



HAL
open science

The EU-FP7 ERA-CLIM2 project contribution to advancing science and production of Earth-system climate reanalyses

Roberto Buizza, Stefan Brönnimann, Leopold Haimberger, Patrick Laloyaux, Matthew J. Martin, Manuel Fuentes, Magdalena Alonso-Balmaseda, Andrea Becker, Michael Blaschek, Per Dahlgren, et al.

► To cite this version:

Roberto Buizza, Stefan Brönnimann, Leopold Haimberger, Patrick Laloyaux, Matthew J. Martin, et al.. The EU-FP7 ERA-CLIM2 project contribution to advancing science and production of Earth-system climate reanalyses. *Bulletin of the American Meteorological Society*, 2018, 99 (5), pp.1003-1014. 10.1175/BAMS-D-17-0199.1 . hal-01661240

HAL Id: hal-01661240

<https://inria.hal.science/hal-01661240>

Submitted on 11 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The EU-FP7 ERA-CLIM2 project contribution

to advancing science and production of Earth-system climate reanalyses

Roberto Buizza (1), Stefan Brönnimann (4), Leopold Haimberger (5), Patrick Laloyaux (1),
Matthew J. Martin (2), Manuel Fuentes (1), Magdalena Alonso-Balmaseda (1), Andreas
Becker (10), Michael Blaschek (5), Per Dahlgren (1), Eric de Boisseson (1), Dick Dee (1),
Marie Doutriaux-Boucher (3), Feng Xiangbo (14), Viju John (2,3), Keith Haines (14), Sylvie
Jourdain (9), Yuki Kosaka (1), Daniel Lea (15), Florian Lemarié (15), Michael Mayer (5),
Palmira Messina (16), Coralie Perruche (8), Philippe Peylin (16), Jounie Pullainen (13), Nick
Rayner (2), Elke Rustemeier (10), Dinand Schepers (1), Roger Saunders (2), Jörg Schulz (3),
Alexander Sterin (7), Sebastian Stichelberger (5), Andrea Storto (12), Charles-Emmanuel
Testut (8), Maria- Antónia Valente (6), Arthur Vidard (15), Nicolas Vuichard (16), Anthony
Weaver (11), James While (2), Markus Ziese (10)

- 1) European Centre for Medium-Range Weather Forecasts (ECMWF; Europe);
- 2) Met Office, UK (UKMO; UK);
- 3) European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT;
Europe);
- 4) University of Bern (Switzerland);
- 5) University of Vienna (Austria);
- 6) Instituto Dom Luiz Faculdade de Ciências da Universidade de Lisboa (Portugal);
- 7) Russian Research Institute of Hydrometeorological Information (RIHMI; Russia);
- 8) Mercator Ocean Societé Civile (MERCATOR; France);

- 24 9) Météo-France (MF; France);
25 10) Deutscher Wetterdienst (DWD; Germany);
26 11) Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique
27 (CERFACS; France);
28 12) Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC; Italy);
29 13) Ilmatieteen Laitos (FMI; Finland);
30 14) The University of Reading (UK);
31 15) Institut National de Recherche en Informatique et en Automatique (INRIA; France);
32 16) Laboratoire des Sciences du Climat et de l'Environnement (LSCE) France.

33

34 *Corresponding author:* Roberto Buizza, European Centre for Medium-Range Weather
35 Forecasts (ECMWF), Shinfield Park, Reading, RG2-9AX, U.K..

36

37 *Submitted to the Bulletin of the American Meteorological Society (BAMS), 7 August 2017.*

38

39 **Capsule Summary**

40 The main goals and activities of the ERA-CLIM2 project are presented, and some of its key
41 results, including the first ensemble of coupled reanalysis of the 20th century, are discussed.

42

43 **Abstract**

44 ERA-CLIM2 is a European Union Seventh Framework Project started in January 2014. It
45 aims to produce coupled reanalyses, which are physically consistent data sets describing the
46 evolution of the global atmosphere, ocean, land-surface, cryosphere and the carbon cycle.
47 ERA-CLIM2 has contributed to advancing the capacity for producing state-of-the-art climate
48 reanalyses that extend back to the early 20th century. It has led to the generation of the first
49 ensemble of coupled ocean, sea-ice, land and atmosphere reanalyses of the 20th century. The
50 project has funded work to rescue and prepare observations, and to advance the data-
51 assimilation systems required to generate operational reanalyses, such as the ones planned by
52 the European Union Copernicus Climate Change Service. This paper summarizes the main
53 goals of the project, discusses some of its main areas of activities, and presents some of its
54 key results.

55

56 1. The ERA-CLIM2 project

57 ERA-CLIM2 (European Reanalysis of Global Climate Observations 2) is a 4-year research
58 project funded by the European Union Seventh Framework Program (EU FP7; Grant
59 Agreement No. 607029; see Appendix for a list of the Consortium's Institutions), that started
60 on the 1st of January 2014 following the successful completion of its predecessor project
61 ERA-CLIM. The research initiated in these two projects underpins a concerted effort in
62 Europe to build the information infrastructure needed to support climate monitoring, climate
63 research and climate services, based on the best available science and observations. ERA-
64 CLIM2 is one of several collaborative research projects designated by the European
65 Commission as precursors to the EU's Copernicus Climate Change Service (C3S). Indeed,
66 ERA-CLIM2 activities on data rescue, satellite data rescue and reprocessing, coupled data
67 assimilation and reanalysis production have had a strong impact on the design and
68 implementation of the C3S. Furthermore, activities performed in ERA-CLIM2 and aiming at
69 improving the assimilation of observations at inter-medium boundaries, i.e. sea surface
70 temperature and sea-ice concentration observations, will certainly benefit ocean and sea-ice
71 reanalyses and operational marine forecast activities performed within the EU's Copernicus
72 Marine Environment Monitoring Service (CMEMS).

73 The aim of the ERA-CLIM2 project is to enable production of state-of-the-art reanalyses of
74 the coupled climate system for the 20th century, following earlier work on coupled and long-
75 term climate reanalysis at the National Center for Atmospheric Research and the National
76 Centers for Environmental Predictions (NCAR and NCEP; *Saha et al.*, 2010, *Compo et al.*,
77 2011), and the European Centre for Medium-Range Weather Forecasts (ECMWF) strategy
78 described in a previous BAMS article (*Dee et al.*, 2014). Climate reanalyses are physically
79 consistent data sets derived from observations that document the recent evolution of the
80 global atmosphere, ocean, land-surface, cryosphere and the carbon cycle. Reanalysis data are

81 generated by a sequential process called data assimilation, which combines first-guess
82 estimates defined by short model forecasts with vast amounts of data from a range of
83 observing platforms (surface, upper-air, satellites). Climate reanalyses provide long (multi-
84 decadal) time series of gridded estimates for many different climate variables, which are used
85 to study past weather events, estimate climatologies, monitor climate change, and supply
86 crucial data on climate needed for science and applications.

87 Reanalysis is a complex activity that requires large computational resources, access to
88 observations from many providers, and expertise in multiple disciplines. During the past three
89 decades, successive reanalyses of the global atmosphere have been produced by NCAR and
90 NCEP (*Kalnay et al., 2006, Saha et al., 2010*), the Japan Meteorological Agency (JMA;
91 *Onogi et al., 2007*), and the National Aeronautics and Space Administration (NASA;
92 *Schubert et al., 1993, Rienecker et al., 2011*), in addition to ECMWF. A global reanalysis
93 extending back to the late nineteenth century was first produced by NOAA in collaboration
94 with the Cooperative Institute for Research in Environmental Sciences (CIRES), using only
95 surface pressure observations and prior estimates of sea-surface temperature (SST) and sea-
96 ice distributions to avoid the effects that large changes in the observing system could have
97 had on the reanalysis (*Compo et al., 2011*).

