Author's Response to the Reviewer 1 Comments on "Reinitialised versus continuous regional climate simulations using ALARO-0 coupled to the land surface model SURFEXv5"

Julie Berckmans[1,2], Olivier Giot[1,2], Rozemien De Troch[1,3], Rafiq Hamdi[1,3], Reinhart Ceulemans[2], and Piet Termonia[1,3]

[1]Royal Meteorological Institute, Brussels, Belgium
[2]Centre of Excellence PLECO (Plant and Vegetation Ecology), Department of Biology, University of Antwerp, Antwerp, Belgium
[3]Department of Physics and Astronomy, Ghent University, Ghent, Belgium

Correspondence to: Julie Berckmans (julie.berckmans@meteo.be)

1 Reply to the Editor

Dear Editor,

  we have prepared a majorily improved version of the manuscript by incorporating all suggestions and critical comments raised by Reviewer 1. Please find attached our response to the referees' comments on our above mentioned manuscript, titled "Reinitialised versus continuous regional climate simulations using ALARO-0 coupled to the land surface model SURFEX". Below mentioned you will find our detailed responses to all the reviewers' comments and suggestions (put in italics and red). We have also explained where and how they were incorporated in the revised manuscipt.

2 Reply to Anonymous Referee #1

We would like to thank the Anonymous Referee #1 for the encouraging and constructive comments, which have improved the manuscript. Below is a list of modifications that we have implemented based on your comments.

  Comments:

  (1) While the reporting of the results has vastly improved there is still very limited discussion on why certain results are obtained with respect to particular simulation configurations. For example why does FS simulate better temperatures during summer? Is this because this simulation benefits from soil moisture memory by allowing the land surface to be fully interactive? The reduced performance in all configurations for precipitation is explained by the wet bias of the forcing data but perhaps more needs to be said on why FS is particularly worse than DRI and CON in winter.

The discussion has been extended in the revised manuscript. The discussion points were either based on own findings, or supported by other literature.

The new version (Page 7 Lines 4-7) reads:
"This is due to compensating effects, as the bias represents an average over the subdomain and might be the result of large negative and large positive biases over different parts of the particular subdomain compensating each other. However, the area-averaged bias gives a good impression of the ranking of the experiments (Kotlarski et al., 2014)."

and (Page 7 Lines 13-14):
"The frequent reinitialisations keep the large scales closer to the ERA-Interim forcing, whereas ALARO and ARPEGE are bound to a cold bias (Voldoire et al., 2013; Giot et al., 2016)."

and (Page 7 Lines 19-20):
"The Alps were characterised by a zero bias on the northern flank and mixed cold and warm bias on the southern flank compensating each other (Table 2)"

and (Page 7 Lines 22-24):
"These positive biases for FS might be related to rapidly decreasing soil moisture values in spring and summer (not shown). The temperature-soil moisture relation is strongest for FS, as this simulation benefits from soil moisture memory by allowing the land surface to be fully interactive with the atmosphere (Koster and Suarez, 2001)."

and (Page 8 Lines 7-10):
"More specifically, the overestimation of winter precipitation was strongest in the Mediterranean and Eastern Europe with values from 35.3% to 108.5% for all simulation modes (Table 3). The large values in the Mediterranean agreed with the large underestimation of 2 m temperature, as this region is characterised by a strong dependence of temperature and precipitation (Faggian, 2015)."

and (Page 8 Lines 11-12):
"The too wet driving field of ERA-Interim was superimposed on the smaller cold bias of FS, suggesting a higher precipitation bias than CON and DRI."

and (Page 8 Lines 22-24):
"Consequently, the summer precipitation was simulated better by FS than CON and DRI. During summer, the influence of the soil moisture memory on the atmosphere is more important, resulting in an improved representation of the precipitation with

FS."

and (Page 8 Lines 29-30):

"Frequent reinitialisations did not allow the land surface to build up a soil moisture memory, resulting in less skill for the representation of the precipitation."

and (Page 9 Lines 26-28):

"The dry climate of the Iberian Peninsula is less dominated by land surface-atmosphere interactions, as soil moisture does not impact the evapotranspiration availability (Seneviratne et al.,2010)."

(2) I was really concerned about the massive bias in the ground heat flux. I think the authors were too quick to dismiss this and should say more about why this bias persists particularly because it ranges from 40 to >100 W m-2. In particular, is the ground heat flux calculated or updated to include the residual energy imbalance (e.g. $G = RN – H – LE$) in order to maintain energy balance. . . because it certainly looks like it is. Perhaps something needs to be said on what steps will be taken next to improve the surface turbulent energy fluxes and their partitioning because the biases for the surface energy balance are quite large. It looks like including (or improving) a soil resistance for soil evaporation and a stomatal resistance for transpiration may improve the excessive evaporation in the model.

We agree that the biases are too large and need to be improved in order to represent well the surface energy budget. A detailed study by Napoly et al. (2016) confirmed the large bias for the ground heat flux in ISBA. The authors presented new parameterisations for a multi-source model, which resulted in a reduced ground heat flux and more energy available for the turbulent fluxes. This suggestion is included in our study.

The new version (Page 11 Lines 15-26) reads:

"The ground heat flux (G) showed improbably high values compared to the observed ones (Table 4). G is dependent on the soil temperature, which was largely overestimated by the land surface model (not shown). The standard version of ISBA, the nature tile of SURFEX, aggregates soil and vegetation properties for each grid cell (Noilhan and Planton, 1989). The net radiation is directly transferred to the ground, causing an inaccurate partitioning of the incoming energy into turbulent and ground heat fluxes (Napoly et al., 2016). An additional parameterisation for the leaf litter on the surface soil impacts this distribution (Wilson et al., 2012). We suggest to include an explicit formulation of the canopy layer (Napoly et al.,2016) and potentially a parameterisation for the forest litter layer (Napoly et al., 2016; Wilson et al., 2012). The implementation of these explicit formulations in ISBA outperformed the representation of the soil temperature of the original ISBA model (Napoly et al., 2016). They showed that the original ISBA model overestimated the G flux amplitude with several 10's of W m-2 during both daytime and nighttime. However, using the distinct surface energy budgets resolved part of the overestimated G by intercepting most of the downward solar radiation, leaving more energy available for turbulent fluxes. Consequently, less net radiation reaches the forest surface, reducing the energy available for the soil conductance Napoly et al. (2016)."

(3) In the first review I suggested to augment the validation of the surface energy balance with gridded observational products such as GLEAM or LandFlux. I'm not convinced by the reason given as to why this request was dismissed. It is difficult to evaluate the skill of the land surface fluxes at the grid cell values from a model that represents here a 20 km2 grid cell average to point observations, particularly if the vegetation 'mix' contrasts to that of the point observations. The biases are quite large and it would be good to check if this is associated with comparing grid cell averages to a point observation and whether the model can at least capture the spatial variability of the domain. Comparing against two points is not very rigorous and given how large the biases are, avoiding comparison to products such as the MERRA reanalysis, GLEAM or LandFlux really undermines the credibility of this model validation study. Could you look at how well the model captures the anomalies to assess more critically the ability of the model to simulate the temporal variability?

We have to admit to dismiss this suggestion to easily last time. The biases of the two stations are indeed too large to build conclusions on this. We have added 5 FLUXNET sites to the validation of the spatial distribution of Bowen Ratio. This gives a better representation of the spatial variability by the model. Even though these 5 stations represent the correct land cover by less than 50%, they are helpful in providing contrasts. The comparison of model grid cell averages to FLUXNET observations was also done by Blyth et al. (2010), Stöckli et al. (2008).

The new version (Page 6 Lines 21-25) reads:

"The model validation was done using grid cell averages compared to point observations, suggesting large differences in the land cover representation. In total, a subset of 7 stations that cover different biome types (Table 1), was selected to demonstrate the spatial variability of the domain by the model. However, the main focus was on the Vielsalm and Collelongo sites (Fig. 1), as their model grid cells represent more than 50% of the corresponding land cover, and cover different climate regimes."

and (Page 12 Lines 8-9):

"In summary, the model presented a good spatial variability of BR, and the agreement with observations was highest for FS. "

In addition, we compared the evapotranspiration (ET) of the model with the LandFLUX product (Fig. **??**). We calculated ET in the model based on the latent heat flux (LE). Next, we aggregated the hourly values to daily values, in order to be comparable with the satellite product that provides daily values based on monthly values. The model data at 20x20 km were upscaled to 1x1 degree, and regridded to the projection of LandFLUX. The results confirm our findings. CON overestimates ET/LE, and FS underestimates ET/LE compared to LandFLUX. DRI provides the best agreement with the observations. Besides, the spatial variability of the model is comparable with the observed one. We will not add this figure to the paper, as it requires more in-depth analysis and a proper description on the features of the LandFLUX dataset. This would serve as content for a follow-up manuscript. We would like to keep the focus of the manuscript to the objective on the potential of the simualation modes. However, we appreciate the author's suggestion on validating with a gridded observational product, as this certainly adds value to point observations.
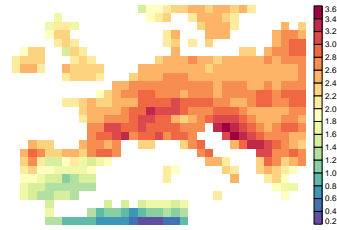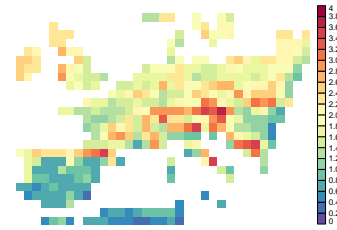
(a) LandFLUX
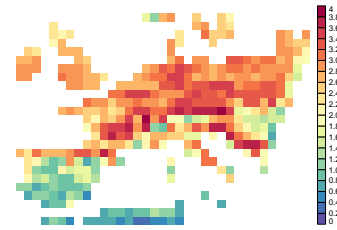
(b) CON
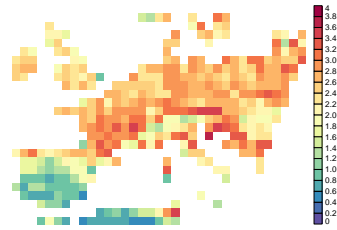
(c) DRI

(d) FS

Figure 1: Fig 1. The daily evapotranspiration (in mm) during JJA of [a] LandFLUX, [b] CON, [c] DRI, [d] FS, over Europe with a horizontal resolution of 1 x 1 degree, over a 10-year period.

