Update on NAO prediction by NMME models,

started as QC on CESM data, recently added to the tank

and

Using the NAO prediction data (82-2010) by 8 models for finger exercises regarding weighting of models

Huug van den Dool 9/8/16

http://www.cpc.ncep.noaa.gov/products/people/wd51hd/ake/nmme/

Has directories for February Z200 EOF 1-4 at lead 0, 1, 2, ….7

CESM	22-Jul-2016 19:31	4K
CFSv2	06-Mar-2015 16:46	4K
CMC1	24-Nov-2014 15:35	4K
CMC2	24-Nov-2014 15:36	4K
CSM4	27-Feb-2015 17:59	4K
FLOA	04-Mar-2015 15:37	4K
FLOB	06-Mar-2015 18:15	4K
FLOR24	06-Mar-2015 16:43	4K
GFDL	24-Nov-2014 15:32	4K
NASA	24-Nov-2014 15:32	4K
NCAR	24-Nov-2014 15:36	4K
OBS	25-Nov-2014 14:31	4K

Skill in the prediction of NAO by ALL NMME Models

~First Observed EOF for February Lead+1, Z200, 1982-2010

	CFSv2	CCSM4	CMC1	CMC2	NASA	GFDL	NMME	FLOR24	CESM	OBS
Skill-AC	40.7	43.3	38.0	54.4	40.2	48.1	50.4	38.9	21.2	
Predictability	50.6	60.8	44.0	62.7	60.1	58.0	52.8	35.7	51.2	
Sd-individual	3.8	3.9	3.9	4.3	4.2	4.3	4.1	3.4	3.3	5.0
SD-ens.mean	2.1	2.7	2.2	3.1	2.9	3.0	2.2	1.5	2.1	

Recent NMME changes: March 2015: CCSM4 and FLOR24 added; June 2016 CESM added

All models have some, even 'decent', skill in the NAO prediction in week 4-8 averaged, which is a bit surprising.

Predictability is higher than the actual prediction skill in all models, but not (much) in NMME.

Huug van den Dool

Skill in the prediction of NAO by ALL NMME Models

~First Observed EOF for February Lead+1, Z200, 1982-2010

	CFSv2	CCSM4	CMC1	CMC2	NASA	GFDL	NMME	FLOR24	CESM	OBS
Skill-AC	40.7	43.3	38.0	54.4	40.2	48.1	50.4/49.1	38.9	21.2	
Predictability	50.6	60.8	44.0	62.7	60.1	58.0	52.8/48.7	35.7	51.2	
PAC3	27.5	22.8	17.6	25.3	12.8	17.2	23.5	10.4	2.5	
Sd-individual	3.8	3.9	3.9	4.3	4.2	4.3	4.1	3.4	3.3	5.0
SD-ens.mean	2.1	2.7	2.2	3.1	2.9	3.0	2.2	1.5	2.1	

Recent NMME changes: March 2015: CCSM4 and FLOR24 added; June 2016 CESM added

Huug van den Dool

Unequal Weights Project Update

Inputs from
Emily Becker
Li-Chuan Chen

Put together by Huug van den Dool 9/8/16

Model M_i (i=1,8) in NMME

CON (t) = $\sum_i \alpha_i M_i(t)$ consolidation, weighted average. α_i does NOT depend on time t.

Equal weight option $\alpha_1 = \alpha_2 = \alpha_3 = ... \alpha_8 = 1/8$; $\sum \alpha_i = 1$

MSE (or BS) minimization ***as principle*** to solve for α_i . Minimize $Q = \sum_t \{\sum_i \alpha_i M_i(t) - O(t)\}^2$

Solve for α_i as per **A** α = b (Equation (1)), where matrix **A** is .., vector α is .. , and vector b is .. .

$a_{11}\ a_{12} ....a_{18}$

. $\quad a_{ij} = \sum_t M_i(t) M_j(t) \quad\quad b_i = \sum_t M_i(t) O(t)$

.

$a_{81}\ a_{82} .... a_{88}$

Sofar: Unconstrained linear regression. ; $\sum \alpha_i \neq 1$

α is a set of regression coefficients, ~ to correlation coefficients (PAC if data is probs in %)

Co-linearity and redundancy among models

Stability (or applicability of α to independent future data)

Regression with constraint: example is ridge regression.

What is ridging of **A ?**, constrain $\sum_i \alpha_i^2$ to be small, why do we do ridging?.

Cross Validation, why do we do that?

CV0 (full sample), CV1, CV3, CV3R…

Safe 'fall-back' limiting options for CV3R when ridging does not yield success:

-) towards equal weights for large ridging - $\sum \alpha_i = 1$? (DelSole 2007; Bayesian)

-) towards weights proportional to skill for large ridging - $\sum \alpha_i = 1$? (Pena+vdDool 2008)

Study in physical units using ensemble means for M_i. Then repeat the whole thing with probabilities.

