# Assessing model dependency in CMIP5 and CMIP6 based on their spatial dependency structure with probabilistic network models

### Catharina E. Graafland\* & José M. Gutiérrez

Instituto de Física de Cantabria (Universidad de Cantabria-CSIC), Santander, Spain

\*catharina.graafland@unican.es







# Assessing model dependency in CMIP5 and CMIP6 based on their spatial dependency structure with probabilistic network models

- Ensembles of Global Climate Models (GCMs) are the main tool available to study the future evolution of climate (e.g. CMIP5 and CMIP6).
- However, the problem of model interdependency (Boe, 2018) poses limitations for the application of multi-model simulations in practical applications.
- There is a need of automatic methods for assessing similarity of GCMs.

# Assessing model dependency in CMIP5 and CMIP6 based on their spatial dependency structure with probabilistic network models

We propose the use of **Directed Probabilistic Networks (DPNs)** 

A multidisciplinary approach for weather & climate

• **DPNs** have proven effective in learning the **spatial structure** underlying climate data using a complex network and encoding it as a multivariate probabilistic model (Graafland et al.,2020).

$$P(X1,\ldots,XN) = P(X1|\Pi X1) \times \ldots \times P(XN|\Pi XN)$$

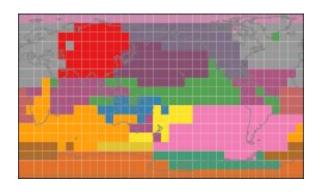
$$= P(X1|\Pi X1) \times \ldots \times P(XN|\Pi XN)$$

This DPN is constructed from **monthly mean surface temperature data** from a reanalysis dataset (ERA-Interim) for the 1981-2010 period interpolated to a 10 degrees grid with bilinear interpolation.

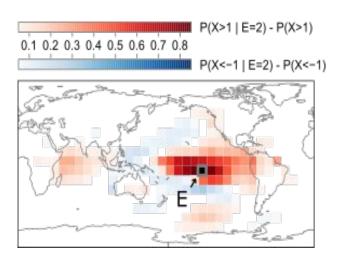
## Santander Meteorology Group A multidisciplinary approach for weather & climate

• DPNs provide the spatial backbone of complex data, modeling spatial dependencies between variables (in contrast to global statistical field measures such as mean, standard deviation) and only include significant dependencies (in contrast to networks based on correlation).

The backbone can be analysed with complex network measures like community structure detection:



but one is also able to extract **probabilistic features** like conditional probabilities given some evidence E.







• We have monthly mean surface temperature data from three reanalysis (ERA-Interim, NCEP-NCAR, and JRA55) from historical simulations from 33 GCMs participating in CMIP5 and from 24 GCMs participating in CMIP6 (Eyring, 2016), for the 1981-2010 period. The original resolutions of the datasets vary from 1 to 3 degrees. All datasets are interpolated to a 10 degrees grid with bilinear interpolation.



- We have monthly mean surface temperature data from three reanalysis (ERA-Interim, NCEP-NCAR, and JRA55) from historical simulations from 33 GCMs participating in CMIP5 and from 24 GCMs participating in CMIP6 (Eyring, 2016), for the 1981-2010 period. The original resolutions of the datasets vary from 1 to 3 degrees. All datasets are interpolated to a 10 degrees grid with bilinear interpolation.
- We compute Climate DPNs corresponding to the reanalysis and CMIP6 simulations and assume that similar CMIP5/CMIP6/Reanalysis models yield similar dependency structures and close DPNs:



- We have monthly mean surface temperature data from three reanalysis (ERA-Interim, NCEP-NCAR, and JRA55) from historical simulations from 33 GCMs participating in CMIP5 and from 24 GCMs participating in CMIP6 (Eyring, 2016), for the 1981-2010 period. The original resolutions of the datasets vary from 1 to 3 degrees. All datasets are interpolated to a 10 degrees grid with bilinear interpolation.
- We compute Climate DPNs corresponding to the reanalysis and CMIP6 simulations and assume that similar CMIP5/CMIP6/Reanalysis models yield similar dependency structures and close DPNs:
- We use a simple **metric to measure the distance between the DPNs**: the Bhattacharya distance between the underlying probabilistic models.

