Signal and noise in regime systems

A hypothesis on the predictability of the NAO

K. Strommen, T. Palmer (Oxford University)

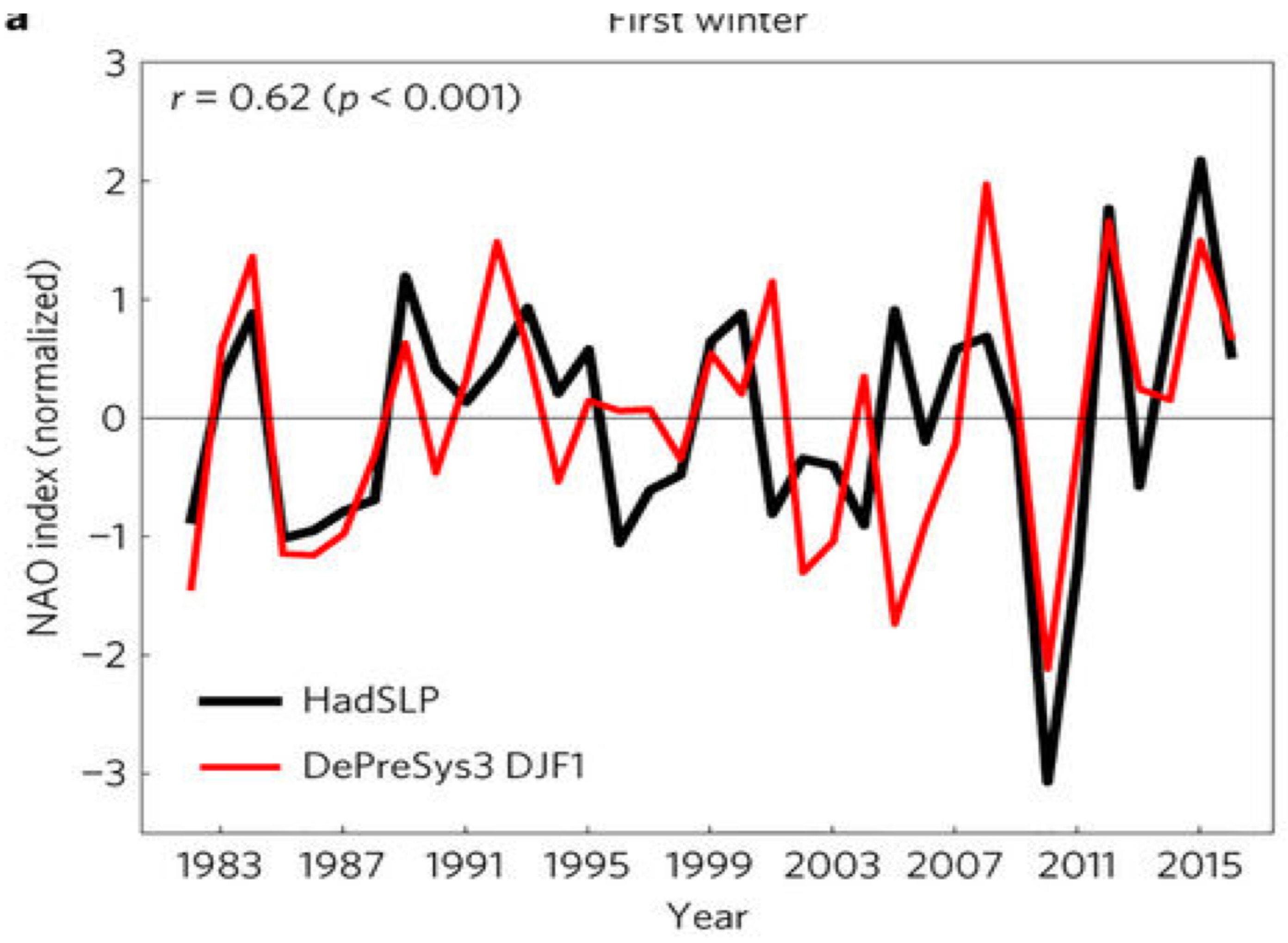
Thanks to the Met Office for providing DePreSys3 data

Motivation

- Growing evidence of a `**signal-to-noise paradox**’ in winter NAO forecasts.
- See Scaife et al (2014), Eade et al (2014), Dunstone et al (2016), Siegert et al (2016), Scaife and Smith (2018), Baker et al (2018), …
- **MY GOAL** → Convince you that paradox may be due to models not representing North Atlantic regimes properly.

Motivation

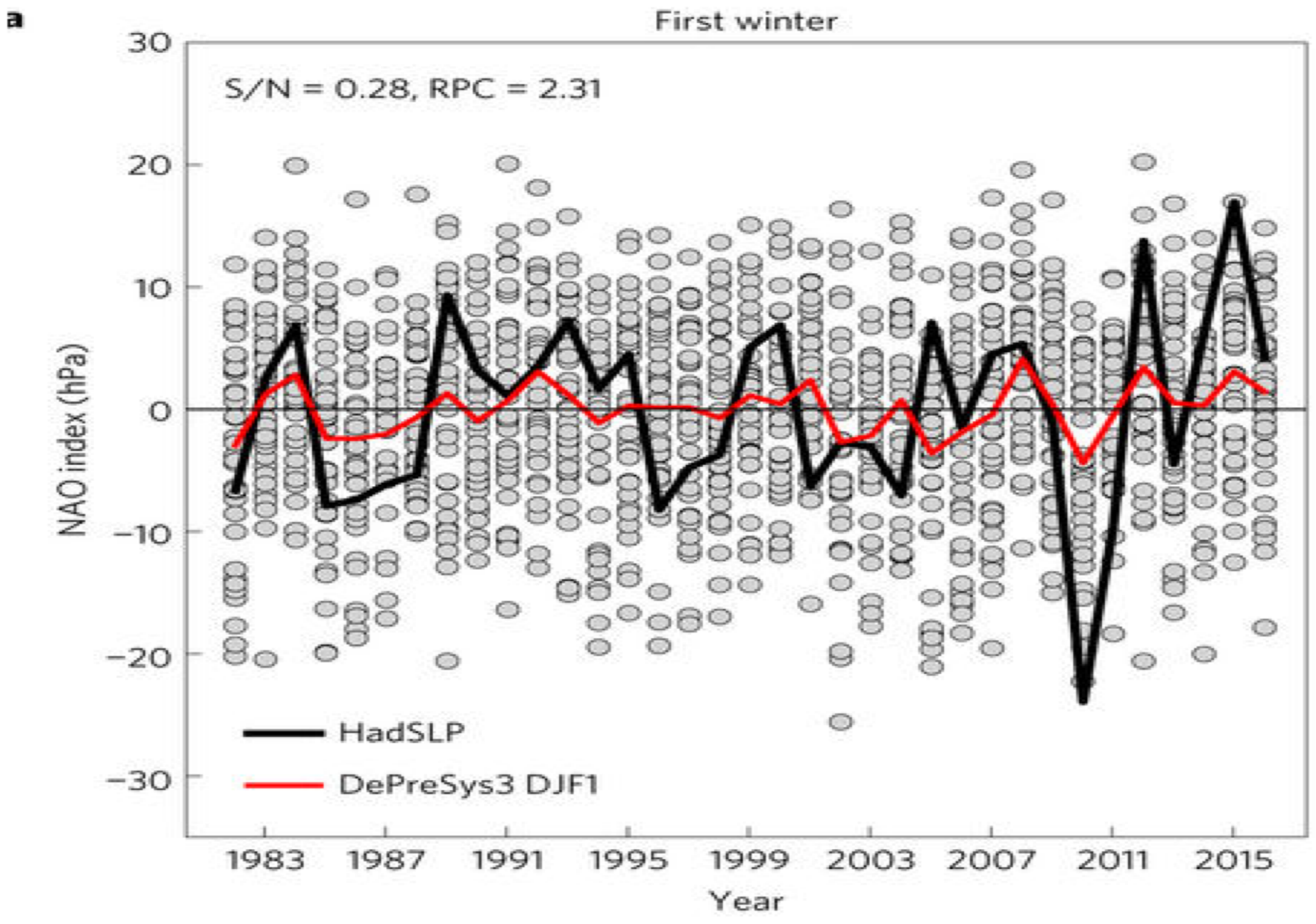
- **Dunstone et al (2016): our canonical example**