98 ERA-CLIM2 has contributed to advancing reanalysis science and development in four main
99 areas:

- 100 i. *Observation data rescue and post-processing*: activities under this theme include a
101 large effort on data rescue for historic in-situ weather observations around the world,
102 and substantial work on the reprocessing of satellite climate data records and enabling
103 the use of historical satellite data for reanalysis;

- 104 ii. *Data assimilation methods*: activities under this theme aim to progress the
105 development and testing of ‘coupled assimilation methods’, capable of including
106 observations from different Earth-system components (land surface, ocean, sea-ice,
107 atmosphere, chemical components, ...) to produce a more consistent estimate of the
108 Earth-system evolution, especially at the surface;
- 109 iii. *Reanalysis production*: activities under this theme aim to generate the innovative
110 reanalysis data-sets, such as the first European coupled ocean-land-atmosphere
111 reanalysis of the 20th century, and to provide access to the reanalysis data;
- 112 iv. *Evaluation and uncertainty estimation*: activities under this theme aim to assess the
113 reanalyses’ quality and how products differ from previous uncoupled products, and to
114 develop methods for estimating uncertainty in reanalyses.

115 Hereafter, we will discuss briefly some of these activities.

116

117 2. Reanalysis as a tool to monitor the climate

118 The Earth’s climate has traditionally been studied by statistical analysis of observations of
119 particular weather elements such as pressure, temperature, wind and rainfall. These
120 meteorological observations are temporally and spatially incomplete and are often presented
121 in terms of long-term averages to identify evidence of climate change. Reanalyses provide a
122 more complete source of data to understand and monitor the climate. In a reanalysis, the
123 weather observations collected in past decades are fed into a modern forecasting system
124 designed to assimilate observations from existing and planned platforms (e.g. forthcoming
125 satellite instruments), which provides a physically consistent description of the Earth system.
126 By constantly correcting a model simulation towards the observations, a reanalysis combines
127 the advantages of the model first guess and the available data (further details in section 4).
128 Therefore, reanalyses are physically consistent, spatially complete and encompass many

129 variables for which observations are not always available. Irregular and intermittent
130 observation sampling throughout the reanalysis period, especially during early periods, might
131 however prevent the reanalysis dataset from showing temporal homogeneity at both global
132 and regional scale (e.g. *Ferguson and Villarini, 2012*; see also Section 7).

133 Since its creation in 1975, ECMWF has been a key player in the production of reanalyses.
134 The initial focus was on producing atmospheric reanalyses covering the modern observing
135 period, from 1979. The first of these reanalyses, FGGE (First GARP Global Experiment,
136 where GARP stands for Global Atmospheric Research Program), was produced in the 1980s
137 at the Geophysical Fluid Laboratories (GFDL; *Ploshay et al., 1992*), followed by ERA-15
138 (European Reanalysis, version 15-years), ERA-40 (*Uppala et al., 2005*) and ERA-Interim
139 (*Dee et al., 2011*). The next reanalysis in this series, ERA5, is now in production after many
140 years of research and development.

141 Generating a reanalysis for climate monitoring is very challenging because it needs to be
142 extended further back in time when the observing system was very sparse, especially before
143 the availability of satellite data from the 1970s onwards, and even more so before the arrival
144 of radiosonde measurements in the 1930s. To tackle the unavoidable issue of the ever-
145 changing observational network, the European Union started the ERA-CLIM (European
146 Reanalysis of Global Climate Observations) project, to investigate data selection for
147 reanalyses covering the whole 20th century. As part of the ERA-CLIM project, ECMWF
148 produced the uncoupled atmospheric reanalysis ERA-20C, which covers the period January
149 1900 to December 2010 (*Poli et al., 2016*). The ERA-20C reanalysis assimilated only
150 conventional observations of surface pressure and marine wind, obtained from well-
151 established climate data collections. ERA-20C delivered three-hourly products describing the
152 spatial and temporal evolution of the atmosphere, land surface and waves.

153 As part of the FP7 ERA-CLIM2 project, the reanalysis capabilities developed in the ERA-
154 CLIM project have been extended to the ocean and sea-ice components. A new assimilation
155 system (CERA, the Coupled European Reanalysis system) has been developed to
156 simultaneously ingest atmospheric and ocean observations in the coupled Earth system model
157 used for ECMWF's ensemble forecasts (*Laloyaux et al., 2016, 2017*). This approach accounts
158 for interactions between the atmosphere and the ocean during the assimilation process and
159 has the potential to generate a more balanced and consistent Earth system climate
160 reconstruction (Figure 1). CERA has also been found to reproduce better the observed
161 negative SST-precipitation relationships on monthly timescales due to the resolving of
162 atmospheric feedbacks on SST (Figure 2). Efforts are being made to investigate whether this
163 improvement will improve the prediction of final precipitation (*Feng et al., 2017b*). One of
164 the key deliverables of the ERA-CLIM2 project is CERA-20C, the first ten-member
165 ensemble of coupled climate reanalyses of the 20th century. It is based on the CERA system,
166 which assimilates only surface pressure and marine wind observations as well as ocean
167 temperature and salinity profiles. The data are openly available from the ECMWF platform
168 <http://apps.ecmwf.int/datasets/>.

169 There is now a strong need for detailed information of CO₂ fluxes and carbon pools from the
170 climate modelling community who want to understand and quantify the carbon cycle at
171 global and regional scales, and from policy makers and citizens who want to take well-
172 informed decisions on CO₂ emissions at regional and local scales. For this reason, ERA-
173 CLIM2 is also producing associated global reanalyses of carbon fluxes and stocks using
174 terrestrial biosphere and ocean biogeochemistry models, which are forced by the CERA-20C
175 reanalysis.

176 A new version of the CERA system has also been developed based on a higher resolution
177 coupled model with the full observing system. This system is now used to produce the

178 CERA-SAT (the Coupled European ReAnalysis of the SATellite era) reanalysis which covers
179 the period 2008-2016.

180 The availability of atmospheric, oceanic and coupled reanalyses allows for new advanced
181 coupled diagnostics of the global energy cycle. *Mayer et al. (2017)* improved the classic
182 method of evaluating the vertically integrated atmospheric energy budget such that it is
183 independent of reference temperature and consistent with diagnosed ocean heat budgets.
184 *Pietschnig et al. (2017)* demonstrated the value of comparing the net energy transport through
185 all major Arctic Ocean gateways from ocean reanalyses with independent mass-consistent
186 transport estimates from instrumented mooring observations.

187

188 3. Observation data rescue

189 Reanalyses efforts strongly depend on observations of the atmosphere, land surface, the
190 ocean and cryosphere. Observations are assimilated into a coupled general circulation model
191 in order to produce the reanalysis, but observations are also used in several other steps. They
192 are used to constrain the boundary conditions, to calibrate certain relations and to validate the
193 final product. Particularly when going back in time, not all observations are available. A large
194 fraction of historical meteorological observations has never been digitised because the data
195 have thus far not been considered valuable. Even in the rather recent past, the availability of
196 satellite products (and the computer code to read and process these data) is an issue that needs
197 to be addressed.

198 Major efforts were therefore undertaken in ERA-CLIM2 to collect and make available
199 observations for reanalyses (*Brönnimann et al., 2017*). Such an undertaking requires a much
200 broader vision than the production of one reanalysis. Availability of historical observations

201 becomes a legacy, and producing reanalyses or other data products must be seen as a
202 continuous effort.

203 Within ERA-CLIM2, millions of radiosonde, pilot balloon and other ascending instruments'
204 profiles were digitised, which allows describing the third dimension of the atmosphere back
205 to the early 20th century. Although the radiosonde data were not yet incorporated into the
206 CERA-20C reanalysis produced in ERA-CLIM2, they were used for the validation of
207 reanalysis products. A test reanalysis that included the historical upper-air data and
208 demonstrated the potential benefits was performed for the period 1939-1967 (*Hersbach et al.*,
209 2017). By the end of the project, all digitized upper air data will be available in assimilation-
210 ready format with bias adjustments for radiosonde temperatures (*Haimberger et al.*, 2012)
211 extending back to 1939. Future reanalysis efforts will be able to incorporate this vast amount
212 of upper-air data and will thus build on the ERA-CLIM2 efforts.

