NSF Expeditions in Computing

Outreach Opportunities for Expedition's Data-Driven Discoveries

Fredrick Semazzi

North Carolina State University

fred_semazzi@ncsu.edu

http://climatechange.cs.umn.edu

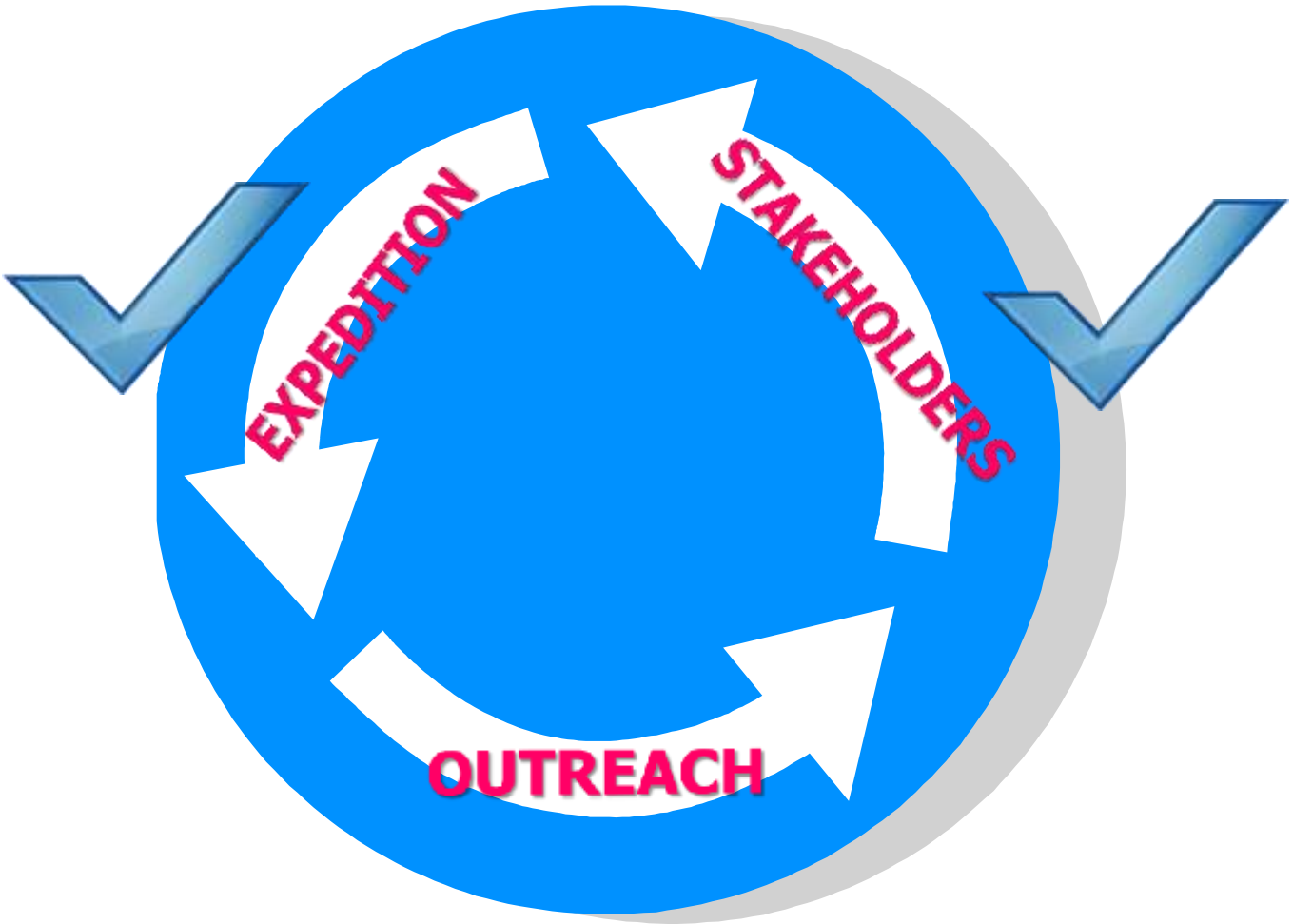
Outline

- Importance of outreach activities
- Framework for outreach interface
- Emerging outreach partnerships & activities
- Examples of opportunities for future expedition outreach activities

Importance of Expedition's Outreach Activities

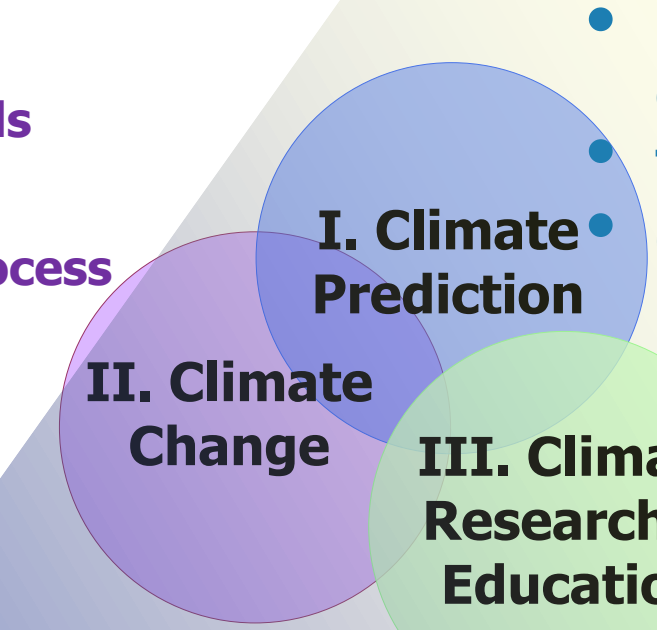
- Expedition project's novel data-driven approaches can contribute to climate science and climate services
- The Expedition outreach strategy is to create bridges and special partnerships with research and operational climate communities to maximize the benefits

Framework for Outreach Interface



Framework for Outreach Interface Based on Use Cases)

- Model validation tools
- Formal Publications
- IPCC assessment process



- Forecasting algorithms; climate algorithms & indices
- Toolkit Portals for RCCs
- Compatible Formats

- Extremes; monsoons; teleconnections; seasonal climate prediction
- Journal publications, reports

NSF Expedition

Expedition Data Driven Discovery
Use (Application) Case: Climate Teleconnections

Computer Science

A graph-based approach to find teleconnections in climate data. Conference on Intelligent Data Understanding (*CIDU) 2011 Vol 6, Issue 3,158 –179, June 2013*

