Predicting Ecological Responses to Climate Variability with a Dynamic Bayesian Network Model

Neda Trifonova, Mandy Karnauskas and Chris Kelble

RSMAS

CIMAS

NOAA

Motivation

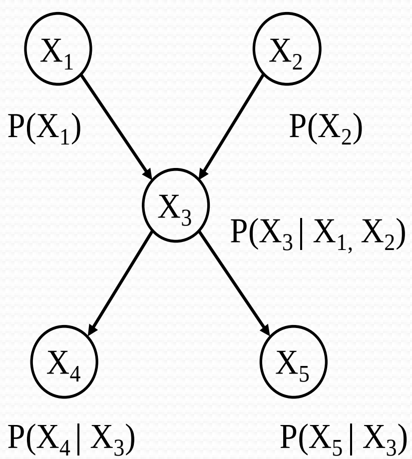
- Gulf of Mexico is an ecologically and economically important dynamic ecosystem
- Interactions with natural and anthropogenic factors

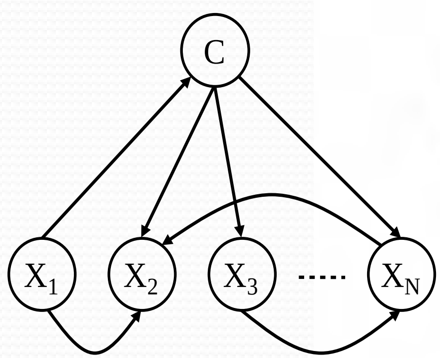
- Application of explorative, data-driven techniques
- Evaluation and implementation

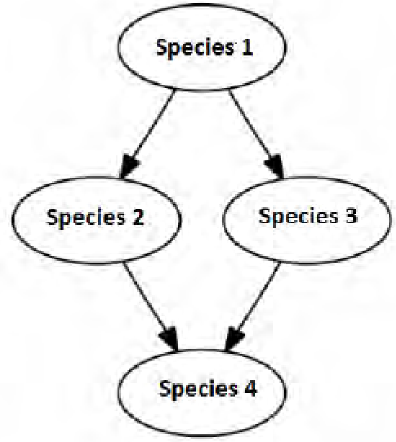
- Potential response of the system to pressure
- Sustainability and management

Bayesian Networks

- Describes the joint distribution over a set of variables, $X_1 \ldots X_N$ by exploiting conditional relationships represented by a graph
- Strength of the relationships shown by a conditional probability table
- Inference algorithms to ask 'What if?' questions







(a) DAG

Sp2	Sp3	p(Sp4=F)	p(Sp4=T)
F	F	0.9	0.1
T	F	0.2	0.8
F	T	0.2	0.8
T	T	0.1	0.9

(b) Conditional probabilities for Species 4

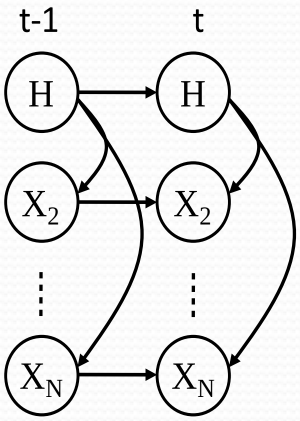
Bayesian Networks for Classification & Feature Selection & Forecasting

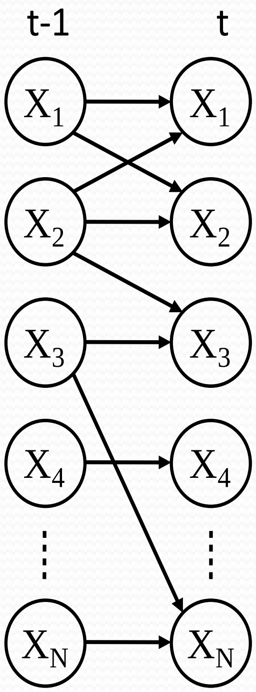
- Nodes that can represents class labels or variables at "points in time"

- Also hidden variables via EM

- Inter and Intra slice connections

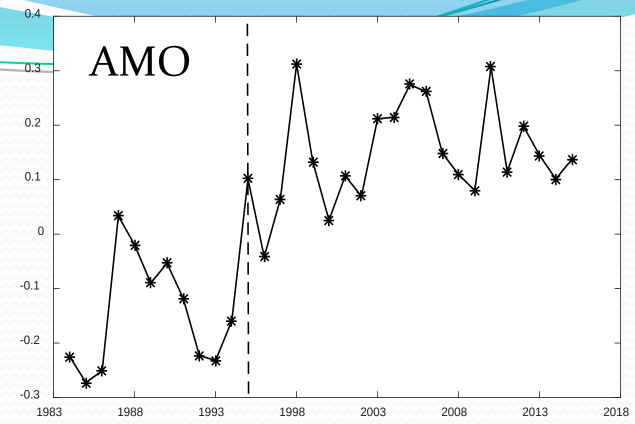
- Predict future observations given all the observations up to the present time: $y_{1:t} = (y_1,\ldots,y_t)$