213 Surface pressure and mean-sea-level-pressure data were also digitised for several countries in
214 both the Northern and Southern Hemisphere (NH and SH), some with sparse observation
215 networks, and sent to the International Surface Pressure Databank (ISPD). These will be
216 assimilated in forthcoming ISPD versions to be used in future reanalyses inputs. Many other
217 land surface, daily and sub-daily, observations of temperature, relative humidity, surface
218 wind, cloud cover, precipitation, evaporation and sunshine duration were rescued, subjected
219 to Quality Control procedures and can be used for reanalyses comparison purposes
220 (*Brönnimann et al.*, 2017).

221 Snow is an important component of the climate system, at the interface of the land-surface,
222 vegetation and atmosphere. It is highly relevant for various fields such as ecology, water
223 resources, transport, and tourism. By digitising large amounts of historical snow data and
224 combining this with satellite products, ERA-CLIM2 generated snow products for various

225 uses, including (but not limited to) reanalyses. Snow courses are specified paths of a few km
226 length around a location: along this path, different snow properties are measured regularly.
227 Thanks to ERA-CLIM2, snow course about 1.2 million snow course observations have been
228 collected from close to 400 stations compiled and made available.

229 Satellite data are the backbone of today's reanalysis data sets, but they became only available
230 in the mid-1960s and were built for the purpose of weather monitoring. The use in reanalysis
231 requires a quantification and correction of long term effects due to systematic differences
232 between satellite instruments of the same kind and changes in the characteristics of satellites
233 and performance of sensors during their operational lifetime in space. A reprocessing of the
234 data that applies the corrections is fundamental to serve the generation of physically
235 consistent data records of geophysical variables by reanalysis. In ERA-CLIM2, efforts were
236 put into re-processing of satellite data of infrared and microwave radiances from
237 geostationary imagers and microwave sounders, radio occultation bending angle profiles for
238 several satellites and atmospheric motion vectors from different instruments in geostationary
239 and polar orbit. In addition, as part of satellite data rescue activities (*Poli et al., 2017*)
240 radiative transfer calculations for some early satellite instruments were enabled to allow for
241 monitoring and assimilation of the satellite measurements using the circulation model which
242 also supports the characterisation and correction of instrument issues. Indeed, a more
243 comprehensive use of early satellite data is expected to improve future reanalyses.

244

245 4. Data assimilation methods for reanalysis

246 The ERA-CLIM2 reanalyses have been produced using a state-of-the-art data assimilation
247 system, capable of combining observations in the atmosphere and ocean with a coupled
248 ocean-land-atmosphere model. Part of the work within ERA-CLIM2 was devoted to improve

249 such a data-assimilation system, testing also new methods and ideas that could be used in
250 future reanalysis productions.

251 A crucial part of the coupled climate system is the interface between the ocean/sea-ice and
252 atmosphere (*Feng et al., 2017*). In the existing CERA system, the sea surface temperature is
253 constrained to follow a global observational analysis, such as HadISST2 (*Titchner and*
254 *Rayner 2014*) or OSTIA (*Donlon et al. 2012*), which are calculated externally. Enabling the
255 next CERA system to assimilate directly the wealth of high quality satellite SST observations
256 should allow the system to combine them with the temperature profile data and the coupled
257 model background in a more consistent manner, thereby improving the accuracy of the
258 reanalysis. Care has to be taken to deal properly with biases in the satellite data that could
259 otherwise introduce spurious trends. In addition, the in situ ocean data are sparse, particularly
260 in the early part of the 20th Century.

261 Figure 3 shows an example of the results from using a new method for effectively spreading
262 sparse observational information in the data assimilation. Developments have also been made
263 to improve the assimilation of sea-ice concentration during the satellite era. Sea-ice
264 concentration estimates have error characteristics which make it difficult to assimilate in most
265 data assimilation algorithms so techniques have been tested which transform the sea-ice
266 concentration into a form which has Gaussian errors.

267 The ocean data assimilation system used in CERA is a state-of-the-art three-dimensional
268 variational system called NEMOVAR (Nucleus for European Modelling of the Ocean
269 Variational assimilation system). It is now a very flexible framework for assimilating data
270 into the NEMO ocean model, and includes developments to allow the ocean data to be
271 assimilated using sophisticated techniques similar to those used in the atmosphere (*Weaver et*
272 *al., 2016*). The system can now be configured to use information from an ensemble of model

273 runs to improve the assimilation, which is a more sophisticated system than the algorithm
274 used in the production of the ERA-CLIM2 reanalyses. In the present CERA system, the
275 ocean bias correct scheme is not applied during the ocean data assimilation as it normally
276 needs to be estimated from a priori run. However, now with CERA-20C being finished, it
277 becomes practically possible to implement this scheme throughout the whole 20th century.
278 Implementing the bias correction scheme will avoid the ocean model to produce a spurious
279 reaction to adjust the imbalance between the ocean and atmosphere initial conditions. Efforts
280 are also being made to investigate if the implementation will benefit the atmosphere analysis.
281 Furthermore, a four-dimensional version of the ocean assimilation system has been developed
282 and is under testing.

283 Currently the coupled model is introduced at the outer-loop level in the CERA system by
284 coupling ECMWF's Integrated Forecasting System (IFS) for the atmosphere, land and waves
285 to the NEMO model for the ocean and to the LIM2 model for sea ice. This means that air-sea
286 interactions are continuously taken into account when observation misfits are computed and
287 when the increments are applied to the initial condition. This allows feedback between the
288 ocean and atmosphere models (*Laloyaux et al.* 2016, 2017). Investigations into whether the
289 atmosphere observations could be used to directly correct the ocean model, and vice versa,
290 have been carried out in ERA-CLIM2 (*Storto et al.*, 2017; *Pellerej et al.*, 2016). This
291 technique, known as “strongly coupled data assimilation”, could allow even better use of
292 sparse observations. There are many open research questions in the development of strongly
293 coupled data assimilation, so ERA-CLIM2 partners were involved in the organisation of a
294 Coupled Data Assimilation workshop in October 2016, sponsored by WMO, to discuss these
295 with the wider international research community.

296 Land and ocean carbon reanalyses are also being produced as part of ERA-CLIM2 and the
297 techniques used to produce them are being developed for use in future reanalyses. Various

298 new data streams have been tested for improving the accuracy of the land carbon reanalysis
299 both through improved state-estimation, and through improving knowledge of the model
300 parameters (*Peylin et al.*, 2016). The methods used to couple ocean biogeochemical models
301 to the physical ocean/atmosphere reanalysis system have also been investigated in order to
302 provide improved information about the ocean carbon cycle.

303

304 5. Reanalysis and society

305 ERA-CLIM2 demonstrated that reanalyses can be used to understand past events and provide
306 valuable information to present day society. A demonstration case was discussed by
307 *Brugnara et al.*, (2017), who examined the weather conditions in December 1916, in the
308 middle of the First World War, when a massive snow fall event in the Southern Alps
309 triggered countless avalanches, which killed thousands of soldiers and civilians. This event
310 was studied using dynamical downscaling of the earlier ERA-20C reanalysis (the uncoupled
311 reanalysis produced by the FP7 project ERA-CLIM, the precursor of ERA-CLIM2) in
312 combination with historical observations. By looking at reanalysis data (Figure 4), the
313 atmospheric conditions that led to such catastrophic events could be understood: a blocking
314 flow situation, moisture transport from the warm Mediterranean Sea towards the Alps, and a
315 rapidly rising snow line, leading to a dangerous “rain-on-snow” situation. Historical events
316 that are captured by the ERA-CLIM reanalyses can thus inform present day risk management.

