Horizon 2020

Call: H2020-MSCA-IF-2016
(Marie Skłodowska-Curie Individual Fellowships)

Topic: MSCA-IF-2016

Type of action: MSCA-IF-EF-ST
(Standard EF)

Proposal number: 741379

Proposal acronym: CAMBIO

Deadline Id: H2020-MSCA-IF-2016

Table of contents

Section	Title	Action
1	General information	
2	Participants & contacts	
3	Budget	
4	Ethics	
5	Call-specific questions	

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 previous steps in the submission wizard.

Proposal ID **741379**   *Acronym* **CAMBIO**

1 - General information

Topic	MSCA-IF-2016
Call Identifier	H2020-MSCA-IF-2016
Type of Action	MSCA-IF-EF-ST
Deadline Id	H2020-MSCA-IF-2016

Acronym   CAMBIO

Proposal title   Climate dAta Mining for a Better InfOrmation

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

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   *Climatology and climate change*

Descriptor 2   *Sustainability sciences, environment and resources*

Descriptor 3   *Machine learning, statistical data processing and applications using signal processin*   [Remove]

[Add]

Free keywords   *climate services, climate forecasts, data mining*

Proposal ID **741379**	*Acronym* **CAMBIO**

Abstract

Climate plays an important role in a many sectors of our society and then many important decisions, from the timing to reduce the effects of a drought to the best way to manage a water basin, are made by using "the best information" we have on past, present and future climate. This information comes from multiple heterogeneous sources: weather stations, satellites and numerical models. In the last decade, the quality and the resolution of the models has increased and is also increased the time-scale in which the models are able to provide "useful" information. In fact, nowadays climate forecasts — at particular conditions — may predict the climate for a specific region for the incoming season, creating a vast range of opportunities for all the sectors that need to make decisions at seasonal time-scales. But those improvements have also led to new challenges: climate information at seasonal time-scales is inherently probabilistic and it is normally produced by ensemble simulations, increasing drastically the complexity of the climate datasets. Nevertheless, the use of tools to deal with large-scale data are seldom applied to climate data management.

This project has the aim to take advantage of the latest improvements in data science to perform data mining of large climate datasets easily and in a streamlined way. It proposes a big data processing framework able to extract the maximum amount of information from multiple datasets, applying machine learning and data dimensionality reduction methods on large-scale climate data. This framework will be used as a basis to carry out: 1) a systematic analysis of the information content of seasonal climate forecasts hindcasts with respect to observational datasets and 2) data-driven discovery of the best predictors to model a meteorological variable. Finally, to demonstrate the potential impact of the proposed approach, a set of case-studies will be chosen to evaluate the added value of the enhanced climate information.

Remaining characters   5

Has this proposal (or a very similar one) been submitted to a Horizon 2020 Marie Skłodowska-Curie Individual Fellowship call?	○ Yes ◉ No

Proposal ID **741379**	Acronym **CAMBIO**

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) confirms:

- to have carried out the self-check of the financial capacity of the organisation on https://ec.europa.eu/research/participants/portal/desktop/en/organisations/lfv.html or to be covered by a financial viability check in an EU project for the last closed financial year. Where the result was "weak" or "insufficient", the applicant (future beneficiary) confirms being aware of the measures that may be imposed in accordance with the H2020 Grants Manual (Chapter on Financial capacity check); or	◯
- is exempt from the financial capacity check being a public body including international organisations, higher or secondary education establishment or a legal entity, whose viability is guaranteed by a Member State or associated country, as defined in the H2020 Grants Manual (Chapter on Financial capacity check); or	◉
- as sole participant in the proposal is exempt from the financial capacity check.	◯

5) 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.

According to Article 131 of the Financial Regulation of 25 October 2012 on the financial rules applicable to the general budget of the Union (Official Journal L 298 of 26.10.2012, p. 1) and Article 145 of its Rules of Application (Official Journal L 362, 31.12.2012, p.1) applicants found guilty of misrepresentation may be subject to administrative and financial penalties under certain conditions.

Personal data protection

Your reply to the grant application will involve the recording and processing of personal data (such as your name, address and CV), which will be processed pursuant to Regulation (EC) No 45/2001 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data. Unless indicated otherwise, your replies to the questions in this form and any personal data requested are required to assess your grant application in accordance with the specifications of the call for proposals and will be processed solely for that purpose. Details concerning the processing of your personal data are available on the privacy statement. Applicants may lodge a complaint about the processing of their personal data with the European Data Protection Supervisor at any time.

Your personal data may be registered in the Early Warning System (EWS) only or both in the EWS and Central Exclusion Database (CED) by the Accounting Officer of the Commission, should you be in one of the situations mentioned in:
-the Commission Decision 2008/969 of 16.12.2008 on the Early Warning System (for more information see the Privacy Statement), or
-the Commission Regulation 2008/1302 of 17.12.2008 on the Central Exclusion Database
(for more information see the Privacy Statement) .

Proposal ID **741379**   *Acronym* **CAMBIO**

List of participants

#	Participant Legal Name	Country
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 of the organisation

Street   Calle Jordi Girona 31

Town   BARCELONA

Postcode   08034

Country   Spain

Webpage   www.bsc.es

Legal Status of your organisation

Research and Innovation legal statuses

Public body ……………………………………… yes   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

Academic Sector …………………………………… yes

NACE Code: 72 - Scientific research and development

Does this participant deliver doctoral degrees that are recognised as such by the relevant national authorities?   ○ Yes   ◉ No

Department(s) carrying out the proposed work

Department 1

Department name | Earth Sciences   ☐ not applicable

☐ Same as organisation address

Street | 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.

Proposal ID **741379**	Acronym **CAMBIO**	Short name **BSC**

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.

Researcher ID http://orcid.org/0000-0002-5457-3045

Last Name* DE FELICE

Last Name at Birth

First Name(s)* Matteo

Gender* ◉ Male ○ Female

Title Dr.

Country of residence* Italy

Nationality* Italy

Nationality 2

Date of Birth (DD/MM/YYYY) 29/03/1982

Country of Birth* Italy

Place of Birth Rome

Contact address

Current organisation name ENEA

Current Department/Faculty/Institute/Laboratory name Department of Sustainability, Climate Impacts & Modelling Laboratory

☐ Same as organisation address

Street Via Anguillarese 301

Postcode/Cedex 00123

Town Rome

Phone +390630484582

Country Italy

Phone2 / Mobile +xxx xxxxxxxx

E-Mail* matteo.defelice@enea.it

Proposal ID **741379**	*Acronym* **CAMBIO**	*Short name* **BSC**

Qualifications

University Degree	Date of award (DD/MM/YYYY)	21/12/2007
Doctorate (in progress)	Date of award (DD/MM/YYYY)	
Doctorate	Date of award (DD/MM/YYYY)	28/03/2011
Full time postgraduate research experience	Number of months	84
Other Academic qualifications	Date of award (DD/MM/YYYY)	

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, until the call deadline (14/09/2016).

Period from	Period to	Duration (days)	Country
10/09/2009	14/09/2016	2562	Italy
	Total	2562	

Proposal ID **741379**	*Acronym* **CAMBIO**	*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* **Oriol**

Last name* **Jorba**

E-Mail* **oriol.jorba@bsc.es**

Position in org. Atmospheric Composition Group Manager

Department Earth Sciences

☐ Same as organisation address

Street Jordi Girona, 29

Town Barcelona

Post code 08034

Country Spain

Website

Phone 934134050

Phone 2 *+xxx xxxxxxxx*

Fax *+xxx xxxxxxxx*

Other contact persons

First Name	Last Name	E-mail	Phone
Dorota	CHMIELEWSKA	dorota.chmielewska@bsc.es	+34 934134082

Proposal Submission Forms

Proposal ID **741379**		*Acronym* **CAMBIO**

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	0,976	24	108921,60	14400,00	12000,00	19200,00	15600,00	170121,60
Total				24	108921,60	14400,00	12000,00	19200,00	15600,00	170121,60

Partner Organisation from Third Country does not sign the Grant Agreement, does not recruit the researcher and does not directly claim costs from the action. The entire EC contribution is transmitted to the Host organisation located in Members States or Associated Countries.

Proposal ID **741379**	Acronym **CAMBIO**

4 - Ethics issues table

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	

Proposal ID **741379**	Acronym **CAMBIO**

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

Proposal ID **741379**   Acronym   **CAMBIO**

5 - Call specific questions

Eligibility Researcher (future fellow)

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

Other Questions

For communication purposes only, the REA asks for permission to publish the name of the researcher (future fellow) should the proposal be retained for funding.