- 35 years
- 40 ensemble members

Motivation

- **Dunstone et al (2016): our canonical example**



Motivation

- Summarized by three numbers:
  - $Corr(EnsMean, Obs) = 0.62$ (*actual skill*)
  - $Corr(EnsMean, Mem) = 0.18$ (*potential predictability*)
  - $RPC = 2.31$

the **`signal-to-noise paradox'**

Motivation

- Siegert et. al (2016) → could be a result of the model having too weak a signal.
- So `paradox’ might be due to poorly propagated teleconnections
- Used a simple linear-regression style statistical model.

Motivation

- On the other hand….

- Many studies suggest the existence of **regimes** are playing a part in modulating North Atlantic variability.

- **How might we expect signal-to-noise ratios to look like in a more non-linear regime system?**

Outline of talk

1. Construct a simple statistical model of the NAO based on regimes.

2. Show how the three metrics (actual skill, potential predictability and RPC) behave in this model.

3. Show that `signal-to-noise paradox’ is normal behavior in this system if the NWP model has bad persistence.

0. What is RPC?

- By definition (Eade et al 2014):

$$RPC = \frac{PC(Obs)}{PC(Mod)}$$

Proportion of variance of NAO predictable in the **real** world

Proportion of variance of NAO predictable in the **model** world

- Eade et al. (2014) provide a **lower bound** for this quantity amenable to computation.

0. What is RPC?

- If the ensemble size is sufficiently large, then their estimate simplifies to just

$$RPC \approx \frac{Corr(EnsMean, Obs)}{Corr(EnsMean, Mem)}$$

Actual skill

Potential predictability

- Essentially uses the ensemble mean as a proxy for the real world.

1. A regime view of the NAO

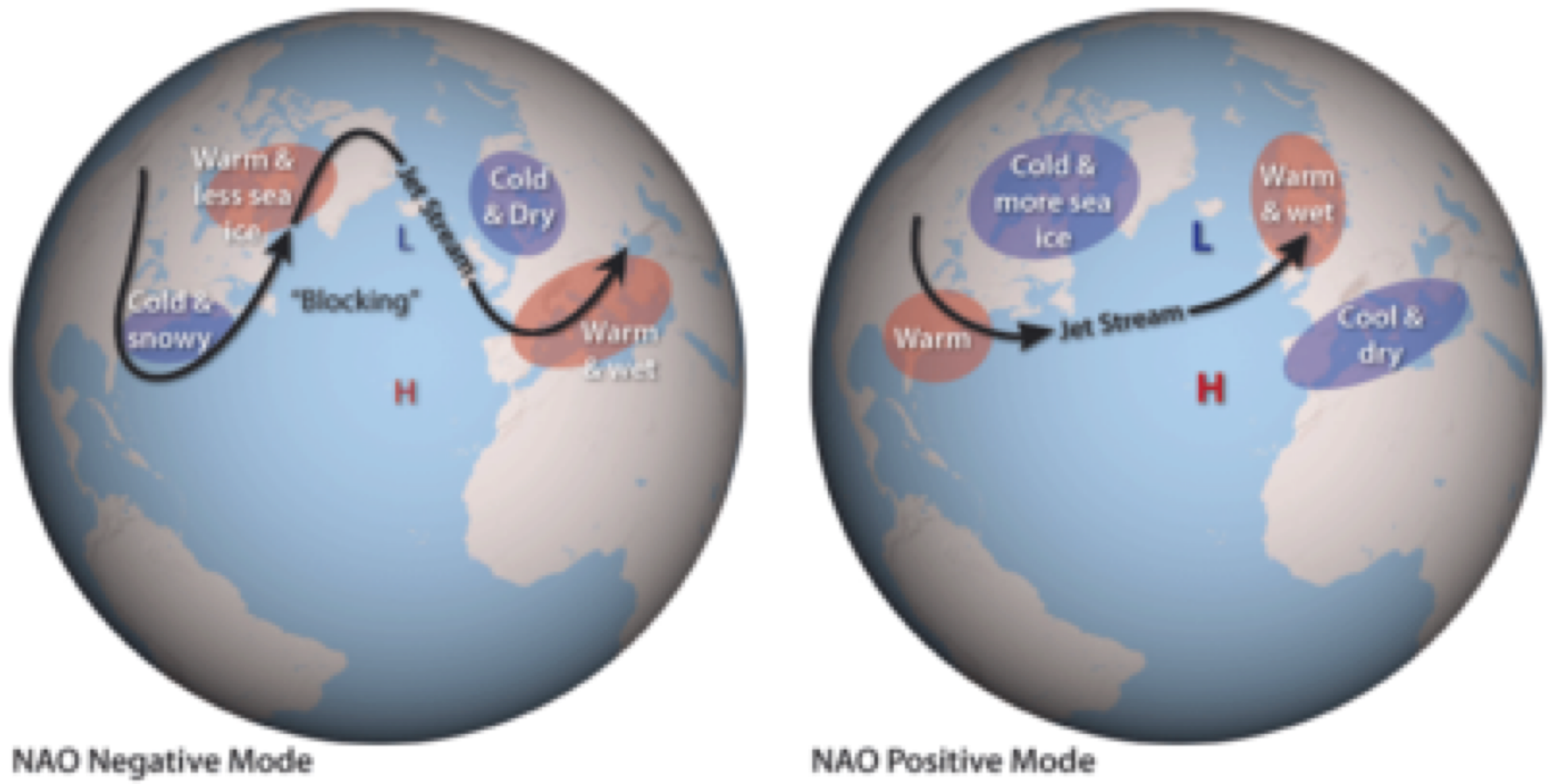
- Studies suggest the North Atlantic has anywhere between 2 and 4 distinct regimes (jet stream location, geopotential height patterns, …)

- Because 2 is an easy number to deal with, we will take a 2-state view of the North Atlantic. Probably too idealized???