317

318 6. A new approach: ensembles of reanalyses to estimate confidence

319 The accuracy of any physical measurement is limited. Furthermore, the spatial resolution of
320 measurements and of assimilated gridded data is limited, and thus there may be deviations of
321 the grid point values from the true values. Confidence (or uncertainty) is best described by

322 the probability distribution of these deviations. The distribution itself can be characterized by
323 a few parameters such as standard deviation, or by a limited-size ensemble of realizations
324 drawn from the distribution, from which the user may derive statistics. The latter approach
325 consumes much more data storage but also leaves more choices for informed users.

326 With reanalysis products getting more mature, it is now possible to quantify their accuracy
327 and assign to the reanalysis data a confidence level. In ERA-CLIM2, this has been achieved
328 by applying an ensemble approach, based on several complete realizations of all quantities.

329 Both the 20th century reanalysis CERA-20C, and the upcoming coupled reanalysis of the
330 satellite era, CERA-SAT, consist of ten realizations, run in parallel. The ten members can be
331 used to estimate a range of possible states for all the reanalysis output variables. They have
332 been generated using an Ensemble of Data Assimilation (EDA) in the atmosphere (*Isaksen et*
333 *al.* 2010; *Bonavita et al.* 2016) and with perturbing the positions of in-situ ocean
334 observations, the air-sea fluxes and the SST following ECMWF's Ocean ReAnalysis System-
335 5 (ORAS5) (*Zuo et al.* 2016). The SST perturbations lead to variations in important output
336 parameters such as surface precipitation. Figure 5 shows how CERA-20C precipitation has
337 improved over the earlier 20th century uncoupled reanalysis ERA-20C (*Poli et al.*, 2016) over
338 less well observed land areas, which exhibited a strong under-forecast at the higher quantiles
339 in Africa and at the Monsoon areas for the 0.9 quantiles (*Rustemeier et al.*, 2017). Figure 5
340 also shows how the calculated precipitation varies within the ensemble. For example, for
341 strong precipitation episodes over Africa (upper right of panel a) the spread has been
342 increased by about 20%, making the ensemble more reliable.

343

344 7. CERA-20C: the first European 110-year coupled ocean-land-atmosphere reanalysis
345 The accurate representation of variability on inter-annual and decadal time scales is a
346 requirement for climate applications of reanalysis data, such as reconstructing the time
347 evolution of the atmosphere surface temperature and of the ocean heat content (Figure 6).
348 Climate signals in reanalyses are inevitably affected by changes in the global observing
349 system and by the presence of time-varying biases in models and observations. To build
350 confidence in climate change information derived from reanalyses, it is important that
351 information about the data assimilation methodology, the forecast model, and the input
352 observations are made available. It is also necessary to compare results based on reanalyses
353 (ECMWF CERA-20C, NOAA 20CR, ECMWF ORA-20C), with results obtained using more
354 traditional, observation-only climate datasets (CRUTEM4 two-meter temperature, GPCC
355 precipitation data, EN4 ocean temperature), and to test whether the climate signals in CERA-
356 20C are robust to different analysis methodologies. CERA-20C data are been made freely
357 available from the ECMWF web site (see: [https://www.ecmwf.int/en/research/climate-
358 reanalysis/cera-20c](https://www.ecmwf.int/en/research/climate-reanalysis/cera-20c)) precisely to favour these comparisons.

359

360 8. Summary and conclusions

361 ERA-CLIM2 is a European Union Seventh Framework project started in January 2014,
362 which involves 17 organisations (see Appendix A). It aims to produce coupled reanalyses, i.e.
363 physically consistent data sets describing the evolution of the global atmosphere, ocean, land-
364 surface, cryosphere and the carbon cycle. The main contributions of the ERA-CLIM2 project
365 to climate science have been to rescue and re-process past conventional and satellite data,
366 improve the capacity for producing state-of-the-art climate reanalyses that extend back to the
367 early 20th century, along with uncertainties, and generate unique and extremely valuable data

368 sets. One of the main deliverables of ERA-CLIM2 has been CERA-20C, the first European
369 coupled reanalysis of the 20th century. CERA-20C is now being used to generate a land
370 (water and energy) and a carbon (land and ocean) component (CERA-20C/Land and CERA-
371 20C/Carbon). At the time of writing, the production of the CERA-SAT reanalysis has started:
372 the aim is to complete the period 2008-to-date by the end of the project (December 2017).
373 Thanks to ERA-CLIM2, many older data have been rescued and post-processed, and are
374 delivered to relevant database providers so that they can be used in future reanalysis.
375 Furthermore, new assimilation methods (e.g. use of a stronger coupling method between the
376 ocean and the atmosphere and the direct assimilation of SST data) developed and tested
377 within the project are planned to be integrated and used in future reanalysis productions.
378 Understanding climate change is highly dependent on the availability of global satellite and
379 conventional observational data in the atmosphere, the land and the ocean and sea-ice, and
380 the development of coupled ocean-land-atmospheric models and assimilation systems that
381 can ingest these data. A continuous cycle of research and development activities in data-
382 assimilation, of data rescue and observation re-processing, production and diagnosis and
383 evaluation is required to improve future reanalyses, so that they can provide a better, closer-
384 to-reality image of the time evolution of the Earth system.

385 Here are two examples of why we need a continuous stream of investments in the two areas
386 mentioned above:

- 387 • Within ERA-CLIM2 millions of new observation records, mostly made before the
388 International Geophysical Year 1958, have been discovered, rescued, digitized and
389 prepared to be inserted in appropriate data sets so that they can be used in future
390 reanalysis production. As this work has been progressing, new data are discovered,
391 but unfortunately due to lack of resources and time (the project will finish at the end

392 of 2017), they will not be rescued, digitized, quality controlled and prepared to be
393 inserted in the data sets.

394 • As part of the ERA-CLIM2 work, the possibility to assimilate directly sea-surface
395 temperature observations has been explored and tested in prototype systems.

396 Furthermore, the possibility to use ensemble methods to estimate flow-dependent
397 background error statistics within the ocean has been developed and tested. Neither of
398 these advances could be included in the current ERA-CLIM2 reanalyses, since tested
399 software was not ready in time for production.

400 Preliminary assimilation experiments have shown that the amount and quality of those data
401 justify a full reanalysis, using earlier satellite observations and all conventional (surface and
402 upper-air) data, back to the early 20th century (*Hersbach et al.*, 2017). Such a reanalysis
403 would realize the potential of the data collected and would lead to a much better description
404 of the climate evolution over the last century. However, the only way to be able to continue
405 these essential activities would be to fund them either through a new stream of European
406 projects, or directly as part of the European Union Copernicus Climate Change Service (C3S)
407 activities.

408

409

410 9. Appendix A: the ERA-CLIM2 Consortium

411

412 The ERA-CLIM2 Consortium included 17 organisations:

- 413 1) European Centre for Medium-Range Weather Forecasts (ECMWF; Europe);
- 414 2) Met Office, UK (UKMO; UK);
- 415 3) European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT;
- 416 Europe);
- 417 4) University of Bern (Switzerland);
- 418 5) University of Vienna (Austria);
- 419 6) Instituto Dom Luiz Faculdade de Ciências da Universidade de Lisboa (Portugal) ;
- 420 7) Russian Research Institute of Hydrometeorological Information (RIHMI; Russia);
- 421 8) Mercator Ocean Société Civile (MERCATOR; France);
- 422 9) Météo-France (MF; France);
- 423 10) Deutscher Wetterdienst (DWD; Germany);
- 424 11) Centre European de Recherche et de Formation Avancée en Calcul Scientifique
- 425 (CERFACS; France) ;
- 426 12) Centro Euro-Mediterraneo Sui Cambiamenti Climatici (CMCC; Italy);
- 427 13) Ilmatieteen Laitos (FMI; Finland);
- 428 14) Universidad del Pacifico (Chile);
- 429 15) The University of Reading (UK);
- 430 16) Institut National de Recherche en Informatique et en Automatique (INRIA; France) ;
- 431 17) Université de Versailles Saint-Quentin-en-Yvelines (France).

