

Causal Discovery Techniques for Studying Arctic-Midlatitude Connections

Elizabeth Barnes
Colorado State University

Imme Ebert-Uphoff
Colorado State University



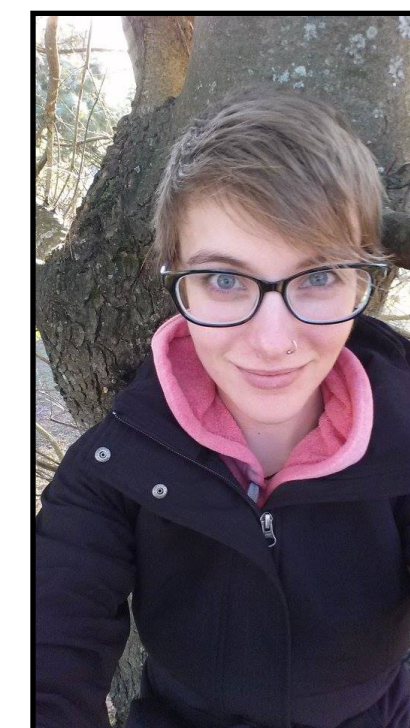
Savini Samarasinghe
Colorado State University



Marie McGraw
Colorado State University



Isla Simpson
NCAR



Bryn Ronalds
Colorado State University

Another tool for our toolbox

Lagged Regression/Correlations

- highly biased by autocorrelation
- does not provide direction of causality
- indirect connections via a 3rd actor can complicate things
- correlation \neq causation

Targeted Model Experiments

- only quantify effects in isolation
- does not allow for feedbacks
- requires the model adequately simulates relevant processes
- results can differ across models

Forecasting Approach

- pathways may not be easily pulled-apart
- requires the model adequately simulates relevant processes
- results can differ across models

Another tool for our toolbox

Lagged Regression/Correlations

- highly biased by autocorrelation
- does not provide direction of causality
- indirect connections via a 3rd actor can complicate things
- correlation \neq causation

Targeted Model Experiments

- only quantify effects in isolation
- does not allow for feedbacks
- requires the model adequately simulates relevant processes
- results can differ across models

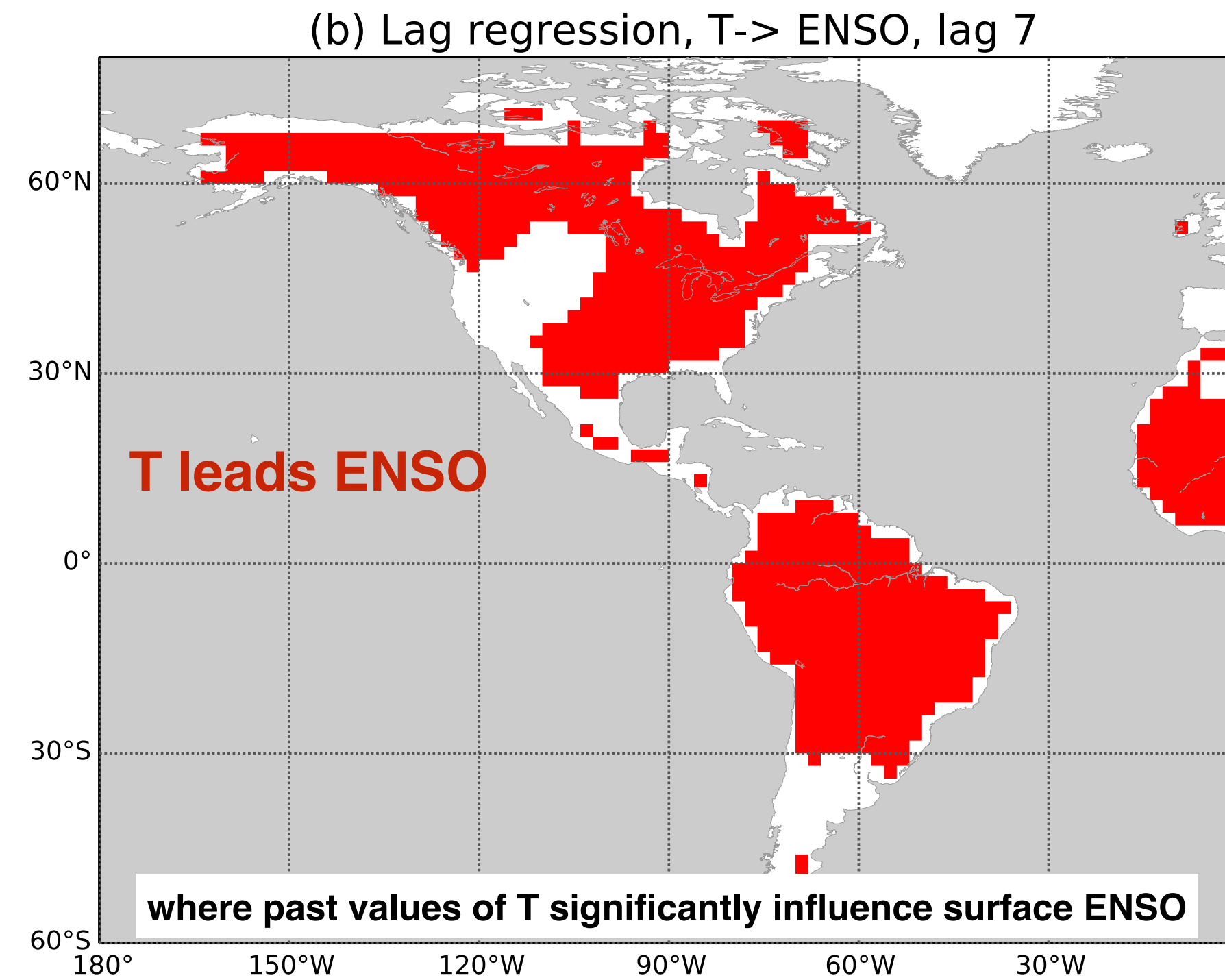
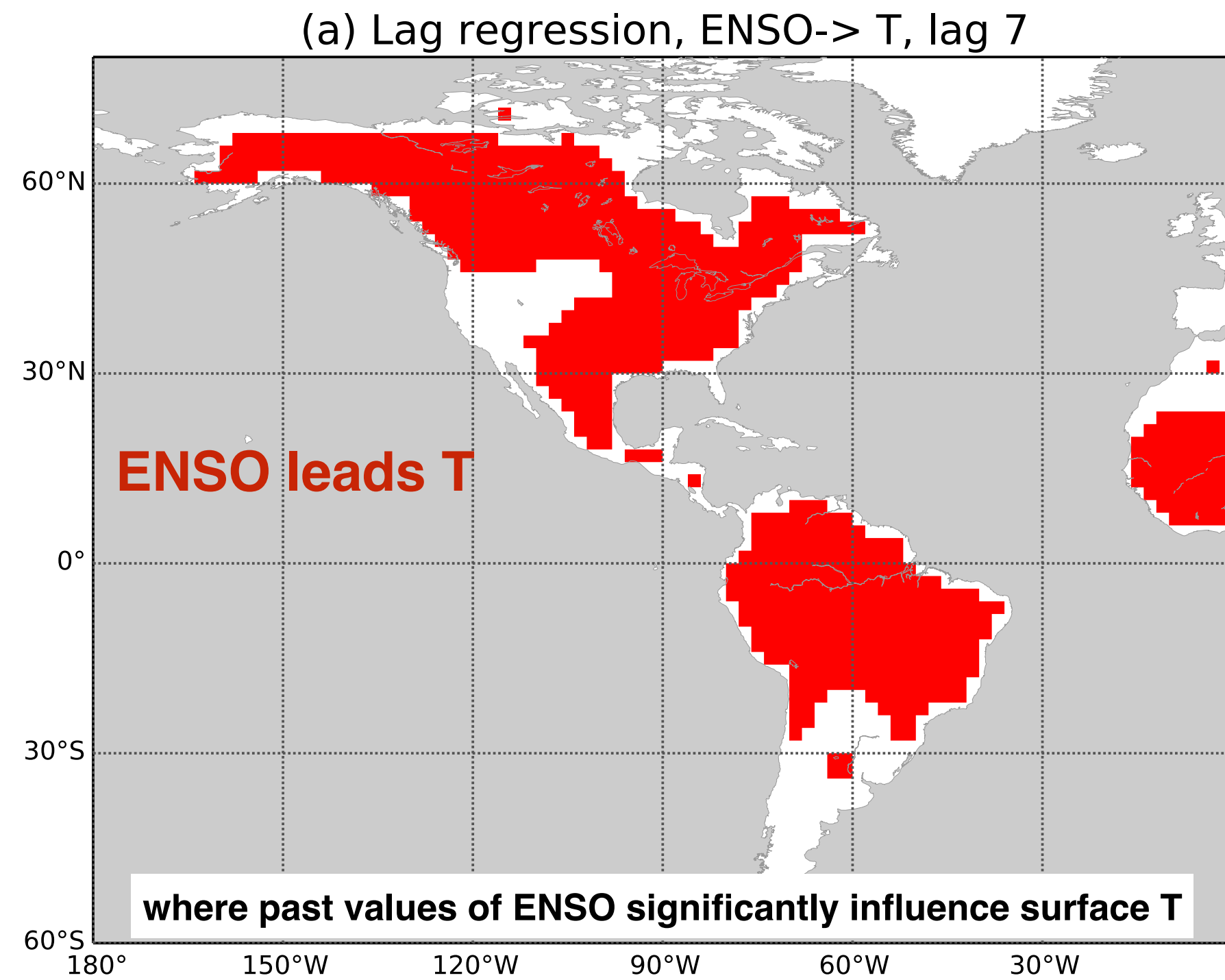
Forecasting Approach

- pathways may not be easily pulled-apart
- requires the model adequately simulates relevant processes
- results can differ across models

Causal Discovery Techniques

- ✓ uses rigorous definitions of causality (although there are different definitions)
- ✓ allows for feedbacks
- ✓ can compare connection strengths
- ✓ can analyze the observations and model simulations identically
- ✓ provides direction of connections
- ✓ ties to forecasting/prediction

Lagged regressions can get you in trouble



McGraw and Barnes (submitted)

Granger Causality



Clive Granger (1934-2009)

2003 Nobel Prize in Economic Sciences

*the extent to which X provides information on Y
beyond what is already provided by Y itself*

Granger Causality



Clive Granger (1934-2009)

2003 Nobel Prize in Economic Sciences

*the extent to which X provides information on Y
beyond what is already provided by Y itself*

Step 1: $Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \dots + a_k Y_{t-k} + \epsilon_{Y,t}$

Step 2: $Y_t = c_1 Y_{t-1} + c_2 Y_{t-2} + \dots + c_k Y_{t-k} + b_1 X_{t-1} + b_2 X_{t-2} + \dots + b_k X_{t-k} + \epsilon_{X,t}$

Granger Causality



Clive Granger (1934-2009)
2003 Nobel Prize in Economic Sciences

*the extent to which X provides information on Y
beyond what is already provided by Y itself*

use Y to predict itself

Step 1: $Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \dots + a_k Y_{t-k} + \epsilon_{Y,t}$

Step 2: $Y_t = c_1 Y_{t-1} + c_2 Y_{t-2} + \dots + c_k Y_{t-k} + b_1 X_{t-1} + b_2 X_{t-2} + \dots + b_k X_{t-k} + \epsilon_{X,t}$

Granger Causality



Clive Granger (1934-2009)
2003 Nobel Prize in Economic Sciences

*the extent to which X provides information on Y
beyond what is already provided by Y itself*

use Y to predict itself

Step 1: $Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \dots + a_k Y_{t-k} + \epsilon_{Y,t}$

use X to predict Y

Step 2: $Y_t = c_1 Y_{t-1} + c_2 Y_{t-2} + \dots + c_k Y_{t-k} + b_1 X_{t-1} + b_2 X_{t-2} + \dots + b_k X_{t-k} + \epsilon_{X,t}$

Granger Causality



Clive Granger (1934-2009)

2003 Nobel Prize in Economic Sciences

*the extent to which X provides information on Y
beyond what is already provided by Y itself*

use Y to predict itself

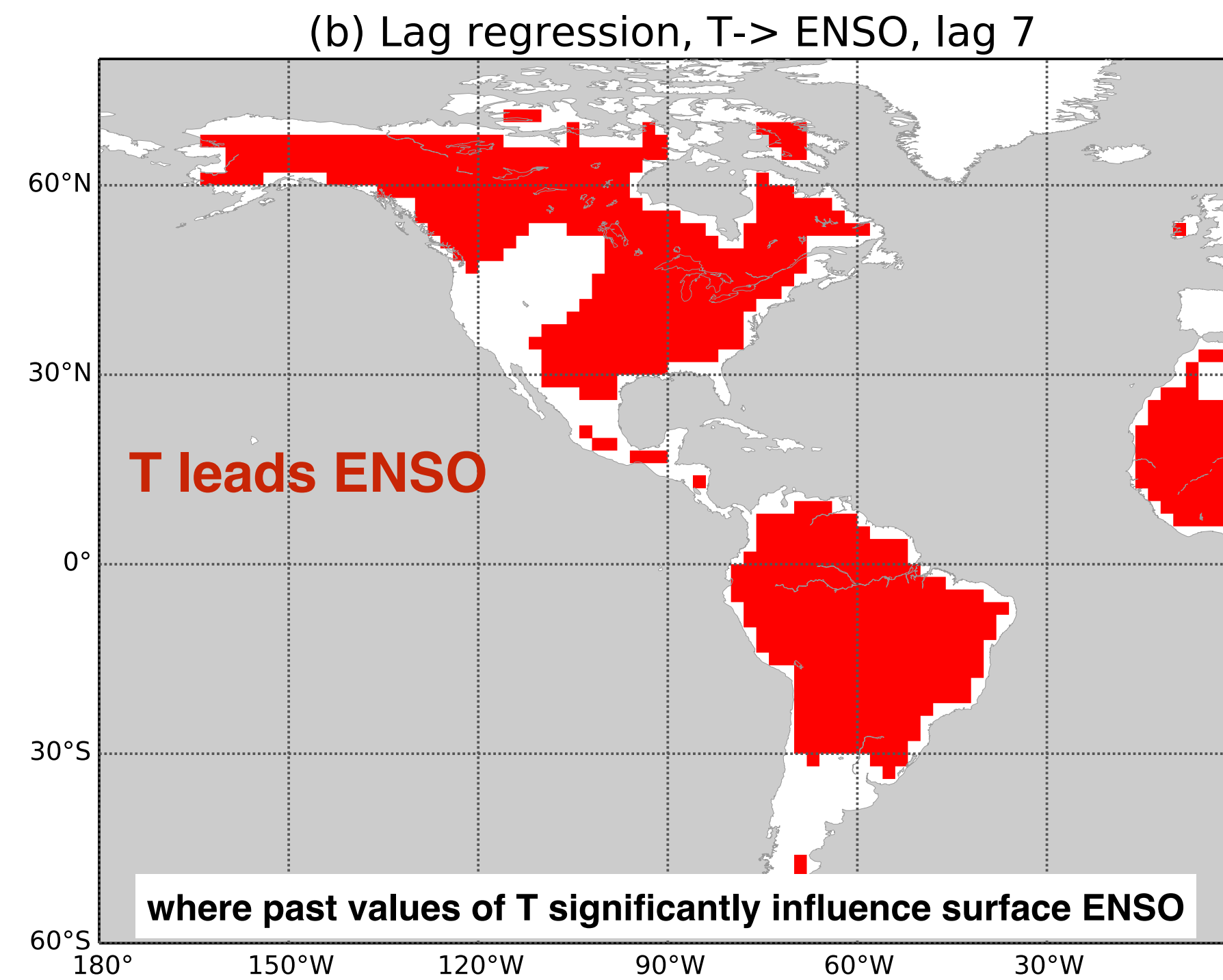
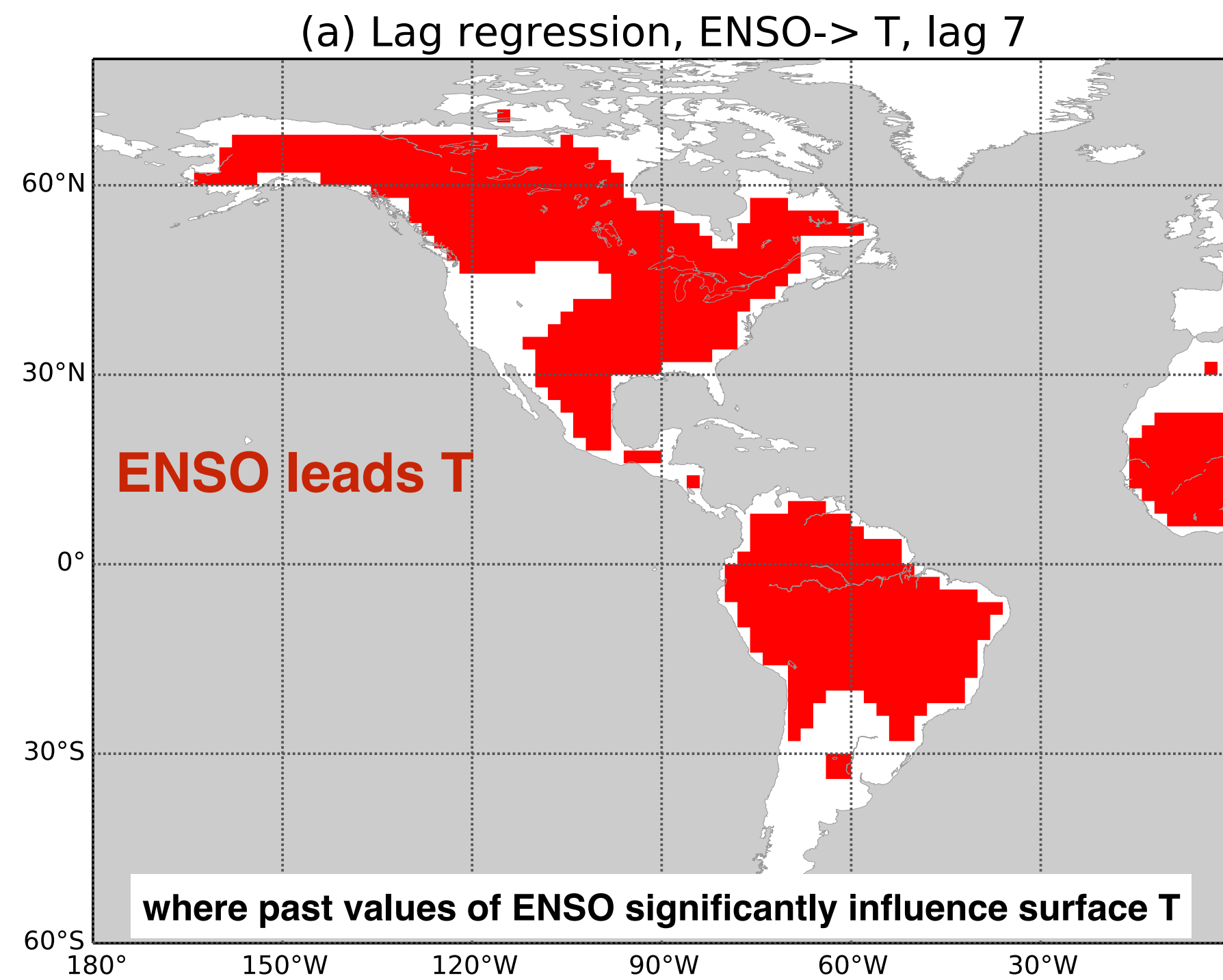
Step 1: $Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \dots + a_k Y_{t-k} + \epsilon_{Y,t}$

use X to predict Y

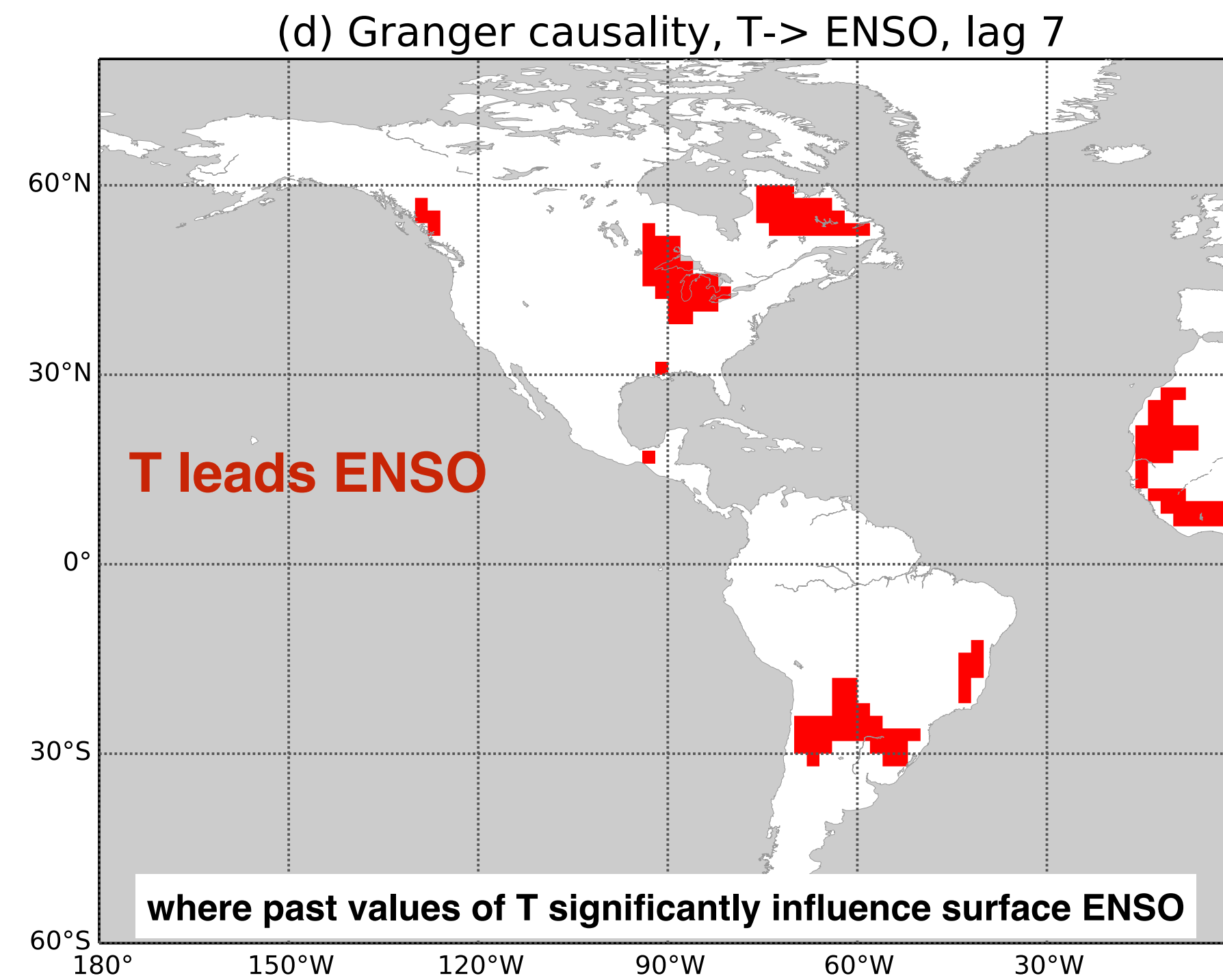
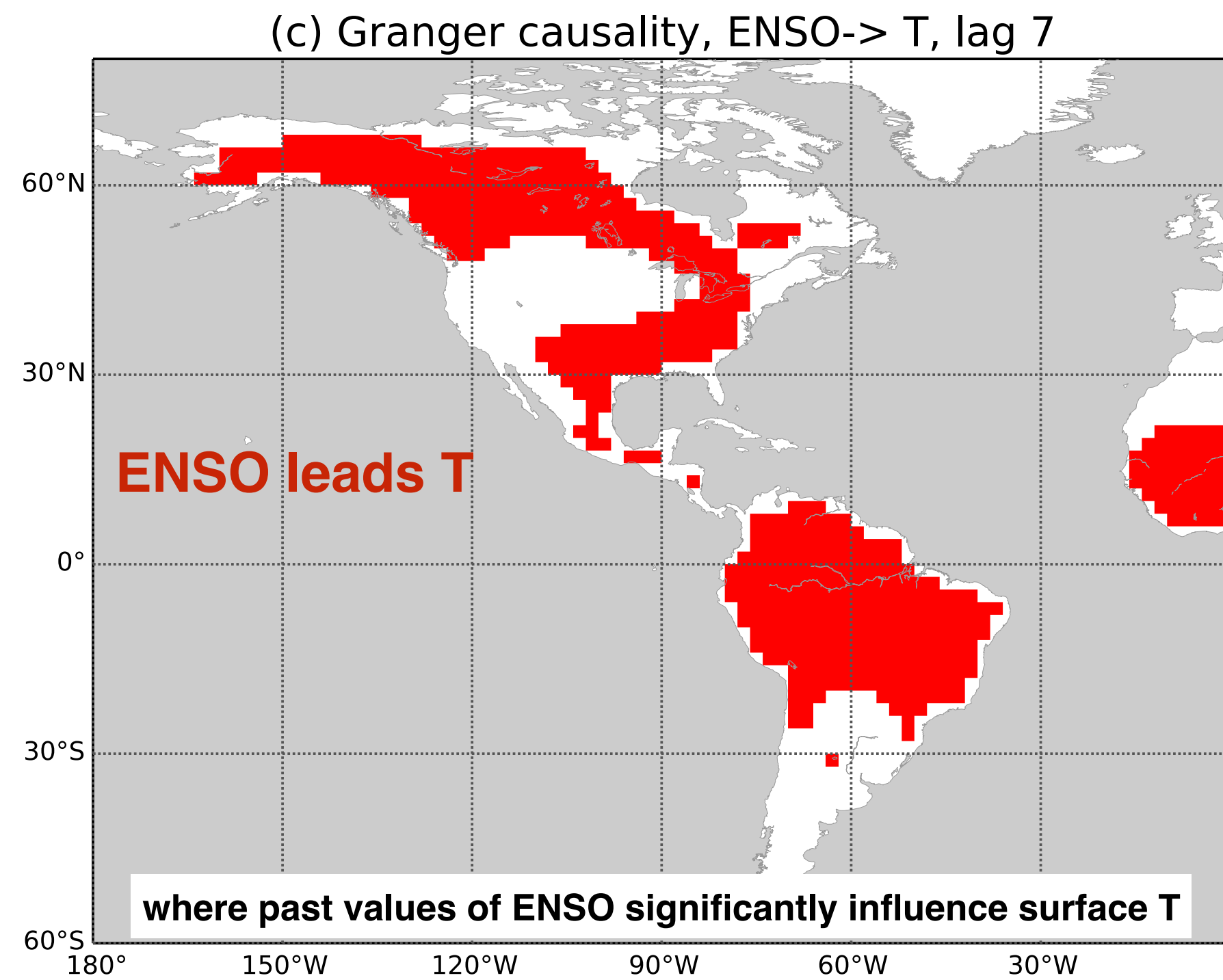
Step 2: $Y_t = c_1 Y_{t-1} + c_2 Y_{t-2} + \dots + c_k Y_{t-k} + b_1 X_{t-1} + b_2 X_{t-2} + \dots + b_k X_{t-k} + \epsilon_{X,t}$

- there exists at least one significant **b** (t-test)
- all of the **b** terms collectively add power to the regression (F-test)

Lagged regression



Granger causality



Observed Feedback between Winter Sea Ice and the North Atlantic Oscillation

COURTENAY STRONG* AND GUDRUN MAGNUSDOTTIR

Department of Earth System Science, University of California, Irvine, Irvine, California

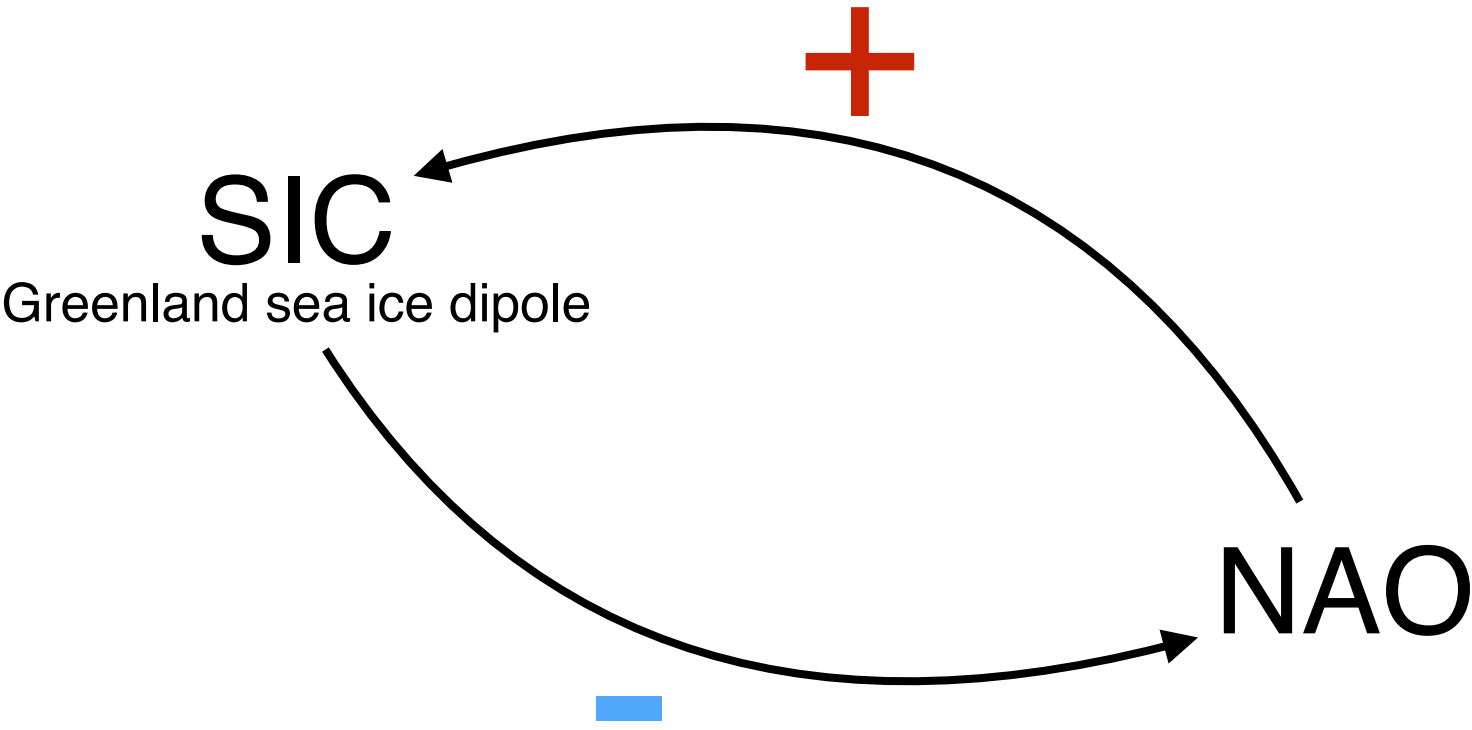
HAL STERN

Department of Statistics, University of California, Irvine, Irvine, California

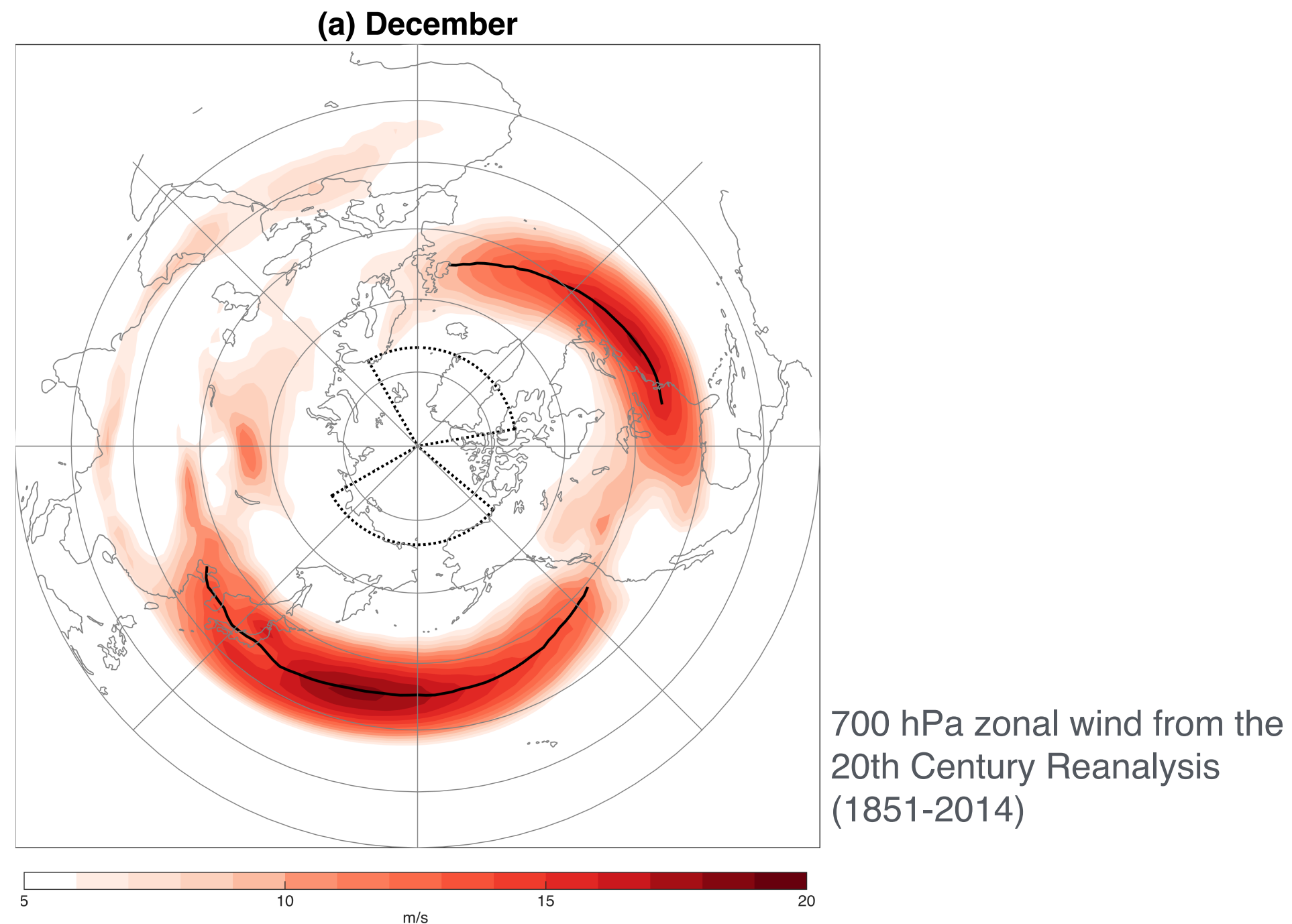
(Manuscript received 20 February 2009, in final form 17 June 2009)

ABSTRACT

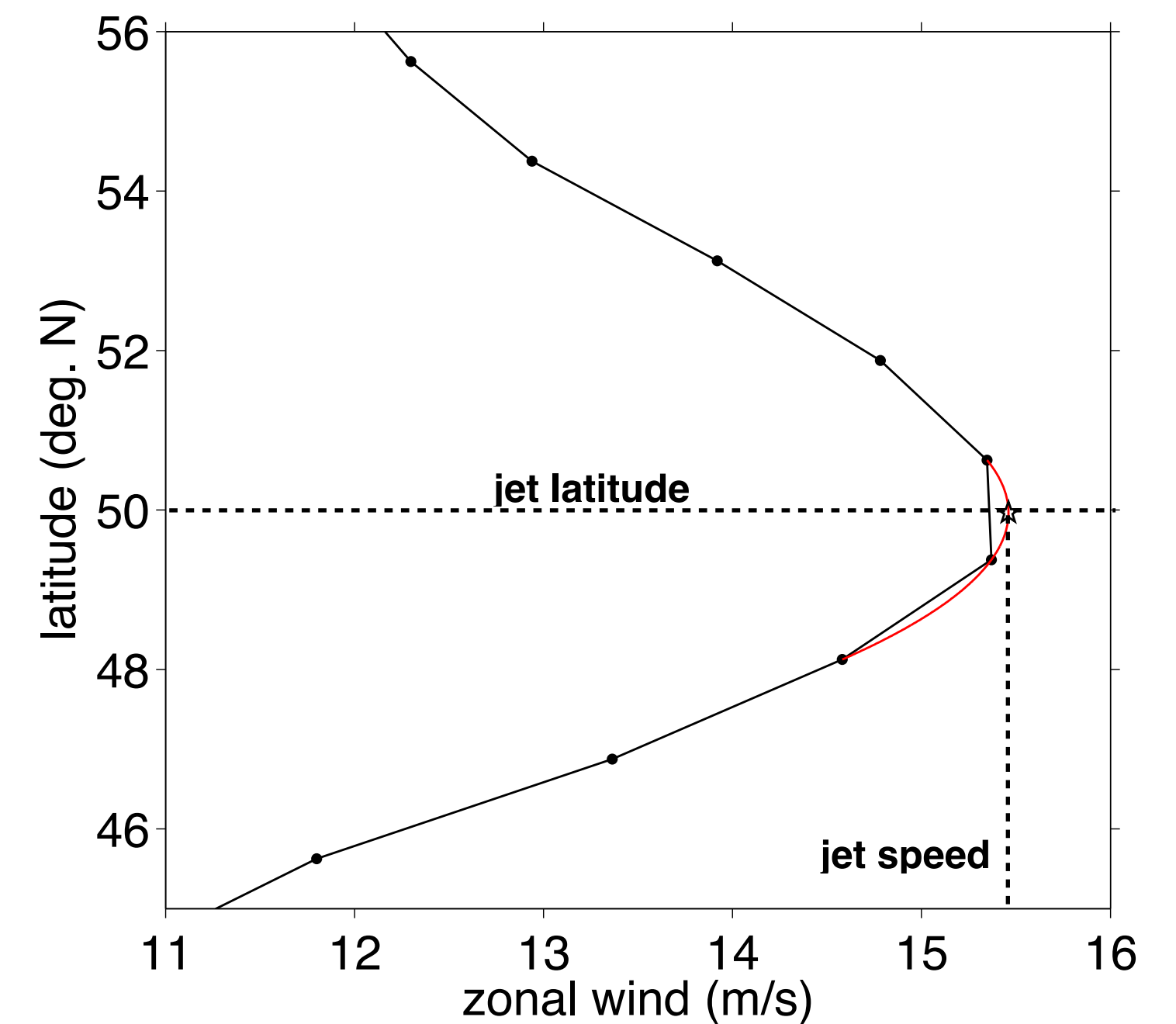
Feedback between the North Atlantic Oscillation (NAO) and winter sea ice variability is detected and quantified using approximately 30 years of observations, a vector autoregressive model (VAR), and testable definitions of Granger causality and feedback. Sea ice variability is quantified based on the leading empirical orthogonal function of sea ice concentration over the North Atlantic [the Greenland Sea ice dipole (GSD)], which, in its positive polarity, has anomalously high sea ice concentrations in the Labrador Sea region to the southwest of Greenland and low sea ice concentrations in the Barents Sea region to the northeast of Greenland. In weekly data for December through April, the VAR indicates that NAO index (N) anomalies cause like-signed anomalies of the standardized GSD index (G), and that G anomalies in turn cause oppositely signed anomalies of N . This negative feedback process operates explicitly on lags of up to four weeks in the VAR but can generate more persistent effects because of the autocorrelation of G . Synthetic data are generated with the VAR to quantify the effects of feedback following realistic local maxima of N and G , and also for sustained high values of G . Feedback can change the expected value of evolving system variables by as much as a half a standard deviation, and the relevance of these results to intraseasonal and interannual NAO and sea ice variability is discussed.



