Evaluation of Leading Modes of Climate Variability in the CMIP Archives

Using the Climate Variability Diagnostics Package

J Fasullo, A Phillips, C Deser
CGD, NCAR

Journal of Climate, 2020 doi:10.1175/JCLI-D-19-1024.1
Science Questions

Have models improved over time? If so, how? What biases persist?

What are the temporal and spatial characteristics of model biases?

What modes are well-constrained by the observational record relative to model error?
The Climate Variability Diagnostics Package

Phillips et al. 2014, EOS
http://www.cesm.ucar.edu/working_groups/CVC/cvdp/

El Niño / Southern Oscillation (ENSO): Niño3.4 SSTA (Deser et al. 2012)

Pacific Decadal Oscillation (PDO): (Mantua et al. 1997)

Northern Annular Mode (NAM): Hurrell and Deser (2009)

Southern Annular Mode (SAM): Thompson and Wallace (2000)
The Climate Variability Diagnostics Package

Phillips et al. 2014, EOS
http://www.cesm.ucar.edu/working_groups/CVC/cvdp/

**El Niño / Southern Oscillation (ENSO):** Niño3.4 SSTa (Deser et al. 2012)

**Pacific Decadal Oscillation (PDO):** (Mantua et al. 1997)

**North Atlantic Oscillation (NAO):** Hurrell and Deser (2009)
**Northern Annular Mode (NAM):** Hurrell and Deser (2009)

**Southern Annular Mode (SAM):** Thompson and Wallace (2000)

All modes evaluated for 1920-2018 except SAM (post 1950).

Sensitivity of mode patterns to precise timeframe is small.

Impact of data uncertainties (e.g. SAM pre1950) however can be large.
Observations

**Sea Surface Temperature**
ERSSTv5 and HADISST
(Huang et al. 2017, Rayner et al. 2003)

**Sea Level Pressure**
ERA20C-ERAI, CERA20C-ERAI – not HadSLP2

**Surface Air Temperature**
BEST, GISTEMP
(Rhode et al. 2013, Lenssen et al. 2019)
CMIP Simulations

**CMIP3**: 16 models*; 77 simulations *(Meehl et al. 2007)*

**CMIP5**: 29 models*; 197 simulations *(Taylor et al. 2012)*

**CMIP6**: 22 models*; 426 simulations *(Eyring et al. 2016)*

**CESM1 Large Ensemble**: 40 members *(Kay et al. 2015)*

**MPI Grand Ensemble**: 40 of 100 members *(Maher et al. 2019)*

*using one model (newest non-ESM) from each modeling center to avoid overweighting

To estimate intrinsic noise

---

Evaluating Modes of Climate Variability in Models: Fasullo et al. 2020 *J. Climate*
Mode Intrinsic Variability

For which modes can a meaningful assessment be made?

Requirements for model evaluation:

1) Agreement between **observations** must be $>>$ model and obs.

2) **Range of intrinsic variability** must be $<<$ model structural uncertainty.
**Mode Intrinsic Variability**

*For which modes can meaningful assessments be made?*

Requirements for model evaluation:

1) Agreement between **observations** must be $>>$ model and obs.

2) Range of intrinsic **variability** must be $<<$ model structural uncertainty.

**ENSOSLP** (ENSO-$T_S$ is noisy)

El Niño / La Niña $T_S$ HOVs

PDO (IPO is noisy, obs uncer.)

NAO/NAM/SAM used cautiously. PNA not used due to mode-swapping issue.

Mean pattern correlations of key modes vs observations (bars)

$2\sigma$ inter-member spread of ensembles

MPI (dark bar), CESM (light bar)

CMIP5 $2\sigma$ range (green), Obs (red dots)
Methods: Bias EOF Decomposition

A method for characterizing model differences

PC analysis applied to mode bias patterns.

Observations also included in the PC analysis to provide an observational reference.

Incorporating multiple members of each model can be used to estimate intrinsic variability.