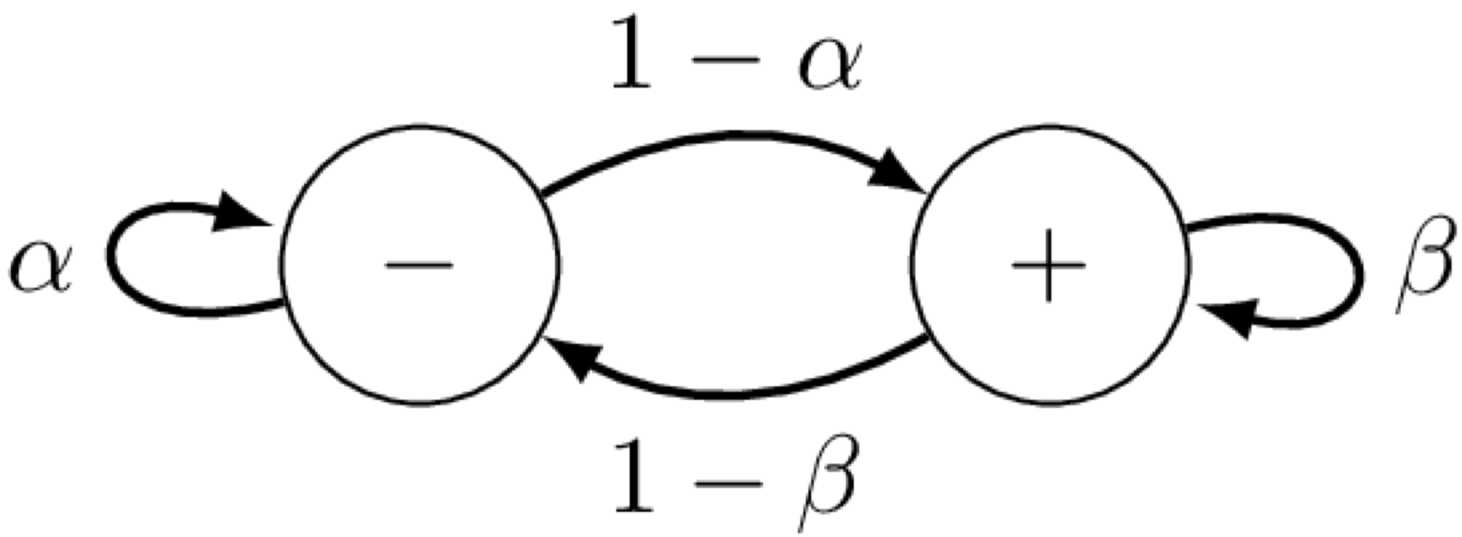
- Can easily be expanded to more states (work in progress).

