

Spy4Cast: a Python Tool for statistical seasonal forecast based on Maximum Covariance Analysis. Review: Anonymous Referee #2

Duran-Fonseca, Pablo and Rodriguez-Fonseca, Belen

In this response we have gone through the comments made by Anonymous Referee #2. We appreciate the effort at revising the manuscript and we have taken into account the corrections and the suggestions of the referee.

1 Major Comments

Given the current scope of the software, I strongly recommend providing thorough documentation and a user manual. Additional examples of its application would further enhance its value to the community.

Thank you for your comment. Anonymous referee #1 had a similar comment regarding documentation. To address it we made a clearer reference to the documentation (currently changed to [readthedocs.io](https://spy4cast.readthedocs.io): <https://spy4cast.readthedocs.io>). We have also improved the user manual at <https://github.com/pabloduran016/Spy4CastManual> by adding `Tutorial.ipynb` and other examples you are welcome to explore, which shows in detail part of the functionality that Spy4Cast implements. In the new published version, Spy4Cast repository has a folder `examples` which contains itself the user manual.

Furthermore, as a comprehensive tool for statistical seasonal forecasting, the software would benefit from referencing influential prior work that incorporates statistical analysis and cross-validation. This would provide a more complete methodological context. A quick Google search yields references such as 'Cross-Validation in Statistical Climate Forecast Models' by Barnston and van den Dool (1987). I would encourage the authors to conduct a more thorough search to identify additional relevant studies.

Thanks for the comment. We have added a more methodological context, including references as the one suggested (which in fact is Michaelsen [1987]), but also others in which these methodologies have been applied. An extended description of the methodology is also included. In this list of references, we include the work of Wilks [2014] in which the author compares different methodologies including CCA and MCA, suggesting the use of MCA. Also we cite Michaelsen [1987], Elsner and Schmertmann [1994], Johansson et al. [1998]

Since this software provides a ready-to-use implementation of an existing method rather than introducing a new approach, what are its future prospects? Do you have plans to expand its functionality or broaden its applicability?

Thanks for the comment. Spy4Cast offers a new approach regarding its implementation, not the methodology. It is true that the way it was written in the manuscript could lead to misunderstandings and sounded pretentious. We have indicated that in the revised manuscript.

The idea behind the creation of this API was to apply it to a broad number of datasets to assess predictability. We are in the era of DATA and we can test many data sets to assess predictability of different target variables. This let us to better explore teleconnections and impacts and build hypothesis. Thus, we need a easy to run methodology that allows us to perform simulations in a fast and easy way, including

the selection of regions and seasons, the preprocessing, application of the MCA, crossvalidation and plotting. There are many applications that can be done, that we have indicated in the new version (with a link to some examples available in the manual). For example:

1. Identify coupled patterns in climate data (all examples)
2. Evaluate the ACC as a function of the lag between the predictor and predictand field using different start dates and forecast times (`Analysis_of_Predictability.py`)
3. Assess stationarity of the teleconnections (`Pacific_Impact_European_Rainfall.ipynb`).

In addition, other applications could be the assessments of the role of different drivers on the skill of a target variable to be predicted, the comparison between the skill of different simulations provided by different CMIP models, the comparison with the skill of dynamical forecasts etc.

In the new version of the manuscript, we have included these different applications of the MCA

2 Detailed comments

Line 1: Did you mean a dimension reduction technique?

We have changed the sentence to “Maximum Covariance Analysis (MCA) is a reduction dimension technique”

Line 10: How do you test model sensitivity to particular years? Did you say this in the manuscript?

