ERA-Interim/Land: A global land water resources dataset

Gianpaolo Balsamo[(1)], Clement Albergel[(1)], Anton Beljaars[(1)], S. Boussetta[(1)], Hannah Cloke[(2)], Dick Dee[(1)], Emanuel Dutra[(1)], Joaquín Muñoz-Sabater[(1)], Florian Pappenberger[(1)], Patricia de Rosnay[(1)], Tim Stockdale[(1)], Frederic Vitart[(1)]

[1] European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK.

[2] University of Reading, UK.

Submitted to HESS on 10th of November 2013

Revised on March 2014

Corresponding author: gianpaolo.balsamo@ecmwf.int

European Centre for Medium-Range Weather Forecasts (ECMWF) Shinfield Park, RG2 9AX Reading, UK

Tel: +44 (0) 118 9499 246

Abstract- The ERA-Interim/Land is a global land-surface dataset covering the period 1979–2010 and describing the evolution of the soil (moisture and temperature) and snowpack. ERA-Interim/Land is the result of a single 32yr simulation with the latest ECMWF land surface model driven by meteorological forcing from the ERA-Interim atmospheric re-analysis and precipitation adjustments based on GPCP v2.1 with horizontal resolution of about 80km and 3-hourly frequency. ERA-Interim/Land preserves closure of the water balance and includes a number of parameterisations improvements in the land surface scheme with respect to the original ERA-Interim dataset, which makes it more suitable for climate studies involving land water resources. The quality of ERA-Interim/Land, assessed by comparing with ground-based and remote sensing observations is discussed. In particular, estimates of soil moisture, snow depth, surface albedo, turbulent latent and sensible fluxes, and river discharges are verified against a large number of sites measurements. ERA-Interim/Land provides a global integrated and coherent water resources estimate that is used also for the initialization of numerical weather prediction and climate models.

1 Introduction

Multi-model land-surface simulations, such as those performed within the Global Soil Wetness Project (Dirmeyer 2011; Dirmeyer et al. 2002, 2006), combined with seasonal forecasting systems have been crucial in triggering advances in land-related predictability as documented in the Global Land Atmosphere Coupling Experiments (Koster et al. 2011, 2009, 2006). The land-surface state estimates used in those studies was generally obtained with offline model simulations, forced by 3-hourly meteorological fields from atmospheric reanalyses, and combined with simple schemes to address climatic biases. Bias corrections of the precipitation fields are particularly important to maintain consistency of the land hydrology. The resulting land-surface data sets have been of paramount importance for hydrological studies addressing global water resources (e.g. Oki and Kanae 2006). A state-of-

50 the-art land-surface reanalysis covering the most recent decades is highly relevant to foster
51 research into intra-seasonal forecasting in a changing climate, as it can provide consistent land
52 initial conditions to weather and climate models.

53 In recent years several improved global atmospheric reanalyses of the modern era from 1979
54 onwards have been produced that enable new applications of offline land-surface simulations.
55 These include ECMWF's Interim reanalysis (ERA-Interim, Dee et al. 2011) and NASA's
56 Modern Era Retrospective-analysis for Research and Applications (MERRA, Rienecker et al.
57 2011). Simmons et al. (2010) have demonstrated the reliability of ERA-Interim near-surface
58 fields by comparing with observations-only climatic data records. Balsamo et al. (2010a)
59 evaluated the suitability of ERA-Interim precipitation estimates for land applications at
60 various time-scales from daily to annual over the conterminous US. They proposed a scale-
61 selective rescaling method to address remaining biases based on Global Precipitation
62 Climatology Project monthly precipitation data (GPCP, Huffman et al. 2009). This method
63 "calibrates" the monthly precipitation amount addressing the issue of non-conservation
64 typical of data assimilation systems, as analysed in Berrisford et al. (2011). Szczypta et al.
65 (2011) have evaluated the incoming solar radiation provided by the ERA-Interim reanalysis
66 with ground-based measurements over France. They showed a slight positive bias, with a
67 modest impact on land-surface simulations. Decker et al. (2012) confirmed these findings
68 using flux tower observations and showed that the land-surface evaporation of ERA-Interim
69 compared favourably with the observations and with other reanalyses.

70 Offline land-surface only simulations forced by meteorological fields from reanalyses are not
71 only useful for land-model development but can also offer an affordable mean to improve the
72 land-surface component of reanalysis itself. Reichle et al. (2011) have used this approach to
73 generate an improved MERRA-based land-surface product (MERRA-Land,
74 http://gmao.gsfc.nasa.gov/research/merra/merra-land.php). Similarly we have produced ERA-
75 Interim/Land, a new global land-surface data set associated with the ERA-Interim reanalysis,

by incorporating recent land model developments at ECMWF combined with precipitation bias corrections based on GPCP v2.1.

To produce ERA-Interim/Land, near-surface meteorological fields from ERA-Interim were used to force the latest version of the HTESSEL land-surface model (Hydrology-Tiled ECMWF Scheme for Surface Exchanges over Land). This scheme is an extension of the TESSEL scheme (van den Hurk et al. 2000) that was used in ERA-Interim, which was based on a 2006 version of ECMWF's operational Integrated Forecasting System (IFS). HTESSEL includes an improved soil hydrology (Balsamo et al. 2009), a new snow scheme (Dutra et al. 2010), a multi-year satellite-based vegetation climatology (Boussetta et al. 2013a), and a revised bare-soil evaporation (Balsamo et al. 2011; Albergel et al. 2012a). The majority of improvements in ERA-Interim/Land in Northern hemisphere are attributed to land parameterization revisions, while the precipitation correction is important in Tropics and Southern hemisphere.

The next section describes the various data sets used for production and verification of ERA-Interim/Land. Section 3 describes the offline land-surface model integrations. Section 4 presents the main results on verification of land-surface fluxes, soil moisture, snow, and surface albedo. The land-surface estimates from ERA-Interim/Land are a preferred choice for initializing ECMWF's seasonal forecasting system (System-4, Molteni et al. 2011), as well as the monthly forecasting system (Vitart et al. 2008), since both systems make use of ERA-Interim/Land scheme. A summary and recommendation for the usage of the ERA-Interim/Land product is reported in the conclusions.

2 Dataset and methods

The experimental set-up makes use of offline (or stand-alone) land simulations, which represents a convenient framework for isolating benefits and deficiencies of different land surface parameterizations (Polcher et al. 1998). In addition, in terms of computational cost,

given the complexity of the coupling with the atmosphere, offline simulations are much more cost-effective (faster) to run than a full atmospheric-land assimilation system.

In this study, offline runs are performed both at the global and point scales. All the 3-hourly meteorological forcing parameters were linearly interpolated in time to the land surface model integration time step of 30 minutes. The land-use information has been derived from the United States Geophysical Survey - Global Land Cover Classification (USGS-GLCC) and the United Nations - Food and Agriculture Organization (UN-FAO) data set at the same resolution as the forcing data. A comprehensive description of the land surface model and the ancillary datasets is given in the IFS documentation (2012, Part IV, chapters 8 and 11, http://www.ecmwf.int/research/ifsdocs/CY37r2/index.html).

2.1 Validation and supporting datasets

The quality of ERA-Interim/Land builds upon an error correction methodology applied to the forcing precipitation and on a comprehensive verification applied to the different compartments of the water and energy cycle at the surface. In the following the datasets entering the ERA-Interim/Land generation and its verification are briefly presented. The datasets ERA-Interim and GPCP v2.1 are support to the generation of ERA-Interim/Land while the other datasets are used for the validation of the water cycle components (water storage terms and fluxes)**.**

2.1.1 ERA-Interim meteorological reanalysis

ERA-Interim (Dee et al. 2011) is produced at T255 spectral resolution (about 80 km) and covers the period January 1979 to present, with product updates approximately 1 month delay from real-time. The ERA-Interim atmospheric reanalysis is built upon a consistent assimilation of an extensive set of observations (typically tens of millions) distributed worldwide (from satellite remote sensing, in-situ, radio-sounding, profilers, etc.) and by the analysis step that combines observations and Earth system model a-priori information in a

127 statistically optimal manner. In ERA-Interim two analyses per day are performed at 00 and 12
128 UTC times and serves as initial conditions for the subsequent forecasts. To create a
129 continuous time series of meteorological forcing therefore an opportune combination of
130 analyses and forecasts has been produced as detailed in the Fig. 1. The weather forecast's
131 spin-up effects are typical of fields such as precipitation and radiation fluxes, for which the
132 first hours after the analysis are subject to some initial shock problem. The atmospheric
133 forcing data was gridded on the original reduced Gaussian grid (with a resolution of 0.7° at
134 the Equator) with a 3-hour time interval. ERA-Interim precipitation and radiation fields
135 (incoming long- and short-wave components) are generated by the forecast model and stored
136 as 3-hourly accumulations. To avoid possible spin-up effects of precipitation and radiation (as
137 documented in Kållberg 2011) on the offline land surface simulations, the 3-hourly surface
138 fluxes correspond to the 09-21h forecast intervals from initial conditions at 00 and 12 UTC.
139 ERA-Interim temperature, surface pressure, humidity and wind fields are instantaneous values
140 representative of the lowest model level corresponding to a height of 10m above the surface
141 and are extracted from the 03-12 forecast-range intervals and from both 00 and 12 UTC runs.
142 The schematic representation in Fig. 1 shows how the continuous meteorological forcing is
143 generated for a given day. The difference in the choice of forecast range selected for
144 instantaneous and accumulated fluxes is motivated by the spinup effect being a concern
145 mainly for the latter. The forecasts are then concatenated to produce a continuous 3-hourly
146 meteorological forcing data set that can be used to drive land surface simulations. The ERA-
147 Interim 3-hourly precipitation is rescaled to match the GPCP monthly averages, as detailed by
148 Balsamo et al. (2010a).

149 ### *2.1.2 GPCP v2.1 precipitation*

150 The GPCP dataset merges satellite and rain gauge data from a number of satellite sources
151 including the Global Precipitation Index, the Outgoing long-wave radiation Precipitation
152 Index (OPI), the Special Sensor Microwave/Imager (SSM/I) emission, the SSM/I scattering,
153 and the TIROS Operational Vertical Sounder (TOVS). In addition, rain gauge data from the

154 combination of the Global Historical Climate Network (GHCN) and the Climate Anomaly
155 Monitoring System (CAMS), as well as the Global Precipitation Climatology Centre (GPCC)
156 dataset which consists of approximately 6700 quality controlled stations around the globe
157 interpolated into monthly area averages, are used over land. Adler et al. (2003) detail the
158 datasets and methods used to merge these data.

159 Compared to earlier versions the version 2.1 of the GPCP used in this study takes advantage
160 of the improved GPCC gauge analysis and the usage of the OPI estimates for the new SSM/I
161 era. Thus, the main differences between the two versions are introduced by the use of the new
162 GPCC full data reanalysis (Version 4) for 1997-2007, the new GPCC monitoring Product
163 (version 2) thereafter and the recalibration of the OPI data to a longer 20-year record of the
164 new SSM/I-era GPCP data. Further details on the new version can be found in Huffman et al.
165 (2009).

166 *2.1.3 FLUXNET land energy fluxes*

167 FLUXNET is a global surface energy, water, and CO_2 FLUX observation NETwork and it is a
168 collection of existing regional networks (Baldocchi et al. 2001, http://fluxnet.ornl.gov).
169 Available observational data for the year 2006 from the Boreal Ecosystem Research and
170 Monitoring Sites (BERMS, Betts et al. 2006), and the Coordinated Energy and water cycle
171 Observations Project (CEOP) were also used in this study.

172 As part of the CEOP program, reference site observations from the Amazonian region also
173 belonging to the LBA experiments (the Large Scale Biosphere-Atmosphere Experiment in
174 Amazonia) are available for scientific use. In this study, observations are taken from flux
175 towers located within an evergreen broadleaf forest (Manaus) and a woody savannah region
176 (Brasilia).

177 The FLUXNET observations used in this study are part of the LaThuile dataset, which
178 provides flux tower measurements of latent heat flux (LE), sensible heat flux (H) and net
179 ecosystem exchange (NEE) at high temporal resolution (30 min to 60 min). For verification

purposes, hourly observations from the year 2004 were selected from the original observational archive (excluding gap filled values) with high quality flag only (see Table 1).

2.1.4 ISMN soil moisture observing network

In-situ soil moisture observations are valuable to evaluate modelled soil moisture. In the recent years huge efforts were made to collect observations representing contrasting biomes and climate conditions. Some of them are now freely available on the Internet such as data from The International Soil Moisture Network (ISMN, Dorigo et al. 2011, 2013, http://ismn.geo.tuwien.ac.at/). The ISMN is a new data-hosting centre where globally available ground-based soil moisture measurements are collected, harmonized and made available to users. This includes a collection of nearly 1000 stations (with data from 2007 up to present) gathered and quality controlled at ECMWF. Albergel et al. (2012a, b, c) have used these data to validate various soil moisture estimates produced at ECMWF, including from ERA-Interim as well as from offline land simulations. Data from 6 networks are considered for 2010: NRCS-SCAN (Natural Resources Conservation Service - Soil Climate Analysis Network) and SNOTEL (short for SNOwpack TELemetry) over the United States, with 177 and 348 stations, respectively; SMOSMANIA (Soil Moisture Observing System-Meteorological Automatic Network Integrated Application) with 12 stations; REMEDHUS (REd de MEDición de la HUmedad del Suelo) in Spain with 20 stations, the Australian hydrological observing network labelled OZNET with 38 stations; and AMMA (African Monsoon Multidisciplinary Analyses) in western Africa with 3 stations. Data at 5 cm are used and the year 2010 is retained for the comparison. Table 3 gives a full list of reference for each network as well as the main statistical scores for the comparison.