Setup: interested in importance of circulation seasonality



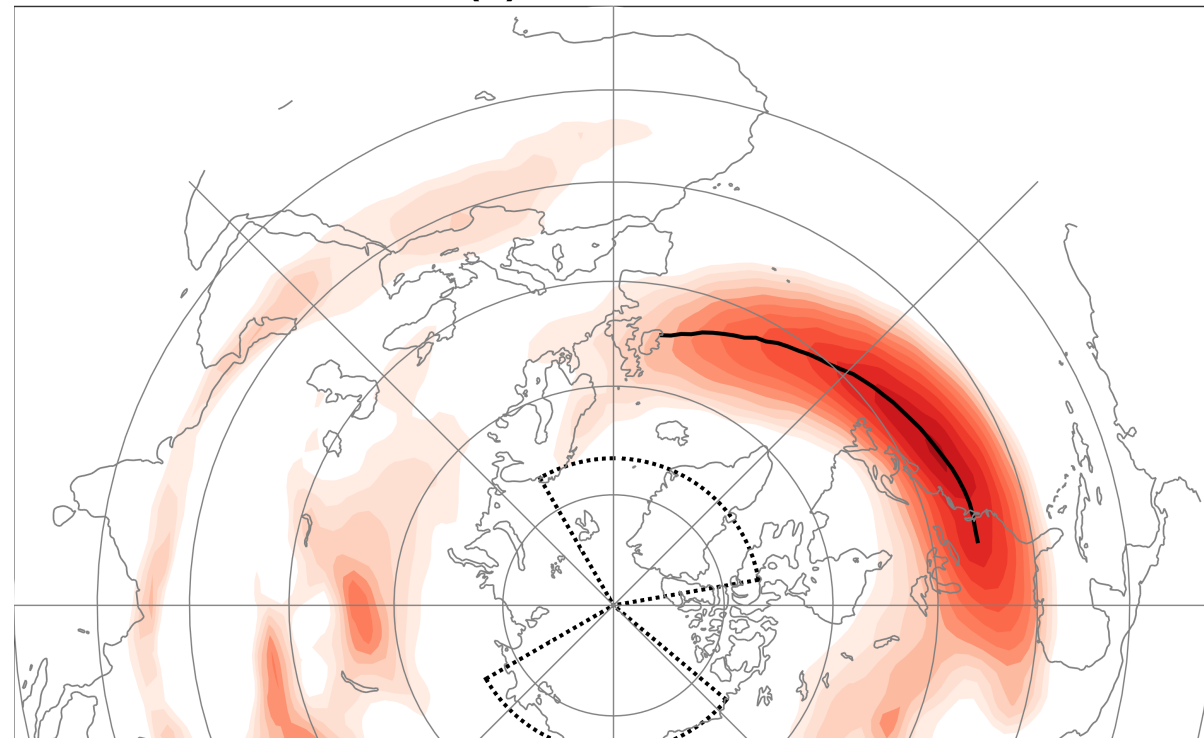
- CMIP5 (historical + RCP8.5: *detrended*)
- daily data (averaged into 10-day chunks*)
- 700 hPa zonal wind
- 850 hPa air temperature



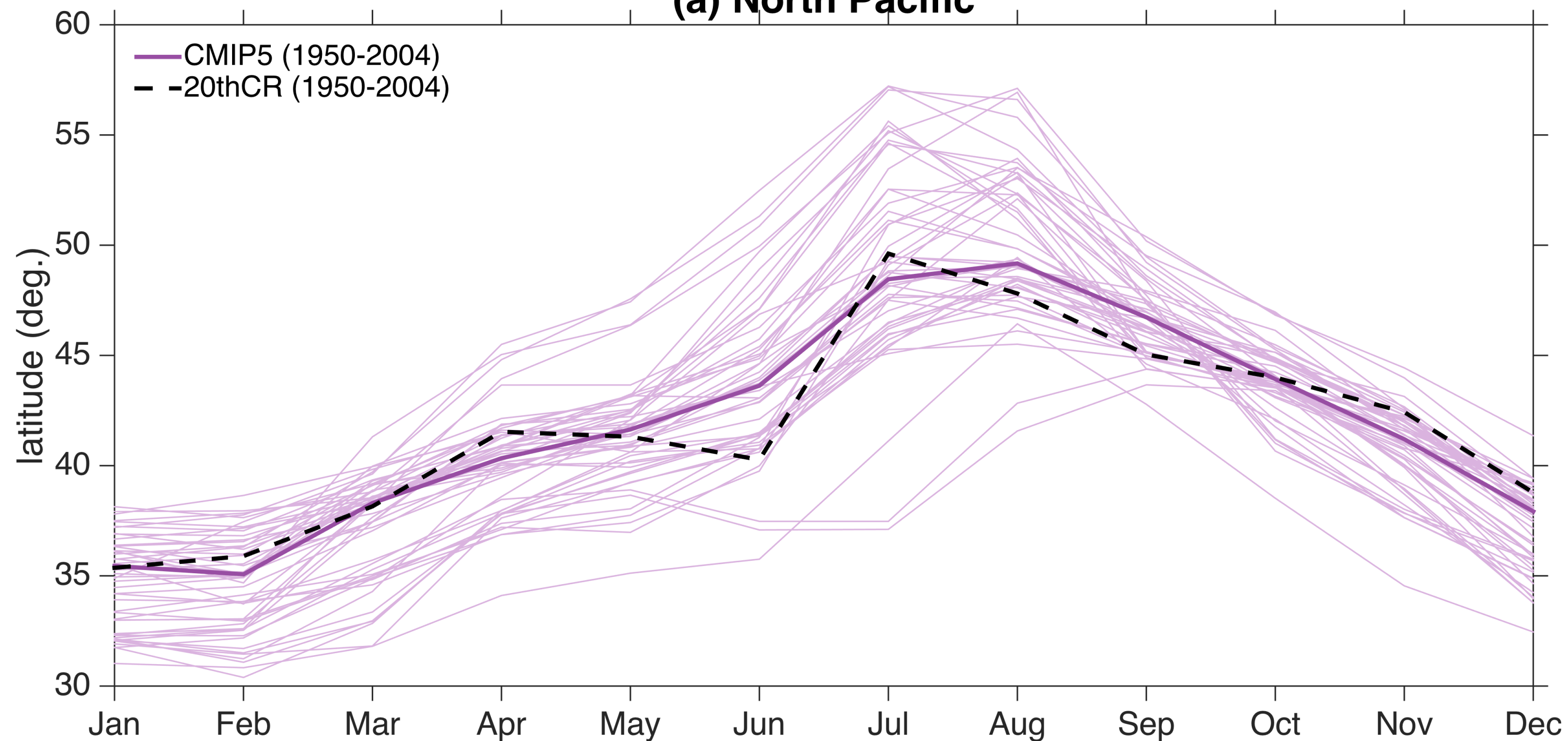
*results are not sensitive to 10-days, e.g. could be 5-days

Setup: interested in importance of circulation seasonality

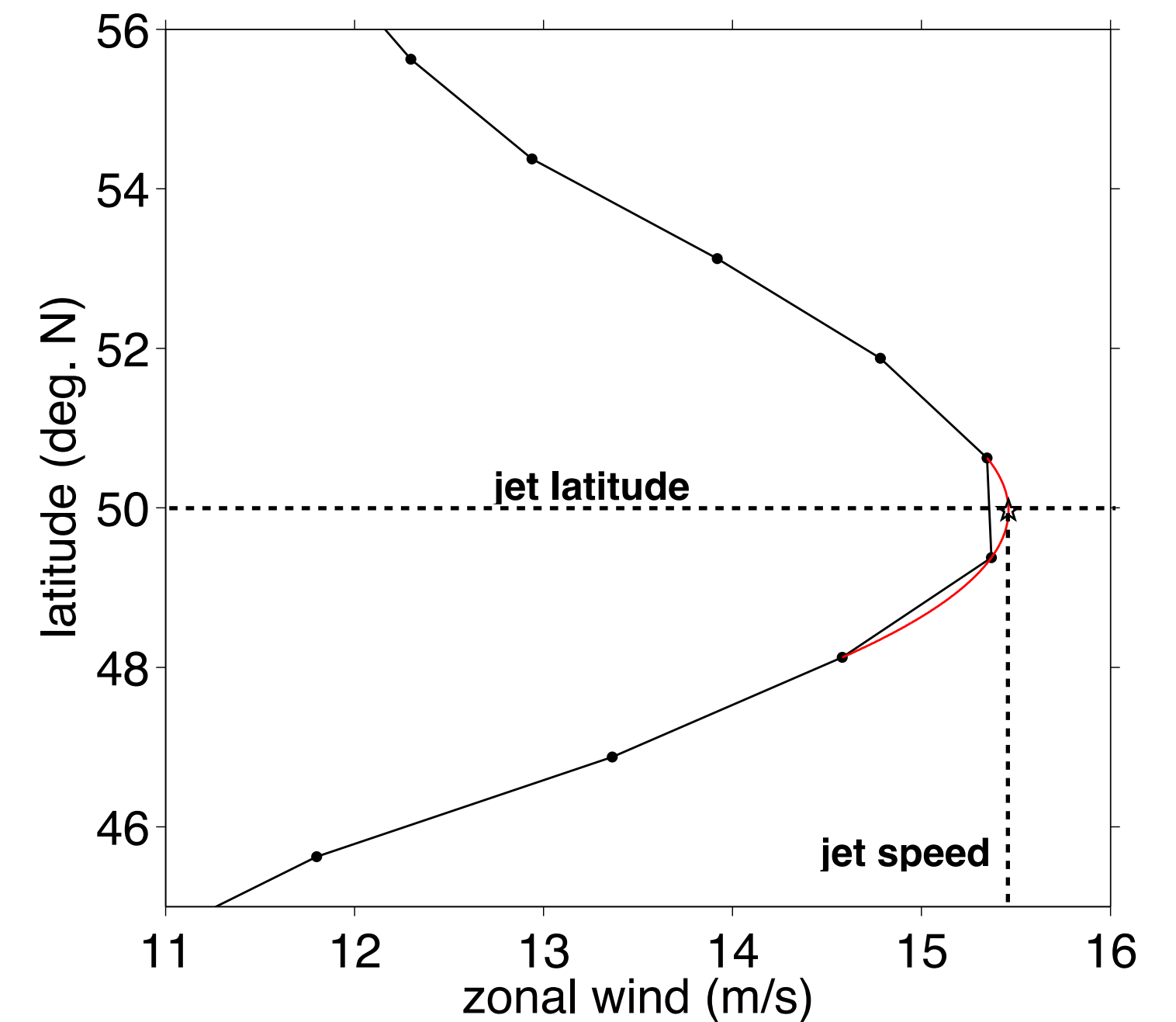
(a) December



(a) North Pacific



- CMIP5 (historical + RCP8.5: *detrended*)
- daily data (averaged into 10-day chunks*)
- 700 hPa zonal wind
- 850 hPa air temperature



*results are not sensitive to 10-days, e.g. could be 5-days

Granger Causality



Clive Granger (1934-2009)
2003 Nobel Prize in Economic Sciences

*the extent to which information of **ARCTIC TEMP** provides information on the **JET** beyond what is already provided by the **JET** itself*

Granger Causality



Clive Granger (1934-2009)
2003 Nobel Prize in Economic Sciences

*the extent to which information of **ARCTIC TEMP** provides information on the **JET** beyond what is already provided by the **JET** itself*

$\mathcal{J} = \text{JET}$

$T = \text{ARCTIC TEMP}$

Granger Causality



Clive Granger (1934-2009)

2003 Nobel Prize in Economic Sciences

*the extent to which information of **ARCTIC TEMP** provides information on the **JET** beyond what is already provided by the **JET** itself*

$$\mathcal{J} = \text{JET}$$

$$T = \text{ARCTIC TEMP}$$

Step 1: $\mathcal{J}_t = a_1 \mathcal{J}_{t-1} + a_2 \mathcal{J}_{t-2} + \dots + a_k \mathcal{J}_{t-k} + \epsilon_{\mathcal{J},t}$

Step 2: $\mathcal{J}_t = c_1 \mathcal{J}_{t-1} + c_2 \mathcal{J}_{t-2} + \dots + c_k \mathcal{J}_{t-k} + b_1 T_{t-1} + b_2 T_{t-2} + \dots + b_k T_{t-k} + \epsilon_{T,t}$

Granger Causality



Clive Granger (1934-2009)
2003 Nobel Prize in Economic Sciences

*the extent to which information of **ARCTIC TEMP** provides information on the **JET** beyond what is already provided by the **JET** itself*

$$\mathcal{J} = \text{JET}$$

$$T = \text{ARCTIC TEMP}$$

Step 1: $\mathcal{J}_t = a_1 \mathcal{J}_{t-1} + a_2 \mathcal{J}_{t-2} + \dots + a_k \mathcal{J}_{t-k} + \epsilon_{\mathcal{J},t}$

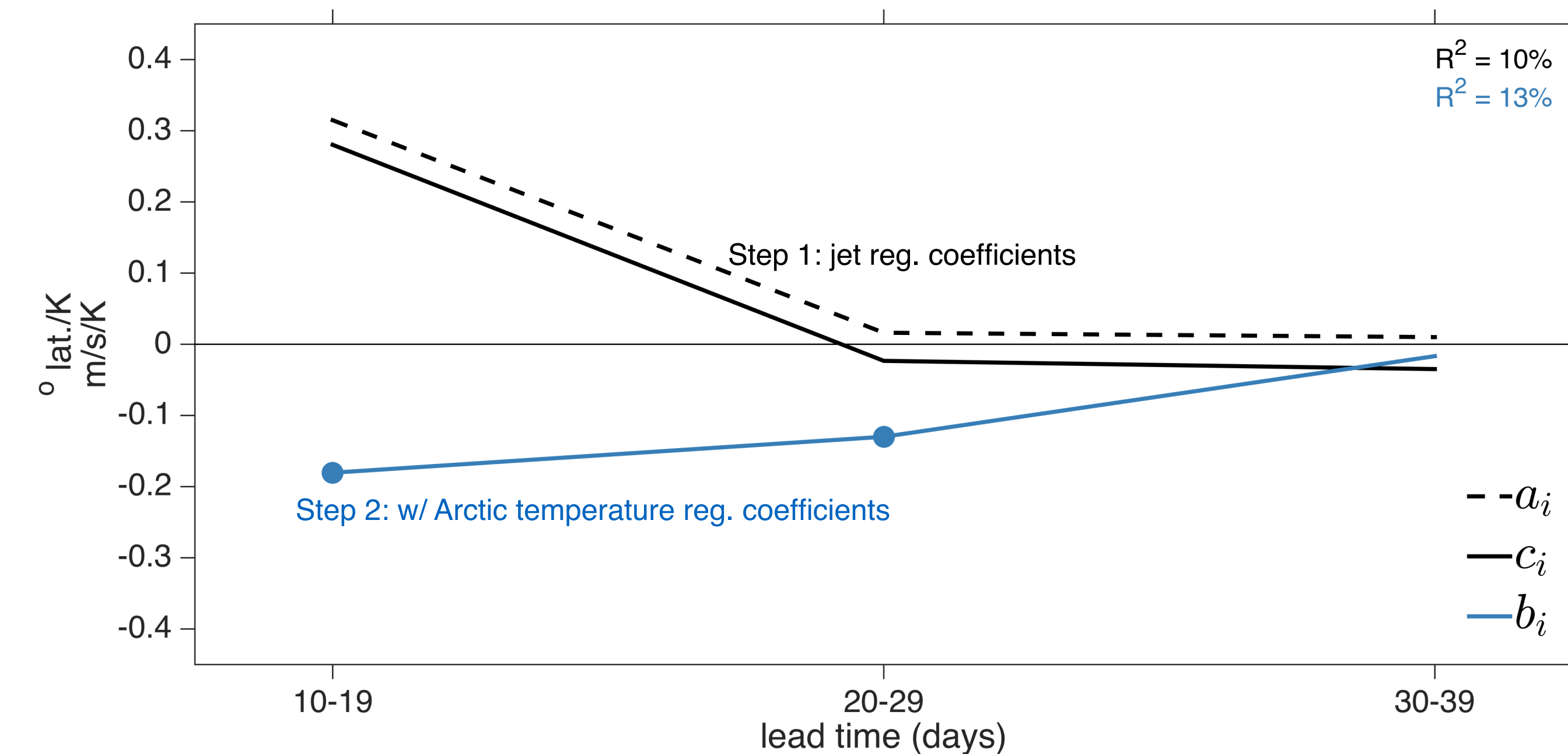
Step 2: $\mathcal{J}_t = c_1 \mathcal{J}_{t-1} + c_2 \mathcal{J}_{t-2} + \dots + c_k \mathcal{J}_{t-k} + b_1 T_{t-1} + b_2 T_{t-2} + \dots + b_k T_{t-k} + \epsilon_{T,t}$

how much?

Applying Granger Causality

Step 1: $\mathcal{J}_t = a_1 \mathcal{J}_{t-1} + a_2 \mathcal{J}_{t-2} + \dots + a_k \mathcal{J}_{t-k} + \epsilon_{\mathcal{J},t}$

Step 2: $\mathcal{J}_t = c_1 \mathcal{J}_{t-1} + c_2 \mathcal{J}_{t-2} + \dots + c_k \mathcal{J}_{t-k} + b_1 T_{t-1} + b_2 T_{t-2} + \dots + b_k T_{t-k} + \epsilon_{T,t}$

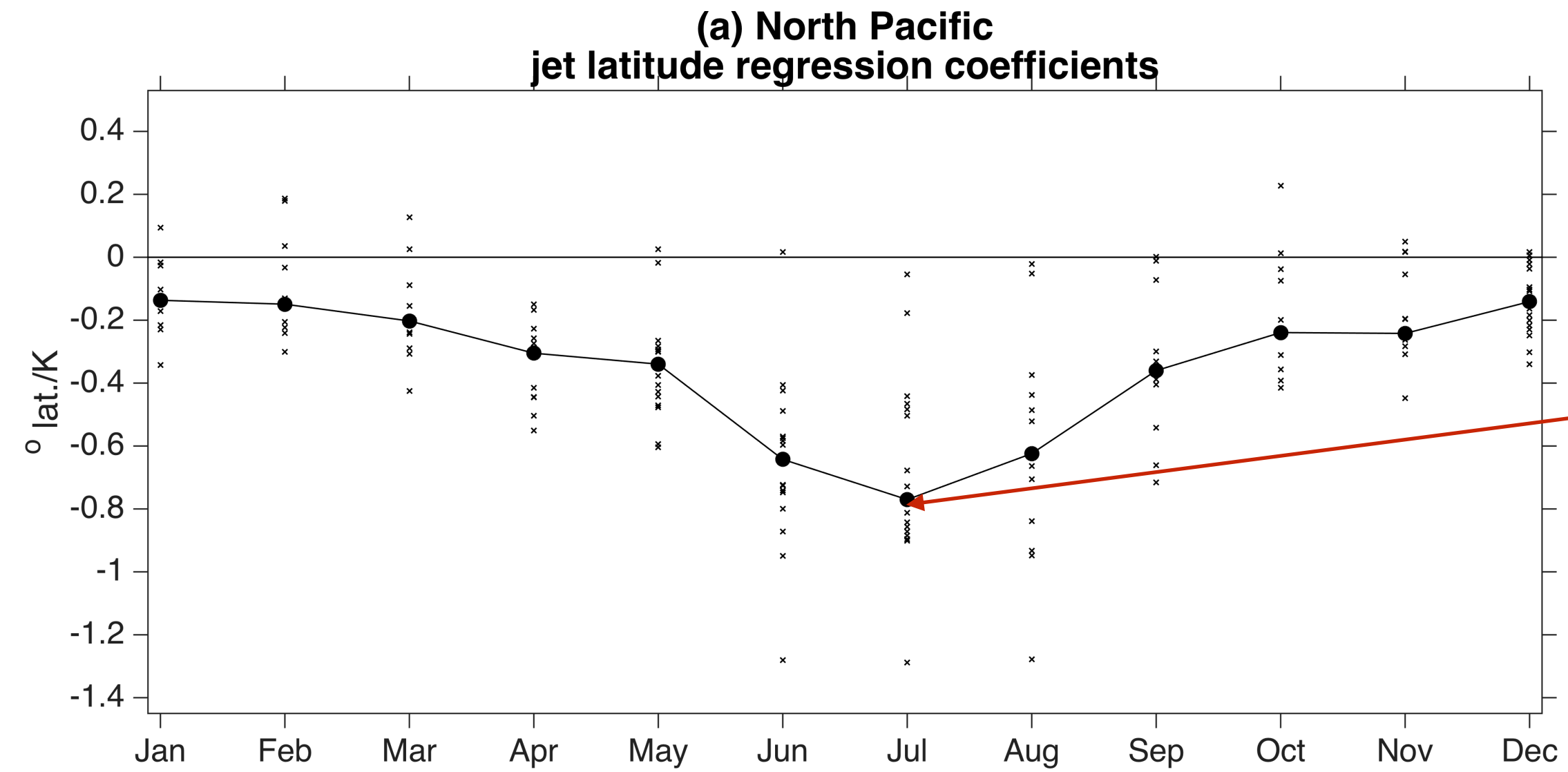


- jet shifts equatorward when Arctic was warm 10-30 days earlier
- first two coefficients are significant

An example Granger causality calculation for North Pacific jet latitude in March from the CanESM2 model. For this example, we use the combined Historical + RCP8.5 time series.

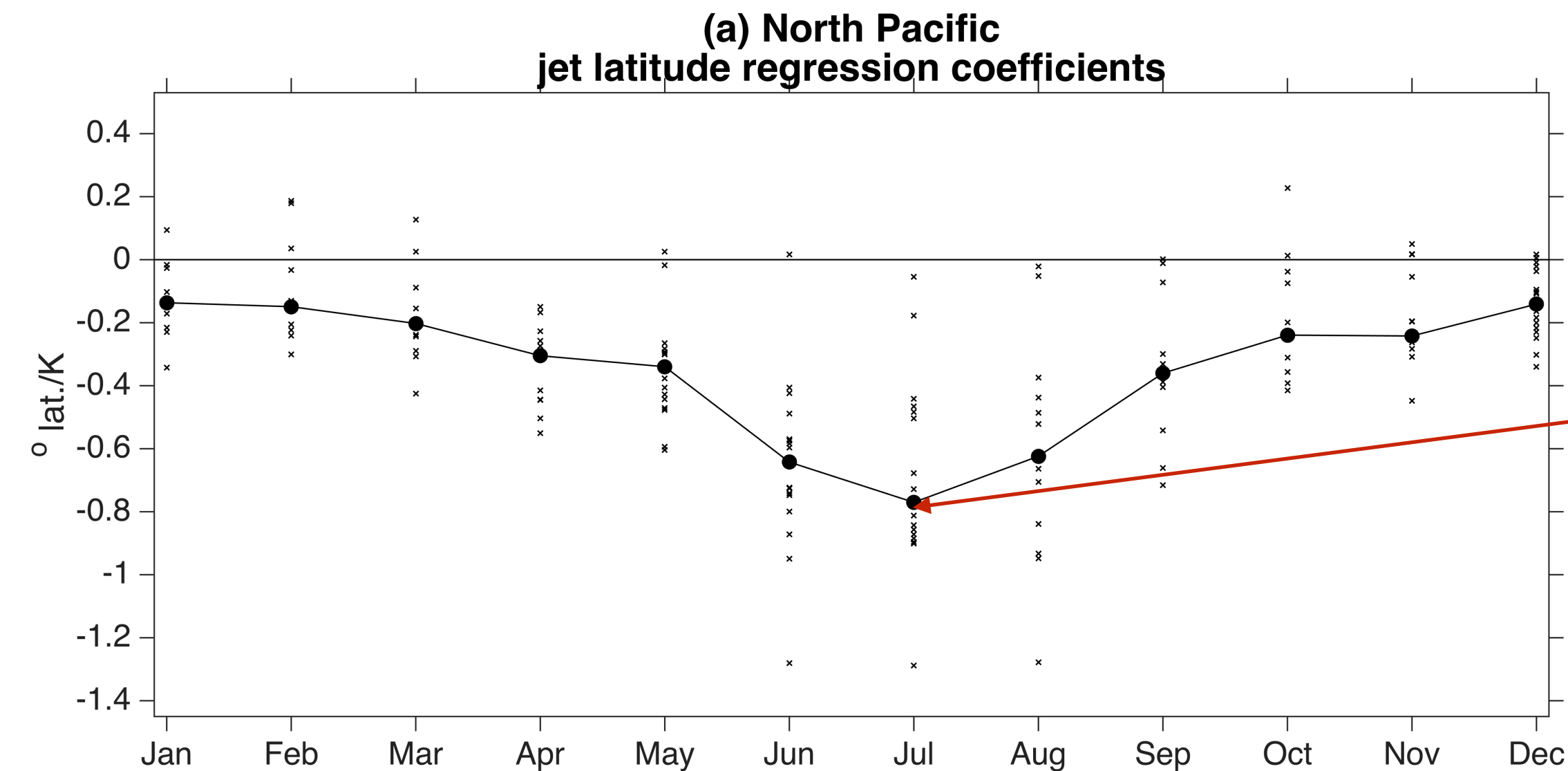
Barnes and Simpson (submitted)

How Much?: **jet position**

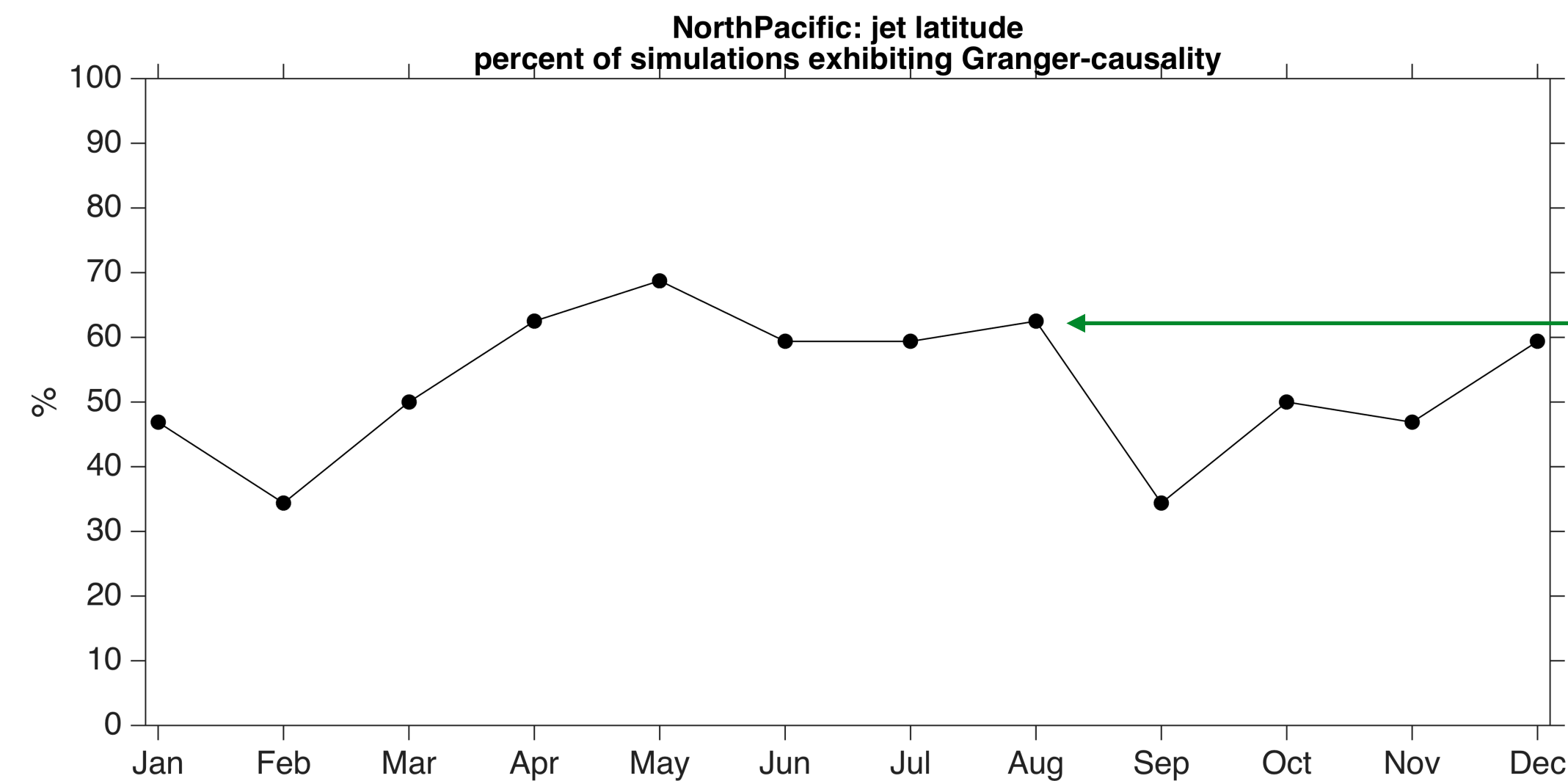


jet shifts **equatorward** more in warm months

How Much?: **jet position**

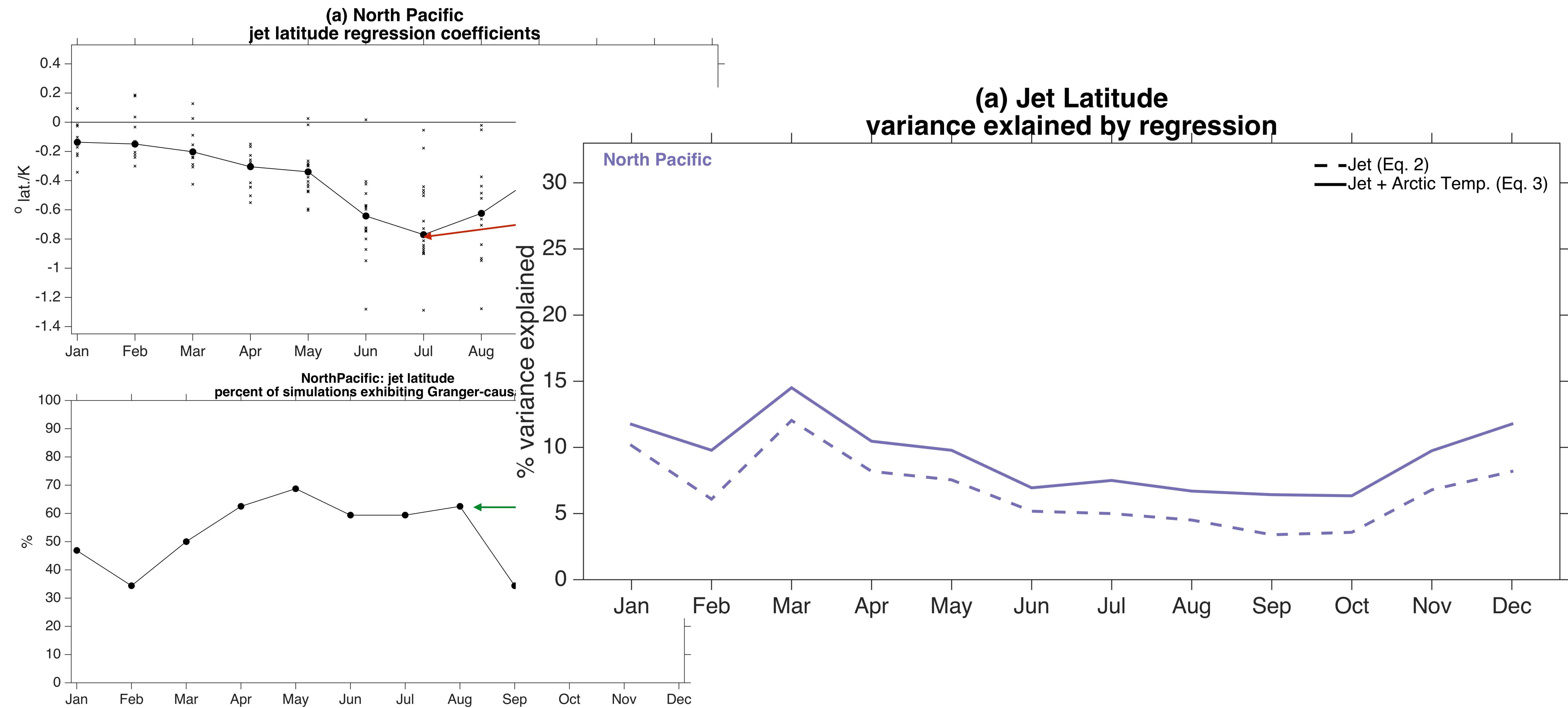


jet shifts **equatorward** more in warm months



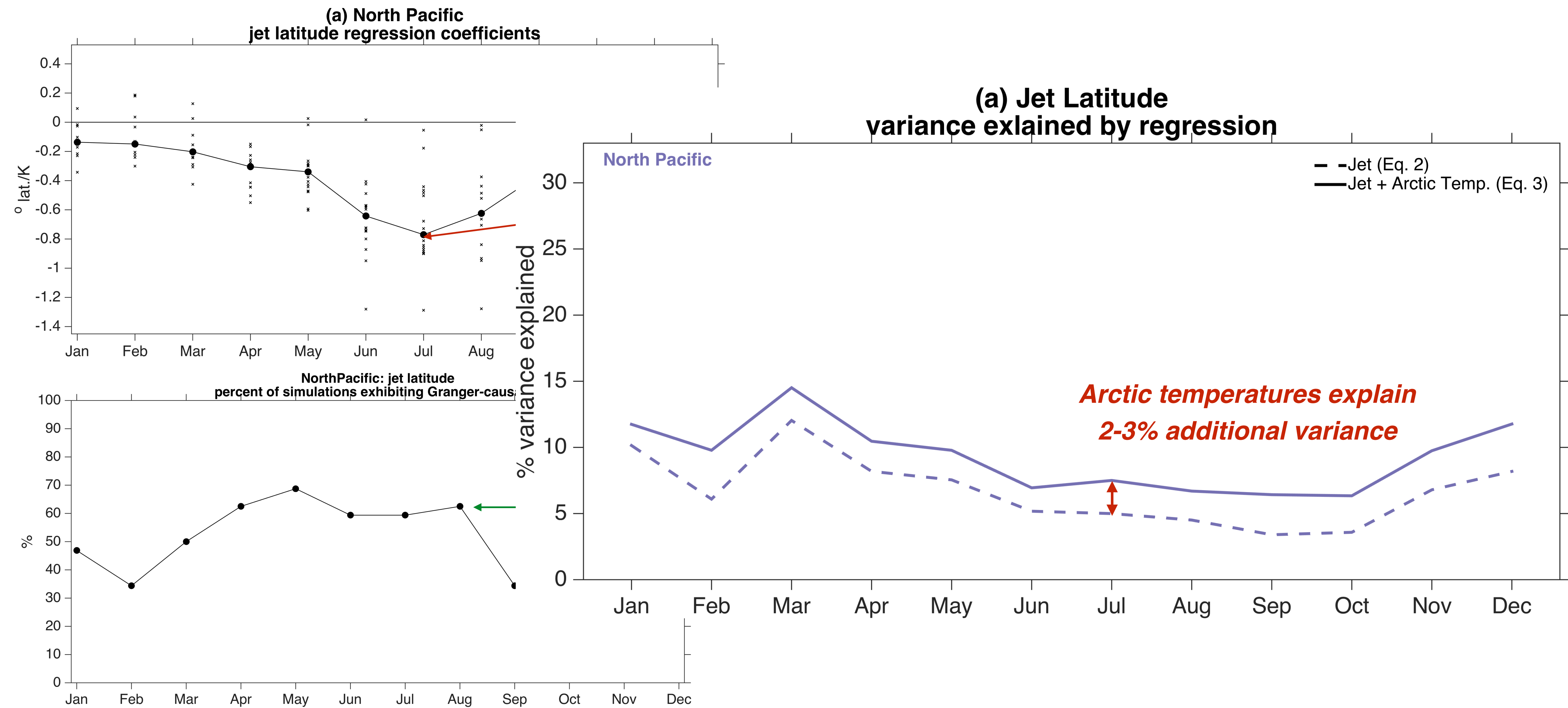
40-60% of simulations exhibit Granger-causality

How Much?: jet position



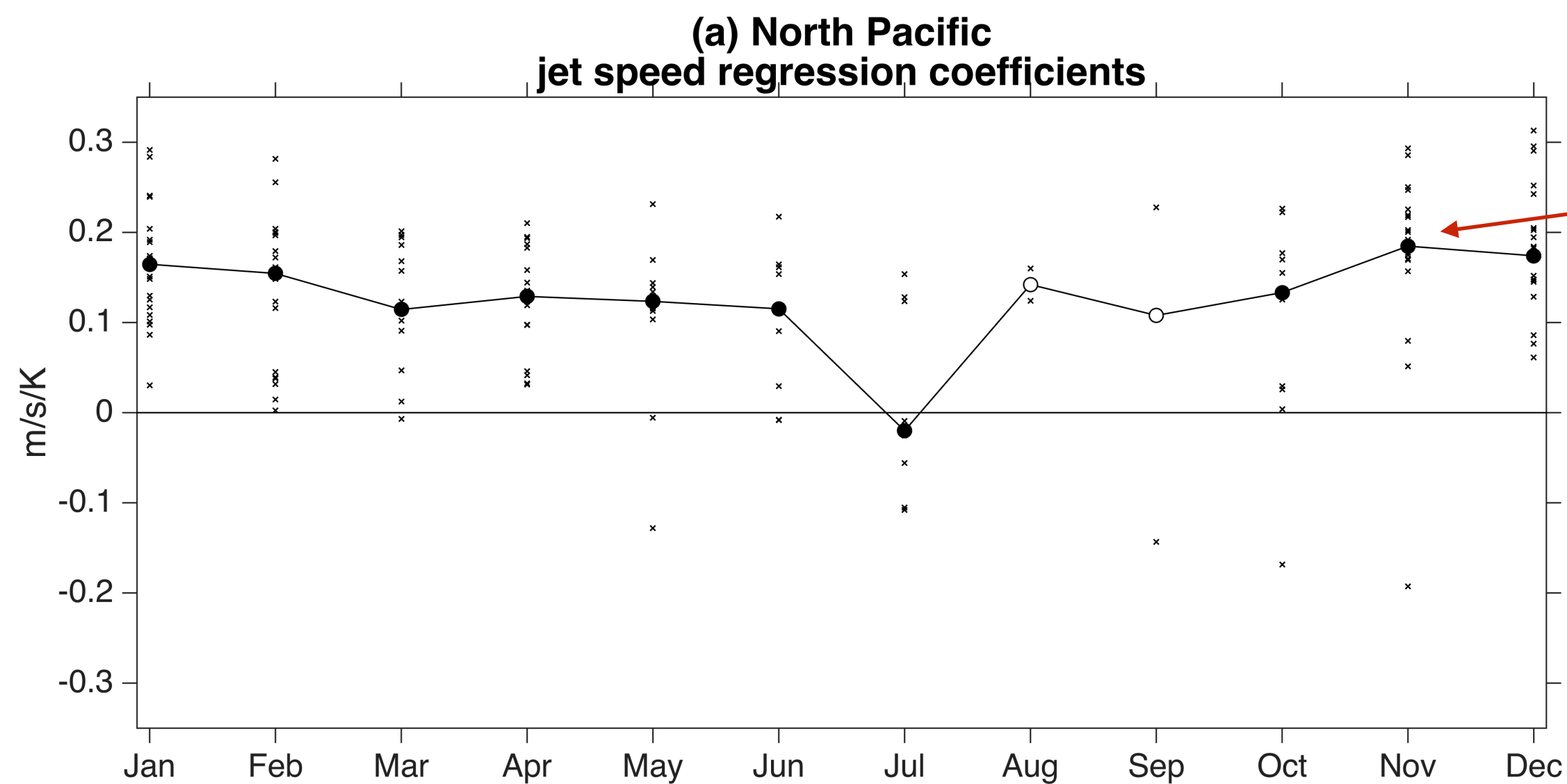
Barnes and Simpson (submitted)