A multidisciplinary approach for weather & climate

## Assessing model dependencies with DPNs

- We have monthly mean surface temperature data from three reanalysis (ERA-Interim, NCEP-NCAR, and JRA55) from historical simulations from 33 GCMs participating in CMIP5 and from 24 GCMs participating in CMIP6 (Eyring, 2016), for the 1981-2010 period. The original resolutions of the datasets vary from 1 to 3 degrees. All datasets are interpolated to a 10 degrees grid with bilinear interpolation.
- We compute Climate DPNs corresponding to the reanalysis and CMIP6 simulations and assume that similar CMIP5/CMIP6/Reanalysis models yield similar dependency structures and close DPNs:
- We use a simple **metric to measure the distance between the DPNs**; the Bhattacharya distance between the underlying probabilistic models.

DPN 1

$$P_{M1} = p(W|X, Z)p(Z|X, Y)p(Y)p(X)$$

DPN 2

$$P_{M2} = p(W|X,Z)p(Z|X)p(Y)p(X)$$



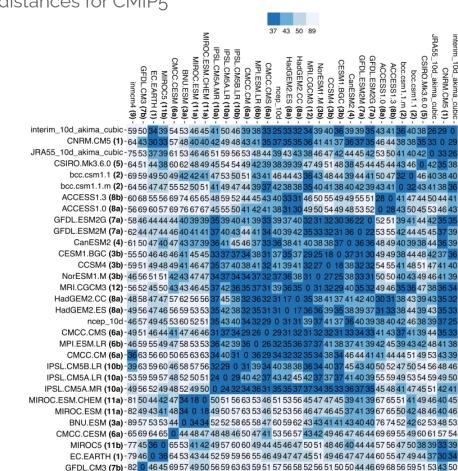
Bhattacharya distance between DPNs 
$$d_B(P_{M1}, P_{M2}) = -\log \int_{\mathcal{R}} \sqrt{p_{M1}(x)p_{M2}(x)} dx$$





A multidisciplinary approach for weather & climate

• Example: resulting sub-matrix of pairwise Bhattacharya distances for CMIP5

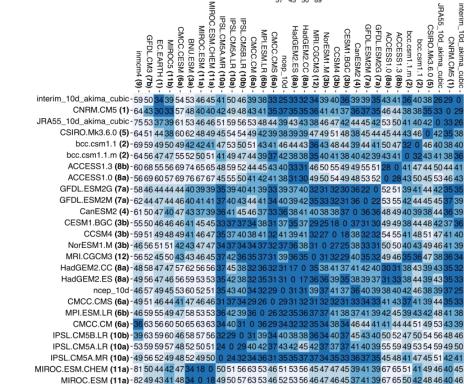


inmcm4 (9) - 0 82 79 77 65 89 82 81 49 53 39 36 46 49 46 49 48 56 46 59 55 61 62 58 56 60 64 69 64 75 64 59

A multidisciplinary approach for weather & climate

• Example: resulting sub-matrix of pairwise Bhattacharya distances for CMIP5.

 Matrix of pairwise distance between DPNs of CMIP5 and Reanalysis.