Kawale, J., S. Liess, A. Kumar, M. Steinbach, P. Snyder, V. Kumar, A. R. Ganguly, N. F. Samatova, and F. Semazzi, 2013

Climate Science

ENSO and SAM Interactions over the Tasman Sea: Implications for Regional Climate. To be submitted to Journal of Climate14 August 2013

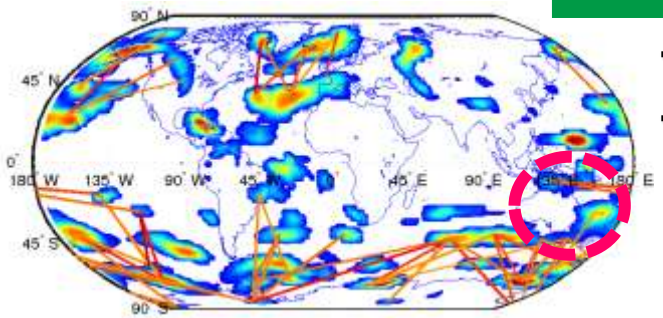
Stefan Liess, Arjun Kumar, Peter K. Snyder, Jaya Kawale, Fredrick Semazzi, Auroop Ganguly, Nagiza Samatova, Vipin Kumar

Use case

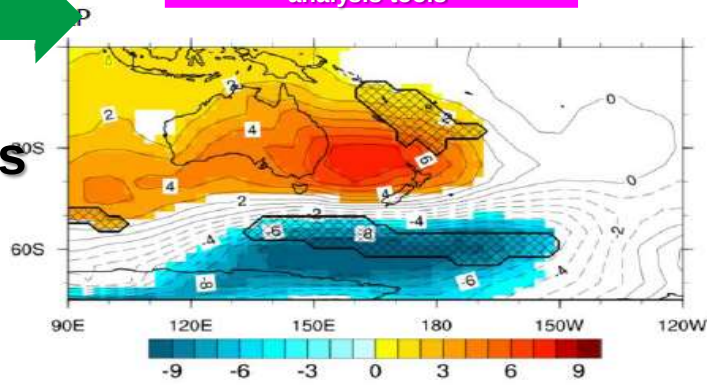
Teleconnections

Shared Reciprocal Nearest Neighbors (SRNN) Method

Eliassen-Palm flux wave propagation (u) momentum flux composite, climate diagnostic analysis tools



Tasman Sea Teleconnections



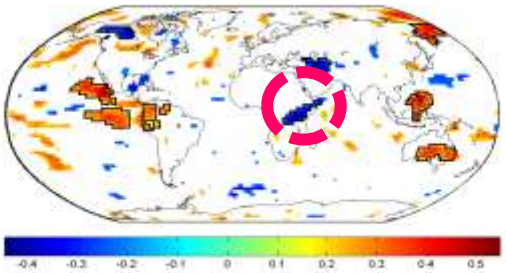
Expedition Data Driven Discovery
Use Case: Atlantic Hurricanes

Computer Science

Forecaster: Forecast Oriented Feature Elimination-based Classification of Adverse Spatio-Temporal Extremes," ACM Eighteenth Conference on Information and Knowledge Management, p. , vol. , (2011).

Zhengzhang Chen, Tatdow Pansombut, William Hendrix, Doel Gonzalez, Frederick Semazzi, Alok Choudhary Vipin Kumar, Anatoli V. Melechko, Nagiza F. Samatova

'Forecaster' Machine Learning Method



Use case

Atlantic Hurricanes

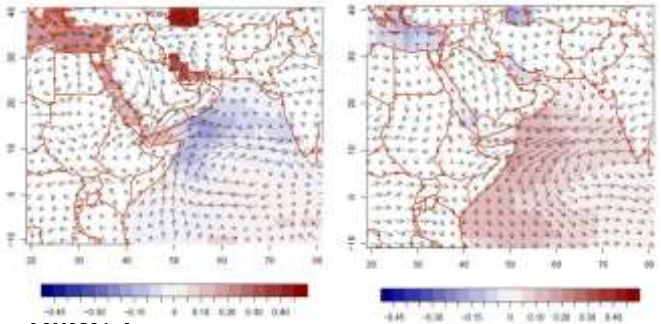
New Atlantic GHA Hurricanes Index

Climate Science

The Role of the Greater Horn of Africa in Modulating Atlantic Hurricane Variability. To be submitted to Journal of PNAS, 14 August 2013

Fredrick H. M. Semazzi, Pascal F. Waniha, Gonzalo A. Bello, Jitendra K. Harlalka, Kara A. Smith, Zhengzhang Chen, Vipin Kumar, and Nagiza F. Samatova

Uniqueness, causality, composite analysis tools



Some Cultural Differences

- Complex climate networks
- Data driven vs hypothesis driven
- Publication challenges

- conferences vs journals philosophy
- review cycle
- algorithm vs physical understanding

Use Case
Seasonal Climate Prediction for Meningitis

I. Climate Prediction

Motivation: Google.org-NCAR-NCSU Meningitis Project

- **Meningitis is a serious infectious disease affecting 21 countries; kills up to 100s of thousands in one year:**
  - **300 million people at risk**
  - **700,000 cases in the past 10 years**
  - **10-50 % case fatality rates**

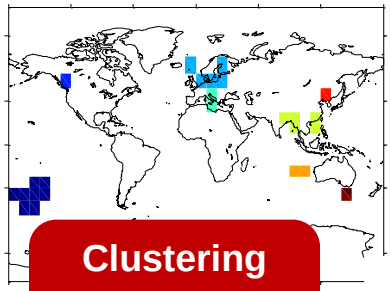
African Meningitis Belt

Methodology for data-driven Discovery of Climate Indices

.