How Much?: jet position

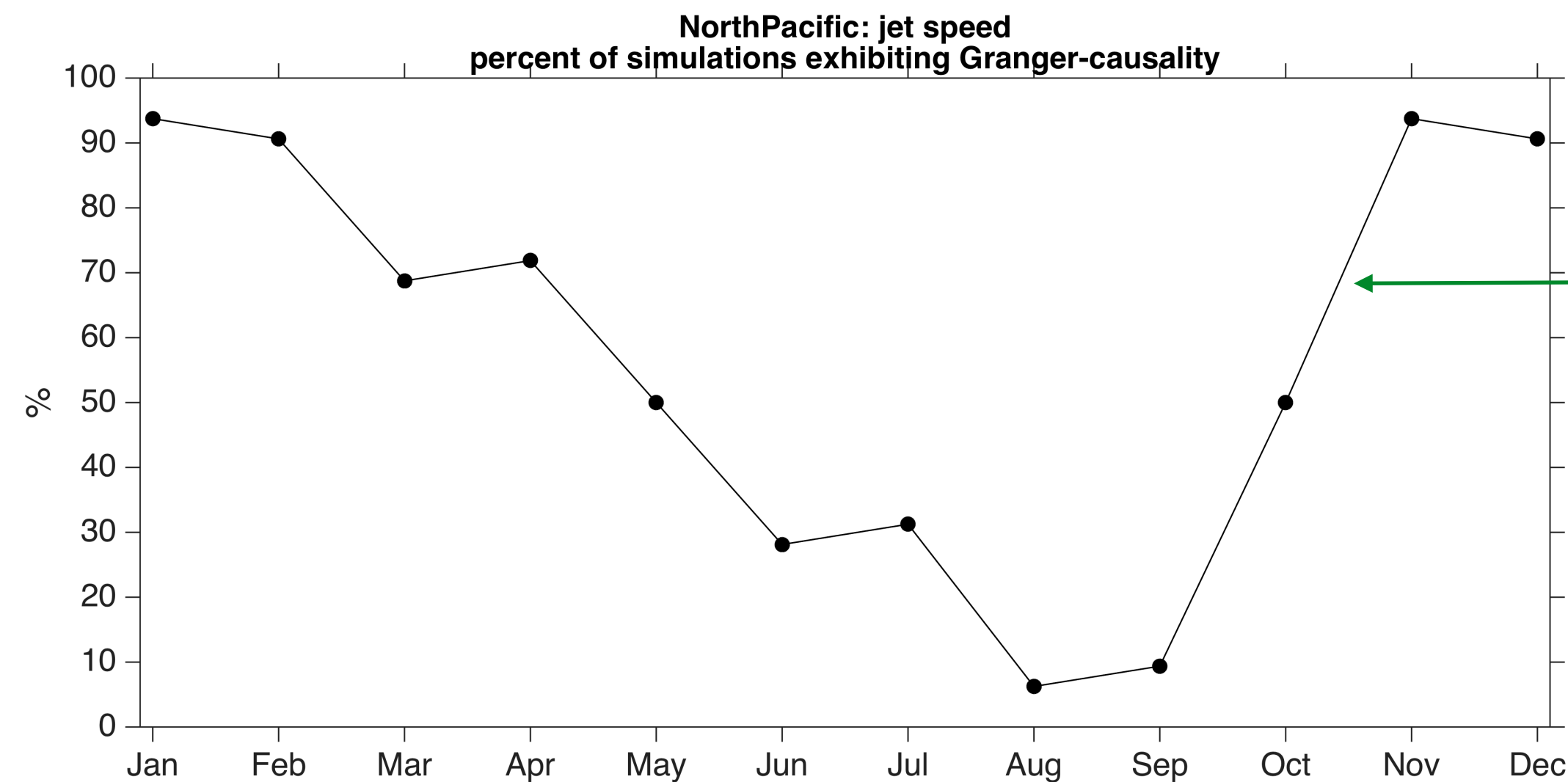


Barnes and Simpson (submitted)

How Much?: jet speed

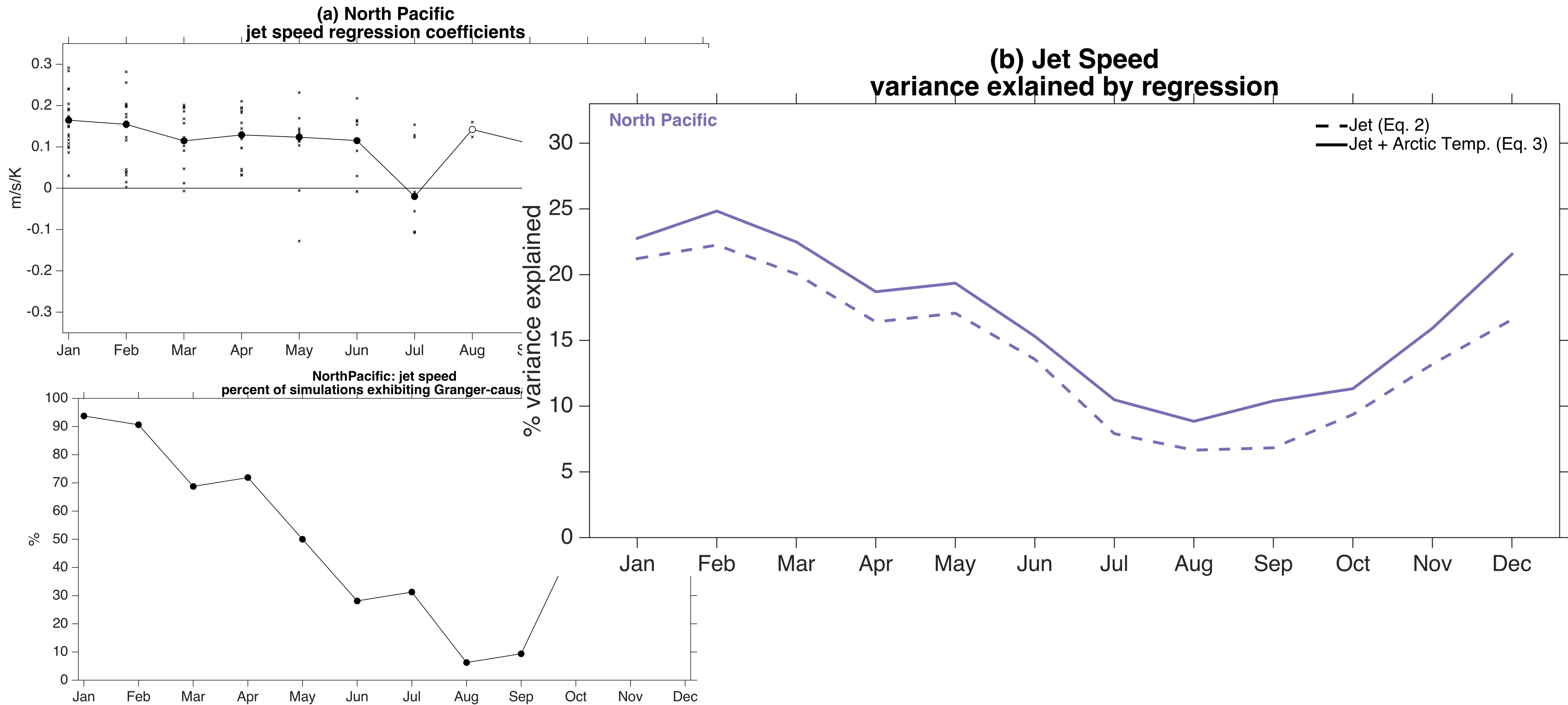


jet strengthens in most months



large seasonality in number of models exhibiting Granger-causality

How Much?: jet speed



Barnes and Simpson (submitted)

Granger Causality



Clive Granger (1934-2009)
2003 Nobel Prize in Economic Sciences

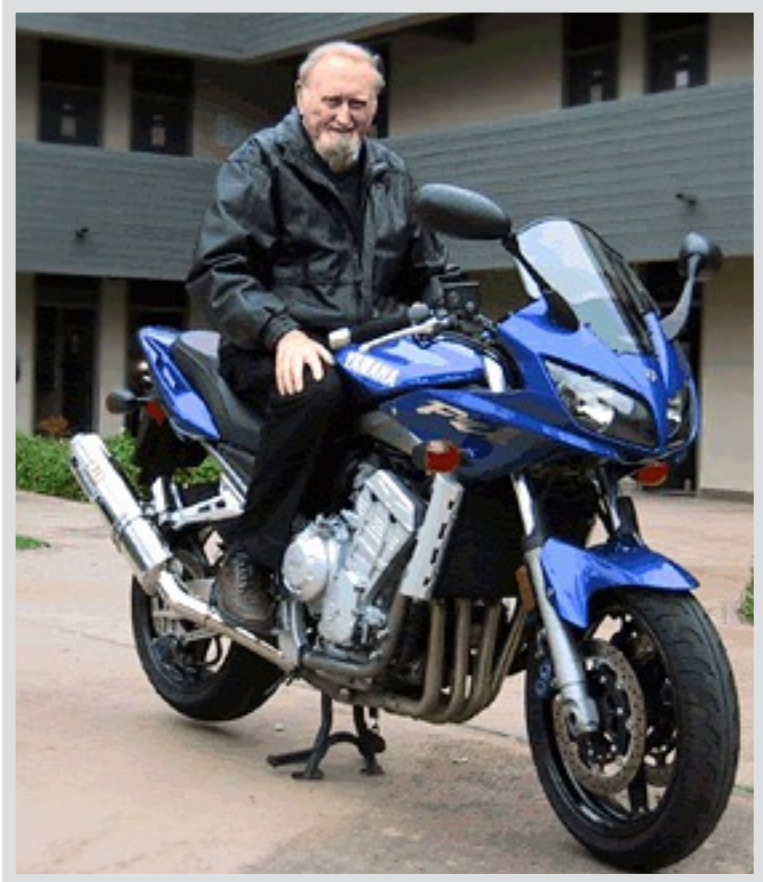
*the extent to which information of **ARCTIC TEMP** provides information for **WINDS** beyond what is provided by the **WINDS***

\mathcal{U} = ZONAL WINDS T = ARCTIC TEMP

Step 1: $\mathcal{U}_t = a_1\mathcal{U}_{t-1} + a_2\mathcal{U}_{t-2} + \dots + a_k\mathcal{U}_{t-k} + \epsilon_{\mathcal{U},t}$

Step 2: $\mathcal{U}_t = c_1\mathcal{U}_{t-1} + c_2\mathcal{U}_{t-2} + \dots + c_k\mathcal{U}_{t-k} + b_1T_{t-1} + b_2T_{t-2} + \dots + b_kT_{t-k} + \epsilon_{T,t}$

Granger Causality



Clive Granger (1934-2009)
2003 Nobel Prize in Economic Sciences

*the extent to which information of **ARCTIC TEMP** provides information for **WINDS** beyond what is provided by the **WINDS***

\mathcal{U} = ZONAL WINDS T = ARCTIC TEMP

Step 1: $\mathcal{U}_t = a_1\mathcal{U}_{t-1} + a_2\mathcal{U}_{t-2} + \dots + a_k\mathcal{U}_{t-k} + \epsilon_{\mathcal{U},t}$

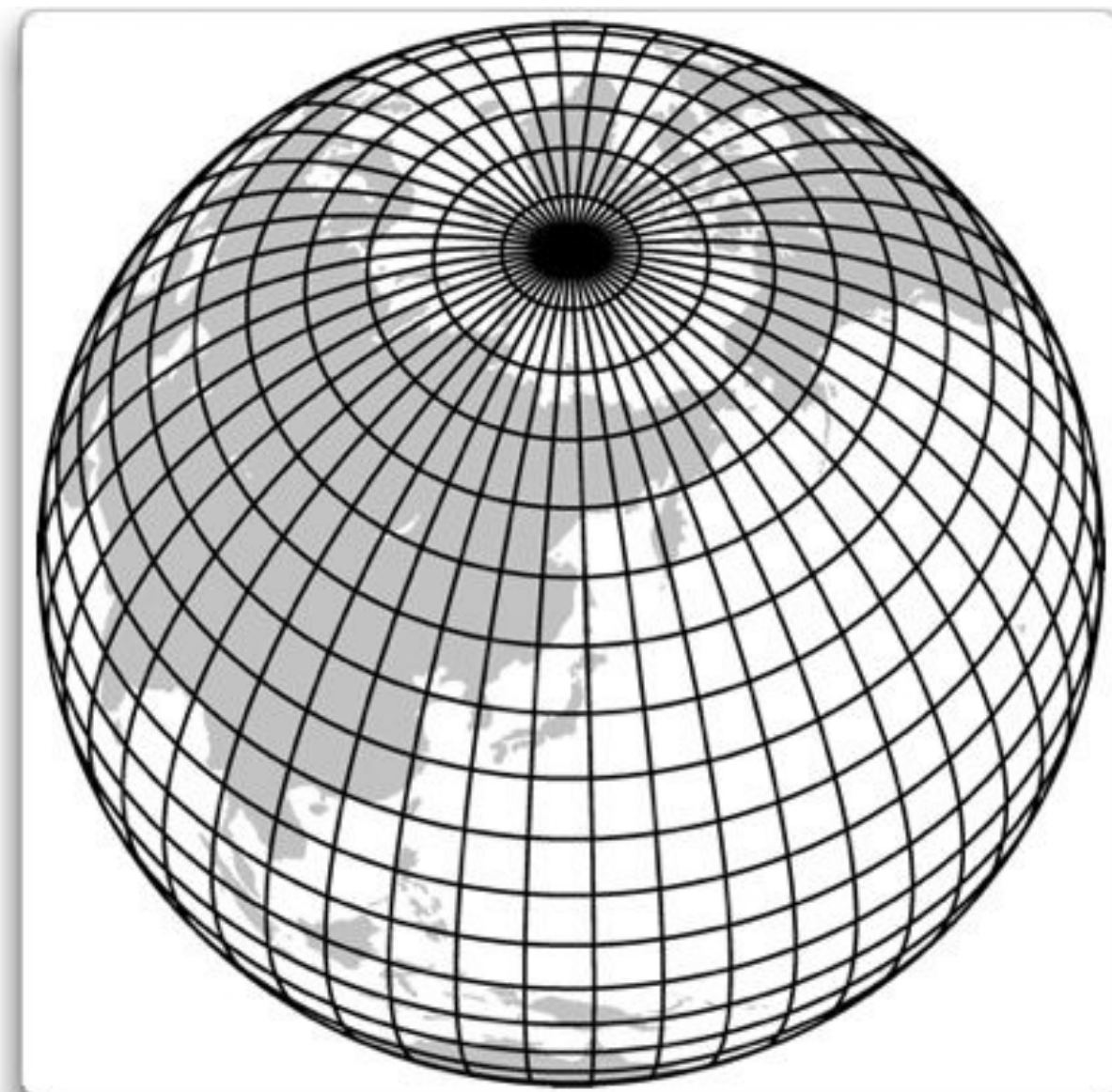
Step 2: $\mathcal{U}_t = c_1\mathcal{U}_{t-1} + c_2\mathcal{U}_{t-2} + \dots + c_k\mathcal{U}_{t-k} + b_1T_{t-1} + b_2T_{t-2} + \dots + b_kT_{t-k} + \epsilon_{T,t}$

how much?

Granger Causality



Clive Granger (1934-2009)
2003 Nobel Prize in Economic Sciences



*the extent to which information of **ARCTIC TEMP** provides information for **WINDS** beyond what is provided by the **WINDS***

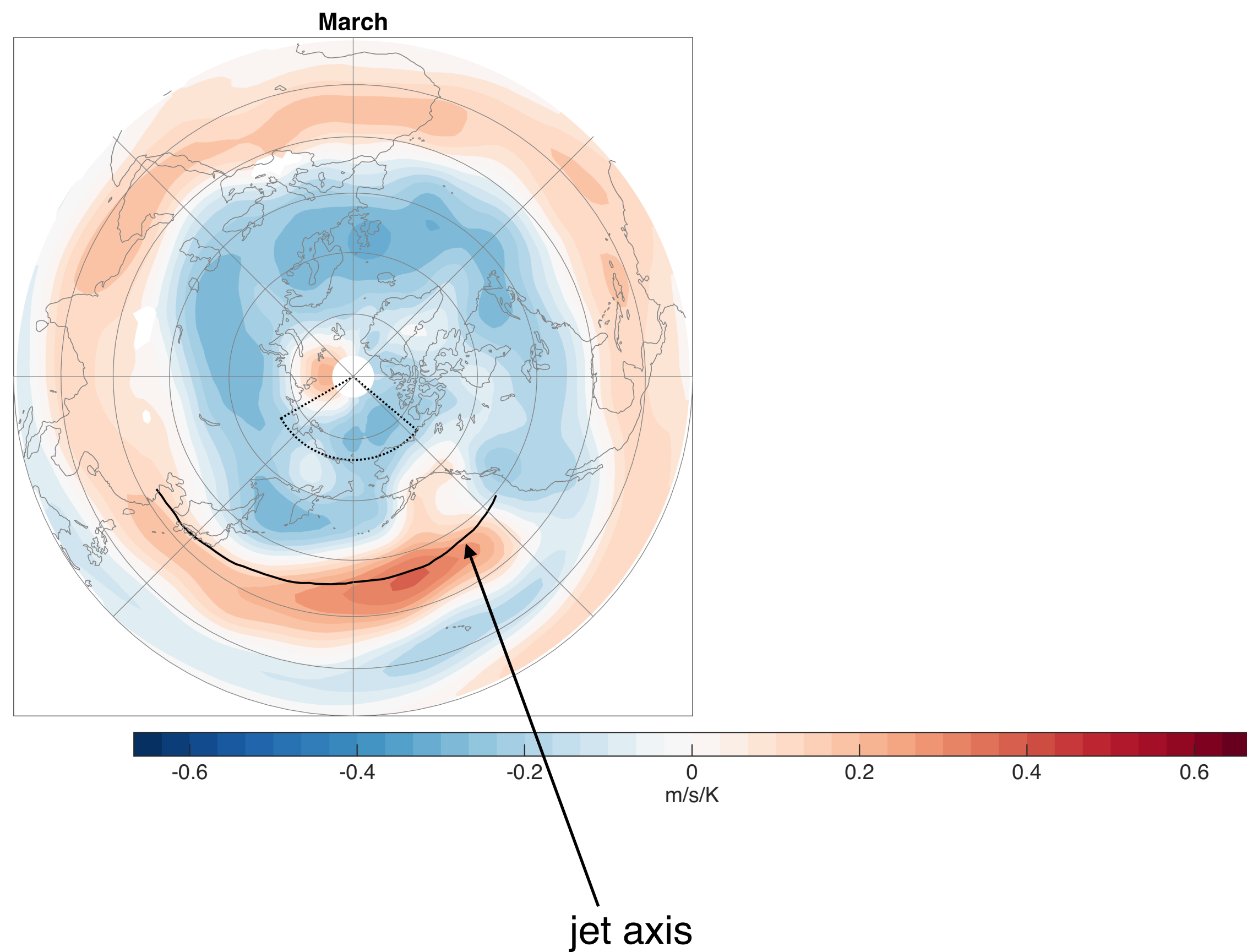
\mathcal{U} = ZONAL WINDS T = ARCTIC TEMP

$$\mathcal{U}_t = a_1\mathcal{U}_{t-1} + a_2\mathcal{U}_{t-2} + \dots + a_k\mathcal{U}_{t-k} + \epsilon_{\mathcal{U},t}$$

$$\mathcal{U}_t = c_1\mathcal{U}_{t-1} + c_2\mathcal{U}_{t-2} + \dots + c_k\mathcal{U}_{t-k} + \boxed{b_1T_{t-1} + b_2T_{t-2} + \dots + b_kT_{t-k}} + \epsilon_{T,t}$$

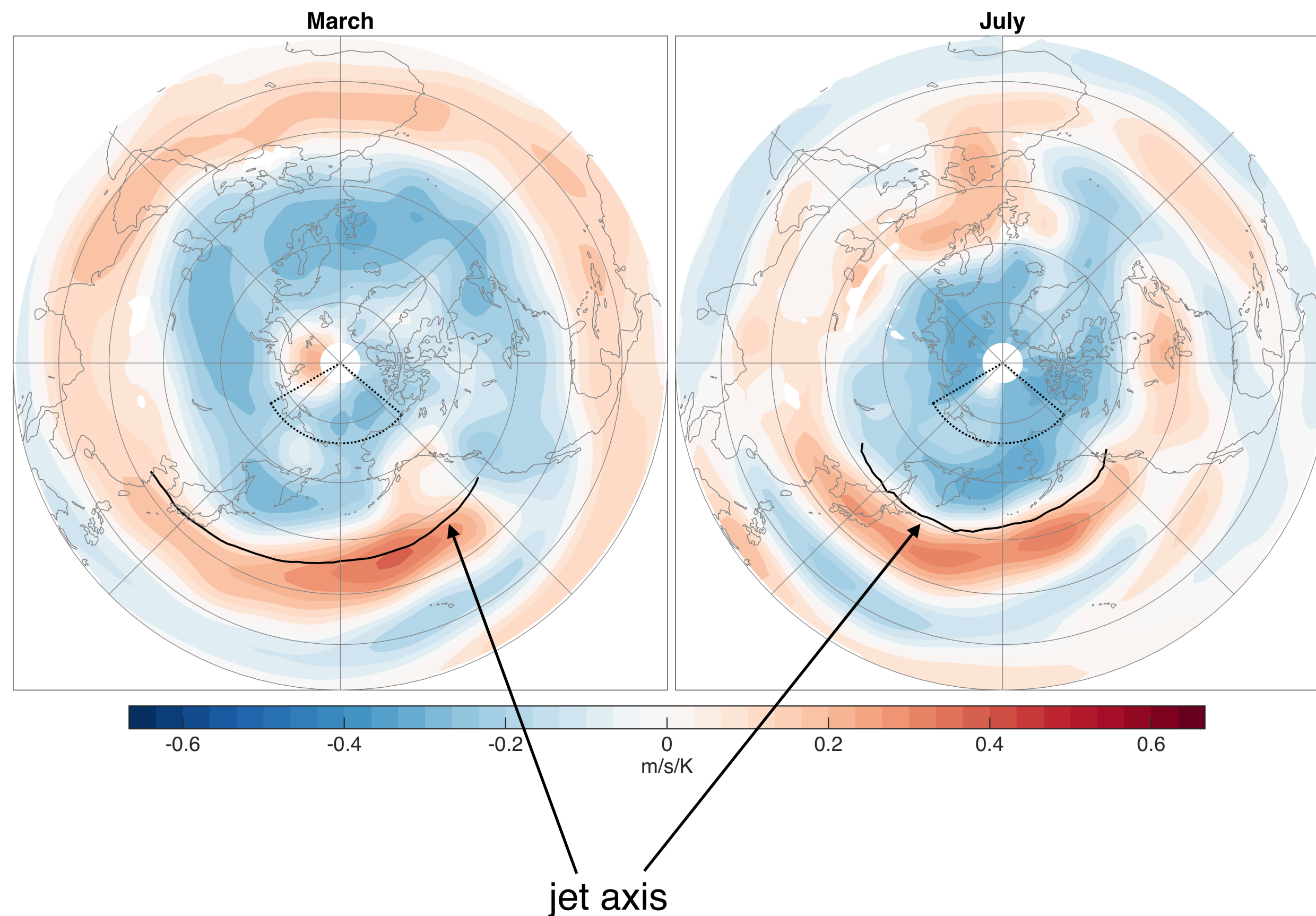
how much?

Looking closer at the seasonality



colors = CMIP5 model mean regression coefficients (via Granger-causality)

Looking closer at the seasonality

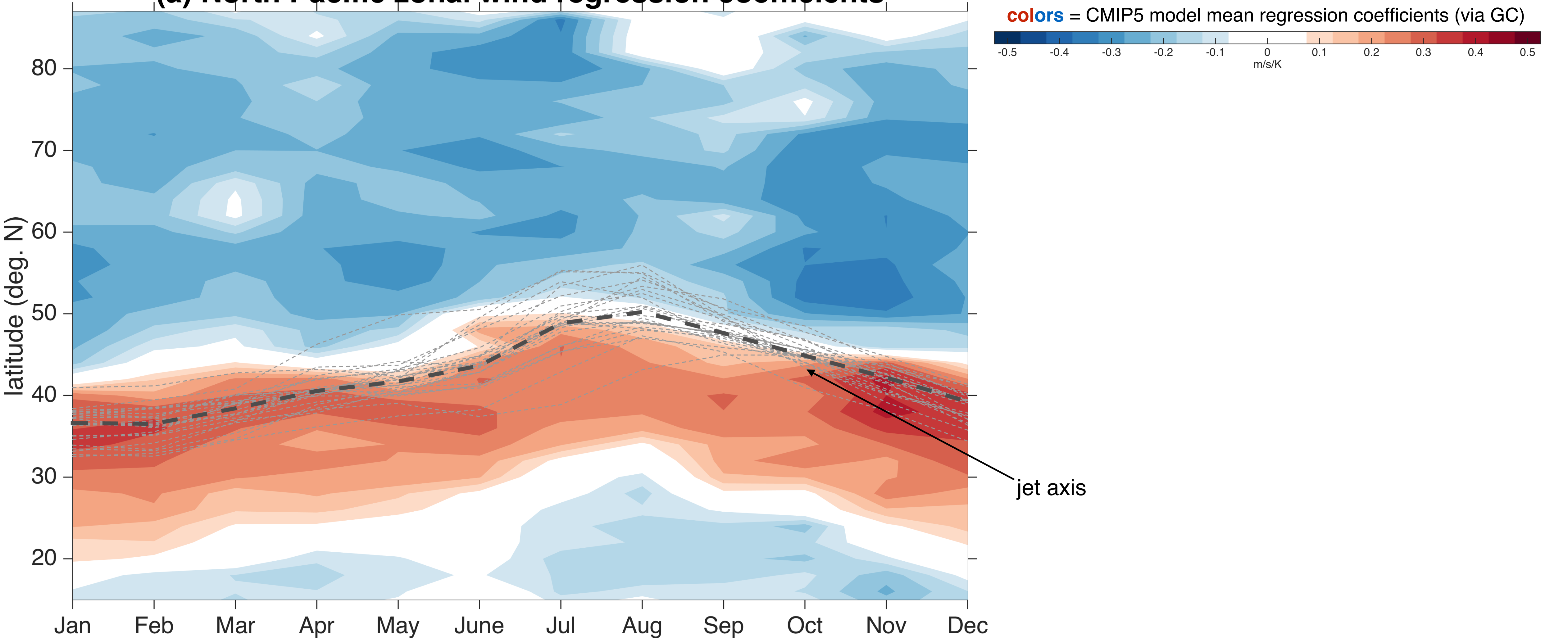


- jet shifts with the seasonal cycle
- wind anomalies remain relatively fixed in latitude
- this has implications for GCM biases (and can be quantified)

colors = CMIP5 model mean regression coefficients (via Granger-causality)

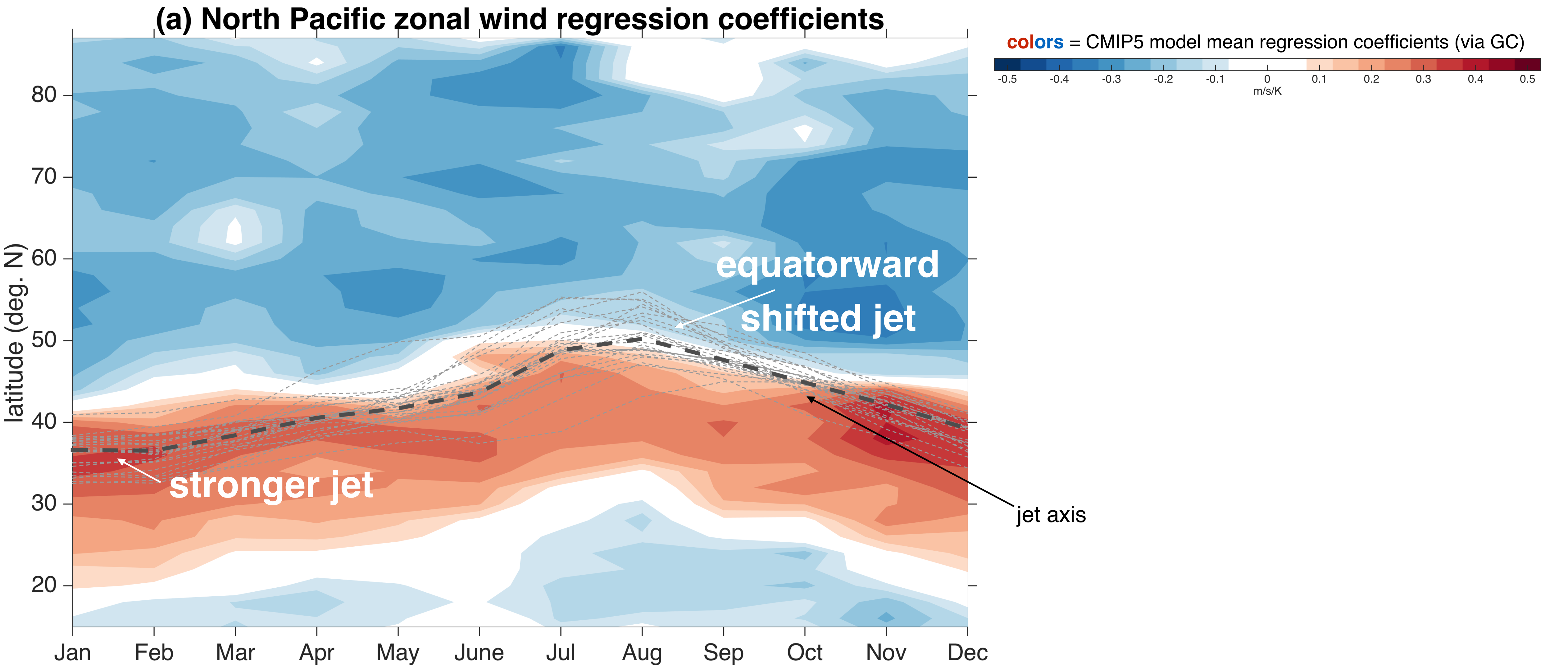
Looking closer at the seasonality

(a) North Pacific zonal wind regression coefficients



Barnes and Simpson (submitted)

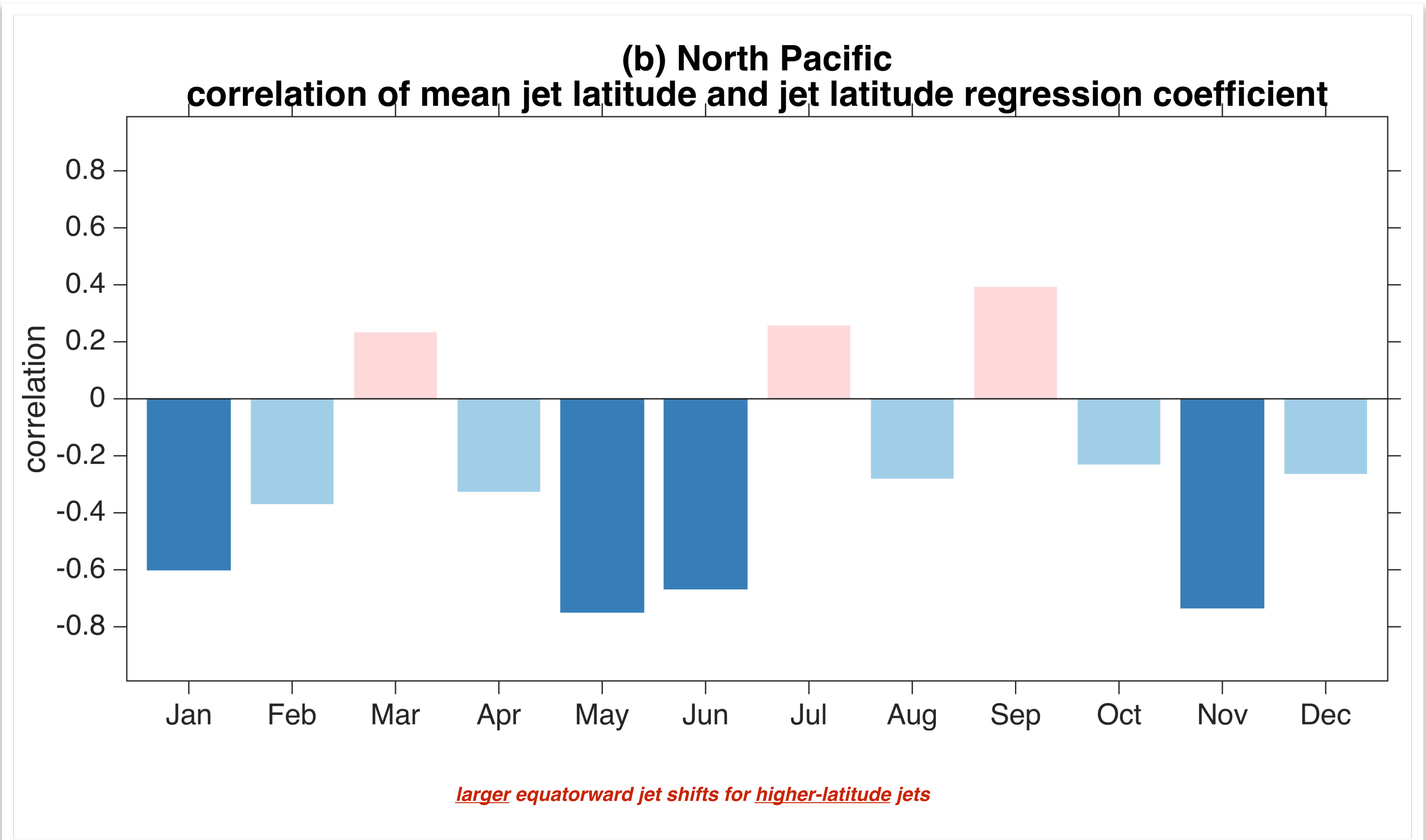
Looking closer at the seasonality



Barnes and Simpson (submitted)

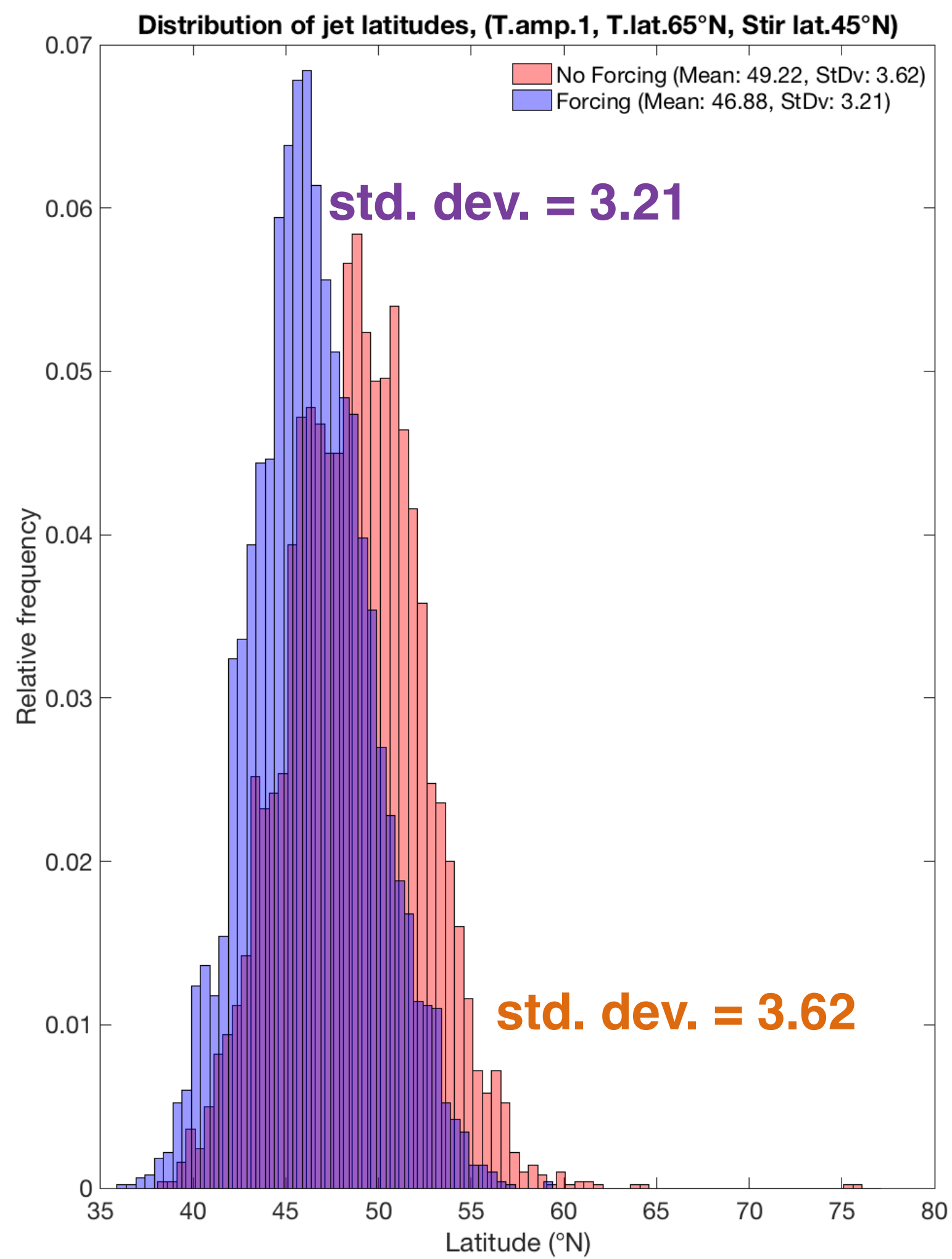
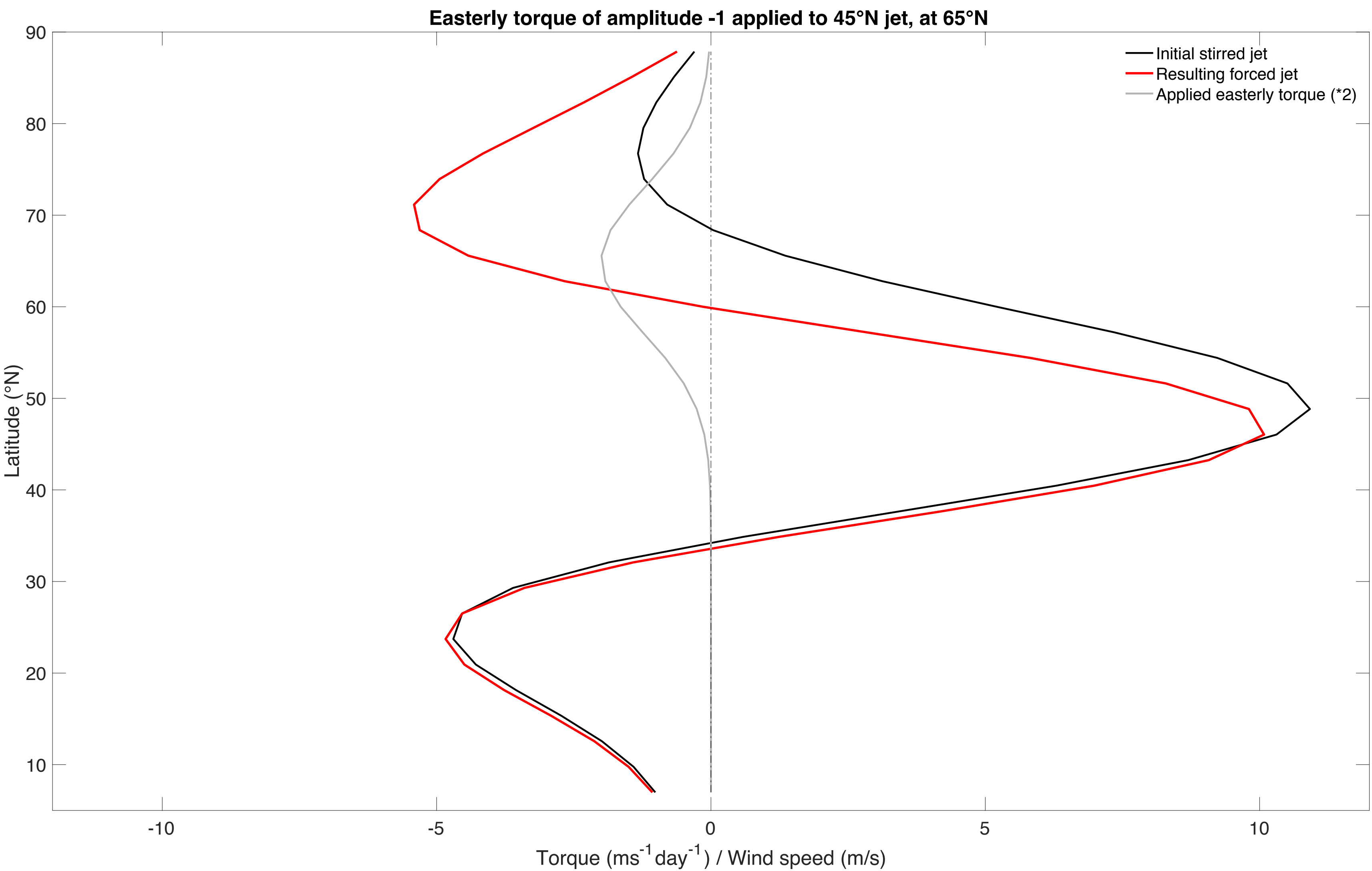
Models with higher latitude jets shift further

