Please check our wiki for help on navigating the form.

Horizon 2020

Call: H2020-MSCA-IF-2019

(Marie Skłodowska-Curie Individual Fellowships)

Topic: MSCA-IF-2019

Type of action: MSCA-IF-EF-ST

(Standard European Fellowships)

Proposal number: 897575

Proposal acronym: ARTIST

Deadline Id: H2020-MSCA-IF-2019

Table of contents

How to fill in the forms

The administrative forms must be filled in for each proposal using the templates available in the submission system. Some data fields in the administrative forms are pre-filled based on the steps in the submission wizard.

1 - General information

Topic	MSCA-IF-2019	Type of Action	MSCA-IF-EF-ST
Call Identifier	H2020-MSCA-IF-2019	Deadline Id	H2020-MSCA-IF-2019

Acronym: ARTIST

Proposal title: ARTificial Intelligence for Seasonal forecast of Temperatures

Note that for technical reasons, the following characters are not accepted in the Proposal Title and will be removed: < > " &

Duration in months: 24

Scientific Area: ENV - Environmental and Geosciences (ENV)

Please select up to 5 descriptors (and at least 3) that best characterise the subject of your proposal, in descending order of relevance.

Descriptor 1: Meteorology, atmospheric physics and dynamics

Descriptor 2: Natural hazards

Descriptor 3: Machine learning, data mining, statistical data processing and applications

Descriptor 4: Artificial intelligence, intelligent systems, multi agent systems

Free keywords: *seasonal prediction, ensemble machine learning, deep learning, land-atmosphere interaction, land surface drivers, statistical model, dynamical model, hybrid forecast system*

Please choose the scientific area and descriptors carefully, and in order of importance, since this will guide the REA in the selection of experts for proposal evaluation and the allocation of proposals to experts. To help you select the most relevant area for your proposal, please consult the Guide for Applicants which provides a breakdown of each scientific area into a number of descriptors.

*Abstract**

Seasonal forecasts are critical tools for early-warning decision support systems that can help reduce the related risk associated to hot or cold weather, droughts, wind discontinuances, and other events that can strongly affect a multitude of socio-economic sectors.
The ARTIST project aims at improving our knowledge of climate variability and predictability, focusing on the role of unexplored drivers, to finally enhance the performance of current seasonal prediction systems. This effort is meant to reduce uncertainties and make forecasts efficiently usable by regional met-services and the socio-economic world. An innovative statistical/dynamical hybrid model will be designed through the synthesis of (a) a state-of-the-art dynamical seasonal prediction system and (b) an ensemble of new-generation statistical methods, often referred to collectively as “machine learning”. Such a hybrid approach may become critical to improve climate forecasts, because it combines the theoretical foundation and interpretability of physical modeling with the power of artificial intelligence that can reveal unknown or overcomplex spatio-temporal features.
Within the bounds of the MSCA fellowship, ARTIST will focus on seasonal prediction of both temperature anomalies and extremes in Europe, but its scalable nature can make it applicable across a wide range of variables and geographical areas. Besides the employment of artificial intelligence, a strength of the action stands in the use of local land surface predictors to instruct the model. Together with standard large-scale drivers such as North Atlantic surface temperatures, Arctic sea ice and others, the empirical model will be trained by unconventional drivers such as soil moisture, snow cover, vegetation indexes and land use, that are expected to significantly amplify prediction skill due to their impact on land-atmosphere feedback mechanisms.

Remaining characters 90

Has a similar proposal in terms of research objectives been submitted to a Horizon 2020 Marie Sklodowska-Curie Individual Fellowship call? ○ Yes ◉ No

Declarations

1) The applicant (future beneficiary) declares to have the explicit consent of all partner organisations (if applicable) on their participation and on the content of this proposal.	☒
2) The information contained in this proposal is correct and complete.	☒
3) This proposal complies with ethical principles (including the highest standards of research integrity — as set out, for instance, in the European Code of Conduct for Research Integrity — and including, in particular, avoiding fabrication, falsification, plagiarism or other research misconduct).	☒

4) The applicant (future beneficiary) hereby declares:

- it is fully eligible in accordance with the criteria set out in the specific call for proposals; and	☒
- it has the financial and operational capacity to carry out the proposed action.	☒

The applicant (future beneficiary) is only responsible for the correctness of the information relating to his/her own organisation. Where the proposal to be retained for EU funding, the applicant (future beneficiary) will be required to present a formal declaration in this respect.

Note:

For **multi-beneficiary applications**, the coordinator vouches for its own organization and that all other participants confirmed their participation and compliance with conditions set out in the call. If the proposal is retained for funding, each participant will be required to submit a formal declaration of honour confirming this.

False statements or incorrect information may lead to administrative sanctions under the Financial Regulation 2018/1046.

Personal data will be collected, used and processed in accordance with Regulation 2018/1725 and the **Funding & Tenders Portal privacy statement**.

Please be however aware that, to protect EU financial interests, your data may be transferred to other EU institutions and bodies and be registered in the EDES database. Data in the EDES database is also subject to Regulation 2018/1725 and the EDES privacy statement.

2 - Participants & contacts

#	Participant Legal Name	Country	Action
1	BARCELONA SUPERCOMPUTING CENTER - CENTRO NACIONAL DE SUPERCOMPUTACION	Spain	

2 - Administrative data of participating organisations

Future Host Institution

PIC	***Legal name***
999655520	*BARCELONA SUPERCOMPUTING CENTER - CENTRO NACIONAL DE SUPERCOMPUTACION*

Short name: BSC

Address

Street Calle Jordi Girona 31

Town BARCELONA

Postcode 08034

Country Spain

Webpage www.bsc.es

Specific Legal Statuses

Legal person ..... yes

Non-profit ..... yes

International organisation ..... no

International organisation of European interest ..... no

Secondary or Higher education establishment ..... no

Research organisation .....yes

Small and Medium-sized Enterprises (SMEs) .....no

Public body .....yes

Academic Sector ..... yes

Department(s) carrying out the proposed work

Department 1

Department name: Earth Sciences ☐ not applicable

☐ Same as proposing organisation's address

Street: Calle Jordi Girona 29

Town: Barcelona

Postcode: 08034

Country: Spain

If the location of the Department carrying out the proposed work is not the same as the location of the Host Institute, please note that although the proposal submission system calculates the budget of the project based on the location of the Host Institute, the budget of the project for the grant agreement will be calculated by using the country coefficient of the location of the Department carrying out the proposed work.

Researcher

The name and e-mail of the Researcher and Supervisor are read-only in the administrative form, only additional details can be edited here. To give access rights and contact details of contact persons, please go back to Step 4 of the submission wizard and save the changes.

Last Name*	Materia	Last Name at Birth	
First Name(s)*	Stefano	Gender*	(•) Male ( ) Female
Title	Dr.	Country of residence*	Italy
Nationality*	Italy	Nationality 2	
Date of Birth (DD/MM/YYYY)	24/08/1979	Country of Birth*	Italy
		Place of Birth	Bologna

Contact address

Current organisation name	Centro Euro-Mediterraneo sui Cambiamenti Climatici		
Current Department/Faculty/Institute/ Laboratory name	Climate Simulations and Predictions		
☐ Same as organisation address			
Street	Via C. Berti Pichat, 6/2		
Postcode/Cedex	40127	Town	Bologna
Phone	+39 0510301603	Country	Italy
Phone2 / Mobile	+xxx xxxxxxxxx		
E-Mail*	stefano.materia@cmcc.it		
ORCID	0000-0001-5635-2847		
Researcher ID			
Other ID	*Please enter the type of ID here*	*Please enter the identifier number here*	

The maximum length of the identifier is 11 characters (ZZZ-9999-2010) and the minimum length is 9 characters (A-1001-2010).

Qualifications

Doctorate Date of (expected) award	Select the exact date (DD/MM/YYYY)	09/06/2009
Doctorate start date	Select the exact date (DD/MM/YYYY)	
University Degree giving access to PHD*	Date of award (DD/MM/YYYY)	21/03/2005

Place of activity/place of residence (previous 5 years - most recent one first)

Indicate the period(s) and the country/countries in which you have legally resided and/or had your main activity (work, studies, etc) during the last 5 years up until the deadline for the submission of the proposal.

Please fill in this section without gaps. Short stays (as defined in the Guide for Applicants) shall not be listed in this box.

Period from	Period to	Duration (days)	Country
11/09/2014	11/09/2019	1827	Italy
	Total	1827	

Proposal ID **897575** *Acronym* **ARTIST** *Short name* **BSC**

Supervisor

The name and e-mail of the Researcher and Supervisor are read-only in the administrative form, only additional details can be edited here. To give access rights and contact details of contact persons, please go back to Step 4 of the submission wizard and save the changes.

Title: Dr. Sex: (•) Male ( ) Female

First name*: **Markus** Last name*: **Donat**

E-Mail*: **markus.donat@bsc.es**

Position in org.: Climate Prediction Group Co-Leader

Department: Climate Prediction Group at Earth Sciences Department

[ ] Same as organisation address

Street: Calle Jordi Girona 29

Town: Barcelona Post code: 08034

Country: Spain

Website:

Phone: +34 934054290 Phone 2: *+xxx xxxxxxxxx* Fax: *+xxx xxxxxxxxx*

Other contact persons

First Name	***Last Name***	***E-mail***	***Phone***
Dorota	Jouet	dorota.jouet@bsc.es	*+34 934134082*

3 - Budget

Is the Researcher eligible for family allowance?* (•) Yes ( ) No

Participant Number	Organisation Short Name	Country	Country Coefficient	Number of Months	Researcher Unit Cost			Institutional Unit Cost		Total
					Living Allowance	Mobility Allowance	Family Allowance	Research, training and networking costs	Management and Overheads	
1	BSC	ES	0,954	24	111732,48	14400,00	12000,00	19200,00	15600,00	172932,48
Total					111732,48	14400,00	12000,00	19200,00	15600,00	172932,48

4 - Ethics

1. HUMAN EMBRYOS/FOETUSES		Page
Does your research involve Human Embryonic Stem Cells (hESCs)?	○ Yes ◉ No	
Does your research involve the use of human embryos?	○ Yes ◉ No	
Does your research involve the use of human foetal tissues / cells?	○ Yes ◉ No	
2. HUMANS		**Page**
Does your research involve human participants?	○ Yes ◉ No	
Does your research involve physical interventions on the study participants?	○ Yes ◉ No	
3. HUMAN CELLS / TISSUES		**Page**
Does your research involve human cells or tissues (other than from Human Embryos/ Foetuses, i.e. section 1)?	○ Yes ◉ No	
4. PERSONAL DATA		**Page**
Does your research involve personal data collection and/or processing?	○ Yes ◉ No	
Does your research involve further processing of previously collected personal data (secondary use)?	○ Yes ◉ No	
5. ANIMALS		**Page**
Does your research involve animals?	○ Yes ◉ No	
6. THIRD COUNTRIES		**Page**
In case non-EU countries are involved, do the research related activities undertaken in these countries raise potential ethics issues?	○ Yes ◉ No	
Do you plan to use local resources (e.g. animal and/or human tissue samples, genetic material, live animals, human remains, materials of historical value, endangered fauna or flora samples, etc.)?	○ Yes ◉ No	
Do you plan to import any material - including personal data - from non-EU countries into the EU?	○ Yes ◉ No	
Do you plan to export any material - including personal data - from the EU to non-EU countries?	○ Yes ◉ No	
In case your research involves low and/or lower middle income countries, are any benefits-sharing actions planned?	○ Yes ◉ No	
Could the situation in the country put the individuals taking part in the research at risk?	○ Yes ◉ No	