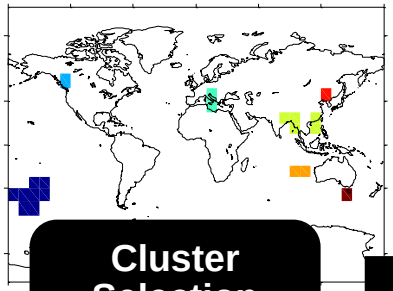
Feature Selection

Select coordinates with a statistically significant correlation with the response for multiple climate variables



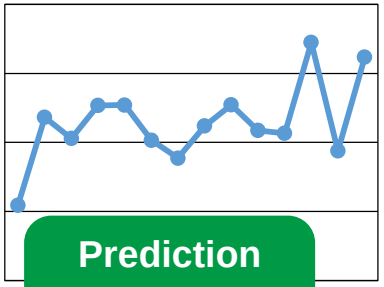
Clustering

Merge selected coordinates with a statistically significant pairwise correlation for multiple climate variables



Cluster Selection

Select clusters for each climate variable with a statistically significant correlation with the response



Prediction

Use selected clusters to predict the response

Use Case: Relative Humidty Seasonal Climate Prediction for Meningitis

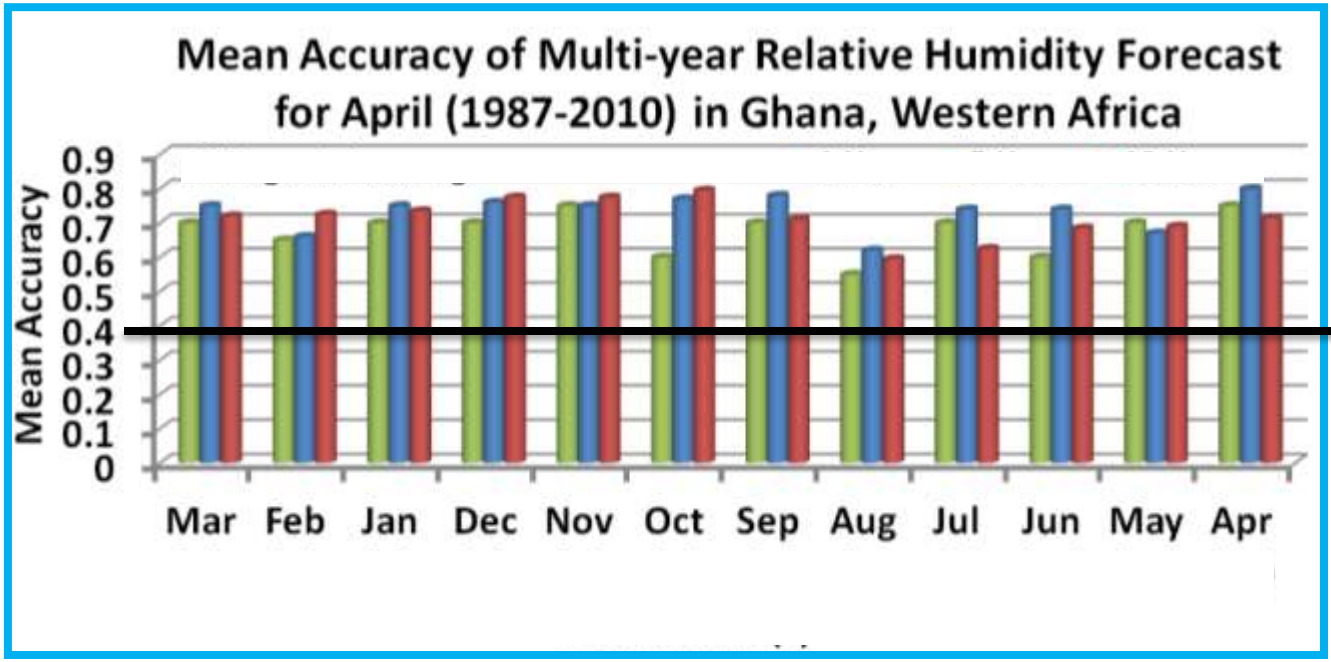
I. Climate Prediction

Application to Climate-Meningitis Problem over West Africa

Applied hierarchical classification and supervised feature selection algorithms to select multivariate features and locations of the predictors to predict sfc RH.

Outreach Opportunity

- Toolkit Portals (e.g., seasonal forecasting algorithms)
- Compatible Formats & Products (e.g., climate indices)

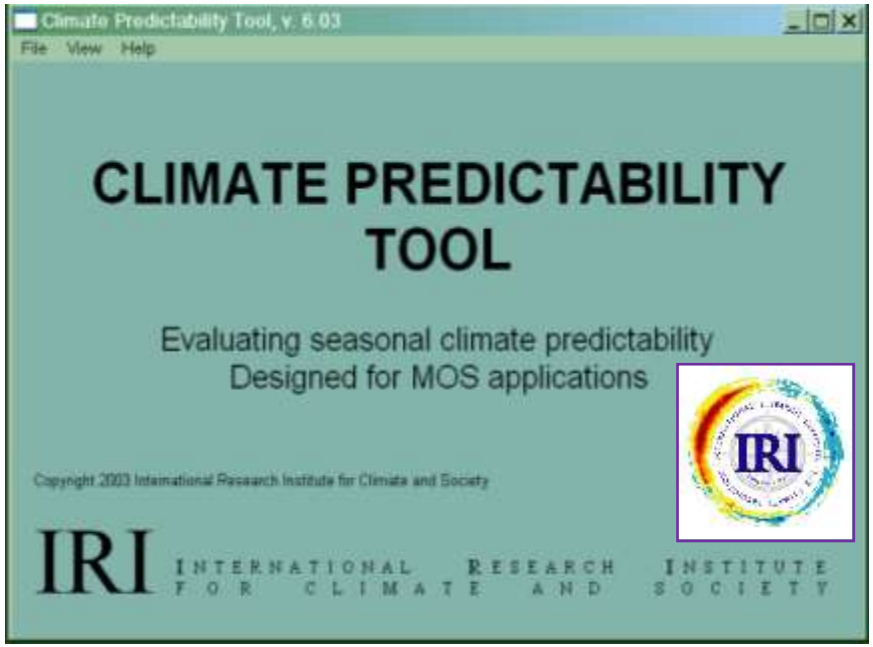


Climotological baseline

Climate Predictability Tool (CPT) is an easy-to-use Windows-based software package for making downscaled seasonal climate forecasts, that was designed specifically for use at WMO RCCs

I. Climate Prediction

Specifically, CPT is designed to produce statistical forecasts of seasonal climate using either the output from a GCM, or fields of sea-surface temperatures.