1. Does the researcher (future fellow) give this permission?   ◉ Yes ○ No

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

Proposal ID **741379**	Acronym **CAMBIO**

Data management activities

A new focus within Horizon 2020 is data management, for example through the use of Data Management Plan (DMP).

DMPs detail what data the project will generate, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved.

The use of a DMP is required for projects participating in the Open Research Data Pilot in the form of a deliverable in the first 6 months of the project (possible updates during the project).

Other projects are invited to submit a DMP if relevant for their planned research.

Are data management activities relevant for your proposed project?	○ Yes	● No

Open Research Data Pilot in Horizon 2020

All applicants can participate in the Pilot on Open Research Data in Horizon 2020[1] on a voluntary basis. This Pilot aims to improve and maximise access to and re-use of research data generated by actions.

Participants in the Pilot will be invited to formulate a Data Management Plan (DMP). DMPs detail what data the project will generate, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved.
Participating in the Pilot is flexible in the sense that it does not mean that all research data needs to be open. Rather, projects can define certain datasets to remain closed via a Data Management Plan (DMP).

Please note that participation in this Pilot does not constitute part of the evaluation process. Proposals will not be evaluated favourably because they participate in the Pilot on a voluntary basis.

We wish to participate in the Pilot on Open Research Data in Horizon 2020 on a voluntary basis	○ Yes	● No

[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.*

DOCUMENT 1

START PAGE

MARIE SKŁODOWSKA-CURIE ACTIONS

Individual Fellowships (IF)
Call: H2020-MSCA-IF-2016

PART B

"CAMBIO"
Climate dAta Mining for a Better InfOrmation

This proposal is to be evaluated as:

Standard EF

TABLE OF CONTENTS

List of Participating Organisations

Participants	Legal Entity Short Name	Academic (tick)	Non-academic (tick)	Country	Dept./ Division / Laboratory	Supervisor	Role of Partner Organisation
Beneficiary							
Barcelona Supercomputing Center	BSC	x		Spain	Department of Earth Sciences	Dr Oriol Jorba Casellas	Host organisation

1. Excellence

1.1 Quality and credibility of the research/innovation action (level of novelty, appropriate consideration of inter/multidisciplinary and gender aspects)

1.1.1 Introduction / State-of-the-art/ Objectives / Overview

Introduction

Nowadays, the scientists agree that we are living in the "Era of Data" where the data scarcity has been replaced by abundance in almost all the scientific fields (from theoretical physics to social science). Climate scientists are observing a dramatic rise of data availability with a steady increase both in spatial and temporal resolution: to give a clear example, data archived at ECMWF (the European Centre for Medium-Range Weather Forecasts, the main European climate & weather data provider) grows by about 120 TB each day and currently their storage archives 107 PB of data, about the 240% of 2014[1]. Furthermore the use of climate ensembles, a system consisting of different possible realisations of a climate model, has multiplied the amount of climate and weather data generated. This increase of complexity has not led to a more widespread use of tools to manage and analyse the data more efficiently, despite the fact that other scientific areas have experienced the so-called "Big Data Revolution" and the resulting blossoming of a plethora of software libraries, frameworks and tools to analyse, process and visualize large-scale datasets.

The first goal of this project is to develop a large-scale climate data processing framework in order to carry out operations in a streamlined and efficient way on climate large datasets. Once this data management platform has been set up, the project will focus on the demonstration of the potentiality of data-driven approaches on climate data analysis, extracting the maximum information from a very large number of possible predictors for a climate variable (or index). This project is an answer to the urgent need of our society to improve the decision-making in all the cases when climate plays a critical role. Furthermore, it fits with the European strategies of DG CLIMA like the 2020 Climate and Energy package.

State of art

Currently, there are many options for large-scale data processing frameworks, from general-purpose solutions like Apache Spark to other solution focused on specific applications (e.g. COMP Superscalar). Those tools have become common in some domains but in climate-related applications they are rarely used, one of the few applications that can be found is for example the NASA/JPL SciSpark.

In the last decades, many methods for regression and classification have been created. Tree based methods, like Classification and Regression Trees (CARTs), have the advantage of being at the same time powerful and easy to interpret. Recently, the Random Forests algorithm (proposed in 2001 by Leo Breiman) have become a very common tool thanks to its ability to cope with the high-dimensionality of input data and thus will be considered here as one of the main methods to classify and predict climate variables.

One of the rare applications of ML techniques to climate data can be found in the work by Kamale et al.[2] on the automatic discovery of long-distance climate phenomena. Dimensionality reduction techniques are widely used in climate science to find the most informative component of a dataset. Principal Component Analysis (PCA, also called EOF Analysis) is a well-established method that can be easily found already implemented in most of the scientific software and programming libraries. ICA (Independent Component Analysis) is instead a non-linear method used to separate non-Gaussian statistically-independent signals into underlying informational components[3].

Objectives

CAMBIO project aims to propose an innovative framework to deal with large-scale climate datasets. The main objectives will be the following:

1) To develop a climate data processing layer on the top of an existing large-scale data processing software;

---

[1] These statistics are as of June 2016 (see http://www.ecmwf.int/en/about/media-centre/key-facts-and-figures)

[2] Kawale, Jaya, Michael Steinbach, and Vipin Kumar. "Discovering Dynamic Dipoles in Climate Data." In SIAM International Conference on Data Mining (*SDM)* (2011): 107-118.

[3] Probably the most comprehensive introduction on ICA can be found in Hyvärinen, Aapo, Juha Karhunen, and Erkki Oja. Independent component analysis. Vol. 46. John Wiley & Sons, 2004.

2) Implement a framework that find the optimal set of predictors (observations or forecasts) for a meteorological variable by using deterministic/probabilistic (information-theory based) metrics and machine learning techniques (such as Random Forests);

3) Improve the present skill of dynamical seasonal forecasts in predict discrete events with respect to the decision-making for the various sectors.

To summarize, the methodology proposed and developed within CAMBIO project has a twofold goal: enhance the extraction of information from climate datasets in an innovative and efficient way and demonstrate the added value of those information for a set of real-world decision-making models. The expected major outcomes are a set of software libraries (released with an open source license) to be used to deal with large-scale climate datasets to carry out analysis and data mining and to facilitate the exchange of ideas between data science and climate science communities.

1.1.2   Research Methodology and Approach

CAMBIO project will focus on the transformation of data into information starting from a vast collection of climate data (observations, reanalyses, climate model outputs). This means that the entire work, despite it will be based on the solid foundation of information theory and statistical modelling, will be particularly data intensive.

The project will be undertaken in two stages. The first stage (WP1) develops a processing layer able to deal with large climate datasets in an efficient way. The datasets will be of three different types: observational datasets (data from weather stations, satellites, possibly interpolated on a regular grid), climate reanalyses and climate forecasts, normally provided by ensemble prediction systems. Data will be obtained by public repositories and, when possible, from the Copernicus Climate Data Store (CDS), which data should be freely available from a set of distributed data repositories. Furthermore, the source of seasonal forecasts will be the hindcasts of APCC/ClIPAS[4], NMME (North-American Multi-Model Ensemble), the project FP6 ENSEMBLES and a subset of the latest seasonal forecast models from the WMO Global Producing Centres for long-range forecasts. Preliminary agreements with data providers have been already established.

Stage two (WP2 and WP3) involves using machine-learning techniques to find the optimal set of meteorological predictors to estimate and predict a specific variable, also considering the impact of this improved information for the decision-making.

The applicant's research method is to assess the amount of information (i.e. data that can be used to make decisions) contained in climate forecasts and to extract it with the maximum efficiency. The measures that will be used to evaluate the quantity (and quality) of information will be the traditional error measures and skill scores and, particularly important when working with probabilistic forecasts, information theory based measures (entropy, mutual information, redundancy score).

While the metrics will be the way to quantify the information, the tools that will allow to reduce data dimensionality and extract information content will be linear and nonlinear decomposition methods (PCA and ICA). Those methods will allow to reduce the data redundancy of climate ensembles and at the same time to decompose the original data fields into minimally correlated and statistically independent patterns. Then, the exploration of the optimal predictors will be performed with search metaheuristics and machine-learning (ML) techniques for feature selection and statistical modelling.

1.1.3   Originality and innovative aspects

CAMBIO project will reduce a gap within climate science and information science communities fostering scientific collaboration and the consequent exchange of ideas and tools. One of the main expected advancements of CAMBIO project is a methodology able to ingest a large amount of climate data in order to find the optimal set of predictors for a specific climate event. The proposed approach will consists of an unprecedented application of data mining and machine learning techniques to climate data and in particular to ensemble climate forecasts.