7. ENVIRONMENT & HEALTH and SAFETY		Page
Does your research involve the use of elements that may cause harm to the environment, to animals or plants?	○ Yes ◉ No	
Does your research deal with endangered fauna and/or flora and/or protected areas?	○ Yes ◉ No	
Does your research involve the use of elements that may cause harm to humans, including research staff?	○ Yes ◉ No	
8. DUAL USE		**Page**
Does your research involve dual-use items in the sense of Regulation 428/2009, or other items for which an authorisation is required?	○ Yes ◉ No	
9. EXCLUSIVE FOCUS ON CIVIL APPLICATIONS		**Page**
Could your research raise concerns regarding the exclusive focus on civil applications?	○ Yes ◉ No	
10. MISUSE		**Page**
Does your research have the potential for misuse of research results?	○ Yes ◉ No	
11. OTHER ETHICS ISSUES		**Page**
Are there any other ethics issues that should be taken into consideration? Please specify	○ Yes ◉ No	

I confirm that I have taken into account all ethics issues described above and that, if any ethics issues apply, I will complete the ethics self-assessment and attach the required documents. ☒

How to Complete your Ethics Self-Assessment

5 - Call specific questions

Eligibility Researcher (future fellow)

1. Were you in the last 5 years in military service? ○ Yes ◉ No

2. Did you spend time on procedures for obtaining refugee status (according to the 1951 Geneva Refugee Convention and the 1967 Protocol) in a Member State or Associated Country? ○ Yes ◉ No

Other Questions

1. For communication purposes only, the European Commission REA asks for permission to publish the name of the researcher (future fellow) should the proposal be retained for funding. Does the researcher (future fellow) give this permission? ◉ Yes ○ No

2. Some national and regional public research funding authorities run schemes to fund MSCA applicants that score highly in the MSCA evaluation but which cannot be funded by the MSCA due to their limited budget. In case this proposal could not be selected for funding by the MSCA, do the researcher and supervisor consent to the European Commission disclosing to such authorities the results of its evaluation (score and ranking range) together with their names and contact details, non-confidential proposal title and abstract, proposal acronym, and host organisation? ◉ Yes ○ No

3. Is there a secondment in Member States or Associated Countries envisaged in Part B of this proposal? ◉ Yes ○ No

In which sector is the secondment in Member States / Associated Countries foreseen?

☒ Academic ☐ Non Academic

Do you already know the organisation to which this secondment will be? ◉ Yes ○ No

Name: Max Planck Institute for Biogeochemistry

Country: Germany

Extended Open Research Data Pilot in Horizon 2020

If selected, applicants will by default participate in the Pilot on Open Research Data in Horizon 2020[1] , which aims to improve and maximise access to and re-use of research data generated by actions.

However, participation in the Pilot is flexible in the sense that it does not mean that all research data needs to be open. After the action has started, participants will formulate a Data Management Plan (DMP), which should address the relevant aspects of making data FAIR – findable, accessible, interoperable and re-usable, including what data the project will generate, whether and how it will be made accessible for verification and re-use, and how it will be curated and preserved. Through this DMP projects can define certain datasets to remain closed according to the principle "as open as possible, as closed as necessary". A Data Management Plan does not have to be submitted at the proposal stage.

Furthermore, applicants also have the possibility to opt out of this Pilot completely at any stage (before or after the grant signature). In this case, applicants must indicate a reason for this choice (see options below).

Please note that participation in this Pilot does not constitute part of the evaluation process. Proposals will not be penalised for opting out.

We wish to opt out of the Pilot on Open Research Data in Horizon 2020.	○ Yes	◉ No

Further guidance on open access and research data management is available on the Funding & Tenders portal: http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm and in general annex L of the Work Programme.

[1] *According to article 43.2 of Regulation (EU) No 1290/2013 of the European Parliament and of the Council, of 11 December 2013, laying down the rules for participation and dissemination in "Horizon 2020 - the Framework Programme for Research and Innovation (2014-2020)" and repealing Regulation (EC) No 1906/2006.*

1. Excellence

1.1 Quality and credibility of the research/innovation project; level of novelty, appropriate consideration of inter/multidisciplinary and gender aspects

1.1.1 Introduction

Seasonal forecasts are critical tools for early-warning decision support systems that can help reduce the related risk associated to hot or cold weather[1], droughts[2], wind discontinuances[3], and other events that can strongly affect a multitude of socio-economic sectors.

Empirical/statistical approaches and numerical/dynamical methods are the two main categories employed to formulate seasonal forecasts. Advances in statistical prediction are often associated with the enhancement of understanding that usually leads to improve dynamical prediction, and vice versa[4]. Thereby, the two approaches are frequently combined to increase the robustness of forecast for temperature[5], precipitation[6], tropical cyclones[7], and other anomalous events. However, seasonal forecast in mid-latitudes is still affected by large uncertainties, making its value and usability controversial[8].

The **ARTIST** project aims at improving our knowledge of climate variability and predictability, focusing on the role of unexplored drivers, to finally enhance the performance of current seasonal prediction systems. This effort is meant to reduce uncertainties and make forecasts efficiently usable by regional met-services and the socio-economic sectors. An innovative statistical/dynamical hybrid model will be designed through the synthesis of (a) a state-of-the-art dynamical seasonal prediction system and (b) an ensemble of new-generation statistical methods, often referred to collectively as "machine learning". Such a hybrid approach may become critical to improve climate forecasts, because it combines the theoretical foundation and interpretability of physical modeling with the power of artificial intelligence that can reveal unknown or overcomplex spatio-temporal features.

Within the bounds of the MSCA fellowship, **ARTIST** will focus on seasonal prediction of both temperature anomalies and extremes in Europe, but its scalable nature can make it applicable across a wide range of variables and geographical areas. Besides the employment of artificial intelligence, a major innovation of the action stands in the use of local land surface predictors to instruct the model. Together with standard large-scale drivers such as North Atlantic surface temperatures, Arctic sea ice and others, the empirical model will be trained by unconventional drivers such as soil moisture, snow cover, vegetation indexes and land use, that are expected to significantly amplify prediction skill due to their impact on land-atmosphere feedback mechanisms.

1.1.2 State-of-the-art

Seasonal forecasting began in the 1950s using statistical or empirical techniques, when large scale atmospheric patterns were first identified, together with the relationships between atmospheric variability and ocean temperature anomalies[9]. Back in the 1980s, seasonal prediction was based on lag correlations with observed upper atmosphere geopotential height anomalies[10] or analogs[11]. Analog techniques typically choose years with similar boundary forcings and use composites of those chosen years as the forecast. More recently, statistical or empirical forecasts have often identified relationships between weather patterns and ocean anomalies, relying on linear regressions or variations such as canonical correlation analysis (CCA).

---

[1] Haustein, K., Otto, F., Uhe, P., Schaller, N., Allen, M., Hermanson, L., Christidis, N., McLean, P., and Cullen, H. (2016) Real-time extreme weather event attribution with forecast seasonal SSTs. *Environmental Research Letters*, 11(6):064006.

[2] Dutra, E., Di Giuseppe, F., Wetterhall, F., and Pappenberger, F. (2013) Seasonal forecasts of droughts in African basins using the standardized precipitation index. *Hydrology and Earth System Sciences*, 17(6):2359, 2013.

[3] Lledó, L., Bellprat, O., Doblas-Reyes, F. J., & Soret, A. (2018). Investigating the effects of Pacific sea surface temperatures on the wind drought of 2015 over the United States. *Journal of Geophysical Research: Atmospheres*, *123*(10), 4837-4849.

[4] Doblas-Reyes, F. J., García-Serrano, J., Lienert, F., Biescas, A. P., & Rodrigues, L. R. (2013). Seasonal climate predictability and forecasting: status and prospects. *Wiley Interdisciplinary Reviews: Climate Change*, *4*(4), 245-268.

[5] Manzanas, R., Gutiérrez, J., Fernández, J., van Meijgaard, E., Calmanti, S., Magariño, M., Cofiño, A., and Herrera, S. (2017) Dynamical and statistical downscaling of seasonal temperature forecasts in europe: Added value for user applications. *Climate Services*.

[6] Tippett, M. K., Goddard, L., and Barnston, A. G. (2005). Statistical–dynamical seasonal forecasts of central- southwest Asian winter precipitation. *Journal of climate*, 18(11):1831–1843.

[7] Murakami, H., Villarini, G., Vecchi, G. A., Zhang, W., and Gudgel, R. (2016). Statistical–dynamical seasonal forecast of North Atlantic and us landfalling tropical cyclones using the high-resolution GFDL FLOR coupled model. *Monthly Weather Review*, 144(6):2101–2123.

[8] Soares, M. B., & Dessai, S. (2016). Barriers and enablers to the use of seasonal climate forecasts amongst organisations in Europe. *Climatic Change*, *137*(1-2), 89-103.

[9] Next generation earth system prediction: Strategies for subseasonal to seasonal forecasts. Washington, DC: *National Academies Press*.

[10] Wagner, A. J. (1989). Medium- and long-range forecasting. *Weather and Forecasting*, 4, 413–426.

[11] Livezey, R. E., & Barnston, A. G. (1988). An operational multi-field analog anti-analog prediction system for United States seasonal temperatures. System design and winter experiments. *Journal of Geophysical Research*, 93A, 10953–10974.

In the last few years, climate science has begun to push artificial intelligence to the limit, in order to fully exploit spatial and temporal structures in the data, and in this context Reichstein et al.[12] described perspectives and opportunities of deep learning for earth system science problems. Deep learning[13] is a subset of machine learning that allows for self-training to perform tasks, by exposing multi-layered neural networks to vast amounts of data, ultimately extracting hidden spatio-temporal features. As an example, deep learning is able to address video and motion prediction (widely used for programming self-driving cars), a problem that shares remarkable similarities to climate forecasting. Both problems aim at predicting a future state starting from a measurable initial condition, exploiting temporal-dynamical properties as well as spatial structures. Predicting the position of two walking people in the next frame of a video can be done by an algorithm informed on (a) the velocity at which the two are currently walking and (b) the spatial context across which they are moving (e.g. roads, pavements, flowerbeds, buildings, etc.). Likewise, a temperature seasonal prediction cannot disregard the so-called memory effect of the ocean temperature anomalies, that will long persist influencing atmospheric patterns for the forthcoming months (temporal property). At the same time, it needs to acknowledge e.g. the spatial distribution of the local snow cover, because soils covered in snow will be slowly provided with water able to infiltrate in depth, and this wetness will affect temperatures in the months to come (spatial feature).

Dynamical predictions are based on the assumption that large-scale and long-lasting anomalies will convey predictive skill to seasonal forecast. In particular, the ocean provides memory to the climate system[14], enabling fully coupled models to produce skilful predictions up to one year ahead or more. Since the cardinal work of Fennessy and Shukla[15], the role played by land surface processes has received growing attention in the seasonal prediction community: initial state of snow cover, soil wetness, and other land surface variables can have a major influence on seasonal mean circulation, and may carry predictability to the system up to six months. In addition, it has been recently demonstrated that initialization of sea ice may convey predictability through persistence or advection of sea-ice anomalies and their interactions with the ocean and atmosphere[16]. Such a multi-way interactive system cannot neglect the use of fully coupled atmospheric-ocean models, complemented by interactive land surface, sea ice and stratosphere components. These prediction systems have shown enhanced skill with respect to those relying on atmospheric models with prescribed SST boundary conditions[17], and also to statistical forecasts, at least in the Tropics[18].