432

433 10. References

- 434 - Brönnimann, S., Allan, R., Buizza, R., Bulygina, O., Dahlgren, P., Dee, D., Gomes, P.,
435 Jourdain, S., Haimberger, L., Hersbach, H., Poli, P., Pulliainen, J., Rayner, N., Schulz, J.,
436 Sterin, A., Stickler, A., Valente, MM. A., Ventura, M. C., Wilkinson, C., 2017: Observations
437 for reanalysis. *Bull. Amer. Met. Soc.*, in preparation.
- 438 - Brugnara, Y., Brönnimann, S., Zamuriano, M., Schild, J., Roher, C., Segesser, D.M., 2017:
439 Reanalysis sheds light on 1916 avalanche disaster. *ECMWF Newsletter*, **151**, 28-34
440 (doi:10.21957/h9b197).
- 441 - Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Matsui, N., Allan, R.J., Yin, X., Gleason
442 Jr, B.E., Vose, R.S., Rutledge, G., Bessemoulin P., Brönnimann, S., Brunet, M., Crouthamel,
443 R.I., Grant, A.N., Groisman, P.Y., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J.,
444 Maugeri, M., Mok, H.Y., Nordli, Ø., Ross, T.F., Trigo, R.M., Wang, X.L., Woodruff, S.D.,
445 and Worley, S.J., 2011: The Twentieth Century Reanalysis Project. *Q. J. R. Meteorol.*
446 *Soc.*, **137**, 1-28.
- 447 - Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae,
448 U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg,
449 L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L.,
450 Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi,
451 M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay,
452 P., Tavolato, C., Thépaut, J.-N. and Vitart, F., 2011: The ERA-Interim reanalysis:
453 configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**,
454 553–597 (doi:10.1002/qj.828).

455 - Dee, D. P., Balmaseda, M. A., Balsamo, G., Engelen, R., Simmons, A. J. and J.-N. Thépaut,
456 2014: Toward a consistent reanalysis of the climate system. *Bull. Amer. Meteor. Soc.*, **95**,
457 1235-1248 (doi:10.1175/BAMS-D-13-00043.1).

458 - Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., & Wimmer, W. (2012).
459 The operational sea surface temperature and sea ice analysis (OSTIA) system. *Remote*
460 *Sensing of Environment*, **116**, 140-158.

461 - Feng, X. and K. Haines, 2017. Uncertainties of sea surface and air temperature in the
462 CERA-20C coupled reanalysis ensemble. *Q. J. R. Meteorol. Soc.*, submitted.

463 - Ferguson, C. R., Villarini, G., 2012: Detecting inhomogeneities in the Twentieth Century
464 Reanalysis over the central United States. *J. Geophys. Res.*, **117**, 2156-2202.

465 - Hersbach, H., S. Brönnimann, L. Haimberger, M. Mayer, L. Villiger, J. Comeaux, A.
466 Simmons, D. Dee, S. Jourdain, C. Peubey, P. Poli, N. Rayner, A. Sterin, A. Stickler, M. A.
467 Valente, S. Worley, 2017: The potential value of early (1939-1967) upper-air data in
468 atmospheric climate reanalysis. *Quart. J. Roy. Meteorol. Soc.*, **143**, 1185-1196
469 (doi:10.1002/qj.3020).

470 - Kalnay E, Kanamitsu M, Kirtler R, Collins W, Deaven D, Gandin L, Iredell, M.,
471 Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins W., Janowiak
472 J., Mo, K.C., Ropelewski C., Wang J., Leetma, A., Reynolds, R., Jenne, R., Joseph, D.,
473 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteorol. Soc.*, **77**, 437 –
474 471.

475 - Laloyaux, P., Balmaseda, M., Dee, D., Mogensen, K. and Janssen, P, 2016: A coupled data
476 assimilation system for climate reanalysis. *Quarterly Journal of the Royal Meteorological*
477 *Society*, 142, 65-78.

478 - Laloyaux, P., de Boisseson, E. and Dahlgren, P, 2017: CERA-20C: An Earth system
479 approach to climate reanalysis. *ECMWF Newsletter*, **150**, 25-30).

480 - Mayer, M., L. Haimberger, J. M. Edwards, P Hyder, 2017: Towards consistent diagnostics
481 of the coupled atmosphere and ocean energy budgets. *J. Climate*, accepted for publication.

482 - Onogi K., Tsutsui J., Koide H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto,
483 T., Yamazaki, N., Kamahori, H., Takahashi, K., Kadokura, S., Wada, K., Kato, K., Oyama,
484 R., Ose, T., Mannoji, N. and Taira R., 2007: The JRA-25 Reanalysis. *J. Meteor. Soc. Japan*,
485 **85**, 369 – 432.

486 - Pellerej, R., A. Vidard, and F. Lemarie. Toward variational data assimilation for coupled
487 models: first experiments on a diffusion problem. *Arima, CARI 2016* (preprint), 2016.
488 [Available online at <https://hal.archives-ouvertes.fr/hal-01337743>].

489 - Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P. J., Kuppel, S., Koffi, E. N.,
490 Kane, A., Maignan, F., Chevallier, F., Ciais, P., and Prunet, P., 2016: A new stepwise carbon
491 cycle data assimilation system using multiple data streams to constrain the simulated land
492 surface carbon cycle, *Geosci. Model Dev.*, **9**, 3321-3346 (doi: 10.5194/gmd-9-3321-2016).

493 - Pietschnig, M., M. Mayer, T. Tsubouchi, A. Storto, L. Haimberger, 2017: Comparing
494 reanalysis-based volume and temperature transports through Arctic Gateways with mooring-
495 derived estimates. *J. Geophys. Res. Oceans*, under review.

496 - Ploshay, J. J., Stern, W. F., and Miyakoda, K, 1992: FGGE reanalysis at GFDL. *Mon. Wea.*
497 *Rev.*, **120**, 2083-2108.

498 - Poli, P., D. Dee, R. Saunders, V. O. John, P. Rayer, J. Schulz, K. Holmlund, D. Coppens,
499 D. Klaes, J. E. Johnson, A. E. Esfandiari, I. V. Gerasimov, E. B. Zamkoff, A. F. Al-Jazrawi,
500 D. Santek, M. Albani, P. Brunel, K. Fennig, M. Schröder, S. Kobayashi, D. Oertel, W.

501 Döhler, D. Spänkuch, and S. Bojinski, 2017: Recent Advances in Satellite Data Rescue.
502 *Bull. Amer. Meteorol. Soc.*, **91**, 1471-1484.

503 - Poli, P., H. Hersbach, D.P. Dee, P. Berrisford, A.J. Simmons, F. Vitart, P. Laloyaux, D.G.
504 Tan, C. Peubey, J. Thépaut, Y. Trémolet, E.V. Hólm, M. Bonavita, L. Isaksen, and M. Fisher
505 (2016) ERA-20C: An Atmospheric Reanalysis of the Twentieth Century. *J. Climate*, **29**,
506 4083–4097 (DOI:10.1175/JCLI-D-15-0556.1).

507 - Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich,
508 M.G., Schubert, S.D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A.,
509 da Silva, A., Gu. W., Joiner, J., Koster, R.D., Lucchesi, R., Molod, A., Owens, T., Pawson,
510 S., Pegion, P., Redder, C.R., Reichle, R., Robertson, F.R., Ruddick, A.G., Sienkiewicz, M.,
511 and Woollen, J., 2011: MERRA – NASA’s Modern-Era Retrospective Analysis for Research
512 and Applications. *J. Climate*, **24**, 3624-3648 (DOI: 10.1175/JCLI-D-11-00015.1).