Another advancement will be the development of a novel processing layer for an existing large-scale computing framework to analyse and explore climate data. In general, the whole project will develop the global research on the study of dynamical seasonal climate forecasts and their application to real-world problems, with a close link with the WMO GFCS (Global Framework on Climate Services) and Copernicus Programme.

---

[4] The Asian-Pacific Economic Cooperation (APEC) Climate Center (APCC)/Climate Prediction and its Application to Society (CliPAS) project started in 2005 with the objective to develop a multi-model ensemble prediction system for seasonal and sub-seasonal time scales.

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

Exposure to the highly dynamic research environment at the BSC will allow the applicant to hugely advance his maturity and widen his research interests and intellectual curiosity. The project content, and the applicant profile, will place him in the ideal condition to take full advantage of all benefits and expertise that BSC has to offer. The training objectives are the following:

1) Thanks to the expertise of Dr. Jorba Casellas, the candidate will significantly improve his knowledge in climate forecasting and the management of large-scale climate data to deal with complex applications. The applicant will have the opportunity to learn from Dr. Jorba Casellas and its main collaborators what are the main challenges in producing high-quality climate data and using it at operational level, having the occasion to examine the workflow of complex models like the NMMB/BSC-CTM (a weather model applied to air quality developed in collaboration with NCEP) and its use for real-world application (for example the prediction of efficiency of concentrated solar power plants). The formative background of Dr. Jorba Casellas is very extensive and has demonstrated a very successful experience in training postdoctoral scientists;

2) The multidisciplinary environment in BSC and the expertise in visualization & computer sciences (the applicant will also collaborate with the Scientific Visualization group at BSC) will provide a unique approach, constructing synergies and new networks among top worldwide researchers, giving the opportunity to gain precious knowledge on computer sciences and data exploration;

3) Working in close collaboration with other scientists at BSC and benefiting of the personalized programmes that the BSC provides, the applicant will have many opportunities to enhance his complementary skills. These will range from writing scientific papers and conference participation in the field of climate science up to contribute to the preparation of scientific proposals: the project, facilitating the exchange of ideas between people with different backgrounds, will also give the occasion to plan the participation to European project calls and initiatives on Big Data and Climate or on Climate Services.

The applicant will also disseminate the results of the project to the other scientists at BSC with two objectives: 1) to promote the discussion with other researchers on the project's topics discussing the future research lines in the Earth Sciences Department; 2) to support the research activities of post-docs and Ph.D. students at BSC involved in research relevant to the topics of the project.

The applicant's past experiences on climate services projects and his inter-disciplinary background will help to consolidate the capability of the BSC Department of Earth Sciences to apply climate forecasts in multiple sectors. Moreover, the applicant will transfer expertise to the host, stemming from his experience in machine learning applied to climate data-based applications, to this end the host will also benefit of the applicant's collaboration contacts of companies and research institutes working on machine learning and Big Data.

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

Qualifications and experience of the supervisor

Dr. Oriol Jorba is a leader of the BSC-ES Atmospheric Composition group since 2008. He has participated in projects funded by the European Commission on air quality, specifically in aerosols, (e.g. ACTRIS2) and in the application of atmospheric modeling in HPC (e.g. IS-ENES2). He has lead the research project on the development of the multiscale chemical weather forecasting system NMMB/BSC-CTM which is the official model used by the Barcelona Dust Forecast Center (BDFC), the World Meteorological Organization (WMO) Regional Meteorological Center specialized on Atmospheric Sand and Dust. As a Spanish representative member of the management committee of COST Actions ES1002 and ES1004, he is part of the International Technical Meeting on Air Pollution Modelling and its Application (ITM) scientific committee since 2012. He maintains regular collaborations on aerosol modelling with the NASA Goddard Institute for Space Studies, National Center for Environmental Prediction, Finish Meteorological Institute, among others. He is a scientific reviewer of the Scientific Commission of the Spanish National Research Program. His research expertise include high resolution mesoscale meteorology and air quality, development of online meteorology-chemistry models, boundary layer, atmospheric chemistry studies and environmental impact assessment. He has co-authored more than 50 papers in international scientific journals, over 100 communications in international conferences and directed 6 PhD thesis.

Oriol Jorba has successfully supervised 4 PhD students since 2008 and nowadays is advisor of 2 PhD students. Additionally, he has been the mentor of two national FPI PhD fellowships, he has been part of the research group in

charge of one Marie Curie Intra-European Fellowships in the call FP7-PEOPLE-2013-IEF, and act as experienced researcher of one Marie Skłodowska-Curie Individual Fellowship in the call H2020-MSCA-IF-2015.

Hosting arrangements

The applicant will have access to the supercomputing facilities of the BSC. The sole facility and expense to allow the applicant to work will be the purchase of a desktop computer, which will be provided at the start of the project. Software tools are already present in the host institution if required, however most are expected to be free and open-source. BSC also has a well-furnished library and subscription to on-line journals. The applicant will be transferred to the BSC and work within the BSC-ES department. This will permit direct contact with colleagues and the supervisor through regular team meetings. Synergies with other BSC departments will be fostered, the applicant will also establish a solid collaboration with the BSC Computer Sciences department, especially with Rosa Badia and the group of Workflows and Distributed Computing, which will integrate the applicant's knowledge on Big Data, data-intensive research and machine learning and with the Scientific Visualization group of the Computer Applications in Science & Engineering department. Furthermore, the applicant will take advantage of the international networking opportunities offered by BSC, for example ERA4CS (European Research for Climate Services). BSC's Project Management Office will be a contact point for the researcher during the entire duration of the fellowship and will be responsible for a preparation of the Grant Agreement and Fellowship Agreement which will be based on the Horizon 2020 contractual rules of the European Union and the Spanish fiscal and social security laws. Furthermore, a Career Development Plan will be drafted by the applicant and the Supervisor before the start date of the project in order to ensure that the project's objectives and training activities are defined and delivered within a project timeframe. The plan will describe in detail the researcher's training and career needs, including training on transferable skills, planning for publications and participation in conferences. That will allow the researcher and the Supervisor to monitor the progress being made in these areas, and to take corrective measures in case any delays occurred. The training objectives in the plan will be monitored and updated regularly in meetings undertaken with the supervisors, discussing research progress and identifying potential difficulties and ways of overcoming them.

1.4 Capacity of the researcher to reach or re-enforce a position of professional maturity/independence

The applicant's computer science background and his latest research activities on the field of applied climate science allow him to deal with the complexities behind the inter-disciplinary research at the basis of CAMBIO project. The applicant started his research activities on applied climate science in 2011, before he was working on the research related to optimization and forecasting of buildings' energy consumption. Although he moved to a new research field, since 2012 he started working on two European projects on Climate Services and in one of them (FP7 SPECS) he is coordinating the work package related to pilot applications. Currently he is also the ENEA PI for the ECEM project which aims to contribute to the Copernicus Sectorial Information Systems in the Energy sector. During the last five years the applicant has published more than 12 journal papers and two book chapters both on climate and energy topics.

During the recent years, he took any available opportunity to discuss the main challenges in climate science to engineers and computer scientists and he carried out several training and teaching events in Renewable Energy and Climate Data Science. The applicant has also established fruitful collaborations with private companies and public utilities, for example he supervised the service contract with TERNA (Italian electricity transmission grid operator), organising regular meetings with managers and engineers to discuss the potentialities in including climate information into their operational activities. More details about his research and professional activities can be found later in Section 4.

The CAMBIO project will be a main step for the applicant to strengthen his position as a cross-disciplinary research scientist on climate and data sciences. This project would give him opportunity to work in close contact with top-level scientists in one of the very few European research centres that has both a large computer science and earth science departments. The fellowship would give him the occasion to enlarge his research network including both computer and climate scientists.

2. Impact

2.1 Enhancing the potential and future career prospects of the researcher

This fellowship would allow the applicant to deepen his understanding on the use of climate forecasts and their limits and potentialities, especially considering ensemble forecasting. The collaboration with BSC and Dr. Jorba

Casellas would help him to gain a wide view of forecasting and at the same time to promote an interesting discussion about the links between climate science and data mining. Moreover, given that the BSC is one of the most important European centres on HPC and Big Data, the applicant will have the possibility to establish a knowledge exchange between the climate science and the data science communities, with the possibility of a continuous interaction with experts on data analytics and distributed computing (for example, the Barcelona Spark Meetup is organized by BSC scientists and it has more 1400 members) as well as leading experts in climate forecasts and earth system models.