Most of the inherent seasonal skill is derived from the accurate prediction of coupled atmospheric–oceanic phenomena, predominantly related to El Niño Southern Oscillation (ENSO)[19,20]. ENSO is the dominant mode of global climate variability and the main process contributing to predictability on seasonal times scales. The large spatial shifts in tropical Pacific rainfall associated with ENSO lead to large-scale changes in global circulation (e.g.[21,22]). Therefore, realistic representation of the associated physical processes and teleconnections by the dynamical model is key factor for getting the correct signal in the forecasts[23].

However, in the extra-Tropics only a fraction of climate variability may be ascribed to ENSO, and dynamical models produce mixed results in the prediction of mid- and high-latitude regions, especially far from the tropical Pacific Ocean. For example, Weisheimer and Palmer[24] recently marked Southern Europe as a region where seasonal predictions of temperatures are "marginally useful" for decision making, based on the European Centre (ECMWF) System 4. In addition, many physical processes in current land surface models are still coarsely parameterized, and the initialization of land surface is affected by the lack of available records of soil moisture. In this way, despite the

[12] Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., & Carvalhais, N. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195.
[13] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436.
[14] Luo, J. J., Masson, S., Behera, S., Shingu, S., & Yamagata, T. (2005). Seasonal climate predictability in a coupled OAGCM using a different approach for ensemble forecasts. *Journal of climate*, *18*(21), 4474-4497.
[15] Fennessy, M. J., & Shukla, J. (1999). Impact of initial soil wetness on seasonal atmospheric prediction. *Journal of Climate*, *12*(11), 3167-3180.
[16] Guemas, V., Blanchard-Wrigglesworth, E., Chevallier, M., Day, J. J., Déqué, M., Doblas-Reyes, F. J., ... & Koenigk, T. (2016). A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales. *Quarterly Journal of the Royal Meteorological Society*, 142(695), 546-561.
[17] Kug JS, Kang IS, Choi DH. Seasonal climate predictability with tier-one and tier-two prediction systems. *Climate Dynamics* 2008, 31:403 – 416.
[18] Rodrigues, L. R. L., Doblas-Reyes, F. J., & dos Santos Coelho, C. A. (2014). Multi-model calibration and combination of tropical seasonal sea surface temperature forecasts. *Climate Dynamics*, *42*(3-4), 597-616.
[19] Cane M, Zebiak SE, Dolan SC. (1986) Experimental forecasts of El Niño. *Nature*, 321:827–832.
[20] Barnston, A. G., Van den Dool, H. M., Zebiak, S. E., Barnett, T. P., Ji, M., Rodenhuis, D. R., ... & Kousky, V. E. (1994). Long-lead seasonal forecasts—where do we stand?. *Bulletin of the American Meteorological Society*, *75*(11), 2097-2114.
[21] Philander, S. G. H. (1983). El Nino southern oscillation phenomena. *Nature*, *302*(5906), 295.
[22] Alexander, M. A., Bladé, I., Newman, M., Lanzante, J. R., Lau, N. C., & Scott, J. D. (2002). The atmospheric bridge: The influence of ENSO teleconnections on air–sea interaction over the global oceans. *Journal of Climate*, *15*(16), 2205-2231.
[23] Scaife, A. A., Karpechko, A. Y., Baldwin, M. P., Brookshaw, A., Butler, A. H., Eade, R., ... & Smith, D. (2016). Seasonal winter forecasts and the stratosphere. *Atmospheric Science Letters*, *17*(1), 51-56.
[24] Weisheimer, A., & Palmer, T. N. (2014). On the reliability of seasonal climate forecasts. *Journal of the Royal Society Interface*, *11*(96), 20131162.

proved importance of land initial state for mid-latitude seasonal prediction quality[25], initialization of the surface component has sometimes shown marginal benefits[26].

Systematic comparison of the statistical and dynamical approaches shows advantages and disadvantages for both, and it has been demonstrated that stake-holders may extract the best information out of the forecasts by optimally combining statistical and dynamical predictions[27], in order to provide the most trustworthy advice for decision support in a large spread of sectors[28]. Until recently, hybridization efforts have either focused on statistical post-processing and calibration of dynamical models[29], or fusion of the two approaches through, e.g., Bayesian methods that give greater weights to better performing models[30]. To the candidate's best knowledge, no attempts to hybridize traditional forecast systems with statistical methods based on data mining tools have been yet undertaken.

1.1.3 Objectives and expected impacts

The main goals of **ARTIST** are:

(a) strengthening our understanding of climate variability and predictability at the seasonal time-scale, with particular regard to the role played by land surface-atmosphere interactions;

(b) developing a scalable and portable artificial intelligence approach for climate forecast problems, relevant for prediction of field anomalies and extremes, that will eventually improve climate models and seasonal prediction skill.

The project foresees major scientific impacts: the extraction of data spatio-temporal features will improve our understanding of hidden physical processes, that can be incorporated into global climate models and used to widen insights about land-atmosphere interactions.

Important socio-economic impacts are as well expected: the hybrid seasonal forecast system is trusted to enhance prediction skill, eventually making seasonal forecast a widely used early-warning tool for decision-making processes, with the ultimate aim to reduce the anthropic impact on the environment and improve human's life and economy.

1.1.4 Research methodology and approach

ARTIST proposes an integrated process-based and data-driven approach intended to improve performance of current seasonal forecast systems for the prediction of near-surface temperature anomalies and extremes across Europe. This target will be achieved through smart detection of antecedent large-scale and local land surface drivers for the prediction of seasonal temperature anomalies and extremes. A number of statistical models will be built to link drivers with temperature forecasts, by training models with different machine and deep learning approaches that can exploit spatio-temporal structures in the data. This ensemble of empirical models will be combined with the European Community Earth system prediction model EC-Earth[31], for the generation of a hybrid statistical/dynamical forecast system. A long time-series is required for the robustness of the identified statistical relationships, and the hybridization requires that the dynamical model runs on the same time range. **ARTIST** will benefit from the long-lasting EC-Earth hindcast set available at BSC, therefore the project will rely on an already performed re-forecast, avoiding costly runs. The **ARTIST** project aims at designing a scalable methodology that could, eventually, be applied to other models, predictands (e.g. precipitation or winds) and different regions across the world.

The project is organized into three main work packages, complemented by a scientific and coordination management work package (WP0):

WP0. Scientific and coordination management

This work package will be devoted to project coordination and management: progress monitoring, periodic review, identification of risk of underachievement and contingency planning, internal communication among the host and the secondment will be the main activities internal to the WP. WP0 will ensure sound legal, contractual, financial and administrative management of the project, in compliance with the funded proposal and the MSCA guidelines.

WP1. Identification of external drivers

---

[25] Materia, S., Borrelli, A., Bellucci, A., Alessandri, A., Di Pietro, P., Athanasiadis, P., ... & Gualdi, S. (2014). Impact of atmosphere and land surface initial conditions on seasonal forecasts of global surface temperature. *Journal of Climate*, *27*(24), 9253-9271.

[26] Prodhomme, C., Doblas-Reyes, F., Bellprat, O., & Dutra, E. (2016). Impact of land-surface initialization on sub-seasonal to seasonal forecasts over Europe. *Climate dynamics*, *47*(3-4), 919-935.

[27] Kim, H. M., & Webster, P. J. (2010). Extended-range seasonal hurricane forecasts for the North Atlantic with a hybrid dynamical-statistical model. *Geophysical Research Letters*, *37*(21).

[28] Lowe, R., Barcellos, C., Coelho, C. A., Bailey, T. C., Coelho, G. E., Graham, R., ... & Rodó, X. (2014). Dengue outlook for the World Cup in Brazil: an early warning model framework driven by real-time seasonal climate forecasts. *The Lancet infectious diseases*, *14*(7), 619-626.

[29] Wilks, D. S. (2009). Extending logistic regression to provide full-probability-distribution MOS forecasts. *Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling*, *16*(3), 361-368.

[30] Coelho, C. A. S., Pezzulli, S., Balmaseda, M., Doblas-Reyes, F. J., & Stephenson, D. B. (2004). Forecast calibration and combination: A simple Bayesian approach for ENSO. *Journal of Climate*, *17*(7), 1504-1516.

[31] Hazeleger, W., Severijns, C., Semmler, T., Ştefănescu, S., Yang, S., Wang, X., ... & Bougeault, P. (2010). EC-Earth: a seamless earth-system prediction approach in action. *Bulletin of the American Meteorological Society*, *91*(10), 1357-1364.

An initial selection will be made to recognize precursors of near surface temperatures across Europe (e.g. North Atlantic surface temperatures, Arctic sea ice cover, and others), using process-knowledge and literature, but also tailored machine learning tools, such as efficient algorithms[32], to find the most suitable set of predictors. The drivers identified in this preliminary phase include large-scale predictors as well as local land surface predictors (e.g. land use, local snow cover, vegetation indices, albedo and soil moisture), and in this way two complementary setups of input drivers are used to train the statistical model.

WP2. Model training and evaluation

An ensemble of different machine learning and deep learning techniques from various families (e.g. tree-based models[33,34], deep[35] and conventional[36] neural networks, Kernel methods[37] and others) will be explored to link drivers to predictands. Drivers are the anomaly of the predictors described in WP1, in the month preceding the season to be predicted (i.e. drivers of December are used as predictors for the JFM season, and so on). Predictands are near surface temperature anomalies and extremes (heat waves and cold spells), with the latter that will be defined in M2.1.

The capability of data-driven methods to extrapolate seasonal temperatures and extremes will be performed through a rigorous cross-validation. WP2 will make use of a variety of metrics to evaluate the machine learning performance, such as anomaly correlation coefficient (ACC), coefficient of determination (R^2), model efficiency (MEF), root mean square error (RMSE) among others.

The set of afore-mentioned data-driven techniques will be used to train the predictands again, by using the same set of predictors with the exclusion of land surface drivers. One by one, the local land drivers are removed and the models re-fit, in order to evaluate the fraction of predictability imparted by the land surface component as a whole and by each of the land variables. To assess the effect of land surface drivers on the predictability of temperature anomalies and extremes, the metrics used to evaluate the different models in WP2 are run again and compared.

WP3. Hybrid dynamical-statistical forecast system

The statistical-based model is combined with EC-Earth, a fully coupled dynamical forecast system, for the construction of the hybrid prediction model. Different hybridization techniques are systematically tested in order to exploit the ensemble nature of both the dynamical and statistical models, to finally find the best suit for the project purposes. Explored methodologies will include smart ensemble subsampling, weighted combinations, and an additional training, that will connect drivers and predictands generated by the dynamical forecast through the machine and deep learning techniques developed in WP2.

A range of probabilistic techniques (e.g. reliability, ranked probability skill score (RPSS), relative operating characteristic (ROC) score, and others) will be used for comparison with EC-Earth forecast.