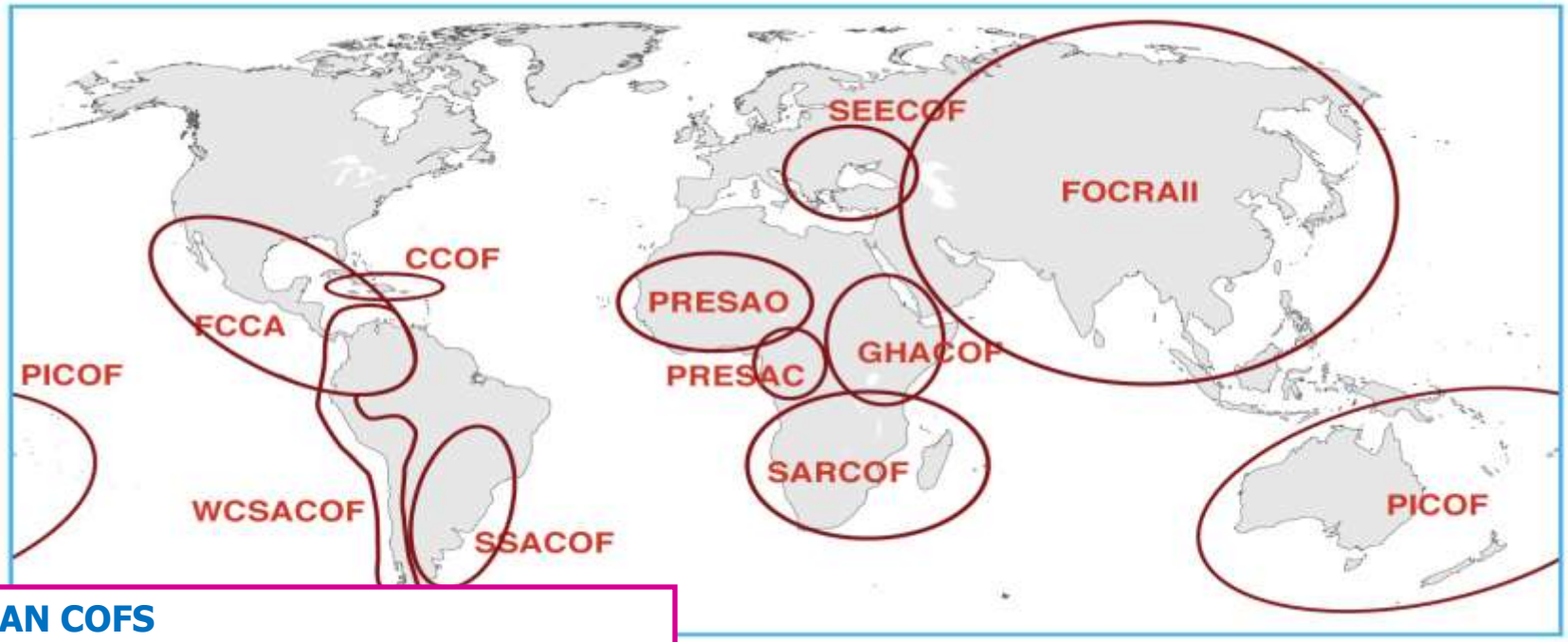


- Predictions are based on canonical correlation analysis.
- The user provides their own predictand data in the form of seasonal data in ASCII format.
- The predictors are SST fields or GCM outputs, although indices can be used.
- Rigorous cross-validation procedures are enforced, and extensive verification information is provided.

CPT will be upgraded using Expedition project downscaling methods (greater skill) instead of canonical correlation analysis.

WMO Regional Climate Centers (RCCs) Serve *billions of people.*

I. Climate Prediction



AFRICAN COFS

1. **GHACOF: Greater Horn of Africa COF**
2. **SARCOF: Southern Africa COF**
3. **PRESAO: Western Africa**
4. **COFPRESAC: Central Africa CO**

REST OF THE THE WORLD

1. **FOCRAII: Forum on regional Climate monitoring, assessment and prediction for Regional Association II (Asia)**
2. **SSACOF: Southeast of South America COF**
3. **WCSACOF: Western Coast of South America COF**
4. **CCOF: Caribbean COF**
5. **FCCA: Foro Regional del Clima de América Central**
6. **PICOF: Pacific Islands COF**
7. **SEECOF: SouthEastern Europe COF**

Climate Change Community Interface

II. Climate Change

Outreach Opportunities

- Model validation tools
- Formal Publications
- IPCC assessment process

Climate Research Community Interface
WCRP Grand Challenges

III. Climate Research & Education

1. **Prediction of extremes**
2. **Decadal predictions**
3. **Availability of fresh water in a changing climate**
4. **Seasonal prediction of tropical cyclone landfall**
5. **Prediction of Monsoons**

Basic Climate Research: Hurricane Use Case

III. Climate Research & Education

Complex Networks vs Traditional Method

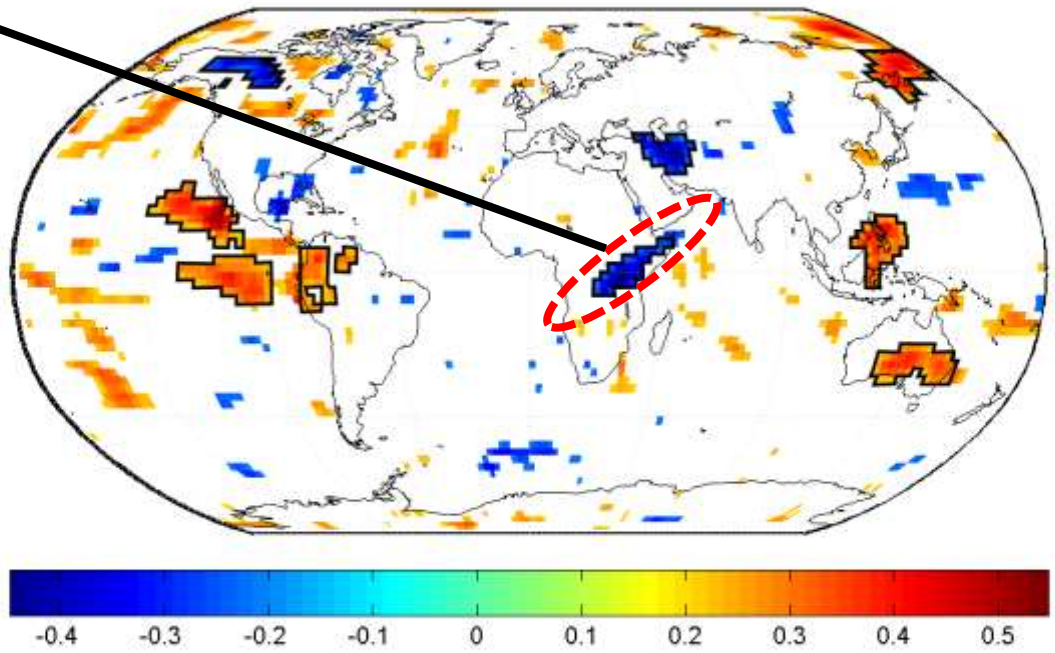
FORECASTER Performance on North Atlantic Hurricanes