The applicant will also benefit of the BSC personalised training programmes on technology transfer, technical writing, leadership and communication skills. Due to the multidisciplinary environment where the applicant will have the opportunity to work during this project, he will gain new skills among which climate ensembles' development, and scientific data visualisation, etc. Moreover, moving in a place with a very high cultural diversity would also have a positive impact on the applicant's way of thinking, with the possibility to approach the scientific questions from a variety of perspectives.

The CAMBIO project will be a milestone in the researcher's career and thanks to the research-through-training objectives, the fellow will have an opportunity to grow as a researcher and develop his leadership and entrepreneurial skills. Needless to say, that the project would help the applicant to consolidate his research career laying the groundwork for research grants and subsequently, will build a cross-disciplinary research group on the same topics as project CAMBIO. After completing the fellowship, the researcher will be in a good position to continue and develop further his research ideas by applying for European Research Council (ERC) Starting Grant and establish his independent research group.

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

CAMBIO project has a strong connection with the EU research strategy on climate that can be summarised in three points: 1) it will demonstrate another approach to generate economic value from climate data and, more in general, from the research funded by H2020 projects on climate; 2) the project will contribute to enhance the quality of climate services improving the quality of climate information (this is a direct link to the Challenge 3.3 of the European R&I Roadmap on Climate Services); 3) the improved decision-making here analysed can be applied to the main challenges regarding the ambitious goals of the Europe 2020 Strategy on Climate & Energy (e.g. improved management of renewable energy).

It also worth mentioning that in the long time perspective, the project might be profitable for the EU society due to the improved quality of the decisions and the consequent reduction of climate-related risks in many sectors like agriculture (e.g. improved skill in predicting extreme events) and tourism (e.g. better management of touristic seasons). The capability of use in a better way the available climate information for the incoming months naturally leads to a positive economic impact in two ways: 1) Reduced risk means minor losses derived by adverse climate events; 2) an improved decision-making will provide a competitive advantage for all the companies working in the sectors where climate plays a critical role.

Dissemination activities of CAMBIO will have the first objective to show the potentialities in using the latest advances in climate and information science for decision-making in various sectors. At the same time the communication activities described below will engage scientists and promote inter-disciplinary discussions on the most important research topics.

The applicant will organize at least three seminars are planned for the project: two seminars, hosted at BSC, will be held from the applicant before the start of WP2 (M10, at the end of WP1) and at the end of the project to summarise the outcomes and outline the research questions still open. Moreover, there will be a webinar at M6 to describe the goal of the project and discuss the first steps to start to engage developers and scientists.

Scientific results will be disseminated in two ways: by scientific publishing and conferences, and by using the Internet and the social media (the Communication Team at BSC will be able to provide further support and advice to the researcher on the dissemination activities which might be taken during the project).

While the first way is the traditional path to consolidate the academic reputation and to disseminate science among specialists and researchers, the World Wide Web and the social media will allow the applicant to reach directly the non-specialists and to "tune" the dissemination activities for each different target. Discussion about the main topics of the projects and its outcomes will be also proposed on scientific collaborative blogs (e.g. EGU Blogs, in particular "Earth and Space Science Informatics" blog) and on the applicant's personal research blog.

The publication of a scientific paper is planned and it is included as project deliverable (D3.1). Findings will also be presented at conferences chosen among the most related to Climate Science and Machine Learning, such as

European Geosciences Union (EGU) General Assembly (estimated: April 2018 and April 2019), American Geophysical Union (AGU) Fall Meeting (estimated: December 2018) and the International Workshop on Climate Informatics (estimated September 2018)

The outcomes of the project, in particular the implications in the use of climate ensembles for practical purposes, will be presented during a MedCOF (WMO Mediterranean RCOF) event and during the annual meetings of the Climate Services Partnership. Finally, the applicant will participate at the Spanish Supercomputing Network (RES) events.All the scientific results will be published openly (green open access), the scientific publications will be submitted to OpenAire and open data will be implemented whenever possible. Furthermore, all the tools and packages developed during this project will be made available open-source by hosting them directly on the project web page or via public repositories (e.g. CRAN or GitHub).

Exploitation of results and intellectual property

The main focus of CAMBIO project is on proposing an innovative method to process and analyse available climate data, taking advantage of all the developed technology and the research of the last years on data science and machine learning. Furthermore, one third of the project (WP3, from M17 to M24) will be dedicated to demonstrate how this innovative framework would impact the decisions made in many sectors addressing not only the question "What will be the benefits derived by this project?" but also "How much benefit you could get" (more details are available in Section 3.1). The source code developed during the project will be available open source (under license Apache 2.0), hosted on a public repository (e.g. GitHub). The Apache 2.0 license allows the users, with the proper attribution, to modify and use the code for personal or commercial purposes. In this way the applicant would maximise the potential impact of the project both on the research and commercial sides, fostering further collaborative research and, at the same time, making possible to companies to integrate the source code into their products.Before the end on the project, a specific website for the project CAMBIO, hosted on BSC website or on applicant's personal webpage, will be created to provide a quick access to the source code, tutorials and documentation and, most important, keep track of the applications based on CAMBIO source code.

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

Besides fostering important scientific collaborations among different scientific communities, CAMBIO project represents an important step forward of the European scientific community in the Climate Services with a strong link with Copernicus Climate Change Service and GFCS. Project goals will make more clear outside the academic sector what would be the main advantages in using the state-of-art of climate information to make real-world effective decisions. In fact, except for part of the scientific climate community, there is a low awareness on what the climate forecasts are and what are their limitations and potentialities for helping the society to deal with climate variability. This is the reason why project CAMBIO gives importance to the communication both with the scientific community and the general public.

The BSC, as the host institution of the project, offers very good opportunities to the applicant to disseminate the project results and reach the General Public. At least two news are intended to be delivered, one at the beginning of the project as a press release, and another one at the project end showing its outcomes.

Public engagement is of critical importance for CAMBIO project: considering the difficulties in handling the communication between two different scientific communities (each of them with its vocabulary and rules) the applicant will also take the occasion to disseminate the results making them accessible to non-specialists, to this end the applicant will also organise a Reddit scientific AMA to engage non-specialists in conversation about the project underpinning science. The communication activities will also have the goal to improving the public's understanding of science at the basis of climate forecasts. Fact sheets and illustrative videos for the general public will be prepared at the end of the activities of each work package.

The applicant will also participate in the activities for the European Researchers' Night that will be organised in Spain with the aim to explain what this project can really do for society and to promote research careers to young people.Communication targeted to specialists (e.g. climate scientists) will be implemented through: 1) Talks at scientific conferences (see Section 2.2); 2) Scientific collaborative blogs; 3) Twitter (an account dedicated to the project will be created to communicate projects' updates and interact with other scientists); 4) Illustrative video describing the projects and its scientific outcomes. All the communication and dissemination activities will be coordinated by WP4 (see Section 3).

The following table summarises the dissemination and communication activities.

What	How	Target	When
Updates about main projects steps	Website and social media	Specialists and non-specialists	Start and ending of new activities, main events
Media (figures, interactive charts)	Website and social media	Non-specialists	End of each Work Package
Deliverables and scientific reports	Website	Specialists	When available
Discussions about main research issues	Twitter, collaborative blogs	Specialists and non-specialists	-
Q&A session	Reddit scientific AMA	Non-specialists	End of the project M24
Webinar on ML tools and R packages for climate scientists	Webinar platforms (hosted by BSC)	Students	M6
Illustrative videos	Website and social media	Specialists and non-specialists	After deliverable D2.2 (M16)

# 3. Quality And Efficiency Of The Implementation

3.1   Coherence and effectiveness of the work plan

CAMBIO project is structured in four main work packages as summarized in the following Table:

WP Number	WP Title	Start month	End month
WP1	Development of a Large-Scale Climate Data Processing Layer	1	10
WP2	Optimization of Climate Data Information	11	16
WP3	Improving the decision-making	17	24
WP4	Dissemination and outreach	5	24

Two seminars (M10 and M24) will be held at the host institution by the applicant to discuss and disseminate CAMBIO outcomes. Moreover, a webinar, targeted to developers and climate data scientists, will be given at M6 to present the actual development of the data processing platform.