1.1.5 Advancement in the research field and interdisciplinary aspects of the project

The fellow, Stefano Materia, is proposing the definition of a methodology that will improve our understanding of the European seasonal variability and predictability by quantifying the effects of different predictors, some of which yet unexplored. The synthesis between a statistical approach based on machine and deep learning and a fully-coupled seasonal prediction system will lead to a definite improvement of forecast skill. Hybrid strategies for seasonal forecasts have been widely examined in climate science, however this project allows for at least two major innovations in the way this would be done:

1) The employment of artificial intelligence to build relationships between predictors and predictand are at their infancy in the seasonal forecast science, and for the first time a multitude of machine learning techniques are chosen in order to build an ensemble of empirical prediction models;
2) The inclusion of local land surface variables, beside large-scale predictors, as fosterers of predictability has not previously been tested in empirical forecast models.

Moreover, the effect of land surface will not be limited to the role of soil moisture, but potentially highly impacting fields such as the local snow cover and vegetation state will be taken into account. Also, a quasi-static variable such as land-use will be employed to explore temperature spatial variability in a region where natural vegetation is spaced apart by cultivated areas and large conurbations.

This project will combine the main competences of Stefano in dynamical seasonal forecast and land surface-atmosphere interactions, with the expertise of the Climate Prediction Group (CPG) at the Barcelona Supercomputing

[32] Jung, M., and Zscheischler, J. (2013) A Guided Hybrid Genetic Algorithm for Feature Selection with Expensive Cost Functions. *Procedia Computer Science*, 18, 2337-2346, doi: 10.1016/j.procs.2013.05.405.
[33] Breiman, L.: Random Forests, Mach. Learn., 45, 5–32, doi:10.1023/A:1010933404324, 2001
[34] Jung, M., Reichstein, M., & Bondeau, A. (2009). Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model. *Biogeosciences*, *6*(10), 2001-2013.
[35] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Dieleman, S. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484.
[36] Haykin, S. (1994). *Neural networks: a comprehensive foundation*. Prentice Hall PTR.
[37] Shawe-Taylor, J., & Cristianini, N. (2004). *Kernel methods for pattern analysis*. Cambridge University Press.

Centre (BSC) in climate prediction, extremes and statistical approaches applied to the Earth system. Besides this interdisciplinarity falling within the big umbrella of climate science, the combination of unexplored drivers, increased computational power, and the recent advances in artificial intelligence offers exciting new opportunities for expanding our knowledge about climate from data analysis. In particular, in this project tools from the field of machine learning will be developed and adapted to climate science with the help of both the principal host and the secondment (the Max Planck Institute for Biogeochemistry, MPI-BGC).

Hence, **ARTIST** turns out to be a highly interdisciplinary project, that combines different aspects of climate science (predictability, extremes, land-atmosphere interaction) with artificial intelligence. Stefano is committed to highlight how the automatic extraction of abstract features from data may overcome the skill constraints of dynamical forecast systems.

1.2 Quality and appropriateness of the training and of the two-way transfer of knowledge between the researcher and the host

This project has all the required qualifications to be beneficial for both the career advancement of Stefano Materia and for the Earth Sciences Department (ESD) at the BSC.

The fellow will benefit from the renowned expertise of his supervisor, Dr. Markus Donat, who is an authority in climate extremes, predictability and variability. In his studies, Markus has most often post-processed observational data through a large variety of statistical analysis. These combined competences will be of service to many stages of the **ARTIST** project, specifically in the model training phase, and in the choice of the hybridization methods to combine the statistical approach with the numerical model. Stefano will acquire unquestionable advantages from this supervision, through the reinforcement of his skills in data analytics and new knowledge about statistical approaches. He will as well develop a new perspective about seasonal predictability and forecast, with respect to the more classical dynamical viewpoint where he comes from.

Stefano will work closely with the CPG, with whom he already has collaborated during a profitable three-month visiting program (February - May 2019). The extended interdisciplinarity, the international environment and the genuine collaborative atmosphere that can be breathed in the Nexus building will boost his knowledge through interesting scientific discussions.

The afore-mentioned training-through-research activities will be carried out by means of the cutting-edge laboratory equipment and powerful computing tools at BSC, and will benefit from the direction of Prof. Francisco Doblas-Reyes, director of the ESD, who will co-supervise the project.

BSC is the ideal institution for a mid-career scientist like Stefano Materia, since they offer tailored training activities for the development and strengthening of important skills intersecting the scientific career: Stefano is committed to attend the courses, arranged in the framework of the BSC's Diploma of Excellence in Research Skills, on soft skills (in particular planning and time organization), technical skills (with particular emphasis on project management and networking), outreach skills and leadership. Besides, participation in the PATC courses (https://www.bsc.es/education/training/patc-courses), available free of charge at the BSC, will boost his coding skills and expand his knowledge on the HPC technologies. The classes on disciplines related to big data process and analytics are expected to have a tremendous impact on the effectiveness of the project, and will ensure the development of new ideas. Finally, BSC offers Spanish and Catalan classes to all foreigner employees and visitors, providing an excellent opportunity to connect the new-comers to the local culture and to communicate their research to a non-scientific audience. The most suitable training activities will be described in details, and regularly updated, in the Career Development Plan (see Section 3.2) which will be prepared with the supervisor at the start of the fellowship. This intense coaching activity, together with a strong involvement in the financial management of his project, is expected to guide Stefano through the action and ultimately his career, boosting his abilities in writing and leading projects and possibly heading his own research group.

ARTIST includes a secondment at the Max Planck Institute for Biogeochemistry in Jena (MPI-BGC), where the fellow will integrate his proficiency in climate modelling with the relatively new field of deep learning applied to geophysics: his mentor, Dr. Martin Jung, is in fact internationally esteemed for applying artificial intelligence to earth system science. Acquiring competence in such an interdisciplinary matter will act as a springboard for a profusion of activities that involve the use of seasonal forecasts for application in various economic sectors.

In turn, Stefano will share with both groups in Barcelona and Jena his skills on land-atmosphere interaction and their impact on sub-seasonal to seasonal predictability. Moreover, at CMCC he led numerous projects that implemented seasonal forecasts into decision-making processes of private companies, and this expertise will be beneficial to the Earth System Services group at BSC as well. He foresees a series of seminars on this topic, where he will disclose his experience on how to transform climate data into useful information for the end-users. A seminar in Jena is expected as well during the secondment.

Finally, all codes and methods developed during this project will be organized in a GIT repository, and as such provide a legacy for further improvements to be used at BSC and MPI-BGC, as well as the wider scientific community, beyond the lifetime of the project. To ensure portability of the developed methodology, all codes will make use of Bash and Python scripting, and will be accompanied by detailed documentation that will allow other researchers to take full advantage of the final products. Stefano will bring new knowledge to both institutions through its hosts and visiting experiences, and will share his existing and future contact network as a result of this project.

1.3 Quality of the supervision and of the integration in the team/institution

1.3.1 Expertise of the supervisors and mentor

Dr. Markus Donat is co-leader of the Climate Prediction group at the BSC, and an internationally recognised expert in studying mechanisms and model representation of climate extremes and variability. Markus has published more than 75 peer-reviewed journal articles, nine of these in Nature-family journals, and has contributed to the IPCC 5th Assessment Report. His work has been cited 3,216 times (according to Scopus as of August 2019), and he has an h-index of 29. Markus is Associated Investigator with the Australian Research Council Centre of Excellence for Climate Extremes, and Associated expert with the World Meteorological Organization (WMO) Commission of Climatology (CCl) Expert Team on Data Development and Stewardship. Based on his achievements he has been awarded the World Climate Research Program (WCRP) / Global Climate Observing System (GCOS) International Data Prize 2017. Markus has a vast experience in supervising students (e.g. he has supervised four PhD students) and post-docs including two MSCA-IF fellows, and will supervise Stefano's fellowship at the BSC.

The co-supervisor, ICREA **Prof. Francisco Doblas-Reyes**, is the director of BSC Earth Sciences Department. He is a worldwide expert in the development of seasonal-to-decadal climate prediction systems and has more than 20 years of experience in weather and climate modelling, climate prediction, as well as the development of climate services. Prof. Doblas-Reyes serves on scientific panels of the World Climate Research Programme (WCRP) and the World Weather Research Programme (WWRP) and has authored and co-authored more than 100 peer-reviewed papers on climate modelling and prediction, as well as on climate services. He is a member of the European Network for Earth System modelling HPC Task Force and has participated in numerous national and European FP4 and FP7 projects. Prof. Doblas-Reyes is also involved in four Horizon 2020 Collaborative projects and C3S_512 *Copernicus Climate Change* contract as a Principal Investigator, has supervised or is currently supervising four Marie Skłodowska-Curie Individual Fellowships awarded at BSC.

The mentor **Dr. Martin Jung** leads the group of "Global Diagnostic Modelling" at the Max Planck Institute for Biogeochemistry in Jena. Since his PhD, he has become an expert in machine learning applications for land-atmosphere interaction problems. Dr. Jung has been the Principal Investigator of the FLUXCOM initiative, aiming at generating machine learning based global land-atmosphere energy and carbon fluxes based on in-situ observations and remote sensing. He is a member of the scientific steering committee of the Coupled Climate-Carbon Cycle Model Intercomparison Project (C4MIP) under the global Climate Model Intercomparison Project Phase 6 (CMIP6) of the WCRP and team member of the WCRP SurFlux Project, aiming at reducing uncertainties of land-atmosphere fluxes. Dr. Jung was actively involved in more than ten research projects funded by the European Union or the European Space Agency where he led a work package in the EU FP7 CLIMAFRICA, H2020 BACI, and ESA CCI RECCAP2 projects. Dr. Jung published 87 papers, many of which in high-profile journals (Nature, Science, PNAS) and two as first author in Nature. By August 2019 his publications received 8528 citations and he was identified as a highly cited researcher (top 1% in the field) by Clarivate Analytics in 2018. He has been advising six PhD students and four PostDocs. His academic network reaches across disciplines of machine learning, land-atmosphere interactions, land surface modelling, among others, with numerous collaborations going on.

1.3.2 Integration in the host team and premises

In Barcelona, Stefano Materia will be hosted by the CPG co-led by Dr. Markus Donat, and will be provided with a workspace including workstation. He will have full access to the BSC's computer infrastructure and to the Mare Nostrum supercomputer. Stefano and Markus will meet once a week to routinely discuss the state of advancement of the project, idea sharing and updates on literature. Markus' scientific guidance will be of certain benefit for the efficiency and relevance of the work, and the two will have frequent coffee talks as well. Every second week, the two will gather with Prof. Doblas-Reyes for updates on work advancement.

Stefano will take advantage of the stimulating environment at the CPG, where has already collaborated during a three-month visit during spring 2019, culminated in the draft of an article now in preparation. Two more groups of the ESD will be involved in Stefano's research. The Computational Earth System group will supply assistance on computational infrastructure, data download and postprocessing, provide guidance on the technical issues arising from the work, and supply assistance for an efficient use of the HPC resources. The Earth System Services group

will support Stefano in the communication and dissemination of **ARTIST**, and it will be involved in the discussion on making the project outcomes usable for the society and economic world.

One of the best chances for Stefano to integrate into the research activities of the team is throughout attendance of the regular ESD internal seminars, where scientific knowledge is shared across the different groups, and the bi-monthly CPG lunch meetings, functional to the exchange of technical and scientific issues and methodological approaches within the group. An ESD internal seminar will be held by Stefano as a deliverable of **ARTIST**. Stefano's inclusion in the department life will be eased by a specific policy internal to BSC, that every year approves and co-finances an activity plan proposed by the staff. The aim is to promote relationships between staff from different departments through their participation in leisure activities outside working hours, and therefore foster cohesion among people and promote the center's values.