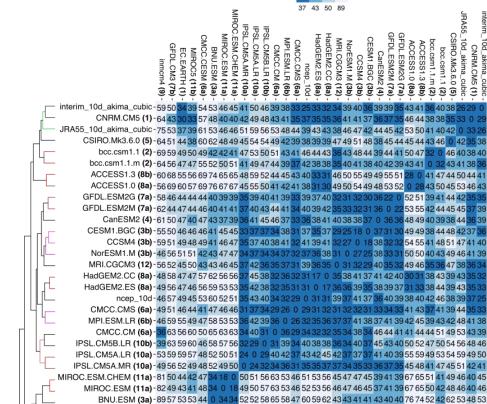


BNU.ESM (3a) -89 57 53 53 44 0 34 34 52 52 58 65 58 47 60 59 62 43 43 41 41 43 40 40 76 74 52 42 62 53 48 53 CMCC.CESM (6a) -65 69 64 65 0 44 48 74 84 84 65 0 47 41 53 56 57 43 42 49 46 47 46 44 69 69 55 49 60 61 57 54 MIROCS (11b) -77 45 36 0 65 53 41 42 49 57 60 60 49 44 45 46 47 50 51 48 46 40 44 44 57 66 47 50 38 39 33 39 EC.EARTH (1) -79 46 0 36 64 53 43 44 52 59 59 56 55 46 49 47 47 47 55 14 94 64 74 74 44 60 55 47 49 44 37 30 34 GFDL.CM3 (7b) -82 0 46 45 69 57 49 50 56 59 63 63 59 51 57 56 58 25 56 51 50 50 44 46 69 68 56 59 51 53 43 50 immcm4 (9) -0 82 79 77 65 89 82 81 49 53 39 36 46 49 46 49 48 56 46 59 55 61 62 58 56 60 64 69 64 75 64 59

A multidisciplinary approach for weather & climate

• Example: resulting sub-matrix of pairwise Bhattacharya distances for CMIP5

- Matrix of pairwise distance between DPNs of CMIP5 and Reanalysis.
- Hierarchical clustering of close DPNs

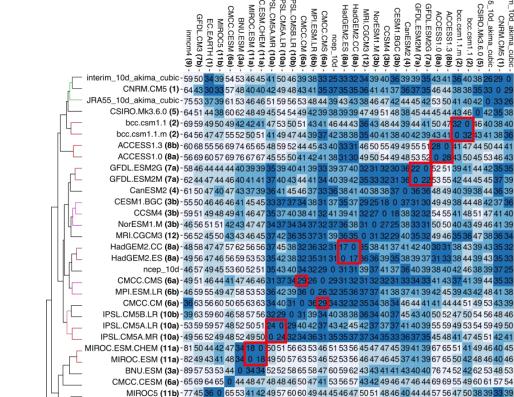


CMCC.CESM (6a) -65 69 64 65 0 44 48 74 88 46 50 47 41 53 56 57 43 42 49 46 47 46 44 69 69 55 49 60 61 57 54 MIROCS (11b) -77 45 36 0 65 53 41 42 49 57 60 60 49 44 45 46 47 50 51 48 46 40 44 44 57 56 47 50 38 39 33 39 EC.EARTH (1) -79 46 0 36 64 53 43 44 52 59 59 56 55 46 49 47 47 47 55 14 94 64 74 74 44 60 55 47 49 43 47 30 34 GFDL.CM3 (7b) -82 0 46 45 69 57 49 50 56 59 63 63 59 51 57 56 58 52 56 51 50 50 44 46 69 68 56 59 51 53 43 50 immcm4 (9) -0 82 79 77 65 89 82 81 49 53 39 36 46 49 46 49 48 56 46 59 55 61 62 58 56 60 64 69 64 75 64 59

A multidisciplinary approach for weather & climate

• Example: resulting sub-matrix of pairwise Bhattacharya distances for CMIP5

- Matrix of pairwise distance between DPNs of CMIP5 and Reanalysis.
- Hierarchical clustering of close DPNs
- Uncover interdependencies
  - same institute



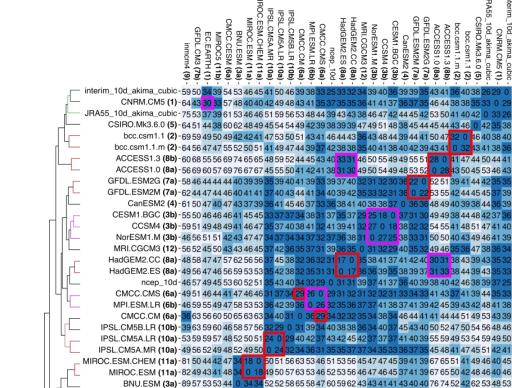
EC.EARTH (1) -7946 0 36 6453 43 44 52 59 59 56 55 46 49 47 47 45 51 49 46 47 47 44 60 55 47 49 44 37 30 34 GFDL.CM3 (7b) -82 0 4645 69 57 49 50 56 59 63 63 59 51 57 56 58 52 56 51 50 50 44 46 69 68 56 59 51 53 43 50 inmcm4 (9) -0 82 79 77 65 89 82 81 49 53 39 36 46 49 46 49 48 56 46 59 55 61 62 58 56 60 64 69 64 75 64 59