All the workplan is summarized in the Gantt diagram shown below.
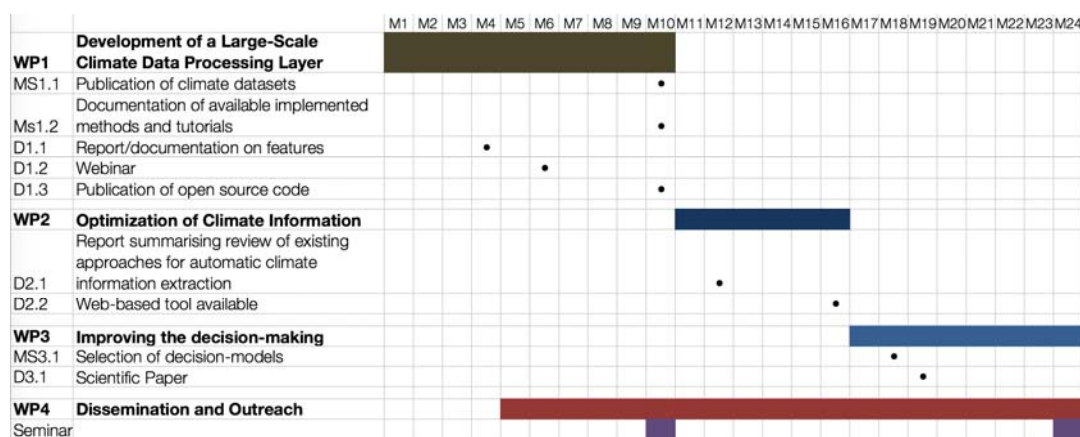
3.1.1   WP DESCRIPTION

WP1 – Development of a Large-Scale Climate Data Processing Layer

This work package has two objectives:
1. Implementation of a software framework to manage and access in an easy and efficient way large climate datasets
2. Development of a an application programming interface (API) to perform a set of data mining and processing methods to large-scale climate datasets in an efficient way

Task 1.1: Implement a scalable platform based on an existing distributed computing infrastructure to process large climate datasets fast and easily. The platform will be able to access the most common science data file formats such as NetCDF. A selection of climate observations, reanalyses and forecasts will be made available in the developed platform to be used in WP2 and WP3.

Task 1.2: The algorithms that will be used for WP2 and WP3 will be made available through an API in order to carry out data mining operations on the available datasets. Where possible, well-tested and optimised algorithms already implemented (e.g. in CRAN R packages and Python Scikit) will be preferred. Furthermore, also a set of deterministic and probabilistic measures will be available to evaluate the quality.

Milestones & Deliverables

- MS1.1: Publication of climate datasets (M10)
- MS1.2: Documentation of available implemented methods and tutorials (M10)
- D1.1: Report/documentation on features (M4)
- D1.2: Webinar (M6)
- D1.3: Publication of open source code (M10)

WP2 – Optimization of Climate Data Information

The primary objectives within this work package is a systematic analysis on the best way to extract useful information from a set of heterogeneous climate datasets to predict meteorological variables on a specific spatial/temporal domain.

Task 2.1: Review of existing approaches in scientific literature where data mining techniques are used to maximise the information extraction from a set of climate data.

Task 2.2: Systematic analysis of the maximum information extraction for the most used meteorological variables (i.e. 2-m temperature, wind speed, precipitation, solar radiation, geopotential height) at sub-continental level. For each variable an automatic discovery of the optimal set of predictors able to estimate/predict it for each geographical domain will be carried out. Both deterministic and probabilistic approaches will be applied, using the metrics and algorithms implemented in WP1. Given that a large amount of data is expected to be produced during this step, a web-based interactive tool to explore the results will be developed with the support of the BSC Visualization & Post-process group.

Milestones & Deliverables

- D2.1: Report summarising review of existing approaches for automatic climate information extraction (M12)
- D2.2: Development of a web application with the possibility to explore the optimal predictors according to different measures (deterministic or probabilistic), spatial and temporal domains, meteorological variables (M16)

WP3 – Improving the decision-making

This work package aims to evaluate the added-value of the results obtained in WP2 from an user perspective. A set of case-studies of climate-driven decision-making in the important sectors will be obtained based on decision-models that are already available in scientific works and past/ongoing experiences on Climate Services. The sectors will be chosen according to the GFCS priority areas. The capability of the enhanced information of contributing to the decision-making will be assessed.

Task 3.1: Review of existing and available decision-models that can be used to evaluate the added value of the enhanced information. The selection of at least two decision-making models will be made according to several variables: 1) the possibility to provide a realistic assessment of the added-value; 2) the possibility to ask for a feedback from experts and/or stakeholders; 3) easiness of implementation of the decision-model.

Task 3.2: The added value of the enhanced information provided will be assessed using the selected decision-making models driven both by observed and refined information climate data.

Milestones & Deliverables

- MS3.1: Review about the selected decision-models (M18)
- D3.1: Scientific paper on the advantages in combining different sources of information through data mining techniques to decision-making (M19)

WP4 – Dissemination and outreach

The primary objective of this Work Package is to disseminate and promote the research activities of CAMBIO project. The dissemination activities are described more in detail in Section 2.2.1 and 2.2.2

Task 4.1: Preparation of a public website (hosted on the applicant's research domain) where all the main steps of CAMBIO are described providing links to the web application developed (D2.2) and to the open-source code implemented and used (see D1.3)

Task 4.2: Design of training material (tutorials and webinars) on the use of machine learning methods for climate scientists.

3.2 Appropriateness of the allocation of tasks and resources

The project has a strong focus on large-scale datasets and thus storage and computing resources must be accurately managed. At the beginning of the project, a meeting with the Computer Science department will be held to discuss the roadmap discussing any potential issue related to storage and computing. Furthermore, throughout the project the applicant will keep the Grid Computing and Clusters group updated about the project workplan in order to prevent any bottleneck. The 24-month length of the project is consistent with its objectives, with the first half more centred on the development of the climate information processing layer and the second part instead revolving around the enhancement of those information and their use for a set of specific real-world case-studies. The applicant will take advantage of existing open source software in order to focus his work on the most innovative part of the project development (e.g. interface with climate data and integration with machine learning algorithms)

The progress of the project will be checked during the periodic meetings with the supervisor.

3.3 Appropriateness of the management structure and procedures

The project will be managed through bi-weekly meetings to ensure full coherence between the applicant's research and the general objectives of the research group. The applicant will work in close contact with the supervisor of this project and present progress and results to the scientific staff on a regular basis, through reports and presentations. This will allow the evaluation of his progress against the work plan, and seek assistance in a timely and periodic manner if the necessity arises. Regular meetings will take place involving the rest of BSC members to ensure an adequate integration of this activity into the rest of the research carried out in the BSC. A Project Manager will support the researcher in all the legal, financial and administrative arrangements needed and work in close contact with Education, Human Resources and Communications departments for all training arrangements and dissemination and outreach activities. Finances of the project will be managed in accordance with Marie Skłodowska-Curie funding rules signed at the grant agreement and will follow the already established processes in the centre.

The Barcelona Supercomputing Center (BSC) is dedicated to providing high-quality postdoctoral training to future scientific leaders. It draws upon its experience in developing specific technical and scientific skills, as well as the complementary skills required for efficient research execution and communication. BSC has a strong support for the management and correct development of the fellowship, including 1) Research Project Manager who will be responsible for the Grant Agreement Negotiation process in case the project has been successfully selected for funding. In addition, the Project Manager will be a contact point for the fellow during the entire duration of the grant and will ensure that all necessary documents are provided and duly signed (Fellowship Agreement, Grant Agreement, Career Development Plan). As a contact point of the Beneficiary, the PM will ensure that all report are submitted within a timeframe indicated by the Research Executive Agency; 2) Communication team that will help the fellow to use the communication channels that hold most impact; 3) Finances department for using the available resources in an optimal way and being responsible for the financial report (Form C) at the end of the project; 4) Operations team, for any support the fellow might need regarding supercomputing and software, and 5) Human Resources – will ensure that the Employment Contract is provided and signed by the Host Institution and the Fellow. Additionally, a complex introduction and information about the current training activities will be provided to the fellow. In addition to the above, the Career Development Plan will be drafted and approved by both – the fellow and the supervisor, including additional training and activities proposed by the Human Resources and education department.