In Jena, Stefano will work closely to Dr. Martin Jung and the Global Diagnostic Model group at MPI-BGC. He will be supplied with a workstation and necessary facilities, and have access to the large database of global data-driven land-atmosphere energy fluxes developed in the department, including the machine learning codes used for generating it and related technical assistance.

1.4 Potential of the researcher to reach or re-enforce professional maturity/independence during the fellowship

A few of Stefano's competences will contribute to the success of the action, and in turn these talents will be strengthened during the course of the fellowship.

First, Stefano has a strong multi-disciplinary background. His master in Environmental Science allows him a broad view over earth science problems and a capacity of thinking about environmental issues that goes beyond the physical process. The PhD in Geophysics (with a thesis facing the problem of river routing in general circulation models) has built up his physical background and technical skills, and allowed him to familiarize with climate models. His seven-month experience as a post-doc at the Center for Ocean, Land and Atmosphere studies (COLA), working with Prof. Paul Dirmeyer, was crucial for the establishment of Stefano's expertise in land-air interaction and its role on predictability. His work has a rather significant impact: his 11 published papers are cited 189 times (Scopus). Four additional papers are under review and four more are in preparation. **ARTIST** will add artificial intelligence to Stefano's multi-disciplinary expertise portfolio, and the combination of existing skills and new competences will contribute establishing himself as a leading researcher in the field of statistical-dynamical predictions. This fellowship will enlarge Stefano's international profile and collaboration network in the fields of climate prediction and land surface impact on seasonal predictability, and the awaited research papers will increase his reputation.

Second, he is highly experienced in climate forecast science. Since taking the lead of a work package in the FP7 Climafrica project, he has been dealing with seasonal forecast problems from different points of view: the study of processes imparting predictability, verification of forecast quality, translation of usable information for end-users. He is active in the subseasonal-to-seasonal prediction (S2S) community, through an ongoing collaboration with the International Research Institute for Climate and Society (IRI), co-leader of the S2S prediction project (http://s2sprediction.net/); he is work package leader of the Copernicus C3S Operational Production of Seasonal Forecast tender (C3S_330), and in charge of land initialization for the CMCC seasonal prediction system[38]; he is the CMCC delegate for the WMO Mediterranean Climate Outlook Forum (MEDCOF) activities. Through this fellowship, he will expand his expertise to the empirical prediction sphere and become an expert of the synergies between physical and data-driven models, with the ultimate goal of mastering hybrid modelling approaches.

Third, he has developed solid management skills, both by supervising PhD students and Post-Docs and by writing successful proposals that allowed him to manage project financial and personnel resources. Right now, he is in charge of two ongoing contracts to provide seasonal forecast services to private companies in the energy and agribusiness sectors, for an amount of about Euro 200.000. This grant will allow Stefano to fully grow his management competence and become a mature researcher, ready for heading his own group in a research institute providing operational seasonal forecasts or in a company developing climate services.

2. Impact

2.1 Enhancing the future career prospects of the researcher after the fellowship

Stefano Materia's final goal is to contribute making seasonal forecast a stable instrument for decision-making processes, used for a sustainable planning of crop calendars, a more efficient use of energy, an optimized schedule of transports, with the ultimate aim to reduce human's impact on the environment. In his vision, this fellowship

[38] Sanna, A., Borrelli, A., Athanasiadis, P., Materia, S., Storto, A., Navarra, A., ... & Gualdi, S. (2016). CMCC-SPS3: the CMCC Seasonal Prediction System 3. *CMCC Research Paper*, (RP0285).

provides the instruments to achieve this goal, and acquired expertise will open career opportunities both in a research institute and in a private entity.

ARTIST is the chance for Stefano to immensely expand his knowledge and experience about seasonal forecasts: the fellowship envisages the incorporation of an avant-garde approach to climate science, aimed at broadening insights on variability and predictability at the seasonal time-scale. At the same time, the opportunity to work on a variety of machine learning techniques will give Stefano a wide overview of the applications of artificial intelligence in climate prediction problems. Last but not least, the project foresees a visit to IRI which includes a training week on seasonal forecast (Section 2.2), that will give a new perspective on the needs of end-users in developing countries.

In addition, being at BSC is an immeasurable opportunity to be up-to-date on innovative and advanced topics. The training activities on disciplines related to big data process and analytics, which are strictly related to the purposes of **ARTIST**, will provide Stefano with a new viewpoint on ways of generating value from information, a topic that will be central in the future knowledge economy. Moreover, BSC is the Spanish HPC facility, and working closely to a highly qualified team of Parallel Computing specialists as well as attending classes on computational infrastructure and software, will make him more efficient in his work and research after completing the fellowship.

The combination of these acquired competences will widen Stefano's career opportunities, and will make him ready to head an innovative line of research at a major research institute in Europe or a consulting company developing climate services.

2.2 Quality of the proposed measures to exploit and disseminate the project results

With the help of the BSC Project Dissemination Unit and Communication Department, a dissemination and communication plan will be designed at the beginning of the project. The plan will cover activities shared with the scientific community, public and private users and the general public, and allocate time for communication in order to ensure continuity. The Earth System Services group, that over the years has built a strong network of stakeholders using climate services produced within the BSC, will provide support on finding the best targets for the results of **ARTIST**. The opportunity to involve stakeholders to test the methodology and supply feedback will be evaluated.

Part of the dissemination action will be dedicated to scientific production: two research papers will be submitted on high impact, open access peer reviewed journals, such as Journal of Climate, Environmental Research Letters, Nature Geoscience or the recently issued Nature Machine Intelligence, and a third one will be drafted for submission immediately after the end of the fellowship. These works will discuss (1) an ensemble of machine learning methods for seasonal forecast in Europe (2) the role imparted by land-surface drivers in the seasonal predictability of temperature anomalies and extreme in Europe, (3) a hybrid statistical-dynamical model for seasonal forecast of temperatures in Europe. Given the original and interdisciplinary nature of the proposed action, additional research articles will likely come out during the fellowship.

During the fellowship, Stefano will participate in international conferences such as the AGU and EGU, which have recently started sessions on machine learning in earth system science: other than disseminating results and advancement of his project, he will increase his collaboration network and share ideas for further opportunities. He will reach out to the MEDCOF community, which is keen to find solutions for the improvement of seasonal forecast quality. He will attend to targeted international workshops, such as those organized by the ECMWF on predictability and sub-seasonal to seasonal forecasting.

Finally, Stefano foresees a two-week visit at IRI: one week will be spent at their headquarters in Palisades (NY), hosted by Dr. Ángel Muñoz, and one week in a Latin American country, taking active participation in one of the "Adapting Agriculture to Climate Today, for Tomorrow" (ACToday) training weeks. In fact, Ángel leads the Latin American component of the ACToday Columbia World Project, which aims to enhance the availability and value of climate information in national policy, planning, management and other decision-making processes. The public attending this training activity is the perfect target for dissemination of the results of **ARTIST**, because the envisioned improvement of prediction skill will open wide opportunities in areas where dynamical forecasts still hold scarce quality, by anticipating and managing climate impacts to improve human welfare and the environment.

2.3 Quality of the proposed measures to communicate the project activities to different target audiences

The Communication Department will support Stefano with the communication part of the plan, targeted to the general public to enlarge the audience of data science and seasonal forecasts, and disclose their opportunities and challenges. Hence, the unit will be involved in the drafting of a dedicated posting plan, in order to communicate project activities, participation in workshops, article main outcomes, news about related research and other relevant actions to Stefano's personal and BSC Twitter accounts.

The department will continuously provide support on how to communicate complex research in a plain-sailing way for non-specialists, for example by making use of visual communication, which will favor the participation of Stefano in a number of appointments:

1) The European Research Night, an event that has recently gained vast attention and attracted 1.5 million visitors in 2018. A specific program funded by the MSCA promotes the dissemination of fellows, in order to display the value of public funding to the general public and aims at inspiring students from primary and secondary schools and a new generation of researchers;
2) Barcelona Science Night (https://www.barcelona.cat/barcelonaciencia/en/13th-science-festival), a festival, coordinated by the Science Barcelona programme, which involves research centers, universities, companies, associations and teaching professionals. Many activities are organized to bring people closer to science and technology through a grand two-day event;
3) Data Beers (http://www.databeersbcn.com/?i=2) regular gatherings, where professionals and students interested in data science, a topic very relevant to **ARTIST**, can exchange knowledge and share experience;
4) Days of Science and MareNostrum open days, mainly targeted to high schools and undergraduate students.

Knowledge of Spanish language, improved thanks to the free classes taken at BSC, will facilitate Stefano in the accomplishment of this non-trivial task.

3. Quality and efficiency of the implementation

3.1 Coherence and effectiveness of the work plan, including appropriateness of the allocation of tasks and resources

The very first days will be dedicated to the drafting of the communication and dissemination plan (M0.1). This document is crucial for planning the main communication activities, which will continue for the entire length of the project, and dissemination actions that will be constantly updated. In parallel, a Career Development Plan (see Section 3.2) containing project objectives, personal goals, list of publication, will be drafted and updated twice a year.

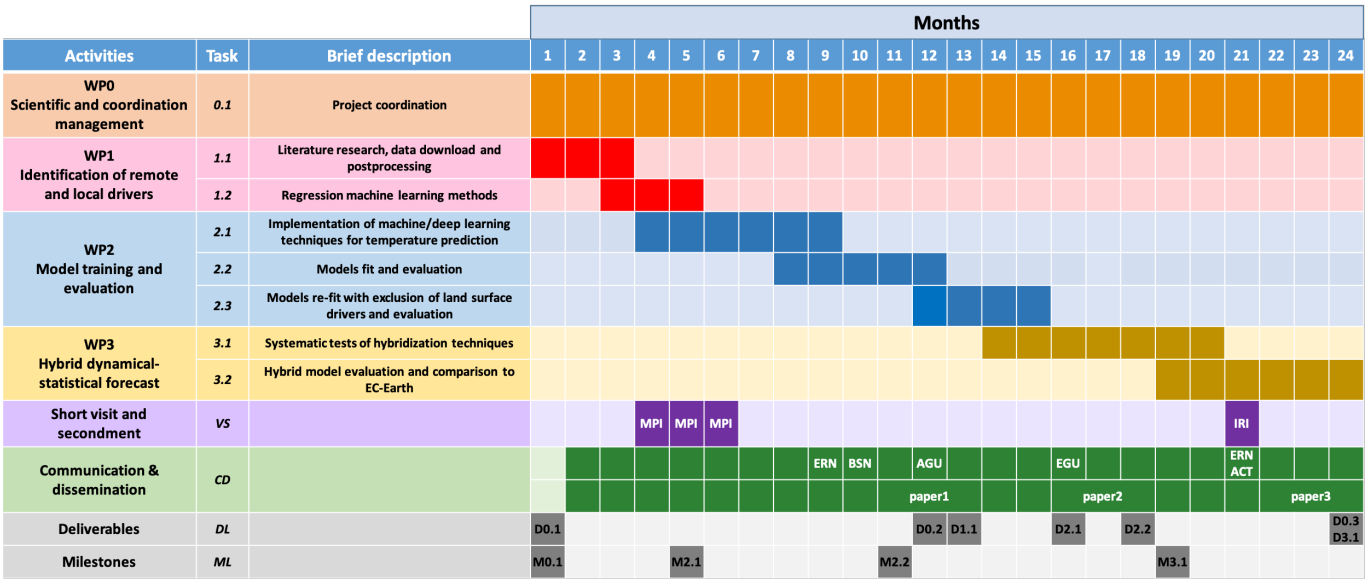


Figure 1. GANTT chart. Task and milestones are specified in the text. Deliverables. 0.1: writing of the career development plan; 0.2: first year report; 0.3 final report; 1.1 submission of paper 1; 2.1 ESD internal seminar; 2.2: submission of paper 2; D3.1: draft of paper 3. Communication and dissemination activities assume that project starts in January 2021: ERN = European Research Night, BSN = Barcelona Science Night, AGU (EGU) = American (European) Geophysical Union annual meetings; ACT = training during the short visit at IRI.