Evaluating Modes of Climate Variability in Models: Fasullo et al. 2020 J. Climate
Results: PDO (TS Regression on PC1)

**Obs**: WP – EP/Aleutian Low (AL)

**Model Std**: Largest at high lat (AL), deep tropics

**Mean Bias**: cor w/mean (too weak) particularly in tropics; reduction across generations.

**EOF1**: 32%; correlates with mean pattern (strength, r=-0.89)

**EOF2**: 16%; horseshoe structure in NP

**PCs**: Improvement in PC1 in CMIP6, not PC2. Upper tercile of CMIP3 scores > CMIP5.
Results: PDO (TS Regression on PC1)

**Obs:** WP – EP/Aleutian Low (AL)

**Model Std:** Largest at high lat (AL), deep tropics

**Mean Bias:** cor w/mean (too weak) particularly in tropics; reduction across generations.

**EOF1:** 32%; correlates with mean pattern (strength, $r=-0.89$)

**EOF2:** 16%; horseshoe structure in NP

**PCs:** Improvement in PC1 in CMIP6, not PC2. Upper tercile of CMIP3 scores > CMIP5.
Results: ENSO TS HOV
(Composite El Niño)

**Obs:** Warm anom 170E (max 135W) transition to cold early in Year 1

**Model Std:** largest in EP, also 165E @Yr0-1

**Mean Bias:** cor w/mean, too much ENSO variance. Improves in CMIP5/6.

**EOF1:** 45%; correlates with mean pattern (strength, r=-0.87), two CMIP3 much too biennial, most models not enough

**EOF2:** 19%; relates to warm phase duration; not systematic

Results: ENSO TS HOV  
(Composite El Niño)

**Obs:** Warm anom 170E (max 135W) transition to cold barely in Year 1  

**Model Std:** largest in EP, also 165E @Yr0-1  

**Mean Bias:** cor w/mean, too much ENSO variance. Improves in CMIP5/6.  

**EOF1:** 45%; correlates with mean pattern (strength, r=-0.87), some CMIP3 too biennial, most models not enough  

**EOF2:** 19%; relates to warm phase duration; not systematic  

Mean Overall and PDO scores improve across CMIP generation.

Magnitude of improvement exceeds the spread of internal variability (some dependence of spread on model exists).

Spread within CMIP generations is considerable, particularly for CMIP3. Improvement characterized to first order by elimination of worst performing CMIP3 models. Best models have also improved.
CVDP Score Summary

CMIP3 (77) CMIP5 (182) CMIP6 (250)

**Overall**: Detectible progressive improvement, worse models improve most.

**PDO**: Detectible progressive improvement, worse models improve most. $\sigma_{GE} > \sigma_{LE}$

**ENSO**: Progressive apparent improvement in teleconnections and HOVs, particularly wrt worst CMIP3 models. $\sigma_{GE} >> \sigma_{LE}$

**NAM**: Little discernible systematic change. Perhaps some improvement in mean model performance. Weak observational constraint.

**SAM**: High scores generally but little discernible systematic change.

Evaluating Modes of Climate Variability in Models: Fasullo et al. 2020 *J. Climate*
**ENSO/PDO Spectra**

**Obs:** Peak ENSO power in 2.5-6 yr band; considerable power across timescales. Peak PDO power greater at >10 yr.

**CMIP3:** a large spread in simulated Niño3.4 power 2.5-6 yrs, median model near obs except < 2.5 and > 10 yrs.

**CMIP5:** outlier models in CMIP3 have improved. Agreement with obs pretty good.

**CMIP6:** increased power in most cases, with obs lying at the edges of ensemble at times.

**GE/LE:** large model dependence of power, spread in GE/LE in instances > CMIP. Need to be used cautiously.

**No evidence that models systematically underestimate low frequency climate variability.**
CONCLUSIONS

Have models progressed in their representation of internal variability?

Yes – CVDP scores have generally increased across model generations for many modes (PDO/ENSO). S/N for extratropical modes is poor however and any apparent improvements are a challenge to distinguish from noise.
CONCLUSIONS

Have models progressed in their representation of internal variability?