Risk Management

The approach for managing risks for the CAMBIO project includes a methodical process by which the applicant identifies, scores, and ranks the various risks. The most likely and highest impact risks are shown in the following table (values for Likelihood and Impact are in the range 1-4):

Description	Likelihood	Impact	Contingency and Mitigation Actions
Delay in the deployment and test of the developed data processing framework on the BSC distributed computing facilities	1	3	Deployment of the framework on a single workstation and reduction of the size of the datasets (e.g. change resolution or reduction of ensemble members) to deal with the reduced computing power
Higher needs of computing time than expected during WP2	1	3	Reduce the number of variables and/or domains considered focusing only on the most relevant for WP3

3.4   Appropriateness of the institutional environment (infrastructure)

The Barcelona Supercomputing Center — Centro Nacional de Supercomputación (BSC-CNS) is the National Supercomputing Facility of the Spain. BSC-CNS manages MareNostrum III with 48,128 cores and 1.1 Pflops capacity, one of the most powerful supercomputers in Europe. The mission of BSC-CNS is to research, develop and manage information technology in order to facilitate scientific and technological progress. At the BSC-CNS, more than 350 researchers and students from more than 40 different countries perform research in Computer Sciences, Life Sciences, Earth Sciences and Computational Applications in Science and Engineering. This multi-disciplinary approach and the combination of world-leading researchers and HPC experts with state-of-the-art HPC resources make BSC-CNS a unique research institution and the best place to host CAMBIO project with its cross-disciplinary and data-centered approach. The involved research groups will be the Atmospheric Composition (coordinated by the supervisor Dr. Oriol Jorba), Earth System Services and Computational Earth Sciences from the Earth Sciences Department (ESD), Grid Computing and Clusters from Computer Sciences Department (CSD) and the Data Visualization group of the CASE department. The applicant will collaborate with many scientists at BSC: Dr. Oriol Jorba (ESD) that will supervise the project during the 24 months of its duration with regular meetings, Kim Serradell (ESD) and Maria Rosa Badia (CSD), Fernando Cucchietti (CASE).

DOCUMENT 2

4. CV of the Experienced Researcher.

Summary of Achievements

The applicant is staff scientist since 2011 in the ENEA Climate Modelling & Impacts laboratory. He has published 16 journal papers (two currently under review), 3 book chapters and 13 peer-reviewed conference papers. He has been involved in 5 European projects as researcher and WP leader and he is the ENEA PI for a Copernicus Climate Change Service project ECEM. The applicant was also the supervisor of the service contract for the Dispatching Office of TERNA S.p.A. (Italian national electricity grid operator).

Background

The applicant is a Computer and Automation Engineer with a strong background in machine learning and statistical modelling. Since his master degree, he has focused his research on the application of machine learning and modelling to energy-related problems in an inter-disciplinary environment collaborating with civil engineers, architects and climate scientists. The applicant got his Ph.D. European Label in Computer Science and Automation at the University of Rome "Roma Tre" with a thesis titled "Application of Computational Intelligence to Energy Systems". During his formation, the applicant attended several international conferences and training courses.

Up to now, he has carried out his research activity at ENEA (National Agency for New Technologies, Energy and Sustainable Economic Development) at the **Energy Efficiency Unit** and, since 2011, he is staff scientist at the **Climate Impacts and Modelling (CLIM)** laboratory. He collaborates with the Department of Informatics and Automation of the University of Rome "Roma Tre" and with other research institutes in Italy (University of Milan, EURAC Academy) and abroad (University of Birmingham, University of Cantabria).

During his professional activity, the applicant has been involved in various national and European projects and currently he is involved in four EU FP7/H2020 projects: SPECS (Seasonal-to-decadal climate Prediction for the improvement of European Climate Services) where he is the WP leader of the Work Package on Pilot Applications, EUPORIAS (EUropean Provision Of Regional Impact Assessment on a Seasonal-to-decadal timescale) where he is working with World Food Program and other partners in developing an operational Climate Service, SINGULAR (Smart and Sustainable Insular Electricity Grids Under Large-Scale Renewable Integration) and CRESCENDO.

The applicant is also the PI for ENEA for two European projects: 1) Copernicus Climate Change (C3S) ECEM project; 2) Interreg V-A Italy-France (Maritime) STRATUS

Education

2007-2011: Ph.D. European Label at the Department of Informatics and Automation (DIA), University of Rome "Roma Tre", in collaboration with **ENEA**. Thesis title: *Application of Computational Intelligence to Energy Systems*

2004-2007: Laurea Magistrale (Msc) in Informatics and Automation Engineering, University of Rome "Roma Tre", in collaboration with **ENEA**. Thesis topic: *Evolutionary Neural Networks with Complex Network Topology for Dynamic Systems Modelling*. Grade: 110/110

2001-2004: Laurea Primo Livello (BSc) in Informatics Engineering, University of Rome "Roma Tre"**,** in collaboration with **Alenia Space**. Thesis topic: *Porting of simulation models from XMATH to MATLAB*. Grade: 110/110 Cum Laude.

Work Experience

2011-present: Staff Scientist at ENEA Climate Impacts and Modelling (CLIM) Laboratory. Research activity focused on Climate Services and Energy and Meteorology. Organization committee of two EUPORIAS Masterclass on Climate Services. Supervisor of the service contract for TERNA S.p.A. (Italian Transmission Grid Operator) for 2012-2015. ENEA PI for ECEM project (2015-2016). Involved in 5 FP7 and H2020 projects and in the ENEA activity for the Italian Ministry of the Environment.

2007-2010: Research Assistant at ENEA, Energy Efficiency Department and Department of Informatics and Automation (DIA), University of Rome "Roma Tre". Research activity on computational intelligence methodologies for energy process, development of optimization systems for renewable energy management, and control methods for building-energy plants.

Involvement in Research Projects

1) Interreg V-A Italy-France STRATUS, 2016-2018. **Role**: PI for ENEA, involved in the design of the project and in the implementation of a climate service for tourism sector in Italy and France.
2) Copernicus Climate Change Programme - ECEM (*European Climatic Energy Mixes*), 2015-2016. **Budget**: ENEA 66k€. **Role**: PI for ENEA, involved in the application of seasonal climate forecasts for the energy sector.
3) H2020 - CRESCENDO (*Coordinated Research in Earth Systems and Climate: Experiments, kNowledge, Dissemination and Outreach)*: **Budget**: 14.3 M€ (ENEA 259k€). **Role**: researcher involved in climate data analysis
4) FP7 - SINGULAR (Smart and Sustainable Insular Electricity Grids Under Large-Scale Renewable Integration), 2013-2016: **Budget**: 3.6 M€ (ENEA 87 k€). **Role:** Researcher involved in WP2 on the development of Renewable Energies forecasting tools
5) FP7 - EUPORIAS (European Provision Of Regional Impact Assessment on a Seasonal-to-decadal timescale) 2012-2016: **Budget**: 8.9 M€ (ENEA 371 K€). **Role:** researcher involved in tasks 12.2-12.4 on the assessment of stakeholders' needs on Climate Services and in the development in an operational Climate Services in collaboration with the World Food Program, Ethiopian Government and other European research institutions. Also in the organising committee of the two EUPORIAS Masterclass on Climate Services held in 2015 and 2016.
6) FP7 - SPECS (Seasonal-to-decadal climate Prediction for the improvement of European Climate Services), 2013-2016: **Budget**: 8.2 M€ (ENEA 401 k€). **Role:** WP Leader for the Pilot Applications (use of the climate forecast for solar and wind power production); involved in seasonal forecasts verification and on the analysis of multi-model ensembles.
7) FP7 - CLIM-RUN (Climate Local Information in the Mediterranean region Responding to User Needs), 2011-2014: Project coordinated by ENEA with the objective of contributing in the creation of a Climate Services Network in the Mediterranean area. **Budget**: 3.5 M€ (ENEA 401 k€). **Role:** researcher, involved in the creation of the Case Studies portal.

Research Visits

6-9/2009, 7-10/2010, 3/2012: Visiting Researcher at CERCIA (The Centre of Excellence for Research in Computational Intelligence and Applications), University of Birmingham, UK. Research activity on forecasting of electricity demand on different scales (building, regional, national). Supervisor: Prof. Xin Yao. The visits have been funded by: EU IntelliCIS COST Action IC0806, ENEA and CERCIA.

Invited Talks and Seminars

- Severo Ochoa Excellence Programme seminar lecture at BSC
- 2nd CLIM-RUN School on Climate Services, Trieste, Italy: two days lessons on machine learning and data mining
- INTELLICIS COST Action annual meeting, Aachen, Germany: invited talk on ENEA research activities

Education and Mentoring

- Teaching for Fuzzy Control course at Department of Informatics and Automation (DIA), University of Rome "Roma Tre" (2011/12, 2012/13 and 2013/14) and tutoring of students' projects
- Outside examiner for many (>10) master theses at DIA, University of Rome "Roma Tre", tutoring for two MSc students at DIA and for a student from University Politehnica of Bucharest.