Spy4Cast analyses the model sensitivity to particular years, including a diagnosis of the stability of the obtained modes to particular outliers. When we do a one-year out crossvalidation, we are repeating the MCA without the target year. In this way, we can store the information of the mode, and test the results (scf, ruv, U, V, RUY, RUZ etc) without that year. This let us know this sensitivity. We have further explained this in the new version. An example of this is jupyter notebook `Pacific_Impact_Sahelian_Rainfall.ipynb` where we use `version=2` to plot crossvalidation (section *Analyze the importance of the first 3 modes*)

Line 11: How would you test modes to particular outliers? Did you mention this? Or is this implied by testing different batch of years?

This is explained in the previous answer and we have further explained this in the new version of the manuscript

Line 14: Is the software fully documented? As it stands, it would benefit from additional work to develop a more comprehensive manual.

The software is fully documented. It also has a manual and a list of examples in the source directory. It also has a small suite of unit tests to ensure safe scalability.

Line 20-25: SST and SLP patterns which are highly ‘correlated’. You need to clarify an important distinction. Your description of the coupling between SST and SLP refers to only one phase of the Southern Oscillation or ENSO. However, it applies to both El Niño and La Niña, not just El Niño. The current wording suggests that only El Niño is being considered.

Thanks for the comment. In the new version of the manuscript we have changed it to: “For one of the phases of the expansion coefficients, the SST pattern will have the structure of El Niño and the SLP pattern will have the structure of the Southern Oscillation in its negative phase; which is the co-varying

SLP pattern forced by El Niño. The opposite will be obtained for the other phase of the expansion coefficients”

Line 28: a baseline for seasonal forecasts? Why

We want to apologize for not having further developed this idea in the previous manuscript. As dynamical models suffer for having biases, this statistical tool can serve to identify the dominant patterns of variability in the observations and help to correct or select members in operational predictions. In this way, one could combine the use of the statistical forecast provided by Spy4Cast and those provided for the different members of the dynamical seasonal forecast, calculating the ensemble mean of those members correlated with the statistical prediction.

Line 30-33: Machine learning methods are rapidly evolving, and you should reference more recent studies to support this statement. For example, Toride et al. (2024, <https://arxiv.org/abs/2404.15419>) demonstrates the use of neural networks to identify physical relationships and find predictability.

In this new version we have discussed other machine learning and deep learning techniques including references and discussing the advantages and disadvantages of their use.

Line 41-42: Instead of using the phrase ‘a new paradigm,’ it would be more accurate to reference earlier studies identifying the connection between ENSO and other tropical basins. For example, the connection between ENSO and the Indian Ocean Dipole (IOD) was first identified by Saji et al. (1999): A Dipole Mode in the Tropical Indian Ocean (Nature). This study demonstrated how the IOD can influence ENSO dynamics and has been foundational in the field.

Thanks for the comments. As we use Spy4CAST to predict ENSO using the Equatorial Atlantic as predictor, we mentioned just this paradigm. Nevertheless, we have considered convenient to also include other tropical basins impacting ENSO, as the one suggested by the referee. Our methodology could also be used and, in fact, following Saji et al (1999) we are also going to include in the new version the prediction of ENSO from the Indian Ocean.

Line 46-47: reliable? The citation at the end of the sentence is incomplete.

Thanks for the comment. We want to apologize as we did not include the complete reference, which is now included and it is Counillon et al. [2021].

Line 56: ‘... not designed to assess stationarity’ seems to contradict Line 250: ‘Spy4Cast is able to perform a validation methodology to look for non-stationary relations.’ Line 260 as well.

We have corrected this sentence and also provided an example of an assessment of stationarity, as in Rodríguez-Fonseca et al. [2016].

Line 59: fix the reference

Thank you for the comment, we will fix it.

Line 69: unit tests?

Thank you for the comment. The spelling is fixed now.

Line 76: Section 4 is an example of using Atlantic to predict Pacific SSt. However, if you have the Sahelian rainfall example, I think it would be useful to include in your manual/documentation and showcase different settings and functionalities of your work.

Thank you for pointing that out. All the examples are included in the documentation and in the manual. Also, in the new version of the manuscript, the snippets of code have been removed to improve readability.