WP1 is dedicated to drivers detection: the first months will be devoted to literature research, data downloading and post-processing (Task1.1). Starting from month 4, simple regression analysis for identification of sensible predictors and then, with the help of the Diagnostic Group at MPI-BGC (secondment), regression machine learning methods for variable selection problems with many candidate predictors (T1.2), will be carried out. During the secondment, the fellow will also begin implementing machine and deep learning techniques to predict seasonal temperatures from the identified drivers (T2.1, WP2). This relevant task will continue in Barcelona, and on month 8 it will be flanked by model training and evaluation (T2.2), which will last until the end of year 1. T2.3, where the statistical models will be re-fit and evaluated again after the exclusion of the land surface drivers, will occupy the beginning of year 2, and then WP3 will take over. After a long phase for testing the most suitable hybridization techniques (T3.1), the hybrid model will be implemented, tested, and its skill compared to that of EC-Earth (T3.2).

Milestones are crucial decision points in the action: at M2.1 the extreme (heat wave and cold spell) indexes will be selected, while at M2.2 the ensemble of empirical models will be finalized. At M3.1, after months of testing and research, Stefano will finally choose the hybridization technique to build the statistical/dynamical prediction system.

The planning of the project in terms of person-months has been carefully thought, and accounts for possible delays and re-planning. Given the novelty of the topic, sufficient time in the beginning was secured for literature research and data collection, and long time was reserved for testing before selecting the most suitable tools and techniques. MPI-BGC has the expertise, experience, infrastructure, data and codes for a fast transfer of knowledge about machine and deep learning. The remarkable competence of the team involved in this project on data analytics, prediction problems, land atmosphere interaction, together with Stefano's enthusiasm, will guarantee the success of **ARTIST**.

3.2 Appropriateness of the management structure and procedures, including risk management

The BSC Project Management Office (PMO) has a broad experience with European projects and fellowships, and will take on a decisive role in the management monitoring of the fellowship, by giving assistance with bureaucratic procedures, ensuring smooth communication between departments and scheduling/reminding activities. Stefano Materia will be assigned an experienced project manager to support him in financial arrangements and administration, especially in the management of the Institutional Unit Costs that may cover participation in national and international training activities, workshops and conferences, networking events, dissemination actions, publications, work trips related to the fellowship. The financial progress will be monitored by completing the financial statements, including explanation on the use of resources.

The fellow work in close contact with the Education, Human Resources and Communications departments for training arrangements and dissemination and outreach activities. In all these regards, the BSC is an ideal institute for visitors, fellows and employees: the implementation of measures to foster research careers and strengthen internal training, based on the principles of the European Charter for Researchers and the Code of Conduct for Recruitment, lead to the prestigious Human Resources Excellence in Research award.

At the very beginning of the MSCA fellowship, Stefano and Markus will gather to draft a Career Development Plan, with the supervision of the PMO. This document will describe the fellow's short and long-term career objectives, project goals and activities such as publications or trainings. Every six months, targeted fellow-supervisor meetings will be arranged to revise and update the plan.

The project manager will also support Stefano in addressing and finding solution to possible issues arising during the fellowship. At the moment, the management and scientific risk can be considered low, and not serious enough to endanger the success of **ARTIST**. The following table summarizes potential risks and mitigation actions to be undertaken to limit their impact; likelihood comes from personal perception, experience and literature.

Identified risk	Likelihood	WP affected	Solution or mitigation actions
Too short observational dataset for one or more land-surface variables	3	1,3	Inclusion of reanalysis products, after validation versus available observational data
Loss of data or data corruption on one of the BSC archive disks	1	all	Opening of a ticket for priority reruns
Delays in learning and implementation of MPI-BGC codes and datasets	1	all	One-month prolongation of the secondment

Table 1. Risks that could arise from the project and possible solutions or mitigation measures. Likelihood is expressed on a scale of 1 to 5, where 1 is *very unlikely* and 5 is *very likely*.

3.3 Appropriateness of the institutional environment (infrastructure)

Both the BSC and MPI-BGC provide outstanding HPC infrastructures, computational resources, and IT support to perform demanding machine learning computations, crucial for the successful implementation of **ARTIST** (detailed description of the institute infrastructures can be found in Section 5). The MPI-BGC stores large amounts of code that was customized to their file systems, and will provide support for transfer of both knowledge and infrastructure. At BSC, the high-quality Computational Earth Science group will provide assistance for the optimization of new computational codes, development of tools to launch and process results, and will help Stefano Materia to overcome technical issues. Most of the data required for this project (i.e. observational datasets and EC-Earth outputs) are already downloaded and available in the BSC archive disks; the group will be of support to easily obtain any data not currently available. The BSC has extensive experience in hosting fellows and providing them project manager support. The combination of HPC facilities, high quality user support, and experience in hosting fellows will provide Stefano with a perfect environment and a robust infrastructure to guarantee a successful completion of the action.

Part B-2 Section 4 - CV of the experienced researcher

Stefano Materia (ORCID ID 0000-0001-5635-2847)
Centro Euro-Mediterraneo sui Cambiamenti Climatici
Viale Carlo Berti Pichat, 6
40127 Bologna (Italy)

Nationality: Italian
Birth date: 24/08/1979
stefano.materia@cmcc.it
https://www.cmcc.it/people/materia-stefano

Mother tongue Italian
Other languages English (proficient), Spanish (intermediate)

Dr. Stefano Materia graduated in Geophysics in June 2009. He spent his 2.5-year post-Doc between Bologna and Boulder, at the National Center for Atmospheric Research (NCAR), where he worked with Prof. Bill Large and Prof. Joe Tribbia. Since 2012 he has been at CMCC in Bologna, working both in the Climate Service and Climate Simulations and Predictions teams, headed by Dr. Silvio Gualdi. Stefano's scientific production has decreased after 2015 for two main reasons. First, he started being strongly committed in the design and development of climate services for the agribusiness and renewable energy sectors: through these service activities, a few of these still ongoing, he secured financial and human resources but paper writing was moved to the background. Second, in 2013 he founded and headed a start-up named Olio & Sale, aimed at building and organizing food chains in hidden corners of Italy for the commercialization of high-quality food. The huge commitment required to develop its own company forced Stefano to work only part-time as a researcher. Olio & Sale was sold at the end of 2017. Stefano's scientific production is recovering, with four papers in review (two at their final stage of revision) and four that will be submitted by the beginning of 2020.

WORK EXPERIENCES

01-01-2017 – present **Research scientist (CMCC)**

- Work Package leader in the Copernicus C3S framework (Operational Production of Seasonal Forecast C3S_330)
- Project designer and Principal Investigator in three founded Climate Service commissions with private companies:
  - ENEL Green Power SPA (2017-2018): 100.000 euro
  - Ferrero SPA (2017-2019): 72.000 euro
  - ENEL Green Power SPA (2019-2020): 25.000 euro
- Co-designer and Principal Investigator in the EuropeAid project LCBAB
- Task leader in the ERA4CS project MEDSCOPE
- Collaborator in the H2020 projects:
  - PRIMAVERA (WP3)
  - CRESCENDO (WP1, WP4)
  - CLARA (WP3)
- Co-advisor of the PhD thesis of Dr. Paola Marson and Ms. Ritika Kapoor (in progress)
- Personal research on:
  - climate predictability at different time-scales (from sub-seasonal to decadal)
  - impact of land surface initialization in climate predictions
  - land surface – atmosphere interaction and coupling

01-01-2012 – 31-12-2016 **Junior Scientist (CMCC)**

- Work Package leader for the FP7 project Climafrica
- Work Package co-leader for the FP7 project INDOMARECLIM
- Principal Investigator for the Climate-KIC project Wat-Ener-Cast
- Co-advisor of the PhD thesis of Dr. Claudine Wenhaji

10-06-2009 – 31-12-2011 **Post-Doc (CMCC/NCAR)**

- River discharge representation and impacts in climate models

OTHER WORK EXPERIENCES

07-01-2014 – present **Teacher (University Ca' Foscari, Venezia)**

- PhD on Science and Management of Climate change. Course title: General Circulation Models

12-02-2013 – 31-12-2017 **Founder and CEO (Olio & Sale SRL)**

- Development of supply chains for the production and commercialization of high-quality food products. Company sold at the end of 2017.

02-05-2005 – 31-03-2006 **Collaborator (ARPA Emilia Romagna)**

- Development and implementation of a dataset for disposal of waste sludge

EDUCATION AND TRAINING

01-01-2006 – 09-06-2009 **PhD student (University of Bologna)**

- PhD in Geophysics (XXI doctorate cycle)

13-09-2007 – 31-03-2008 **PhD visiting student (Centre for Ocean Land and Atmosphere Studies, COLA)**

- Visiting at the Land Surface team working with Dr. Paul Dirmeyer

09-11-1998 – 21-03-2005 **University of Bologna**

- Degree course in Environmental

PERSONAL SKILLS

Dissemination and Communication skills

- Participation in more than 25 international conferences and workshops (WCRP, AMS, EMS, EGU, AOGS, IUGG among others).
- Authored and co-authored 11 publications in peer reviewed scientific journals (citations = 189, h-index = 8)
- Media interviews: Bologna Today, Il Resto del Carlino (newspapers), TeleNorba (TV)
- Communication activities in schools (Istituto Tecnico Aldini Valeriani, Liceo Scientifico E. Fermi) symposia, and events organized by private foundations (Un pozzo di scienza, 2008; La scienza in piazza, 2011; Il giardino delle imprese, 2017; for the Marino Golinelli Foundation) and public entities (European Night of Researchers 2016-2018)

Organizational and Management skills

- Supervision or co-supervision of three PhD students and two master students
- Co-authored 11 publications in peer reviewed scientific literature (citations: ~ 189, h-index=8)
- Supervision of Post-Doc researchers involved in H2020 projects (CRESCENDO and PRIMAVERA) and contracts with private companies (Ferrero and ENEL).

Computer and technical skills

- GrADS, Python, Matlab, Shell scripting, CDO, NCO, LaTeX, Linux / Mac OS

ADDITIONAL INFORMATION

Honours and awards

- 2010 ISSNAF fellowship (65.000 USD), for a two-year post-doc between Italy and the United Stated sponsored by the Italian Ministry of Environment

Invitations

- Invited Visiting Scientist: Université of Québéc á Montreal (talk and research work, July 2019). Talk title: *Near-surface temperature response to extreme soil conditions*.
- Invited Visiting Scientist at the Barcelona Supercomputing Center (talk and research work, February - May 2019). Talk title: *A multi-model approach for subseasonal prediction of spring cold spells: potential benefits for the agribusiness sector*.
- Invited Visiting Scientist: International Research Institute for Climate and Society – Columbia University (talk, January 2018). Talk title: *The value of sub-seasonal to seasonal forecasts to face extreme events affecting hazelnut production in Northern Turkey*
- Invited speaker: HyMeX Drought and Water Resources Workshop (Apr 2016). Talk title: *Seasonal forecast of droughts: an overview*.