(4) There are still sections of text where the language / sentence composition is awkward. Rather than note them all here I attach these comments in an annotated PDF of the revised manuscript. These are my suggestions for further refinement of the manuscript text and include some minor requests for further clarification.

We appreciate the reviewer's suggestions for the further refinement of the manuscript. They have been included. The revised manuscript with marked up changes has been added as attachment.

3 Manuscript version with highlighted changes is supplemented.

Reinitialised versus continuous regional climate simulations using ALARO-0 coupled to the land surface model SURFEXv5

Julie Berckmans[1,2], Olivier Giot[1,2], Rozemien De Troch[1,3], Rafiq Hamdi[1,3], Reinhart Ceulemans[2], and Piet Termonia[1,3]

[1]Royal Meteorological Institute, Brussels, Belgium
[2]Centre of Excellence PLECO (Plant and Vegetation Ecology), Department of Biology, University of Antwerp, Antwerp, Belgium
[3]Department of Physics and Astronomy, Ghent University, Ghent, Belgium

Correspondence to: Julie Berckmans (julie.berckmans@meteo.be)

Abstract. ~~For the simulation of the regional climate with limited area models, the common method for dynamical downscaling is the continuous approach with initial and lateral~~ Dynamical downscaling in a continuous approach using initial and boundary conditions from ~~the reanalysis or the~~ a reanalysis or a global climate model is a common method for simulating the regional climate. The simulation potential can be improved by applying an alternative approach of reinitialising the atmosphere, combined with either a daily reinitialised or a continuous land surface. We evaluated the dependence of the simulation potential on the running mode of the regional climate model ALARO coupled to the land surface model SURFEX, and driven by the ~~European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Re-Analysis (~~ERA-Interim ~~) data~~reanalysis. Three types of downscaling simulations were carried out for a 10-year period covering 1991 to 2000, over a Western European domain at 20 km horizontal resolution: (1) a continuous simulation of both the atmosphere and the land surface; (2) a simulation with daily reinitialisations for both the atmosphere and the land surface; and (3) a simulation with daily reinitialisations of the atmosphere while the land surface is kept continuous. The results showed that the daily reinitialisation of the atmosphere improved the simulation of the 2 m temperature for all seasons. It revealed a neutral impact on the daily precipitation totals during winter, but the results were improved for the summer when the land surface was kept continuous. The behaviour of the three model ~~simulations~~ configurations varied among different climatic regimes. Their seasonal cycle for the 2 m temperature and daily precipitation totals was very similar for a Mediterranean climate, but more variable for temperate and continental climate regimes. Commonly, the summer climate is characterised by strong interactions between the atmosphere and the land surface. The results for summer demonstrated that the use of a daily reinitialisated atmosphere improved the representation of the partitioning of the surface energy fluxes. Therefore, we recommend to use the alternative approach of the daily reinitialisation of the atmosphere for the simulation of the regional climate.

1 Introduction

The first long-range simulation of the general circulation of the atmosphere dates back to 1956 (Phillips, 1956). Today it is still the primary tool for global climate projections. However, due to limiting computer resources, the current horizontal resolution

of 100-200 km is still ~~coarse. A higher resolution and more spatial details~~ too coarse to resolve sufficient detail for regional climate projections. Finer spatial resolution that resolves the land surface heterogeneity can be obtained by nesting a regional climate model (RCM), over a smaller domain, ~~into~~ within a coarse-resolution global climate model (GCM). This is also referred to as dynamical downscaling. ~~The~~ A GCM or global reanalysis data product provides the large-scale meteorological and surface fields to the RCM as initial and lateral boundary conditions. The global features are thus translated into regional and local conditions over the region of interest (Giorgi, 2006). Hence, RCMs ~~allow to run~~ enable climate simulations over a smaller domain with ~~higher~~ finer scale horizontal resolution and with ~~an affordable computing cost~~ less expensive computational cost than running a GCM at the same resolution.

Since the late ~~60's~~ 60s, the numerical weather prediction (NWP) community ~~uses~~ has used high-resolution limited area models. The numerical approach was first used for a regional climate simulation by Dickinson et al. (1989). Their climate simulation used the NWP model in forecasting mode with ~~short-term~~ 3-5 daily reinitialisations of the initial conditions. To be able to run them without these short-term reinitialisations, the regional climate community applied monthly to multidecadal simulations, with only one single initialisation of the large-scale fields and frequent updates of the lateral boundary conditions (Giorgi and Mearns, 1999). These so-called long-term continuous simulations required improvements in the representation of physical processes in the RCMs. ~~This~~ The continuous simulation is still the most common in the RCM community (Leung et al., 2003). ~~Nonetheless, the~~ However, by applying the continuous approach, the simulated large-scale fields deviate from the driving lateral boundary conditions ~~, by applying the continuous approach~~ (von Storch et al., 2000).

The accuracy of the dynamical downscaling has improved by using short-term reinitialisations ~~(Kotlarski et al., 2012; Qian et al., 2003; L~~ ~~All these authors showed the advantage of using short-term reinitialisations by reducing systematic errors~~ to reduce systematic errors (Kotlarski et al., 2012; Qian et al., 2003; Lo et al., 2008; Lucas-Picher et al., 2013). However, only few ~~authors~~ in the RCM community adopted this method, mainly because of its higher computational costs. Most studies (Kotlarski et al., 2012; Qian et al., 2003; Lo et al., 2008) dealing with the evaluation of reinitialised versus continuous climate simulations, ~~covered~~ examined only short time periods ~~. The 24-hourly reinitialised simulation of the precipitation, in particular of the precipitation pattern, improved as compared to the continuous simulation (Kotlarski et al., 2012). This last mentioned analysis covered only a short time period, one month in 2002 during a~~ of one month to one year. Using a daily reinitialisation, Kotlarski et al. (2012) showed improvements in the prediction of precipitation for a case study of a large flooding event in the Elbe river catchment ~~.~~ in August 2002. Changing the period of reinitialisation, from monthly to 10-daily, a reduction in systematic errors has been shown for precipitation when using the 10-day reinitialisation (Qian et al., 2003). Even in a 20-year RCM simulation forced by reanalysis data, the sequence of events was better preserved by using ~~short-term~~ daily reinitialisations (Lucas-Picher et al., 2013).

A model approach with short-term reinitialisations demands additional simulation time at each reinitialisation start. This time is required to reach dynamical equilibrium between the lateral boundary conditions and the internal model physics and dynamics (Giorgi and Mearns, 1999). Beyond 24 hours small perturbations in the initial conditions of the atmosphere have only limited impact on the simulation potential (Anthes et al., 1989). In contrast to the atmosphere, the surface takes a longer

time to reach dynamical equilibrium with the overlaying atmosphere, ~~in the order of~~ from a few weeks to several seasons, depending on the depth of the soil layer.

The surface interacts with the climate through the soil moisture and soil temperature, by influencing the surface energy budget (Giorgi and Mearns, 1999). The soil moisture controls the partitioning of the incoming energy into ~~a~~ latent and sensible heat flux. The soil moisture limitation on the evapotranspiration is largest during ~~the~~ summer (Seneviratne et al., 2010). The availability of soil moisture for evapotranspiration is determined by the 2 m temperature (Jaeger et al., 2009). As the land surface-atmosphere interactions play a crucial role in the representation of the current and future climate (Seneviratne et al., 2010), it is important to validate the model with ~~ground observations. The FLUXNET database provides data on~~ observations. In site measurements can provide valuable estimates of the surface energy fluxes~~, based on eddy covariance measurements~~. More specifically, FLUXNET establishes a global network of eddy-covariance towers measuring these fluxes (Baldocchi et al., 2001).

The objective of this study was to evaluate the simulation potential of three regional climate downscaling approaches with different update frequencies of the initial conditions: (1) a continuous simulation of both the atmosphere and the land surface; (2) a simulation with daily reinitialisations for both the atmosphere and the land surface; and (3) a simulation with daily reinitialisations of the atmosphere while the land surface is kept continuous. We used the ALARO model to dynamically downscale the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Re-Analysis (ERA-Interim, Dee et al., 2011). Within this study, ALARO was coupled to the ~~land surface model of~~ Météo-France SURFace Externalisée land surface model (SURFEX, Masson et al., 2013). We evaluated the mean 2 m temperature and mean daily total precipitation by comparing with the 0.22° ECA&D E-OBS dataset (Haylock et al., 2008), and the surface energy fluxes by comparing with the FLUXNET database (Baldocchi et al., 2001). The analysis covered a 10-year period from 1991 to 2000, for a domain encompassing Western Europe.

The models, experimental design and observational datasets are described in section 2. The results for the mean surface parameters are covered in section 3. Section 4 demonstrates the results with respect to the surface energy budget. Finally, conclusions are given in section 5.