513 - Rustemeier, E., Ziese, M., Meyer-Christoffer, A., Schneider, U., Finger, P., Becker, A.
514 (2017): Uncertainty assessment of the ERA-20C reanalysis based on the monthly in-situ
515 precipitation analyses of the Global Precipitation Climatology Centre. In prep for submission
516 to *J. Hydrometeor.*

517 - Saha S., Moorthi S., Pan H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R.,
518 Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Hou, Y.-T.,
519 Chuang, H.-Y., Juang, H.-M.H., Sela, J., Iredell, M., Treadon, R., Kleist, D., van Delst, P.,
520 Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H.,
521 Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K.,
522 Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen
523 Y., Han, Y., Cucurull, L., Reynolds, R.W., Rutledge, G., Goldberg, M., 2010: The NCEP
524 Climate Forecast System Reanalysis. *Bull. Amer. Meteorol. Soc.*, **91**, 1015-1057.

525 - Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Ziese, M. (2015):
526 GPCC Full Data Reanalysis Version 7.0 at 1.0°: Monthly land-surface precipitation from
527 rain-gauges built on GTS-based and historic data. 2015. (DOI:
528 10.5676/DWD_GPCC/FD_M_V7_100).

529 - Schubert, S.D., Rood, R., and Pfaendtner, J., 1993: An assimilated dataset for Earth science
530 applications. *Bull. Am. Meteorol. Soc.*, **74**, 2331-2342.

531 - Storto, A., M. J. Martin, B. Deremble, S. Masina, 2017: Strongly coupled data assimilation
532 experiments with linearized ocean-atmosphere balance relationships. *Mon. Wea. Rev.*,
533 submitted.

534 - Titchner, H. A., & Rayner, N. A. (2014). The Met Office Hadley Centre sea ice and sea
535 surface temperature data set, version 2: 1. Sea ice concentrations. *Journal of Geophysical*
536 *Research: Atmospheres*, **119**(6), 2864-2889.

537 - Uppala, S. M., P. W. Kallberg, A. J. Simmons, U. Andrae, V. Da Costa Bechtold, M.
538 Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N.
539 Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. Van De
540 Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher,
541 M. Fuentes, S. Hagemann, E. Hólm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne,
542 A. P. McNally, J.-F. Mahfouf, J.-J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A.
543 Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, J. Woollen, (2005): The ERA-40
544 re-analysis. *Q.J.R. Meteorol. Soc.*, **131**, 2961–3012 (doi:10.1256/qj.04.176).

545 - Weaver, A.T., Tshimanga J., Piacentini A., 2016. Correlation operators based on an
546 implicitly formulated diffusion equation solved with the Chebyshev iteration. *Q. J. Roy.*
547 *Meteorol. Soc.*, **142**: 455-471.

548

549 11. Figures captions

550 Figure 1. High-pass filtered sea surface temperature (colour) and wind stress (contour) for
551 ERA-20C (left) and CERA-20C (right) over the period 01/04/73 to 27/03/74. CERA-20C
552 represents the Tropical Instability Waves thanks to the ocean dynamics and the atmosphere is
553 responding accordingly with the surface wind stress sensitive to the ocean TIWs. In ERA-
554 20C, there are no TIWs and wind stress signals.

555

556 Figure 2. (a-c) SST-precipitation correlations for their monthly fluctuations in observational
557 data (HadISST2 and GPCP), ERA-20C and CERA-20C (control run), over 1979-2010. Blank
558 areas are where the correlations do not pass the significance test at the 90% confidence level.
559 Note the agreement between observations and CERA-20C in the heavily precipitating
560 regions.

561

562 Figure 3. Examples of sea surface temperature observations (left) and temperature profile
563 observations (right) in °C from Jan 1953 (top panels) and Jan 2010 (middle panels). The
564 bottom plot shows the percentage change in error in SST from assimilating the Jan 1953 data
565 using the new version of NEMOVAR.

566

567 Figure 4. Results from the dynamical downscaling of ERA-20C: (a) total precipitation
568 (shading) on 13 December 1916 and mean freezing level (grey contours; in m). (b) change in
569 snow depth between 5 and 13 December 1916. Circles represent snow observations, red
570 crosses show the locations of documented major avalanches on this day. The military front
571 line in 1916 is shown as red dotted line.

572

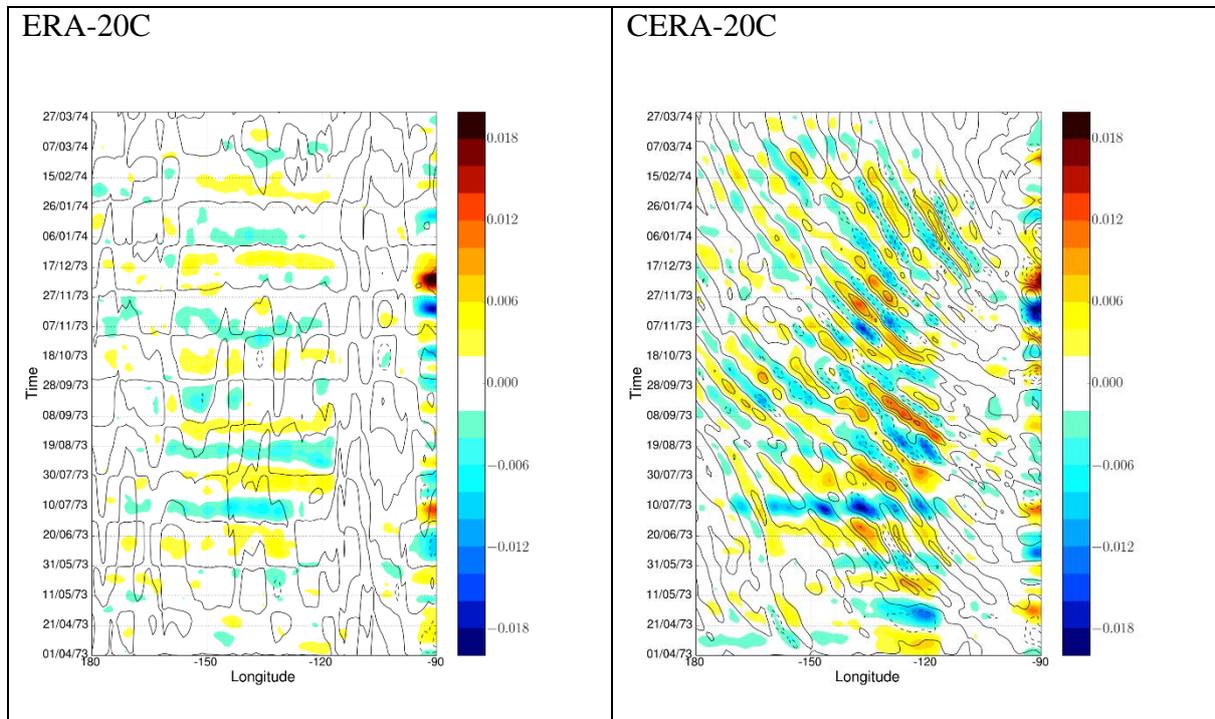
573 Figure 5. Quantile-Quantile Plot of monthly regional mean short term precipitation forecasts
574 from ERA-20C (black) and CERA-20C (grey) against Full Data Monthly V7 (Schneider et
575 al., 2015)GPCC precipitation estimates from rain gauges provided by the Global Precipitation
576 Climatology Centre (GPCC). a) Africa, b) India/Monsoon region,, c) South America,. Time
577 interval considered: 1901-2000 on 1° spatial resolution, monthly temporal resolution.

