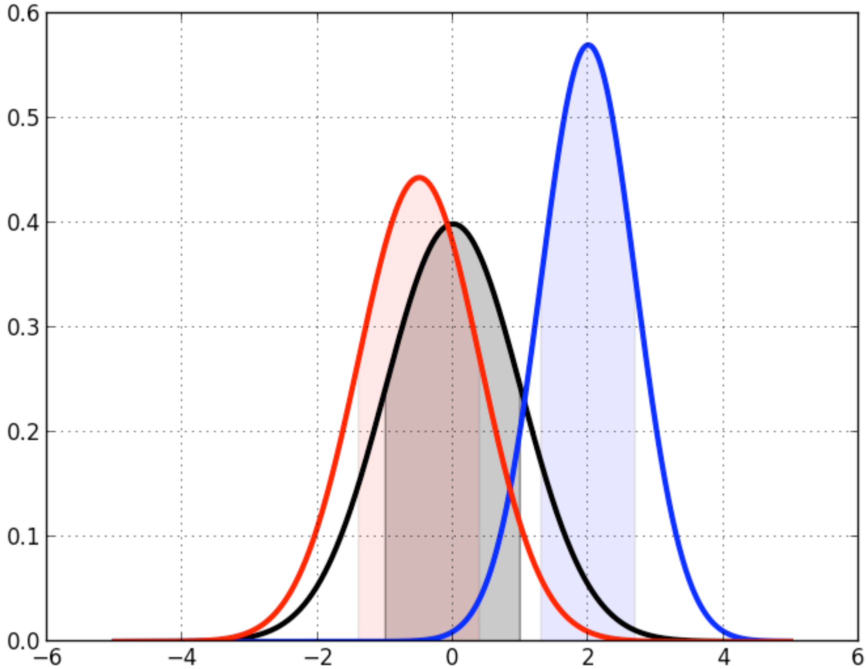



Fraction of temperature variance explained by 90% stations

- Large gaps, even for weekly averages
- Analysis only; doesn't consider forecast performance

# How to pick a station location?

Find the point having the largest impact on a performance measure.




Metric distributions for:

- **Reference (e.g. climatology)**
- **Location #1**
- **Location #2**

Pick location #2 because it reduces variance (uncertainty) the most over the reference.

# Theory I: Scalar Performance Metric *J*

Given: statistics for *J* and the field of interest (*x*) from a reference source

Approximate changes to *J*: $\delta J \doteq \left[\frac{\partial J}{\partial \mathbf{x_0}}\right]^T \delta \mathbf{x_0}$

The change in the variance of *J* is given by:

$$\delta\sigma^2 = \left[\frac{\partial J}{\partial \mathbf{x_0}}\right]^{\mathrm{T}} (\mathbf{B}_{i-1} - \mathbf{B}_i) \left[\frac{\partial J}{\partial \mathbf{x_0}}\right]$$

$\mathbf{B}_i$ is a known function of $\mathbf{B}_{i-1}$ and**, for a single point**, this simplifies to:

$$\delta\sigma^2 = \frac{[\mathrm{cov}(J, x_j)]^2}{\mathrm{var}(x_j) + r_i^2}$$

# The Algorithm

$$\delta\sigma^2 = \frac{[\mathrm{cov}(J, x_j)]^2}{\mathrm{var}(x_j) + r_i^2}$$

1. Evaluate $\delta\sigma^2$ for all potential locations
   - Pick location with largest $\delta\sigma^2$
   - Based on ensemble data drawn from AMPS grids
2. Remove impact of an observation at this location
   - Update the ensemble
3. Go back to 1, and look for the next location
   - Stop when noise dominates (choice no different than random)

Repeat many times (10,000+) to account for uncertainty.