PUBLICATIONS

Published

- Cherchi, A., Fogli, P.G., Lovato, T., Peano, D., Iovino D., Gualdi S., Masina S., Scoccimarro E., **Materia S.,** Bellucci A., Navarra A. (2018). Global mean climate and main patterns of variability in the CMCC-CM2 coupled model. *Journal of Advances in Modeling Earth Systems,* 11(1), 185-209.
- Athanasiadis P. J., Bellucci A., Scaife A. A., Hermanson L., **Materia S.,** Sanna A., Borrelli A., MacLachlan C., Gualdi, S. (2017). A multi-system view of wintertime NAO seasonal predictions. *J. Climate*, 30, 1461-1475.
- Van den Hurk B., Kim H., Krinner G., Seneviratne S. I., Derksen C., Oki T., Douville H., Colin J., Ducharne A., Cheruy F., Viovy N., Puma M. J., Wada Y., Li W., Jia B., Alessandri A., Lawrence D. M., Weedon G. P., Ellis R., Hagemann S., Mao J., Flanner M. G., Zampieri M., **Materia S.,** Law, R. M., Sheffield J., 2016. LS3MIP (v1.0) contribution to CMIP6: the Land Surface, Snow and Soil moisture Model Intercomparison Project – aims, setup and expected outcome. *Geoscientific Model Development*, 9, 2809-2832.
- Bertolini, T., Flechard, C. R., Fattore, F., Nicolini, G., Stefani, P., **Materia, S.,** Valentini R. & Castaldi, S. (2016). DRY and BULK atmospheric nitrogen deposition to a West-African humid forest exposed to terrestrial and oceanic sources. *Agricultural and Forest Meteorology*, *218*, 184-195.
- Bellucci A., Haarsma R., Bellouin N., Booth B., Cagnazzo C., van den Hurk B., Keenlyside N., Koenigk T., Massonnet F., **Materia S.** and Weiss M., 2015. Advancements in decadal climate predictability: the role of non-oceanic drivers. *Reviews of Geophysics*, *53*(2), 165-202.
- Alessandri A, Borrelli A., Cherchi A., Lee J.-Y., Wang, **S. Materia,** and A. Navarra, 2015. Prediction of Indian summer monsoon onset using dynamical sub-seasonal forecasts. *Monthly Wea. Rev.*, 143, 778-793.
- **Materia S.,** Borrelli A., Bellucci A., Alessandri A., Di Pietro P., Athanasiadis P., Navarra A., and Gualdi S., 2014. Impact of Atmosphere and Land Surface Initial Conditions on Seasonal Forecasts of Global Surface Temperature. *J. Climate*, 27, 9253–9271.
- Athanasiadis P., Bellucci A., Hermanson L., Scaife A., MacLachlan C., Arribas A., **Materia S.,** Borrelli A., and Gualdi S., 2014. The Representation of Atmospheric Blocking and the Associated Low-Frequency Variability in Two Seasonal Prediction Systems. *J. Climate*, 27, 9082–9100.
- Fattore F., Bertolini T., **Materia S.,** Gualdi S., Thongo M'Bou A., Nicolini G., Valentini R, de Grandcourt A., Tedesco D., Castaldi S., 2014. Seasonal trends of dry and bulk concentration of nitrogen compounds over a rain forest in Ghana, *Biogeosciences*, 11, 3069-3081.

- **Materia S.,** Gualdi S., Terray L., Navarra A., 2012. The effect of Congo River freshwater discharge on Eastern Equatorial Atlantic climate variability. *Clim Dyn*, 39, 9-10, 2109-2125.
- **Materia S.,** Dirmeyer P.A., Guo Z., Alessandri A., Navarra A., 2010. The sensitivity of simulated river discharge to land surface representation and meteorological forcings. *J. Hydrometeor*, 11, 334-351.

Submitted

- **Materia S.,** Muñoz A.G., Alvarez-Castro M.C., Mason S., Vitart F., Gualdi S. (2019). Cost or loss? Assessing skill and potential value of cold-spell subseasonal forecasts. *Submitted to Weather and Forecasting*
- Peano D., **Materia S.,** Collalti A., Alessandri A., Anav A., Bombelli A., Gualdi S. (2019). Global variability of simulated and observed vegetation growing season. *Submitted to JGR - Biogeosceinces*
- Marson P., **Materia S.,** Nichka D., Tibaldi S., Gualdi S. (2019). A Process-Informed Statistical Framework for the Spatial Distribution and Intensity of Orographic Precipitation. *Submitted to Journal of Hydrology*
- Ruggieri P., Alvarez-Castro, M.C., Athanasiadis P., Bellucci A., **Materia S.**, Gualdi S. (2019). North Atlantic circulation regimes and heat transport by transient eddies. *Submitted to Journal of Climate*

In preparation

- **Materia S.,** Jung M., Bellucci A., Alessandri A., Borrelli A., Weber U., Navarra A., Gualdi S. (in prep.). Precipitation-soil moisture coupling in Sub-Saharan Africa on interannual time scale.
- **Materia S.,** Ardilouze C., Prodhomme C., Donat M.G., Doblas-Reyes F.J., Benassi M., Caron L.P., Ruggieri P., Gualdi S. (in prep.). Summer temperature response to extreme soil conditions in the Mediterranean transitional zone.
- **Materia S.**, Pausata F.S.R. (in prep.). Impacts of a tropical volcano eruption on seasonal and interannual Northern Hemisphere climate extremes
- Jha P.J., **Materia S.**, Bregaglio S., Costa Saura J., Trabucco A. (in prep.). Impact of climate change on Hazelnut production for South East Australia

Part B-2 Section 5 - Capacity of the Participating Organisations

Participating organisations	**Legal Entity Short Name**	**Country**	**Supervisor**	**Role of partner organisation**
Beneficiary				
Barcelona Supercomputing Center – Centro Nacional de Supercomputación	BSC-CNS	Spain	Dr. Markus Donat	
Partner Organisation				
Max Planck Institute for Biogeochemistry	MPI-BGC	Germany	Dr. Martin Jung	Secondment

Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS), Spain	
General description	**The Barcelona Supercomputing Center-Centro Nacional de Supercomputación** (https://www.bsc.es) combines unique high performance computing facilities and in-house top research departments on Computer, Life, and Earth sciences, and in computational applications in science and engineering. It is the main provider of public supercomputing services in Spain, coordinating the Red Española de Supercomputación and representing Spain in international initiatives such as PRACE (http://www.prace-ri.eu/). BSC-CNS was selected in June 2019 by EuroHPC as one of the three institutions to host a pre-exascale supercomputer in the high-capacity supercomputer network that will operate in the EU by 2021. The future MareNostrum5 will be a pre-exascale heterogeneous supercomputer that will achieve a peak performance of 200 Petaflops, which is 18 times more than current MareNostrum4.
Academic organisation	Yes
Role and profile of key persons (supervisor)	**Dr. Markus Donat** is an expert in studying climate extremes and climate variability, mechanisms driving or amplifying extremes, and climate model evaluation focussing on their fidelity to simulate climate extremes. He is a co-leader of the Climate Prediction group at the BSC, after holding Postdoctoral and Senior Research Fellow positions over the past eight years at the University of New South Wales in Sydney (Australia). Markus has published more than 75 peer-reviewed journal articles since 2010, eight of these in Nature-family journals, and has contributed to the IPCC 5th Assessment Report. He is Associated Investigator with the Australian Research Council Centre of Excellence for Climate Extremes, and Associated expert with the World Meteorological Organization (WMO) Commission of Climatology (CCl) Expert Team on Data Development and Stewardship. Based on his achievements he has been awarded the World Climate Research Program (WCRP)/Global Climate Observing System (GCOS) International Data Prize 2017. It is also worth mentioning, that Markus has a vast experience in supervising students and post-docs including two MSCA-IF fellows (Dr. Rachel White, PROTECT fellowship; Dr. Yohan Ruprich-Robert, INADEC fellowship). The ARTIST fellowship will be co-supervised by **Prof. Francisco Doblas-Reyes**, the director of BSC Earth Sciences Department, who is a worldwide expert in the development of seasonal-to-decadal climate prediction systems and has more than 20 years of experience in weather and climate modelling, climate prediction, as well as the development of climate services
Dept./Division / Laboratory	**The Earth Sciences (ES) Department** focuses on the atmosphere-ocean-biosphere system and is structured around four groups with more than 100 researchers and support staff. It is a highly productive scientific entity that has published more than 220 research peer-reviewed articles over the last 5 years, many in high-impact journals. Within the ES Department, **the Climate Prediction Group (CPG)** aims at developing a climate forecast system based on the Earth System Model EC-Earth. The CPG also performs regular assessments of the system's predictive capacity and compares it with other operational and quasi-operational systems in the world. The CPG has a long experience in seasonal to decadal climate prediction, which has been reflected in its active participation to several European projects with a strong component on climate prediction. Of particular relevance was the FP7 project SPECS, led by the BSC, in which specific, innovative global forecast system experiments were coordinated to test hypotheses for the improvement of seasonal to decadal predictions. The CPG currently participates to 10 European and 7 national projects.
Key research facilities, Infrastructure and Equipment	BSC hosts and manages a range of HPC systems, including MareNostrum 4, with 148,176 cores and 13.7 Pflops capacity. Additionally, BSC manages Minotauro, a Sandy Bridge's cluster with NVIDIA GPUs, providing more than 100 TFlops.
Independent research premises?	Yes. All key research facilities, infrastructure, and equipment will be available for the fellow.
Previous and current involvement in research and training programmes	BSC-CNS has a large record of participation and coordination of research and training activities: a dedicated unit of *Education and training* is in charge of the organization of specialized training sessions in HPC and related activities, through the PRACE training courses. Also, BSC is the main beneficiary of a MSCA COFUND program, for postdoc fellows, which implements a training programme **STARS,** Grant Agreement number 754433. At present, BSC-CNS is involved in 105 projects funded by the Horizon 2020 programme and 11 projects are coordinated by the center. The most relevant to ARTIST proposal and hosted by the Earth Sciences department are the following projects: **C3S_51 Lot3 QA4Seas:** *Quality Assessment Strategies for Multi-model Seasonal Forecast;* Copernicus Climate Change contract coordinated by BSC (Prof. Francisco Doblas-Reyes). Total value of the contract: EUR 1,681,759.97. **PROTECT:** *Propagation of atmospheric ROssby waves - connection to prEdictability of Climate exTremes.* MSCA-IF fellowship supervised by Dr. Markus Donat; **CCiCC:** *Climate-Carbon Interactions in the Coming Century,* recently awarded Horizon 2020 project in the Climate Prediction Group with the total funding of EUR 7,784,750.00 **IMPREX:** *IMproving PRedictions and management of hydrological Extremes,* H2020-WATER-2014; EUR 7,996,850.00.
Relevant publications and/or research/innovation products	1. **Donat, M. G.**, A. J. Pitman, S. I. Seneviratne (2017), Regional warming of hot extremes accelerated by surface energy fluxes, *Geophys. Res. Lett.*, **44**, 7011–7019. 2. **Donat, M. G**., Alexander, L. V., Yang, H, and Coauthors (2013). Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. *J. Geophys. Res.*, **118**, 2098-2118. 3. **Donat, M. G.**, G. C. Leckebusch, J. G. Pinto, and U. Ulbrich (2010), Examination of Wind Storms over Central Europe with respect to Circulation Weather Types and NAO phases, *Internat. J. Climatol.*, **30**, 1289 - 1300. 4. Lorenz, R., D. Argüeso, **M. G. Donat**, and Coauthors (2016), Influence of land-atmosphere feedbacks on temperature and precipitation extremes in the GLACE-CMIP5 ensemble, *J. Geophys. Res.*, **121**, 607–623. 5. Oliver, E. C.J. **M. G. Donat**, M. T. Burrows, and Coauthors (2018), Longer and more frequent marine heatwaves over the past century, *Nature Comm.*, **9**, 1324-1336.