2   Model and experimental design

2.1   Model definition

The regional climate model used in this study is the ALARO model version 0, a configuration of the Aire Limitée Adaptation Dynamique Développement International (ALADIN) model with improved physical parameterisations (Gerard et al., 2009)~~.~~, combined with the Application de la Recherche à l'Opérationnel à Meso-Echelle (AROME), first baseline version released in 1998. The ALADIN model is the limited area model version of the global scale Action de Recherche Petite Echelle Grande Echelle Integrated Forecast system (ARPEGE-IFS) (Bubnová et al., 1995; ALADIN International Team, 1997). ARPEGE is a global spectral model, with a Gaussian grid for the grid-point calculation. The vertical discretisation ~~is done according to a~~ uses hybrid terrain-following pressure ~~hybrid coordinate.~~ coordinates. The ALARO-0 model

3

has been developed with the ARPEGE Calcul Radiatif Avec Nebulosité (ACRANEB) scheme for radiation based on Ritter and Geleyn (1992). This ALARO-0 model configuration ~~is being~~ has operated at the Royal Meteorological Institute of Belgium (RMI) for its operational numerical weather forecasts since 2010. The new physical parameterisation within the ALARO-0 model was specifically designed to ~~be~~ run at convection-permitting scales, with a particular focus on an im-

5   proved convection and cloud scheme ~~, developed by Gerard and Geleyn (2005) and further improved by Gerard (2007) and Gerard et al. (2009)~~(Gerard and Geleyn, 2005; Gerard, 2007; Gerard et al., 2009). The ALARO-0 model domain is centered over Western Europe at 46.47° N and 2.58° E with a dimension of 149 x 149 horizontal grid points and spacing of 20 km in both horizontal axes, ~~in~~ with a Lambert conformal projection (Fig. 1). The ~~domain encompasses Western Europe. The~~ model consists of 46 vertical layers with the lowest model level at 17 km and the model top extending up to 72 km.

10   The parameterisation of the land surface in ALARO-0 was initially with the land surface scheme Interaction Soil-Biosphere-Atmosphere (ISBA, Noilhan and Planton, 1989; Noilhan and Mahfouf, 1996). This scheme was designed for NWP and climate models, and describes ~~heat and water exchanges~~ the exchange of energy and water between the low-level atmosphere, ~~the~~ vegetation and the soil surface, by using either a diffusion method (Boone and Wetzel, 1999), or a force restore method based on two or three layers (Noilhan and Planton, 1989). Using the initial setup with ISBA, ALARO-0 has proven its skill for

15   regional climate modelling with daily reinialisations (Hamdi et al., 2012; De Troch et al., 2013). In addition, this setup has been validated for continuous climate simulations and is now contributing to the EURO-CORDEX project (Giot et al., 2016; Jacob et al., 2014). Meanwhile the more recent land surface model SURFEX, with additional parameterisations for ~~other surface typesthan nature~~urban surface types, has been implemented in the ALARO-0 model ~~. With respect to NWP applications, the introduction of~~ (Hamdi et al., 2014). A NWP application with SURFEXv5 within ALARO-0 has shown neutral effects on the

20   winter 2 m temperature and on the vertical profile of the wind speed. However, it has shown positive effects on the summer 2 m temperature, 2 m relative humidity, and resulted in improved precipitation scores compared to the previously used ISBA model (Hamdi et al., 2014). ~~Next to~~ Whereas the validation of ~~this setup for NWP,~~ the implementation of SURFEXv5 within ALARO-0 ~~is highly demanding for~~ has been done in a NWP context, this validation is also required in the context of long-term climate simulations. In this study, SURFEX uses the two-layer force restore method for ISBA. The first layer is the surface

25   superficial layer, that directly interacts with the atmosphere, and the second layer is the combined bulk surface and rooting layer, which is determined at the depth were soil moisture flux becomes negligible for a period of about one week and is thus more important as a reservoir for soil moisture during dry periods (Noilhan and Planton, 1989).

    SURFEX ~~is based on~~ uses a tiling approach ~~. The tiles provide~~ with each tile providing information on the surface fluxes according to the type of surface: nature, town, inland water and ~~ocean~~sea. The initial parameterisation ISBA for the nature

30   tile was conserved, and ~~parameterisations for the other surface tiles were added, such as the~~ the Town Energy Balance ~~scheme~~ (TEB, Masson, 2000) was added as a parameterisation for the town tile. TEB uses a canopy approach with three urban energy budgets for the layers roof, wall and road. The ISBA and TEB schemes were combined, together with parameterisation schemes for inland water and ~~oceans~~seas, and externalised, based on the algorithm of Best et al. (2004). In other words, the code can be used inside a meteorological or climate model, or in stand alone mode. Each tile is divided in different patches, according

35   to the tile type. These patches correspond to the plant functional types described in ECOCLIMAP (Masson et al., 2003).

ECOCLIMAP is a 1 km horizontal resolution global land cover database and assigns the tile fraction and corresponding physical parameters (e.g. LAI, albedo etc.) to SURFEX.

2.2 Experimental design

The regional climate model was driven by initial and lateral boundary conditions provided by the ERA-Interim reanalysis,
5 available at a horizontal resolution of ca. 79 km. ~~The Davies (1976) relaxation zone consisted~~ A relaxation zone of eight grid points ~~irrespective of the resolution~~ was used at the lateral boundaries of the domain (Davies, 1976). The zonal and meridional wind components, atmospheric temperature, specific humidity, surface pressure and ~~surface components were provided~~ soil moisture and soil temperature were updated every 6 model hrs as lateral boundary conditions and interpolated ~~hourly~~ to hourly distributions. They were introduced as initial conditions ~~accross~~ across the domain. A spin-up time was considered for the
10 model to reach equilibrium between the lateral boundary conditions and the internal model physics (Giorgi and Mearns, 1999). ~~For the sake of a good understanding, the following description makes a distinction between~~ Here we use an atmospheric spin-up~~time~~, typically of a few days, and a land surface spin-up~~time~~, typically of a few months to one year. The analysis covered a 10-year period from 00UTC on 01 January 1991 to 00UTC on 01 January 2001. Although the 10-year length is arbitrary, it is sufficiently long to include some inter-annual variability and to generate a reasonable sample of extreme events. The use
15 of a NWP model in a long-term climate setting for the performance of extreme precipitation events for a 10-year period was recently demonstrated (Lindstedt et al., 2015). To evaluate the sensitivity of the model to the update frequency of the initial conditions, three types of downscaling approaches were conducted with ALARO-0 coupled to SURFEXv5 ~~.~~ and are detailed below:

The first downscaling approach was done by simulating the model in a continuous mode for both the atmosphere and the
20 land surface (hereafter called CON ("CONtinuous"), Fig. 2). The model was ~~simulated from~~ started at 00UTC on 01 January 1990, and ran continuously until 00UTC on 01 January 2001. The first year was treated as both atmospheric and land surface spin-up~~time~~, and was excluded from the analysis. The simulations were interrupted and restarted monthly to allow for SSTs only to be updated. Other surface parameters that were updated monthly using the climatological values from ECOCLIMAP were the vegetation fraction, surface roughness length, surface emissivity, surface albedo, sand and clay fractions.

25 In the second downscaling approach, the model was reinitialised daily for both the atmosphere ~~as the~~ and the land surface (herafter called DRI ("Daily ReInitialisation"), Fig. 2). The model started at 12UTC on 01 January 1991, and each reinitialisation ran for 60 hrs. The first 36 hours were treated as atmospheric spin-up~~time, and were~~, and excluded from the analysis. By applying this downscaling approach, the regional model stays close to the ~~driving fields (von Storch et al., 2000). As the driving fields provided daily reanalysed data, a spin-up for the surface was redundant~~ERA-Interim forcing (von Storch et al., 2000).
30 However, the coarse representation of the land surface by the reanalysis is not able to capture the fine scale heterogeneity, particularly the soil moisture and soil temperature.

The third downscaling approach tries to find the best compromise between previous two approaches. The atmosphere was reinitialised daily and the land surface was simulated continuously with one single initialisation (hereafter called FS ("Free Surface"), Fig. 2). This allowed the model to simulate the atmospheric fields close to the ~~driving fields~~reanalysis forcing,

together with a surface in equilibrium state. The model was ~~simulated~~ run from 12UTC on 01 March 1990 until 31 May 1991, and the atmosphere was reinitialised daily for a simulation time of 60 hrs. The first 36 hrs were treated as atmospheric spin-up~~time~~, and were excluded from the analysis. The land surface conditions were kept continuous ~~and joined~~ by adding the land surface conditions for the 24-hr period after the atmospheric spin-up ~~time with the~~ to the land surface conditions of the previous

5 daily simulation. In contrast to the atmospheric spin-up~~time, the~~, the land surface spin-up lasted from 01 March 1990 until 31 May 1990, and this 3-monthly period was excluded from the analysis. Although CON required one year spin-up~~time~~, 3 months were sufficient for the FS deep soil moisture to reach equilibrium state, when starting in March (not shown). The simulations were done in parallel for each year from 1990 to 2000, and the 3 monthly spin-up ~~time~~ was replaced by the analysis of the previous year.