2.1.5 The GTS-SYNOP observing network

The SYNOP (surface SYNOPtic observation) is an operationally maintained datasets under the coordination of the World Meteorological Organization (WMO), which provides daily ground-based observations of the main weather parameters and selected land surface

206 quantities such as snow depth, at a large number of sites worldwide. The snow data are
207 acquired at a minimum frequency of once a day and represent the only quantitative snow-
208 depth measurement on the ground (remote sensing observations have difficulties in
209 representing snow properties). These data are operationally used at ECMWF for the daily
210 global snow analysis as described in Drusch et al. (2004) and de Rosnay et al. (2013a).

211 ### *2.1.6 The satellite surface albedo*

212 The Moderate Resolution Imaging Spectro-radiometer (MODIS) albedo product MCD43C3
213 provided data describing both directional hemispheric reflectance (black-sky albedo) and bi-
214 hemispherical reflectance (white-sky albedo) in seven different bands and aggregated bands.
215 Data from the *Terra* and *Aqua* platforms are merged in the generation of the product that is
216 produced every 8 days, with 16 days acquisition, and available on a 0.05° global grid. The
217 accuracy and quality of the product has been studied by several authors (e.g Roman et al.
218 2009; Salomon et al. 2006). The MODIS product has served as a reference for model
219 validations (e.g. Dutra et al. 2010, 2012; Wang and Zeng 2010; Zhou et al. 2003). In this
220 study, we compare the white-sky broadband shortwave albedo (2000-2010) with ERA-Interim
221 and offline simulations. MODIS albedo was averaged for each month and spatially
222 aggregated to the simulation grid.

223 ### *2.1.7 The GRDC river discharge dataset*

224 The Global Runoff Data Centre (GRDC) operates under the auspice of the World
225 Meteorological Organization and provides data for verification of atmospheric and hydrologic
226 models. The GRDC database is updated continuously, and contains daily and monthly
227 discharge data information for over 3000 hydrologic stations in river basins located in 143
228 countries. Over the GSWP-2 period the runoff data of 1352 discharge gauging stations was
229 available and used for verification of the soil hydrology (Balsamo et al. 2009). Pappenberger
230 et al. (2009) and Balsamo et al. (2010b) used the GRDC daily discharge to evaluate a coupled
231 land surface – river discharge scheme for river flood prediction.

2.2 Land modelling component

ERA-Interim/Land differs from the land component of ERA-Interim in a number of land surface parameterization improvements introduced in the operational ECMWF forecast model since the frozen cycle used in the ERA-Interim reanalysis. The meteorological forcing described in 2.1.3 is used to drive a 11 yr spin-up run (1979 to 1989 included) that serves the purpose of generating plausible initial conditions for the 1st of January (as average of span-up run dates in 1980-1989).

A single continuous 32 yr simulation starting on the 1st of January 1979 is then realised with the latest ECMWF land surface scheme, which includes several updated modelling components. These are briefly described in the following subsections with the highlight of the main components that characterize ERA-Interim/Land performance.

2.2.1 Soil hydrology

A revised soil hydrology in TESSEL was proposed by van den Hurk and Viterbo (2003) for the Baltic basin. These model developments were in response to known weaknesses of the TESSEL hydrology: specifically the choice of a single global soil texture, which does not characterize different soil moisture regimes, and a Hortonian runoff scheme which produces hardly any surface runoff. Therefore, a revised formulation of the soil hydrological conductivity and diffusivity (spatially variable according to a global soil texture map) and surface runoff (based on the variable infiltration capacity approach) were operationally introduced in IFS in November 2007. Balsamo et al. (2009) verified the impact of the soil hydrological revisions from field site to global atmospheric coupled experiments and in data assimilation.

2.2.2 Snow hydrology

A fully revised snow scheme was introduced in 2009 to replace the existing scheme based on Douville et al. (1995). The snow density formulation was changed and liquid water storage in

the snow-pack was introduced, which also allows the interception of rainfall. On the radiative side, the snow albedo and the snow cover fraction have been revised and the forest albedo in presence of snow has been retuned based on MODIS satellite estimates. A detailed description of the new snow scheme and verification from field site experiments to global offline simulations are presented in Dutra et al. (2010). The results showed an improved evolution of the simulated snow-pack with positive effects on the timing of runoff and terrestrial water storage variation and a better match of the albedo to satellite products.

2.2.3 Vegetation seasonality

The Leaf Area Index (LAI), which expresses the phenological phase of vegetation (growing, mature, senescent, dormant), was kept constant in ERA-Interim and assigned by a look-up table depending on the vegetation type; thus vegetation appeared to be fully developed throughout the year. To allow for seasonality, a LAI monthly climatology based on a MODIS satellite product was implemented in IFS in November 2010. The detailed description of the LAI monthly climatology and its evaluation is provided in Boussetta et al. (2013a).

2.2.4 Bare soil evaporation

The bare soil evaporation included in the HTESSEL model in conjunction with the LAI update as reported in Balsamo et al. (2011) has been extensively evaluated by Albergel et al. (2012a) over the US. The evaluation was based on data from the Soil Climate Analysis Network (SCAN) as well as Soil Moisture and Ocean Salinity (SMOS) satellite data. The bare ground evaporation has been enhanced over deserts by adopting a lower stress threshold than for vegetation. This is in agreement with previous experimental findings (e.g. Mahfouf and Noilhan 1991) and results in a more realistic soil moisture for dry lands, as was largely confirmed by Albergel et al. (2012a).

3 Results

283 The quality of ERA-Land builds upon reduced errors in the meteorological forcing and land
284 surface modelling. In the following, selected verification results are included, showing the
285 added value of ERA-Interim/Land in reproducing the main land water reservoirs and fluxes
286 towards the atmosphere and river outlets. The two most active water reservoirs are the root-
287 zone soil moisture (here the top 1m of soil is considered) and the snow accumulated on the
288 ground. These global reservoirs in its median of the distribution calculated on the period
289 1979-2010 are shown in Fig. 2 for soil moisture SM and snow water equivalent SWE (both
290 expressed in mm of water or equivalently in kg m^{-2}). The median of the 32-year SWE valid on
291 the 15th of January is particularly adapted to show the typical values for those dates (a single
292 exceptional year with large snow accumulation leaving the median invariant).

293 The same argument is valid for mid-July SM in which a single exceptional flood will not
294 affect the median. Similarly the 95th percentile of the distribution is shown for comparison in
295 Fig. 3 to illustrate the water resources dynamical range in the past 3 decades associated with
296 snow and unsaturated soil layers and the extent and the magnitude of exceptional events can
297 be appreciated. Figs. 2 and 3 serve the purpose of illustrating the potential of the multi-
298 decadal daily land reanalysis for evaluating typical and extreme value of the global water
299 resources.

300 The evolution of ERA-Interim/Land along the 32 years of this dataset and its differences with
301 respect to ERA-Interim are illustrated in Figs. 4 and 5 for both soil moisture and snow water
302 equivalent. The stability and the differences with respect to ERA-Interim can be appreciated
303 in the plots of Figs. 3a and 4a for snow water equivalent and Figs. 4b and 5b for the top 1-m
304 soil moisture. The snow changes in Fig. 4a are mainly consequence of the new snow scheme
305 and highlight both a snow mass increase in high latitudes and a slight reduction in mid-
306 latitudes. The soil moisture presents large differences in Fig. 5b than can be attributed to the
307 soil hydrology revisions. Fig. 4 is meant to illustrate that ERA-Interim and ERA-Interim/Land
308 are significantly different throughout the 32-year period with respect to land water resources.
309 In these runs observational constraints on the snow and soil water reservoirs such as those

applied by data assimilation is totally absent, however the resulting water reservoirs and the fluxes both towards the atmosphere (heat and moisture) and the river discharges, are shown to improve with respect to the original ERA-Interim output. In the following sections a selection of results to prove the added value of ERA-Interim/Land is presented.

3.1 Land fluxes verification

The land surface fluxes resulting from the offline-driven land simulations are validated against two categories of land-controlled fluxes, the land-atmosphere turbulent heat and moisture and the river discharges.

3.1.1 Latent and Sensible heat flux

The fluxes are measured over 34 FLUXNET, CEOP and BERMS flux-towers, as listed in Table 1. Correlation, mean bias and root mean squared differences are improved using the ERA-Interim/Land surface scheme, indicating a higher skill in reproducing the land atmosphere fluxes.

A detailed evaluation of the ERA-Interim (TESSEL) and ERA-Interim/Land (HTESSEL) surface schemes in offline driven simulations for each site confirms a general improved representation of both the latent and sensible heat fluxes (Fig. 6).

An overall quantitative estimate of the improvements is reported in Table 2. Both Latent and Sensible heat fluxes indicate an average improvement of 8%, when adopting the ERA-Interim/Land surface scheme instead of the ERA-Interim surface scheme, evaluated as root-mean-square-error differences.

3.1.2 River discharge

River discharge is used here to provide an integrated evaluation of the continental water cycle for verifying improvements in the representation of land hydrology. The ERA-Interim/Land discharges are compared to those obtained from ERA-Interim by consideration of their correlation to observed GRDC monthly river discharges clustered by continent. Fig. 7 shows

335 the cumulative distribution function of the correlations between simulated and observed
336 monthly river discharges ERA-Interim/Land (blue dashed line). A general improvement over
337 ERA-Interim (red solid line) is evident since the correlations are higher at all levels in nearly
338 all cases (blue line is nearly always above the red line and area under the blue curve is
339 greater). The improvements on river discharge correlation coefficients (ERA-Interim/Land to
340 GRDC river discharge observations) are averaged on all the continental rivers indicated in
341 each panel of Fig. 6.

342 Although there is still some way to go in effectively representing river discharge in large-
343 scale land surface schemes, modelling cascades can enable bridging the ERA-Interim/Land
344 with river hydrology (Pappenberger et al, 2012). In the current evaluation what is particularly
345 encouraging is the average improvement of river discharge correlations of ERA-Interim/Land
346 over ERA-Interim occurs on all continents that therefore encompass different rivers and water
347 balance regimes.

348 ## *3.2 Land water resources verification*

349 The water reservoirs verification aims at assessing the daily performance of ERA-
350 Interim/Land in reproducing the top metre of soil water content and the snow water
351 equivalent, which are responding to the diurnal, synoptic and seasonal fluxes variations. The
352 deeper and slowly evolving soil moisture layers, such as the water table, are not considered in
353 the present verification since they are not yet properly represented in the model.

354 ### *3.2.1 Soil moisture*

355 The changes in the land surface parameterization have largely preserved the mean annual soil
356 moisture, which ranges around 0.23-0.24 m^3m^{-3} as global land average on the ERA-Interim
357 period. However the spatial variability has greatly increased with the introduction of the
358 revised soil hydrology (Balsamo et al. 2009). In order to verify the soil moisture produced by
359 the offline simulations we make use of the International Soil Moisture Network (ISMN)
360 ground-based observing networks. This has been applied by Albergel et al. (2012b) to

validate soil moisture from both ECMWF operational analysis and ERA-Interim. Offline land surface simulations were also used by Albergel et al. (2012a) to evaluate the new bare ground evaporation formulation mentioned in section 3.2. Considering the field sites of the NRCS-SCAN network (covering the US) with a fraction of bare ground greater than 0.2 (according to the model), the root mean square difference (RMSD) of soil moisture is shown to decrease from 0.118 m^3m^{-3} to 0.087 m^3m^{-3} when using the new formulation in offline experiments (and from 0.110 m^3m^{-3} to 0.088 m^3m^{-3} in operations). It also improves correlations. Fig. 8 illustrates the effect of the model changes for one site located in Utah. ERA-Interim and ERA-Interim/Land soil moisture are shown to illustrate the differences in soil moisture and the contribution of GPCP correction.

In the TESSEL formulation used in ERA-Interim, minimum values of soil moisture are limited by the wilting point of the dominant vegetation type, however ground data indicate much drier conditions, as is clearly observed from May to September 2010. The new soil hydrology and bare ground evaporation allows the model to go below this wilting point so the new analysis is in much better agreement with the observations than in ERA-Interim. The better correlations and reduced RMSD are explained by a more realistic decrease in soil moisture after a precipitation event due to its higher water holding capacity and are attributed to soil hydrology revisions (Balsamo et al. 2009, Balsamo et al. 2011, Albergel et al. 2012a).