1. A regime view of the NAO

- Courtesy of Met Office website

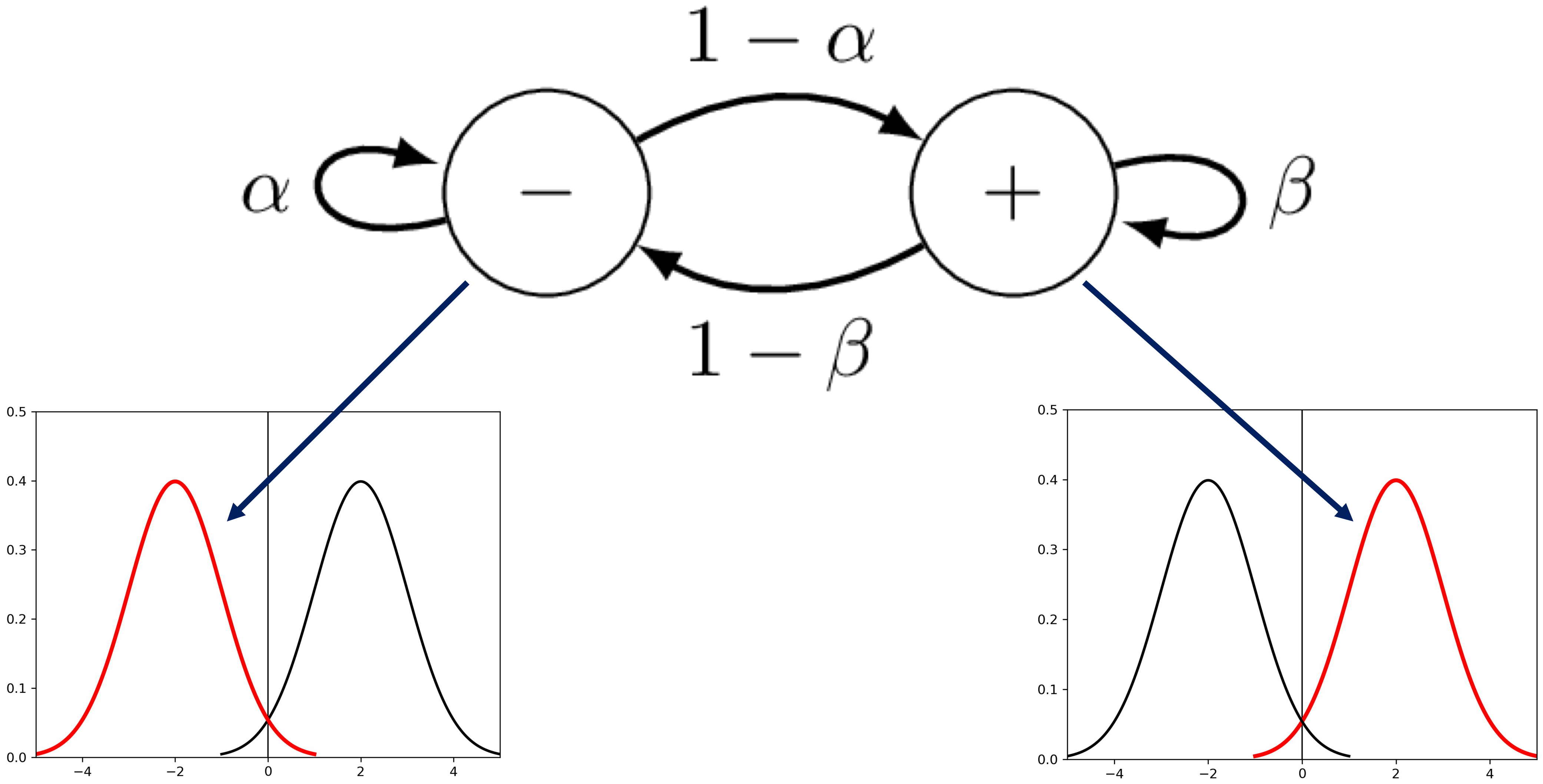


1. A regime view of the NAO

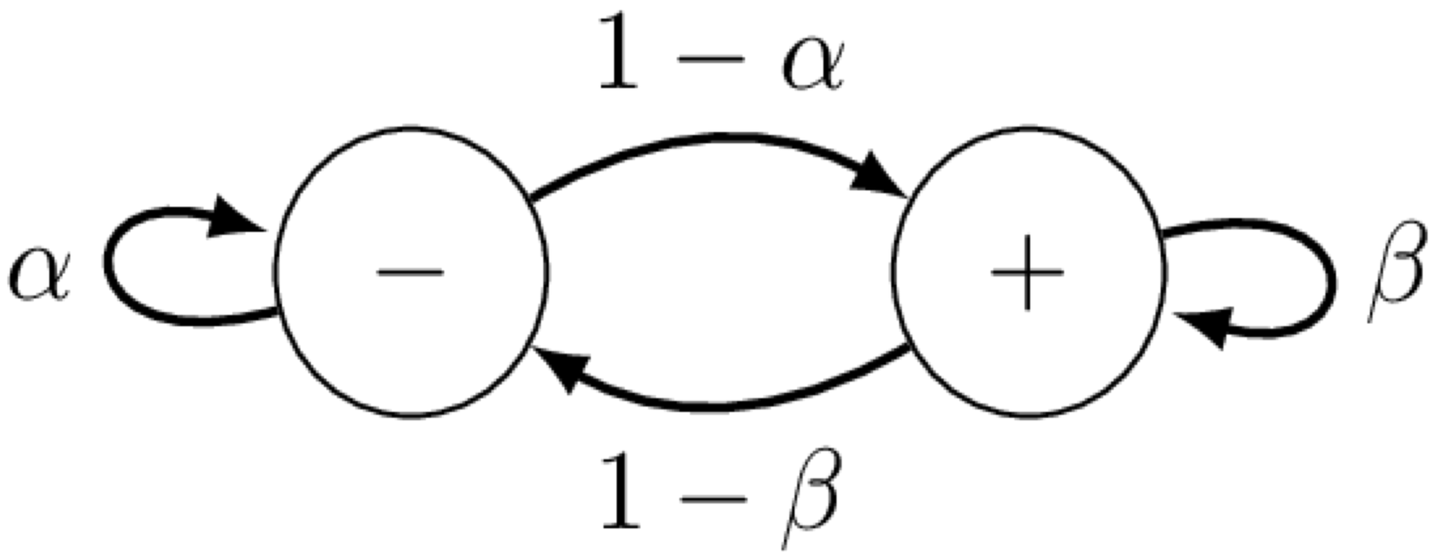


- Each day the atmosphere resides in one of these states, then moves randomly according to the persistence/transition probabilities.

1. A regime view of the NAO



1. A regime view of the NAO



- Persistence probabilities are fixed at the start of a given DJF
- **A DJF mean is then obtained by taking the mean of 90 days sampled using this Markov chain.**

1. A regime view of the NAO

- Predictability is induced by **seasonal deviations** of the two persistence probabilities from their climatological means.
- Such deviations will cause the atmosphere to have a **preferred regime state** during a given DJF → <u>signal in the NAO index.</u>
- **Example**: a preference towards NAO+ regime means more positive daily NAO indices, hence a more positive DJF mean.

2. Representing model skill/error

- How do we represent the imperfect skill that our NWP models have in such a system?

- Need our NWP model to mess up the true persistence probabilities.

2. Representing model skill/error

- Let p_{obs} be the true persistence (of either state), and p_{mod} the corresponding persistence of our NWP model.

- Assumed to be related via a number k, a **regime fidelity parameter**:

$$k = 1 \quad \rightarrow \quad p_{mod} = p_{obs} + \text{noise}$$

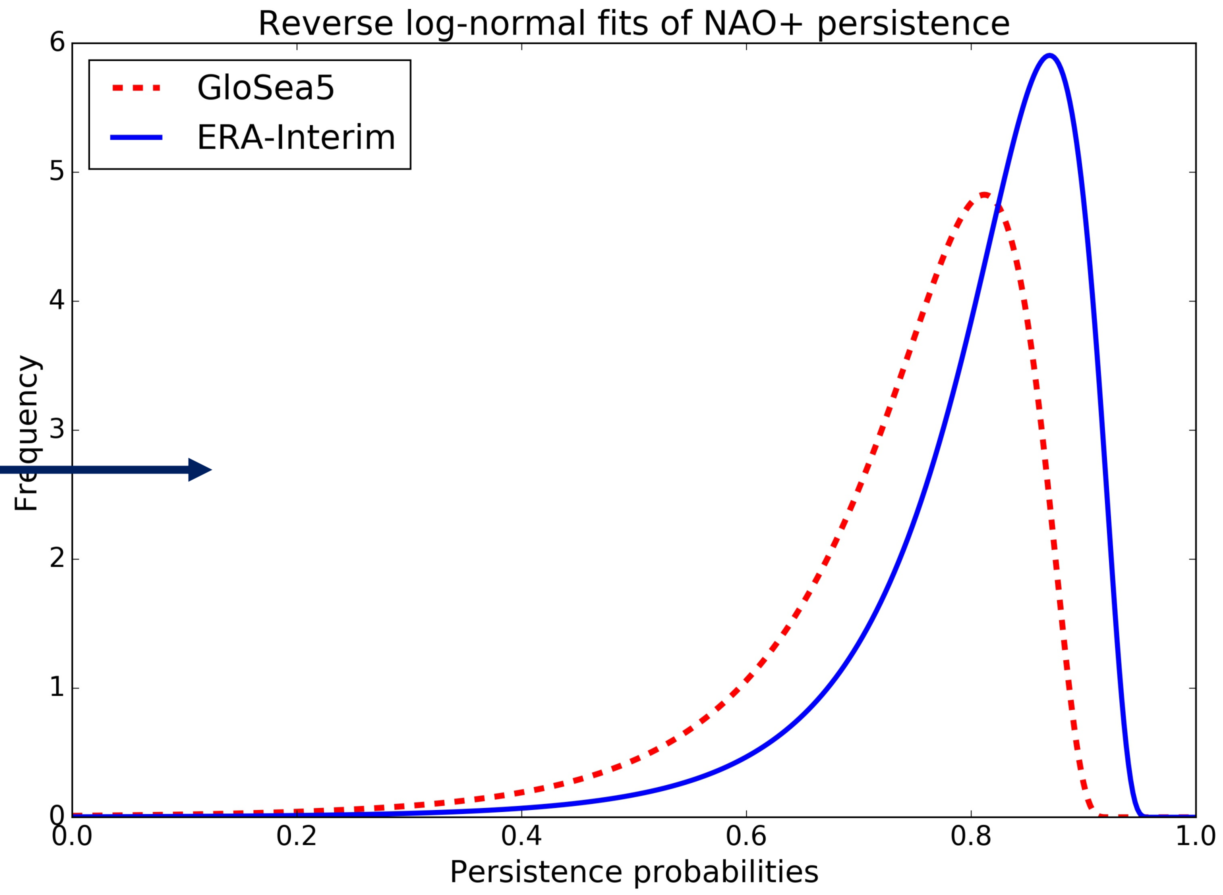
$$k = 0 \quad \rightarrow \quad p_{mod} = 1/2 + \text{noise}$$

2. Representing model skill/error

- In other words, the model error considered is **weak persistence**.

- Known problem in many models!