10 ~~The model output at every 3 hrs was used~~ We use three hourly output for the model evaluation presented in the manuscript. The evaluation of atmospheric variables for winter and summer was done for seven subdomains across Europe, to cover the spatial variability of the domain (Fig. 1). This was in agreement with the subdomains that were used in the EURO-CORDEX community (Kotlarski et al., 2014) and that were defined earlier in the framework of the ~~project~~ "Prediction of Regional scenarios and Uncertainties for Defining European Climate change risks and Effects" (PRUDENCE) project (Christensen

15 et al., 2007). The subdomains used in this study were the British Isles (BI), the Iberian Peninsula (IP), Mid-Europe (ME), France (FR), the Alps (AL), the Mediterranean (MD) and Eastern Europe (EA). ~~For the~~ The subdomains IP, ME, and EA ~~,~~ were chosen due to their diverse climate regimes to evaluate the yearly cycle of the atmospheric ~~variables was evaluated. These selected subdomains covered a range of climatic regimes. Additionally, the surface energy fluxes were analysed.~~ and land surface variables. As land-surface processes play an important role primarily during summer, the model output was stored

20 at every hour for the summer period of June-July-August (JJA) during the 10-year period. We evaluated the partitioning of the sensible and latent heat fluxes by the daily maximum Bowen Ratio (BR, Bowen, 1926) for the summer periods from 1996 to 2000 for the total study domain, and compared ~~the selected FLUXNET stations with their~~ a subset of FLUXNET stations using the corresponding model grid points. The corresponding daily maximum BRs were analysed for the 10-year summer period from 1991 to 2000. When the ~~value is lower (higher) than~~ BR < 1 ~~,~~ the latent heat flux ~~is higher (lower)~~ (LE) is greater than the

25 sensible heat flux ~~.~~ (H). Conversely when BR > 1 LE is less than H. The diurnal cycles of all surface energy fluxes were also analysed and validated against observations.

2.3 Observational reference data

The results of the climate simulations were validated against E-OBS, a daily high-resolution gridded observational dataset (Haylock et al., 2008). The dataset consists of the daily mean temperature, the daily maximum and minimum temperature, and

30 the daily precipitation total. The most recent version v12.0 was selected on the $0.22°$ rotated pole grid, corresponding to a 25 km horizontal resolution in Europe. It covers the period 01 January 1950 to 30 June 2015. ~~With respect to previous versions of E-OBS, some improvements include the new precipitation data series for countries southeast of the Baltic Sea, updated Slovakian series for all variables, updated Croatian series for all variables and a highly extended network for Catalonia, Spain. These improvements also concerned our area of interest and time period of interest.~~ In order to validate the model data, the

6

ALARO-0 data at 20 km horizontal resolution were bilinearly interpolated towards E-OBS at 25 km horizontal resolution and replotted to our study domain. A careful interpretation of E-OBS ~~was~~ is necessary, as this regridded non-homogeneously distributed network ~~applied~~ applies a smoothing out of extreme precipitation and consequently a large underestimation of the mean precipitation (Haylock et al., 2008).

5   For the validation of the surface fluxes ~~distribution~~ in the model, we used measurements from the FLUXNET Level 3 flux tower database (Baldocchi et al., 2001). It provides information on the energy exchange between the ecosystem and the atmosphere. FLUXNET is a global network, and consists of flux towers using the ~~eddy covariance~~ eddy-covariance method to monitor carbon dioxide and water vapor exchange rates, and energy flux densities. No gap-filling has been done and the comparison to the model output was only done at hours when ~~no gaps occurred~~ observational data was available. A number of
10 ~~stations~~ sites were already part of a separate flux measurement network (Aubinet et al., 2000). ~~However, only a few stations provided data for the first operating years covering the period 1996 to 2000. Two FLUXNET stations were selected, that provided data during this period and where the model grid cell represented~~ The model validation was done using grid cell averages compared to point observations, suggesting large differences in the land cover representation. In total, a subset of 7 stations that cover different biome types (Table 1), was selected to demonstrate the spatial variability of the domain by the
15 model. However, the main focus was on the Vielsalm and Collelongo sites (Fig. 1), as their model grid cells represent more than 50% of the corresponding land cover, ~~to show energy fluxes that were representative for the particular land cover . The selected ecosystem towers cover different climatic regimes(Fig. 1): (1) Vielsalm in Belgium, a temperate climate, at an altitude of 491 m with a tower height of 40 m and covered by deciduous broadleaved forest and evergreen coniferous, and (2) Collelongo in Italy, a Mediterranean climate, at an altitude of 1645 m with a tower height of 32 m and mainly covered by deciduous broadleaved
20 forest.~~ and cover different climate regimes.

3   Validation of the mean model state

3.1   Spatial distribution

3.1.1   Daily mean 2 m temperature

The spatial distributions of the 10-year daily mean temperature bias (absolute, (model - observed)) of CON, DRI and FS
25 simulations were compared to E-OBS (Fig. 3), for ~~the~~ winter (DJF: December-January-February) and summer (JJA: June-July-August)~~season. The average biases~~ . The area-averaged bias during winter and summer for CON, DRI and FS for the entire domain as well as for specific subdomains ~~are~~ is presented in Table 2. CON simulated a cold bias in general, except for northern Africa in summer, with a pronounced orographic effect, for both winter and summer (Fig. 3c,d). The cold bias over the entire domain was less pronounced in summer with a value of -0.6 °C compared to the winter bias of -1.8 °C (Table 2). Moreover, the
30 Iberian Peninsula~~was well simulated during summer as compared to E-OBS, resulting in a~~ , Mediterranean and Eastern Europe resulted in a small bias of -0.5 °C during summer as compared to E-OBS (Fig. 3d). This is due to compensating effects, as the bias represents an average over the subdomain and might be the result of large negative and large positive biases over different

7

parts of the particular subdomain compensating each other. However, the area-averaged bias gives a good impression of the ranking of the experiments (Kotlarski et al., 2014). ~~Additionally, the biases of the Mediterranean and Eastern Europe resulted in similar small biases , due to compensating errors as can be seen from (Fig. 3d).~~

5 With respect to CON, DRI demonstrated a reduction of the cold bias during winter and summer, ~~most prominent at~~ particularly for the eastern part of the domain (Fig. 3e,f). ~~This resulted in a smaller~~ The area-averaged bias for Eastern Europe ~~of -0.3 °C and 0.0 °C for DRI relative to CON which had a bias of -1.1 °C and -0.5 °C for~~ was close to zero for DRI during winter and summer ~~respectively~~ (Table 2). ~~Other subdomains showing a~~ , in spite of many significant non-zero bias points for the summer (Fig. 3e,f). A large improvement of the 2 m temperature simulation by DRI ~~, were~~ was also produced for Mid-Europe and the Alps with a winter bias of -0.7 °C and -1.4 °C respectively that is about half of the bias of CON~~, and a summer~~ 10 ~~bias of~~ . For summer the bias decreases further to -0.3 °C (ME) and -0.8 °C ~~, even more than half of the bias of~~ (AL); a 1 °C decrease relative to CON for these subdomains. The frequent reinitialisations keep the large scales closer to the ERA-Interim forcing, whereas ALARO and ARPEGE are bound to a cold bias (Voldoire et al., 2013; Giot et al., 2016).

The performance of the FS simulation was different for winter and summer ~~as compared to CON and DRI~~ (Fig. 3g,h). The simulation of the 2 m temperature during winter was best of all three approaches when using FS. Large parts of the domain 15 resulted in biases close to zero, ~~such as~~ for the British Isles, France, Mid-Europe and Eastern Europe (Fig. 3g). The bias decreased by ca. 1 °C in FS compared to CON for these subdomains (Table 2). During summer, the sign of the bias reversed from negative to positive, except for some isolated areas (Fig. 3h). The Alps were ~~much better presented by FS, resulting in~~ characterised by a zero bias ~~as compared to CON and DRI which showed a bias of -1.8 °C and -0.8 °C~~ on the northern flank and mixed cold and warm bias on the southern flank compensating each other (Table 2). ~~For~~ Large parts of the Iberian Peninsula 20 and the Mediterranean ~~, compensating biases resulted in positive and~~ exhibited a warm bias (Fig. 3h), resulting in positive values close to zero ~~summer biases (Fig. 3h~~ (Table 2). Mid-Europe, France and Eastern Europe were mainly characterised by a positive bias of around 1 °C (Table 2). ~~The summer absolute bias simulated by FS was very similar to CON for the Iberian Peninsula and the Mediterranean, but slightly enhanced for Eastern Europe with ca. 0.6 °C.~~ These positive biases for FS might be related to rapidly decreasing soil moisture values in spring and summer (not shown). The temperature-soil moisture relation 25 is strongest for FS, as this simulation benefits from soil moisture memory by allowing the land surface to be fully interactive with the atmosphere (Koster and Suarez, 2001).

In summary, CON underestimated winter and summer 2 m temperature ~~with~~ by 1-2 °C on average. With respect to CON, DRI ~~and FS showed a~~ showed a general positive effect during winter and summer. ~~In spite of a slight enhancement by FS of the bias during summer for Eastern Europe, the winter bias was improved for most subdomains by using FS. Overall, the~~ 30 Consequently, the use of a daily reinitialised atmosphere improved the representation of the 2 m temperature for both winter and summer compared to a continuous simulation of the atmosphere. The winter bias was further improved for most subdomains for FS compared to CON. For summer, most subdomains experienced a warm bias, in the same order of magnitude as CON. The difference might point at the interaction of the land surface and the atmosphere being stronger with FS, because of the soil moisture memory.