The ability of ERA-Interim/Land and ERA-Interim to reproduce soil moisture is also presented by Fig. 9. This illustrates also the gain in skill in reproducing the observed soil moisture in dry land as a function of vegetation cover. With the RMSD being positive definite and calculated against in-situ soil moisture observation, the differences between RMSD between ERA-Interim/Land and ERA-Interim are measuring improvements realized by ERA-Interim/Land. The RMSD difference is calculated for several vegetation fractions and the improvement is shown to be larger on points with sizeable bare soil. This is a demonstration that the enhanced match to the observed soil moisture is indeed the results of the bare soil evaporation revision as detailed in Albergel et al. (2012a).

The correlation of ERA-Interim/Land soil moisture with the various observed soil moisture networks varies depending on the network selected (Fig. 10). This variation is similar in manner to that seen with ERA-Interim but the correlation is not significantly improved. However, in Fig. 11 a Taylor diagram is used to illustrate a more detailed statistical comparison of ERA-Interim/Land (in red), ERA-Interim (in blue), and in situ observations for 2010. In Fig. 11 the distance to the point marked "In situ" has been reduced with the ERA-Interim/Land, which indicates more realistic soil moisture variability (better reproduction of the standard deviation of observations).

3.2.2 Snow

The verification of snowfields considers two different observational datasets to evaluate the snow evolution in ERA-Interim and ERA-Interim/Land: (i) the SYNOP daily snow depth and (ii) datasets from the former Union of Soviet Socialist Republics (USSR). The 1979-1993 former USSR dataset was used in Brun et al. (2013) to evaluate simulated snow properties, such as density, which is not routinely measured at SYNOP stations. Dutra et al. (2010) attributed the largest improvement in the new snow scheme to the snow density representation. This is confirmed by the verification results on a large number of sites where snow density was measured for the typical Northern latitudes snow season (October to June) average for 1979-1993 period (Balsamo et al. 2012). In ERA-Interim, as well as ERA-Interim-Land, the snow density is not at all constrained by data assimilation due to a lack of observations and therefore it relies solely on the capacity of the land surface model to represent the seasonal evolution, from about 100 kg/m^3 at the beginning of the winter season to more than 300 kg/m^3 towards the end of the snow season.

Simulations of snow water equivalent with and without the GPCP V2.1 rescaling have been evaluated against observations, which are available from 1979 to 1993 over the USSR. A significantly lower bias in this case is obtained without the GPCP rescaling (9.7 mm versus 33.8 mm) confirming the general difficulties in measuring snowfall with gauges. This means

414 that ERA-Interim/GPCP-rescaled is not always beneficial and outperforming ERA-Interim
415 precipitation forecasts at each single location (this is not an easy achievable target).

416 In this case, ERA-Interim original snowfall, without bias correction, lead to higher skill in
417 simulating snow accumulations in this particular verification area and to accurate snow
418 accumulations as confirmed by Brun et al. (2013). Given the difficulty in applying
419 precipitation corrections only partially at this stage it is only possible to document this
420 exception and limitation of the bias correction method and/or the used precipitation datasets.

421 The capacity of detecting the presence of snow on the ground in ERA-Interim/Land is
422 examined using the SYNOP network in more recent years considering two snow seasons
423 2005/06 and 2009/10. Two scores are adopted:

424 (i) SDR = Snow Detection Rate (SDR=1 being the best value) measures the fraction of
425 times the snow fields rightly detect the presence of snow divided by the number of
426 times the SYNOP observation detects snow presence, and

427 (ii) FCA = Fraction of Correct Accuracy (FCA=1 being the best value) measures the
428 fraction of times the snow fields rightly detect the presence or absence of snow in
429 agreement with the SYNOP message (divided by the total amount of stations).

430 The ability of two offline simulations driven by ERA-Interim to represent snow cover was
431 assessed for ERA-Interim surface scheme (control) and ERA-Interim-Land (experiment)
432 offline experiments. Fig. 12 (left) shows the Snow Detection Rate (SDR) function of the snow
433 cover for both ERA-Interim/Land and ERA-Interim configurations and Fig. 12 (right)
434 presents the cumulative distribution function of the SDR for two periods, 2005/06 and
435 2009/10. SDR is much better with ERA-Interim/Land than with ERA-Interim scheme for both
436 periods. For instance, considering the 2005/06 period, while 50% of the SDR is above the
437 value 0.49 for ERA-Interim scheme, 50% of the SDR is above 0.70 for ERA-Interim/Land.
438 Fraction of Correct Accuracy (FCA) are 80 and 86 in 2005/06, 76 and 83 in 2009/10 for
439 ERA-Interim and ERA-Interim/Land surface schemes respectively (Fig. 12). This index is a

robust indicator and is more resilient to model biases compared to SDR, which in case snow abundance may favour a biased snow scheme. The MODIS land surface albedo is used to verify the ERA-Interim/Land, particularly in the snow representation in forest areas (Fig. 13) in Northern Canada and Siberia, where conventional SYNOP observations are generally less informative. Fig. 12c points to a substantially reduced albedo bias in the ERA-Interim/Land attributed to the snow scheme revision described in Dutra et al. (2010) and in particular at the snow-vegetation albedo retuning.

4 Discussion

Dedicated land surface reanalyses, such as the ERA-Interim/Land described and evaluated here, are becoming established added-value products within the reanalysis efforts worldwide (Dee et al. 2013). They allow computationally effective testing of new land surface developments, including improvements to the process representation and parameterisation of the hydrological and biogeochemical cycles that contribute to a fast-track land surface model developments as identified by van den Hurk et al. (2012). Future research into improved representation of the land surface is high priority, and work already underway in this area includes land carbon exchanges (Boussetta et al. 2013b), vegetation inter-annual variability, and hydrological applications such as global water-bodies reanalysis (e.g. Balsamo et al. 2012) and used in applications such as global flood risk assessment (e.g. Pappenberger et al. 2012). More sophisticated rescaling methods (e.g. Weedon et al. 2011, 2014) are envisaged to bias correct the meteorological forcing and to permit a high resolution downscaling of land reanalysis. In addition, consideration of land surface parameterisation uncertainty could be used to further improve predictive skill (e.g. Cloke et al, 2011).

Important developments with advanced land data assimilation methods such as the Extended or Ensemble Kalman Filters (Reichle et al. 2014, de Rosnay et al. 2013b, Drusch et al. 2009) can be combined with offline surface simulations. The experimental equivalence of offline

and atmospheric coupled land data assimilation (Balsamo et al. 2007, Mahfouf et al. 2008) offers also in this case a two orders of magnitude computational saving. This is expected to provide a fast land surface reanalysis as envisaged within the EU-funded ERA-CLIM project, moreover it can open up new possibilities of considering more advanced data assimilation schemes (e.g. Fowler and van Leeuwen, 2012), especially designed for non-linear systems.

The skill of an ERA-Interim/Land variant (with no precipitation readjustment) together with other model-based and remote-sensing datasets for the detection of soil moisture climate trends in the past 30 years is evaluated in Albergel et al. (2013). This study, using the methodology described in this paper, represents an attempt to gain insights on soil water reservoirs and its evolution in response to natural and anthropogenic forcing.

5 Conclusions

This paper documents the configuration and the performance of the ERA-Interim/Land reanalysis in reconstructing the land surface state over the past 3 decades. The ERA-Interim/Land is produced from the ERA-Interim meteorological forcing offline land-surface model simulations. In this paper it has been demonstrated that the ERA-Interim/Land dedicated land surface reanalysis is of added value over the standard land component for the ERA-Interim reanalysis product. The ERA-Interim/Land runs are an integral part of the ERA-Interim on-going research efforts and respond to the need to re-actualize the land surface initial conditions of ERA-Interim, following several model parameterization improvements. The newly produced land-surface estimates benefit from the latest land surface hydrology schemes used operationally at ECMWF for weather, monthly, and seasonal forecasts. The ERA-Interim/Land added value components encompass soil, snow and vegetation description upgrades, as well as a bias correction of the ERA-Interim monthly-accumulated precipitation based on GPCP v.2.1. In the Northern hemisphere the precipitation correction is shown to be

491 effective in reducing the bias over US and rather neutral over Eurasia, while in the tropical
492 land benefits are visible in the river discharge.

493 The new land surface reanalysis has been verified against several datasets for the main water
494 reservoirs, snow and soil moisture, together with the energy and water fluxes that have direct
495 impact on the atmosphere. The verification makes use of both in-situ observations and remote
496 sensing products. Improved match to observations largely attributed to the land surface
497 revisions in the latest ECMWF land surface scheme, is found in the latent and sensible heat
498 fluxes and in soil moisture and snow.

499 The water balance is verified with the observed river discharge from the GRDC river network
500 showing an enhanced correlation to the observations with respect to ERA-Interim as
501 combined effect of the GPCP precipitation correction and the land surface improvements.

502 While river discharges verification is not enough for a global water balance assessment the
503 results from the verification of evaporation fluxes (the other main outgoing land water flux)
504 and of the two main water reservoirs in the soil and snow-pack permit to qualify the ERA-
505 Interim/Land enhanced accuracy as genuine. When water fluxes and water storages terms
506 show consistent indication of improvements there are in fact good grounds to believe that the
507 parameterization changes are real added value and not the result of compensation.

508 Finally, the impact of adopting ERA-Interim/Land as initial condition in retrospective
509 forecasts has also been verified with a generally positive effect of the new land initial
510 condition, more evident in longer lead times of the forecasts (Balsamo et al., 2012).
511 Interannual variability of vegetation state (Leaf-Area-Index) is currently studied at ECMWF
512 in the framework of the EU-FP7 project IMAGINES (http://fp7-imagines.eu) and it is not yet
513 implemented in ERA-Interim/Land.

514 The ERA-Interim/Land dataset has been used operationally at ECMWF since 2010 for the
515 initialization of the past reforecasts needed for the monthly forecasting (Vitart et al. 2008) and
516 the seasonal prediction systems (Molteni et al., 2011).

Ongoing research effort includes the extension of this dataset beyond 2010 using a different dataset for precipitation based on the latest GPCC collections (Weedon, et al. 2014) and application of the described methodology to future ECMWF reanalyses (Dee et al. 2013).

6 Dataset access

The ERA-Interim/Land dataset is freely available and it can be downloaded from:

http://apps.ecmwf.int/datasets/

Acknowledgment- Authors thank R. Riddaway from ECMWF for his valuable comments on the text and C. O'Sullivan and A. Bowen are thanked for their help in improving the figures. Eric Brun is thanked for his interest and precious advise on ERA-Interim/Land snow verification. This work used eddy covariance data acquired by the FLUXNET community that is greatly acknowledged. TU Wien provided the ISMN data for soil moisture verification and we thank them for their important effort. We thank the GRDC for data provision of global river discharge. The ECMWF User-Support team is acknowledged for making the data easily accessible.

References

Albergel, C., W. Dorigo, R. H. Reichle, G. Balsamo, P. de Rosnay, J. Munoz-Sabater, L. Isaksen, R. de Jeu, and W. Wagner: Skill and global trend analysis of soil moisture from reanalyses and microwave remote sensing, *J. of Hydromet.*, **14**, 1259–1277, doi: 10.1175/JHM-D-12-0161.1, 2013.

Albergel, C., Balsamo, G., de Rosnay, P., Muñoz-Sabater, J. & Boussetta, S. A bare ground evaporation revision in the ECMWF land-surface scheme: evaluation of its impact using ground soil moisture and satellite microwave data. *Hydrol. Earth Syst. Sci.*, **16**, 3607-3620, 2012a.

Albergel, C., P. de Rosnay, G. Balsamo, L. Isaksen and J. Muñoz Sabater: Soil moisture analyses at ECMWF: evaluation using global ground-based in-situ observations, *J. Hydromet.*, **13**, 1442–1460, 2012b., also as *ECMWF. Tech. Memo.* No. 651, 2012b.

Albergel C., P. de Rosnay, C. Gruhier, J. Muñoz-Sabater, S. Hasenauer, L. Isaksen, Y. Kerr and W. Wagner: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations, *Remote Sens. Environ.*, **118**, 215-226, doi:10.1016/j.rse.2011.11.017, 2012c.

Albergel, C., Rüdiger, C., Carrer, D., Calvet, J.-C. Fritz, N. Naeimi, V. Bartalis, Z. and Hasenauer S.: An evaluation of ASCAT surface soil moisture products with in situ observations in Southwestern France, *Hydrol. Earth Syst. Sci.*, **13**, 115–124, doi:10.5194/hess-13-115-2009, 2009.

Adler, R.F., G.J. Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, P. Arkin and E.J. Nelkin: The Version 2.1 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979 -Present), *J. Hydrometeor.*, **4**, 1147-1167, 2003.

Baldocchi, D., E. Falge, L. Gu, R. Olson, D. Hollinger, S. Running, P. Anthoni, C. Bernhofer, K. Davis, R. Evans, J. Fuentes, A. Goldstein, G. Katul, B. Law, X. Lee, Y. Malhi, T. Meyers, W. Munger, W. Oechel, K. T. Paw, K. Pilegaard, H. P. Schmid, R. Valentini, S. Verma, T. Vesala, K. Wilson, and S. Wofsy: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, *Bull. Am. Meteor. Soc*., **82**, 2415-2434, doi:10.1175/1520-0477, 2001.