A multidisciplinary approach for weather & climate

• Example: resulting sub-matrix of pairwise Bhattacharya distances for CMIP5

- Matrix of pairwise distance between DPNs of CMIP5 and Reanalysis.
- Hierarchical clustering of close DPNs
- Uncover interdependencies
  - same institute

same atmospheric model

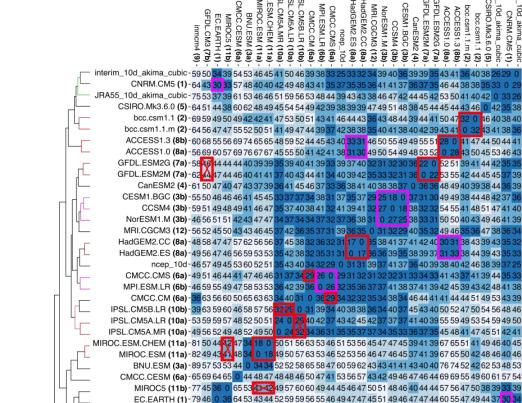


CMCC.CESM (6a) -65 69 64 65 0 44 48 47 48 48 46 50 47 41 53 56 57 43 42 49 46 47 46 44 69 69 55 49 60 61 57 54 MIROCS (11b) -77 45 36 0 65 53 41 42 49 57 60 60 49 44 45 46 47 50 51 48 46 40 44 44 57 56 47 50 38 39 33 39 EC.EARTH (1) -79 46 0 36 64 53 43 44 52 59 59 56 55 46 49 47 47 45 51 49 46 47 47 44 67 56 47 49 44 37 30 34 6FDL.CM3 (7b) -82 0 46 45 69 57 49 50 56 59 63 63 59 51 57 56 58 52 56 51 50 50 44 46 69 68 56 59 51 53 43 50 immcm4 (9) -0 82 79 77 65 89 82 81 49 53 39 36 46 49 46 49 48 56 46 59 55 61 62 58 56 60 64 69 64 75 64 59

A multidisciplinary approach for weather & climate

• Example: resulting sub-matrix of pairwise Bhattacharya distances for CMIP5

- Matrix of pairwise distance between DPNs of CMIP5 and Reanalysis.
- Hierarchical clustering of close DPNs
- Uncover interdependencies
  - same institute
  - same institute +
    significant atmospheric
    model change
  - same atmospheric model

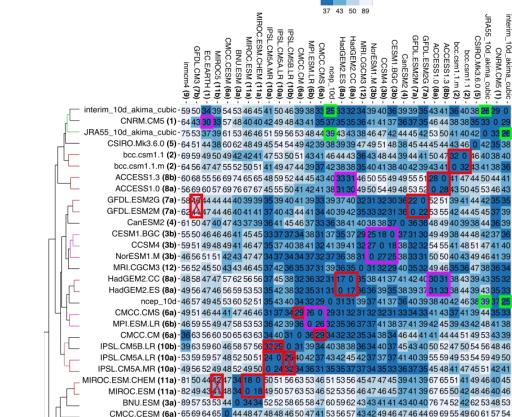


GFDL.CM3 (7b) -82 0 46 45 69 57 49 50 56 59 63 63 59 51 57 56 58 52 56 51 50 50 20 45 69 68 56 59 51 53 43 50 inmcm4 (9) -0 82 79 77 65 89 82 81 49 53 39 36 46 49 46 49 48 56 46 59 55 61 62 58 56 60 64 69 64 75 64 59