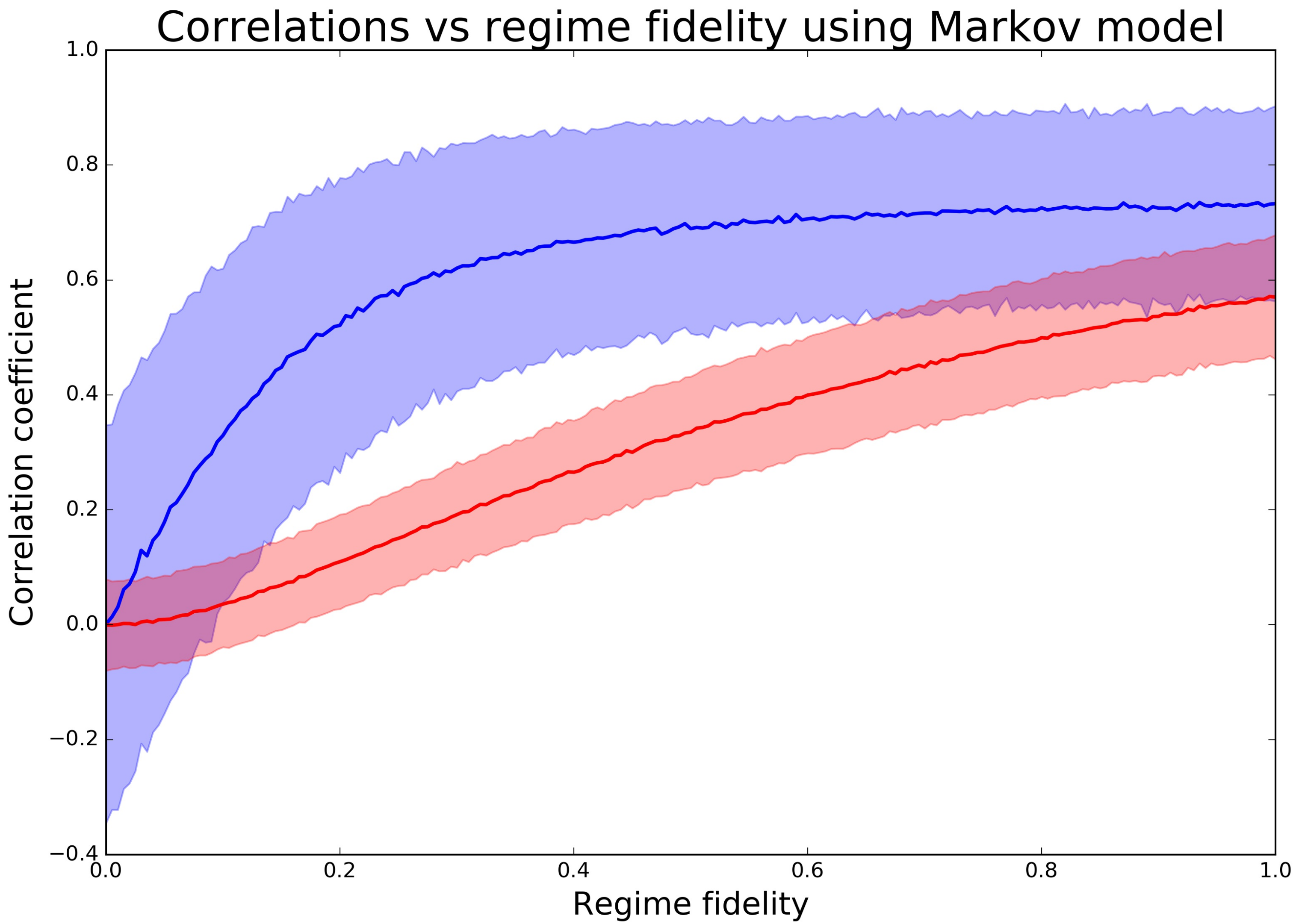
- E.g. GloSea5 (same model as DePreSys3)



3. Results

- Almost all relevant parameters/distributions are fitted to ERA-Interim.
- `Regime fidelity’ parameter k is left free: we let this vary to capture variations in model skill/error.
- Given a choice of k, we simulate 1000s of 35-year long, 40-member hindcasts and see what happens!

3. Results

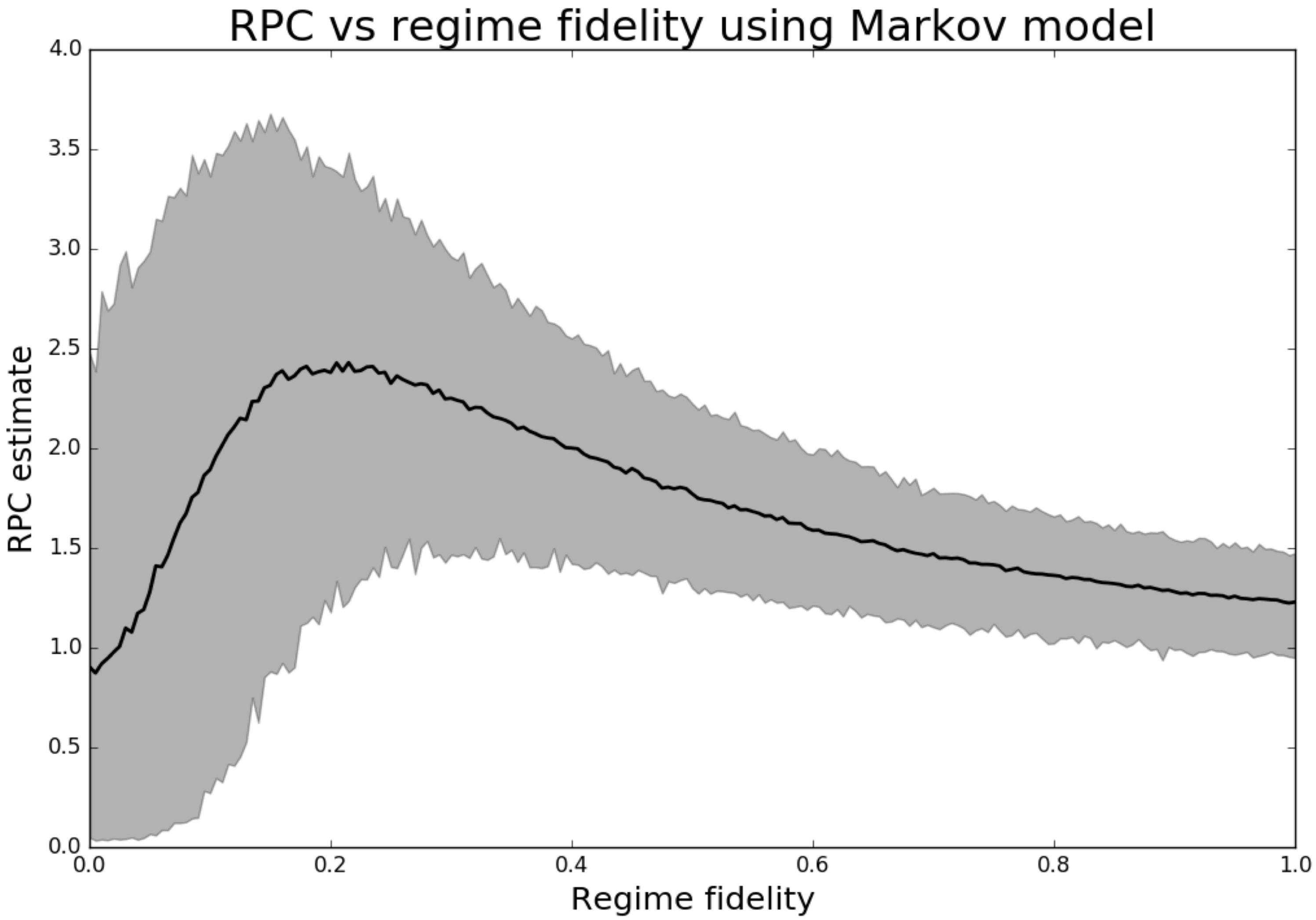


- Actual skill (blue)
- Potential predictability (red)
- Shading = 95% conf. intervals