Balsamo, G., P. Viterbo, A. Beljaars, B. van den Hurk, M. Hirschi, A.K. Betts and K. Scipal:.A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System, *J. Hydrometeor*., **10**, 623–643, 2009.

Balsamo, G., S. Boussetta, P. Lopez and L. Ferranti: Evaluation of ERA-Interim and ERA-Interim-GPCP-rescaled precipitation over the U.S.A., *ERA Report Series*, No. **5**, pp10, 2010a.

Balsamo, G., F. Pappenberger, E. Dutra, P. Viterbo and B. van den Hurk: A revised land hydrology in the ECMWF model: A step towards daily water flux prediction in a fully-closed water cycle, Special issue on large scale hydrology of *Hydrol. Process*., **25**, 1046-1054, doi:10.1002/hyp.7808, 2010b.

Balsamo, G., S. Boussetta, E. Dutra, A. Beljaars, P. Viterbo, B. Van den Hurk: Evolution of land surface processes in the IFS, *ECMWF Newsletter*, **127**, 17-22, 2011.

Balsamo, G., R. Salgado, E. Dutra, S. Boussetta, T. Stockdale and M. Potes:. On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model, *Tellus-A*, **64**, 15829, doi: 10.3402/tellusa.v64i0.15829, 2012.

Berrisford, P., P. Kallberg, S. Kobayashi, D. Dee, S. Uppala, A.J. Simmons, P. Poli and H. Sato: Atmospheric conservation properties in ERA-Interim, *Quart. J. Roy. Meteor. Soc*., **137**, 1381-1399, doi: 10.1002/qj.864, 2011.

Betts, A.K., J.H. Ball, A.G. Barr, T.A. Black, J.H. McCaughey and P. Viterbo: Assessing landsurface-atmosphere coupling in the ERA-40 re-analysis with boreal forest data, *Agric. Forest Meteor.*, **140**, 365-382, doi:10.1016/j.agrformet.2006.08.009, 2006.

Boussetta, S., G. Balsamo, A. Beljaars and J. Jarlan: Impact of a satellite-derived Leaf Area Index monthly climatology in a global Numerical Weather Prediction model, *Int. J. Remote Sensing*, **34** (9-10), 3520-3542, 2013.

Boussetta, S., G. Balsamo, A. Beljaars, A. Agusti-Panareda, J.-C. Calvet, C. Jacobs, B. van den Hurk, P. Viterbo, S. Lafont, E. Dutra, L. Jarlan, M. Balzarolo, D. Papale and G. van der Werf: Natural land carbon exchanges in the ECMWF Integrated Forecasting system: Implementation and offline validation, *J. Geophys. Res.*,**118**, 1-24, 2013.

Brun, E., V. Vionnet, B. Decharme, Y. Peings, R. Valette, F. Karbou and S. Morin: Simulation of northern Eurasian local snow depth, mass and density using a detailed snowpack model and meteorological reanalyses, *J. Hydrometeor.*,**14***, 203-219,* 2013.

Cloke, H. L., A. Weisheimer, F. Pappenberger: Representing uncertainty in land surface hydrology: fully coupled simulations with the ECMWF land surface scheme. *Proceeding of ECMWF, WMO/WGNE, WMO/THORPEX and WCRP Workshop on "Representing model uncertainty and error in numerical weather and climate prediction models"* ECMWF, Reading, UK, June 2011.

Daly, C., R.P. Neilson and D.L. Phillips: A statistical-topographic model for mapping climatological precipitation over mountainous terrain, *J. Appl. Meteor*., **33**, 140-158, 1994.

Daly, C., G.H. Taylor, W.P. Gibson, T.W. Parzybok, G.L. Johnson and P. Pasteris: High-quality spatial climate data sets for the United States and beyond, *Trans. Amer. Society Agric*. Eng., **43**, 1957-1962, 2001.

Decker, M., M.A. Brunke, Z. Wang, K. Sakaguchi, X. Zeng and M.G. Bosilovich: Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF Using Flux Tower Observations, *J. Climate*, **25**, 1916-1944, doi: 10.1175/JCLI-D-11-00004.1, 2012.

Dee, D.P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kallberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J. J. Morcrette, B. K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, F. Vitart: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quart. J. Roy. Meteor. Soc*., **137**, 553-597, doi: 10.1002/qj.828, 2011.

615 Dee, D., M. Balmaseda, G. Balsamo, R. Engelen, and A. Simmons: Toward a consistent
616 reanalysis of the climate system, *Bull. Amer. Meteor. Soc*., *(in press),* 10.1175/BAMS-D-13-
617 00043.1, 2013.

618 Dirmeyer, P.A.: A history and review of the Global Soil Wetness Project (GSWP), *J.*
619 *Hydrometeor.*, **12**, 729-749, doi: 10.1175/JHM-D-10-05010.1, 2011.

620 Dirmeyer, P.A., X. Gao and T. Oki: The second Global Soil Wetness Project - Science and
621 implementation plan, *IGPO Int. GEWEX Project Office Publ. Series* **37**, 65 pp., 2002.

622 Dirmeyer, P.A., X. Gao, M. Zhao, Z. Guo, T. Oki and N. Hanasaki: GSWP-2: Multimodel
623 analysis and implications for our perception of the land surface, *Bull. Amer. Meteor. Soc*., **87**,
624 1381–1397, 2006.

625 Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A.,
626 Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A., and Jackson, T.: The International
627 Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements,
628 Hydrol. Earth Syst. Sci., 15, 1675-1698, doi:10.5194/hess-15-1675-2011, 2011.

629 Dorigo W.A., Xaver A., Vreugdenhil M., Gruber A., Hegyiová A., Sanchis-Dufau A.D.,
630 Wagner W., Drusch M. (2013) Global automated quality control of in-situ soil moisture data
631 from the International Soil Moisture Network. Vadose Zone Journal 12, 3.

632 Douville, H., J.F. Royer and J.-F. Mahfouf: A new snow parameterization for the Meteo-
633 France climate model .1. Validation in stand-alone experiments, *Climate Dyn.*, **12**, 21–35,
634 1995.

635 Drusch, M., D. Vasiljevic, P. Viterbo: ECMWF's Global Snow Analysis: Assessment and
636 Revision Based on Satellite Observations, *J. Appl. Meteorol.*, **43**, 1282-1294, 2004.

637 Drusch M., Scipal K., de Rosnay P., Balsamo G., Andersson E., Bougeault P., Viterbo P.,
638 Towards a Kalman Filter based soil moisture analysis system for the operational ECMWF
639 Integrated Forecast System. Geophys. Res. Lett., 36, 10, L10401,
640 doi:10.1029/2009GL037716, 2009.

641 Dutra, E., G. Balsamo, P. Viterbo, P. Miranda, A. Beljaars, C. Schär and K. Elder: An
642 improved snow scheme for the ECMWF land surface model: description and offline
643 validation, *J. Hydrometeor.*, **11**, 899–916, 2010. Doi: 10.1175/JHM-D-11-072.1

644 Dutra, E., P. Viterbo, P. M. A. Miranda and G. Balsamo: Complexity of snow schemes in a
645 climate model and its impact on surface energy and hydrology, *J. Hydrometeor*., **13**, 521-538,
646 doi: 10.1175/jhm-d-11-072, 2012.

647 Fowler, A., and P. van Leeuwen: Measures of observation impact in non-Gaussian data
648 assimilation. Tellus *A,* **64**. doi:10.3402/ tellusa.v64i0.17192, 2012.

649 Hazeleger, W., C. Severijns, T. Semmler, S. Ştefănescu, S. Yang, X. Wang, K. Wyser, E.
650 Dutra, J. M. Baldasano, R. Bintanja, P. Bougeault, R. Caballero, A. M. L. Ekman, J. H.
651 Christensen, B. van den Hurk, P. Jimenez, C. Jones, P. Kållberg, T. Koenigk, R. McGrath, P.
652 Miranda, T. Van Noije, T. Palmer, J. A. Parodi, T. Schmith, F. Selten, T. Storelvmo, A. Sterl,
653 H. Tapamo, M Vancoppenolle, P. Viterbo, U. Willénand: EC-Earth: A seamless earth-system
654 prediction approach in action, *Bull. Amer. Meteor. Soc*., **91**, 1357-1363, doi:
655 10.1175/2010BAMS2877, 2010.

Huffman G. J., R.F. Adler, D.T. Bolvin and G. Gu: Improving the global precipitation record: GPCP Version 2.1, *Geophys. Res. Lett*., **36**, L17808, doi:10.1029/2009GL040000, 2009.

Kållberg, P.: Forecast drift in ERA-Interim, *ERA Report Series,* No.**10***,* pp. 9. ECMWF, Reading, UK. 2011.

Koster, R.D., Z. Guo, P.A. Dirmeyer, G. Bonan, E. Chan, P. Cox, H. Davies, C.T. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, C.H. Lu, S. Malyshev, B. McAvaney, K. Mitchell, D. Mocko, T. Oki, K.W. Oleson, A. Pitman, Y.C. Sud, C.M. Taylor, D. Verseghy, R. Vasic, Y. Xue and T. Yamada: GLACE: The Global Land-Atmosphere Coupling Experiment. Part I: Overview, *J. Hydrometeor*., **7**, 590-610, 2006.

Koster, R., S. Mahanama, T. Yamada, G. Balsamo, M. Boisserie, P. Dirmeyer, F. Doblas-Reyes, T. Gordon, Z. Guo, J.-H. Jeong, D. Lawrence, Z. Li, L. Luo, S. Malyshev, W. Merryfield, S.I. Seneviratne, T. Stanelle, B. van den Hurk, F. Vitart and E.F. Wood: The contribution of land surface initialization to subseasonal forecast skill: First results from the GLACE-2 Project, *Geophys. Res. Lett*., **37**, L02402, 2009GL041677R, 2009.

Koster, S.P.P. Mahanama, T.J. Yamada, G. Balsamo, A.A. Berg, M. Boisserie, P.A. Dirmeyer, F.J. Doblas-Reyes, G. Drewitt, C.T. Gordon, Z. Guo, J.-H. Jeong, W.-S. Lee, Z. Li, L. Luo, S. Malyshev, W.J. Merryfield, S.I. Seneviratne, T. Stanelle, B.J.J.M. van den Hurk, F. Vitart and E.F. Wood: The second phase of the global land-atmosphere coupling experiment: soil moisture contributions to subseasonal forecast skill, *J. Hydrometeor*., **12**, 805–822. doi: 10.1175/ 2011JHM1365.1, 2011.

Lin Y. and K.E. Mitchell: The NCEP stage II/IV hourly precipitation analyses: Development and applications, *Preprints, 19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., CD-ROM, 1.2, 2005.

Lopez, P. and P. Bauer: 1D+4DVAR assimilation of NCEP stage-IV radar and gauge hourly precipitation data at ECMWF, *Mon. Wea. Rev*., **135**, 2506-2524, 2007.

Mahfouf, J.-F. and J. Noilhan: Comparative study of various formulations of evaporation from bare soil using in situ data, *J. Appl. Meteor*., **30**, 351–362, 1991.

Molteni, F., T. Stockdale, M. Balmaseda, G. Balsamo, R. Buizza, L. Ferranti, L. Magnusson, K. Mogensen, T. Palmer and F. Vitart: The new ECMWF seasonal forecast system (System 4). *ECMWF Tech. Memo.* No. 656, 2011.

Oki, T. and S. Kanae: Global Hydrological Cycles and World Water Resources, *Science*, ***313*** no. 5790 pp. 1068-1072, doi: 10.1126/science.1128845, 2006.

Pappenberger, F., H. Cloke, G. Balsamo, T. Ngo-Duc and T. Oki: Global runoff routing with the hydrological component of the ECMWF NWP system, *Int. J. Climatol*., **30**, 2155-2174, doi:10.1002/joc.2028, 2009.

Pappenberger, F., E. Dutra, F. Wetterhall and H. Cloke : Deriving global flood hazard maps of fluvial floods through a physical model cascade, *Hydrol. Earth Syst. Sci. Discuss*., **9**, 6615-6647, doi:10.5194/hessd-9-6615-2012, 2012.

Polcher, J., B. McAvaney, P. Viterbo, M.-A. Gaertner, A. Hahamann, J.-F. Mahfouf, J. Noilhan, T. Phillips, A. Pitman, C. Schlosser, J.-P. Schulz, B. Timbal, D. Verseghy and Y.

696 Xue: A proposal for a general interface between land surface schemes and general circulation
697 models, *Global and Planetary Change*, **19**, 261-276, 1998.

698 Reichle, R.H., R.D. Koster, G.J.M. De Lannoy, B.A. Forman, Q. Liu, S.P.P. Mahanama and
699 A. Toure: Assessment and enhancement of MERRA land surface hydrology estimates, *J.*
700 *Climate*, **24**, 6322-6338, doi:10.1175/JCLI-D-10-05033.1, 2011.

701 Reichle, R. H., G. J. M. De Lannoy, B. A. Forman, C. S. Draper, and Q. Liu: Connecting
702 Satellite Observations with Water Cycle Variables through Land Data Assimilation:
703 Examples Using the NASA GEOS-5 LDAS, *Surveys in Geophysics*, in press,
704 doi:10.1007/s10712-013-9220-8, 2014.