A multidisciplinary approach for weather & climate

• Example: resulting sub-matrix of pairwise Bhattacharya distances for CMIP5

- Matrix of pairwise distance between DPNs of CMIP5 and Reanalysis.
- Hierarchical clustering of close DPNs
- Uncover interdependencies
  - same institute
  - same institute +
    significant atmospheric
    model change
  - same atmospheric model
  - reanalysis vs reanalysis



MIROC5 (11b) -77 45 36 0 6553 222 49 57 60 60 49 44 45 46 47 50 51 48 46 40 44 44 57 56 47 50 38 39 33 39 EC.EARTH (1) -79 46 0 36 64 53 43 44 52 59 59 56 55 46 49 47 47 45 51 49 46 47 47 44 60 55 47 49 44 37 30 34 GFDL.CM3 (7b) -82 0 46 45 69 57 49 50 56 59 63 63 59 51 57 56 58 52 56 51 50 50 43 46 69 56 59 51 53 43 50

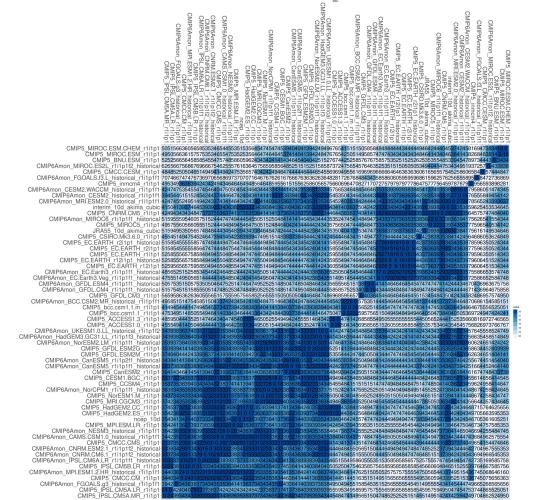
inmcm4 (9) - 0 82 79 77 65 89 82 81 49 53 39 36 46 49 46 49 48 56 46 59 55 61 62 58 56 60 64 69 64 75 64 59



• We extend the analysis adding CMIP6 models to identify significant **model changes** that underwent **CMIP5 GCMs in their transition to CMIP6**.

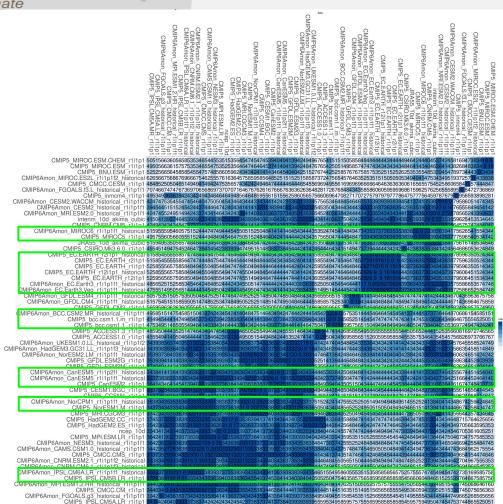
A multidisciplinary approach for weather & climate

 Matrix of pairwise distance between DPNs of CMIP5 and CMIP6 and Reanalysis.



A multidisciplinary approach for weather & climate

- Matrix of pairwise distance between DPNs of CMIP5 and CMIP6 and Reanalysis.
- Under CMIP6 some models hold a similar spatial dependency structure



48474149565150474852525251514551495142414239424970486252495

A multidisciplinary approach for weather & climate

 Matrix of pairwise distance between DPNs of CMIP5 and CMIP6 and Reanalysis.