Training Schools & Courses

- 31/08-4/09/2007: Summer School in Evolutionary Computation and Artificial Life (SECEVITA), Ragusa, Italy
- 20-27/09/2008: International Summer School of Natural Computation (BNC), Bertinoro, Italy
- 8-10/02/2011: HPC Course on GPU Programming, CASPUR, Italy
- 28-30/03/2011: ECMWF Training in Course in Numerical methods, adiabatic formulation of models and ocean wave forecasting

Organisation of Research Activities

- Referee activities for many scientific journals (including Renewable Energy (*Elsevier*), Applied Soft Computing Journal (*Elsevier*), Applied Energy (*Elsevier*), Energy Conversion & Management (*Elsevier*), Journal of Computer Science and Technology (*Springer*)). The full list of verified reviews is available on Publons at https://publons.com/a/590063/
- Reviewing activities ad Program Committee for various international conferences (PPSN, ICANNGA, WCCI, ICANNGA, ICARIS, SSCI)
- Involvement in the creation of a Research Data Alliance Interest Group about "Weather, Climate and Air Quality"

Membership of scientific societies

05/2009 – 12/2015: member of ACM (Association for Computing Machinery)

Other Professional Activities

- Consultancy services for teaching Java and C Programming for private companies (96 hours) and courses funded by Lazio Regional Authority (40 hours)

Publications: Journals (citations obtained from Scopus)

J1. A. Alessandri, F. Catalano, M. De Felice, B. Van Der Hurk, S. Boussetta, G. Balsamo, and P. Miller, "Multi-scale enhancement of climate prediction over land by increasing the model sensitivity to vegetation variability in EC-Earth," *Climate Dynamics*, accepted.

J2. S. Pezzutto, A. Toleikyte, and M. De Felice, "Assessment of the space heating and cooling market in the EU28: a comparison between EU15 and EU13 member states," *International Journal of Contemporary Energy*, Accepted.

J3. F. Catalano, A. Alessandri, M. De Felice, Z. Zhu, and R. B. Myneni, "Observationally based analysis of land-atmosphere coupling," *Earth System Dynamics*, vol. 7, iss. 1, pp. 251-266, 2016.

J4. M. Pierro, F. Bucci, M. De Felice, E. Maggioni, D. Moser, A. Perotto, F. Spada, and C. Cornaro, "Multi-Model Ensemble for day ahead prediction of photovoltaic generation," *Solar Energy*, vol. 134, pp. 132-146, 2016.

J5. S. Pezzutto, R. Fazeli, M. De Felice, and W. Sparber, "Future development of the air-conditioning market in Europe: an outlook until 2020," *Wiley Interdisciplinary Reviews: Energy and Environment*, 2016.

J6. Castelli M., Vanneschi L., De Felice M. (2015) "Forecasting short-term electricity consumption using a semantics-based genetic programming framework: The South Italy case," *Energy Economics*, 47, 37-41. (2 citations)

J7. Iacono, R. and De Felice, M. (2015) "Constructing analytic approximate solutions to the Lane-Emden equation," *Physics Letters A*, 379, 32-33, 1802-1807.

J8. De Felice M., Petitta M., and Ruti, P.M. (2015) "Short-term predictability of photovoltaic production over Italy," *Renewable Energy*, 80, 197-204 (3 citations)

J9. De Felice M., Alessandri A., and Catalano, F. (2015) "Seasonal climate forecasts for medium-term electricity demand forecasting," *Applied Energy*, 137, 435-444 (6 citations)

J10. Alessandri A., De Felice M., Zeng N., Mariotti A., Pan Y., Cherchi A., Lee J., Wang B., Ha K., Ruti P.M., and others (2014) "Robust assessment of the expansion and retreat of Mediterranean climate in the 21st century," *Nature Scientific Reports*, vol. 4. (6 citations)

J11. Iacono, R., & De Felice, M. (2014). Approximate analytic solutions to the isothermal Lane–Emden equation. Celestial Mechanics and Dynamical Astronomy, 118(3), 291-298 (2 citations)

J12. De Felice, M., Alessandri, A., and Ruti, P.M.. (2013). "Electricity Demand Forecasting over Italy: Potential Benefits using Numerical Weather Prediction models," *Electric Power Systems Research*, 104, 71-79. (10 citations)

J13. De Felice, M., & Yao, X. (2011). Short-Term Load Forecasting with Neural Network Ensembles: A Comparative Study [Application Notes]. *Computational Intelligence Magazine, IEEE*, 6(3), 47-56. (36 citations)

J14. Bertini, I., <u>De Felice, M.,</u> Pannicelli, A., & Pizzuti, S. (2011). Soft computing based optimization of combined cycled power plant start-up operation with fitness approximation methods. Applied Soft Computing, 11(6), 4110-4116. (8 citations)

J15. Bertini, I., <u>De Felice, M.,</u> Pannicelli, A., & Pizzuti, S. (2011). Optimizing the start-up operations of combined cycle power plants using soft computing methods. Logic Journal of IGPL, 20(4), 648-656. (2 citations)

J16. Bertini, I., Ceravolo, F., Citterio, M., <u>De Felice, M.,</u> Di Pietra, B., Margiotta, F., ... & Puglisi, G. (2010). Ambient temperature modelling with soft computing techniques. *Solar Energy*, *84*(7), 1264-1272. (8 citations)

Publications: Book Chapters

B1. Ruti, P.M. and <u>De Felice, M.</u> (2013). Climate and Energy Production — A Climate Services Perspective. Elsevier Inc., Academic Press, pp. 117-121.

B2. Athanasiou, A., <u>De Felice, M.,</u> Oliveto, G., & Oliveto, P. S. (2013). Dynamical Modeling and Parameter Identification of Seismic Isolation Systems by Evolution Strategies. In Computational Intelligence (pp. 101-118). Springer Berlin Heidelberg. (2 citations)

B3. Azzini, A., <u>De Felice, M.,</u> & Tettamanzi, A. G. (2012). A Comparison between Nature-Inspired and Machine Learning Approaches to Detecting Trend Reversals in Financial Time Series. In Natural Computing in Computational Finance (pp. 39-59). Springer Berlin Heidelberg.

Publication: Peer-Reviewed Conference Papers (with acceptance rate when available)

C1. Pezzutto S., Toleikyte A., and <u>De Felice M.</u> (2015) "Assessment of the space heating and cooling market in the EU28: a comparison between EU15 and EU13 member states," in REMOO 2015: Technological, Modelling and Experimental Achievements in Energy Generation Systems.

C2. <u>De Felice, M.,</u> Meloni, S., & Panzieri, S. (2011, July). Effect of topology on diversity of spatially-structured evolutionary algorithms. In Proceedings of the 13th annual conference on Genetic and evolutionary computation (pp. 1579-1586). ACM. (38% acceptance rate) (2 citations)

C3. <u>De Felice, M.,</u> & Yao, X. (2011, April). Neural networks ensembles for short-term load forecasting. In Computational Intelligence Applications In Smart Grid (CIASG), 2011 IEEE Symposium on (pp. 1-8). IEEE.

C4. Athanasiou, A., <u>De Felice, M.,</u> Oliveto, G., & Oliveto, P. S. (2011). Evolutionary algorithms for the identification of structural systems in earthquake engineering. In International Conference on Evolutionary Computation Theory and Applications (pp. 52-62). (13% acceptance rate) (3 citations)

C5. Stracquadanio, G., La Ferla, A., <u>De Felice, M.,</u> & Nicosia, G. (2011). Design of robust space trajectories. In Research and Development in Intelligent Systems XXVIII (pp. 341-354). Springer London. (3 citations)

C6. Bertini, I. , <u>De Felice, M.,</u> and Pizzuti, S. (2011). Combining Back-Propagation and Genetic Algorithms to Train Neural Networks for Start-Up Time Modelling in Combined Cycle Power Plants. In ESANN 2010 proceedings, European Symposium on Artificial Neural Networks – Computational Intelligence and Machine Learning, Bruges (Belgium).

C7. Bertini, I., <u>De Felice, M.,</u> Moretti, F., & Pizzuti, S. (2010). Start-up optimisation of a combined cycle power plant with multiobjective evolutionary algorithms. In Applications of Evolutionary Computation (pp. 151-160). Springer Berlin Heidelberg.

C8. Azzini, A., <u>De Felice, M.,</u> & Tettamanzi, A. G. (2010). A study of nature-inspired methods for financial trend reversal detection. In Applications of Evolutionary Computation (pp. 161-170). Springer Berlin Heidelberg. (1 citation)

C9. Azzini, A., <u>De Felice, M.,</u> Meloni, S., & Tettamanzi, A. G. (2009). Soft computing techniques for Internet backbone traffic anomaly detection. In Applications of Evolutionary Computing (pp. 99-104). Springer Berlin Heidelberg.

C10. Ceravolo, F., <u>De Felice, M.,</u> & Pizzuti, S. (2009). Combining back-propagation and genetic algorithms to train neural networks for ambient temperature modeling in Italy. In Applications of Evolutionary Computing (pp. 123-131). Springer Berlin Heidelberg.