Metric	FORECASTER (Expedition)	Webster (Traditional)	Random Forest	Bagging	Boosting
Accuracy (%)	90.3	65.5	76.7	73.3	75.0
HSS	0.85	0.49	0.65	0.60	0.62
PSS	0.85	0.50	0.65	0.63	0.63
GSS	0.84	0.68	0.65	0.67	0.66

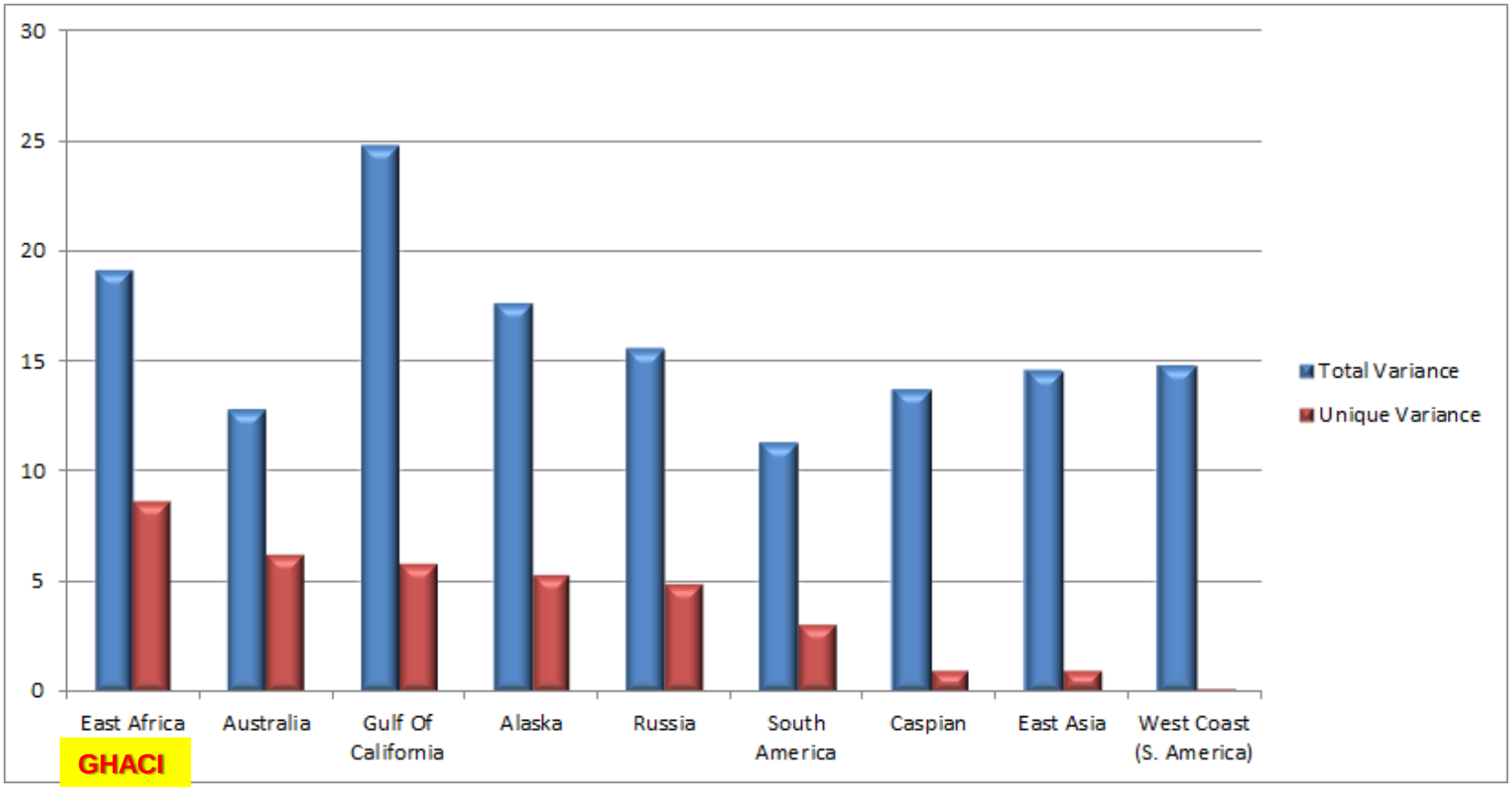
Note: HSS: Heidke score, measures how well relative to a randomly selected forecast; PSS: Peirce score, difference between the hit rate and the false alarm rate; GSS: Gerrity score, occurrences substantially less frequent. [18]

Global heat-map of Pearson Correlation between Atlantic Hurricane Count (HC) and Greater Horn of Africa Climate Index (GHACI)

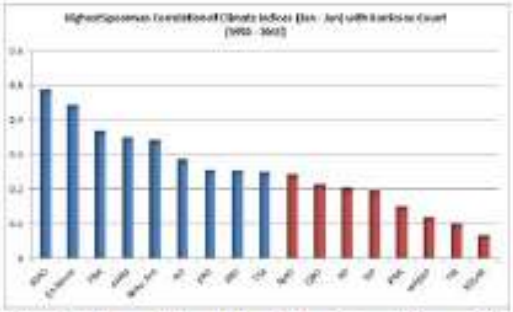
GHACI



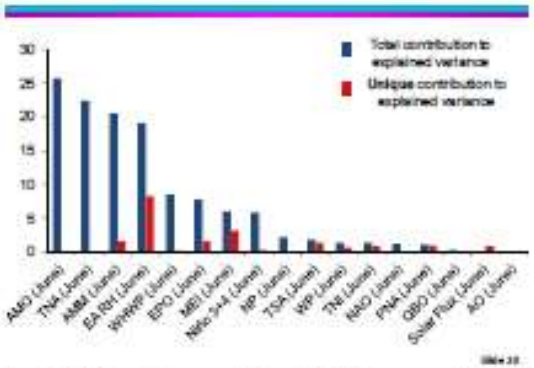
Contribution to explained Variance of Atlantic Hurricane Count (1950-2010) by the Hotspots



New Greater Horn of Africa Index



Correlation of traditional leading CIs and GHACI with the Atlantic Hurricane count (highest from January through June).