3.1.2 Daily accumulated precipitation

The spatial distributions of the 10-year daily accumulated precipitation bias (relative, (model-observed)/observed) of CON, DRI and FS were compared to E-OBS, for the winter and the summer seasons (Fig. 4). The ~~mean~~ relative biases during winter and summer for CON, DRI and FS are presented for the entire domain as well as for the specific subdomains in Table 3. The precipitation pattern of E-OBS during winter displayed highest values of > 3 mm day^{-1} over Portugal, northwestern Spain, western England, Scotland and Ireland, the Adriatic Coast and the northern flanks of the Alps (Fig. 4a,b). During summer, similar amounts of rainfall were concentrated over the Alps and the Carpathians, while lowest values of < 1 mm day^{-1} at the Iberian Peninsula, the Mediterranean and northern Africa.

During winter, all simulations demonstrated a similar spatial variability of the wet bias, except for a dry bias in northern Africa (Fig. 4c,e,g). In general, ALARO was forced towards the too wet driving fields of ERA-Interim (Lucas-Picher et al., 2013), which can explain part of the overestimated precipitation. More ~~particularly~~specifically, the overestimation of winter precipitation was strongest in the Mediterranean and Eastern Europe with values ~~of 46.0 % and~~ from 35.3% ~~respectively to 108.5% for all simulation modes~~ (Table 3). ~~However, the bias averaged over~~ The large values in the Mediterranean agreed with the large underestimation of 2 m temperature, as this region is characterised by a strong dependence of temperature and precipitation (Faggian, 2015). The area-averaged bias for the entire domain was larger for FS in winter with ca. ~~36% compared to less than 25 %~~ 36.3% compared 16.6% and 20.9% for CON and DRI. ~~This corresponded to~~ The too wet driving field of ERA-Interim was superimposed on the smaller cold bias of FS, suggesting a higher precipitation bias ~~of 10-20 % for all specific subdomains and even more than 50% higher for the Mediterranean~~than CON and DRI.

During summer, the simulations showed different spatial variability (Fig. 4d,f,h). ~~Regarding CON, the sumer precipitation bias~~ In comparison to the winter bias, the summer precipitation bias for CON was reduced over the continental part ~~as compared to winter and~~ with positive and negative biases ~~occurred~~ over the southern part of the domain (Fig. 4d). The Mediterranean expressed a high wet bias of 60.5% for CON, but the absolute values in summer were close to zero, as it is characterised by a climate with dry summers (Fig. 4b). The bias pattern over the continental part was very similar for DRI compared to CON during summer, while Southern Europe showed increased wet biases for DRI (Fig. 4f). The Iberian Peninsula, France and the Mediterranean demonstrated a bias of 30.0%, 18.3% and 84.8% respectively compared to 11.5%, 12.0% and 60.7% with CON (Table 3). The performance of FS was similar to CON for Southern and Eastern Europe (Fig. 4h). This ~~contrasted~~ is in contrast to the continental part of the domain, where the precipitation signal reversed ~~and dry biases occurs, though it was rather small~~ relative to CON and DRI and a small dry bias persisted (-7.0% for France, -13.4% for Mid-Europe, -8.2% for Eastern Europe respectively). Consequently, the summer precipitation was simulated better by FS than CON and DRI. During summer, the influence of the soil moisture memory on the atmosphere is more important, resulting in an improved representation of the precipitation with FS.

In summary, the model was characterised by a wet bias in winter and summer. The spatial variability during winter was very similar for all simulations, ~~but during summer~~so the use of a daily reinitialised atmosphere had a neutral impact on the winter precipitation. During summer, the precipitation showed a different behaviour ~~. For~~ with the different simulation modes.

In summer for the southern part of the domain, ~~DRI established increased precipitation biases, while FS was more different to CON~~ precipitation bias experienced a neutral effect with DRI, whereas the precipitation bias increased for the continental part~~, but not so much for the southern part. The use of a daily reinitialised atmosphere in DRI and FS had a neutral impact on the winter precipitation. FS improved the summer precipitation bias~~. Frequent reinitialisations did not allow the land surface to build up a soil moisture memory, resulting in less skill for the representation of the precipitation. However, the precipitation bias improved with FS for the continental part. Therefore, the combination of the daily reinitialised atmosphere together with a continuous surface is crucial in summer to get the best results.

3.2 Mean annual cycle

3.2.1 Daily mean 2 m temperature

To validate specific subdomains within the larger domain on a monthly scale, the mean annual cycles of the downscaled simulations were compared to the observations (Fig. 5). We focused on the following subdomains (Fig. 1): (1) the Iberian Peninsula at the western boundary of the domain with its warm and dry summer climate; (2) Mid-Europe with its temperate climate; and (3) Eastern Europe at the eastern boundary of the domain with its continental climate.

The daily mean 2 m temperature reached ~~about~~ 23 °C ~~in~~ for the Iberian Peninsula, ~~while it raised to~~ and 20 °C ~~in~~ for both Mid-Europe and Eastern Europe (Fig. 5a,b,c). For these selected subdomains, all downscaled simulations presented very similar autumn (SON: September-October-November) temperatures, but underestimated them with respect to E-OBS. Therefore, the autumn temperature is not sensitive to the update frequency of the initial conditions. Regarding the other seasons, the simulations revealed a different behaviour in the representation of the 2 m temperature with respect to the observations.

For the Iberian Peninsula, the 2 m temperature was generally underestimated for all seasons (Fig. 5a). Except for autumn, FS was closer to the observations as compared to CON and DRI, resulting in a yearly mean temperature of 12.5 °C, which was closer to the observed yearly mean temperature of 13.7 °C as compared to 11.6 °C and 11.9 °C by CON and DRI respectively. ~~The~~ Therefore, summer 2 m temperature was well simulated by FS for this subdomain. For Mid-Europe, CON and DRI underestimated the 2 m temperature for all seasons, ~~whereas~~ and FS was very close to the observations from February to May (Fig. 5b). However, FS overestimated the summer 2 m temperature and CON and DRI underestimated the summer 2 m temperature. Still, the yearly mean value of 9.0 °C by FS was very close to the observational mean of 9.3 °C. ~~In contrast to the Iberian Peninsula and Mid-Europe~~For Eastern Europe, DRI and FS demonstrated almost identical behaviour for the simulation of the 2 m temperature ~~for Eastern Europe~~ during winter and spring (MAM: March-April-May) with small biases (Fig. 5c)~~. Their simulation was very close to the observations~~, whereas CON underestimated the 2 m temperature. Similar to Mid-Europe, FS overestimated the summer 2 m temperature with ca. 1 °C and CON underestimated the summer 2 m temperature with ca. 1 °C in Eastern Europe. Yet again, the yearly mean value of 8.5 °C by FS was very similar as compared to the observations with a value of 8.6 °C , while largest differences occurred using CON with a value of 7.5 °C.

In summary, the yearly mean temperature was underestimated by CON for all subdomains. In general, ALARO is bound to a cold bias (Giot et al., 2016). Along the selected subdomains, there were larger differences between the simulations in Mid-

Europe and Eastern Europe compared to the Iberian Peninsula. The dry climate of the Iberian Peninsula is less dominated by land surface-atmosphere interactions, as soil moisture does not impact the evapotranspiration availability (Seneviratne et al., 2010). DRI was able to simulate the 2 m temperature better for Mid-Europe and Eastern Europe as compared to CON for winter, spring, and summer. FS had the best yearly mean 2 m temperature, but the summer 2 m temperature was overestimated by FS for Mid-Europe and Eastern Europe.

3.2.2 Daily accumulated precipitation

Similar to temperature, the monthly means of the daily accumulated precipitation, averaged over the 10-year period, are shown in Fig. 5 for the Iberian Peninsula, Mid-Europe and Eastern Europe. When comparing the observations, the seasonal variability was most pronounced for the Iberian Peninsula, with minimum precipitation values of ca. 0.5 mm day^{-1} during summer, and maximum precipitation values of ca. 3 mm day^{-1} during spring, autumn and beginning of the winter (Fig. 5d). The precipitation in Mid-Europe reached highest values of ca. 3 mm day^{-1} during summer (Fig. 5e). The continental climate of Eastern Europe presented average values of 1 mm day^{-1} for winter and spring, while most rainfall occurred in the summer of ca. 2.5 mm day^{-1} (Fig. 5f).

In general, the agremeent of the simulations was largest during autumn. For the Iberian Peninsula, the seasonal pattern of the downscaled simulations followed the seasonal pattern of E-OBS, despite a general overestimation of the precipitation (Fig. 5d). This overestimation was stronger in winter and in spring for the Iberian Peninsula, and is in agreement with Lucas-Picher et al. (2013). For these two seasons, E-OBS showed an undercatch of the precipitation, which might have amplified the model biases (Rauscher et al., 2010). CON and DRI were closer to the observations than FS in winter and spring, resulting in yearly mean values of 1.9, 2.0, and 2.1 mm day^{-1} respectively for CON, DRI and FS, as compared to the observational mean value of 1.7 mm day^{-1}. In Mid-Europe, the model overestimated the precipitation for most of the year, except for summer (Fig. 5e). During summer, FS showed a large underestimation, whereas CON and DRI showed similar overestimates of precipitation. The precipitation in Eastern Europe (Fig. 5f) was overestimated by the model during most of the year, except for summer (Lucas-Picher et al., 2013), where there is considerable agreement on the estimation of the summer precipitation. The yearly mean precipitation by CON was lowest with 2.0 mm day^{-1} and highest when using FS with 2.1 mm day^{-1}, as compared to 1.6 mm day^{-1} by the observations (Fig. 5f).

In summary, the three downscaling approaches generally overestimated the precipitation over these 3 regions in all seasons except during JJA. On a yearly basis, the differences between CON, DRI and FS were small, but on a monthly basis, the magnitude of differences were regionally dependent with larger differences between the model simulations for Mid-Europe and Eastern Europe compared to the Iberian Peninsula.