Huug

CFSv2	NASA	GFDL	CMC1	CMC2	Flor24	CCSM4	CESM			
<---------------Matrix A---8X8------------------------------------->								α	=	b
4.5	5.3	3.9	3.6	4.5	2.0	4.4	2.8	-0.04		4.3
5.3	8.3	5.7	3.9	6.2	2.9	6.0	4.6	0.16		5.8
3.9	5.7	8.6	3.6	6.0	1.8	5.4	3.3	0.45		7.0
3.6	3.9	3.6	5.0	5.1	1.6	4.3	2.4	-0.21		4.2
4.5	6.2	6.0	5.1	9.4	2.8	6.0	3.8	0.66		8.3
2.0	2.9	1.8	1.6	2.8	2.3	2.4	2.0	0.79		2.9
4.4	6.0	5.4	4.3	6.0	2.4	7.2	3.9	0.26		5.8
2.8	4.6	3.3	2.4	3.8	2.0	3.9	4.2	-1.08		2.2

All elements of A are positive. (Implications for stable solutions).
All elements of vector b are positive (all models have skill)
Why are some diagonal elements (variance) so small (2.3)
Is the solution, alpha vector, credible?

Those correlations (ens mean model i vs ens mean model j are very high.
Beyond heterogeneous predictability
Using only ensemble means (not members, or spread) unhelpful for a stable solution
Even CESM (low correlation to obs) correlates 0.52-0.78 to other models.

100.0	85.8	63.3	75.6	68.6	62.4	77.5	64.3
85.8	100.0	67.1	61.2	70.0	67.6	78.1	78.4
63.3	67.1	100.0	54.5	66.4	40.9	68.5	55.1
75.6	61.2	54.5	100.0	74.9	47.5	72.2	52.2
68.6	70.0	66.4	74.9	100.0	60.9	72.6	60.0
62.4	67.6	40.9	47.5	60.9	100.0	59.1	66.0
77.5	78.1	68.5	72.2	72.6	59.1	100.0	70.6
64.3	78.4	55.1	52.2	60.0	66.0	70.6	100.0

The full sample solution of Equation (1) for the NAO prediction in physical units.
The 8X8 matrix A shows only +ve entries!
Some diagonal elements are quite small. Why?
Rhs of (1), b vector, shows all models correlate +vely with obs
Solution alpha is mildly “off” (like a -1)
The safe solution b_i/a_{ii} indicates ‘non-equal’ weights, 0.08 to 0.19 on full sample.
$b_i/a_{ii} > 1$ for the 6th model . Won’t see that too often.
Is b_i/a_{ii} better? Upon CV3R the RMSE is about the same as equal weights.
How would all this be for prob % inputs, not ensemble mean phys units?? Stay tuned.

<	-	-	-	A	-	-	>		alpha	b	b_i/a_{ii}	$\sum \sim b_i/a_{ii} = 1$
4.5	5.3	3.9	3.6	4.5	2	4.4	2.8		-0.04	4.3	0.95	0.14
5.3	8.3	5.7	3.9	6.2	2.9	6	4.6		0.16	5.8	0.69	0.1
3.9	5.7	8.6	3.6	6	1.8	5.4	3.3		0.45	7	0.82	0.12
3.6	3.9	3.6	5	5.1	1.6	4.3	2.4		-0.21	4.2	0.85	0.12
4.5	6.2	6	5.1	9.4	2.8	6	3.8		0.66	8.3	0.88	0.13
2	2.9	1.8	1.6	2.8	2.3	2.4	2		0.79	2.9	1.29	0.19
4.4	6	5.4	4.3	6	2.4	7.2	3.9		0.26	5.8	0.8	0.12
2.8	4.6	3.3	2.4	3.8	2	3.9	4.2		-1.08	2.2	0.51	0.08

Huug

Example NAO prediction. Target = February 1982-2010, lead =1 month. **RMSE** of EOF#1 projection

Nature of forecast			
Always Climo	4.98 units (the no skill level) (=sd around climo)		
Equal weights (1/8)	4.35 (<4.98, so clear skill)		Phys
Solve (1) just once: Full sample unconstrained weights:	3.86 (if these weights were credible we could beat current NMME)	→ 4.05 upon ridging 5 - 50%	
Solve (1) 29 times, CV1	5.68 (disappointing)	5.68→4.70 upon ridging 5 - 50% (but no skill)	Units
Solve (1) 29 times, CV3R	5.94 (disappointing)	5.94→4.70 upon ridging 5 - 50% (but no skill)	
CV3R Ridge Regression with "safe" limiting options for infinite ridging (Bayesian) In particular 1. Equal weight 2. Weights ~ skill 3. Delmdl option			Huug