Key to interpretation: spatial correlation length scale

- Some locations have information on large regions

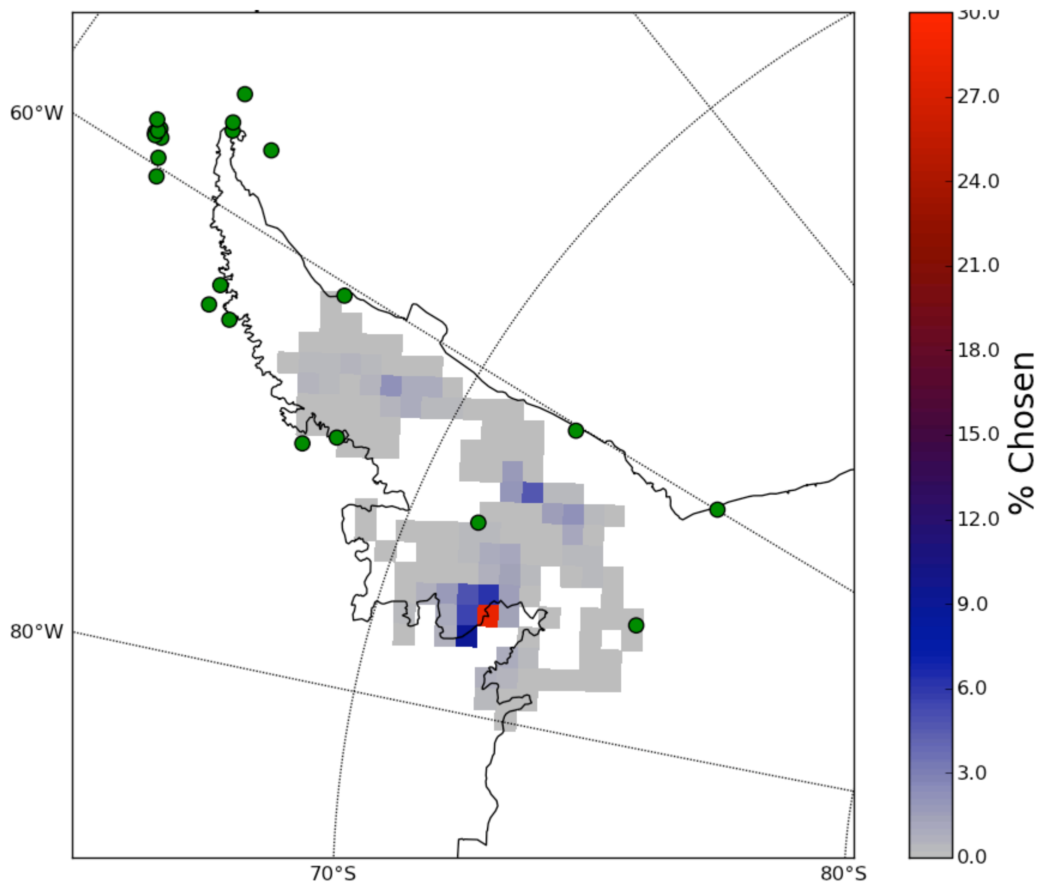
# Results for Spatial Averages

*J*: spatial