578

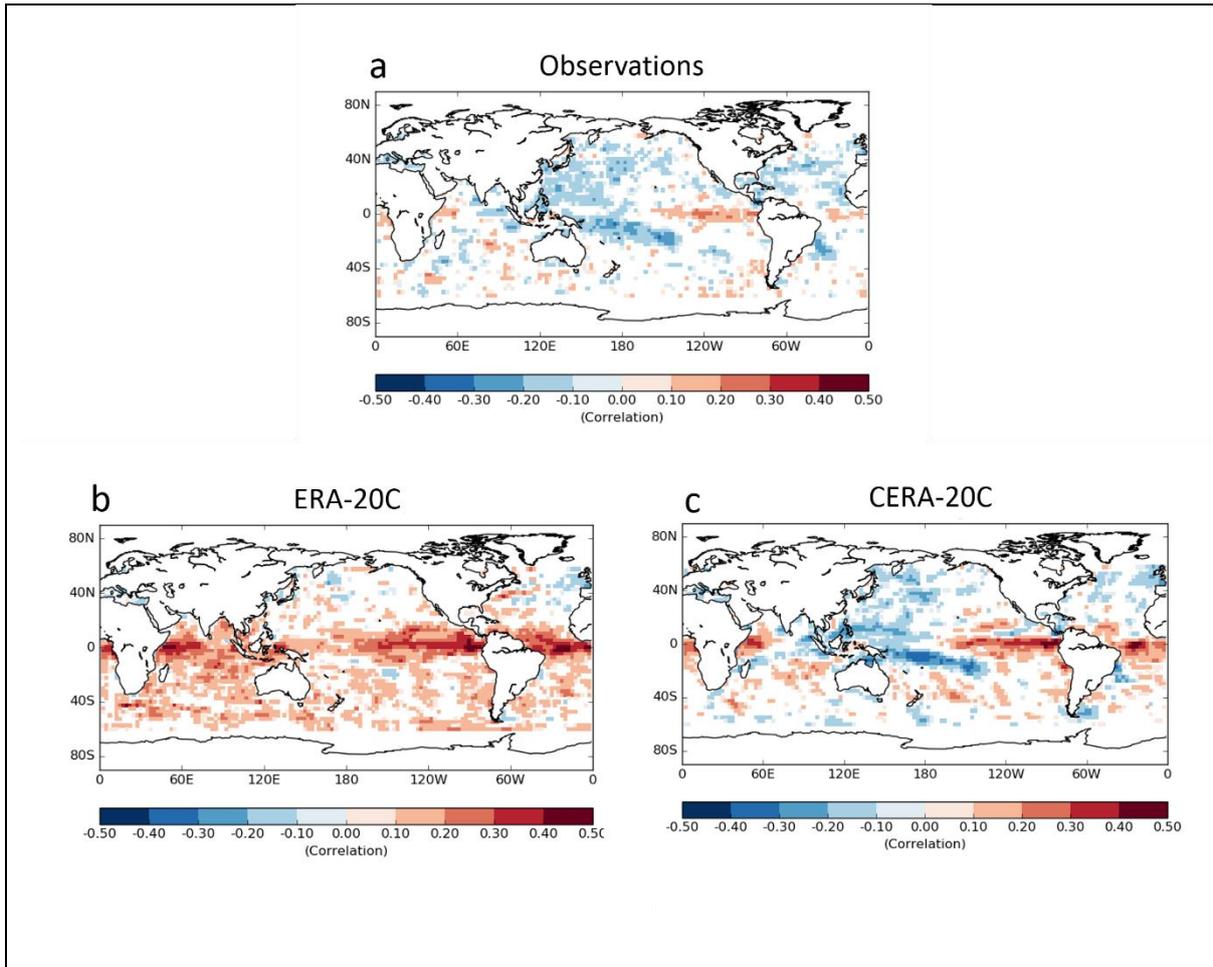
579 Figure 6. The two time series show the evolution of yearly global-mean anomalies relative to
580 the period 1961–1990 for the two-metre temperature (top) and the upper-ocean heat content
581 in CERA20C.

582

583



587 Figure 1. High-pass filtered sea surface temperature (colour) and wind stress (contour) for
588 ERA-20C (left) and CERA-20C (right) over the period 01/04/73 to 27/03/74. CERA-20C
589 represents the Tropical Instability Waves thanks to the ocean dynamics and the atmosphere is
590 responding accordingly with the surface wind stress sensitive to the ocean TIWs. In ERA-
591 20C, there are no TIWs and wind stress signals.

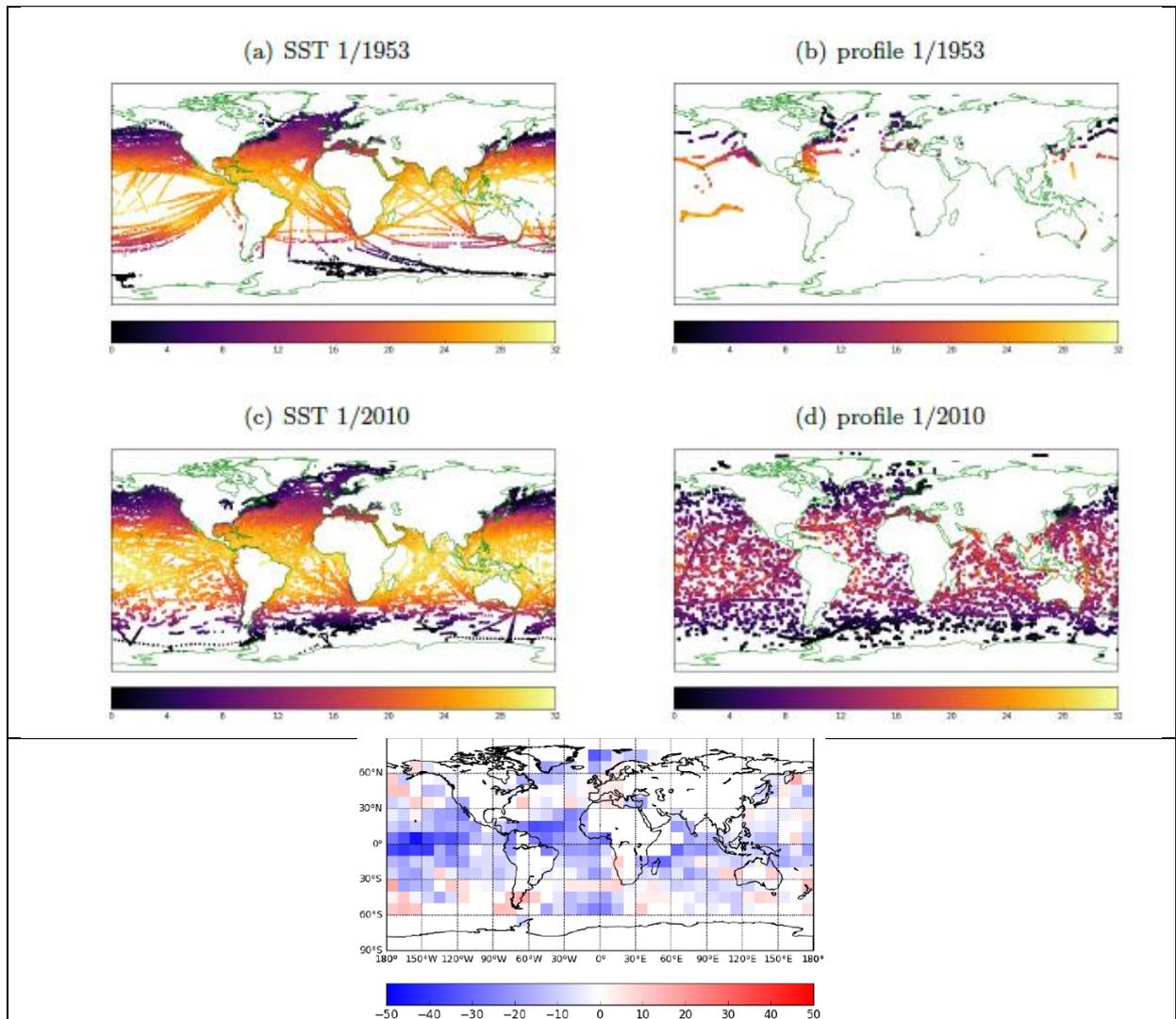


594

595 Figure 2. (a-c) SST-precipitation correlations for their monthly fluctuations in observational
 596 data (HadISST2 and GPCP), ERA-20C and CERA-20C (control run), over 1979-2010. Blank
 597 areas are where the correlations do not pass the significance test at the 90% confidence level.
 598 Note the agreement between observations and CERA-20C in the heavily precipitating
 599 regions.