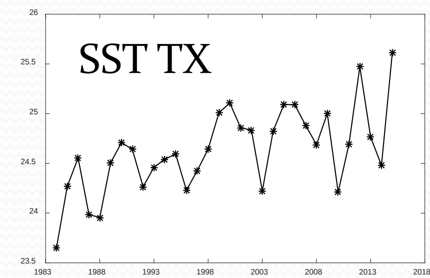


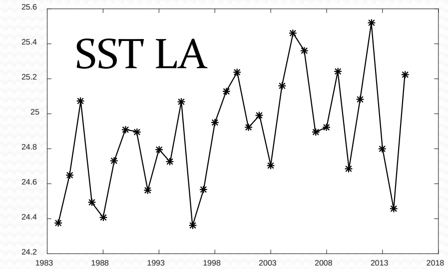


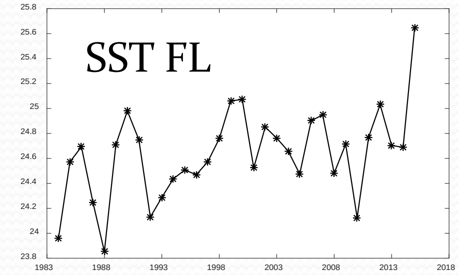
Data

- Temporal data: 1984-2015
- Climate drivers: AMO and SST
- Physical pressures: Hypoxia
- Primary productivity
- Spring and fall zooplankton
- Shrimp recruitment estimates
- Fish recruitment deviations