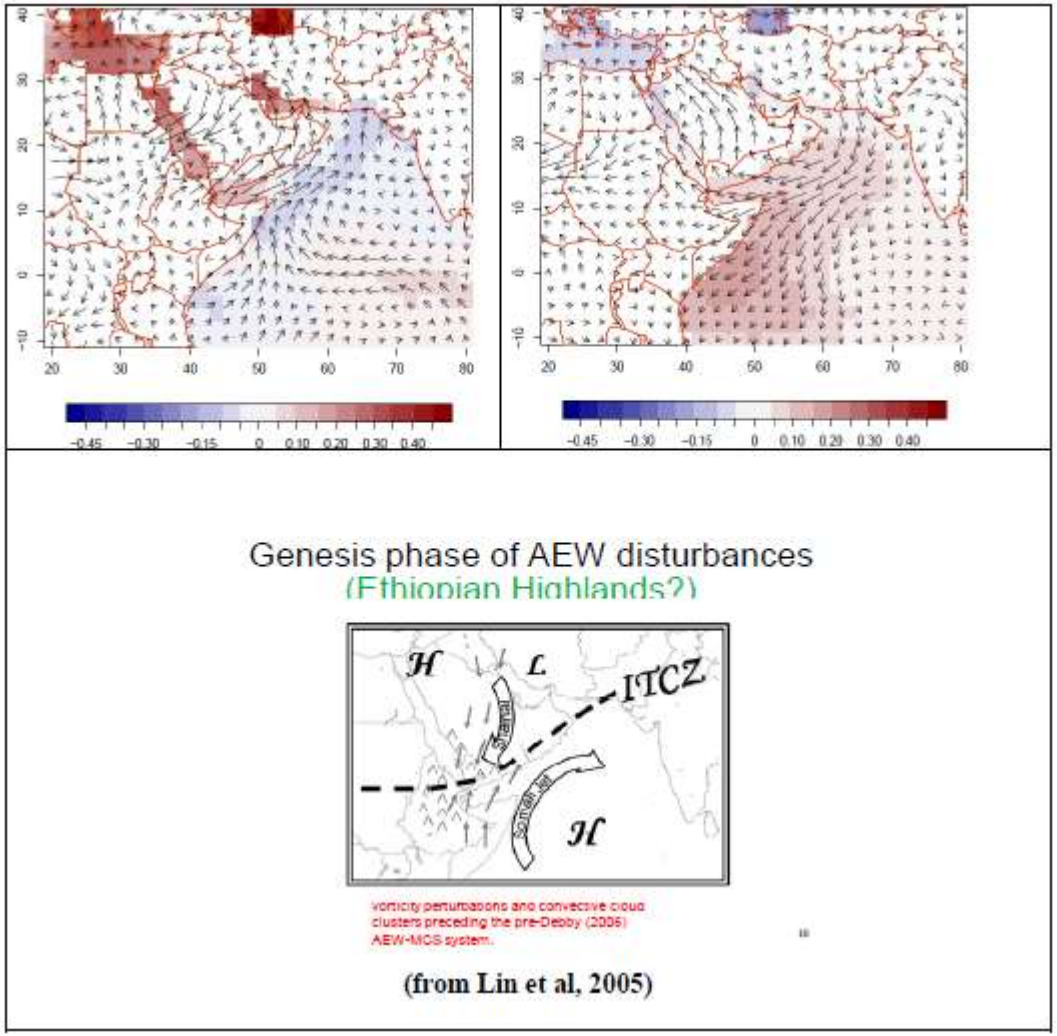


Contribution to explained Variance of Atlantic Hurricane Counts (1950-2010) the traditional climate indices for June. Results for April and May in addition to June are displayed in Table 1.

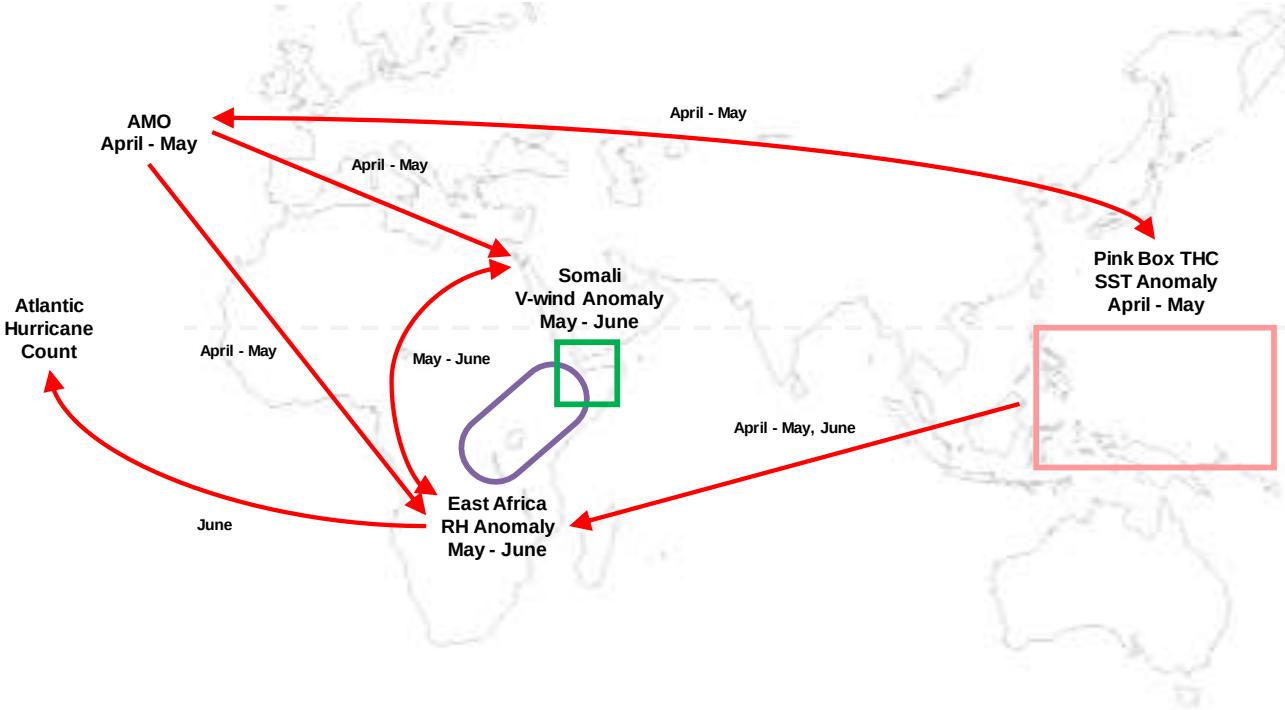
Climate Index	Contribution to explained variance of Atlantic hurricane count					
	Total contribution			Unique contribution		
	April	May	June	April	May	June
AMO	20.98	23.89	25.64	0.15	0.21	0.04
TNA	19.38	22.67	22.37	1.18	0.77	0.01
AMM	11.74	17.05	20.44	3.26	2.33	1.58
EA RH	16.23	12.96	19.08	6.57	6.38	8.29
WHWP	1.46	4.23	8.53	0.05	5.07	0.22
EPO	0.13	0.08	7.78	0.12	0.36	1.59
MEI	0.09	1.4	6.02	0.05	0.42	3.28
Niño 3+4	0.6	3.18	5.87	1.71	3.02	0.31
NP	2.68	2.52	2.23	0.42	2.88	0.04
TSA	5	1.21	1.82	3.95	0.81	1.43
WP	3.08	1.06	1.37	2.22	0.08	0.64
TNI	0.99	1.24	1.36	0.05	0.77	0.81
NAO	0.28	8.47	1.29	0.1	3.57	0.04
PNA	4.26	0.02	1.12	0.09	0.19	0.93
QBO	1.06	0.11	0.32	1.7	0.41	0
Solar Flux	0.04	0.01	0.07	0.14	0.05	0.84
AO	2.66	2.73	0.01	0.21	1.31	0