Models with higher latitude jets shift further

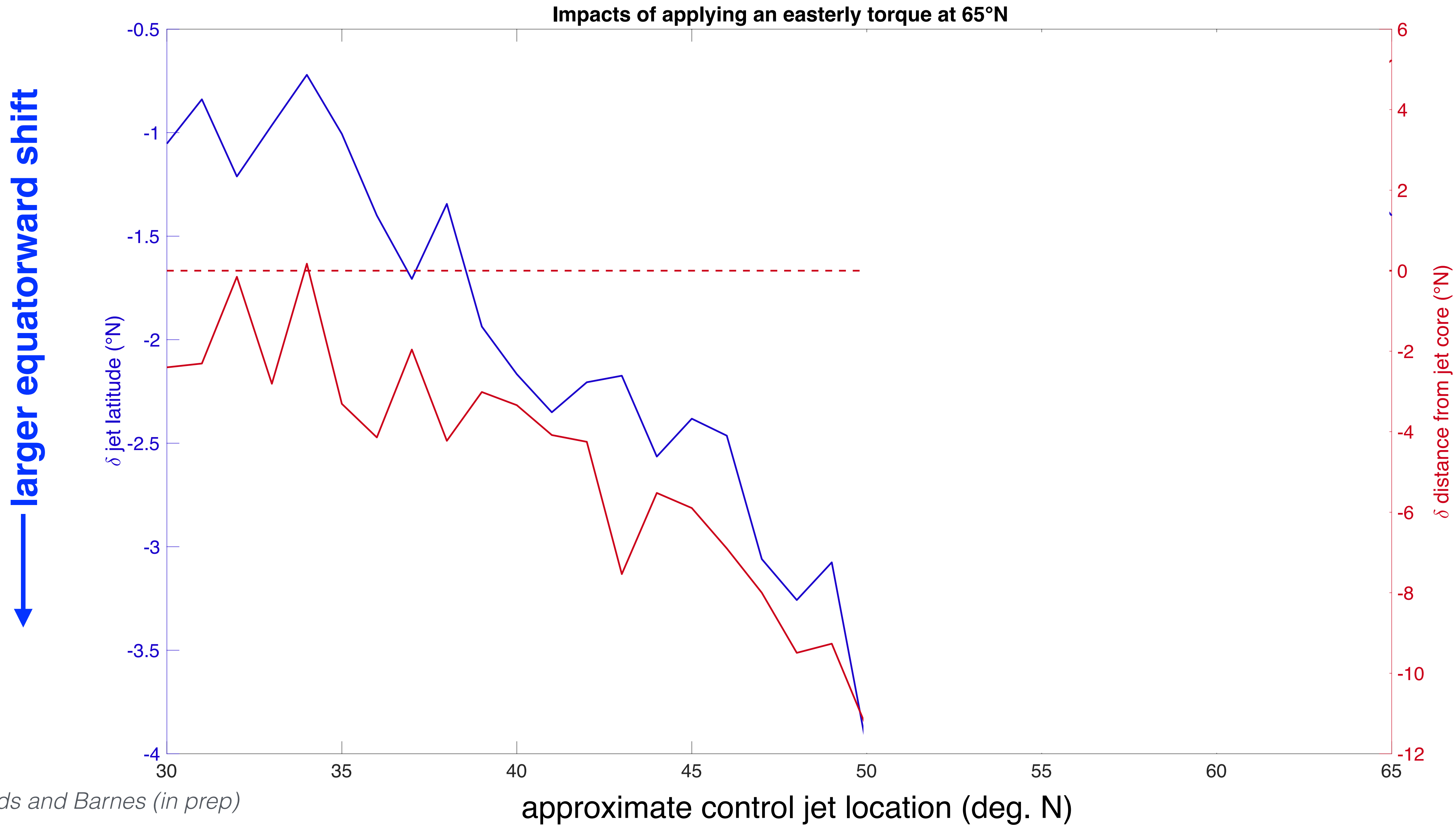


Barnes and Simpson (submitted)

Barotropic Model

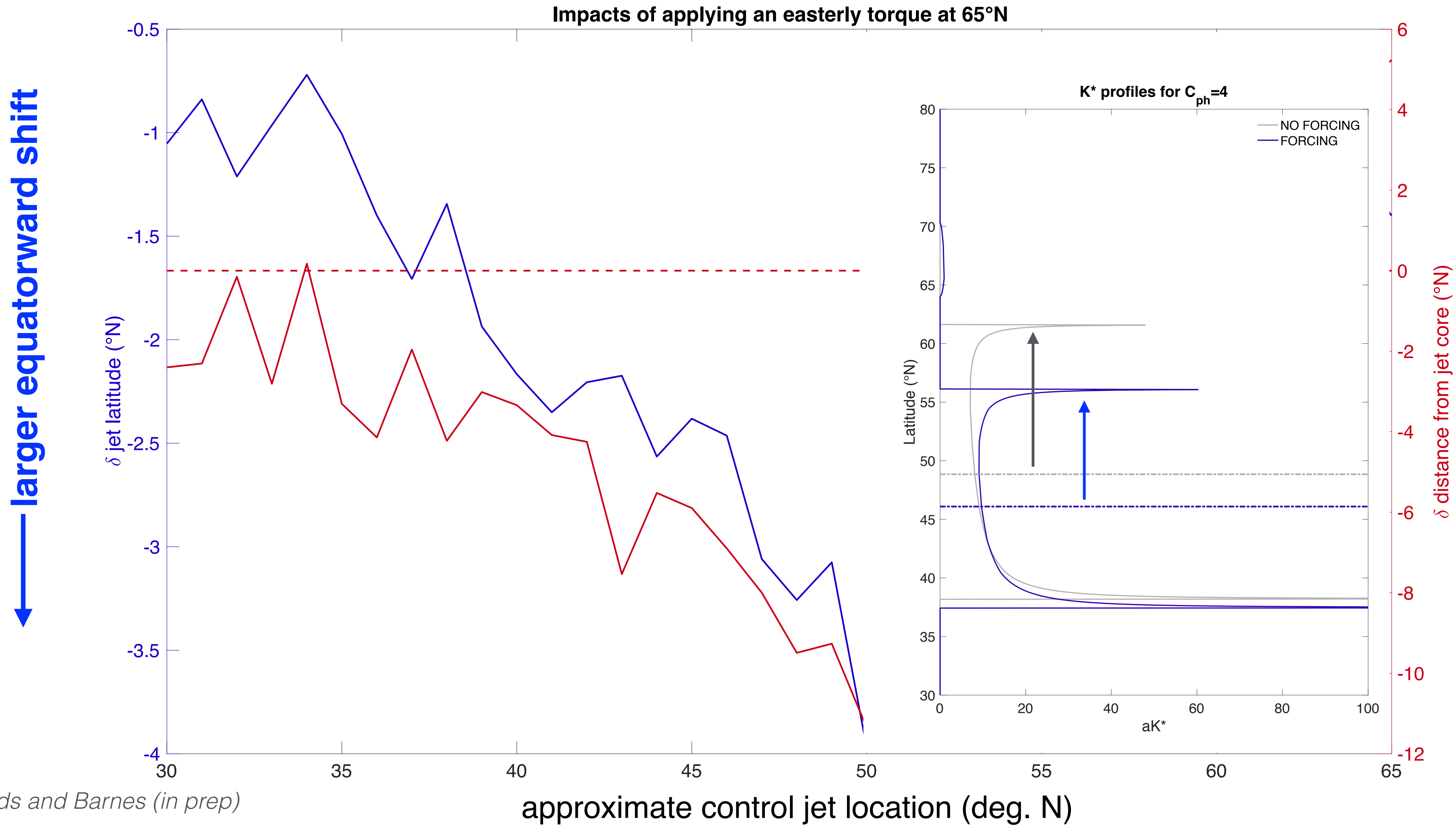


Barotropic Model



Ronalds and Barnes (in prep)

Barotropic Model

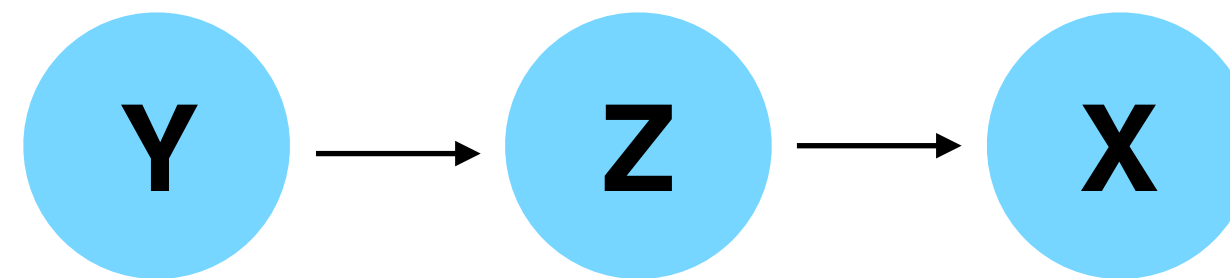


Ronalds and Barnes (in prep)

Causal Effect Networks/Bayesian Networks

Conditional Independence of X & Y

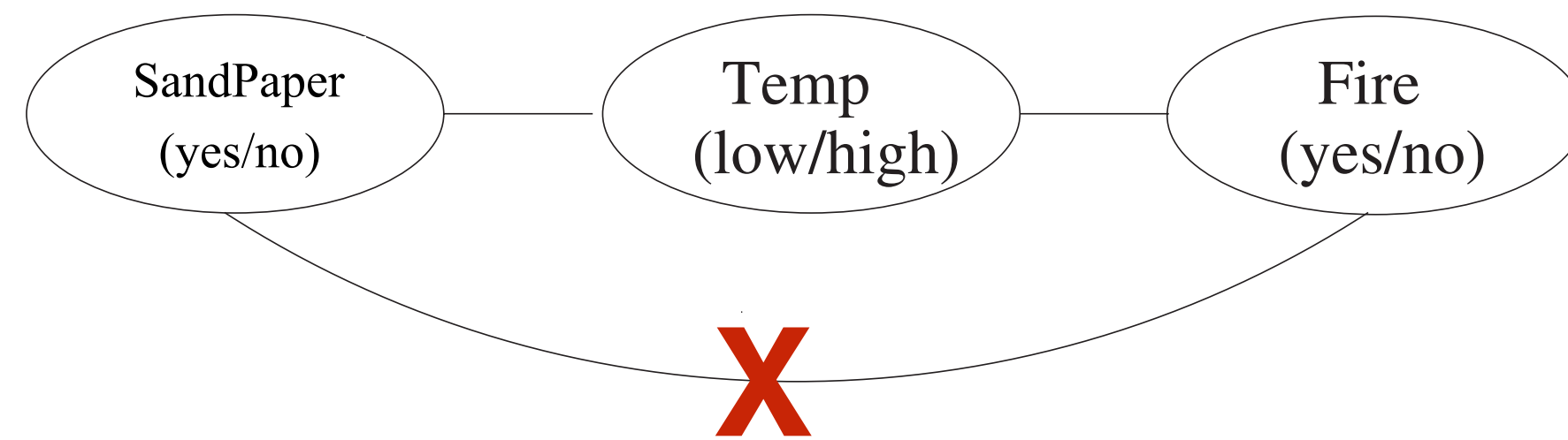
$$\mathbf{Pr}(X = x|Y = y, Z = z) = \mathbf{Pr}(X = x|Z = z)$$



*high correlation between X and Y only
occurs because of the indirect link via Z*

Causal Effect Networks/Bayesian Networks

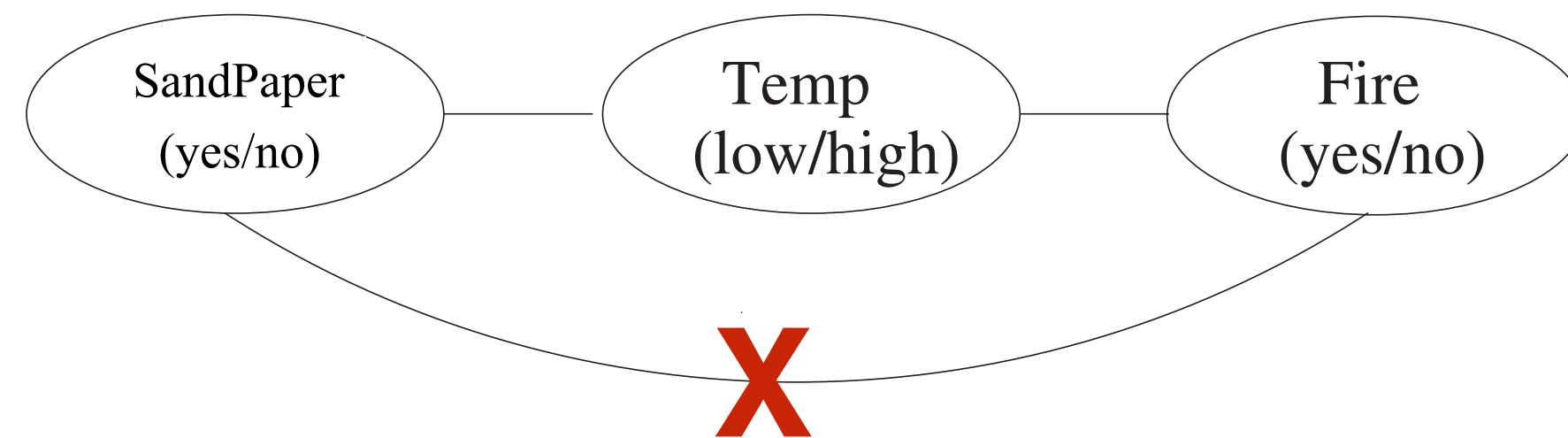
Correlation Graph



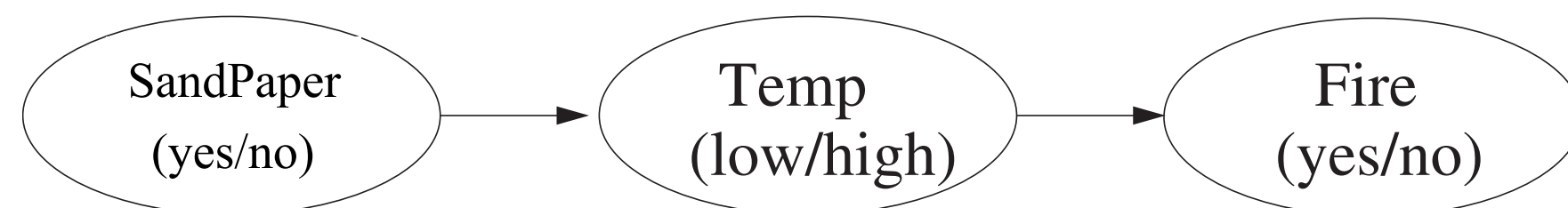
Causal Effect Networks/Bayesian Networks



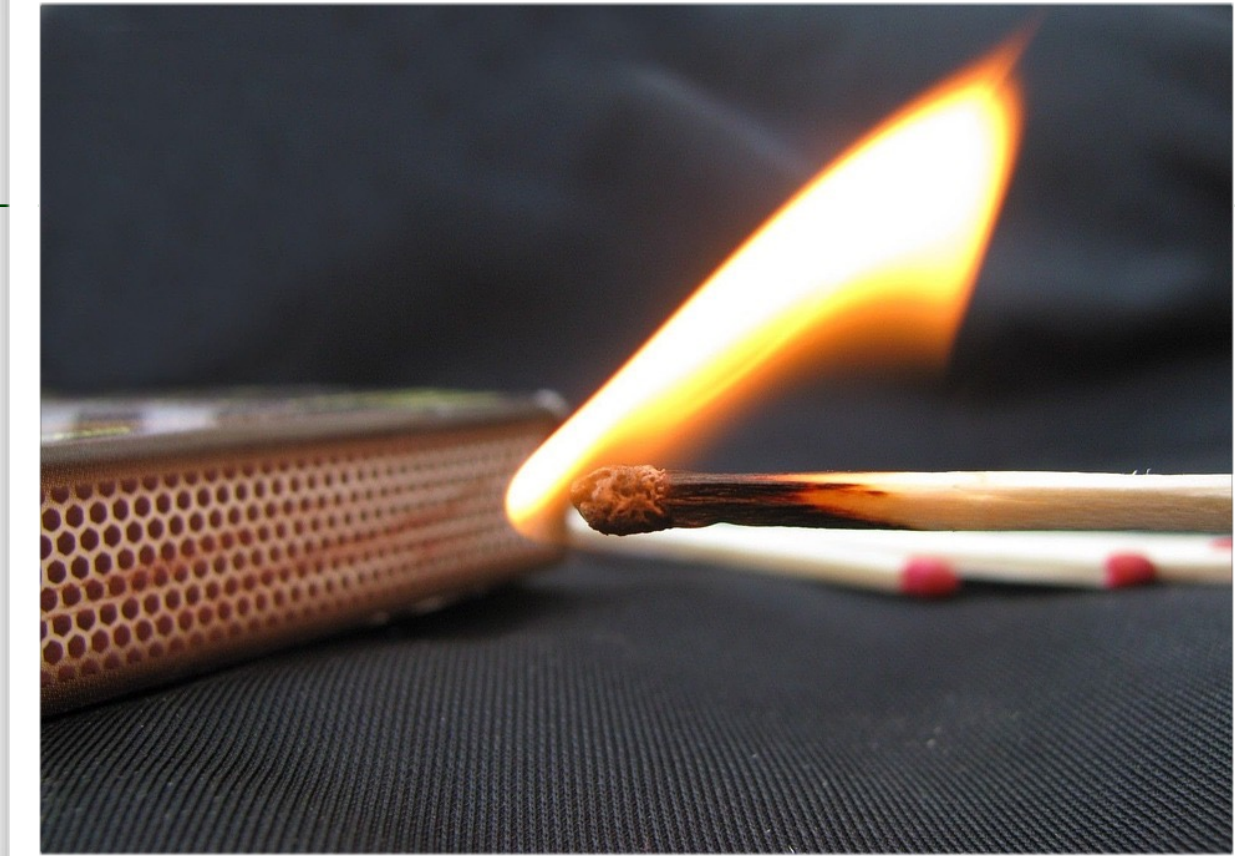
Correlation Graph



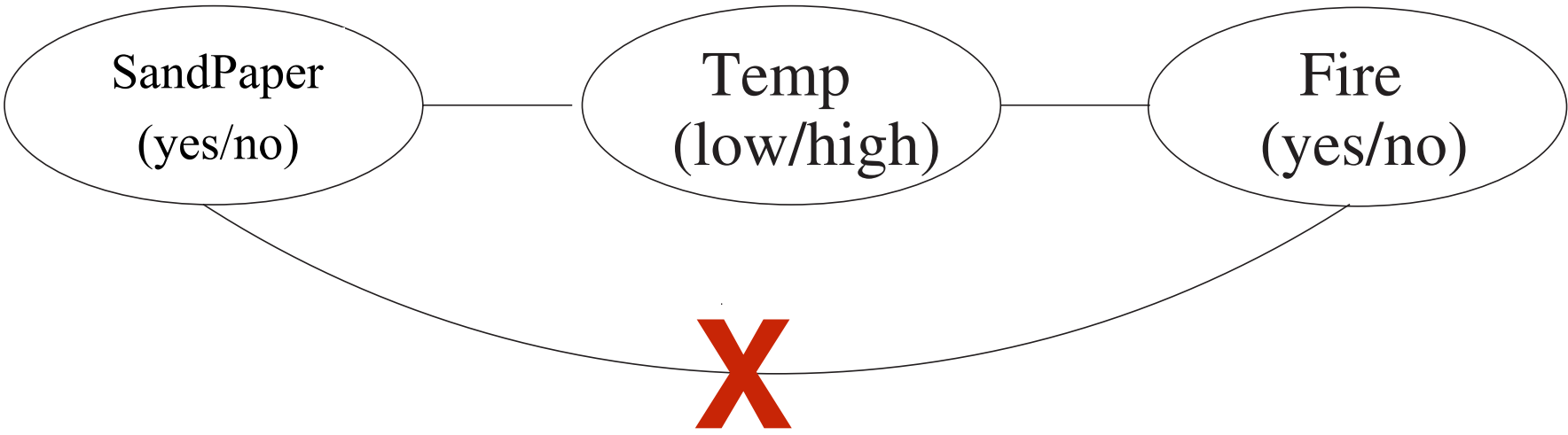
Directed independence graph



Causal Effect Networks/Bayesian Networks



Correlation Graph



Directed independence graph

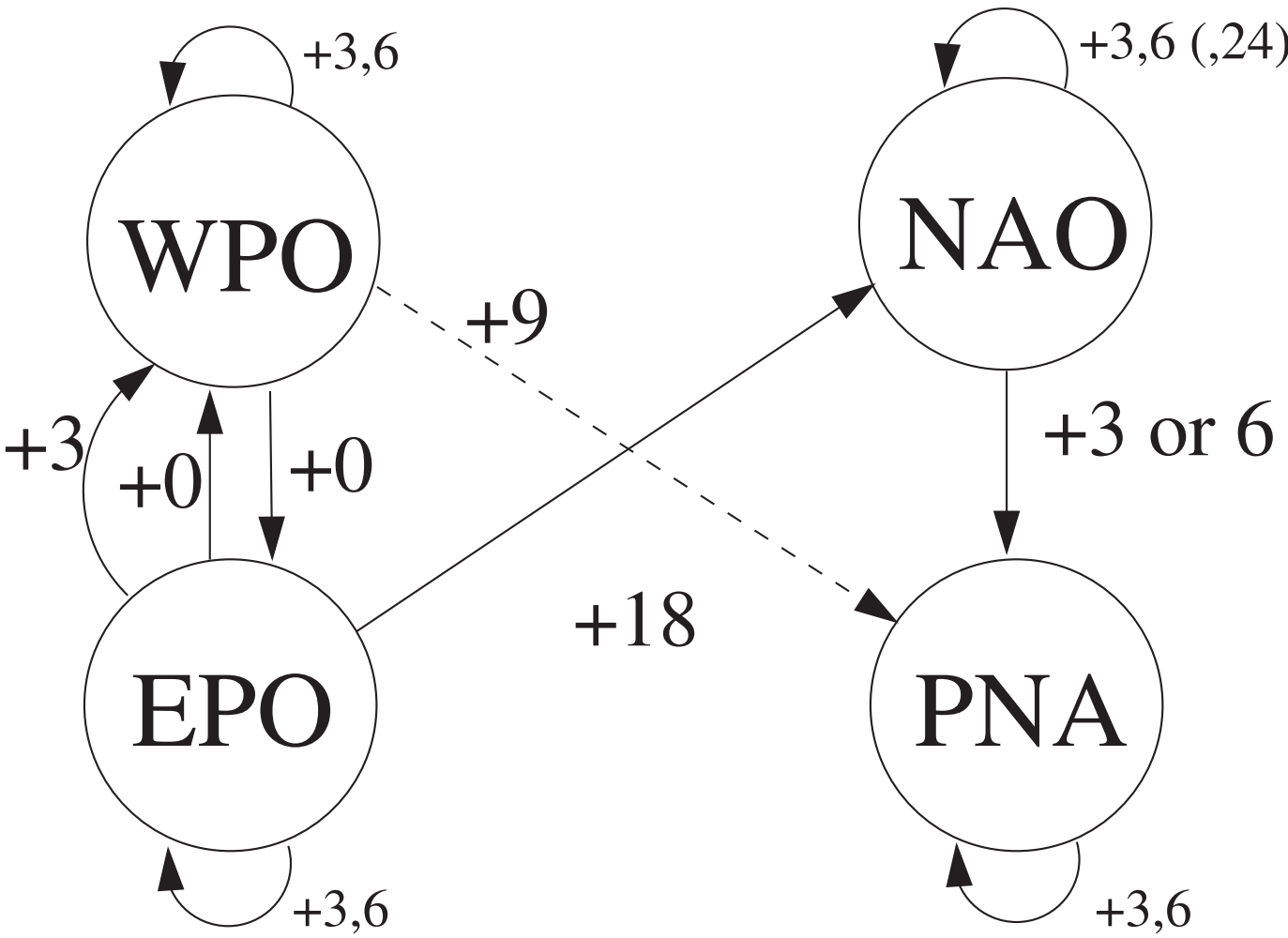
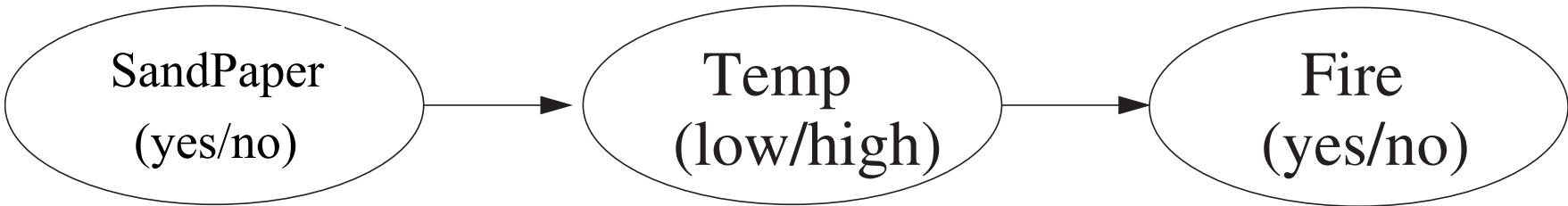
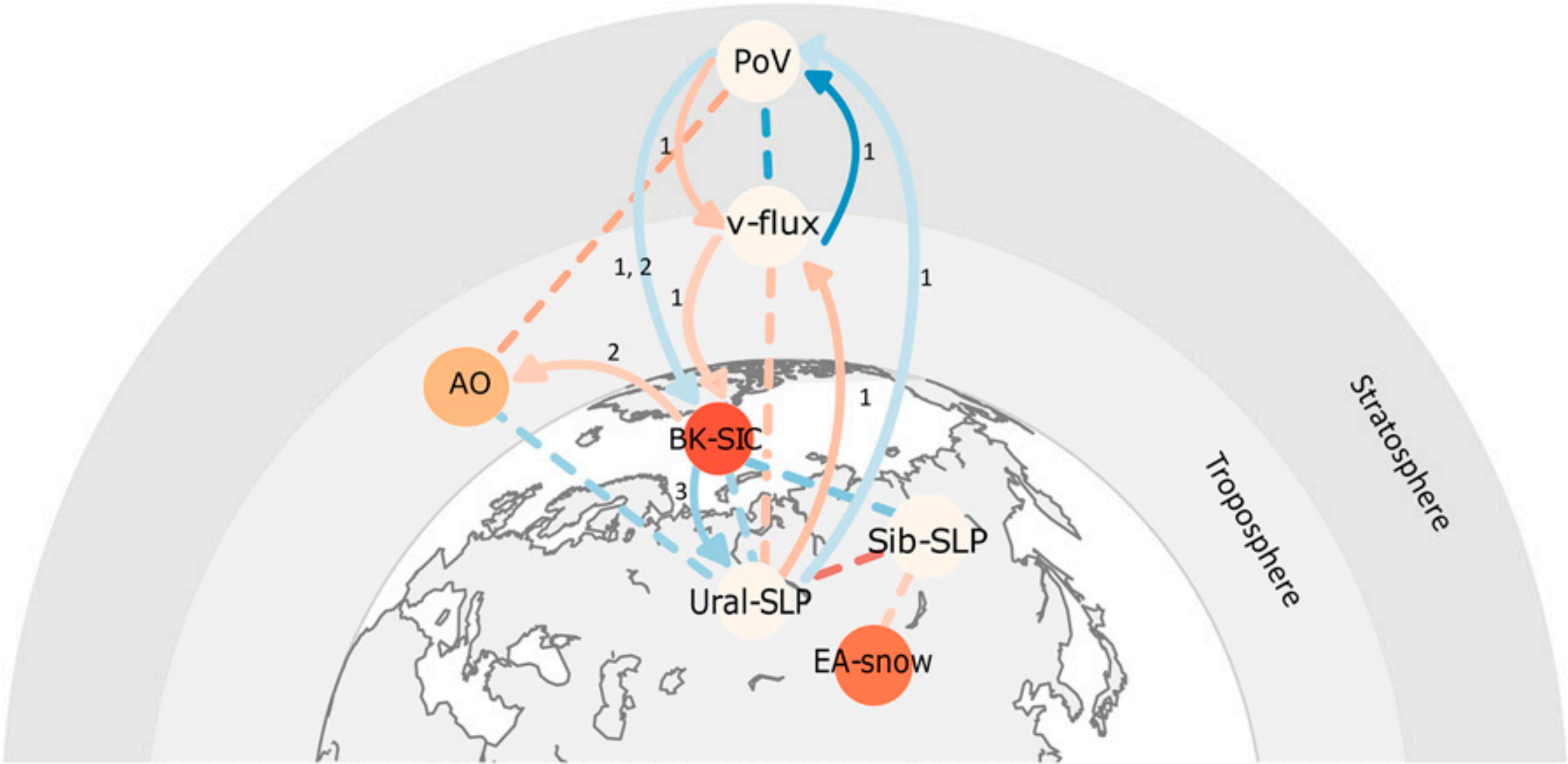


FIG. 7. Summary graph for $D = 3$ and $\alpha = 0.01$. Strong (medium) strength connections are shown as solid (dashed) arrows with corresponding time delays.

Ebert-Uphoff and Deng (2012; JCLI)



Using Causal Effect Networks to Analyze Different Arctic Drivers of Midlatitude Winter Circulation

MARLENE KRETSCHMER

Potsdam Institute for Climate Impact Research, and Department of Mathematics, University of Potsdam, Potsdam, Germany

DIM COUMOU

Potsdam Institute for Climate Impact Research, Potsdam, Germany

JONATHAN F. DONGES

Potsdam Institute for Climate Impact Research, Potsdam, Germany, and Stockholm Resilience Centre, Stockholm, Sweden

JAKOB RUNGE

Potsdam Institute for Climate Impact Research, Potsdam, and Department of Physics, Humboldt University Berlin, Berlin, Germany

(Manuscript received 11 September 2015, in final form 22 January 2016)

ABSTRACT

In recent years, the Northern Hemisphere midlatitudes have suffered from severe winters like the extreme 2012/13 winter in the eastern United States. These cold spells were linked to a meandering upper-tropospheric jet stream pattern and a negative Arctic Oscillation index (AO). However, the nature of the drivers behind these circulation patterns remains controversial. Various studies have proposed different mechanisms related to changes in the Arctic, most of them related to a reduction in sea ice concentrations or increasing Eurasian snow cover. Here, a novel type of time series analysis, called causal effect networks (CEN), based on graphical models is introduced to assess causal relationships and their time delays between different processes. The effect of different Arctic actors on winter circulation on weekly to monthly time scales is studied, and robust network patterns are found. Barents and Kara sea ice concentrations are detected to be important external drivers of the midlatitude circulation, influencing winter AO via tropospheric mechanisms and through processes involving the stratosphere. Eurasia snow cover is also detected to have a causal effect on sea level pressure in Asia, but its exact role on AO remains unclear. The CEN approach presented in this study overcomes some difficulties in interpreting correlation analyses, complements model experiments for testing hypotheses involving teleconnections, and can be used to assess their validity. The findings confirm that sea ice concentrations in autumn in the Barents and Kara Seas are an important driver of winter circulation in the midlatitudes.

1. Introduction

The recent cold winters in North America and Eurasia were characterized by a meandering jet stream pattern that allowed cold Arctic air to reach lower latitudes (Cohen et al. 2014b). Moreover, these winters were

dominated by a negative phase of the Arctic Oscillation index (AO), which is usually associated with pronounced meridional wind patterns, whereas in a positive AO phase strong zonal flow dominates the wind field. Although a negative AO and meandering flow patterns have been linked to surface extremes (Thompson 2001; Coumou et al. 2014; Screen and Simmonds 2014), it is intensively discussed what the mechanisms behind AO variability are.