Yes – CVDP scores have generally increased across model generations for many modes (PDO/ENSO). S/N for extratropical modes is poor however and any apparent improvements are a challenge to distinguish from noise.

*If so, how?*

**ENSO:** 1) the magnitude of teleconnections has strengthened, 2) anomalies that extend to far westward have become smaller, and 3) Year 1-2 transitions have become more realistic (more 2 yr La Niñas).

**PDO:** Connections to the tropics have become stronger.
CONCLUSIONS

Have models progressed in their representation of internal variability?

Yes – CVDP scores have generally increased across model generations for many modes (PDO/ENSO). S/N for extratropical modes is poor however and any apparent improvements are a challenge to distinguish from noise.

If so, how?

ENSO: 1) the magnitude of teleconnections has strengthened, 2) anomalies that extend to far westward have become smaller, and 3) Year 1-2 transitions have become more realistic (more 2 yr La Niñas).

PDO: Connections to the tropics have become stronger.

What biases remain common across models?

Many of the canonical biases, though reduced in magnitude, persist in CMIP6.
END
Methods: Bias Analysis

A central goal is to provide an understanding of patterns that differentiate models, not merely the mean patterns of bias or summary scores.

Taylor diagrams are commonly used: provides metrics on pat. cor. and variance.

But they are opaque regarding regional structures and implications of bias.

Orbe et al. 2020, J Climate, in revision.
Results: ENSO DJF PSL Composite (warm-cold)

**Obs:** WP–EP, Aleutian Low (AL)

**Model Std:** Largest at high lat (AL)

**Mean Bias:** cor w/mean, WP-EP strength; westward displacement of AL; reduction across generations.

**EOF1:** 29%; correlates with mean pattern ($r=-0.96$), relates to weakness

**EOF2:** 17%; westward extent of AL

**PCs:** Improvement in PC1 across all terciles in each generation. PC2 worsens in lower terciles (2,3) of CMIP6.
Results: ENSO TS HOV (La Niña)

**Obs**: Max in EP, cold anom extend to 160E, repeat at Yr 2 (2 yr events)

**Model Std**: Largest in EP, also west to 135E.

**Mean Bias**: Too cold from mid Y0 to mid Y1. Extend too far west. Too warm Y2 – too few 2 yr events.

**EOF1**: 39%; correlates with mean pattern (strength/west ext., r=0.62)

**EOF2**: 18%; timing and duration; not systematic

Results: NAO (DJF)

**Obs:** Icelandic Low (IL) vs Azores High (AH)

**Model Std:** Largest near AH, east of IL.

**Mean Bias:** weakness in pattern, too strong west of IL

**EOF1:** 42% Strength of mean pattern (r=−0.97)

**EOF2:** 26% Relative strength of AH

**PCs:** PC1 tercile (1) in CMIP6 closer to obs. Some reduction in PC2 in tercile (3) across generations. (Noise problematic).

Evaluating Modes of Climate Variability in Models: Fasullo et al. 2020 J. Climate
Results: Northern Annular Mode (DJF)

**Obs:** Arctic vs Azores/Aleutian

**Model Std:** Largest over Europe and north of Canada

**Mean Bias:** suggests weakness of the pattern in many areas.

**EOF1:** 31% relates to Arctic/midlat pattern strength (r=-0.77)

**EOF2:** 22% CoA in Europe

**PCs:** CMIP6 PC1 in upper tercile closer to obs. perhaps some reduction in EOF2
**Results: Southern Annular Mode (DJF)**

**Obs:** CoAs in south IO/Atl. Ant.

**Model Std:** Variance increases with latitude.

**Mean Bias:** @40S from SA eastward.

**EOF1:** 45% N/S dipole about 50S

**EOF2:** 17% Zonal band near 50S, meridional extent of mode.

**PCs:** Changes are generally small compared to intrinsic variability. Improvement in upper tercile for PC1/PC2 suggested. Others in noise.

---

Evaluating Modes of Climate Variability in Models: Fasullo et al. 2020 *J. Climate*