600

601

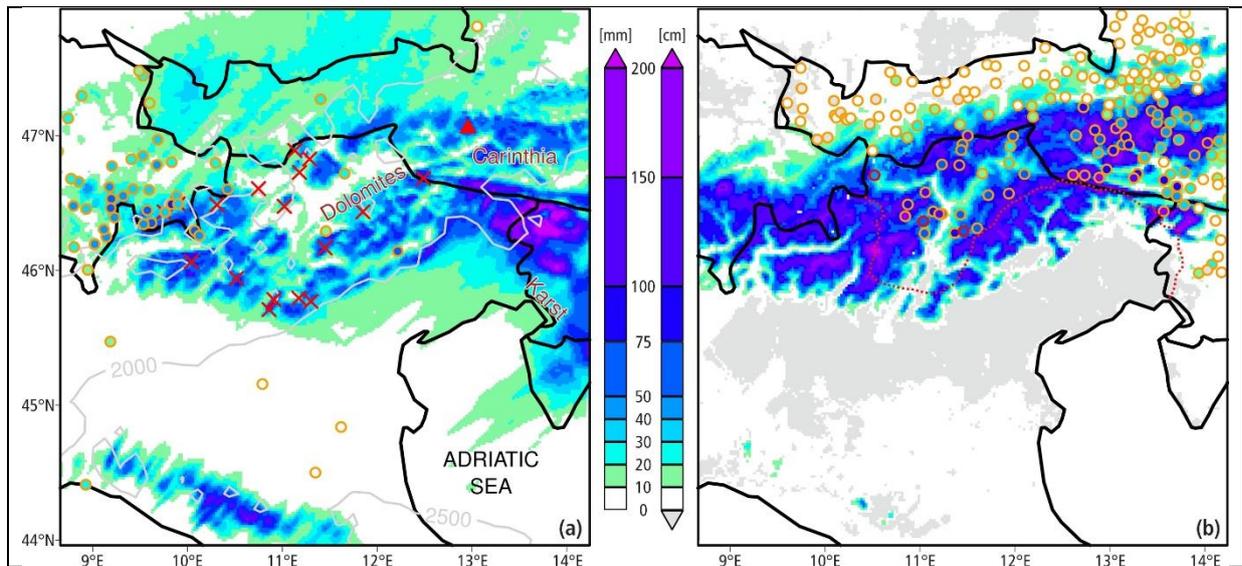


602

603 Figure 3. Examples of sea surface temperature observations (left) and temperature profile
604 observations (right) in °C from Jan 1953 (top panels) and Jan 2010 (middle panels). The
605 bottom plot shows the percentage change in error in SST from assimilating the Jan 1953 data
606 using the new version of NEMOVAR.

607

608

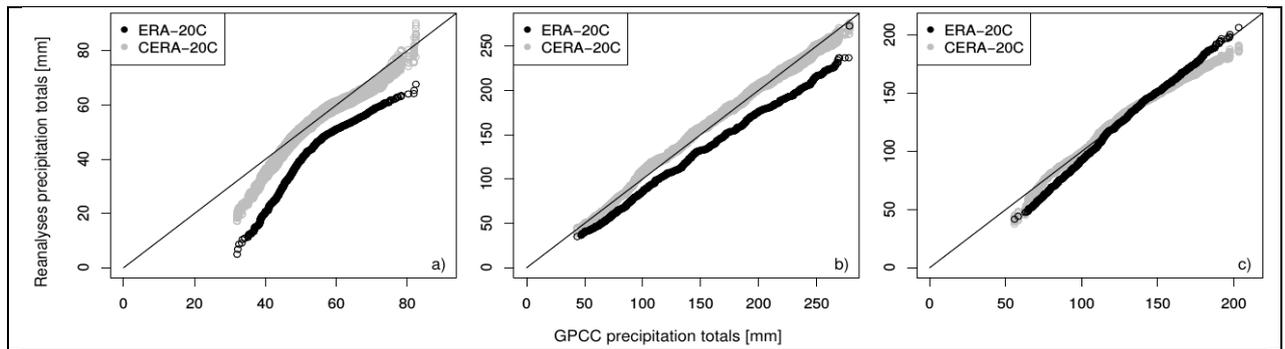


609

610 Figure 4. Results from the dynamical downscaling of ERA-20C: (a) total precipitation
611 (shading) on 13 December 1916 and mean freezing level (grey contours; in m). (b) change in
612 snow depth between 5 and 13 December 1916. Circles represent snow observations, red
613 crosses show the locations of documented major avalanches on this day. The military front
614 line in 1916 is shown as red dotted line.

615

616

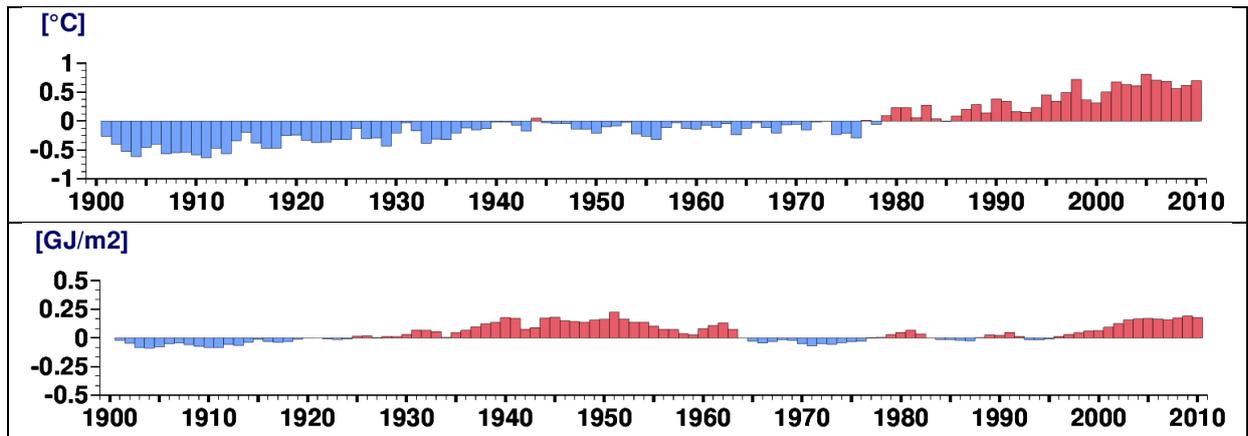


617

618 Figure 5. Quantile-Quantile Plot of monthly regional mean short term precipitation forecasts
619 from ERA-20C (black) and CERA-20C (grey) against Full Data Monthly V7 (Schneider et
620 al., 2015) GPCCC precipitation estimates from rain gauges provided by the Global Precipitation
621 Climatology Centre (GPCCC). a) Africa, b) India/Monsoon region, c) South America,. Time
622 interval considered: 1901-2000 on 1° spatial resolution, monthly temporal resolution.

623

624



625

626 Figure 6. The two time series show the evolution of yearly global-mean anomalies relative to
627 the period 1961–1990 for the two-metre temperature (top) and the upper-ocean heat content
628 in CERA20C.

629