average temperature for three locations

- Peninsula
- West Antarctica
- East Antarctica

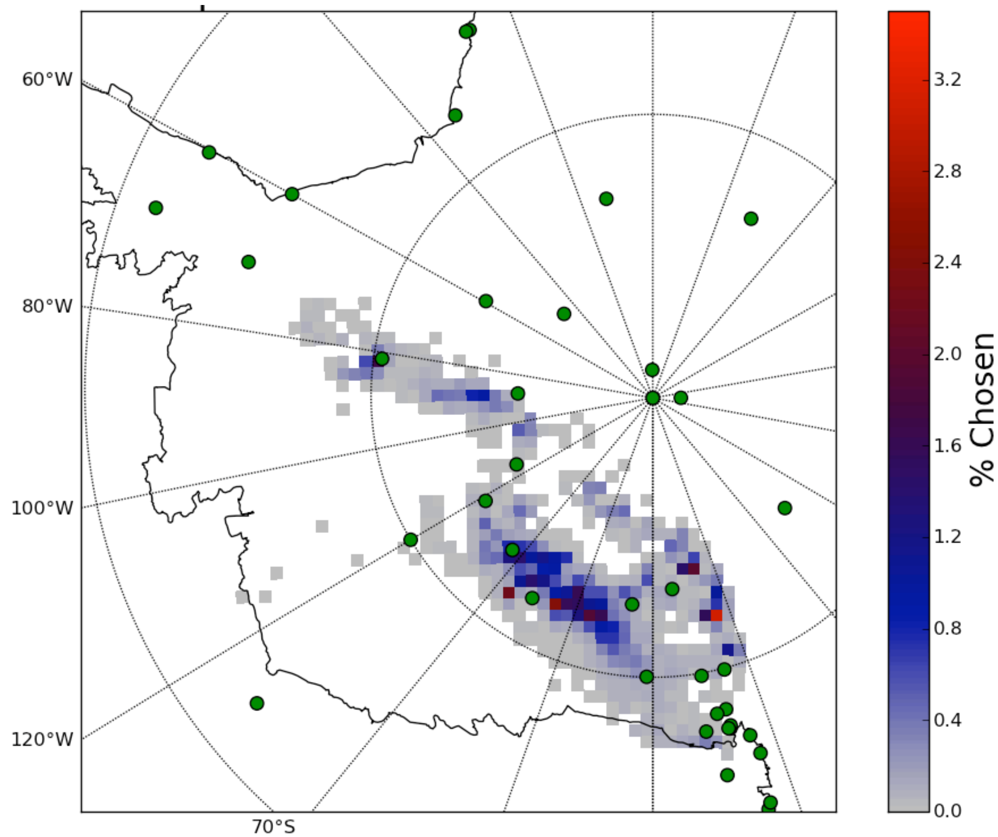
Find first station only (next talk will expand on this)