Total and unique contribution of CI index including the GHACI variability to HC variance for April, May and June.

Composites of the winds at 1.5km and SST anomalies for above normal HC (left) and below normal HC (right); Lin et al (2005) model for the contribution of the East African Low Level Jet (EALLJ) and the Shamal winds on one of the the genesis regions cyclongenesis of the African Easterly Waves (AEWs)

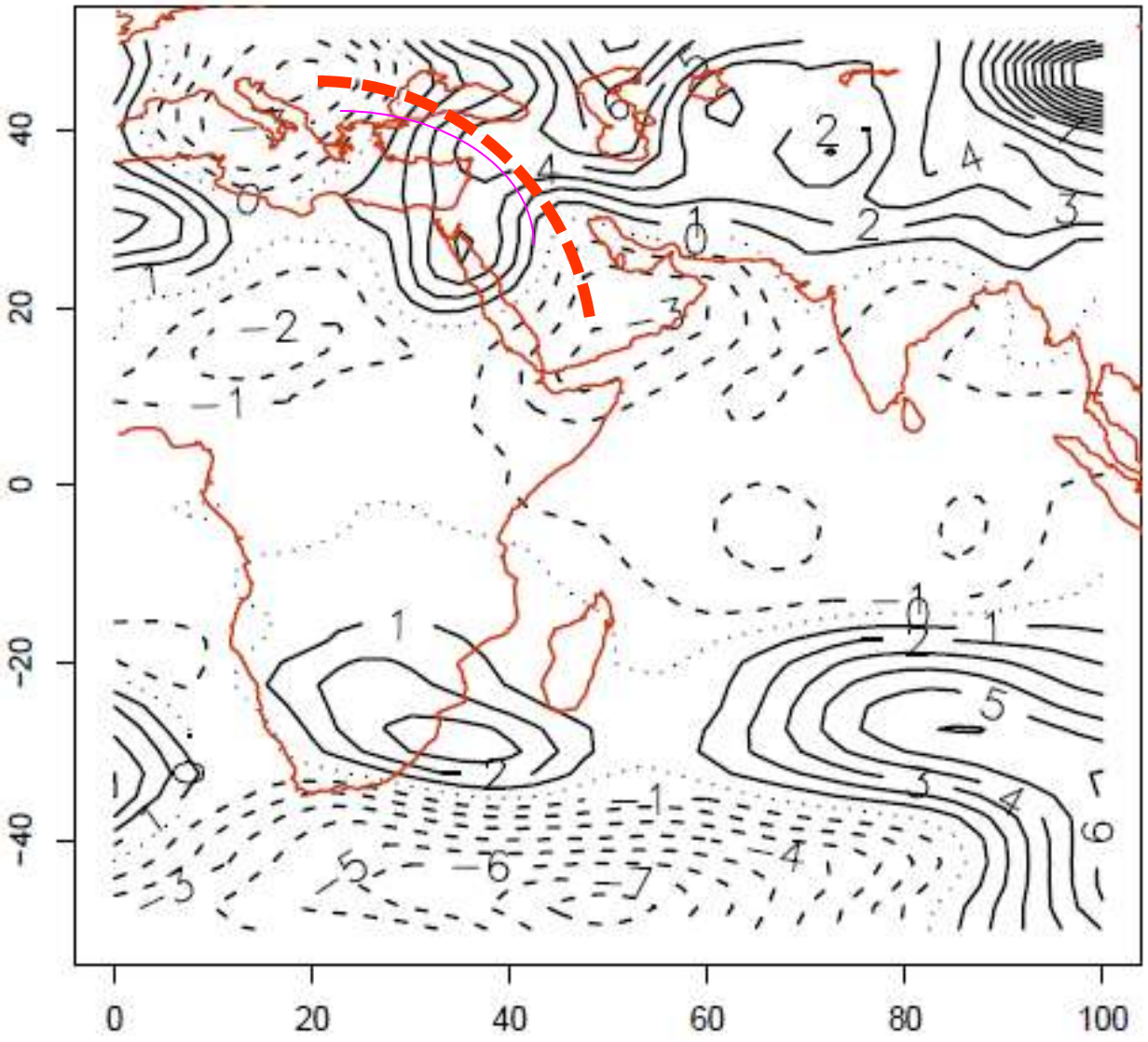


Moderation Pathways Network

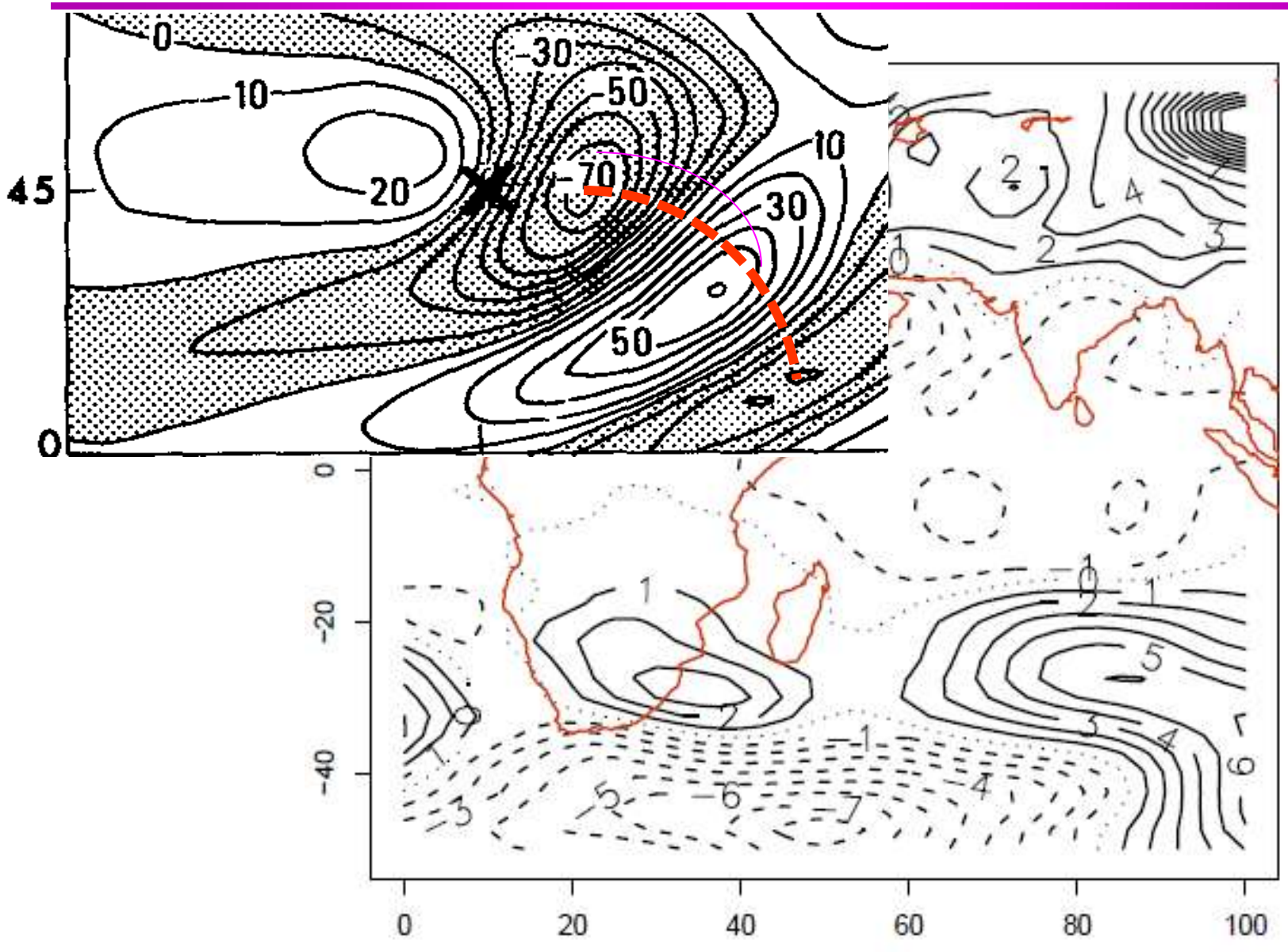


* Using beta coefficients and part correlations to estimate statistical significance with p=0.1

Alps Generated Stationary Rossby Wave Train in the Westerlies



Jan 25, 2012


0
10
20
30
50
70
10
30
50
45
0
2
4
1
0
2
3
1
0
1
2
5
3
4
6
1
0
2
1
3
5
6
7
4
1
6
0
-20
-40
0
20
40
60
80
100

Climate Education Community Interface

III. Climate Research & Education

- **Expedition's discoveries being entrained into educational programs and courses at the project's participating universities**

- **Example: North Carolina State University's new Climate Change & Society Masters degree program (climate-psm.meas.ncsu.edu)**
  - Its curriculum trains students to integrate climate information and methods into the risk management for climate sensitive social-economic sectors.

Examples of Future Expedition Outreach Activities

Outreach Partner	Type of Enabled Outreach
National Centre for Atmospheric Research (NCAR)	**NSF Expedition project's rainfall and humidity prediction for the NCAR/Google project on the vaccination of meningitis epidemics that severely impacts up to *250,000 people* annually; next frontier *data analytics and reduction involving massive PB HPC problems.***