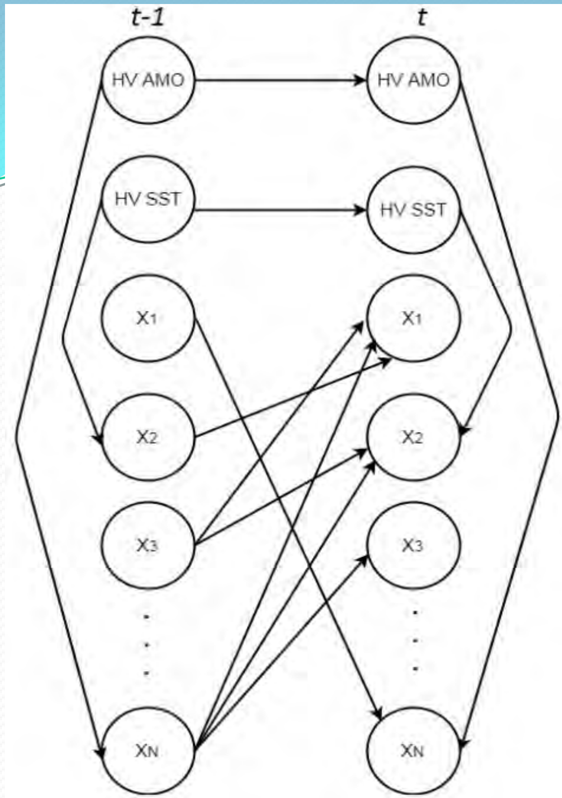


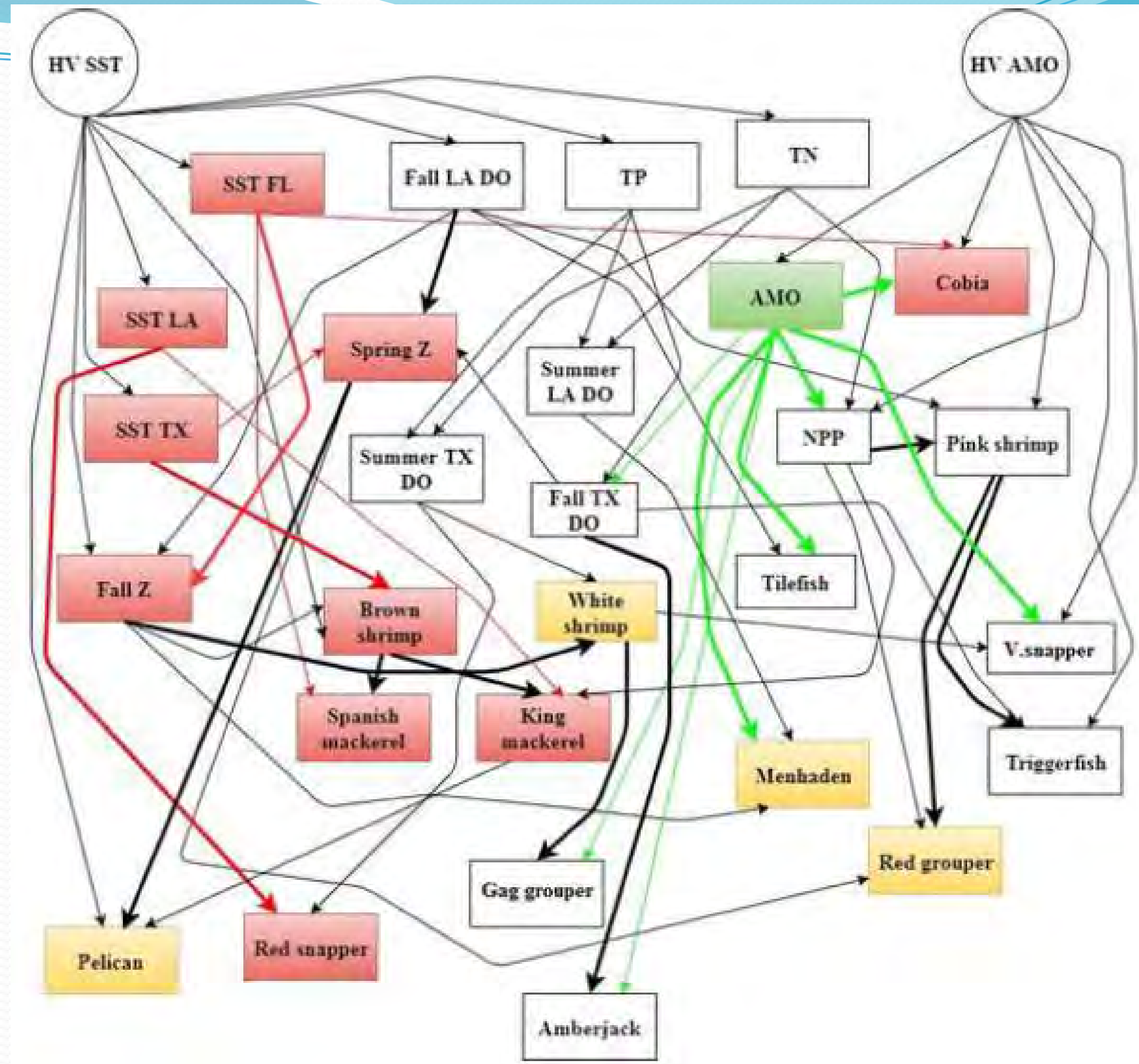




Learning Bayesian Networks

- Hill-climb optimization technique
- The learned BN links represent dependence, these are relationships that are predictive in an informative, not causal aspect
- The Bayesian Information Criterion was used for scoring candidate networks: BIC= $\log P(\Theta) + \log P(\Theta|D) - 0.5k \log(n)$