# Peninsula First Station



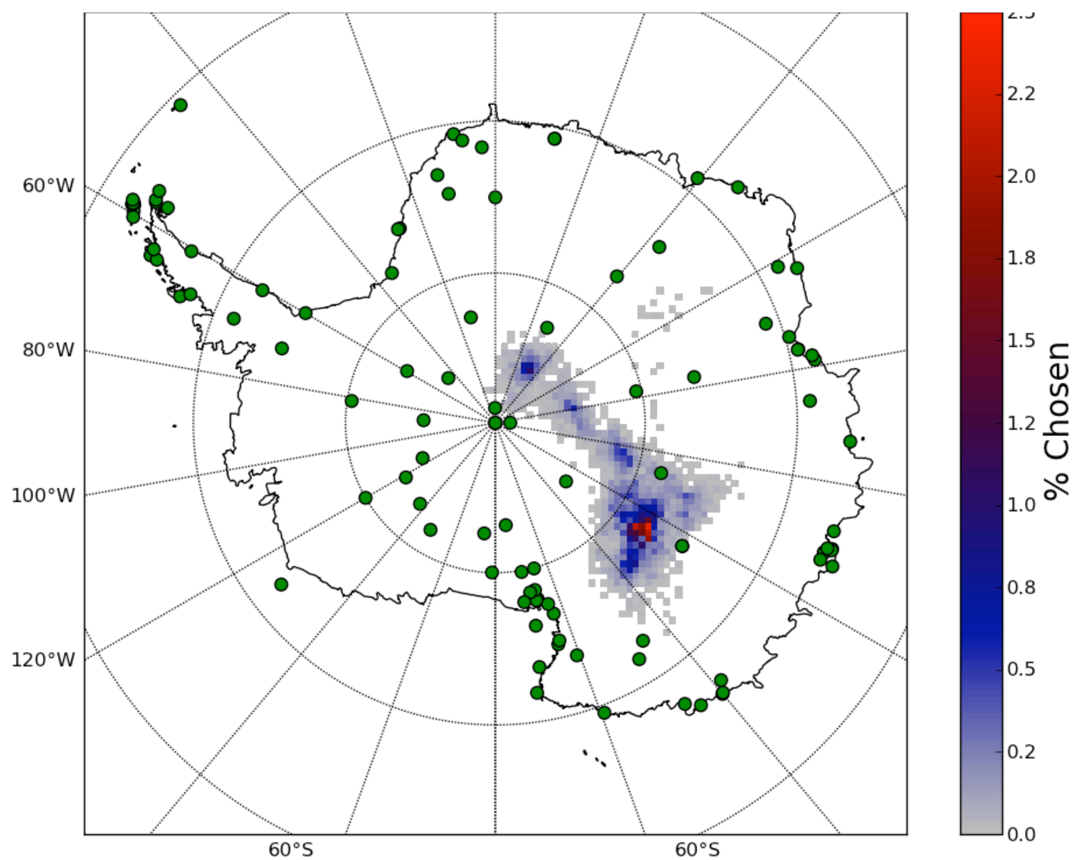

~Palmer Land

# West Antarctica First Station



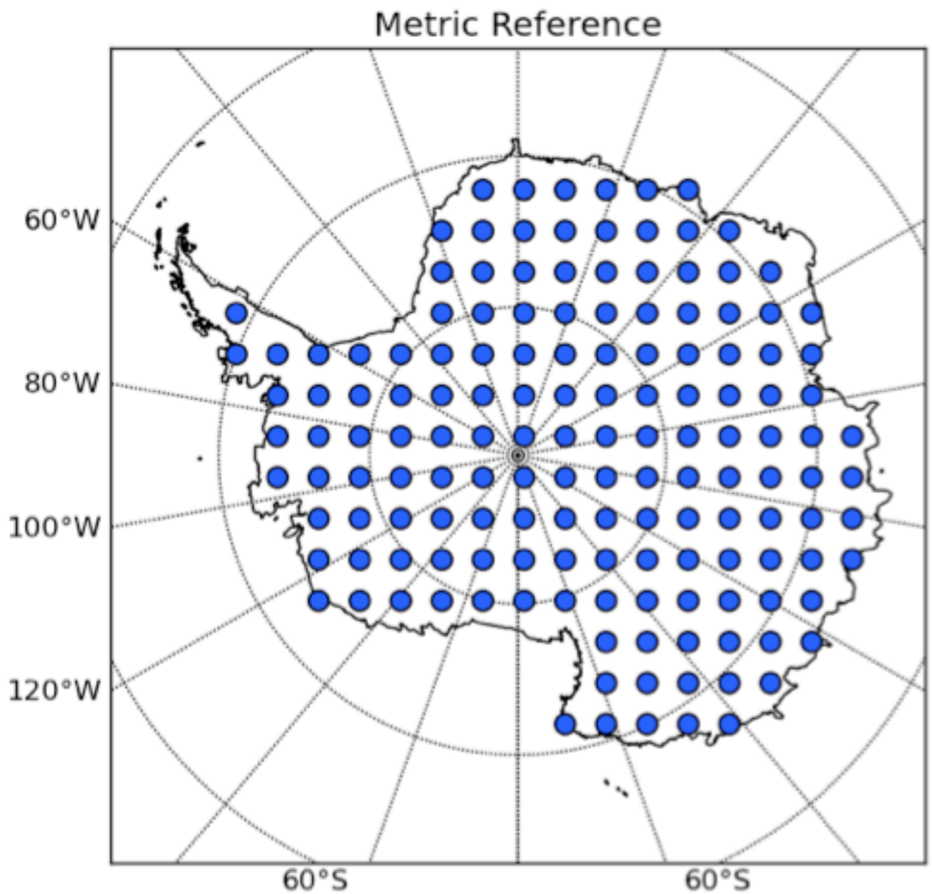

Along flanks of the Transantarctic mountains

# East Antarctica First Station




MegaDunes area. Even more sharply defined when sampling error is treated more carefully.