Line 149: only being able to take monthly data seems limited capability to me.

The model could run with daily data for other applications, but it is specifically designed for seasonal predictability, in which monthly to seasonal means are used. Nevertheless, it is not difficult to change the Preprocess class to do it.

Line 190: How do you determine the sample size? Monthly data are likely highly correlated, i.e., each month is not an independent point, you need to use the effective sample size when you do statistical analysis.

We are not using consecutive montly data. The sample interval is normaly 1 year.

Line 195, 197 and more: What table or listing are you referring to? In general, when referencing your previous paper on which this software is built, I suggest specifying the relevant tables, listings, or figures. This would make it easier for users to trace the code development and better understand the overall concept.

Thank you for your comment. In those lines we refer to the documentation and to the manual. We are aware it is not easy to follow reference this way so in the newer version we have made explicit reference to the documentation and the manual instead of the previous citations.

Line 223: It says 2010 in the listing.

Thank you for pointing it out. we have removed the listings in the revised manuscript.

Line 224: You mention non-stationarity multiple times in the manuscript, but it is unclear how you determine it. Since there are various methods to assess non-stationarity, I recommend specifying the approach you used to ensure clarity.

Thanks for the comment. We think that the text was confusing because one can validate a model for a period, training the model with the data of another period, and find skillful prediction because the signal is stationary. But also, it can happen that, training the model with the data of another period we don't find skillful predictions because the signal is not stationary. So, this is not the way to test stationarity a priori. So, we have re-written the section to better explain the usefulness of the validation and a particular application for non stationary signals. In addition, in this new version of the manuscript we have better specified the way we can assess non-stationarity. There is a parameter associated with each of the modes of variability, ruv, which gives the degree of linear relationship between the expansion coefficients of a particular mode. If this number is not close to 1, that means that the relationship is not robust along the whole period. One of the most common techniques to assess this non-stationary behavior is to calculate moving correlations between the expansion coefficients. In the Github we have included an example about the relation between the anomalous boreal spring rainfall in Europe and Pacific SSTs as the one shown in Rodríguez-Fonseca et al. [2016] to illustrate how can we use Spy4cast to asesss non stationarity: `Pacific_Impact_European_Rainfall.ipynb`

Line 225: What does 'a hot spot' in climate variability studies mean?

Sorry for the error. We meant "hot topic". This is now corrected

Line 241: I think 'can be represented' is a more accurate phrasing.

Thank you, we have rephrased it taking it into account.

Line 246 Us should be in math form?

Thank you for pointing it out.

Line 249: The rest of modes... this statement is misleading and not accurate. Fig 5 seems to say 68% instead of 76%?

Thanks for the comment. We have changed this statement to address this comment.

Line 254: can you say what years? 94 and 91 for example?

Thanks for the comment. It is true that not all the years, for example 1994 and 1995, present skill. We were talking about those years in which the modes presents high values in the expansion coefficients, indicating years in which there was an Atlantic Niño(Niña) and a Pacific La Niña (El Niño). We have rewritten the sentence to better indicate the years we refer to.

Line 260: I am not sure you have explained how to use your software to determine stationarity

Thanks for the comment. It is true that we have not indicated how to asses stationarity. In this new version we have added an example in the Github and in the manual of the API in which the stationarity is assessed. It is indicated in a previous comment.

Line 265: This is not a ‘new’ approach but rather a ready-to-use software implementation of a well-established method for seasonal forecasting.

Thanks for the comment. It is true that the methodology is well established but the python approach not. We have to admit that the sentence was pretentious and we have changed it, indicating, as suggested by the referee that “Spt4CAST is a ready-to-use software implementation of a the MCA methodology to be used to asses seasonal forecast but also non stationarities”

Line 270: What is OFF project again?