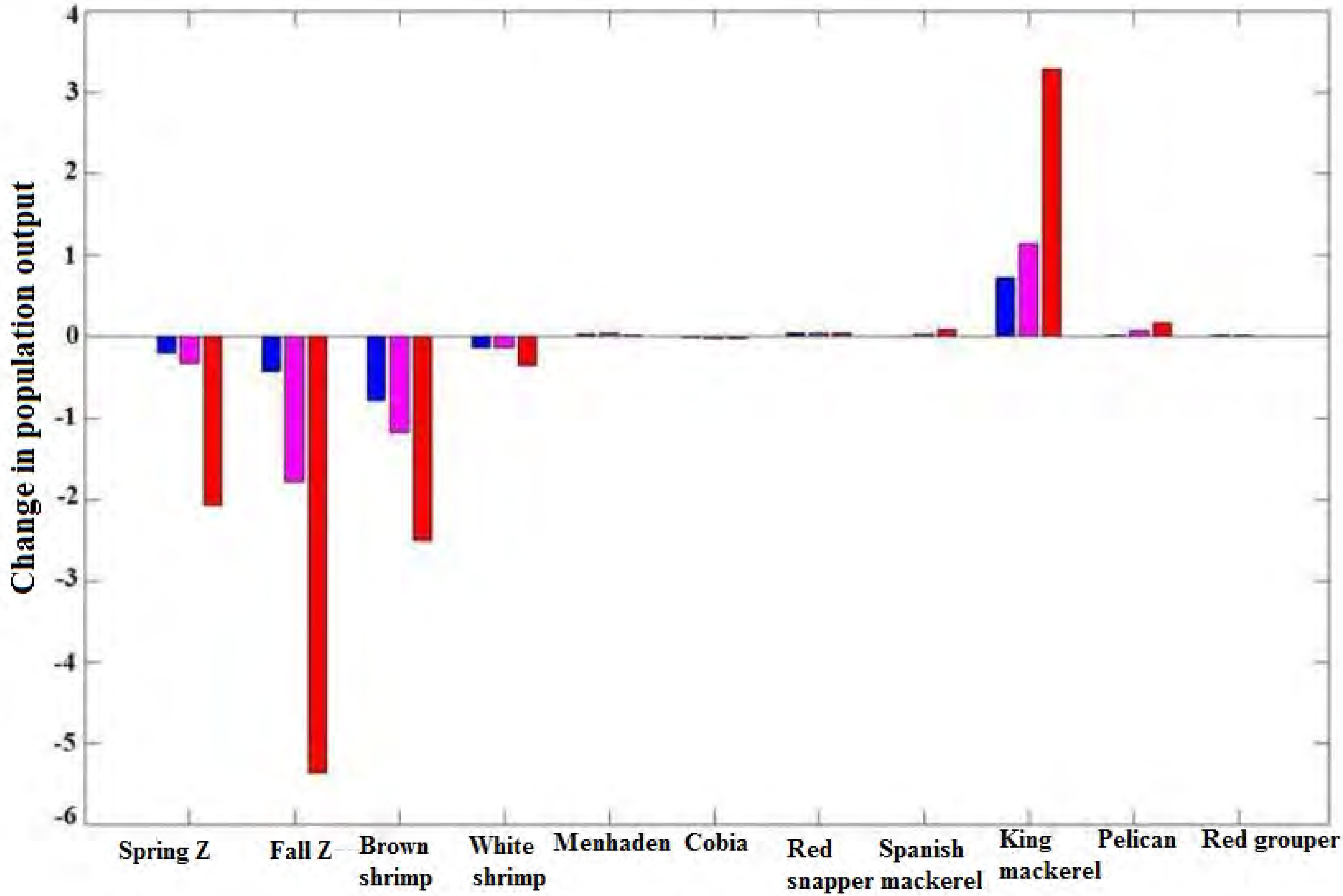




- *Data-driven dynamic BN*
- Nodes- ecosystem states
- Links- potential interactions
- Multiple associations and their changes over time

SST Scenarios and Generating Predictions

- *Baseline* model *vs* SST scenarios: 1.0°C, 1.5°C and 3.0°C
- Given a graphical structure, BNs naturally perform prediction using inference
- X[t] where X = $X_1 \dots X_n$ are the n variables observed along time t
- Non-parametric bootstrap (re-sampling with replacement from the training set) was applied 250 times
- The hidden variables were parameterised using the EM algorithm