705 Rienecker, M. M., M. J. Suarez, R. Gelaro, R. Todling, Julio B., E. Liu, M. G Bosilovich, S.
706 D. Schubert, L. Takacs, G.-K. Kim, S. Bloom, J. Chen, D. Collins, A. Conaty, A. da Silva, W.
707 Gu, J. Joiner, R. D. Koster, R. Lucchesi, A. Molod, T. Owens, S. Pawson, P. Pegion, C. R.
708 Redder, R. Reichle, F. R. Robertson, A. G. Ruddick, M. Sienkiewicz, J. Woollen: MERRA -
709 NASA's modern-era retrospective analysis for research and applications, *J. Climate*, **24**, 3624-
710 3648, doi: 10.1175/JCLI-D-11-00015.1, 2011.

711 Román, M. O., C. B. Schaaf, C. E. Woodcock, A. H. Strahler, X. Yang, R. H. Braswell, P. S.
712 Curtis, K. J. Davis, D. Dragoni, M. L. Goulden, L. Gu, D. Y. Hollinger, T. E. Kolb, T. P.
713 Meyers, J. W. Munger, J. L. Privette, A. D. Richardson, T. B. Wilson, S. C. Wofsy: The
714 MODIS (Collection V005) BRDF/albedo product: Assessment of spatial representativeness
715 over forested landscapes, *Remote Sens. Environ.*, **113**, 2476-2498, 2009.

716 de Rosnay, P., G. Balsamo, C. Albergel J. Muñoz-Sabater and L. Isaksen: Initialisation of
717 land surface variables for Numerical Weather Prediction, *Surveys in Geophysics*, in press, doi:
718 10.1007/s10712-012-9207-x, 2013a.

719 de Rosnay, P., M. Drusch, D. Vasiljevic, G. Balsamo, C. Albergel and L. Isaksen: A
720 simplified Extended Kalman Filter for the global operational soil moisture analysis at
721 ECMWF, *Quart. J. Roy. Meteor. Soc.*, **139**,1199-1213, doi: 10.1002/qj.2023, 2013b.

722 Rubel, F. and K. Brugger: 3-hourly quantitative precipitation estimation over Central and
723 Northern Europe from rain gauge and radar data, *Atmospheric Research*, **94**, 544-554,
724 doi:10.1016/j.atmosres.2009.05.005, 2009.

725 Salomon, J.G., C.B. Schaaf, A.H. Strahler, F. Gao and Y.F. Jin: Validation of the MODIS
726 bidirectional reflectance distribution function and albedo retrievals using combined
727 observations from the Aqua and Terra platforms, *Ieee T. Geosci. Remote*, **44**, 1555-1565,
728 2006.

729 Seneviratne, S.I. and R.D. Koster: A Revised Framework for Analyzing Soil Moisture
730 Memory in Climate Data: Derivation and Interpretation. *J. Hydrometeor.*, **13**, 404-412, 2012.

731 Simmons, A.J., K.M. Willett, P.D. Jones, P.W. Thorne and D.P. Dee: Low-frequency
732 variations in surface atmospheric humidity, temperature and precipitation: Inferences from
733 reanalyses and monthly gridded observational datasets, J. *Geophys. Res.*, **115**, 1-21, doi:
734 10.1029/2009JD012442, 2010.

735 Szczypta, C., J.-C. Calvet, C. Albergel, G. Balsamo, S. Boussetta, D. Carrer, S. Lafont and C.
736 Meurey: Verification of the new ECMWF ERA-Interim reanalysis over France, *Hydrol. Earth*
737 *Syst. Sci*., **15**, 647-666, doi:10.5194/hess-15-647-2011, 2011.

738 van den Hurk, B., P. Viterbo, A. Beljaars and A.K. Betts: Offline validation of the ERA-40
739 surface scheme, *ECMWF Tech. Memo*. No. 295, 2000.

740 van den Hurk, B. and P. Viterbo: The Torne-Kalix PILPS 2(e) experiment as a test bed for
741 modifications to the ECMWF land surface scheme, *Global Planet. Change*, **38**, 165–173,
742 2003.

743 Vitart, F., R. Buizza, M. Alonso Balmaseda, G. Balsamo, J.-R. Bidlot, A. Bonet, M. Fuentes,
744 A. Hofstadler, F. Molteni and T. Palmer: The new VAREPS-monthly forecasting system: a
745 first step towards seamless prediction, *Quart. J. Roy. Meteor. Soc*., **134**, 1789-1799, 2008.

746 Wang, Z. and X. B. Zeng: Evaluation of Snow Albedo in Land Models for Weather and
747 Climate Studies, *J Appl. Meteor. Clim.*, **49**, 363-380, doi: 10.1175/2009jamc2134.1, 2010.

748 Weedon, G.P., S. Gomes, P. Viterbo, W.J. Shuttleworth, E. Blyth, H. Österle, J. C. Adam, N.
749 Bellouin, O. Boucher and M. Best: Creation of the WATCH Forcing Data and its use to
750 assess global and regional reference crop evaporation over land during the twentieth century,
751 *J. Hydrometeor*., **12**, 823–848, 2011.

752 Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best and P.-L. Vidale: The WFDEI
753 meteorological forcing dataset: WATCH Forcing Data methodology applied to ERA-Interim
754 reanalysis data. *Water Resour. Res.* (*submitted*), 2014.

755 Weisheimer, A.: The contribution of the land surface to predictability in the ECMWF
756 seasonal prediction system: The European summer 2003 case, *ECMWF/GLASS Workshop on*
757 *Land Surface Modelling and Data Assimilation and the Implications for Predictability*, 9–12
758 November 2009, ECMWF, Reading, UK, 125-138. 2010.

759 Su, Z., Wen, J., Dente, L., van der Velde, R., Wang, L., Ma, Y., Yang, K., & Hu, Z.: The
760 Tibetan Plateau observatory of plateau scale soil moisture and soil temperature (Tibet-Obs)
761 for quantifying uncertainties in coarse resolution satellite and model products. *Hydrol. Earth*
762 *Syst. Sci*., **15**, 2303-2316, 2011.

763 Zhou, L., R. E. Dickinson, Y. Tian, X. Zeng, Y. Dai, Z.-L. Yang, C. B. Schaaf, F. Gao, Y. Jin,
764 A. Strahler, R. B. Myneni, H. Yu1, W. Wu1, M. Shaikh: Comparison of seasonal and spatial
765 variations of albedos from Moderate-Resolution Imaging Spectroradiometer (MODIS) and
766 Common Land Model, *J. Geophys. Res.*, **108**, D15, 4488, doi: 10.1029/2002jd003326, 2003.

767

Tables

Table 1: List of sites used for the verification of the simulated fluxes, where the biome types are: deciduous broadleaf forest (DBF), evergreen broadleaf forest (EBF), deciduous needle-leaf forest (DNF), evergreen needle-leaf forest (ENF), mixed forest (MF), woody savannahs (WSA), grasslands (GRA), crops (CRO), wetlands (WET).

N	Site	Lat	Lon	Veg Type	N	Site	Lat	Lon	Veg Type
1	sk-oa	53.63	-106.20	DBF	**18**	it-ro2	42.39	11.92	DBF
2	sk-obs	53.99	-105.12	ENF/WET	**19**	nl-ca1	51.97	4.93	GRA
3	brasilia	-15.93	-47.92	WSA/GRA /SH	**20**	nl-haa	52.00	4.81	GRA
4	at-neu	47.12	11.32	GRA	**21**	nl-hor	52.03	5.07	GRA
5	ca-mer	45.41	-75.52	WET	**22**	nl-loo	52.17	5.74	ENF
6	ca-qfo	49.69	-74.34	ENF	**23**	ru-fyo	56.46	32.92	ENF
7	ca-sf1	54.49	-105.82	ENF	**24**	ru-ha1	54.73	90.00	GRA
8	ca-sf2	54.25	-105.88	ENF	**25**	ru-ha3	54.70	89.08	GRA
9	ch-oe1	47.29	7.73	GRA	**26**	se-sk2	60.13	17.84	ENF
10	fi-hyy	61.85	24.29	ENF	**27**	us-arm	36.61	-97.49	CRO
11	fr-hes	48.67	7.06	DBF	**28**	us-bar	44.06	-71.29	DBF
12	fr-lbr	44.72	-0.77	ENF	**29**	us-ha1	42.54	-72.17	DBF
13	il-yat	31.34	35.05	ENF	**30**	us-mms	39.32	-86.41	DBF
14	it-amp	41.90	13.61	GRA	**31**	us-syv	46.24	-89.35	MF
15	it-cpz	41.71	12.38	EBF	**32**	us-ton	38.43	-120.97	MF/WSA
16	it-mbo	46.02	11.05	GRA	**33**	us-var	38.41	-120.95	GRA
17	it-ro1	42.41	11.93	DBF	**34**	us-wtr	45.81	-90.08	DBF

Table 2: Summary of mean latent heat (LE) and sensible heat (H) statistics averaged over the 34 sites (units of W/m^2). Mean correlations R of model fluxes to observations include a 95% Confidence Interval (CI) calculated using a Fisher Z-transform.

Model	LE rmse	LE bias	LE mean R	H rmse	H bias	H mean R
ERA-Interim Land (HTESSEL)	25.14	16.01	0.84 (±0.10)	20.14	-4.87	0.84 (±0.10)
ERA-Interim (TESSEL) scheme	30.42	21.58	0.81 (±0.12)	24.64	-8.90	0.78 (±0.13)

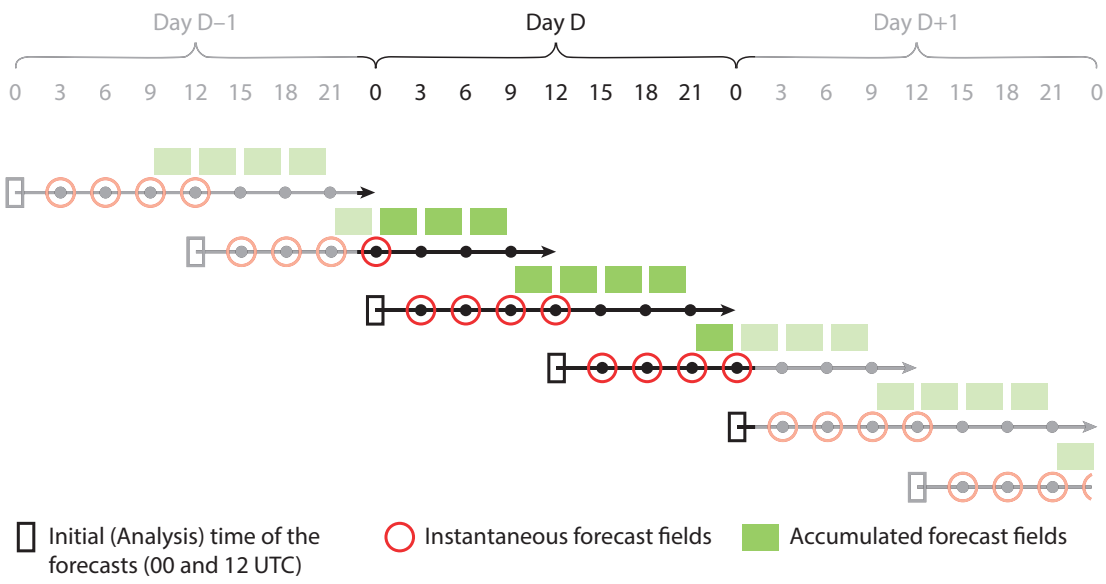
788
789 **Table 3: Comparison of surface soil moisture with in situ observations for ERA-**
790 **Interim/Land [Italic, bold] and ERA-Interim in 2010. Mean correlations (R), bias (in**
791 **situ measurements minus products) root mean square differences (RMSD), normalized**
792 **standard deviation (SDV) and the centred RMSD model and in situ patterns,**
793 **normalized by the in situ standard deviation are given for each network. Scores are**
794 **given for significant correlations with p-values <0.05. For each R estimate a 95%**
795 **Confidence Interval (CI) was calculated using a Fisher Z-transform.**
796

Network (N stations with significant R)	Mean R (95%CI)	Mean Bias (m3m-3)	Mean RMSD (m3m-3)	Mean SDV (σmodel/σobs.)	Mean E
AMMA, W. Africa (3)	***0.63(±0.06)***	***-0.060***	***0.082***	***2.67***	***2.20***
Pellarin et al., 2009	0.61(±0.07)	-0.153	0.154	0.69	0.85
OZNET, Australia (36)	***0.79(±0.05)***	***-0.112***	***0.131***	***1.01***	***0.90***
Smith et al., 2012	0.78(±0.05)	-0.078	0.106	0.55	0.97
SMOSMANIA, France (12)	***0.83(±0.04)***	***-0.080***	***0.108***	***0.83***	***0.95***
Albergel et al., 2008	0.82(±0.05)	-0.037	0.099	0.41	1.20
REMEDHUS, Spain (17)	***0.76(±0.04)***	***-0.152***	***0.175***	***1.57***	***1.40***
Ceballos et al., 2005	0.79(±0.04)	-0.110	0.135	0.84	1.25
SCAN, USA (119)	***0.64(±0.07)***	***-0.078***	***0.130***	***0.95***	***1.48***
Schaefer and Paetzold, 2010	0.62(±0.07)	-0.063	0.110	0.54	1.28
SNOTEL, USA (193)	***0.62(±0.10)***	***-0.045***	***0.115***	***0.78***	***1.27***
Schaefer and Paetzold, 2010	0.69(±0.08)	-0.088	0.123	0.44	1.03

797
798

799 **Figures**

800



801

802 **Fig. 1**: Schematic representation of the ERA-Interim meteorological forecasts concatenation
803 for the creation of the 3-hourly forcing time-series used in ERA-Interim/Land for a given day.
804 Orange circle indicate instantaneous variables valid at their timestamp: 10m temperature,
805 humidity, wind speed, and surface pressure. Green boxes indicate fluxes valid on the
806 accumulation period: surface incoming short-wave and long-wave radiation, rainfall,
807 snowfall.