# Theory II: Optimize Over Many Performance Measures



# Theory II: Many Performance Measures

- Similar to previous, but now **J** = $[J_1, J_2, \ldots, J_m]^T$
  - find stations that optimize ALL *J*s simultaneously
- Approximate changes to **J**:

$$\delta \mathbf{J} = \mathbf{J}(\mathbf{x} + \delta \mathbf{x}) - \mathbf{J}(\mathbf{x}) \approx [D\mathbf{J}(\mathbf{x})]^{\mathrm{T}} \delta \mathbf{x}$$

- The change in the **co**variance of **J** is given by:

$$\delta \mathbf{\Sigma} = D\mathbf{J}(\mathbf{x})^{\mathrm{T}} \left( \mathbf{B}_{i-1} - \mathbf{B}_i \right) D\mathbf{J}(\mathbf{x})$$

- As for the scalar case, for a single point:

$$\delta \mathbf{\Sigma} = [\mathrm{cov}(\mathbf{J}, x_j)\mathrm{cov}(\mathbf{J}, x_j)^{\mathrm{T}}][\mathrm{var}(x_j) + r_i^2]^{-1}$$

- Summary measure of $\delta\Sigma$: trace ("covariance trace" method)

# Theory II: Many Performance Measures

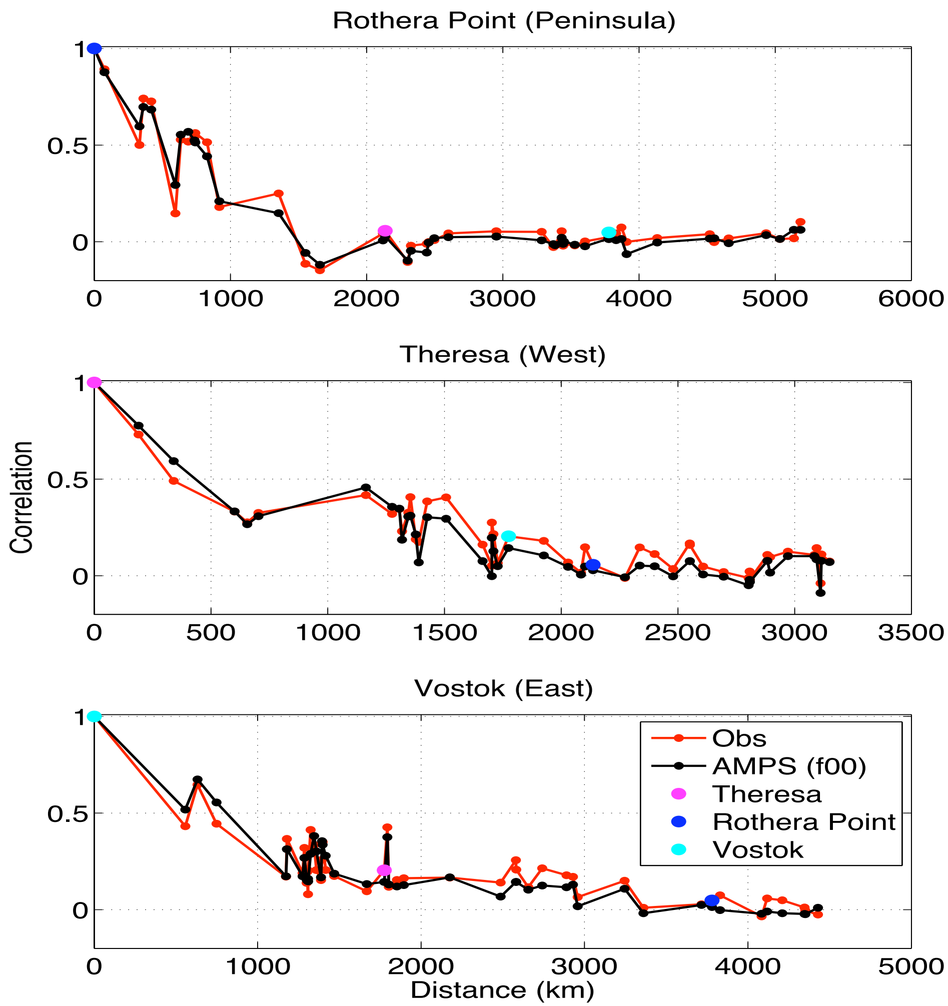
Alternative approach: relative entropy

$$D_{\mathrm{KL}}(p_2||p_1)_{cov} = \frac{1}{2}\left\{\mathrm{tr}\left(\mathbf{\Sigma}_0^{-1}\mathbf{\Sigma}_1\right) - M - \ln\frac{|\mathbf{\Sigma}_1|}{|\mathbf{\Sigma}_0|}\right\}$$