3. Results

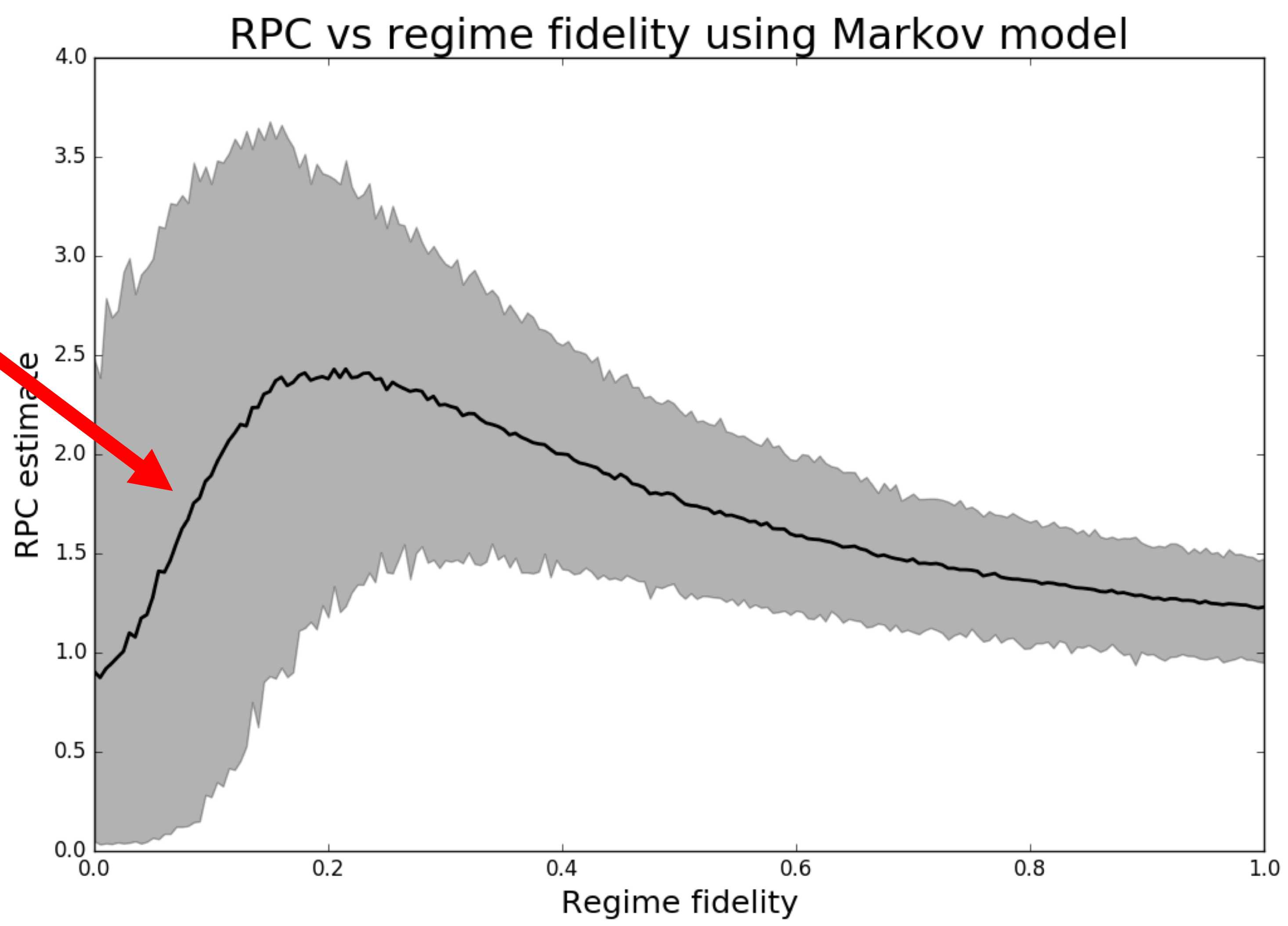
- For a <u>skillful</u> model, RPC estimates will always be > 1 in this system

- Hence true RPC is also > 1



3. Results

- At this low level of skill, 40 ensemble members are not enough to robustly estimate true RPC!
- True RPC goes to ∞ while estimate goes to fixed *finite* limit.
- So true RPC still > 1