4 Validation of surface fluxes

The spatial distributions of the 5-year daily maximum Bowen Ratio (BR)~~of~~, at the time of maximum H and LE, of CON, DRI and FS were compared to FLUXNET observations, for the summer period only (Fig. 6a,b,c). The corresponding spatial distributions of the 10-year daily maximum BR of CON, DRI and FS were evaluated with respect to the results for the 5-year period (Fig. 6d,e,f). The mean diurnal cycles of the surface energy fluxes are illustrated over the 5-year summer period 1996-2000 for the FLUXNET stations of Vielsalm and Collelongo and their corresponding model grid points (Fig. 7, Table 4).

The daily maximum BR showed a strong gradient of increasing values towards the south of the domain (Fig. 6a,b,c). ~~However, large differences appeared for~~ Recall that when the value is lower (higher) than 1, the latent heat flux (LE) is higher (lower) than the sensible heat flux (H). Southern Europe is characterised by dry summers, with a strong control of the soil moisture on the evapotranspiration (Jaeger et al., 2009), and thus lower LE than H. Large differences existed between the three downscaling approaches, particularly for the continental part of the domain. ~~Relatively low~~ This is in agreement with the larger differences between the simulation modes in 2 m temperature for the continental subdomains compared to the southern subdomains. Therefore, the 2 m temperature is correlated to evapotranspiration by the sensitivity of LE on the soil moisture (Seneviratne et al., 2006).

CON had relatively low BR values of 0 to 1 ~~were represented by CON, while~~ for the continenal area, DRI showed BR values of 0.5 to 1 and FS had the highest values of 2 to ~~3 were expressed by FS~~ 3. The larger values presented by FS were in agreement with the warm and dry summer bias. Comparison with the FLUXNET observations showed a general underestimation of BR for CON and DRI, whereas FS estimated well the BR for most of the sites. ~~When the value is lower (higher) than 1, the latent heat flux is higher (lower) than the sensible heat flux.~~ The FLUXNET ~~observations for Vielsalm and Collelongo were displayed, and indicated best agreement with DRI (Fig. 6b), expressed by values of 1.12 and 1.32 respectively (Table 4)~~ sites show typically a closure imbalance of 20% (Wilson et al., 2002), whereas it was assumed that the Bowen ratio is well estimated by the eddy-covariance system (Napoly et al., 2016). Though this validation was based on 5 summer periods only from 1996 to 2000, it was still robust as indicated by the corresponding ~~plot~~ BR for the 10-year summer period from 1991 to 2000 (Fig. 6d,e,f, Table 4). ~~In spite of highest BR values presented by FS, the stations of~~ Considering the two FLUXNET sites Vielsalm and Collelongo, DRI indicates the best agreement with the observed BR (Fig. 6b), with values of 1.12 and 1.32 respectively compared to 0.78 and 1.35 for (Table 4). Despite higher BR values in FS, estimates corresponding to Vielsalm and Collelongo were located ~~into isolated parts of lower BR , indicated by the~~ in regions where the BR was comparatively low, resulting in average values of 0.61 and 0.83 respectively (Table 4).

The ~~net radiation was underestimated for all simulations~~ model perfomed well for the simulation of the daily cycle of the net radation (RN, not shown), even though the model underestimated the values of RN by about 5-10% (Table 4)~~, but this underestimation was larger for Collelongo, which could be related to its complex topography~~. The model generally underestimated H, and overestimated LE ~~.~~ (Fig. 7 and Table 4). The ground heat flux (G) showed ~~much higher values than~~ improbably high values compared to the observed ones (Table 4). G is dependent on the soil temperature, which ~~is~~ was largely overestimated by the land surface model ~~. This is due to the representation of the soil-surface leaf litter in the model.~~

~~Wilson et al. (2012) showed that without~~ (not shown). The standard version of ISBA, the nature tile of SURFEX, aggregates soil and vegetation properties for each grid cell (Noilhan and Planton, 1989). The net radiation is directly transferred to the ground, causing an inaccurate partitioning of the incoming energy into turbulent and ground heat fluxes (Napoly et al., 2016). An additional parameterisation for the leaf litter on the surface soil impacts this distribution (Wilson et al., 2012). We suggest to include an explicit formulation of ~~water and energy exchanges within the residue layer, their surface model overestimated LE, G and soil temperature and underestimated H. As the net radiationand ground heat flux were simulated very similarly for all simulations, they were not shown in Fig. 7.~~ the canopy layer (Napoly et al., 2016) and potentially a parameterisation for the forest litter layer (Napoly et al., 2016; Wilson et al., 2012). The implementation of these explicit formulations in ISBA outperformed the representation of the soil temperature of the original ISBA model (Napoly et al., 2016). They showed that the original ISBA model overestimated the G flux amplitude with several 10's of W m^{-2} during both daytime and nighttime. However, using the distinct surface energy budgets resolved part of the overestimated G by intercepting most of the downward solar radiation, leaving more energy available for turbulent fluxes. Consequently, less net radiation reaches the forest surface, reducing the energy available for the soil conductance Napoly et al. (2016).

For Vielsalm, H was simulated well by DRI and FS during nighttime and daytime, whereas CON underestimated H during daytime (Fig. 7a). The daily maximum H by CON was only 118 W m^{-2}, as compared to 151 and 139 W m^{-2} for DRI and FS respectively (Table 4). ~~Yet again, this~~ This validation was only done for 5 summer periods from 1996 to 2000, but the corresponding daily maximum values for the 10-year summer period 1991-2000 ~~indicated~~ indicate that the 5-year period was representative for the validation of the fluxes (Table 4). The LE was overestimated by all simulations, but the difference with the observations was smallest for DRI, ~~while it was highest~~ due to the frequent land surface reinitialisations, compared to highest values for CON. The daily maximum BR was lower than 1 for all downscaling approaches (Table 4). This means that they all simulated a higher latent than sensible heat flux. Still, DRI and FS showed higher values for BR than CON. Therefore, the partitioning of the surface energy fluxes was better represented by DRI and FS for the station of Vielsalm.

For Collelongo, ~~H was underestimated by the model~~ the model underestimated H during daytime and overestimated H during nighttime, except for DRI which demonstrated a good agreement with the observations. ~~Yet again~~Consequently, the model overestimated LE during daytime, except for DRI. The daily maximum H for DRI of 247 W m^{-2} was close to the observed value of 253 W m^{-2}, whereas CON and FS simulated much lower values of 159 W m^{-2} and 197 W m^{-2} respectively (Table 4). ~~The simulated LE~~ CON showed the largest ~~difference with the observed one using CON~~LE bias. Regarding BR, the simulation by DRI with a value of 1.35 was in very good agreement with the observations. The DRI simulation resulted in the ~~correct~~ least biased partitioning of the surface energy fluxes at Collelongo. ~~CON was not performing well in simulating the correct partitioning, while FS had already much improved as compared to CON~~However, FS showed that frequent atmospheric reinitialisations can add value.

In summary, ~~RN was underestimated by the model, whereas H was underestimated and LE was overestimated~~the model presented a good spatial variability of BR, and the agreement with observations was highest for FS. However, ~~DRI~~ when focusing on the sites that represent at least 50% of the land cover at the corresponding model grid cell, it was DRI that performed well for H at Vielsalm and for LE at Collelongo. For Colellongo, this resulted in ~~a correct~~ the least biased simulation of the

partitioning of the surface energy fluxes, translated into ~~an excellent value for BR . Least well simulated were CON and G.~~ a better BR estimate. The use of a daily reinitialised atmosphere and land surface improved the correct partitioning of the surface energy fluxes~~.~~, whereas the continuous land surface initialisation as in FS could not improve the representation of the surface energy fluxes~~for both stations with respect to DRI. The validation of G was not conclusive, as this parameter needs to be revised with an improved residue layer~~. The high values of G were caused by the large overestimation of the soil temperature and could be solved by implementing additional parameterisations for the canopy layer and soil surface layer. This implementation could alter the energy budget available for the turbulent fluxes, but this lies outside the scope of this study.

5   Conclusions

An assessment of three downscaling approaches has been performed using the regional climate model ALARO-0 coupled to the land surface model SURFEXv5, with lateral and initial boundary conditions from ERA-Interim. The simulations were applied for a 10-year period from 1991 to 2000, for a Western European domain. The performance of ALARO-0 with SURFEX has already been validated for NWP applications (Hamdi et al., 2014), ~~but not yet~~ and here we present an evaluation for long-term climate simulations.

We compared the ~~common~~ commonly used approach of a continuous climate simulation with two alternative ~~aprooaches~~ methods of frequently reinitialising the RCM ~~simulation towards its driving field~~boundary conditions, combined with either a daily reinitialised or continuous land surface. The use of a daily reinitialised atmosphere outperformed the continuous (CON) approach for winter and summer 2 m temperature, and detoriorated the summer precipitation. However, the use of a continuous ~~surface next to~~ land surface (FS) with a daily reinitialised atmosphere improved the summer precipitation ~~with respect to the relative to the full~~ continuous approach. Furthermore, it improved the winter 2 m temperature, whereas it resulted in a neutral impact on the summer 2 m temperature and the winter precipitation, despite a slight deterioration ~~at~~ over the Mediterranean. The SSTs were reinitialised daily together with the atmosphere, as compared to the monthly updated SSTs in the continuous approach.