Spatial correlation reduces the determinant (last term)

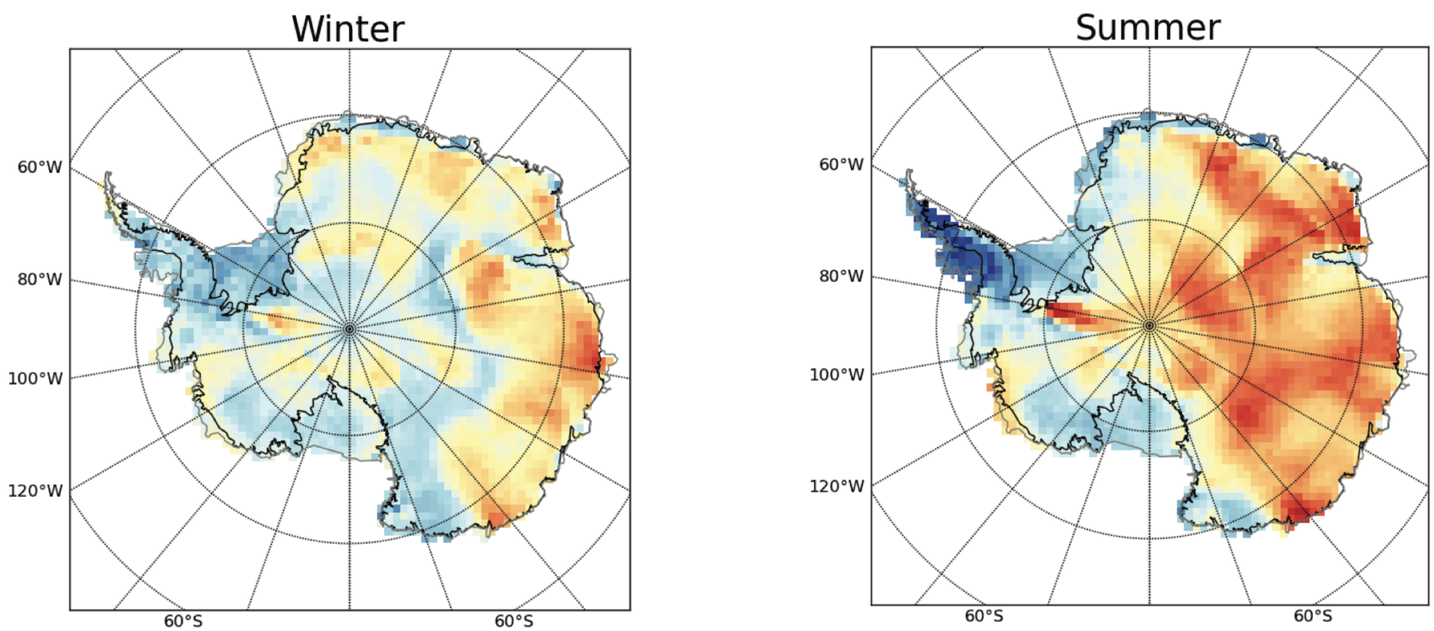
- Penalty on locations that increase correlations
- "covariance trace" favors those locations
- Interpretation benefits from knowing correlations

# Correlation Length Scales

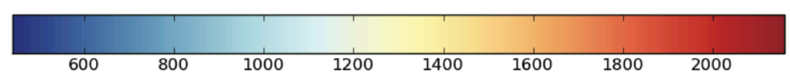



Compare obs and AMPS

# Correlation Length Scales




Winter: April—September
Summer: October—March




Rich patterns. Main signals:

- Longer in summer
- Longer in East Antarctica
- Shortest on Peninsula

# Summary

- Network design provides objective input for
  - New networks (clean slate)
  - Augmenting existing networks
  - Reconfiguring networks
- New extension allows many metrics at once
  - Good for spatial fields; soundings; multiple fields
  - More in next talk
- Current/future work
  - Soundings, and weather forecast improvement
- Please give us feedback on metrics!

Sponsor: National Science Foundation through Award # ANT-1043090