808

809

810

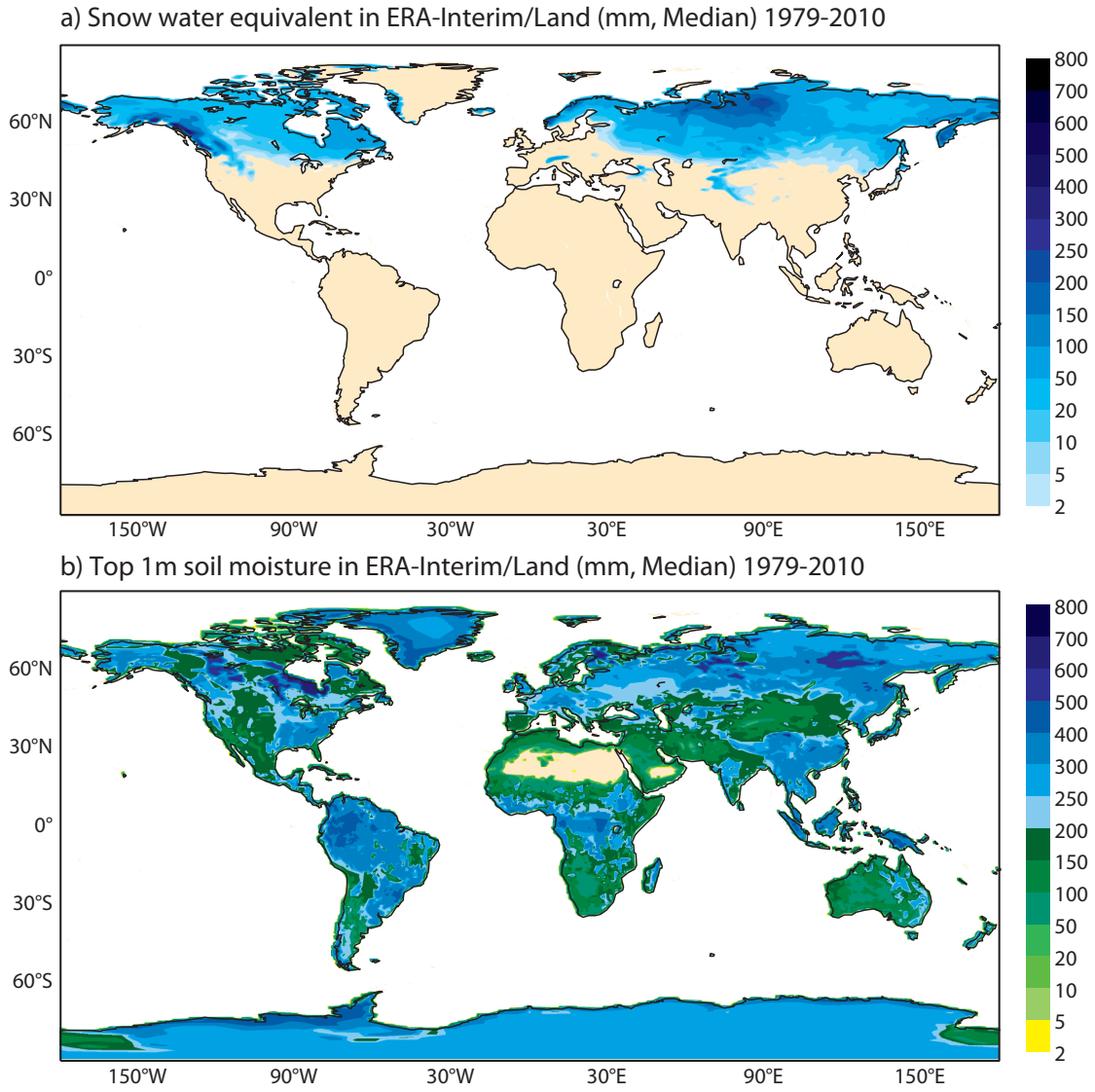


Fig. 2: Median of the land water reservoirs in the 1979-2010 period: (a) Snow Water Equivalent (mm or kg/m2) and (b) Top 1m Soil Moisture (mm or kg/m2), for 2 different dates: (a) 15 January (b) 15 July

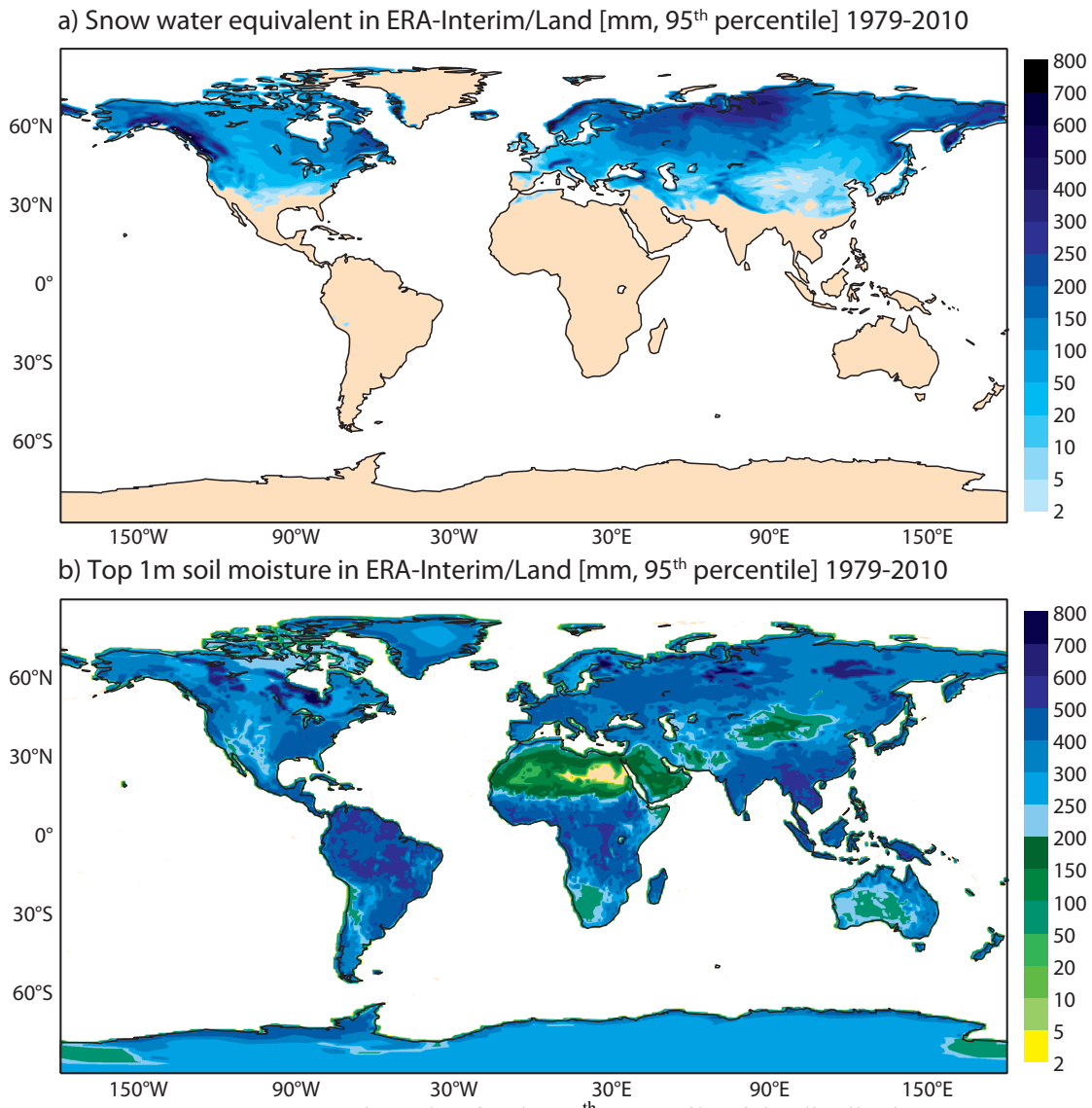


Fig. 3: Same as Fig. 2, but for the 95th percentile of the distribution.

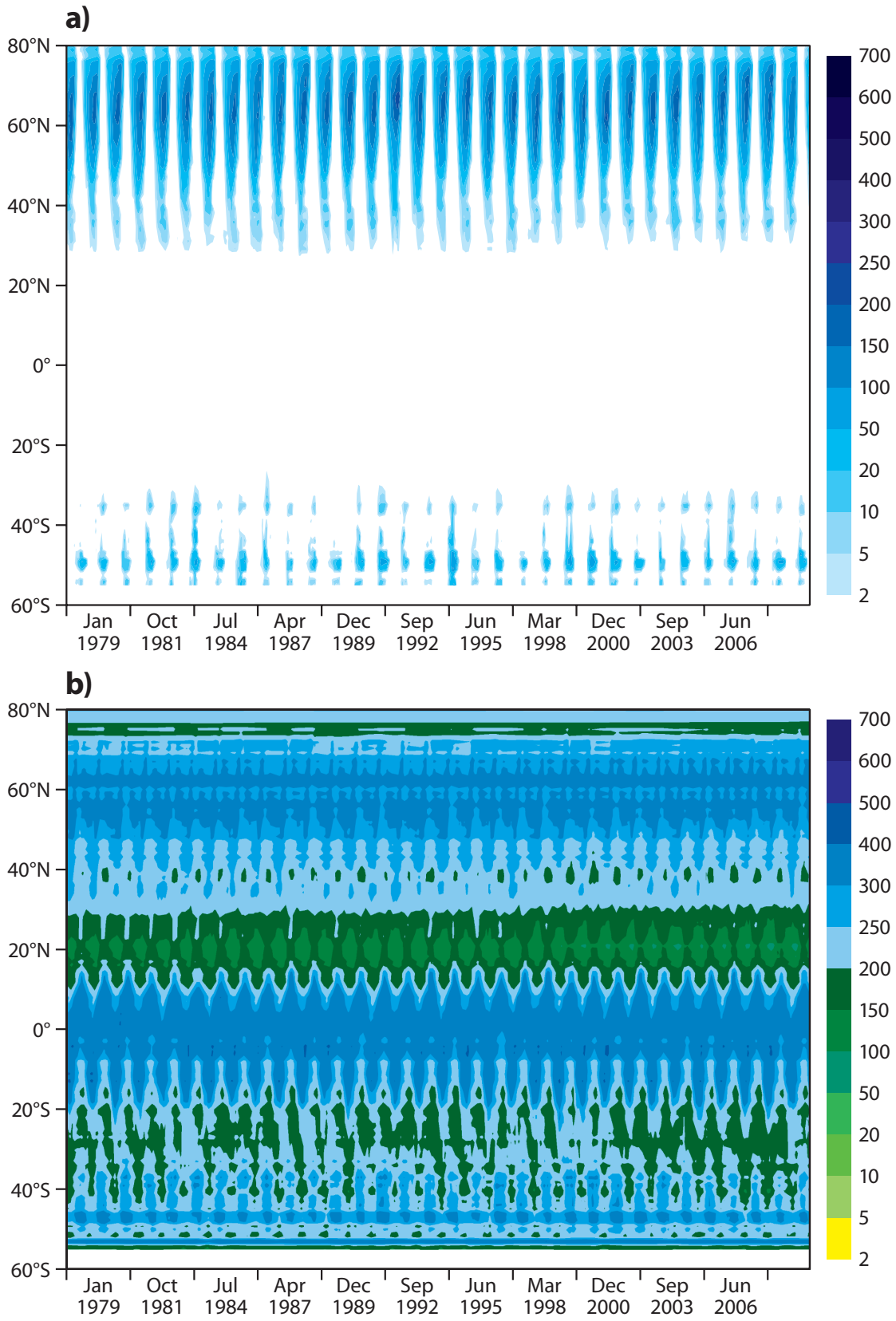


Fig. 4: Hovmöllers of the land water reservoirs for the 1979-2010 period: (s) Snow Water Equivalent (SWE, mm or kg/m^2) and (b) Top 1m Soil Moisture (TCSM, mm or kg/m^2).