The seasonal cycle of the 2 m temperature and precipitation was different for three selected subdomains that covered large climate variability. Both the ~~temperature~~ temperate climate of Mid-Europe and the continental climate of Eastern Europe indicated more seasonal variability than the Mediterranean climate of the Iberian Peninsula. The simulation of the 2 m temperature had improved when applying daily reinitialised atmosphere with continuous land surface, despite an overestimation of the summer 2 m temperature. ~~The model disagreed more for precipitation, because of the forcing towards the too wet driving field of~~ Precipitation biases were larger and are perhaps associated with the tendency for ERA-Interim to be wetter than E-OBS, and the low spatial coverage by the observations in some regions. It was clear that the agreement for the precipitation between the model and the observations was highest during summer, while other seasons showed stronger deviations.

During summer, the interaction between the land surface and the overlaying atmosphere is largest. The 2 m temperature interacts with the soil moisture and influences the partitioning of the surface energy fluxes. The daily reinitialisation of the atmosphere improved the ~~representation of a correct partitioning , though the latent heat was highly overestimated for Vielsalm~~

14

and resulted in a too low value as compared to the FLUXNET observations correct partitioning of the latent and sensible heat flux, although the biases were still quite large to be conclusive. Still, this approach outperformed the use of a continuous simulation. For a more comprehensive analysis, we recommend to include future research will consider including more FLUXNET stations. A more in-depth analysis on the interaction between 2 m temperature, precipitation, and surface energy fluxes can
5 reveal soil-moisture-temperature coupling (Jaeger et al., 2009), but this lies outside the scope of this study.

   In conclusion, this study demonstrated that the approach of a daily reinitialised atmosphere was superior over the full continuous approach. The use of a continuous surface next to a daily reinitialised atmosphere even improved the winter temperature and summer precipitation. The latter approach is highly recommended We recommend using FS in a setup with GCM forcing , as imperfect initial and lateral boundary conditions are applied for climate simulations with ALARO-0.
10

Appendix: References

ALADIN International Team: The ALADIN project: Mesoscale modelling seen as a basic tool for weather forecasting and atmospheric research, WMO Bulletin, 46, 317–324, 1997.

Anthes, R., Kuo, Y., Hsie, E., Low-Nam, S., and Bettge, T.: Estimation of espisodic and climatological skill and uncertainty in regional numerical models, Quaterly Journal of the Royal Meteorological Society, 115, 770–792, 1989.

Aubinet, M., Grelle, A., Ibrom, A., Rannik, U., Moncrieff, J., Foken, T., Kowalski, A. S., Martin, P. H., Berbigier, P., Bernhofer, C., Clement, R., Elbers, J., Granier, A., Grünwald, T., Morgenstern, K., Pilegaard, K., Rebmann, C., Snijders, W., Valentini, R., and Vesala, T.: Estimates of the annual net carbon and water exchange of forests: the EUROFLUX methodology, Advances in Ecological Research, 30, 113–175, 2000.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw U, K., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor and energy flux densities, Bulletin of the Americal Meteorological Society, 82, 2415–2435, 2001.

Best, M. J., Beljaars, A., Polcher, J., and Viterbo, P.: A proposed structure for coupling tiled surfaces with the planetary boundary layer, Journal of Hydrometeorology, 5, 1271–1278, 2004.

Boone, A. A. and Wetzel, P. J.: A simple method for modeling sub-grid soil texture variability for use in an atmospheric climate model, Journal of the Meteorological Society of Japan, 77, 317–333, 1999.

Bowen, I.: The ratio of heat losses by conduction and by evaporation from any water surface, Physical Review Letters, 27, 779–787, 1926.

Bubnová, R., Hello, G., Bénard, P., and Geleyn, J.-F.: Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/Aladin NWP system, Monthly Weather Review, 123, 515–535, 1995.

Christensen, J. H., Carter, T. R., Rummukainen, M., and Amanatidis, G.: Evaluating the performance and utility of regional climate models: the PRUDENCE project, Climate Change, 81, 1–6, 2007.

Davies, H. C.: A lateral boundary formulation for multi-level prediction models, Quaterly Journal of the Royal Meteorological Society, 102, 405–418, 1976.

De Troch, R., Hamdi, R., Van De Vyver, H., Geleyn, J.-F., and Termonia, P.: Multiscale performance of the ALARO-0 Model for simulating extreme summer precipitation climatology in Belgium, Journal of Climate, 26, 8895–8915, 2013.

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

Dickinson, R. E., Errico, R. M., Giorgi, F., and Bates, G. T.: A regional climate model for the Western United States, Climatic Change, 15, 383–422, 1989.

Faggian, P.: Climate Change Projections for Mediterranean Region with Focus over Alpine Region and Italy, Journal of Environmental Science and Engineering, 4, 482–500, 2015.

Gerard, L.: An integrated package for subgrid convection, clouds and precipitation compatible with the meso-gamma scales, Quaterly Journal of the Royal Meteorological Society, 133, 711–730, 2007.

Gerard, L. and Geleyn, J.-F.: Evolution of a subgrid deep convection parametrization in a limited area model with increasing resolution, Quaterly Journal of the Royal Meteorological Society, 131, 2293–2312, 2005.

Gerard, L., Piriou, J.-M., Brožková, R., Geleyn, J.-F., and Banciu, D.: Cloud and precipitation parameterization in a meso-gamma-scale operational weather prediction model, Monthly Weather Review, 137, 3960–3977, 2009.

5 Giorgi, F.: Regional climate modeling: status and perspectives, Journal of Physics, 139, 101–118, 2006.

Giorgi, F. and Mearns, L. O.: Introduction to special section: regional climate modeling revisited, Journal of Geophysical Research, 104, 6335–6352, 1999.

Giot, O., Termonia, P., Degrauwe, D., De Troch, R., Caluwaerts, S., Smet, G., Berckmans, J., Deckmyn, A., De Cruz, L., De Meutter, P., Duerinckx, A., Gerard, L., Hamdi, R., Van den Bergh, J., Van Ginderachter, M., and Van Schaeybroeck, B.: Validation of the ALARO-0

10 model within the EURO-CORDEX framework, Geoscientific Model Development, 9, 1143–1152, 2016.

Hamdi, R., Van de Vyver, H., and Termonia, P.: New cloud and microphysics parameterisation for use in high-resolution dynamical downscaling: application for summer extreme temperature over Belgium, International Journal of Climatology, 32, 2051–2065, 2012.

Hamdi, R., Degrauwe, D., Duerinckx, A., Cedilnik, J., Costa, V., Dalkilic, T., Essaouini, K., Jerczynki, M., Kocaman, F., Kullmann, L., Mahfouf, J.-F., Meier, F., Sassi, M., Schneider, S., Váňa, F., and Termonia, P.: Evaluating the performance of SURFEXv5 as a new land

15 surface scheme for the ALADINcy36 and ALARO-0 models, Geoscientific Model Development, 7, 23–39, 2014.

Haylock, M. R., Hofstra, N., Tank, A. M. G. K., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006, Journal of Geophysical Research, 113, D20 119, 2008.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski,

20 S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P.: EURO-CORDEX: new high-resolution climate change projections for European impact research, Regional Environmental Change, 14, 563–578, 2014.

Jaeger, E., Stöckli, R., and Seneviratne, S.: Analysis of planetary boundary layer fluxes and land-atmosphere coupling in the regional climate

25 model CLM, Journal of Geophysical Research, 114, 1–15, 2009.

Koster, R. D. and Suarez, M. J.: Soil moisture memory in climate models, Journal of Hydrometeorology, 2, 558–570, 2001.

Kotlarski, S., Hagemann, S., Krahe, P., Podzun, R., and Jacob, D.: The Elbe river flooding 2002 as seen by an extended regional climate model, Journal of Hydrology, 472-473, 169–183, 2012.

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E.,

30 Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V.: Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble, Geoscientific Model Development, 7, 1297–1333, 2014.

Leung, L. R., Mearns, L. O., Giorgi, F., and Wilby, R. L.: Regional climate research, Bulletin of the American Meteorological Society, 84, 89–95, 2003.

Lindstedt, D., Lind, P., Kjellström, E., and Jones, C.: A new regional climate model operating at the meso-gamma scale: performance over

35 Europe, Tellus, 67, 1–23, 2015.

Lo, J. C.-F., Yang, Z.-L., and Pielke Sr., R. A.: Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model, Journal of Geophyical Research, 113, 1–16, 2008.

Lucas-Picher, P., Boberg, F., Christensen, J. H., and Berg, P.: Dynamical downscaling with reinitializations: a method to generate finescale climate datasets suitable for impact studies, Journal of Hydrometeorology, 14, 1159–1174, 2013.

Masson, V.: A physically-based scheme for the urban energy budget in atmospheric models, Boundary-Layer Meteorology, 94, 357–397, 2000.

5   Masson, V., Champeaux, J.-L., Chauvin, F., Meriguet, C., and Lacaze, R.: A global database of land surface parameters at 1 km resolution in meteorological and climate models, Journal of Climate, 16, 1261–1282, 2003.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari,

10   M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voldoire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geoscientific Model Development, 6, 929–960, 2013.

Napoly, A., Boone, A., Samuelsson, P., Gollvik, S., Martin, E., Seferian, R., Carrer, D., Decharme, B., and Jarlan, L.: The interactions between Soil-Biosphere-Atmosphere (ISBA) land surface model Multi-Energy Balance (MEB) option in SURFEX - Part 2: Model evaluation for

15   local scale forest sites, Geoscientific Model Development, 2016, 1–40, 2016.

Noilhan, J. and Mahfouf, J.-F.: The ISBA land surface parameterisation scheme, Global and Planetary Change, 13, 145–159, 1996.