 Whereas other CMIP6 models show significant change in spatial dependency structure with respect to their predecessors



## S

#### Santander Meteorology Group

A multidisciplinary approach for weather & climate

CMIP5 models
 GFDL ESM2G and
 GFDL ESM2M did
 not appear close to
 GFDL CM3:

They are build on a distinctively different atmospheric model. (Donner et al. 2011)

CMIP6Amon\_GFDL.ESM4\_r1i1p1f1\_historical-505753515057535050464752514746474952525148504748504846434356555651513819 0 414146464645454747474741424443467043696367575

CMIP6Amon\_GFDL.CM4\_r1i1p1f1\_historical-5157545151585550514748525347484950525251485147495048464445585857525340 0 194344484848474849474849474849474849671436964675858 CMIP5\_GFDL.CM3\_r1i1p1-56595863626360555651575559575658525655515050505044464847486968595657 0 403845464646464551534548435042434582486957574950

CMIP5\_GFDL.ESM2G\_r1i1p1 35393941414038363533363439393740323131323036323322 0 3032335251413944464848424344434242424444535353535343658634452403939

CMIP5 GFDL.ESM2M r1i1p1 374040444243393637343938414039423533343231363535 0 223133335355444247445050434447474645454545444637393433356263465140414

A multidisciplinary approach for weather & climate

CMIP6Amon

 CMIP6 models GFDL ESM4 and GFDL CM4 build on the same atmospherical model that was used for CMIP5 GFDL CM3.

CMIP6Amon\_GFDL.ESM4\_r1i1p1f1\_historica 505753515057535050464752514746474952525148504748504846434356555651513819 0 4141464646454547474747414244434670436963675758

CMIP6Amon\_GFDL.CM4\_r1i1p1f1\_historica 5157545151585550514748525347484950525251485147495048464445585857525340 0 194344484848484849474849414244444671436964675858

CMIP5 GFDL.CM3 r1i1p1 56595863626360555651575559575658525655515050505044464847486968595657 0 403845464646464655153454843504243458248695757495

CMIP5\_GFDL.ESM2G\_r1i1p1-35393941414038363533363439393740323131323036323322 0 30323352514139444648484243444342424444535353535343658634452403939

CMIP5\_GFDL.ESM2M\_r1i1p1-374040444243393637343938414039423533343231363535 0 223133335355444247445050434447474645454545444637393433356263465140414

A multidisciplinary approach for weather & climate

 Interestingly, GFDL ES2MG and GFDL ESM2M were closer to all 3 Reanalysis

models

CMIP6Amon GFDL.ESM4 r1i1p1f1 historical-50575351505753505046475251474 CMIP6Amon GFDL.CM4 r1i1p1f1 historical 51575451515855505147485253 474 849505252514851474950484644455858575253 40 0 1943444848484748494 74 34941 124 4444671436964675858 CMIP5 GFDL.CM3 r1i1p1-5659586362636055565157555957 658525655515050505044464847486968595657

3525141394446484842434443424242444 CMIP5\_GFDL.ESM2G\_r1i1p1 353554442474450504344474746454545<mark>4</mark>54 CMIP5 GFDL.ESM2M r1i1p1 37404044424339363734393841 401 942



#### References:

Boé, J. (2018) Interdependency in Multimodel Climate Projections: Component Replication and Result Similarity. Geophys. Res. Lett. 45, 2771–2779. DOI: 10.1002/2017GL076829

Eyring, V. et al. (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci. Model. Dev. 9, 1937–1958, DOI: 10.5194/gmd-9-1937-2016

Graafland, C.E., Gutiérrez, J.M., López, J.M. et al. (2020) The probabilistic backbone of data-driven complex networks: an example in climate. Sci Rep 10, 11484. DOI: 10.1038/s41598-020-67970-y

Donner, L. J. et al. (2011) The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3. J. Clim. 24, 3484–3519, DOI: 10.1175/2011JCLI3955.1 (2011).

#### **Acknowledgements:**

The authors would like to acknowledge project ATLAS (PID2019-111481RB-IO0) funded by MCIN/AEI (doi:10.13039/501100011033). We also acknowledge support from Universidad de Cantabria and Consejería de Universidades, Igualdad, Cultura y Deporte del Gobierno de Cantabria via the "instrumentación y ciencia de datos para sondear la naturaleza del universo" project for funding this work. C.E..G. acknowledges support from the Spanish Agencia Estatal de Investigación through the Unidad de Excelencia María de Maeztu with reference MDM-2017-0765.