Classical atmosphere dynamic theories relate a meandering jet stream structure to above-normal sea

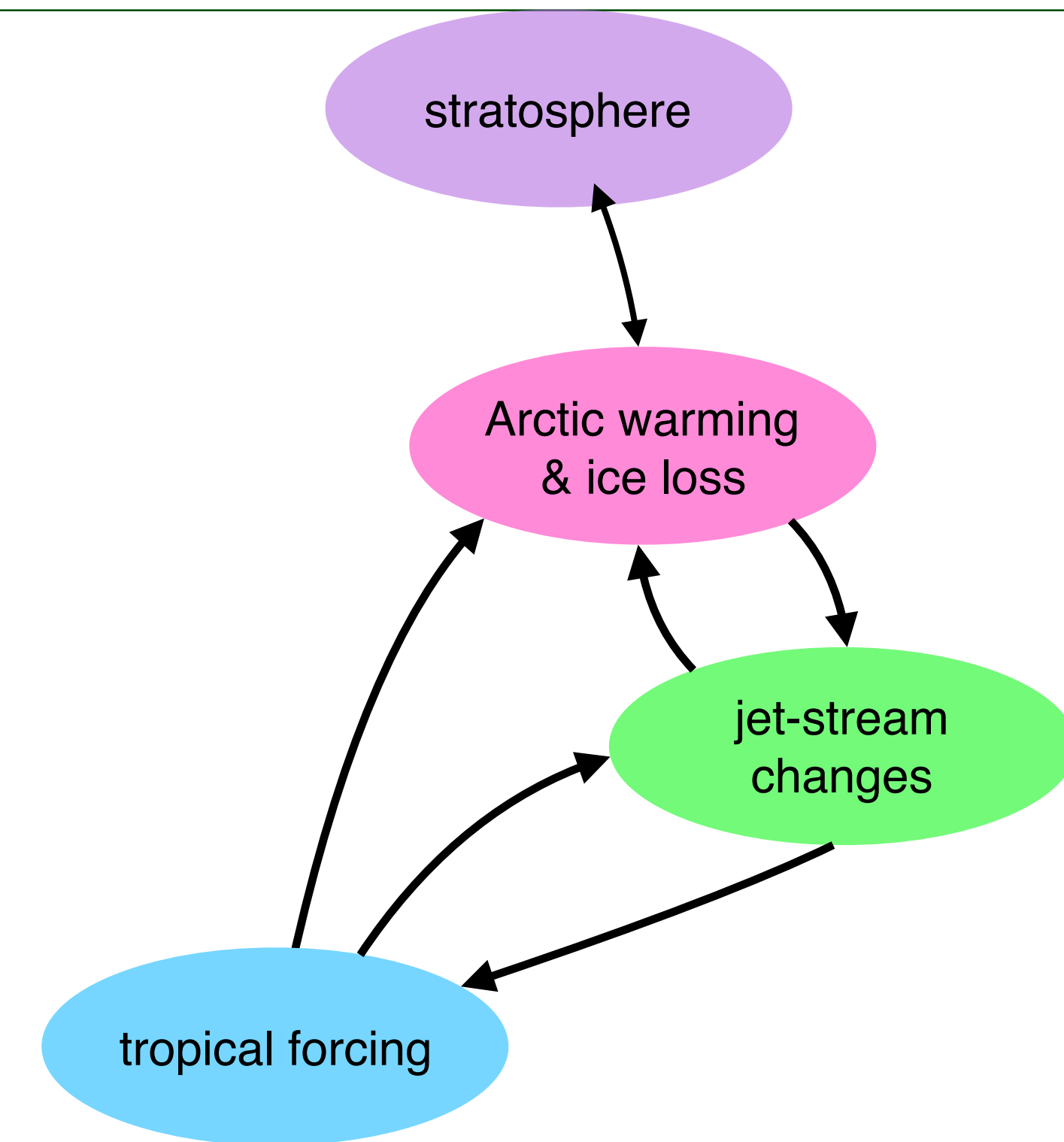
Corresponding author address: M. Kretschmer, Telegrafenberg A62, 14473 Potsdam, Germany.
E-mail: kretschmer@pik-potsdam.de

Another tool for our toolbox

Lagged Regression/Correlations

Targeted Model Experiments

Forecasting Approach



Another tool for our toolbox

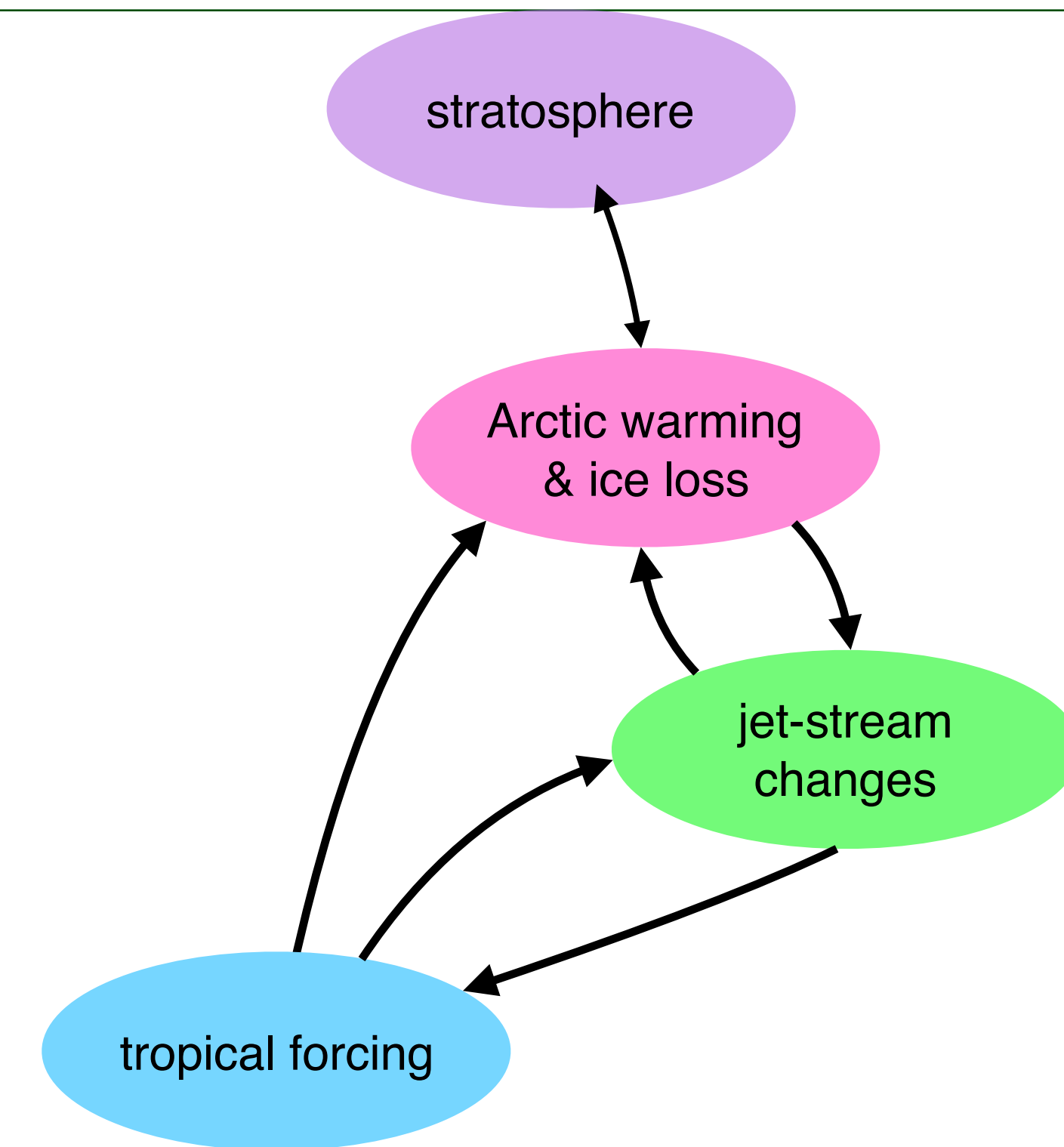
Lagged Regression/Correlations

Targeted Model Experiments

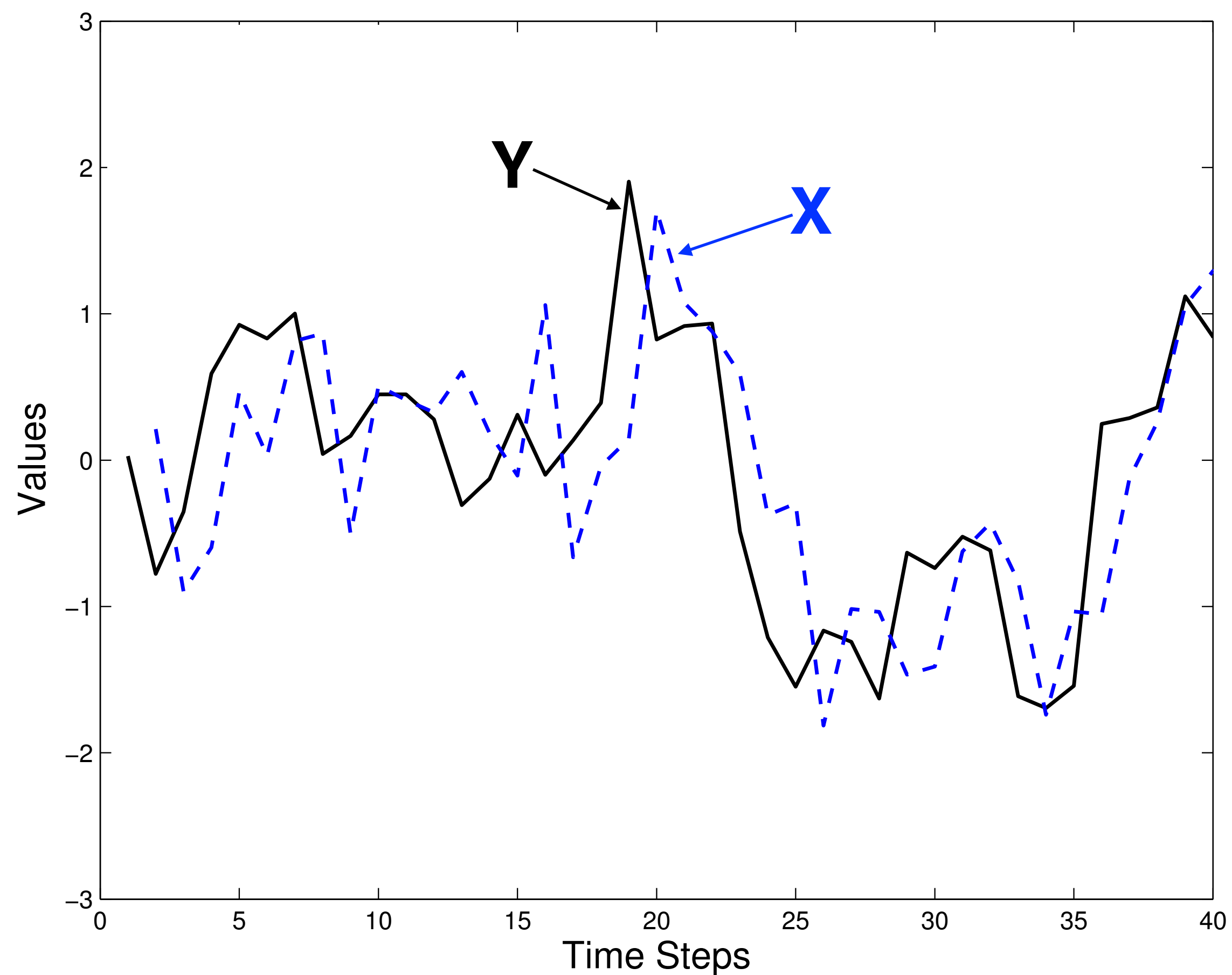
Forecasting Approach

Causal Discovery Techniques

- ✓ robust definitions of causality
- ✓ ties to forecasting/prediction
- ✓ does not suffer from the autocorrelation issues of lagged regression
- ✓ augments targeted model studies
- ✓ **can quantify strong pathways in observations and directly compare to model simulations**
- ✓ **puts pathways in context relative to other drivers and allows for feedbacks**



Lagged regression



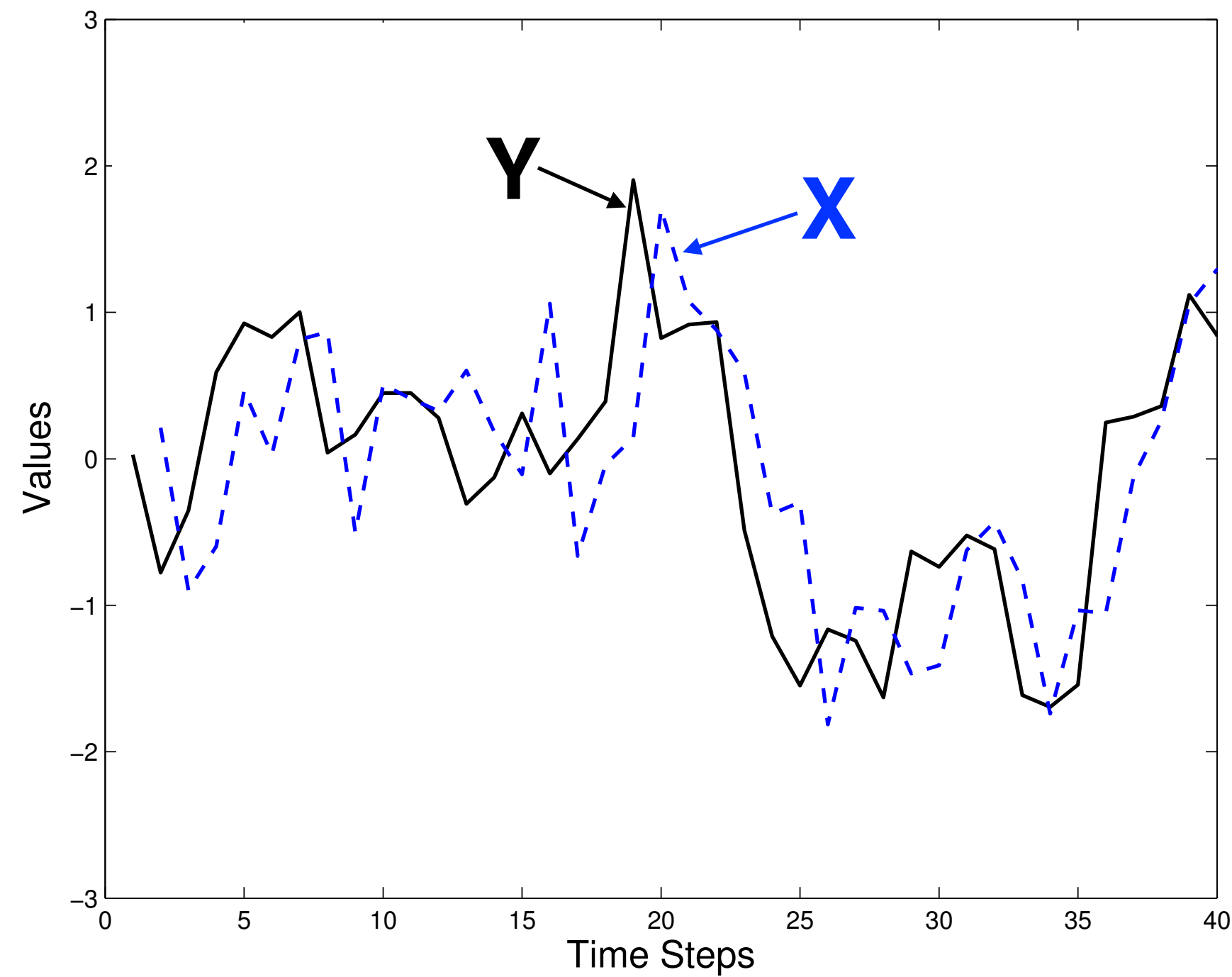
$$Y_t = \overset{\text{memory in } Y}{\alpha} Y_{t-1} + \beta_t$$

red noise (AR1)

$$X_t = Y_{t-1} + \underset{\text{noise in } X}{\epsilon_t}$$

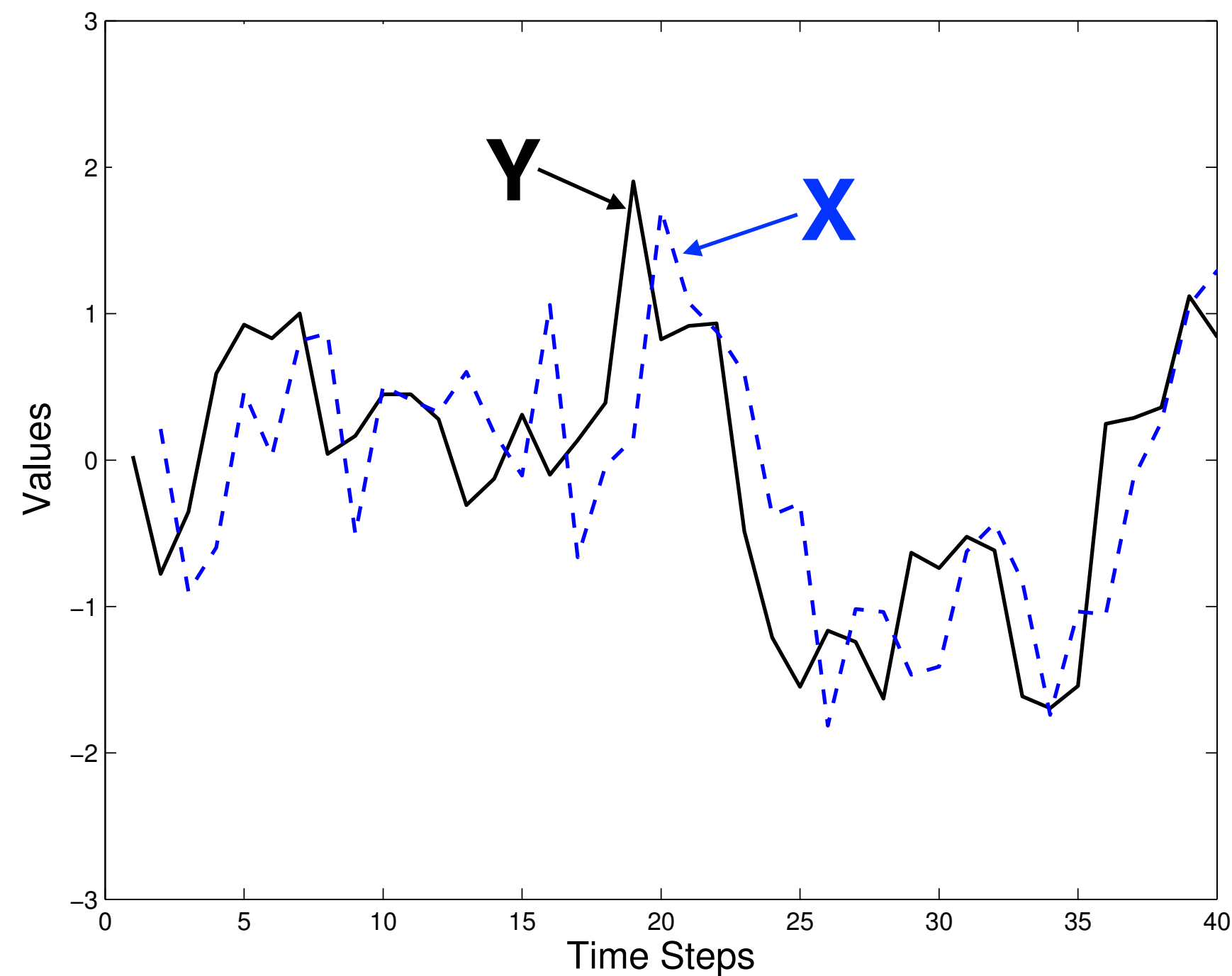
Y drives X one day later
+
noise

Lagged regression



- Pretend we don't know the relationship between X and Y
- **Step 1:** test whether $Y \rightarrow X$ **[right]**
- **Step 2:** test whether $X \rightarrow Y$ **[wrong]**

Lagged regression

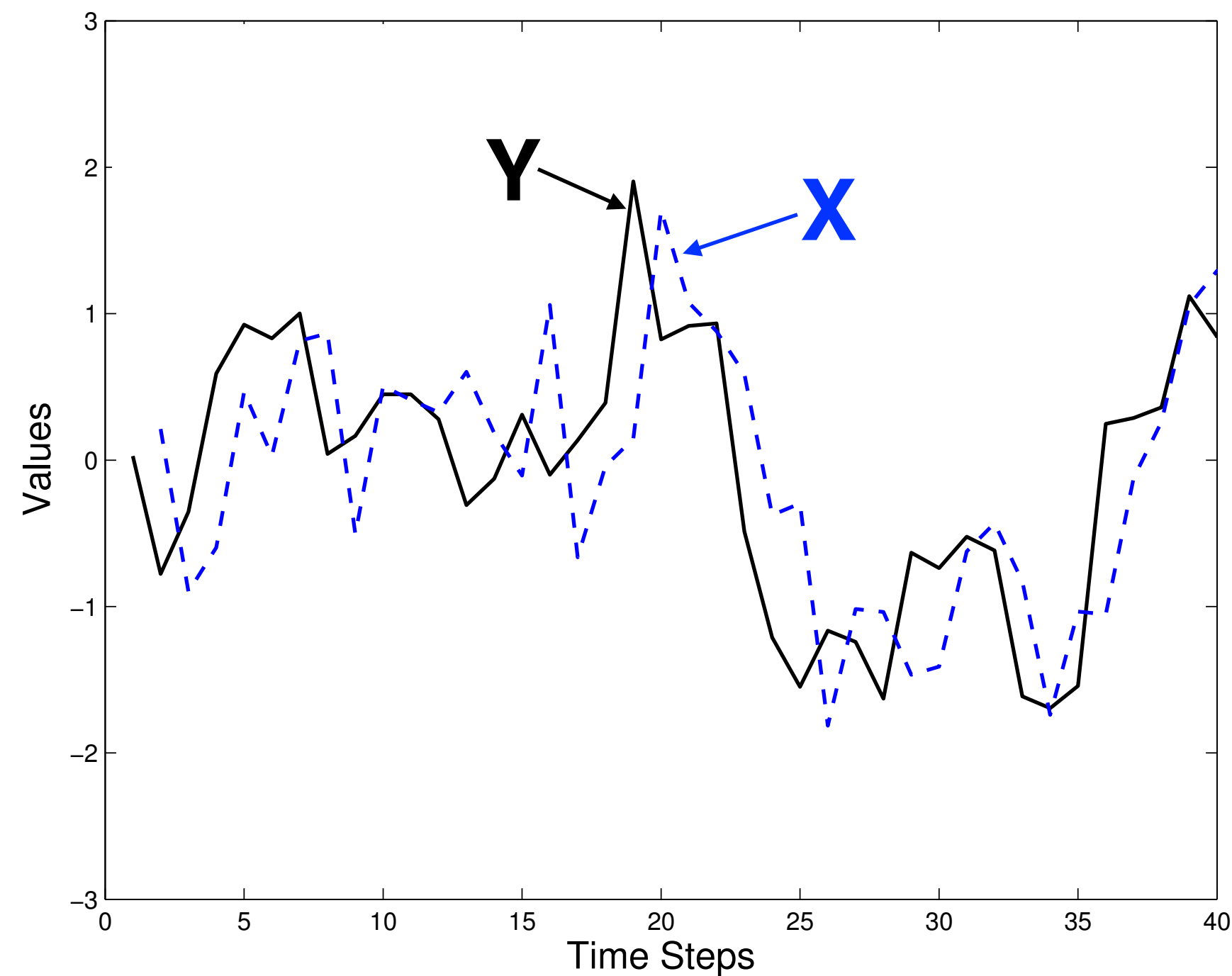


- Pretend we don't know the relationship between X and Y
- **Step 1:** test whether $Y \rightarrow X$ **[right]**
- **Step 2:** test whether $X \rightarrow Y$ **[wrong]**

$$X_t = bY_{t-1} + \epsilon_t$$

$$Y_t = bX_{t-1} + \epsilon_t$$

Lagged regression



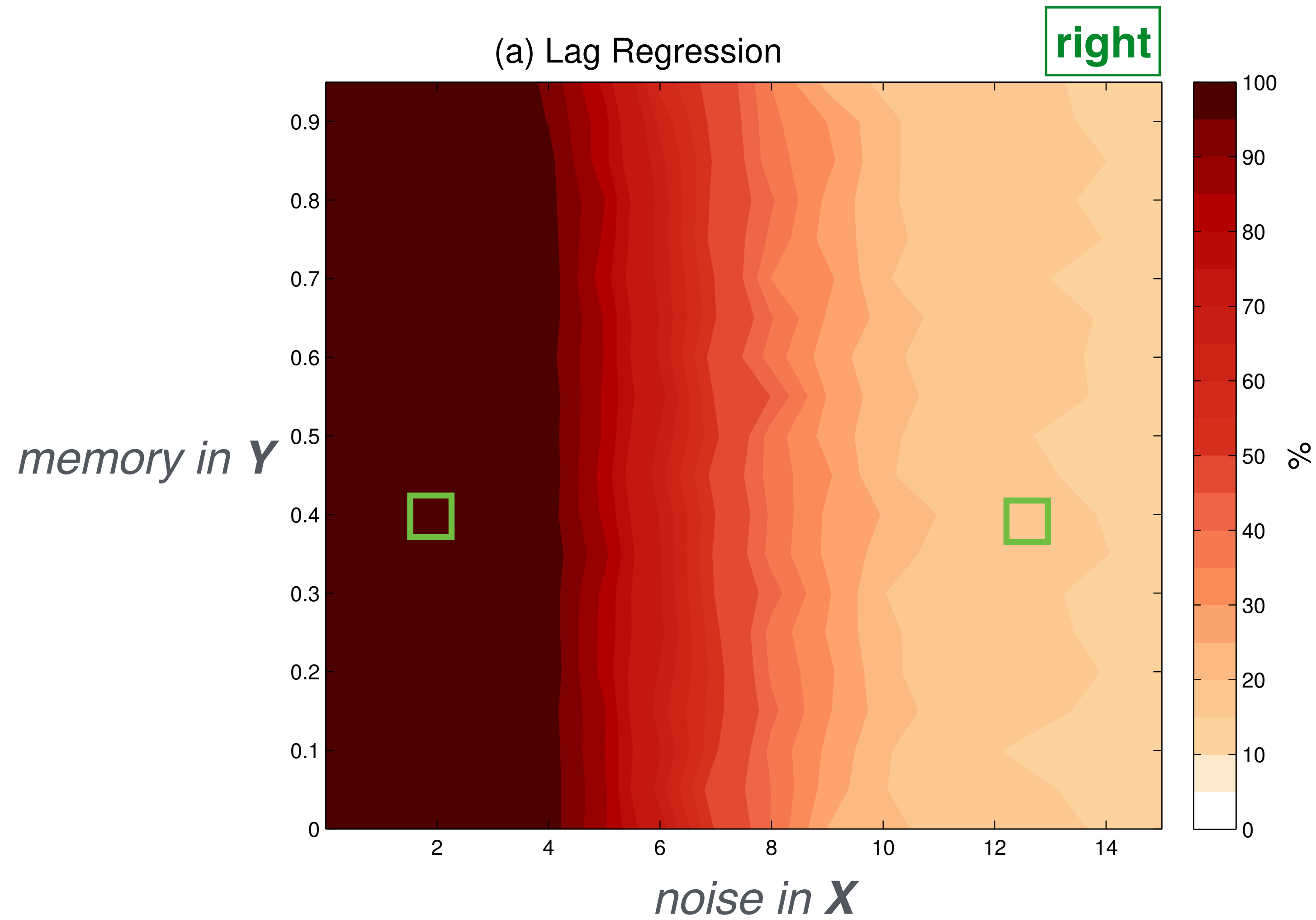
- Pretend we don't know the relationship between X and Y
- **Step 1:** test whether $Y \rightarrow X$ **[right]**
- **Step 2:** test whether $X \rightarrow Y$ **[wrong]**

$$X_t = bY_{t-1} + \epsilon_t$$

how much?

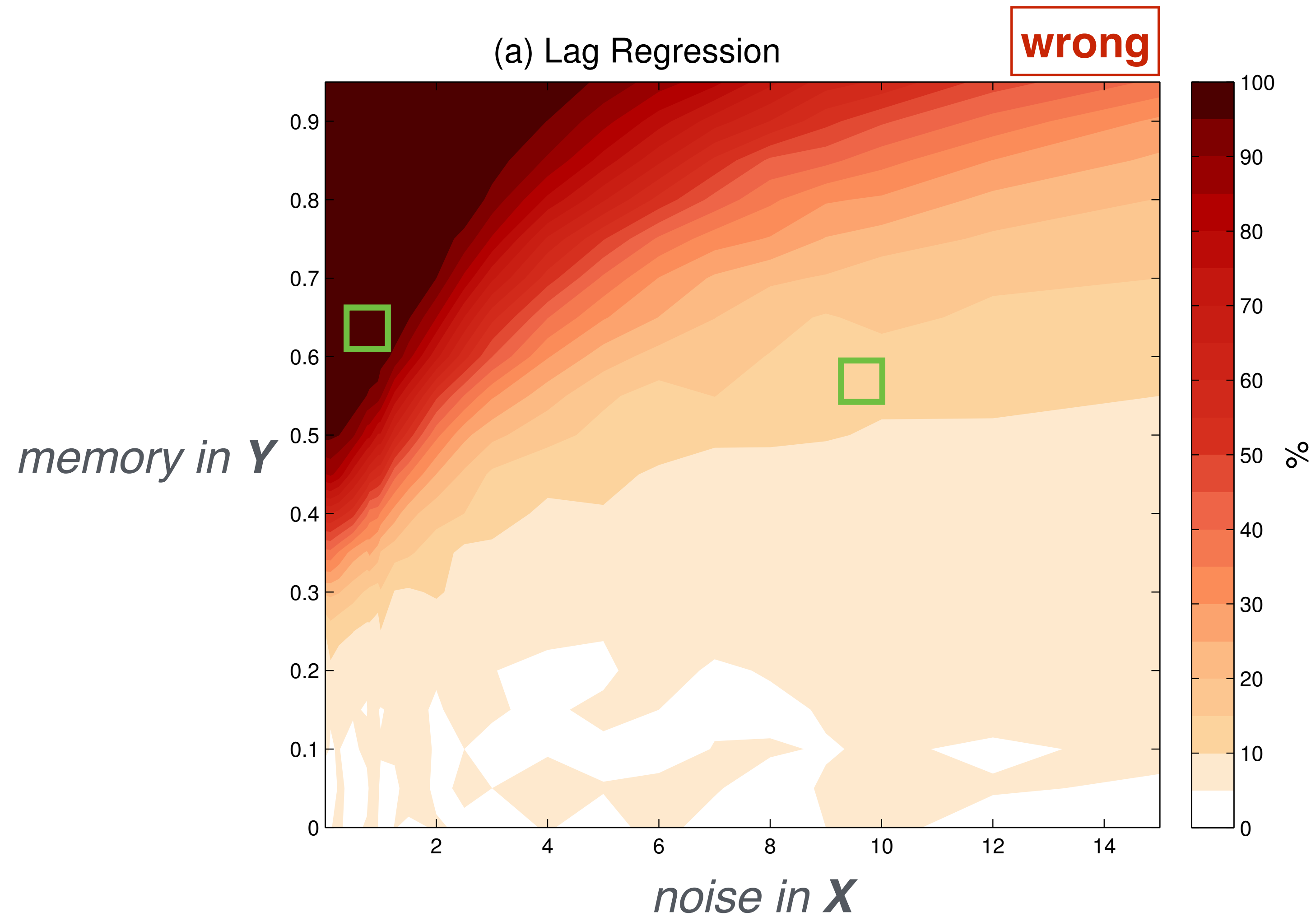
$$Y_t = bX_{t-1} + \epsilon_t$$

Lagged regression



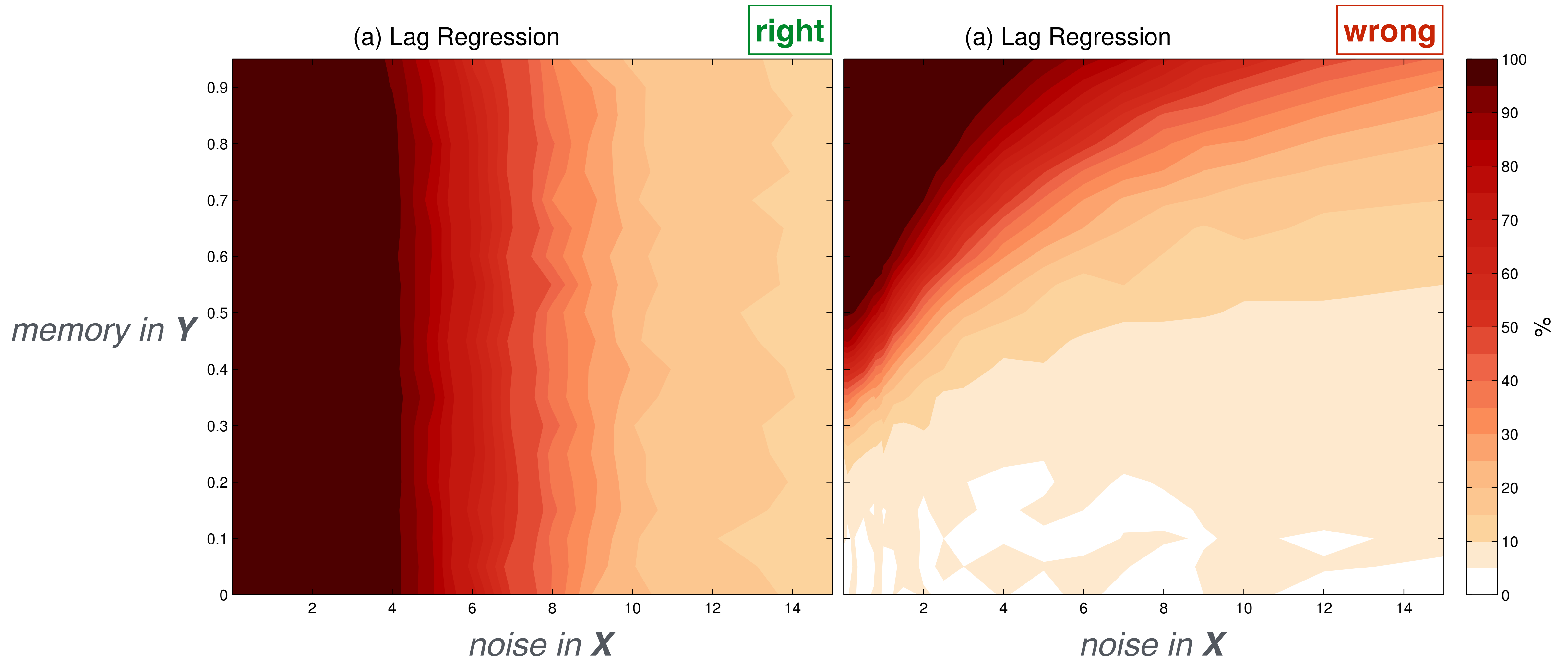
$$X_t = bY_{t-1} + \epsilon_t$$

Lagged regression



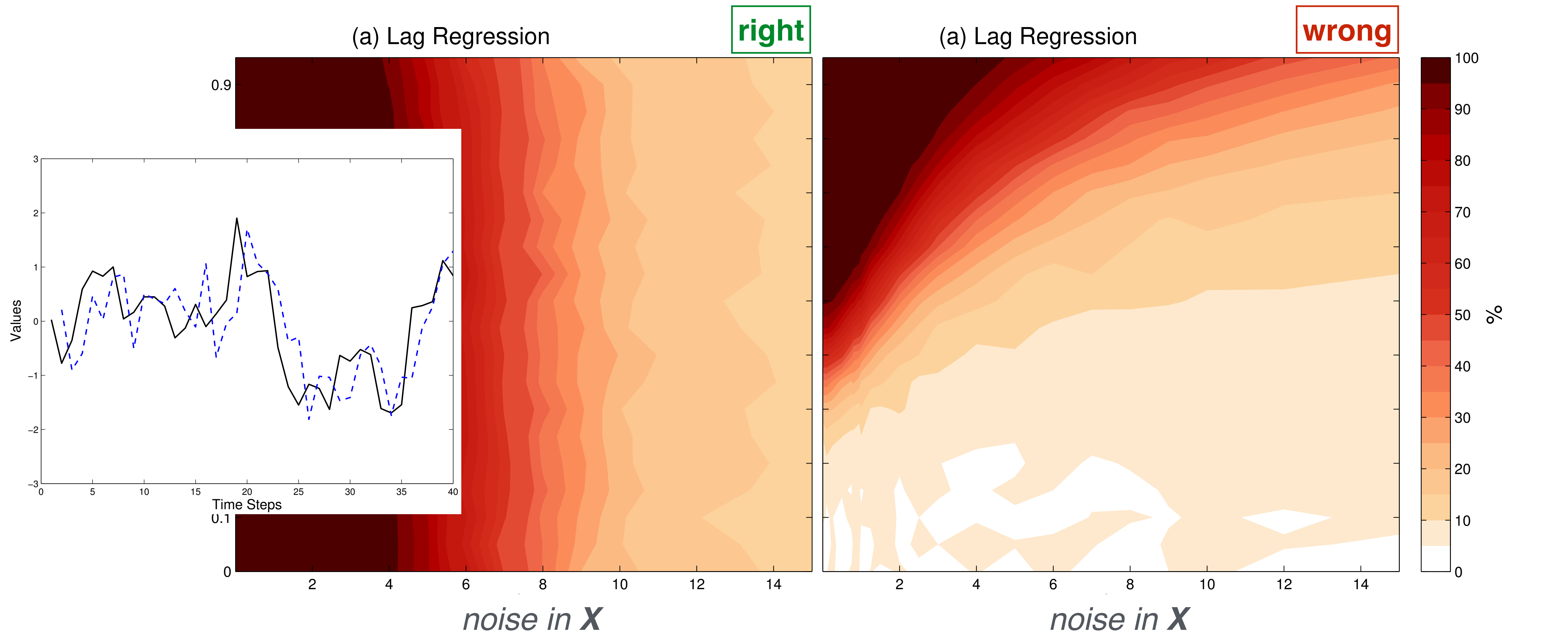
$$Y_t = bX_{t-1} + \epsilon_t$$

Lagged regression



McGraw and Barnes (submitted)

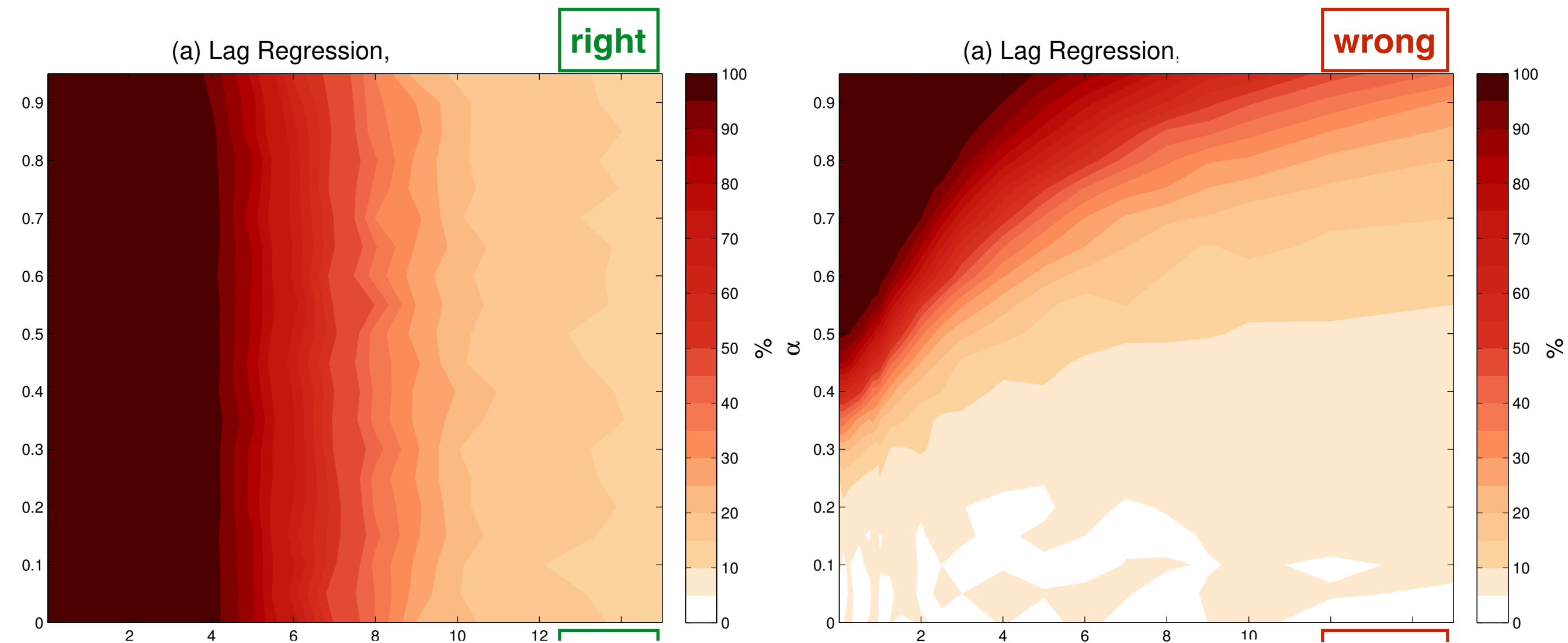
Lagged regression



McGraw and Barnes (submitted)