Example NAO prediction. Target = February 1982-2010, lead =1 month. **BS** of EOF#1 projection

Nature of forecast			
Always Climo	0.226 units (the no skill level) (=sd^2 around climo)		
Equal weights (1/8)	0.210 (Not much skill)		A %
Solve (1) just once: Full sample unconstrained weights:	0.158 (if these weights were credible we could beat current NMME)		
Solve (1) 29 times, CV1			
Solve (1) 29 times, CV3R	0.339 (disappointing)	0.339→0.216 upon ridging 5 - 50% (but no clear skill)	
CV3R Ridge Regression with "safe" limiting options for infinite ridging (Bayesian) **In particular** **1. Equal weight** **2. Weights ~ skill**			Huug

<	-	-	-	A	-	-	>		alpha	b	b_i/a_{ii}	$\sum \sim b_i/a_{ii} = 1$	
0.05	0.05	0.04	0.04	0.04	0.02	0.05	0.04		0.2	0.03	0.63	0.17	
0.05	0.07	0.04	0.04	0.05	0.02	0.06	0.05		-0.13	0.02	0.32	0.09	
0.04	0.04	0.07	0.04	0.04	0.01	0.05	0.04		0.33	0.03	0.42	0.11	
0.04	0.04	0.04	0.06	0.04	0.01	0.04	0.03		0.44	0.04	0.65	0.18	
0.04	0.05	0.04	0.04	0.08	0.02	0.04	0.03		0.7	0.06	0.76	0.21	
0.02	0.02	0.01	0.01	0.02	0.03	0.02	0.02		0.61	0.02	0.56	0.15	
0.05	0.06	0.05	0.04	0.04	0.02	0.07	0.05		-0.55	0.01	0.21	0.06	
0.04	0.05	0.04	0.03	0.03	0.02	0.05	0.06		-0.44	0.01	0.16	0.04	
									PAC				
100	81.2	63.5	77.3	68.6	46.4	77	65.9		31.02				
81.2	100	62.6	61.3	65.5	42.2	84.9	71.3		18.19				
63.5	62.6	100	54.6	51.3	23.2	67.2	57.8		23.48				
77.3	61.3	54.6	100	59.5	23.4	60.4	53.4		33.33				
68.6	65.5	51.3	59.5	100	40.4	57.8	49.4		44.25				
46.4	42.2	23.2	23.4	40.4	100	49.1	58.3		19.49				
77	84.9	67.2	60.4	57.8	49.1	100	78.6		11.49				
65.9	71.3	57.8	53.4	49.4	58.3	78.6	100		8.29				

The matrices for probs for the A tercile

Example NAO prediction. Target = February 1982-2010, lead =1 month. **BS** of EOF#1 projection

Nature of forecast			
Always Climo	0.215 units (the no skill level) (=sd^2 around climo)		B
Equal weights (1/8)	0.200 (Not much skill)		
Solve (1) just once: Full sample unconstrained weights:	0.146 (if these weights were credible we could beat current NMME)		
Solve (1) 29 times, CV1			
Solve (1) 29 times, CV3R	0.303 (disappointing)	0.303→0.204 upon ridging 5 - 50% (but no clear skill)	
CV3R Ridge Regression with "safe" limiting options for infinite ridging (Bayesian) In particular 1. Equal weight 2. Weights ~ skill			Huug

Example NAO prediction. Target = February 1982-2010, lead =1 month. **BS** of EOF#1 projection

Nature of forecast			
Always Climo	0.226 units (the no skill level) (=sd^2 around climo)		
Equal weights (1/8)	0.233 (Worse than no skill)		N
Solve (1) just once: Full sample unconstrained weights:	0.206 (if these weights were credible we have minimal skill)		
Solve (1) 29 times, CV1			
Solve (1) 29 times, CV3R	0.388 (terrible, to be expected)	0.388→0.310 upon ridging 5 - 50% (ridging can't save your)	
CV3R Ridge Regression with "safe" limiting options for infinite ridging (Bayesian) **In particular** **1. Equal weight** **2. Weights ~ skill**			Huug

Final

- NAO is only an example for weighting. Li-Chuan Chen is doing the whole project (global)
- CESM data look OK (except lead 0, watch phase II)
- Using distributional information may help weighting relative to using ensemble means (that type of hope had been dashed before)
- Preliminary impression, based on NAO and on TX rainfall (Li-Chuan): it is tough to calculate trustworthy weights, or to beat Equal weights, or skill-based weight. Some hope in special solutions and sub-sampling by deleting certain models ‘upfront’, when $b < 0$ or weight < 0 .