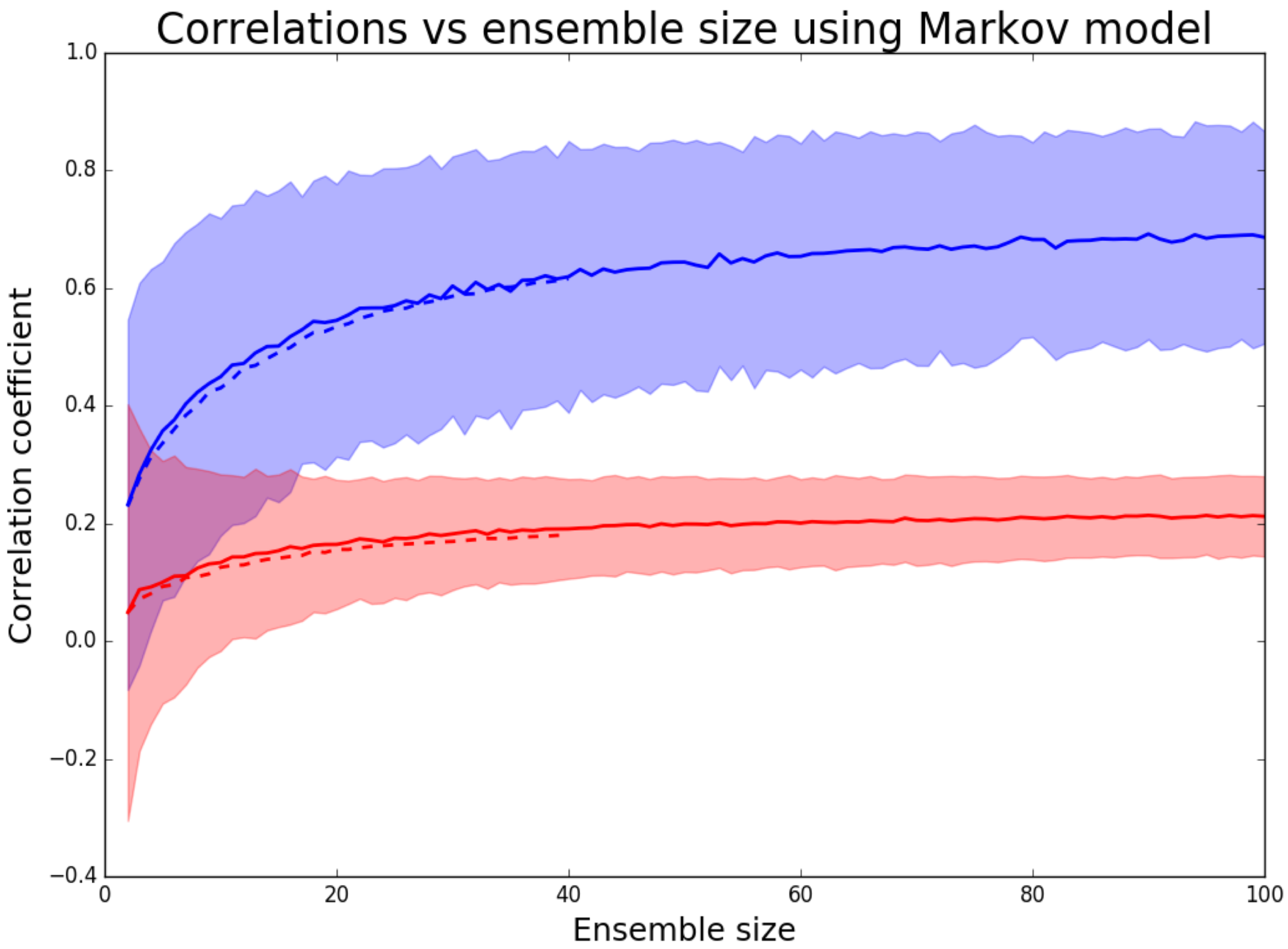
3. Results

- Can we reproduce DePreSys3 values?

- Yes! E.g. with k=0.3, expected values of three key numbers are:
  - $Corr(EnsMean, Obs) = 0.63$
  - $Corr(EnsMean, Mem) = 0.19$
  - $RPC = 2.28$

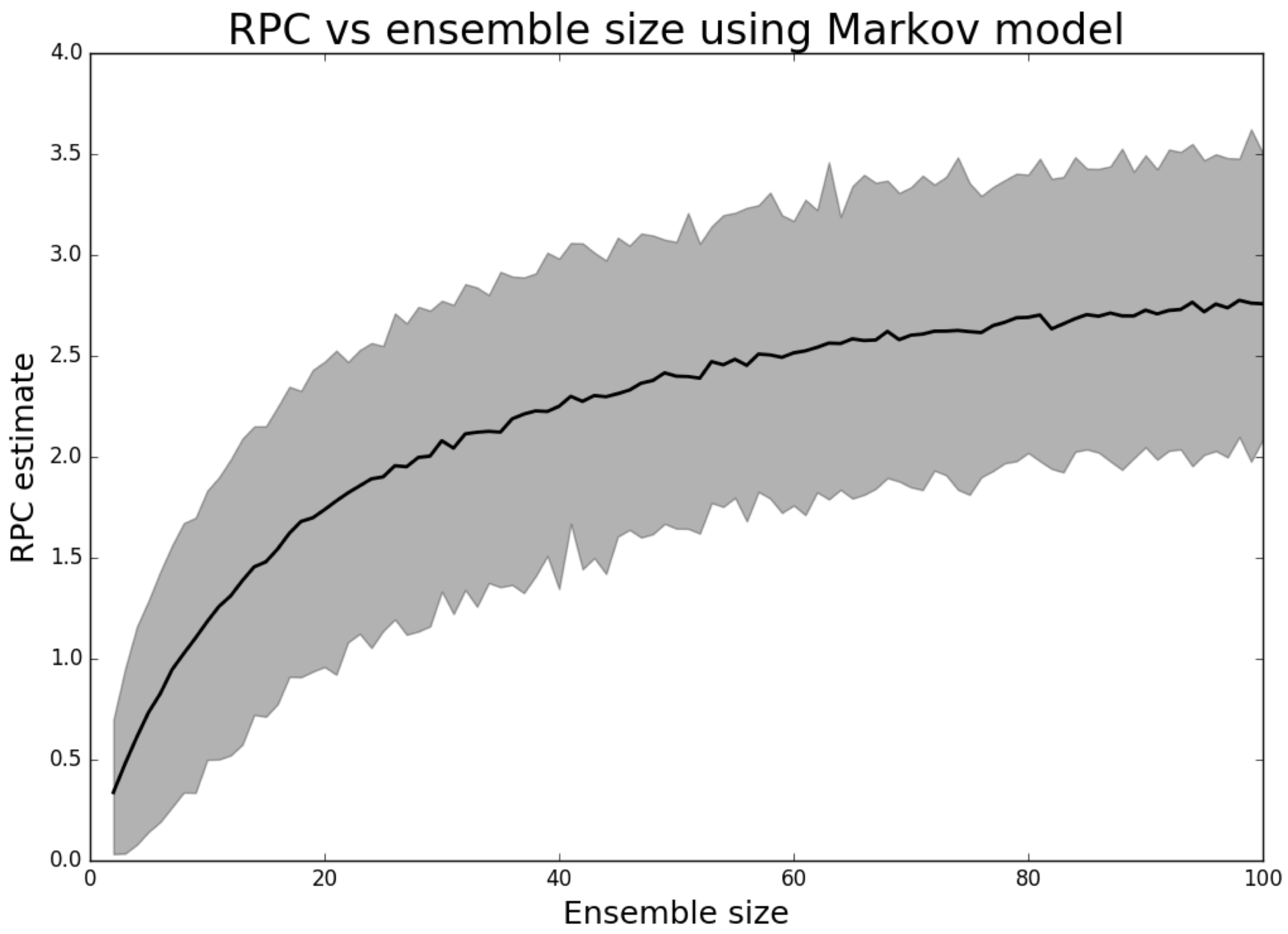
Virtually identical to actual DePreSys3 numbers