Granger causality

memory in Y

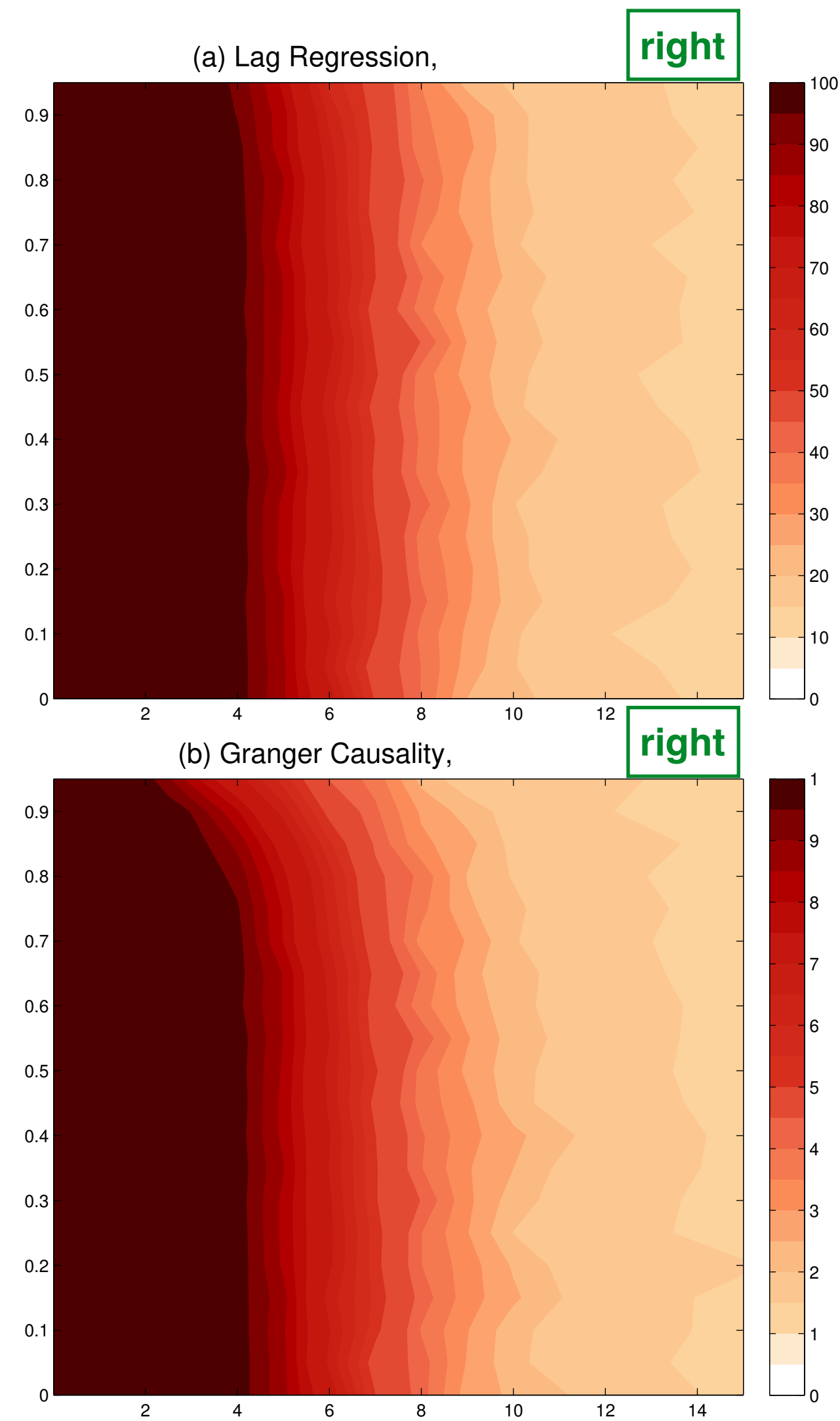


noise in X

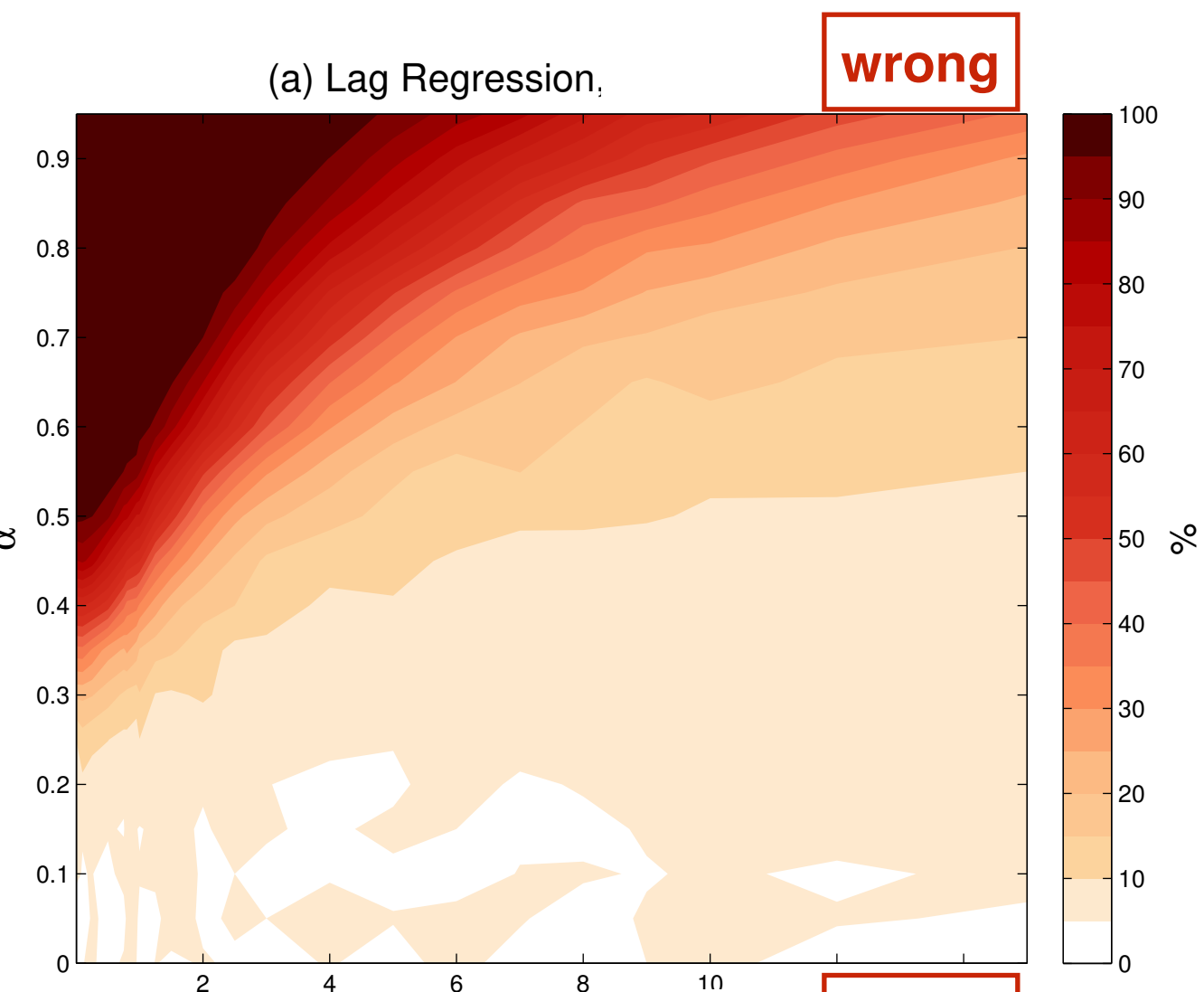
McGraw and Barnes (submitted)

Granger causality

memory in Y



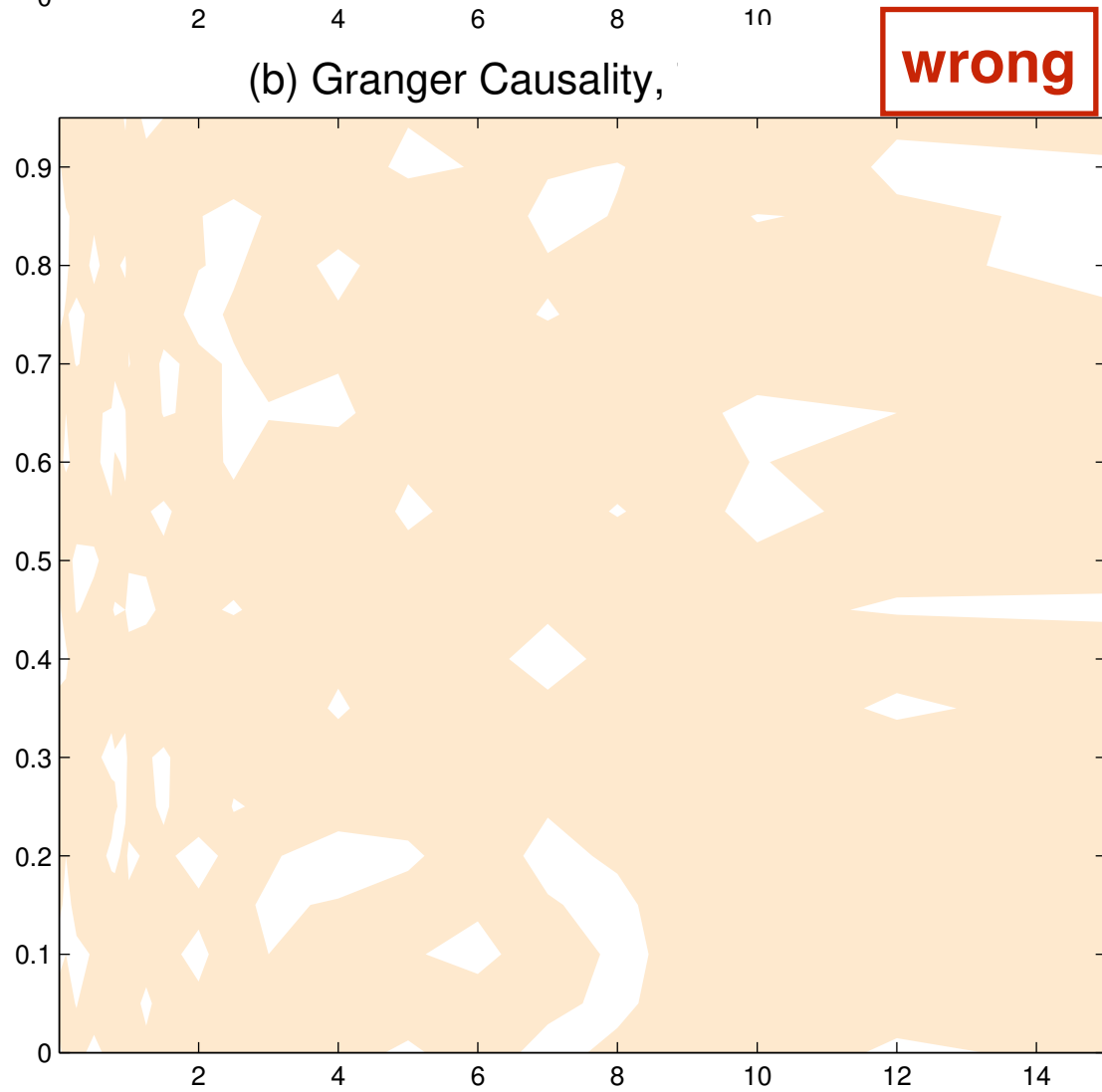
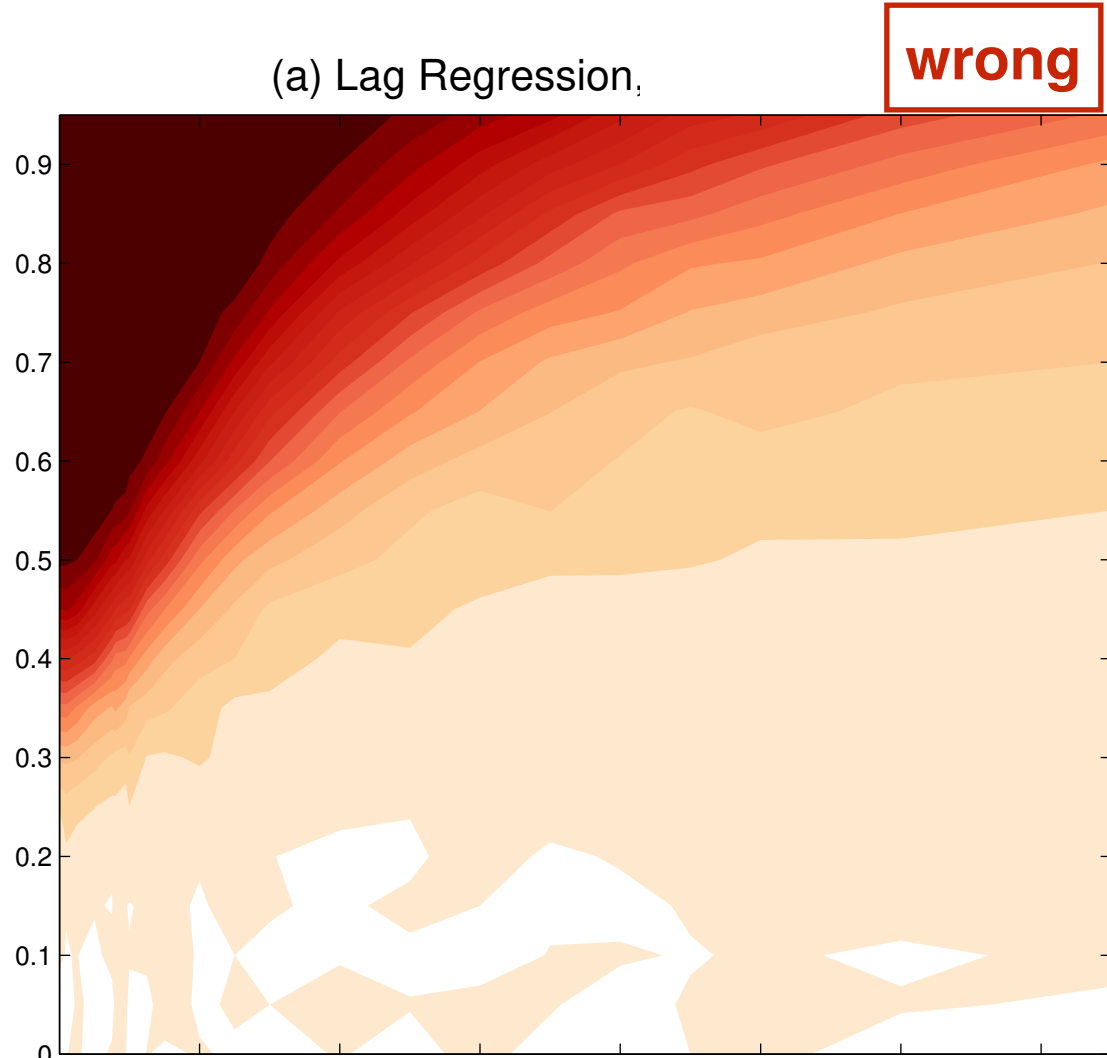
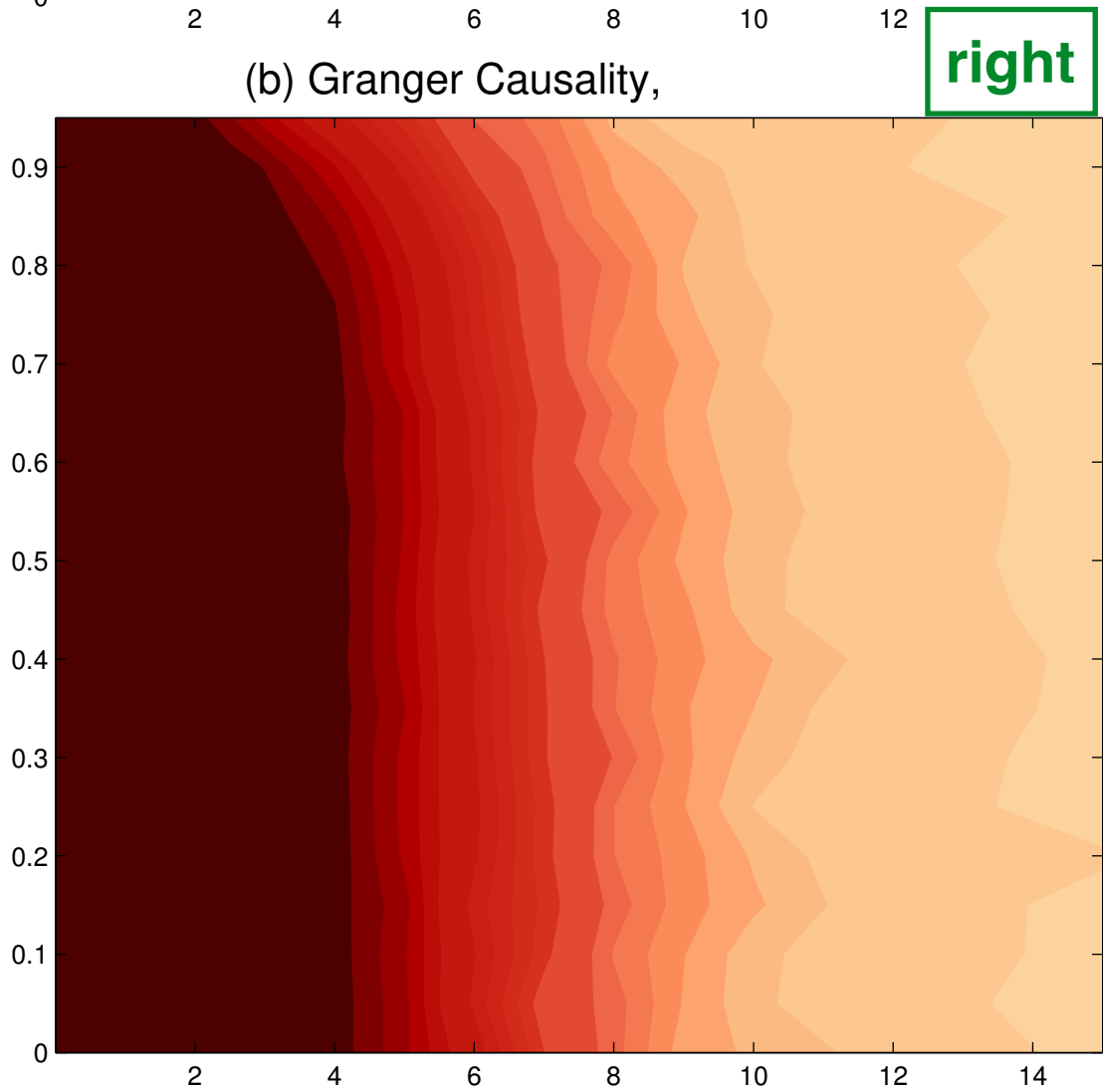
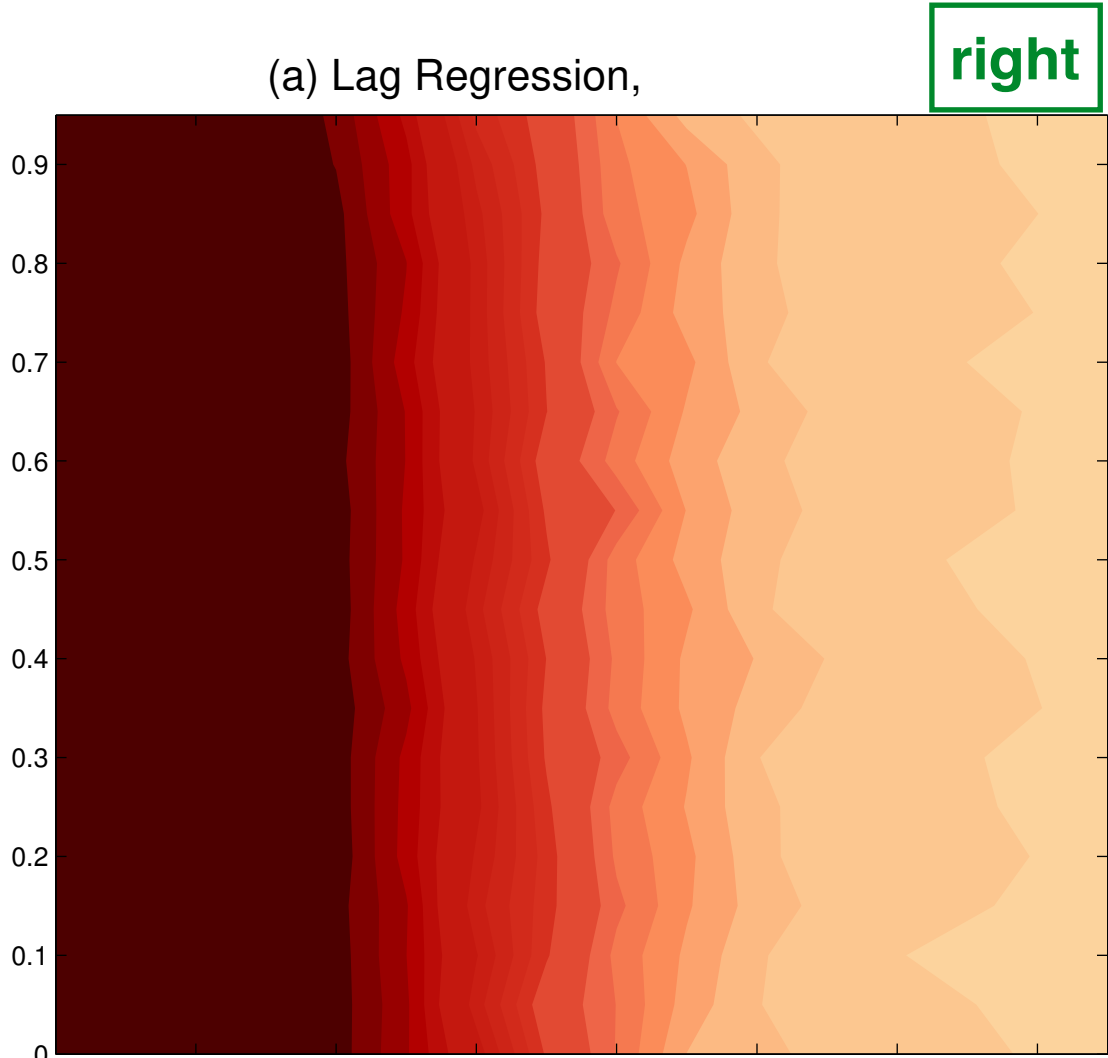
noise in X



McGraw and Barnes (submitted)

Granger causality

memory in Y

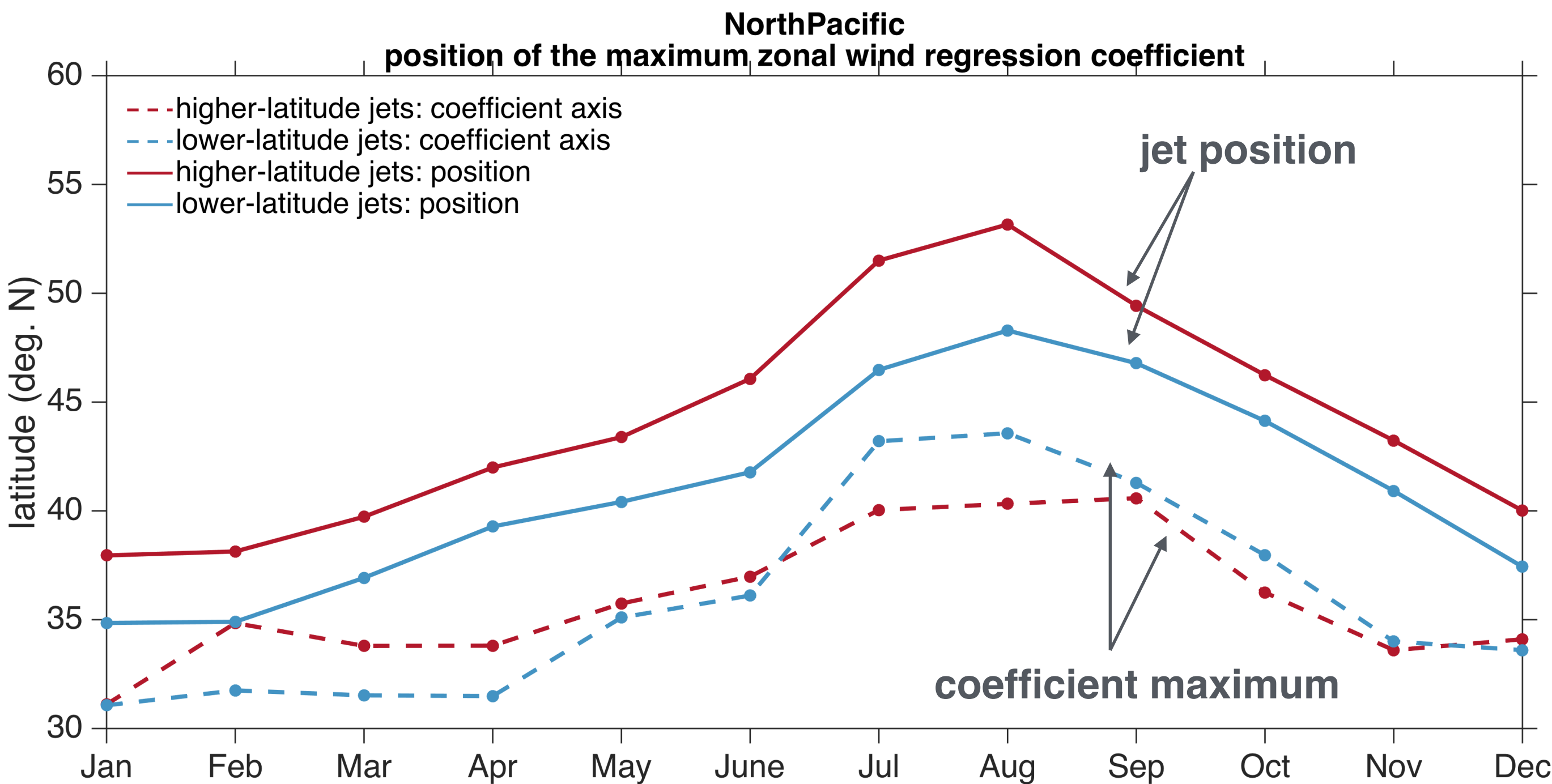


$\alpha = 0.05$

noise in X

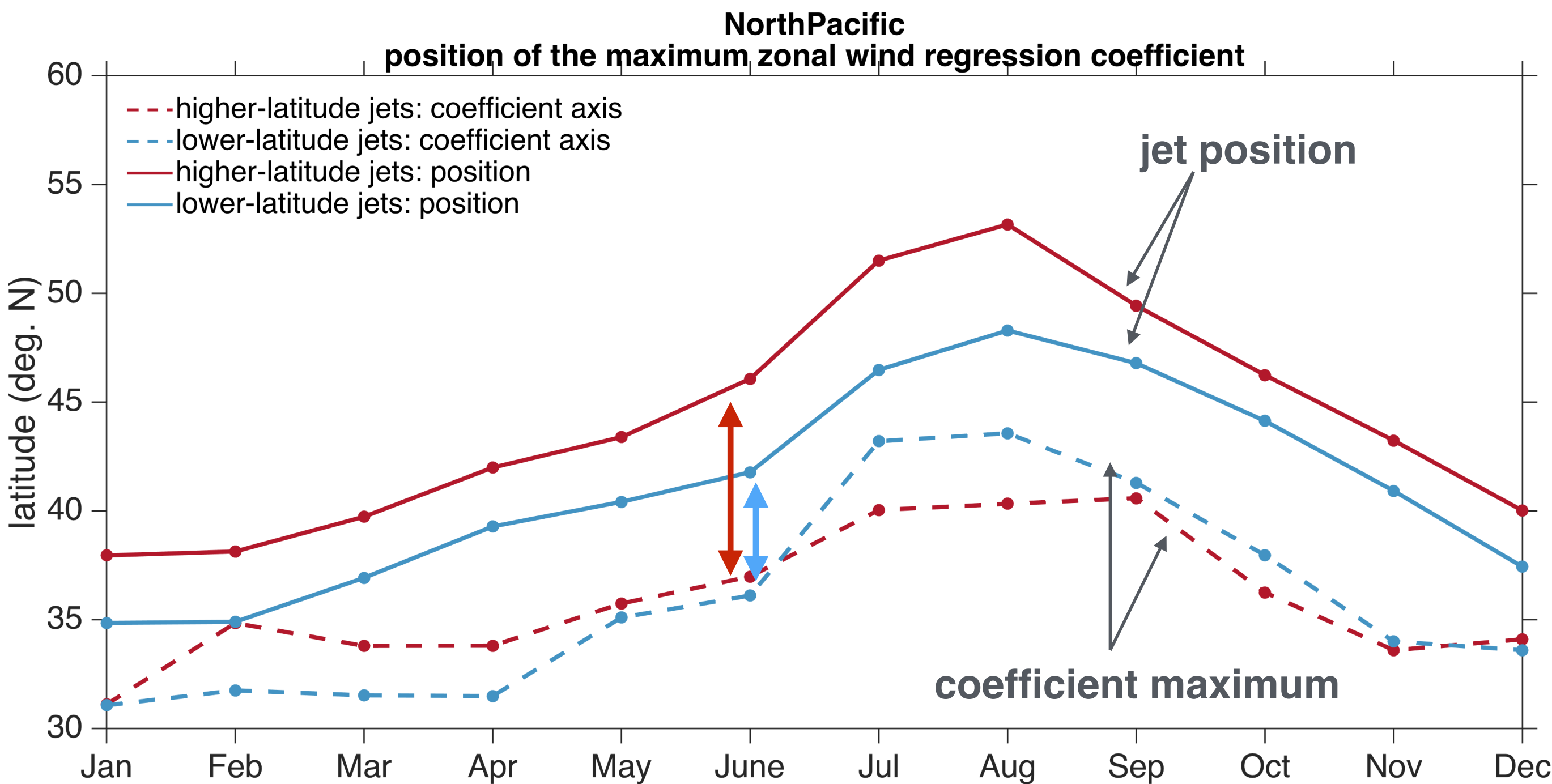
McGraw and Barnes (submitted)

Models with higher latitude jets shift further



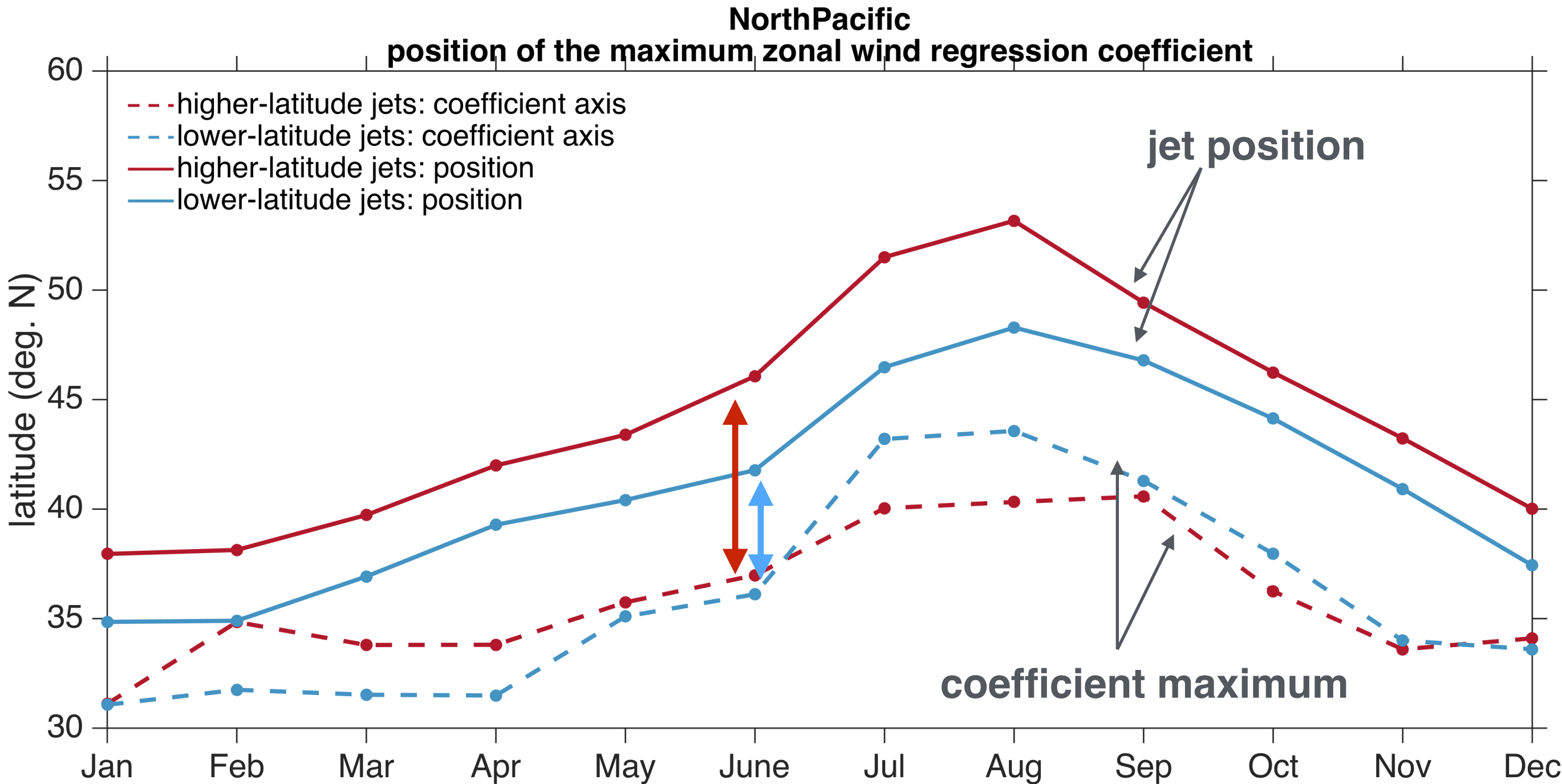
Barnes and Simpson (submitted)

Models with higher latitude jets shift further

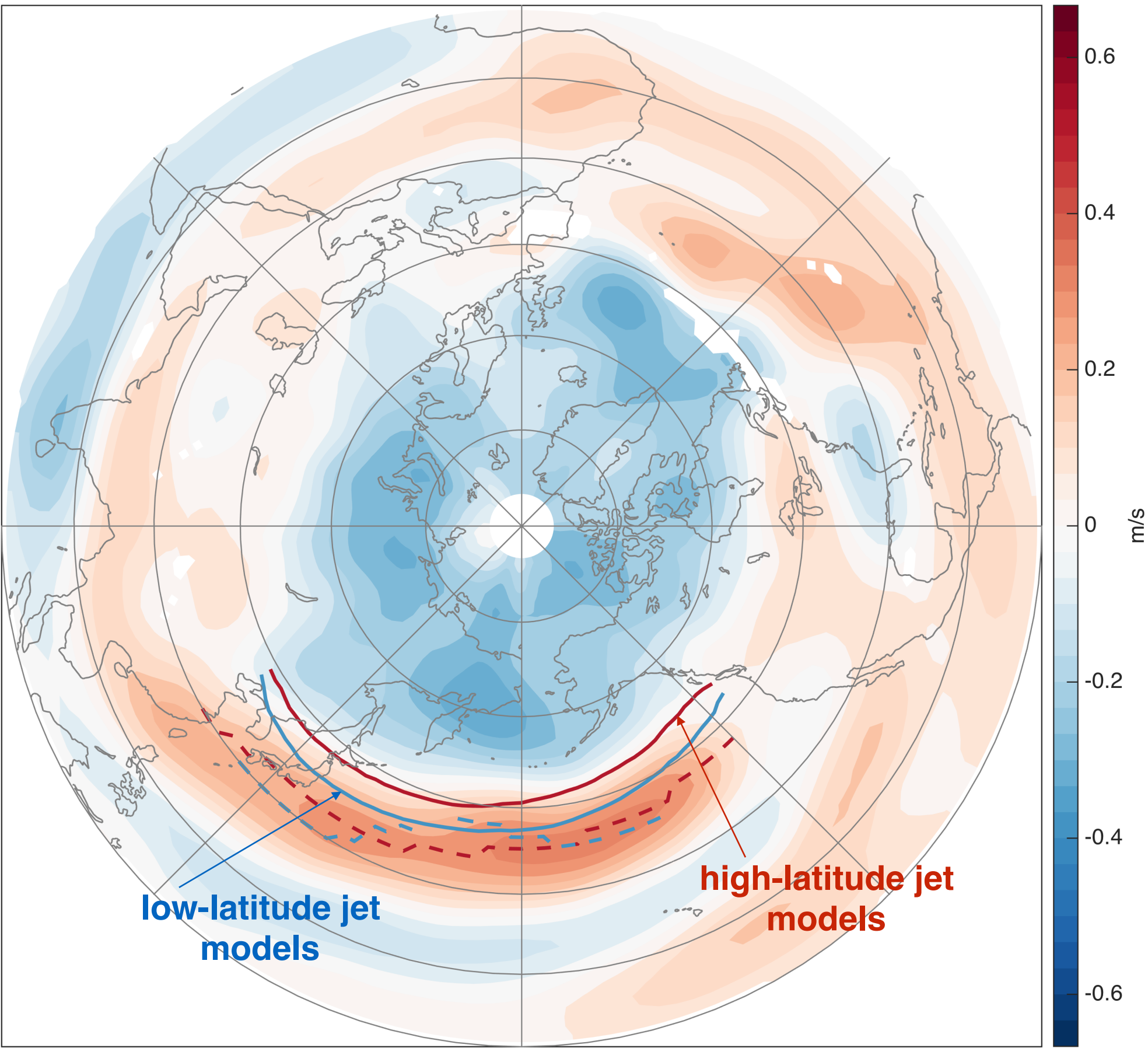


Barnes and Simpson (submitted)