United Nations (UN) World Meteorological Organization (WMO)	**Operationalization of NSF Expedition project experimental seasonal climate prediction methodology for WMO *African Centre of Meteorological Applications for Development* (ACMAD) and IGAD Climate Prediction and Application Centre (ICPAC) (RCCs)**
National Hurricane Centre (NHC)	**Operationalization of NSF Expedition project experimental hurricane prediction (*prospects for lead time greater than ten days before landfall*) methodology for NHC to reduce vulnerability of US coastal population; efficient *hurricane intensity* estimation methods -*50% improvement.***

Examples of Future Expedition Outreach Activities (Cont.)

Outreach Partner	Type of Enabled Outreach
National Climate Data Centre (NCDC)	Creation of *new climate indices* and introduction of NSF Expedition data driven methodologies into the NCDC's analysis tool-kit of the highly multi-dimensional and most voluminous open source global climate data archive; highly efficient graphical methods for *abrupt climate detection*
World Climate Research Program (WCRP)	Fundamental research on *understanding of casual pathways of the Sahelian climate predictability and causality, change detection of Atlantic hurricanes*
Intergovernmental Panel on Climate Change (IPCC)	Trends in rainfall *extremes* over India during last half-century; causality of the Sahelian climate variability; change detection of Atlantic hurricanes.
National Oceanic and Atmospheric Administration (NOAA)	NCAT's Objective *Tropical Cyclone Intensity* Estimation using Analogs of Spatial Features in Satellite Data

Perhaps we can put a smile on this man's face

Thank You