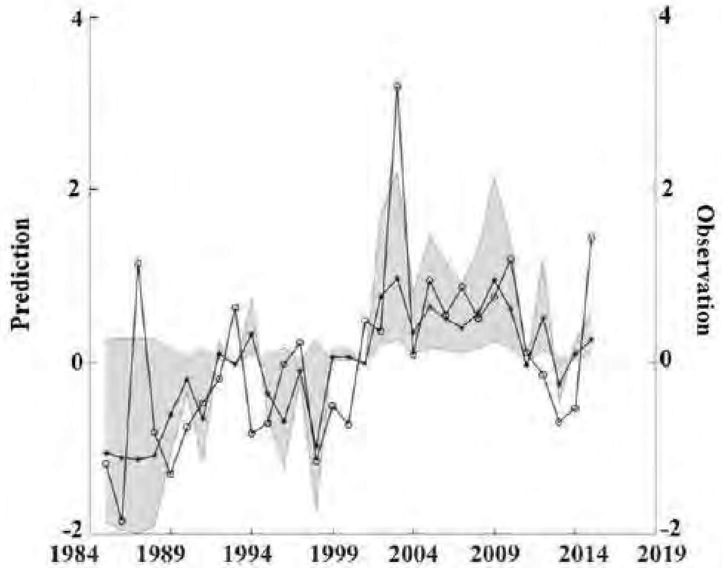


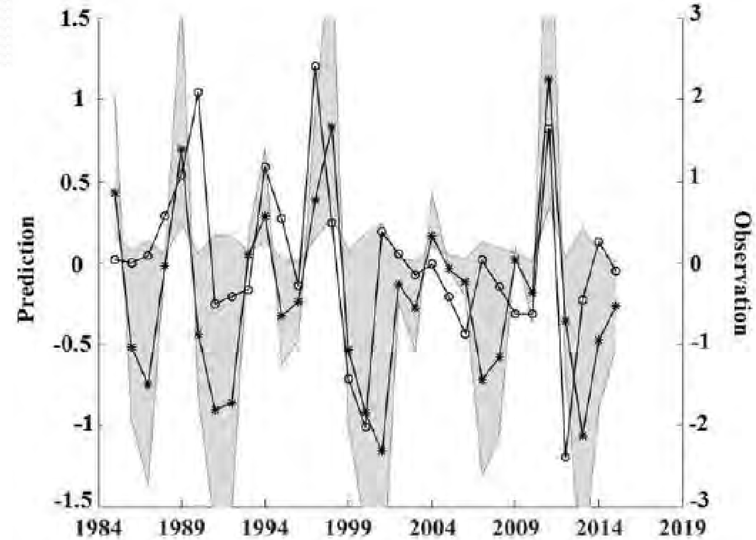
Change in population output
4
3
2
1
0
-1
-2
-3
-4
-5
-6
Spring Z
Fall Z
Brown shrimp
White shrimp
Menhaden
Cobia
Red snapper
Spanish mackerel
King mackerel
Pelican
Red grouper

1.0°C *1.5°C* *3.0°C*

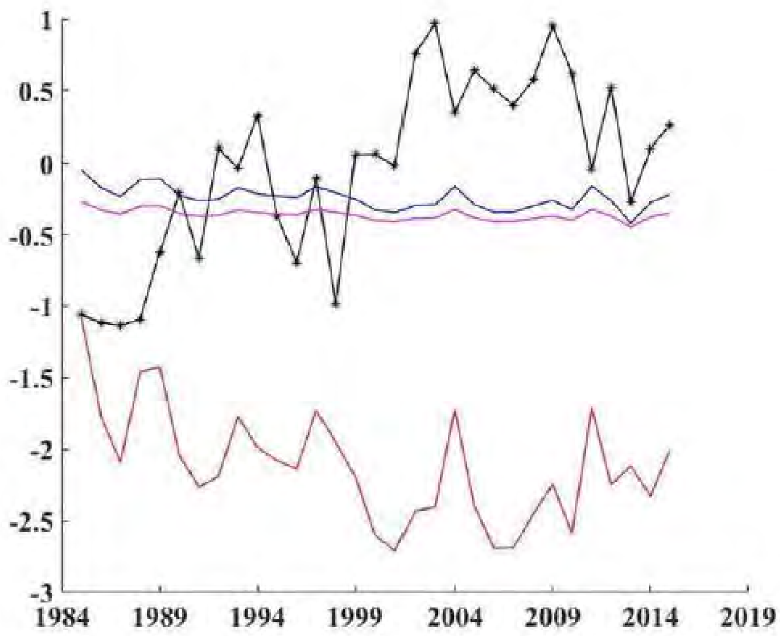
Spring Zooplankton

Fall Zooplankton





Baseline model (*) vs
Original data (○)