Models with higher latitude jets shift further

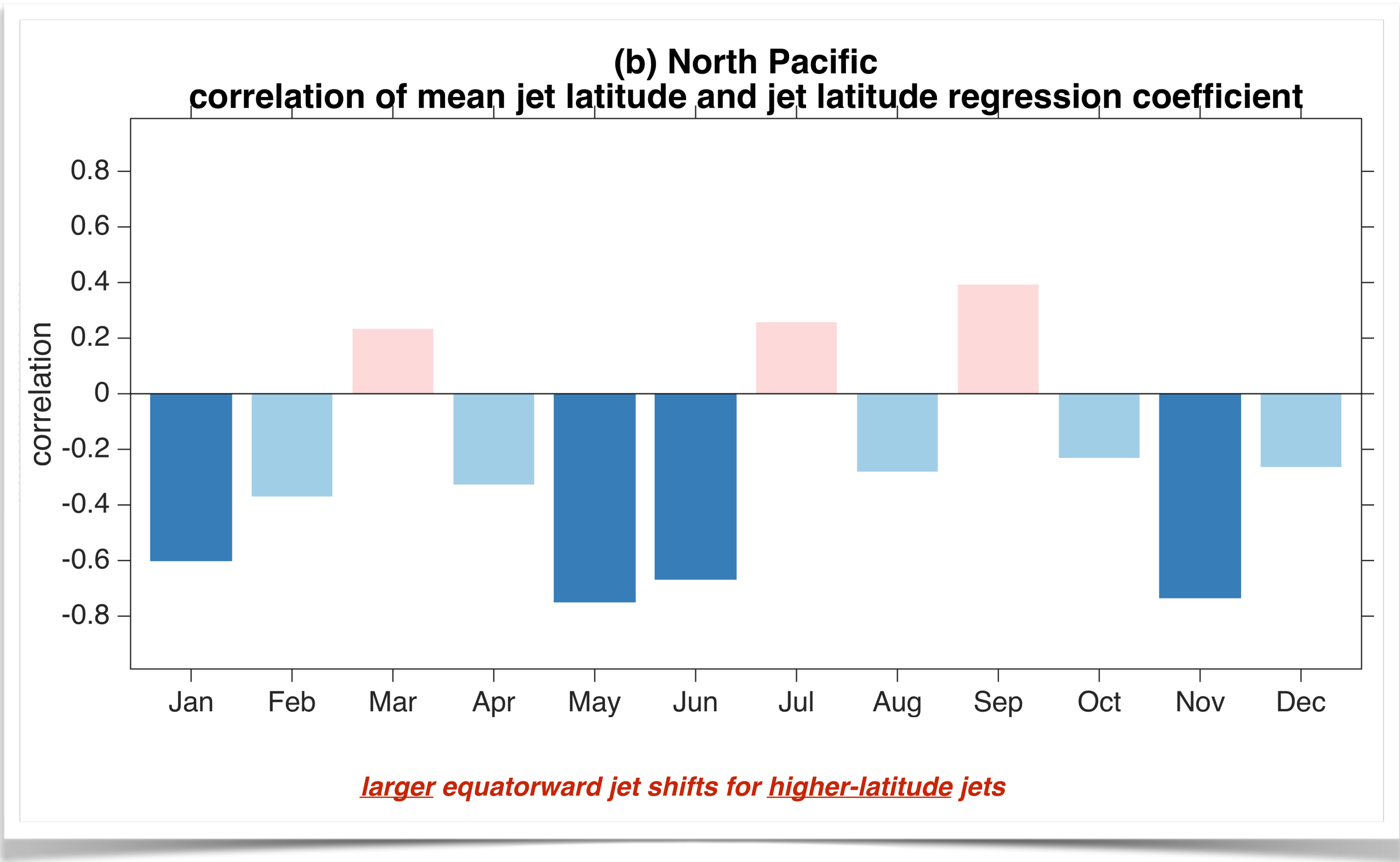


CMIP5 RCP8.5 u700 response to 1K POLE warming for warming in NorthPacific
June

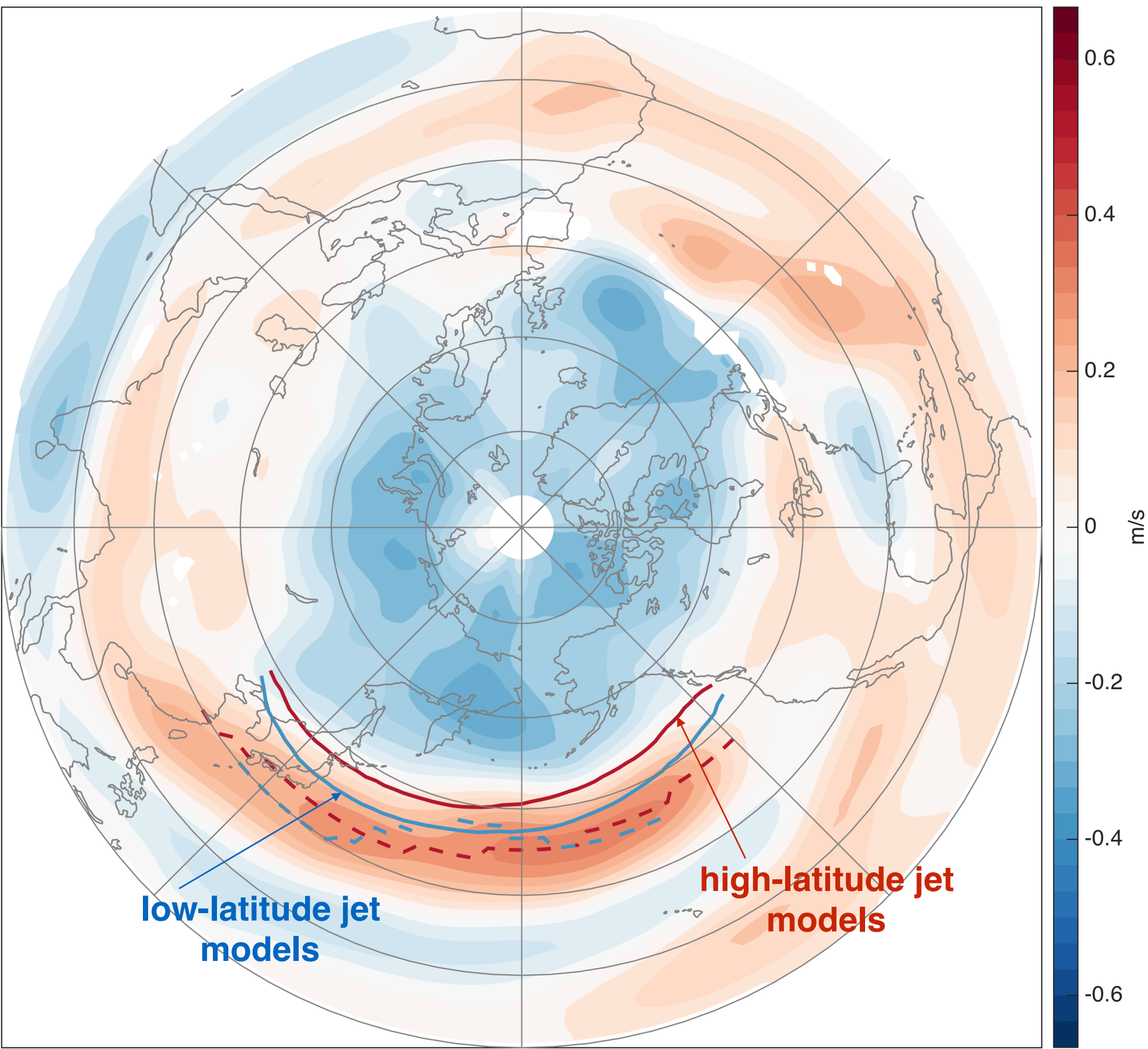


Barnes and Simpson (submitted)

Models with higher latitude jets shift further

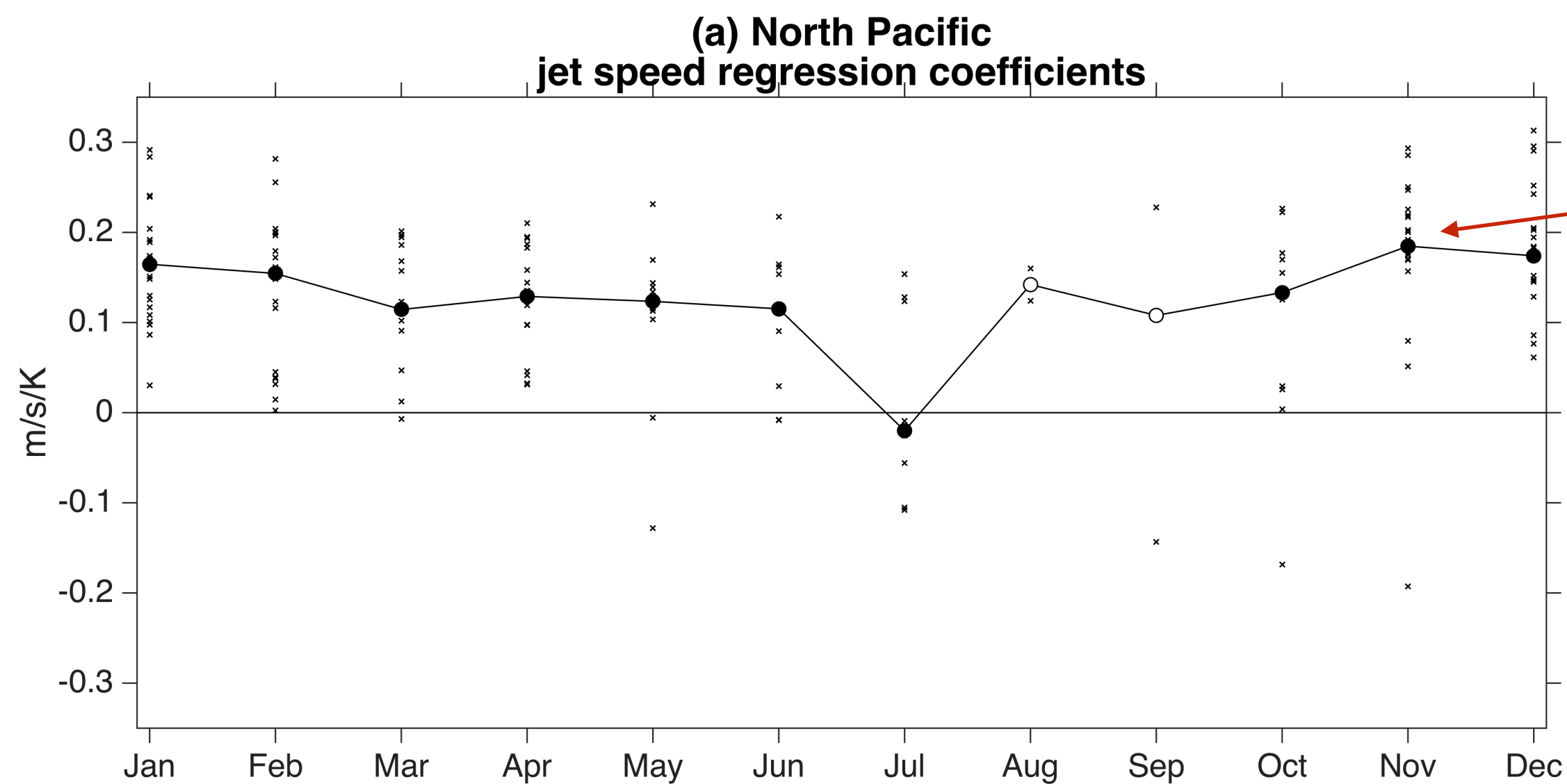


CMIP5 RCP8.5 u700 response to 1K POLE warming for warming in North Pacific June

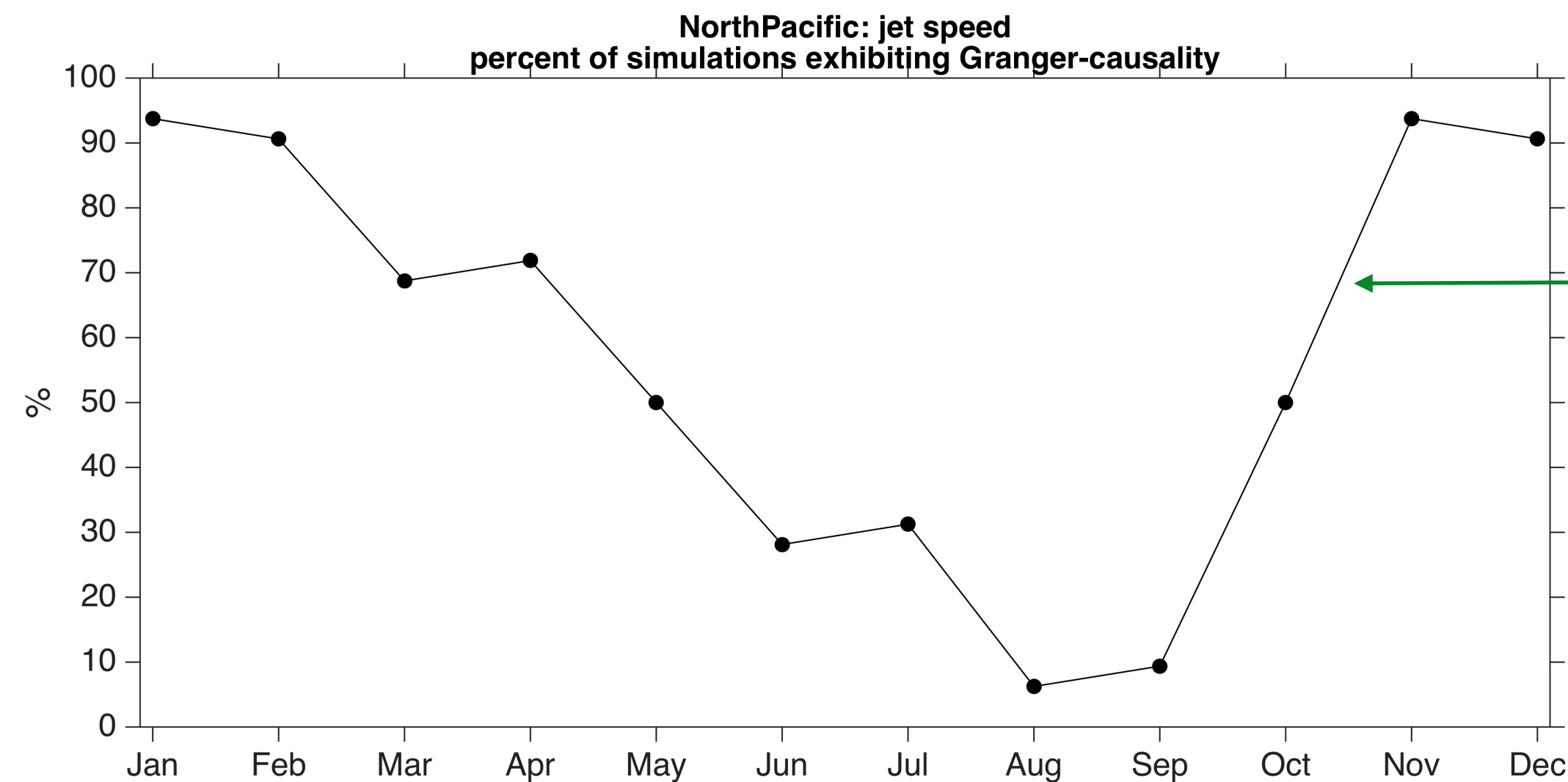


Barnes and Simpson (submitted)

How Much?: jet speed

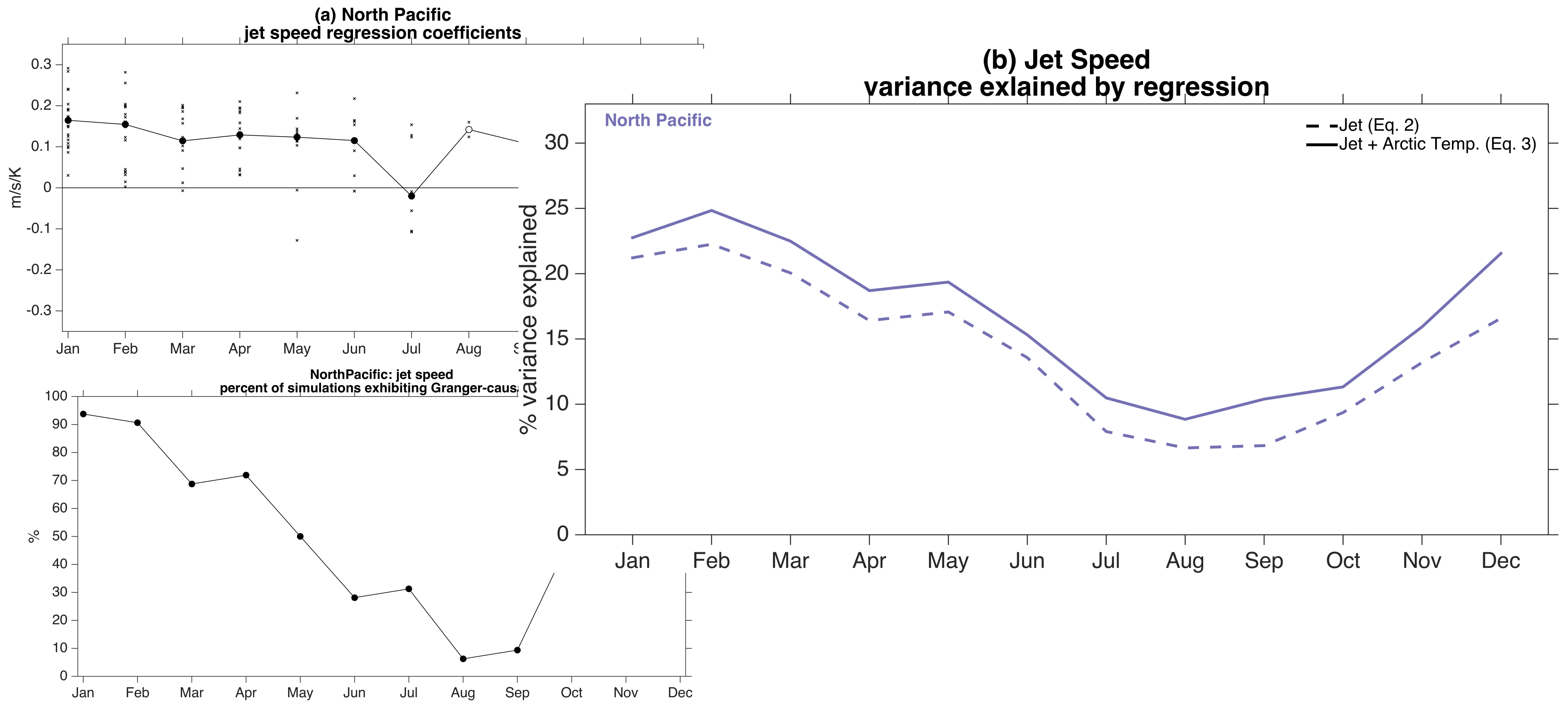


jet strengthens in most months



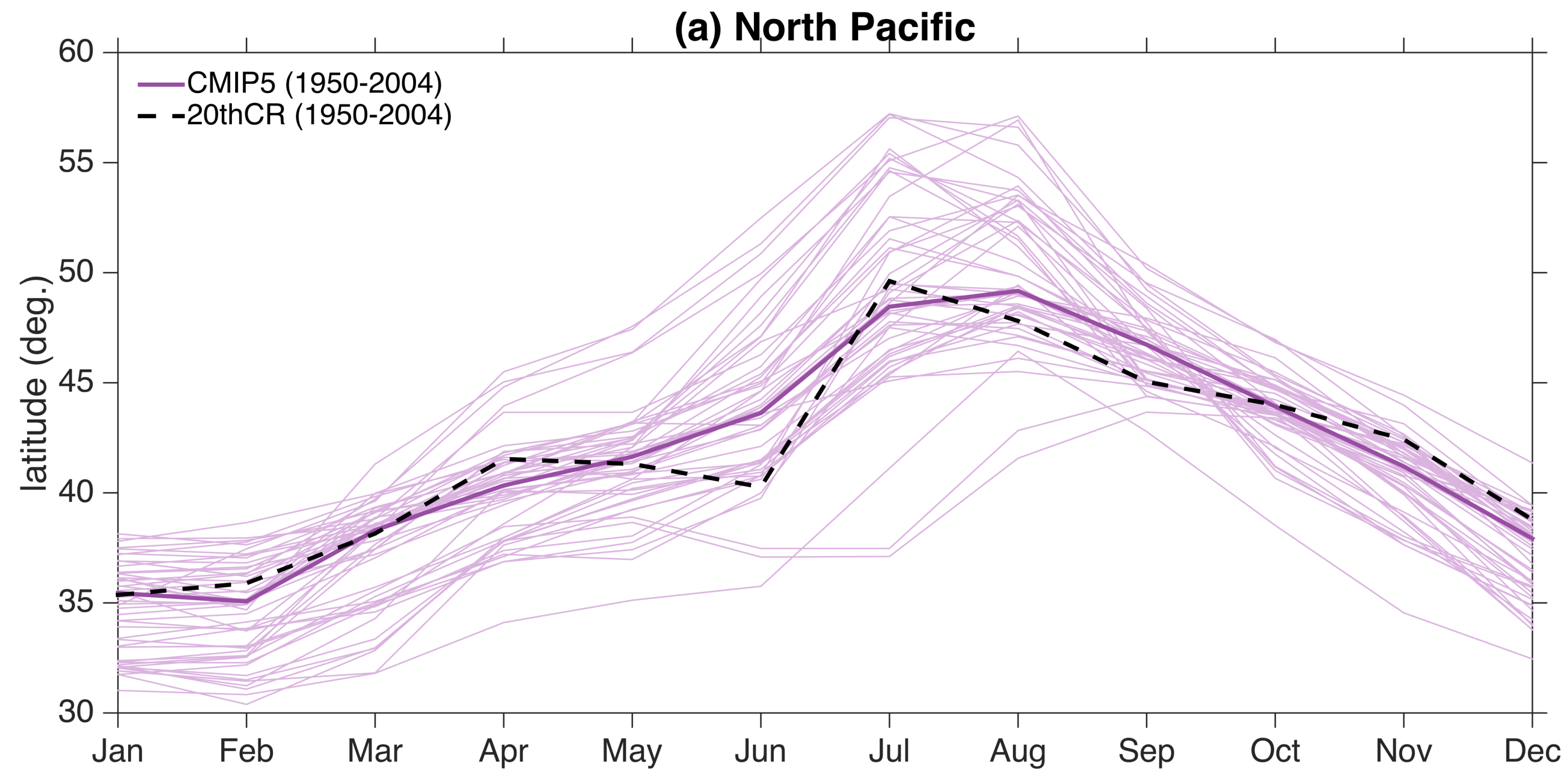
large seasonality in number of models exhibiting Granger-causality

How Much?: jet speed



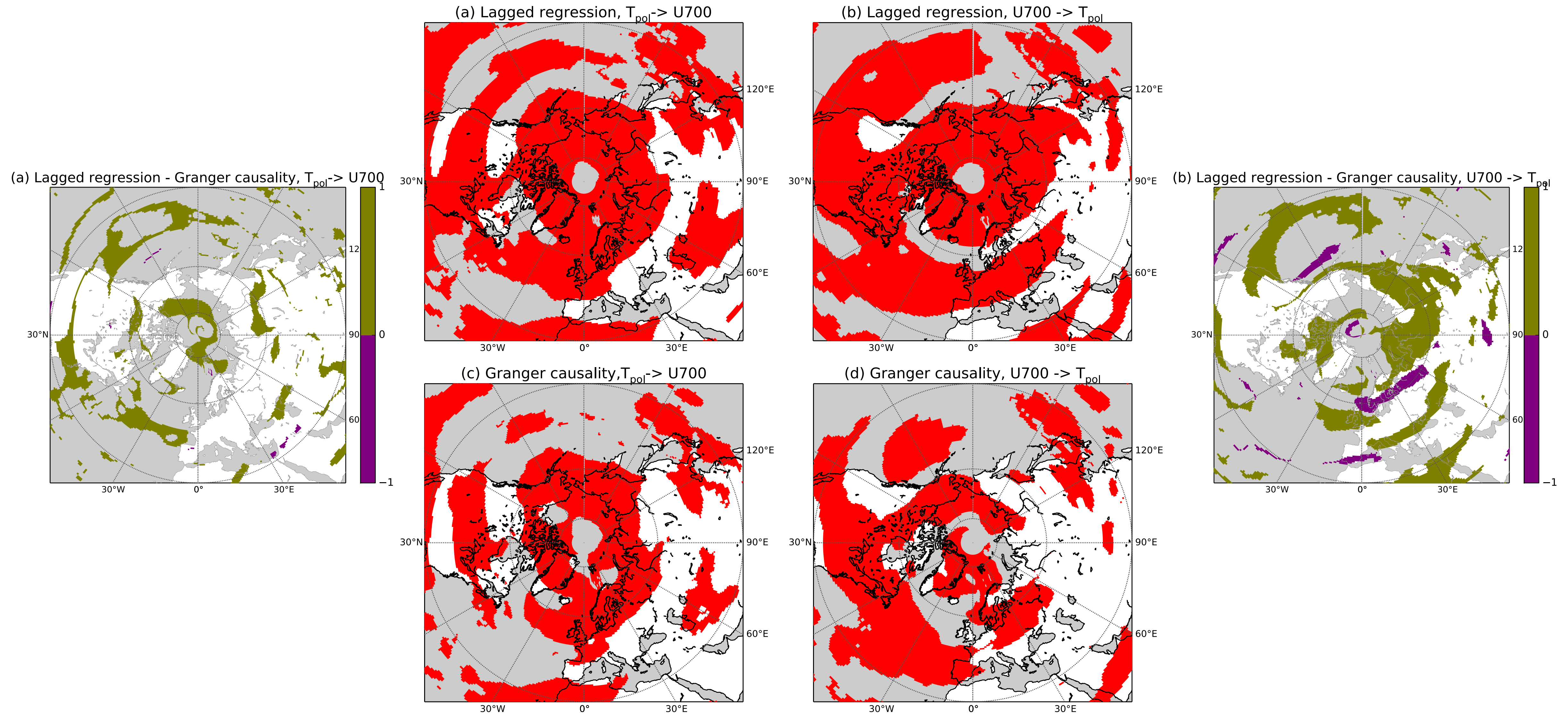
Barnes and Simpson (submitted)

Implications for models biases



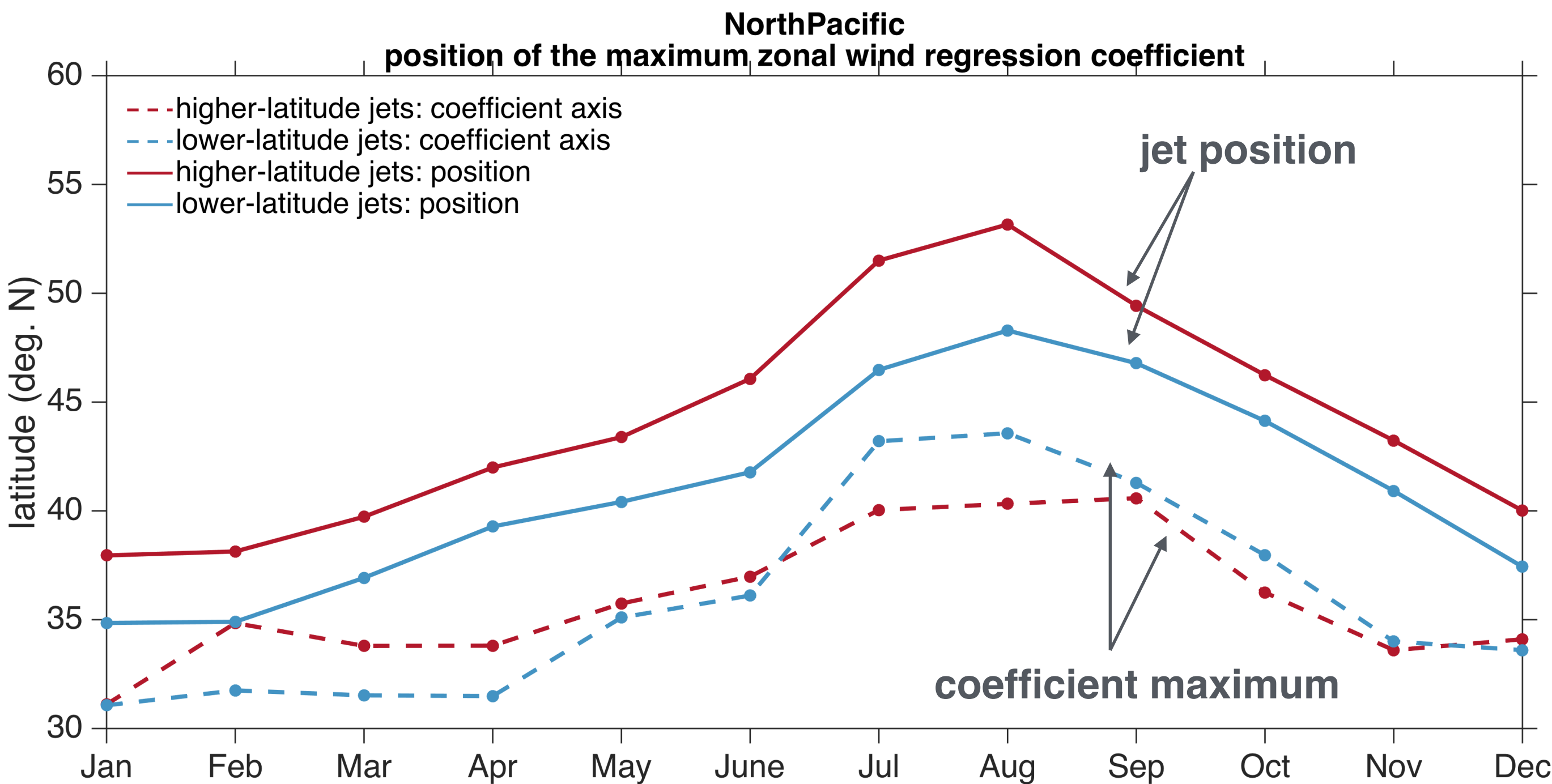
models exhibit seasonal biases
in the mean jet position

Applied to Arctic temperatures and 700hPa zonal wind



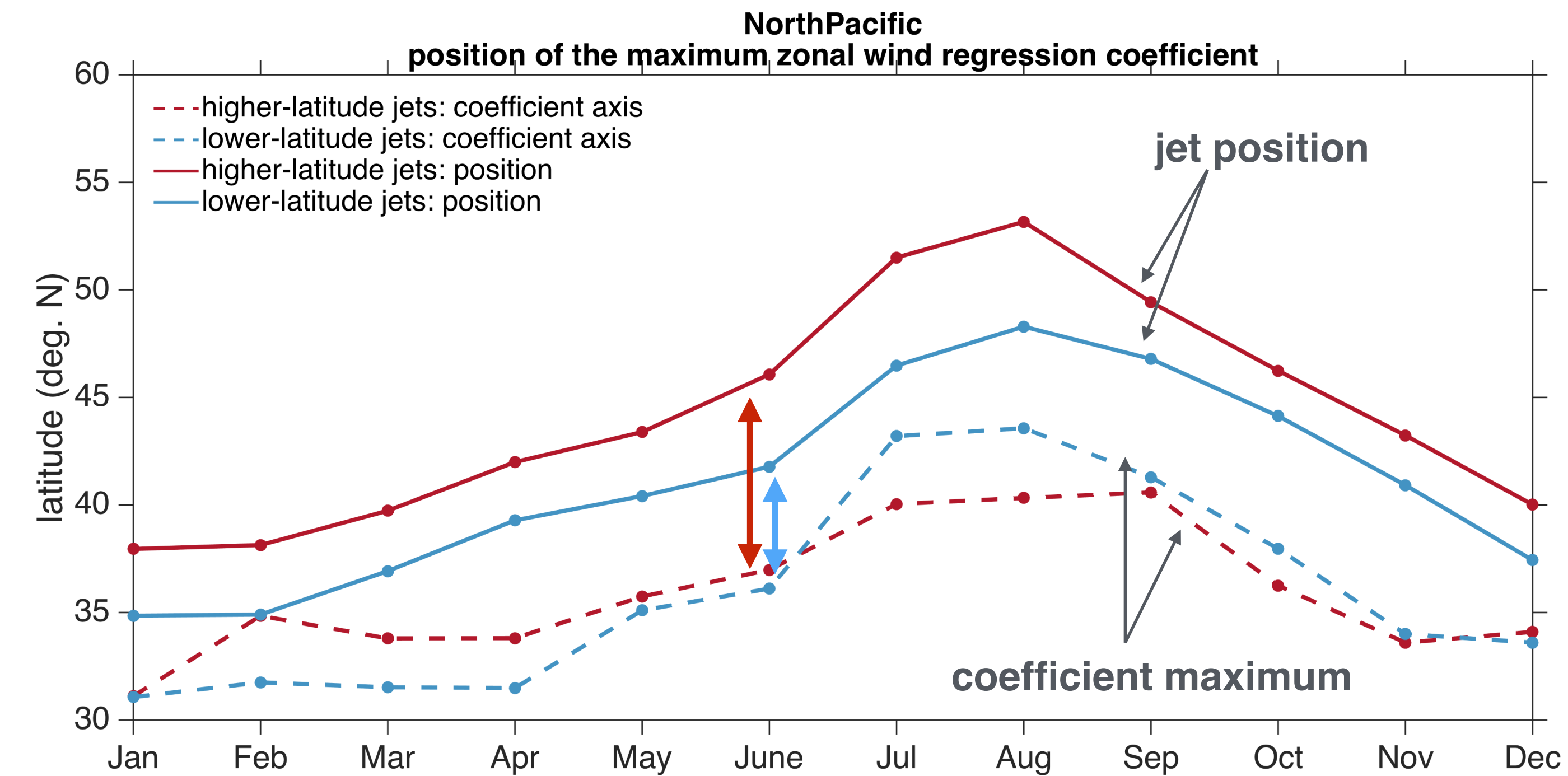
McGraw and Barnes (submitted)

Implications for models biases

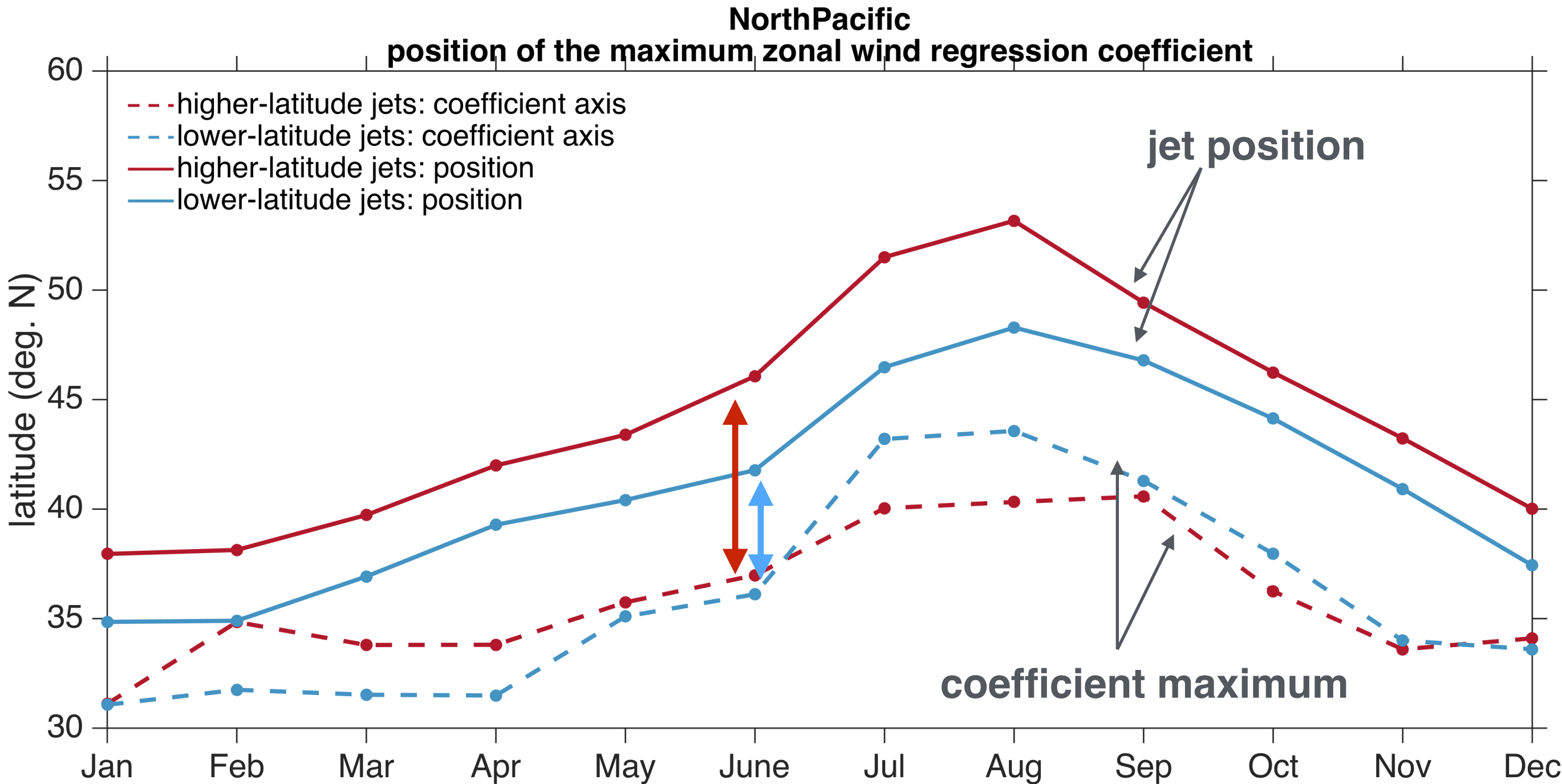


Barnes and Simpson (in prep)