In the new version of the manuscript we have indicated what is this project. ”OFF: Ocean For Future” is a National Project TED2021-130106B-I00) from the European Commission “Next Generation” funds in which predictability is assessed for all CMIP6 model under different scenarios using this tool coupled to the ESMVALTOOL. We have acknowledged this project in this manuscript.

Figure 3 caption: You need to label what year this is. Is it 1997 based on List 4?

Thank you for comment. We have changed the caption to make it clear.

Listing 6: Do you need to ‘import Preprocess’ first in this script?

Yes, it was a typo but, as it has been explained previously, listings have been removed in the new version of the manuscript.

References

- François Counillon, Noel Keenlyside, Thomas Toniazzo, Shunya Koseki, Teferi Demissie, Ingo Bethke, and Yiguo Wang. Relating model bias and prediction skill in the equatorial atlantic. *Climate Dynamics*, 56:2617–2630, 2021.
- Ibrahima Diouf, Roberto Suárez-Moreno, Belen Rodríguez-Fonseca, Cyril Caminade, Malick Wade, Wasila M Thiaw, Abdoulaye Deme, Andrew P Morse, Jaques-André Ndione, Amadou T Gaye, et al. Oceanic influence on seasonal malaria incidence in west africa. *Weather, Climate, and Society*, 14(1): 287–302, 2022.
- James B Elsner and CP Schmertmann. Assessing forecast skill through cross validation. *Weather and Forecasting*, 9(4):619–624, 1994.
- Åke Johansson, Anthony Barnston, Suranjana Saha, and Huug Van den Dool. On the level and origin of seasonal forecast skill in northern europe. *Journal of the Atmospheric Sciences*, 55(1):103–127, 1998.
- Armin Köhl and Andrey Vlasenko. Seasonal prediction of northern european winter air temperatures from sst anomalies based on sensitivity estimates. *Geophysical Research Letters*, 46(11):6109–6117, 2019.

- Jorge Lopez-Parages, Iñigo Gómara, Belén Rodríguez-Fonseca, and Jesús García-Lafuente. Potential sst drivers for chlorophyll-a variability in the alboran sea: A source for seasonal predictability? *Frontiers in Marine Science*, 9:931832, 2022.
- Joel Michaelsen. Cross-validation in statistical climate forecast models. *Journal of Applied Meteorology and Climatology*, 26(11):1589–1600, 1987.
- Belén Rodríguez-Fonseca, Roberto Suárez-Moreno, Blanca Ayarzagüena, Jorge López-Parages, Iñigo Gómara, Julián Villamayor, Elsa Mohino, Teresa Losada, and Antonio Castaño-Tierno. A review of enso influence on the north atlantic. a non-stationary signal. *Atmosphere*, 7(7):87, 2016.
- Eroteida Sánchez García, José Voces Aboy, and Ernesto Rodríguez Camino. Calibration and combination of seasonal forecast over southern europe. 2014.
- P Sinha, UC Mohanty, SC Kar, SK Dash, AW Robertson, and MK Tippett. Seasonal prediction of the indian summer monsoon rainfall using canonical correlation analysis of the ncmrwf global model products. *International journal of climatology*, 33(7), 2013.
- Koen Verbist, Andrew W Robertson, Wim M Cornelis, and Donald Gabriels. Seasonal predictability of daily rainfall characteristics in central northern chile for dry-land management. *Journal of Applied Meteorology and Climatology*, 49(9):1938–1955, 2010.
- Daniel S Wilks. Probabilistic canonical correlation analysis forecasts, with application to tropical pacific sea-surface temperatures. *International journal of climatology*, 34(5), 2014.
- DS Wilks. Northern hemisphere snow cover, indo-pacific ssts, and recent trend as statistical predictors of seasonal north american temperature. *Journal of Applied Meteorology and Climatology*, 54(1):58–68, 2015.
- Eleni Xoplaki et al. *Climate variability over the Mediterranean*. PhD thesis, Verlag nicht ermittelbar, 2002.