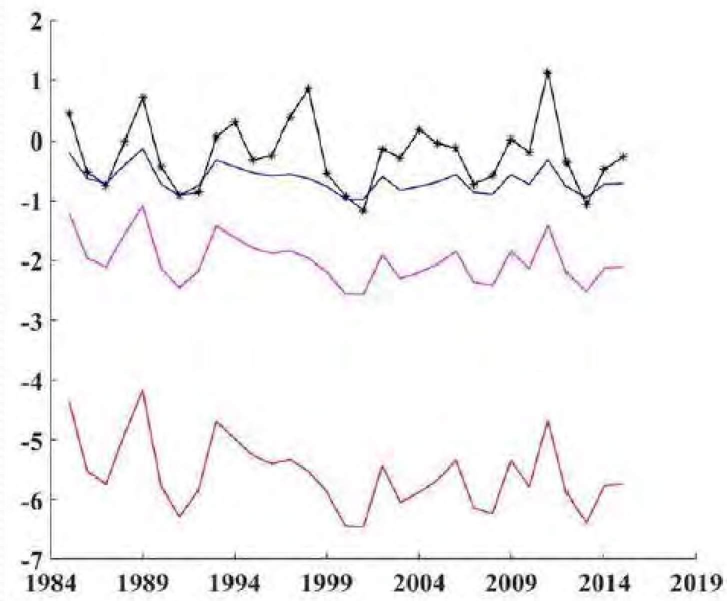


Baseline model (*) vs
SST scenarios

1.0°C

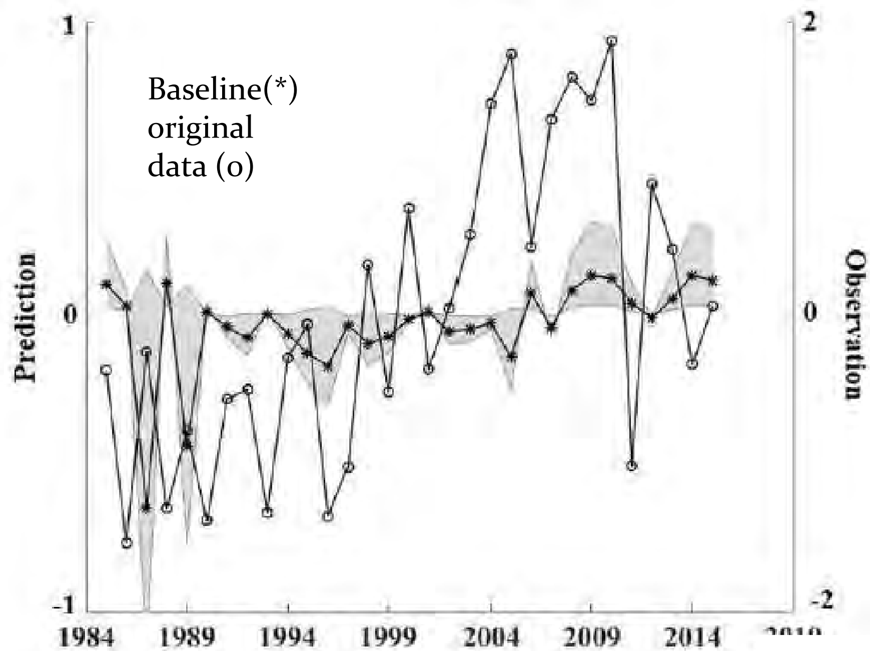
1.5°C

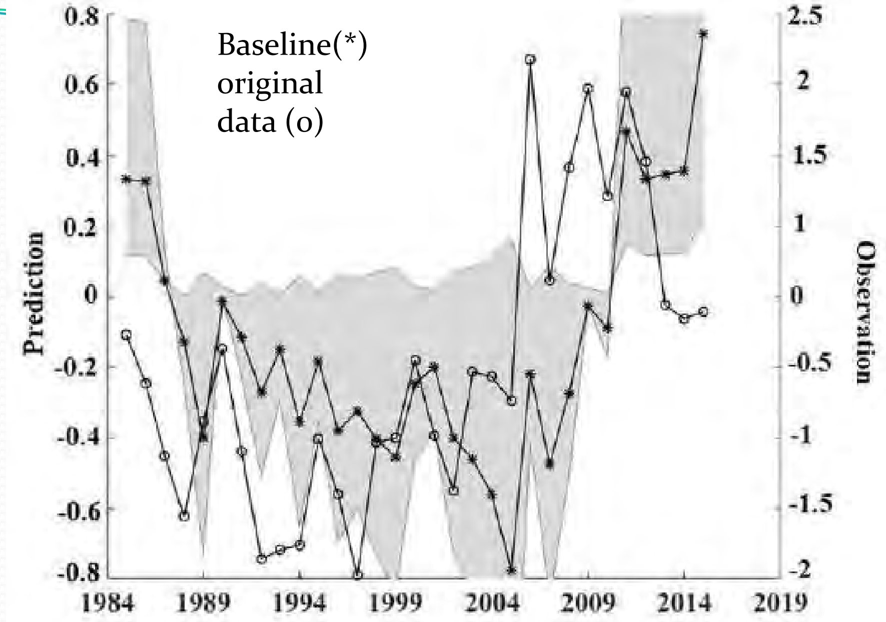
3.0°C

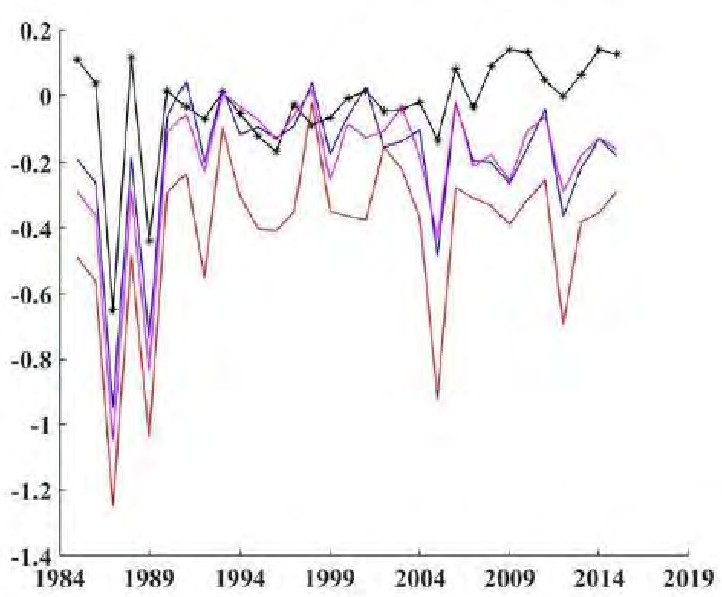


White Shrimp

Brown shrimp



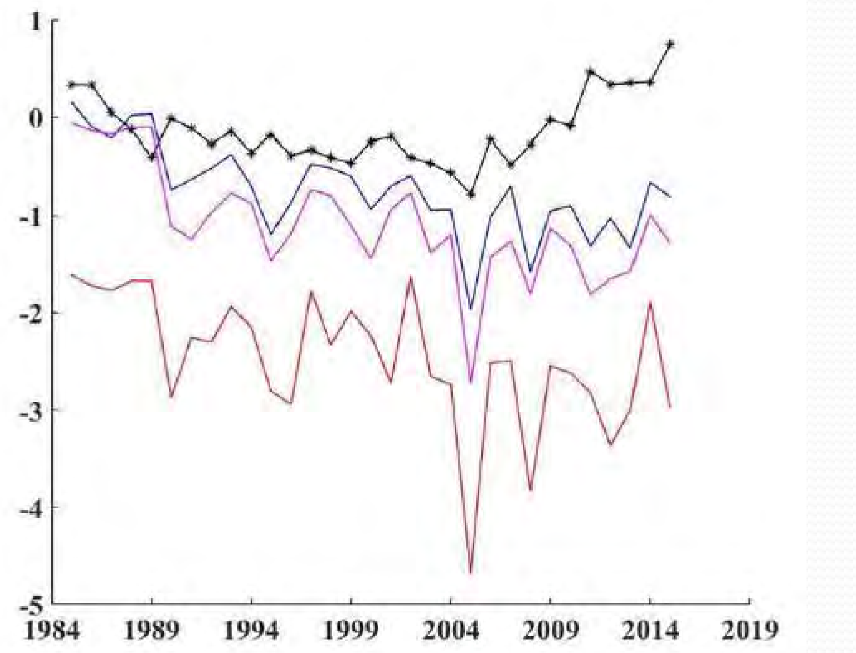




1.0°C

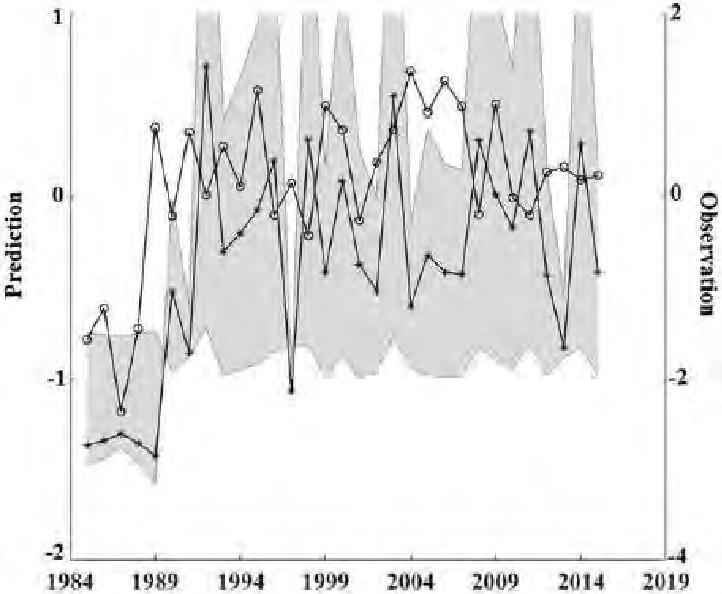
1.5°C

3.0°C

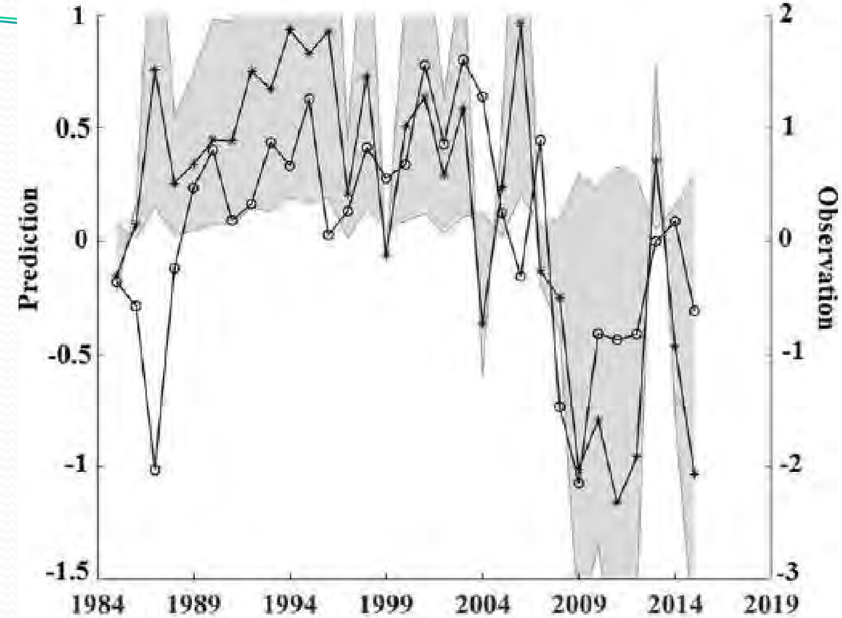


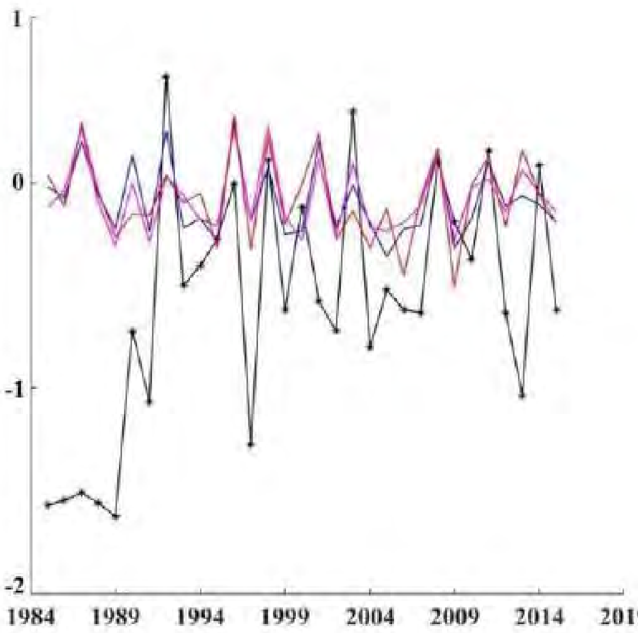