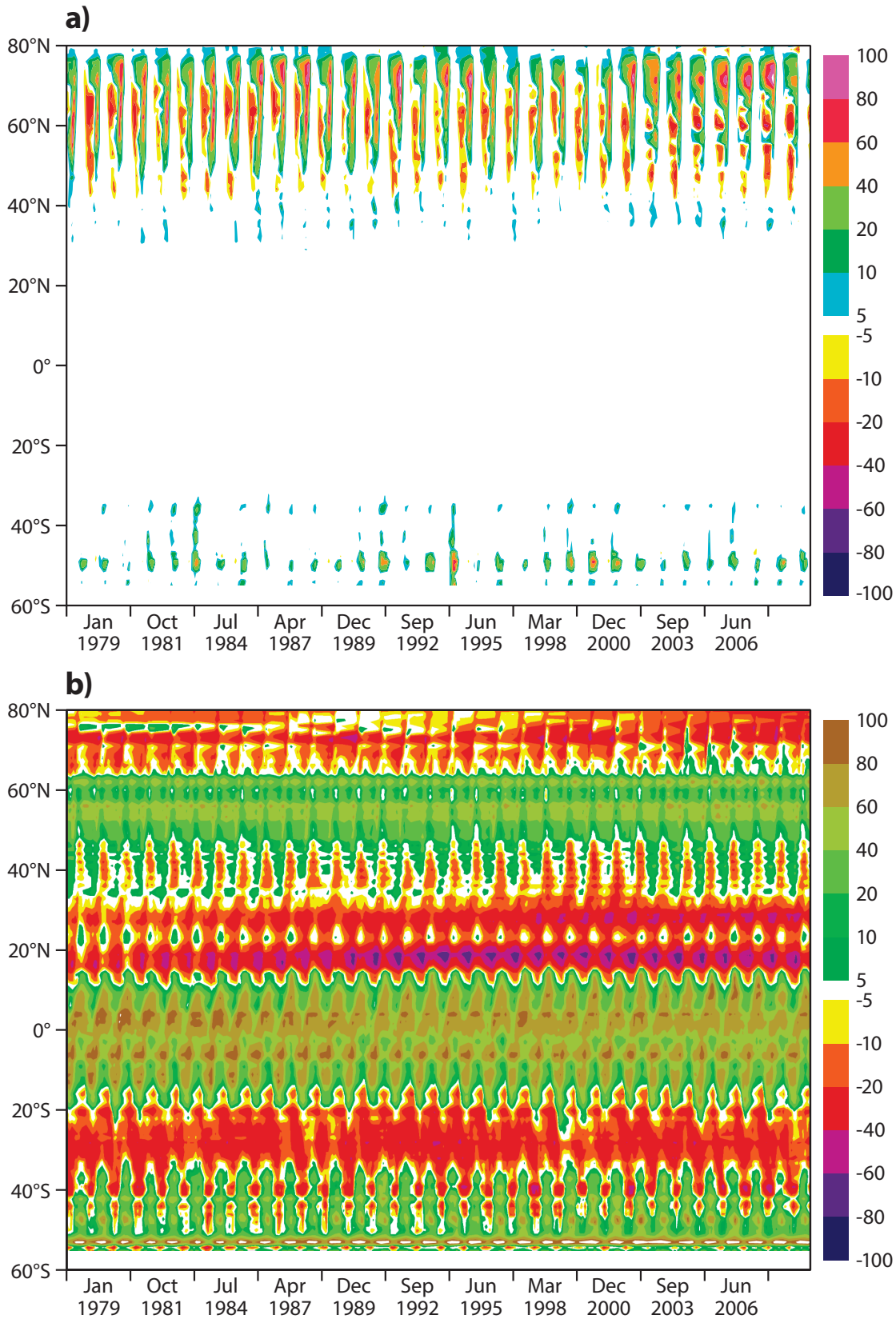


Fig. 5: Same as Figure 4, but for the Differences ERA-Interim/Land minus ERA-Interim

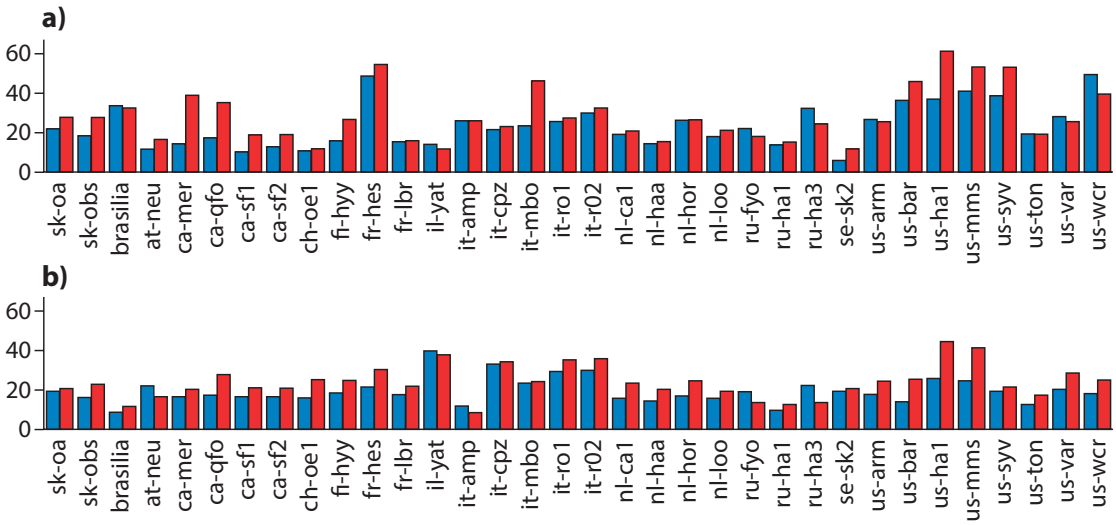


Fig. 6: Root mean square error (W m^{-2})for (a) Latent heat fluxes and (b) Sensible heat fluxes observed at 34 sites (as in Table 1) for ERA-Interim/Land (blue) and ERA-Interim (red) surface schemes.

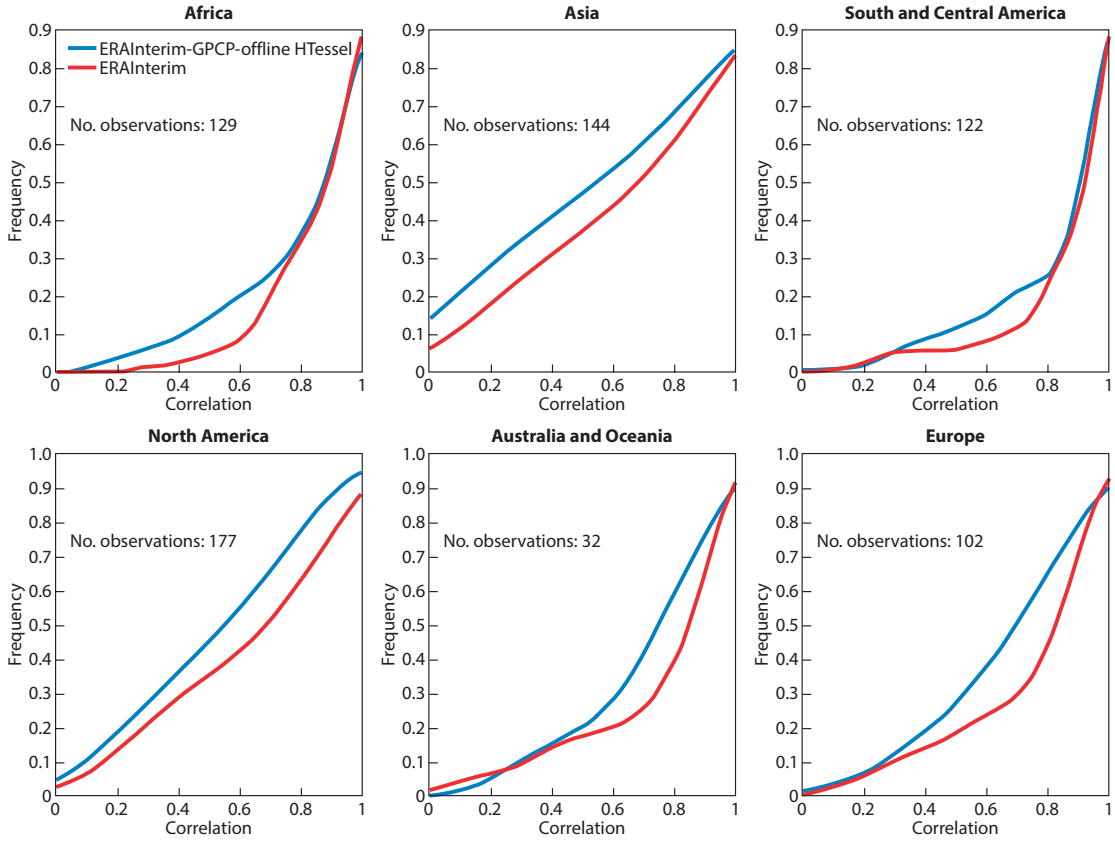
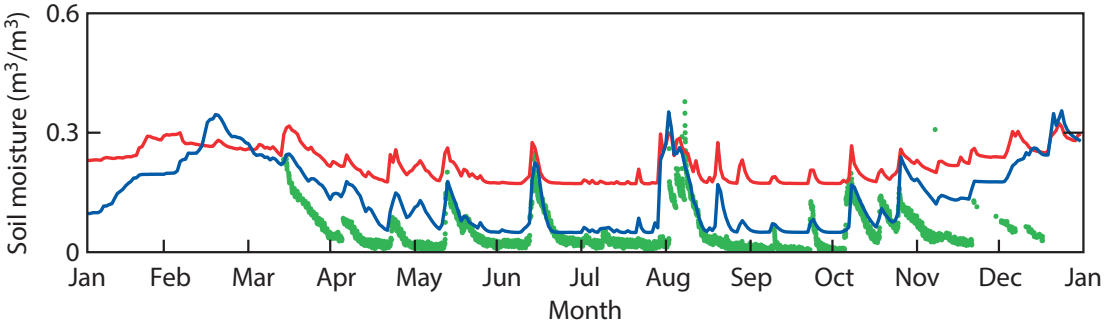


Fig. 7: Cumulative distribution function of river discharge correlations of ERA-Interim (red) and ERA-Interim/Land (blue dashed line) with GRDC data clustered by continents.



- **Fig. 8:** Evolution of volumetric soil moisture at a site in Utah for the year 2010. In-situ observations in green, ERA-Interim estimates in red, and ERA-Land estimates in blue.

-

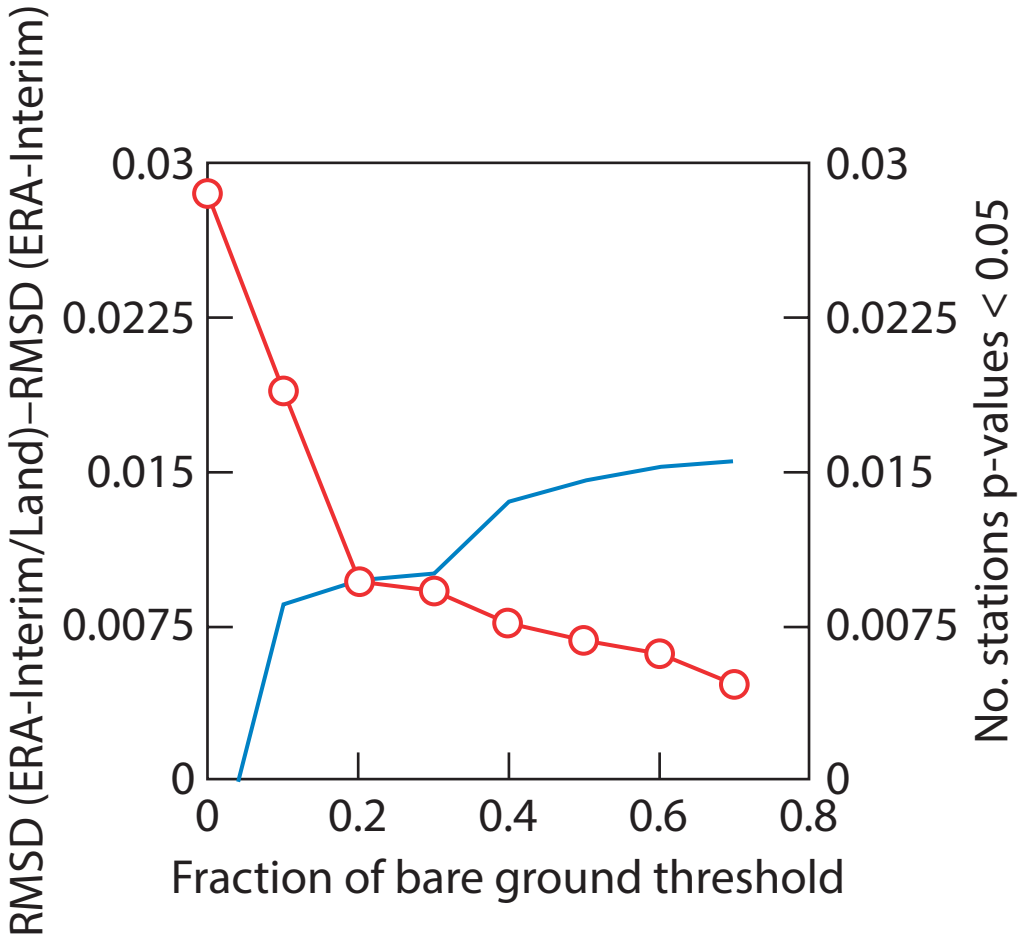


Fig. 9 RMSD difference between ERA-Interim/Land and ERA-Interim (black dots) as a function of the fraction of bare ground (black solid curve, left y-axis), the number of in situ stations with significant correlations is also presented (continues line, right y-axis). The dashed line represents a threshold where the sensitivity to the fraction of bare soil is less pronounced.