3. Results



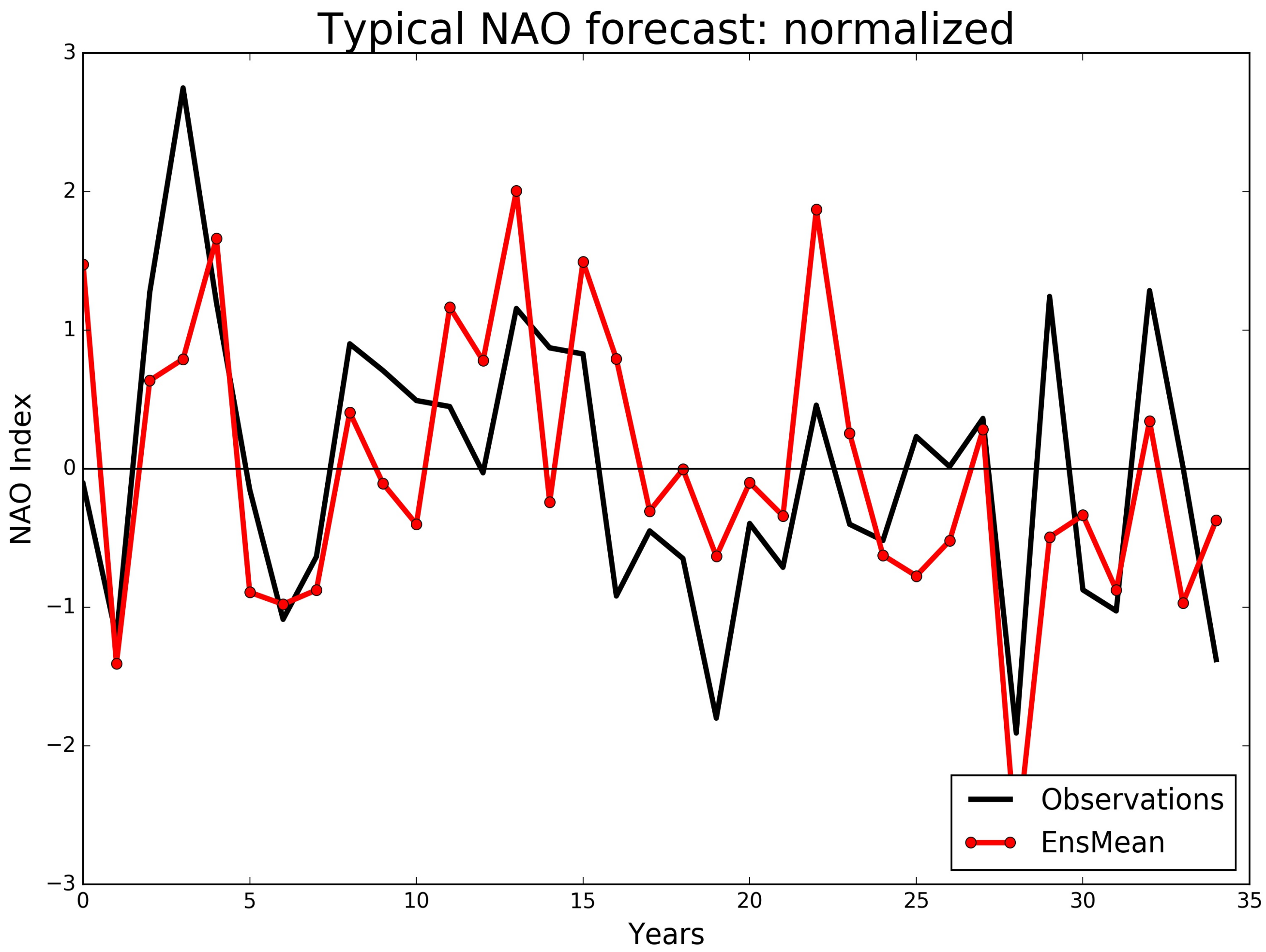
- With 200 ensemble members, correlation saturates at a maximum of around $\mathbf{0.75 \pm 0.2}$
- This is maximum skill that can be expected in this system.

3. Results

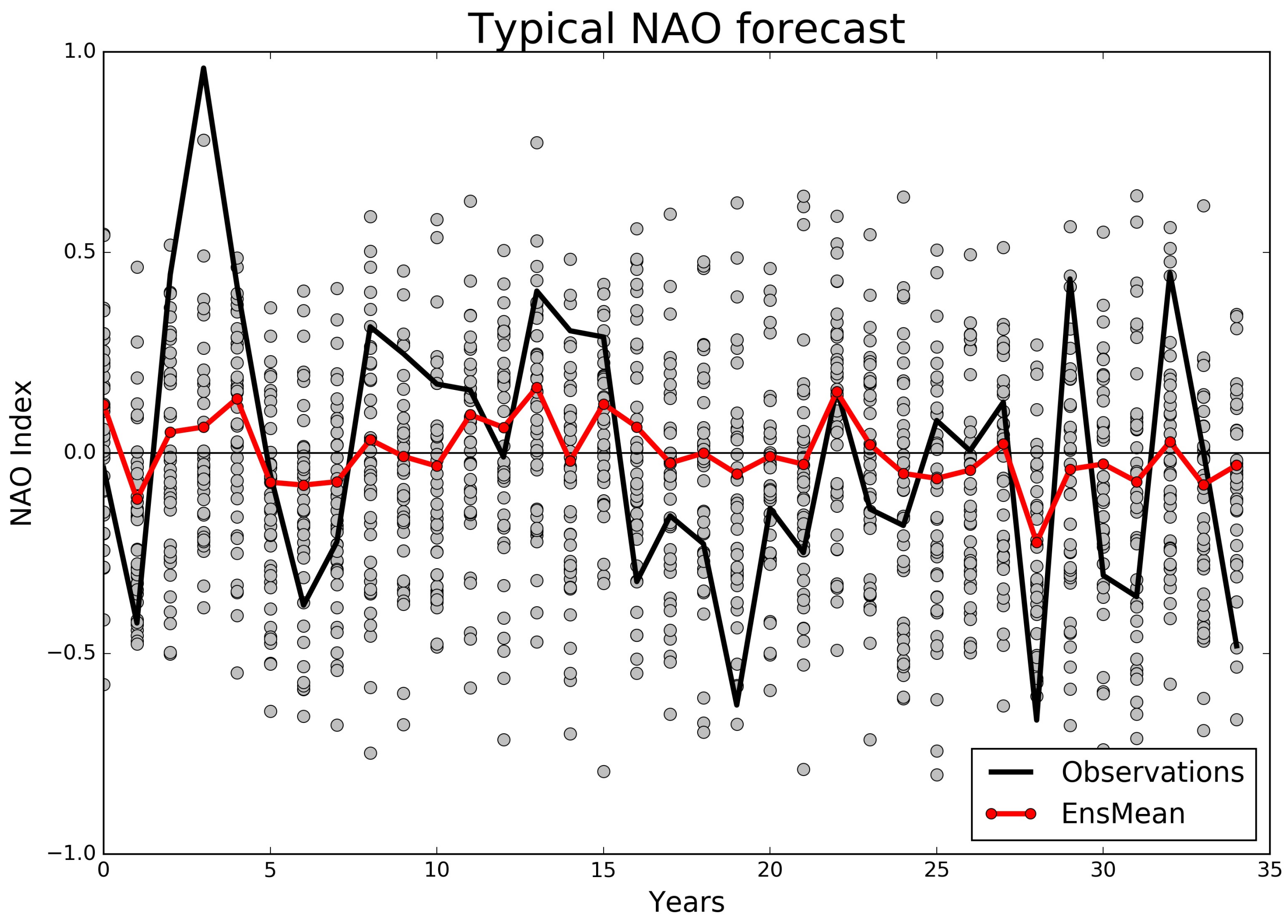


- **More ensemble members → better RPC estimate**
- With 1000 members, get $RPC = 3.2 \pm 0.5$
- Definitely >> 1!

3. Results


Typical NAO forecast: normalized
NAO Index
Years
3
2
1
0
−1
−2
−3
0
5
10
15
20
25
30
35
Observations
EnsMean

3. Results



Typical NAO forecast
NAO Index
Years
1.0
0.5
0.0
−0.5
−1.0
0
5
10
15
20
25
30
35
Observations
EnsMean

CONCLUSION

- \`Signal-to-noise paradox’ is *expected* in a bimodal regime system.
- Predictability in this system comes from persistence probabilities: if your model has systematic problems capturing them, get high RPC, *even with perfect teleconnections!*
- Suggests \`paradox’ may be a result of poor regime structure in NWP models. Known problem in many models.
- Unless the model skill is sufficiently high then a large ensemble (>100 members) is needed to robustly estimate the RPC*.* ***Show caution…***