Max Planck Institute for Biogeochemistry (MPI-BGC), Germany	
General description	The **Max-Planck-Institute for Biogeochemistry (MPI-BGC)** is a research institute of the German MaxPlanck Society (MPG). Its research mission is the investigation of the global biogeochemical cycles and their interaction with the climate system. The institute combines strong observational and process-based studies (soil carbon, plant community structure, nutrition and growth, vegetation-atmosphere fluxes, convective boundary layer) with global scale modelling (e.g. vegetation dynamics, global carbon cycle, aerosol modelling). It is one of the pivotal European Research institutions in its field. In a MPG evaluation of publication citations (2013), MPI-BGC was attributed an "outstanding publication record". In 2010-2015, researchers at the institute have published 155 papers a year, with an increasing trend, in highly respected peer-reviewed (e.g. including in the same period 40 papers in Nature, Science, PNAS).
Academic organisation	Yes
Role and profile of key persons (supervisor)	Dr. Martin Jung leads the group of **Global Diagnostic Modelling (GDM)** at the Max Planck Institute for Biogeochemistry in Jena and will be mentoring Stefano Materia during his secondment, and will support him and train him on the use of in-house datasets, codes and tools. Since his PhD in 2007 he has become an expert in machine learning applications for land-atmosphere interaction problems. His research interests span across disciplines of machine learning, remote sensing, land surface modelling, atmospheric inversion techniques, and in-situ flux measurement community with numerous collaborations going on. He was principal investigator of the FLUXCOM initiative and participated in more than ten projects funded by the EU or the ESA. He has been advising six PhD students and four PostDocs.
Dept./Division / Laboratory	The GDM led by Dr. Jung is part of the Department of Biogeochemical Integration, whose research aims at understanding how the terrestrial biosphere reacts to and exerts feedbacks on ongoing environmental change and variation in atmospheric conditions. GDM develop and analyse global data-driven estimates of land surface carbon and water fluxes as well as vegetation properties. For example, they are integrating in-situ measurements from FLUXNET with satellite remote sensing and meteorological reanalysis using machine learning techniques, to estimate the global spatial and temporal distribution of primary production and evapotranspiration. The department is strongly involved in the coordination and participation in 20 ongoing EU and ESA project, many of which dealing with land-surface monitoring and land-atmosphere interaction.
Key research facilities, Infrastructure and Equipment	The MPI-BGC maintains several experimental sites, equipped with eddy covariance systems (above- and below- canopy), radiometric towers, soil respiration chambers, walk-in lysimeters, electronic dendrometers as well as established campaign protocols for a large number of manual survey measurements. They combine experiments and in-situ long-term observation with Earth measurements gathered by aircraft and satellites across a range of spatial scales, and embrace data-driven machine learning and theory-driven mechanistic modelling to make observation usable in climate models. Computing infrastructure includes the hosting and managing of large-scale databases (e.g. FLUXCOM – a global database of carbon and energy fluxes over land (http://www.fluxcom.org)) and a large multi-processor cluster, suitable for computing-intensive tasks as may be required for the computation of deep learning techniques.
Independent research premises?	YES, the Institute and research groups in the proposal have independent research premises and office space.
Previous and current involvement in research and training programmes	The MPI-BGC has an outstanding international reputation for research, coordinated a series of ECfunded projects e.g. CARBOEUROPE-IP and CARBO-Extreme and was involved in fourteen collaborative FP7 EU projects, as well as in a career Integration grant (FP7-PEOPLE-2012-CIG) of Mathias Göckede (BGS department - Martin Heimann) and an Intra-European Fellowship Marie Curie project (FP7- PIEF-GA-2013-626334) of John Kim (BGP department- Susan Trumbore). Additionally, the institute has a strong commitment to higher education, scientific training and outreach, hosting since 2010 the international Max Planck Research School for Global Biogeochemical Cycles (http://www.imprs-gbgc.de/) and was the coordinating institute for the CarboSchools project (January 2008 to December 2010) with funding from the EU Science in Society programme. In addition, the MPI-BGC has built a searchable infrastructure to guarantee the dissemination of the products to the wider public. MPI-BGC hosts 60 Ph.D. students from 26 countries and operating the International Max Planck Research School for global Biogeochemical Cycles, (IMPRS). In cooperation with Friedrich Schiller University Jena, the MPI-BGC houses a unique and flexible research program in global biogeochemistry and related earth system sciences that grants a broad selection of learning opportunities while still maintaining a with research focus and with a searchable infrastructure to guarantee the dissemination of the products to the wider public. In H2020 MPI-BGC has been awarded with 8 projects, and is currently coordinator of the BACI project (H2020-SPACE-GA 640170), hosts the ERC consolidator grant QUINCY (ERC-2014- CoG_GA 647204), the ETN Grant TRustEE (H2020-MSCA-ITN-2016_GA 721995), the ERC Advanced Grant -14 Constraint (ERC-2015-AdG_GA 695101), the INTAROS Grant (H2020-BG2016-1_GA 727890). The MPI-BGC maintains numerous collaborations in a wide range of geographical locations and scientific disciplines.
Relevant publications and/or research/innovation products	1. Reichstein, M., Camps-Valls, G., Stevens, B., **Jung, M**., Denzler, J., Carvalhais, N., Prabhat (2019). Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743), 195-204. 2. **Jung, M**., Zscheischler, J. (2013). A guided hybrid genetic algorithm for feature selection with expensive cost functions. Procedia Computer Science, 18, 2337-2346. 3. **Jung, M.**, Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., Reichstein, M. (2019). The FLUXCOM ensemble of global land-atmosphere energy fluxes. Scientific Data, 6: 74. 4. **Jung, M.**, Reichstein, M., and Coauthors (2010). Recent decline in the global land evapotranspiration trend due to limited moisture supply. Nature, 467(7318), 951-954. 5. **Jung, M.**, Reichstein, M., Bondeau, A. (2009). Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model. Biogeosciences, 6, 2001-2013.

Part B-2 Section 6 - Ethical Issues

No ethical issues were flagged in the Ethics Issues Table, and thus the proposal meets the EU and national legal and ethics requirements.

Part B-2 Section 7 – Letter of support

Max-Planck-Institut
für Biogeochemie

To whom it may concern

Postfach 10 01 64
07701 Jena
Tel. +49 (0)3641 57-6261
Fax +49 (0)3641 57-7200
E-Mail: mjung@bgc-jena.mpg.de
Web www.bgc-jena.mpg.de
Direktorium
Prof. Susan E. Trumbore, Ph.D.
Prof. Dr. Markus Reichstein (GfD)

Letter of support for the 'ARTIST' proposal by Dr. Stefano Materia

Herewith I express my strong support for the ARTIST proposal by Dr. Materia. The proposed integration of machine learning in seasonal weather prediction is novel and very promising to yield significant improvements that are crucial from a socioeconomic but also scientific perspective. Our department has put large efforts in tailoring machine learning approaches for various environmental questions such as land-atmosphere interactions. We are strongly collaborating with machine learning experts for scrutinizing and exploiting the latest developments in this field for our scientific problems. I am therefore happy to host and mentor Dr. Materia in conducting machine learning related tasks of the proposal. The science carried out in the ARTIST proposal offers various synergies with our activities, for example related to forecasting anomalies of carbon balance and land-atmosphere energy fluxes through improved weather forecasts. Including our global data-driven products of land-atmosphere energy fluxes in the predictor set of ARTIST can further improve the weather prediction skill since those are critical for land-atmosphere feedbacks but uncertain in dynamical models.

I look very much forward to this collaboration and the very exciting scientific questions to tackle.

With best regards,

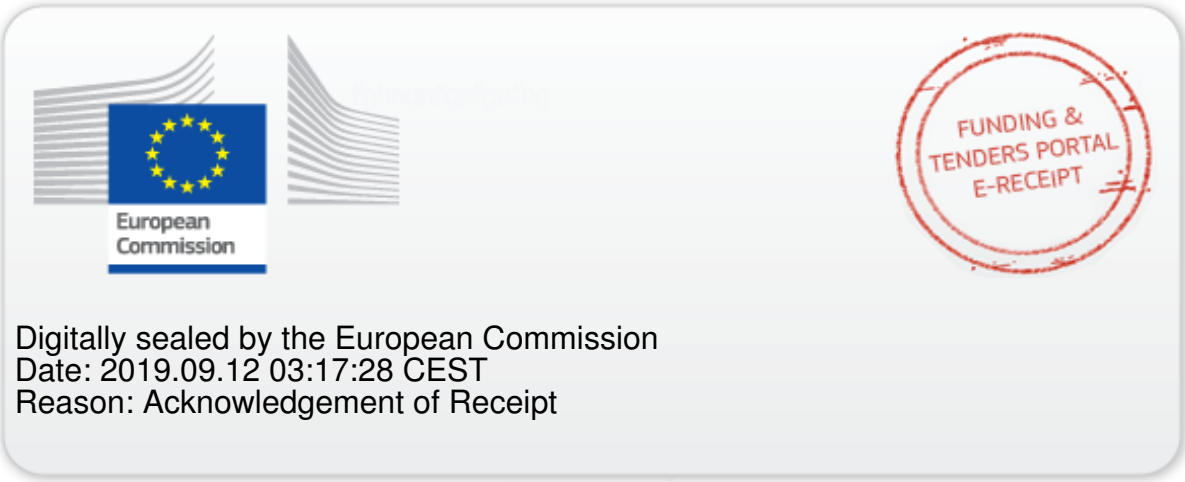
Dr. Martin Jung

Jena, 20th August. 2019

Hans-Knöll-Straße 10
07745 Jena
Germany

Tel.: +49 (0)3641 57-60
Fax: +49 (0)3641 57-70

ID-Nr. DE 129517720
www.bgc-jena.mpg.de



Digitally sealed by the European Commission
Date: 2019.09.12 03:17:28 CEST
Reason: Acknowledgement of Receipt

This electronic receipt is a digitally signed version of the document submitted by your organisation. Both the content of the document and a set of metadata have been digitally sealed.

This digital signature mechanism, using a public-private key pair mechanism, uniquely binds this eReceipt to the modules of the Funding & Tenders Portal of the European Commission, to the transaction for which it was generated and ensures its full integrity. Therefore a complete digitally signed trail of the transaction is available both for your organisation and for the issuer of the eReceipt.

Any attempt to modify the content will lead to a break of the integrity of the electronic signature, which can be verified at any time by clicking on the eReceipt validation symbol.

More info about eReceipts can be found in the FAQ page of the Funding & Tenders Portal.

(https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/support/faq)