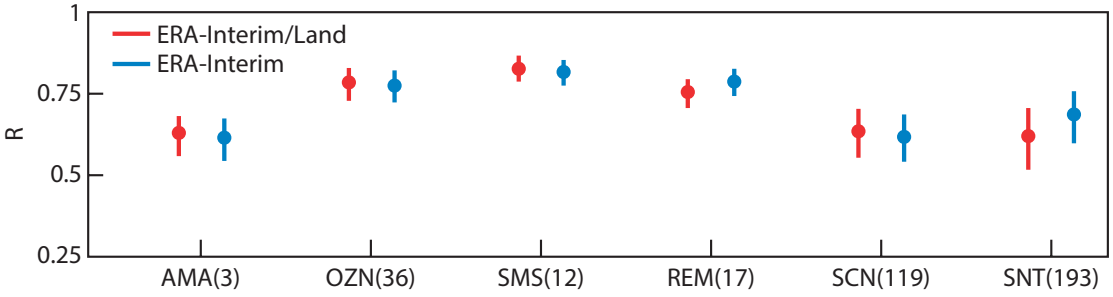


Fig. 10: Correlation with observed ISMN soil moisture networks (as in Table 3) for ERA-Interim/Land (red) and ERA-Interim (orange). Only significant correlations with p-values <0.05 are considered and for each of the observing networks the bars indicate the 95% Confidence Interval calculated using a Fisher-Z-transform.

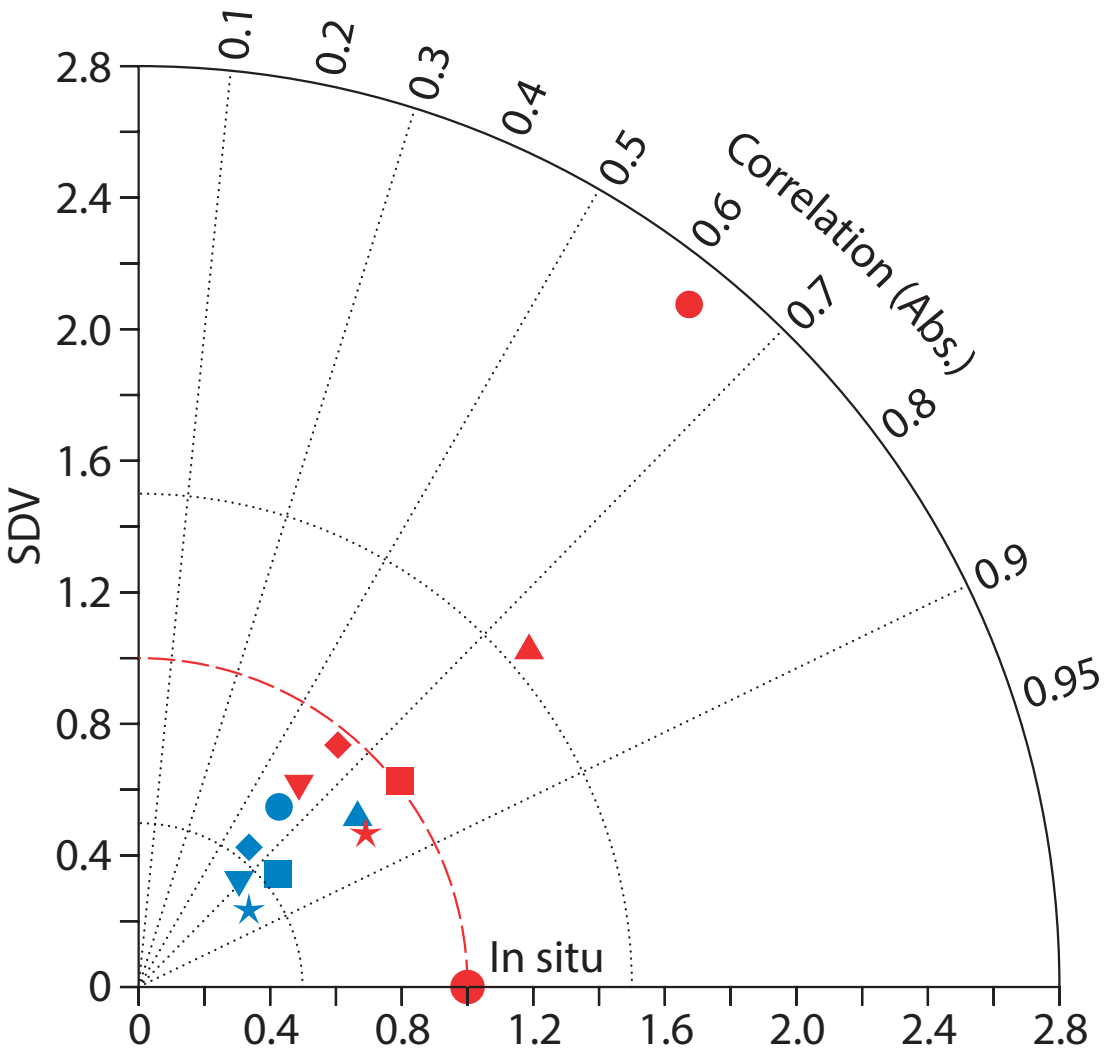


Fig. 11: Taylor diagrams illustrating the statistics from the comparison between ERA-Interim/Land in red and ERA-Interim in blue, compared to situ observations for 2010. Each symbol indicates the correlation value (angle), the normalized SDV (radial distance to the origin point), and the normalized centred root mean square error (distance to the point marked "In situ"). Circles are for the stations of the AMMA network (3 stations), square for that of the OZNET network (36 stations), stars for that of the SMOSMANIA network (12 stations), triangles for that of the REMEDHUS network (17 stations), diamonds for that of the SCAN network (119 stations) and inverted triangle for that of the SNOTEL network (193 stations). Only stations with significant correlations values are considered.

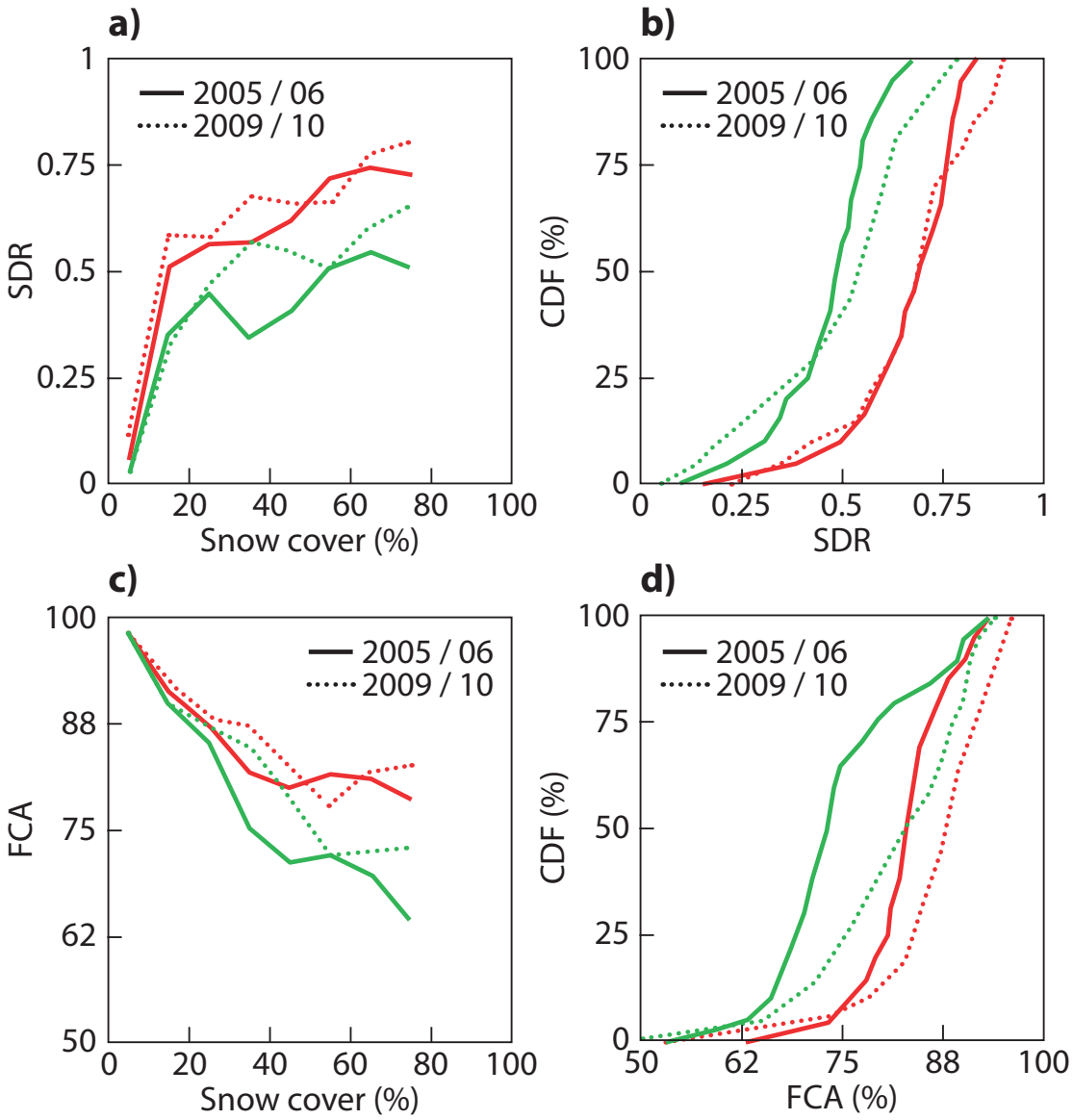
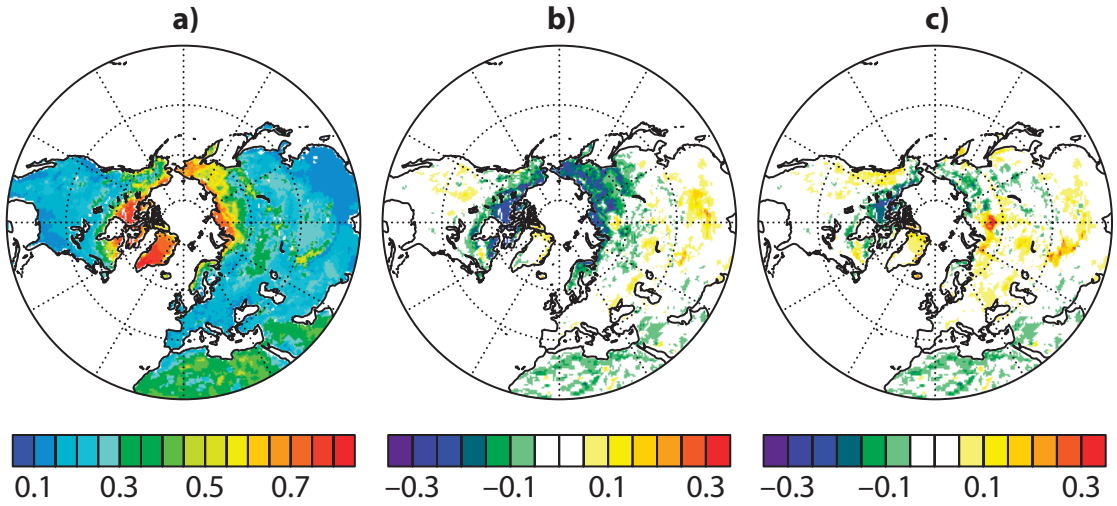


Fig. 12: Snow statistics calculated over Europe for (a) Snow Detection Rate and (b) cumulative distribution function of the Snow Detection Rate for 2005-2006 and 2009-2010 (1st of July to 30th of June), for ERA-Interim/Land (red) and ERA-Interim (green) surface offline simulations. The Fraction of Correct Accuracy function of snow cover (c) and its cumulative distribution function (d) for 2005-2006 and 2009-2010 (1st of July to 30th of June), for ERA-Interim/Land (red) and ERA-Interim (green) surface offline simulations.



872 -
873 **Fig. 13:** Mean observed Northern hemisphere albedo during spring derived from (a)
874 MODIS and differences of (b) ERA-Interim and (c) ERA-Interim/Land.