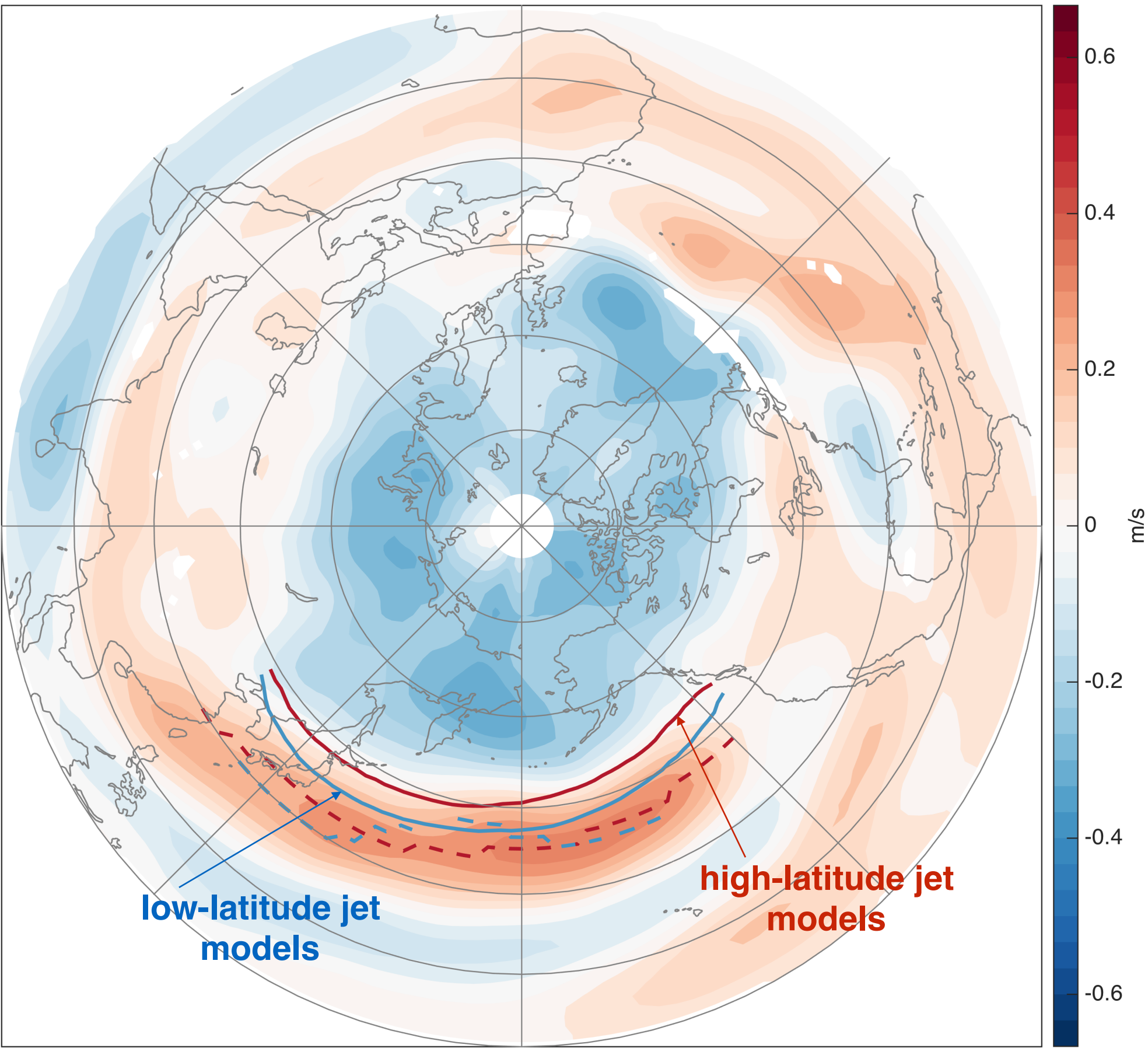
Implications for models biases



Implications for models biases

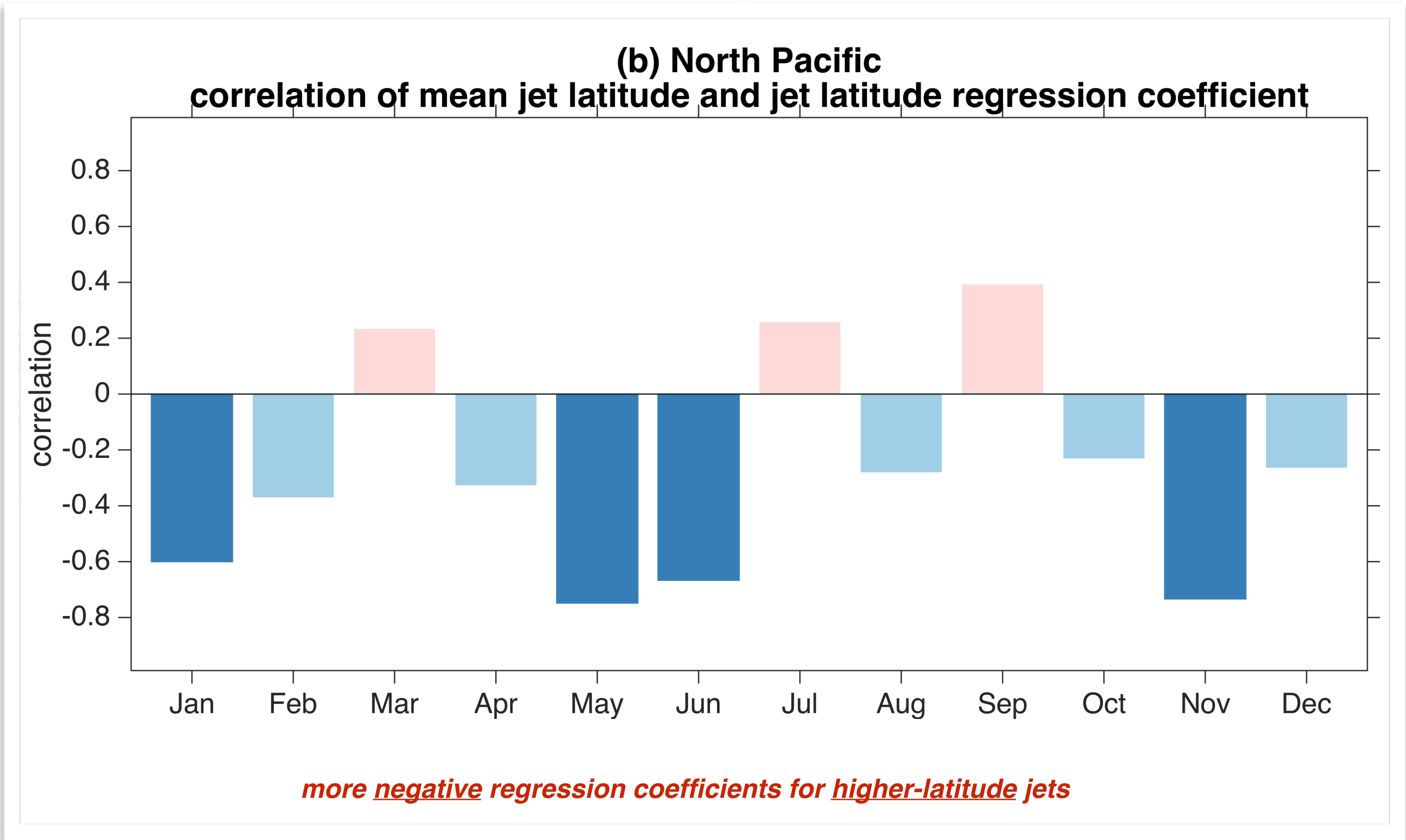


CMIP5 RCP8.5 u700 response to 1K POLE warming for warming in NorthPacific
June

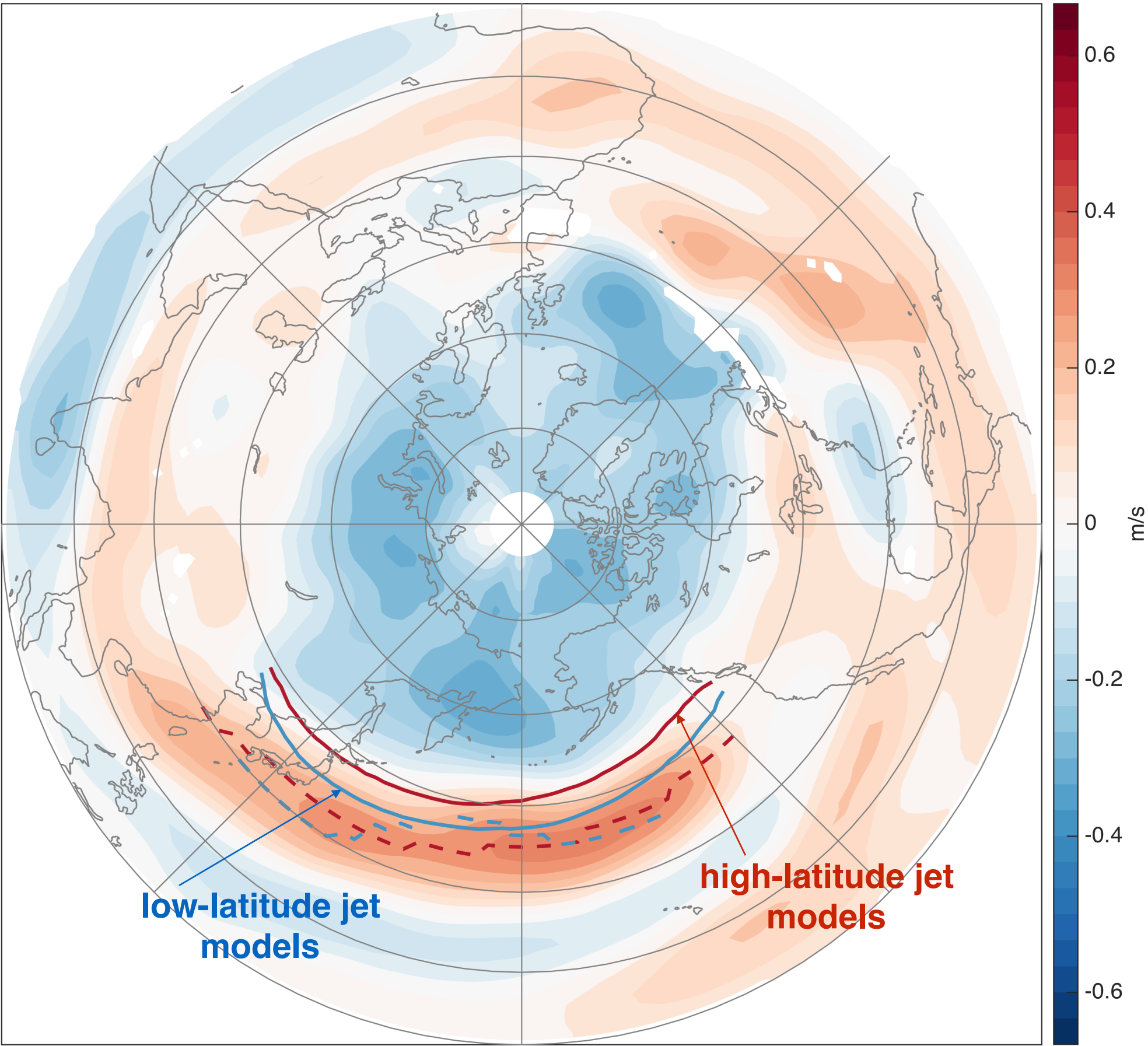


Barnes and Simpson (in prep)

Implications for models biases



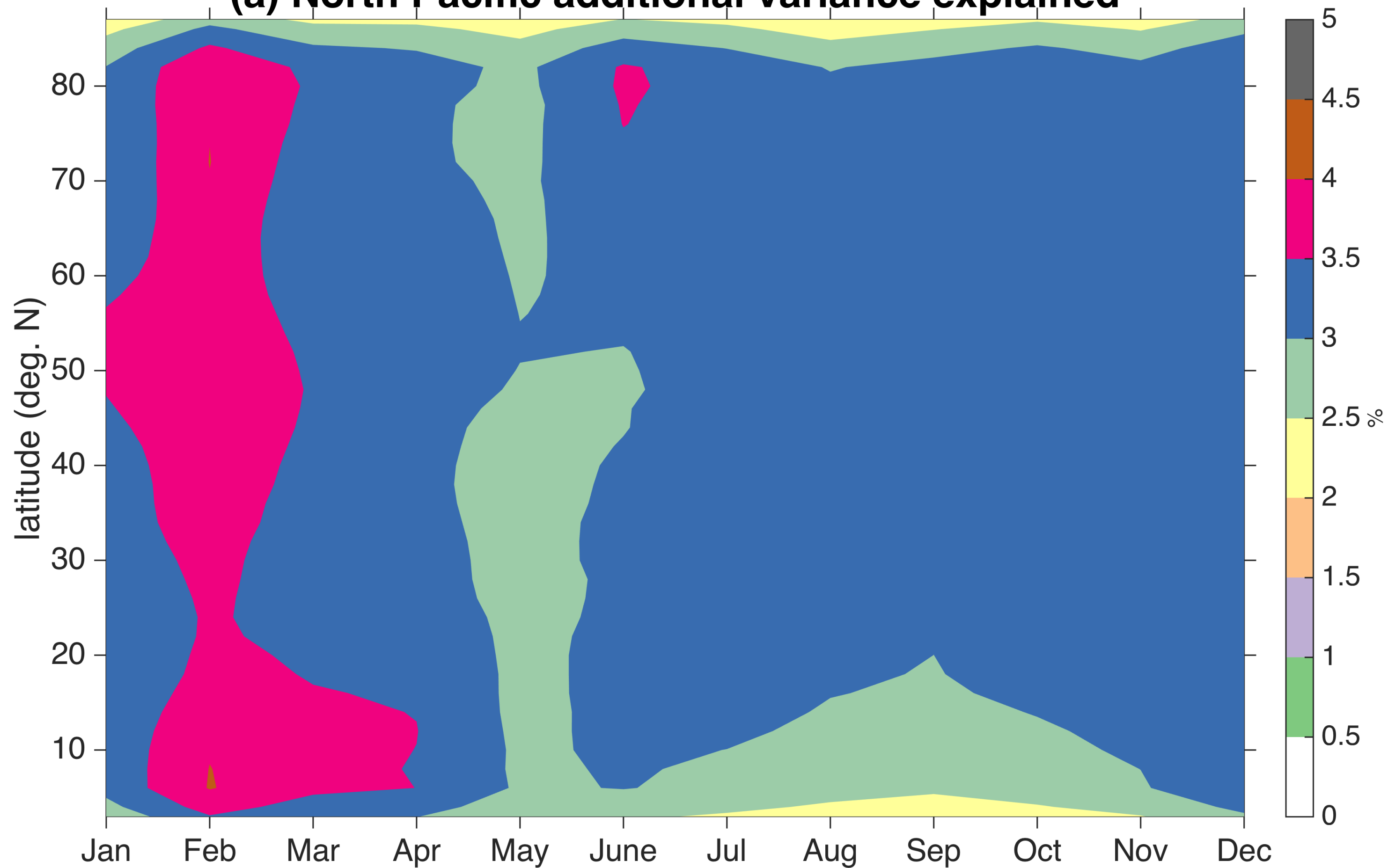
CMIP5 RCP8.5 u700 response to 1K POLE warming for warming in NorthPacific June



Barnes and Simpson (in prep)

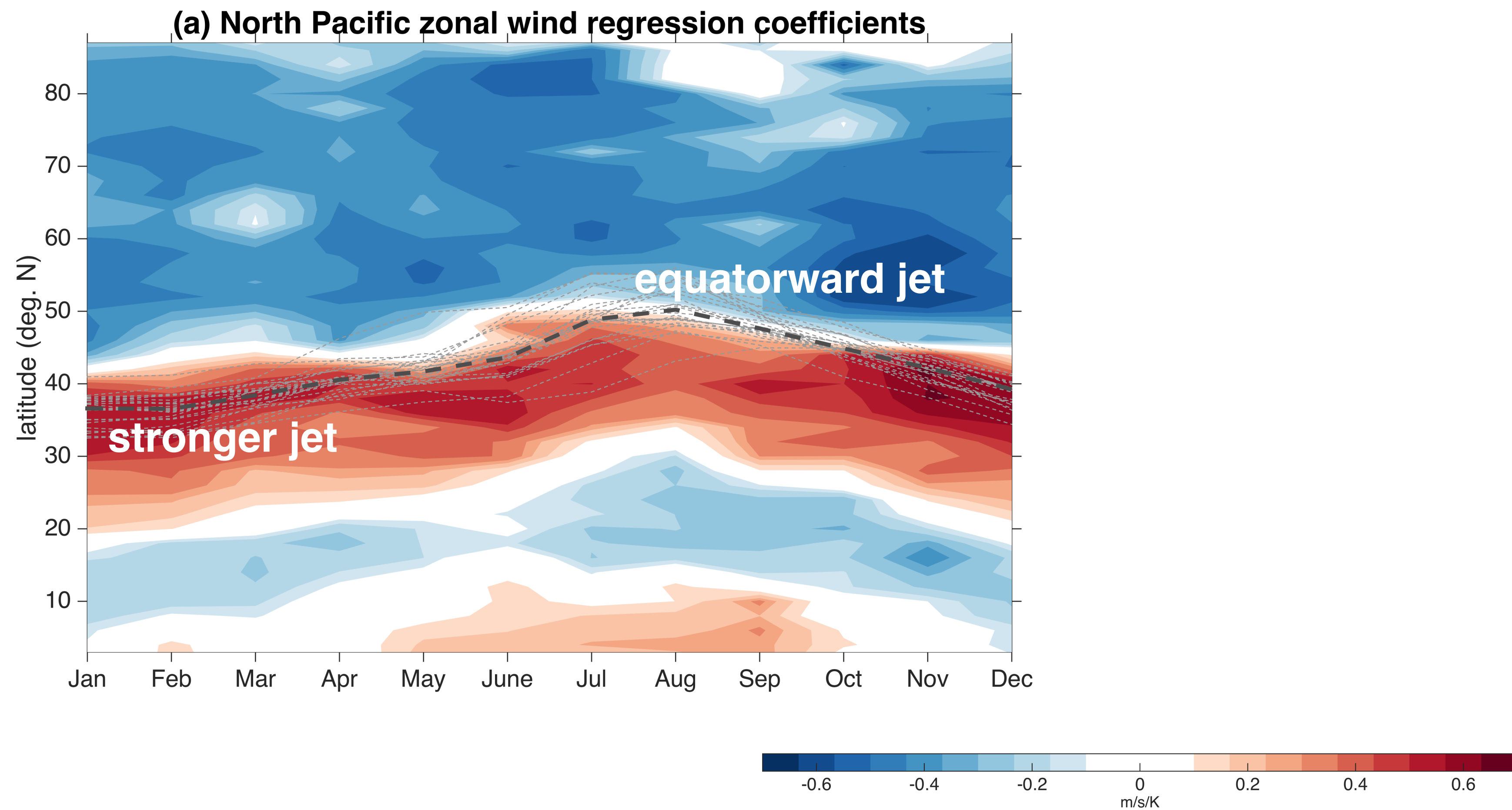
Variance explained in zonal winds

(a) North Pacific additional variance explained



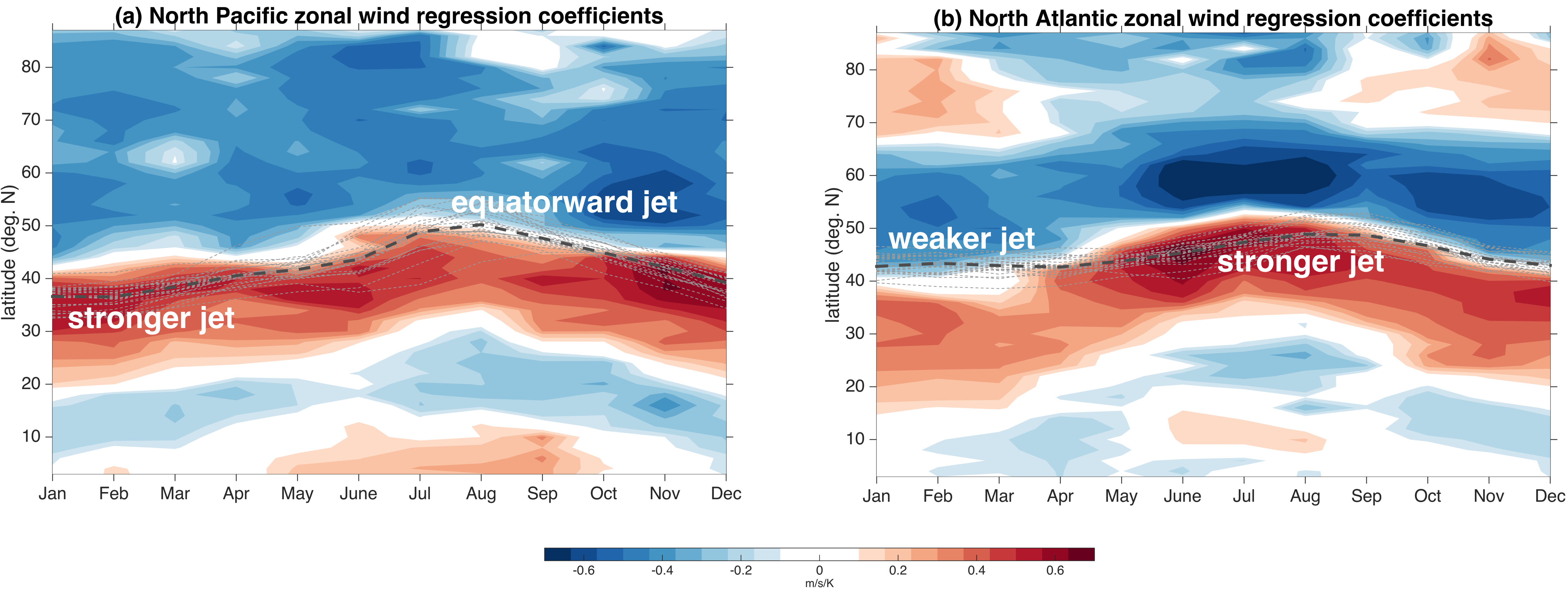
2-4% of additional zonal wind
variance explained by Arctic
temperatures

How Much?: **North Pacific** vs **North Atlantic**



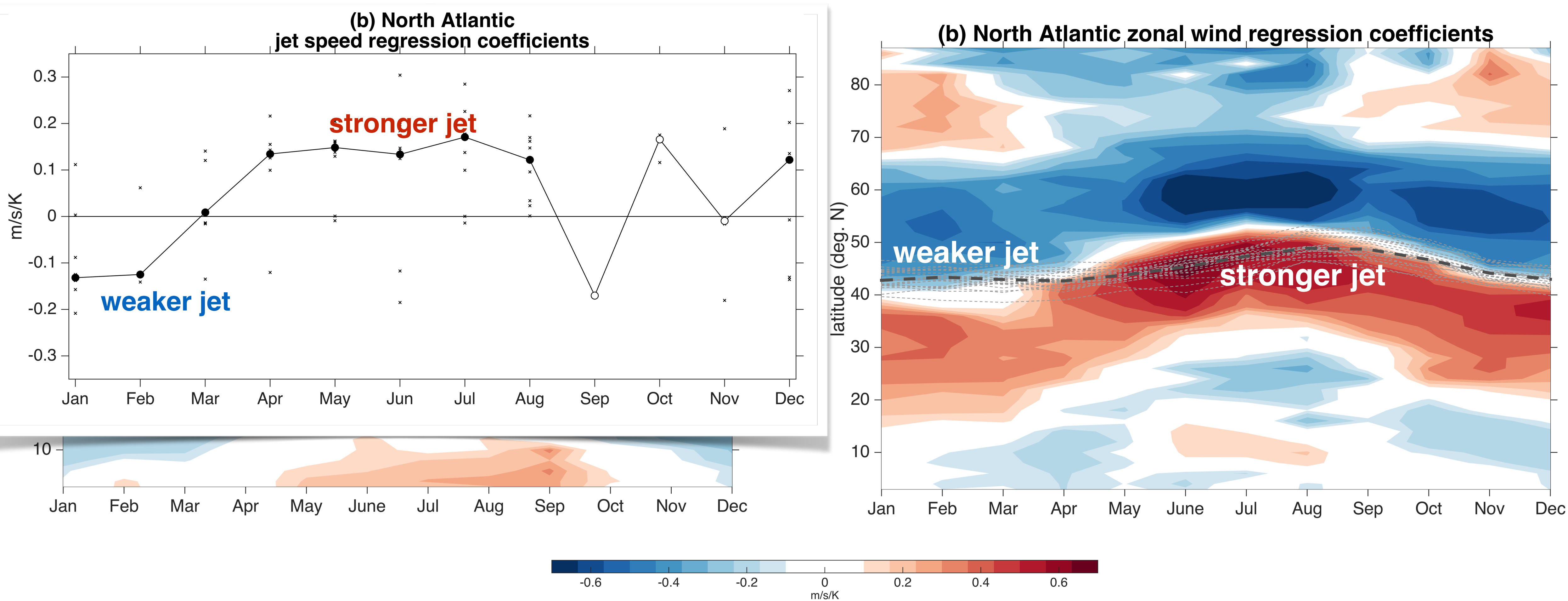
Barnes and Simpson (in prep)

How Much?: North Pacific vs North Atlantic



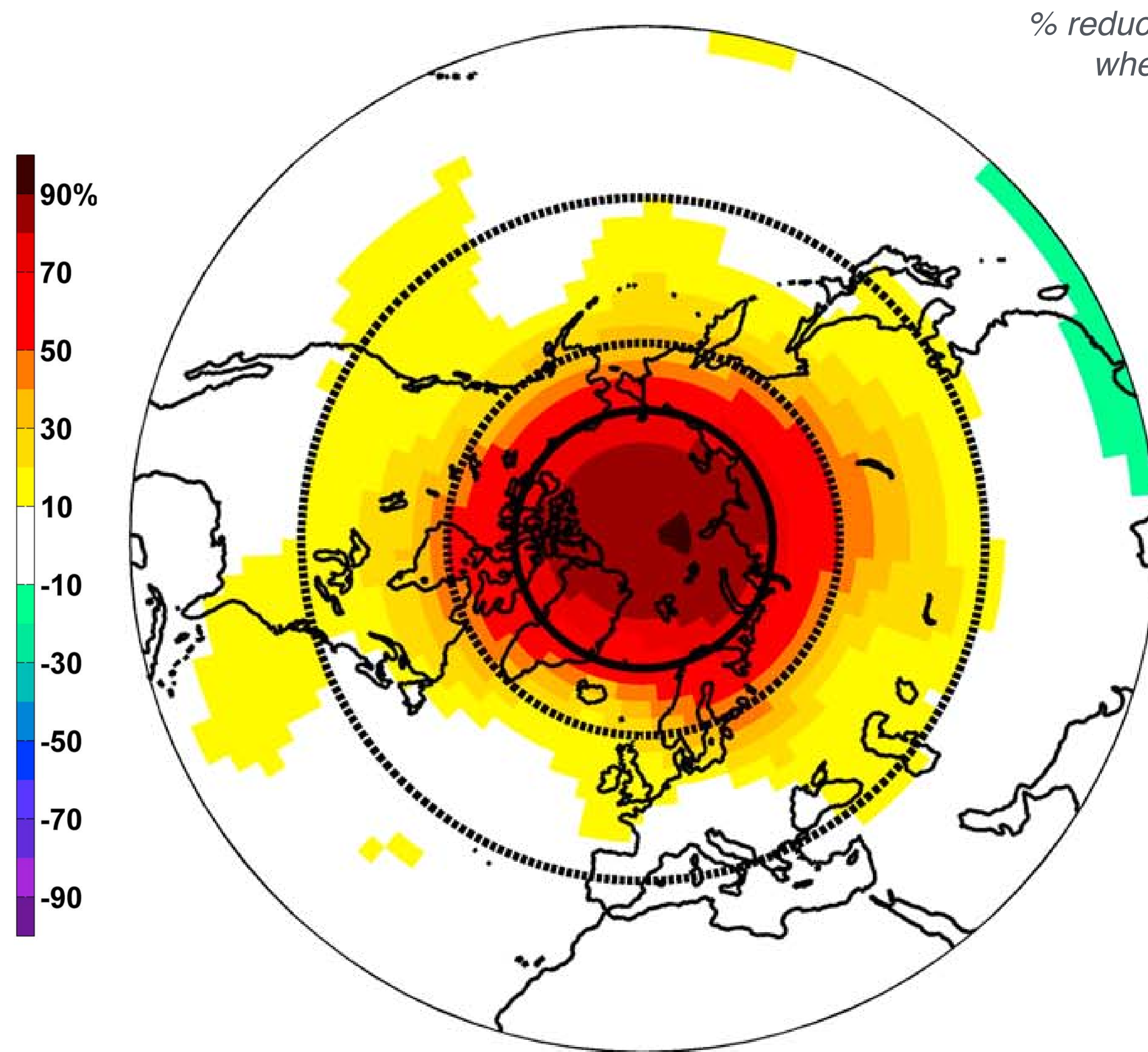
Barnes and Simpson (in prep)

How Much?: North Pacific vs North Atlantic



Barnes and Simpson (in prep)

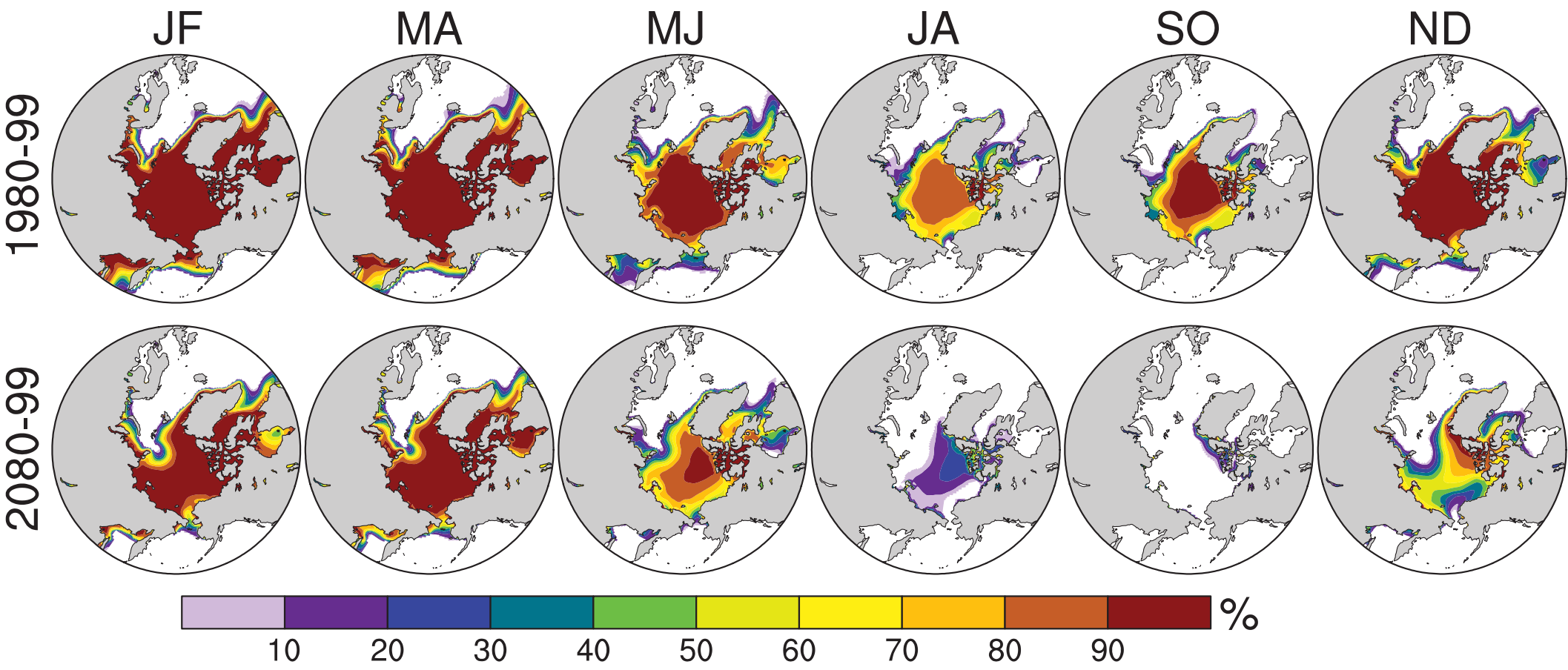
Forecasting evidence



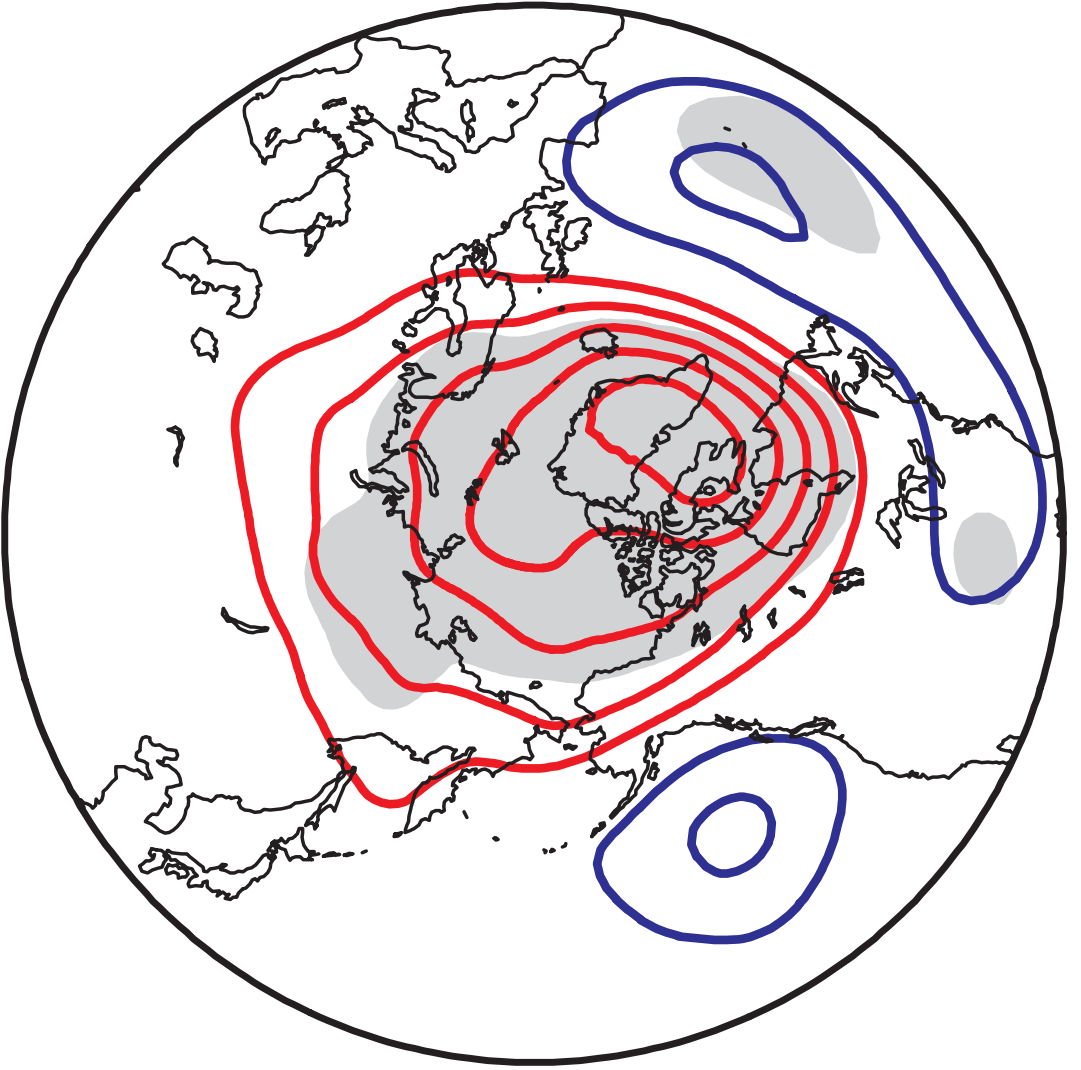
- Forecast experiments with ECMWF model shows that knowledge of the Arctic state can improve forecasts in midlatitudes
- Lowest improvement over the oceans where atmospheric variability is large

Modeling evidence

a) Sea Ice Concentration



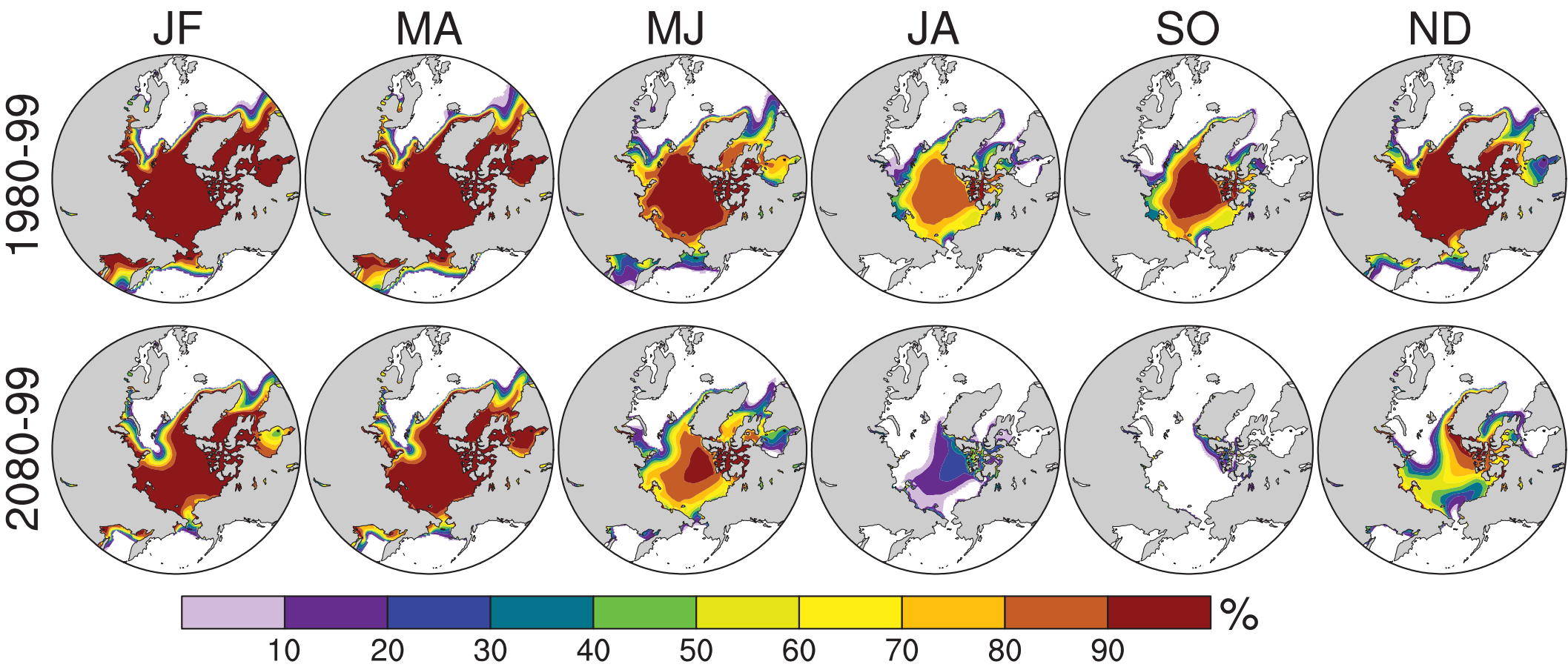
500 hPa geopotential height
response in Jan.-Feb.



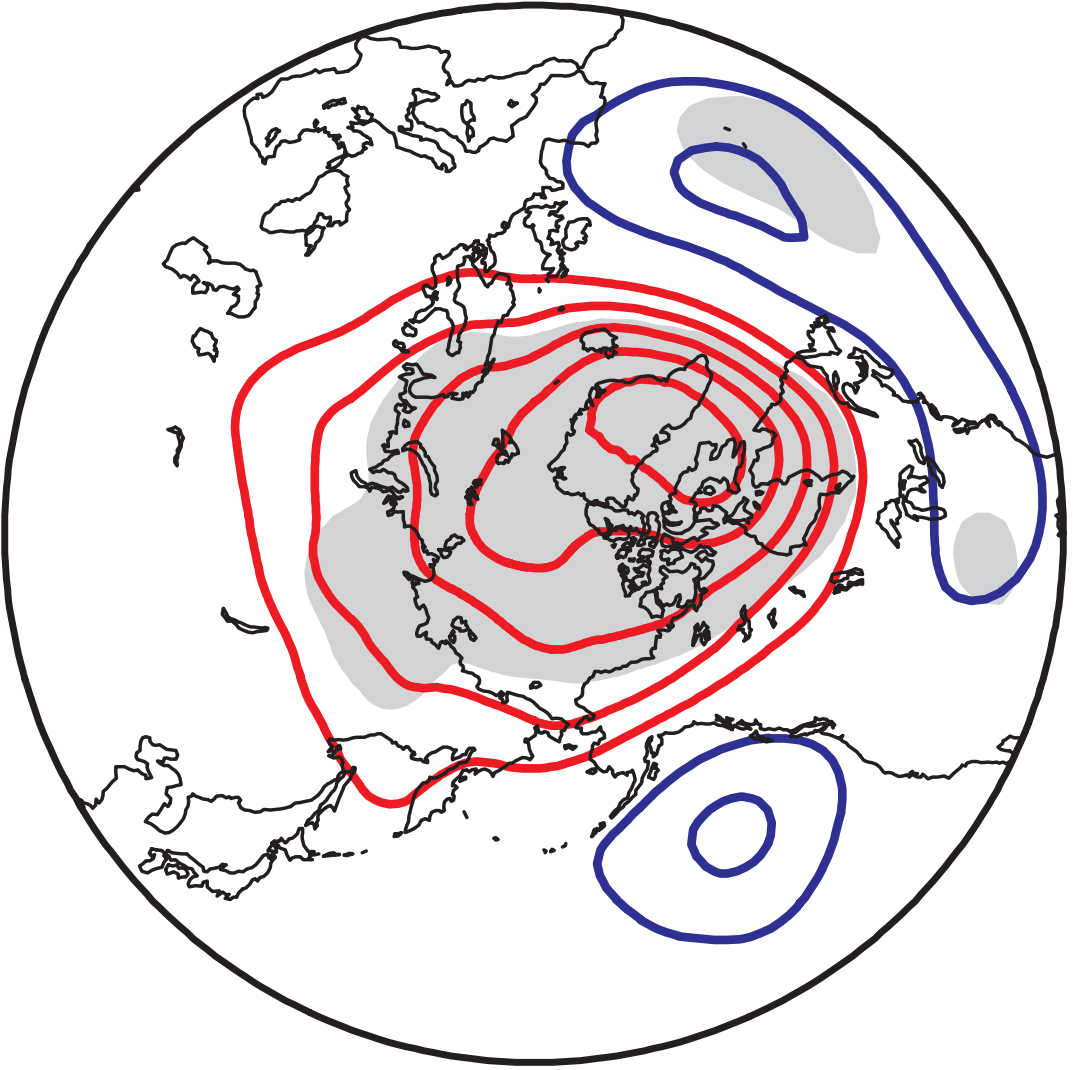
atmosphere-only CAM3 simulations
Deser, Tomas, et al. (2010; JCLI)

Modeling evidence

a) Sea Ice Concentration



500 hPa geopotential height
response in Jan.-Feb.



abstract of Deser et al. (2010):

“The loss of Arctic sea ice is greatest in summer and fall, yet the response of the net surface energy budget over the Arctic Ocean is largest in winter.”

...

“[The circulation] response resembles the negative phase of the North Atlantic Oscillation in February only.”

*atmosphere-only CAM3 simulations
Deser, Tomas, et al. (2010; JCLI)*

Mediators of Surface Temperature Persistence