C11. Ceravolo, F., <u>De Felice, M.</u>, & Pizzuti, S. (2008). Ambient Temperature Modelling through Traditional and Soft Computing Methods. In Hybrid Artificial Intelligence Systems (pp. 322-328). Springer Berlin Heidelberg.

5. Capacity of the Participating Organisations

Beneficiary: Barcelona Supercomputing Center	
General Description	The Barcelona Supercomputing Center (BSC) was established in 2005 and serves as the national supercomputing facility. Currently, it is hosting 1 of the 6 European Tier-0 supercomputers and is among the best supercomputing centres in the world. Its mission is to research, develop and manage information technologies in order to facilitate scientific progress. In terms of attraction of talent, during the period 2011-2015, the BSC has recruited 75 pre-doctoral students, 51 Postdocs and Senior Scientist, 83 technical support staff members and 31 management staff, being currently more than 380 staff members, from around 40 countries. The centre has been awarded with the badge of Human Resources Excellence in Research (HRS4R) in April 2015. The applicant will be enrolled at the Department of **Earth Sciences (ES)**.
Role and Commitment of key persons (supervisor)	Dr. Oriol Jorba is an expert in the development of air quality modeling systems to provide air quality forecasts and perform advanced research on atmospheric chemistry. He is the Head of the Atmospheric Composition group at the BSC-ES. He has been involved in the deployment of the CALIOPE air quality forecasting system and currently he is coordinating the development of the NMMB/BSC-CTM. With his experience on model development, he will supervise the Fellow's work during the entire duration of the project by providing guidance on advanced and efficient use of numerical models in high performance computing environments.
Key Research Facilities, Infrastructure and Equipment	The BSC manages supercomputing facilities, notably MareNostrum. The current MareNostrum BSC is one of the 5 largest computers in Europe, with more than 1 PFlop. Additionally, BSC manages Minotauro, a Sandy Bridge's cluster with NVIDIA GPUs, providing more than 100 TFlops.
Independent research premises?	All BSC's departments have their own research premises. The infrastructure, equipment and key research facilities will be available for the fellow during the entire duration of the project.
Previous Involvement in Research and Training Programmes	The BSC has a successful record of national and international fellowships: since 2008 BSC hosted several Marie-Curie Individual Fellowhips (e.g. EEPPIBM, FP7-PEOPLE-2012-IEF-327899; MDRAF, FP7-PEOPLE-2013-IEF-622662; Noteworthy, that one of the MSC IF has been awarded under Dr. Oriol Jorba supervision in 2015 (INAQUA, H2020-MSCA-IF-2015-707515)) and three ITNs (SCALUS, FP7-PEOPLE-ITN-2008-238808; NEMOH, PF7-PEOPLE-2011-ITN-289976, COPA-GT, FP7-PEOPLE-ITN-2011-290042). BSC also hosts national fellowships, both for early-stage and senior postdoctoral positions: approximately 30 post-docs were awarded since 2006.
Current involvement in Research and Training Programmes	<u>Collaboration with universities:</u> within BSC, there is a large record of collaboration with Universidad Politècnica de Catalunya (UPC) including the Master degree in Environmental Engineering (UPC), associated with BSC Earth Science department. <u>Excellence Programmes and Networks:</u> A number of training activities are organized under the framework of: Severo Ochoa Excellence Programme (Research seminars series); RES (RES training sessions); NVIDIA CUDA/GPU excellence; PRACE (PRACE Advanced Training Center); and H2020-EINFRA-Centers of Excellence for computing applications (*PoP*, Grant Agreement (GA) number: 676553; *ESiWACE*, GA number: 675191; *BioExcel*: GA number: 675728; *NoMaD*, GA number: 676580; *MaX*, GA number: 676598; *EoCoE*, GA number: 676629). <u>Research Fellowships:</u> BSC is currently awarded with 9 early-stage postdoc, 12 senior (5 Ramón y Cajal, 3 I3 and 6 ICREA) and is supporting 4 ITN, 2 RISE and 4 Marie-Curie Individual Fellowships. Noteworthy, two of these Marie-Curie actions *(DPETNA,* H2020-MSCA-IF-2014-655339; *NeTNPPAO,* H2020-MSCA-IF-2015-708063) are currently developed in BSC's Earth Sciences Department, which will host the present Marie-Curie proposal. <u>Current involvement in Research Programmes:</u> **SPECS (Seasonal-to-decadal climate Prediction for the improvement of European Climate Services) FP7-ENV-2012-308378.** This project is coordinated by the BSC and includes 19 Partners (18 European and 1 Brazilian Institutions).Total requested EU contribution for the project: EUR 8,224,862. The detailed project's objectives and work progress can be found here: http://www.specs-fp7.eu/ **EUPORIAS (European Provision Of Regional Impacts Assessments on Seasonal and Decadal Timescales) FP7-ENV-2012-308291** The BSC is a partner for this project with the total budget EUR 243,681.21.
Relevant Publications and/or research/innovation products	1.Badia; A.; O Jorba; A Voulgarakis; D Dabdub; C Pérez García-Pando; A Hilboll; M Gonçalves; Z Janjic, Gas-phase chemistry in the online multiscale NMMB/BSC Chemical Transport Model: Description and evaluation at global scale, Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-141, in review, 2016.2. 2. Spada; M.; O. Jorba; C. Pérez García-Pando; Z. Janjic; J.M. Baldasano. On the evaluation of global sea-salt aerosol models at coastal/orographic sites. Atmospheric Environment. 101, pp. 41 - 48. 2015. 3. Pandolfi; M.; X Querol; A Alastuey; JL Jimenez; O Jorba; D Day; A Ortega; MJ Cubison; A Comerón; M Sicard; others. Effects of sources and meteorology on particulate matter in the Western Mediterranean Basin: An overview of the DAURE campaign. Journal of Geophysical Research: Atmospheres. 119 - 8, pp. 4978 - 5010. 2014. 4. Badia; A.; O Jorba. Gas-phase evaluation of the online NMMB/BSC-CTM model over Europe for 2010 in the framework of the AQMEII-Phase2 project. Atmospheric Environment, 115, pp. 657-669. 2015. 5. Spada; M.; O Jorba; C Pérez García-Pando; Z Janjic; JM Baldasano. Modeling and evaluation of the global sea-salt aerosol distribution: sensitivity to size-resolved and sea-surface temperature dependent emission schemes. Atmospheric Chemistry and Physics. 13 - 23, pp. 11735 - 11755. 2013.

6. Ethical Issues
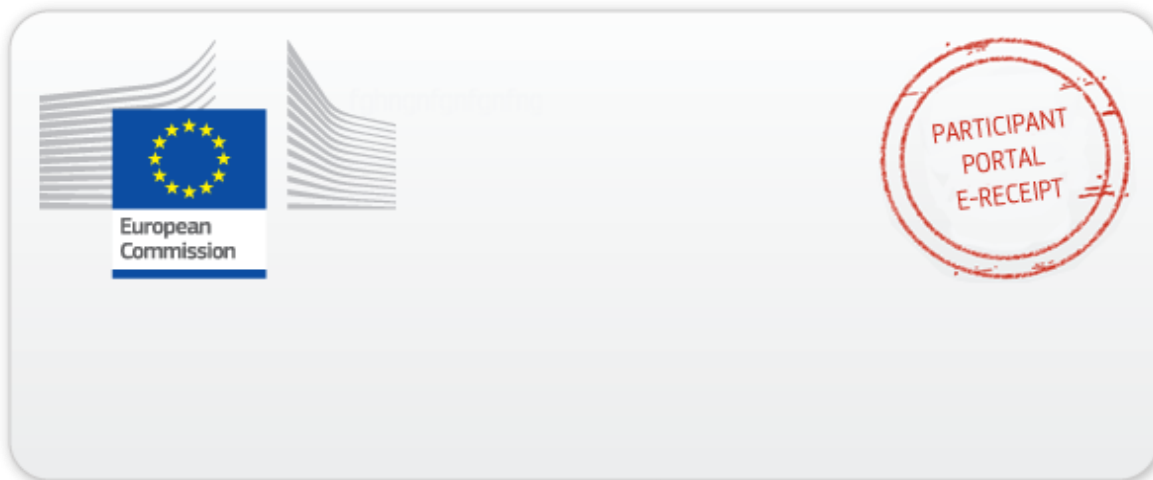
No ethics issues are present in this project.

ENDPAGE

MARIE SKŁODOWSKA-CURIE ACTIONS

Individual Fellowships (IF)
Call: H2020-MSCA-IF-2016

PART B

"CAMBIO"

This proposal is to be evaluated as:

Standard EF

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 Participant 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 Participant Portal. (http://ec.europa.eu/research/participants/portal/page/faq)