Red snapper

King mackerel



Baseline(*)
original
data (o)

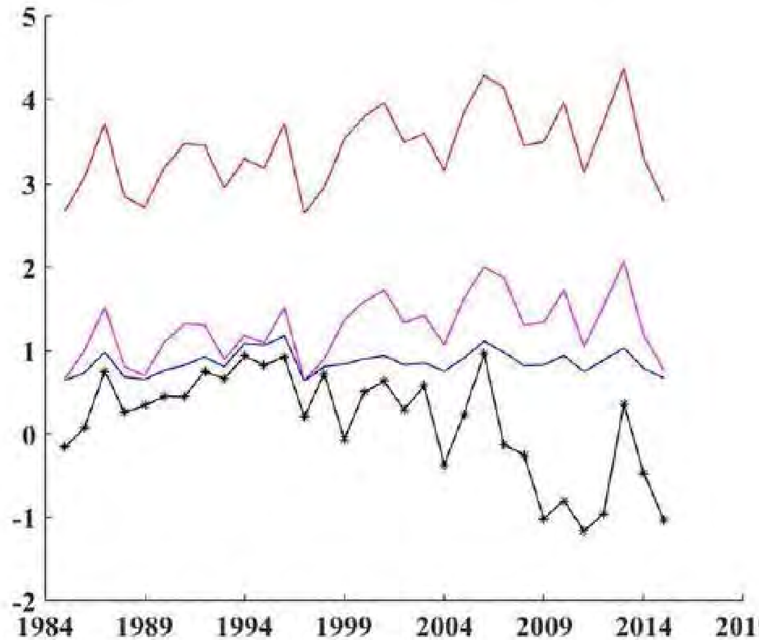




1.0°C

1.5°C

3.0°C



Summary

- An approach that accounts for multiple physical and biological associations and their changes over time
- Variability in ecosystem components to changes in climate
- The data-driven approach provides contrast to other climate prediction methods that are predicated on assumed climate-fish relationships (e.g. NMFS climate vulnerability analysis)
- Network could easily be expanded to include other components of ecosystem (e.g., protected species)
- Relationships are not causal, but model outputs are groundwork for new hypotheses that can be tested

Acknowledgments

- Gulf of Mexico Integrated Ecosystem Assessment Program
- Cooperative Institute for Marine and Atmospheric Studies, University of Miami