Noilhan, J. and Planton, S.: A simple parameterization of land surface processes for meteorological models, Monthly Weather Review, 117, 536–549, 1989.

Phillips, N. A.: The general circulation of the atmosphere: a numerical experiment, Quarterly Journal of the Royal Meteorological Society,

20   82, 123–164, 1956.

Qian, J.-H., Seth, A., and Zebiak, S.: Reinitialized versus continuous simulations for regional climate downscaling, Monthly Weather Review, 131, 2857–2874, 2003.

Rauscher, S. A., Coppola, E., Piani, C., and Giorgi, F.: Resolution effects on regional climate model simulations of seasonal precipitation over Europe, Climate Dynamics, 35, 685–711, 2010.

25   Ritter, B. and Geleyn, J.-F.: A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations, Monthly Weather Review, 120, 303–325, 1992.

Seneviratne, S. I., Lüthi, D., Litschi, M., and Schär, C.: Land-atmosphere coupling and climate change in Europe, Nature, 443, 205–209, 2006.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling., A. J.: Investigating soil moisture-

30   climate interactions in a changing climate: A review, Earth-Science Reviews, 99, 125–161, 2010.

Voldoire, A., Sanchez-Gomez, E., y Mélia, D. S., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A., Chevalier, M., Déqué, M., Deshayes, J., Douville, H., Fernandez, E., Madec, G., Maisonnave, E., Moine, M.-P., Planton, S., Saint-Martin, D., Szopa, S., Tyteca, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., and Chauvin, F.: The CNRM-CM5.1 global climate model: description and basic evaluation, Climate Dynamics, 40, 2091–2121, 2013.

35   von Storch, H., Langenberg, H., and Feser, F.: A spectral nudging technique for dynamical downscaling purposes, Monthly Weather Review, 128, 3664–3673, 2000.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., Field, C., Grelle, A., Ibrom, A., Law, B., Kowalski, A., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valentini, R., and Verma, S.: Energy balance closure at FLUXNET sites, Agricultural and forest meteorology, 113, 223–243, 2002.

Wilson, T., Meyers, T., Kochendorfer, J., Anderson, M., and Heuer, M.: The effect of soil surface litter residue on energy and carbon fluxes in a deciduous forest, Agricultural and forest meteorology, 161, 134–147, 2012.

5

Table 1. Overview of the FLUXNET eddy-covariance sites used in this study

Site and reference	Short	Lon(°E)	Lat(°N)	Alt(m)	Biome type	Years	Climate zone (Köppen)
Vielsalm	BEVie	6.00	50.31	491	Mixed	1996-2000	Maritime temperate (Cfb)
Collelongo	ITCol	13.59	41.85	1645	Deciduous	1996-2000	Humid subtropical (Cfa)
Brasschaat	BEBra	4.52	51.31	15	Deciduous	1997-2000	Maritime temperate (Cfb)
Loobos	NLLoo	5.74	52.17	25	Evergreen	1996-2000	Maritime temperate (Cfb)
Tharandt	DETha	13.57	50.96	320	Evergreen	1996-2000	Maritime temperate (Cfb)
Hesse	FRHes	7.07	48.67	293	Deciduous	1997-2000	Maritime temperate (Cfb)
Le Bray	FRLBr	-0.77	44.72	62	Evergreen	1996-2000	Maritime temperate (Cfb)

Table 2. The daily mean 2 m temperature bias (°C) and RMSE (in brackets) between the downscaled simulations and E-OBS for the total domain and the subdomains (BI, IP, FR, ME, AL, MD, EA) during DJF and JJA for the 10-year period 1991-2000.
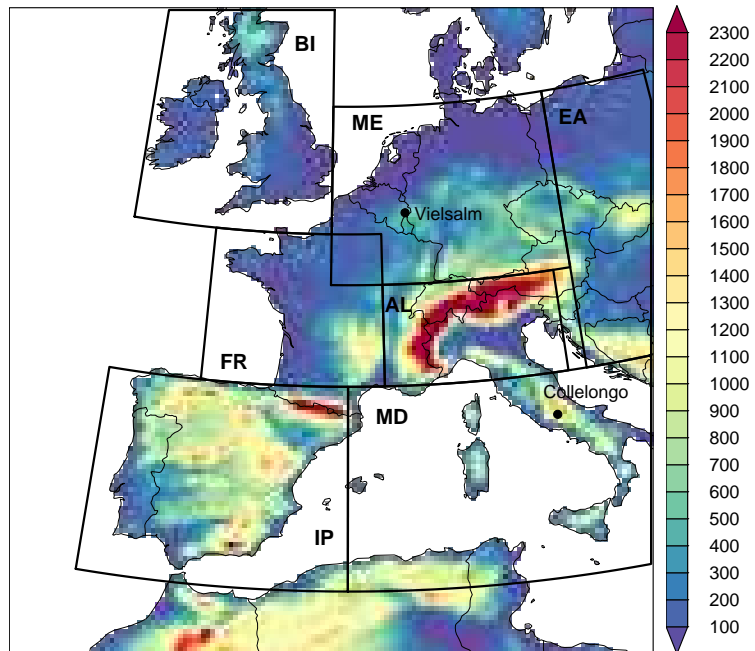
		TOTAL	BI	IP	FR	ME	AL	MD	EA
DJF	CON	-1.8 (2.5)	-1.1 (2.0)	-2.2 (2.7)	-1.5 (2.2)	-1.3 (2.0)	-3.0 (3.8)	-2.4 (3.1)	-1.1 (2.0)
	DRI	-1.2 (2.8)	-1.0 (2.7)	-1.6 (2.7)	-1.2 (2.9)	-0.7 (2.6)	-1.4 (3.4)	-2.1 (3.2)	-0.3 (2.8)
	FS	-1.0 (2.8)	-0.3 (2.8)	-1.3 (2.5)	-0.7 (2.8)	-0.4 (2.6)	-2.1 (3.8)	-1.2 (2.7)	-0.4 (2.8)
JJA	CON	-0.6 (2.0)	-1.7 (2.0)	-0.5 (1.7)	-1.2 (1.9)	-1.3 (1.9)	-1.8 (2.6)	-0.5 (2.0)	-0.5 (1.8)
	DRI	-0.1 (2.3)	-0.9 (2.0)	-0.3 (2.2)	-0.7 (2.4)	-0.3 (2.1)	-0.8 (2.3)	-0.6 (2.3)	0.0 (2.1)
	FS	0.9 (2.7)	-0.7 (2.2)	0.5 (2.4)	1.0 (3.1)	1.3 (2.8)	~~-0.0~~ 0.0 (2.5)	0.7 (2.5)	1.2 (2.8)

21

Table 3. The daily accumulated precipitation bias (%) and RMSE (in brackets) between the downscaled simulations and E-OBS for the total domain and the subdomains (BI, IP, FR, ME, AL, MD, EA) during DJF and JJA for the 10-year period 1991-2000.

		TOTAL	BI	IP	FR	ME	AL	MD	EA
DJF	CON	16.6 (3.8)	4.5 (4.5)	16.1 (4.6)	29.0 (3.6)	25.4 (2.7)	11.2 (4.7)	46.0 (6.2)	35.3 (2.3)
	DRI	20.9 (4.8)	6.6 (5.2)	21.2 (5.6)	26.8 (4.8)	27.9 (3.8)	24.1 (6.3)	41.6 (7.1)	45.7 (3.1)
	FS	36.3 (5.4)	16.9 (5.5)	31.3 (6.2)	38.2 (5.2)	35.7 (4.0)	26.7 (6.7)	108.5 (9.9)	64.1 (3.5)
JJA	CON	12.1 (4.2)	24.7 (4.4)	11.5 (2.9)	12.0 (4.4)	11.9 (5.0)	32.6 (7.3)	60.7 (3.5)	-2.6 (5.4)
	DRI	22.5 (4.7)	27.0 (4.7)	30.0 (3.4)	18.3 (5.1)	8.8 (5.5)	48.2 (8.9)	84.8 (3.8)	6.8 (5.9)
	FS	3.6 (4.5)	17.4 (4.6)	13.0 (3.2)	-7.0 (4.6)	-13.4 (5.1)	23.5 (8.3)	52.4 (3.6)	-8.2 (5.7)

Table 4. The daily maximum surface energy fluxes (Wm^{-2}) averaged over the 5-year JJA period 1996-2000 and the 10-year period 1991-2000 (in brackets).

|		RN	H	LE	G	BR
Vielsalm	OBS	417	151	134	11	1.12
	CON	395 (404)	118 (113)	250 (261)	47 (47)	0.47 (0.43)
	DRI	388 (398)	151 (159)	195 (193)	57 (58)	0.78 (0.82)
	FS	405 (411)	139 (152)	229 (221)	46 (49)	0.61 (0.69)
Collelongo	OBS	538	253	192	-1.39	1.32
	CON	480 (481)	159 (147)	270 (289)	111 (108)	0.59 (0.51)
	DRI	496 (494)	247 (232)	183 (194)	143 (140)	1.35 (1.19)
	FS	501 (498)	197 (191)	236 (247)	111 (110)	0.83 (0.77)

Figure 1. The total domain on 20 km horizontal resolution and the subdomains (BI, IP, FR, ME, AL, MD, EA) based on the subdomains selected in the EURO-CORDEX framework. The color represents the orography (m) in the ALARO+SURFEX setup. The two ~~black dots represent the~~ FLUXNET stations focused on in this study are Vielsalm (~~Belgium~~maritime temperate climate) and Collelongo (~~Italy~~humid subtropical climate).
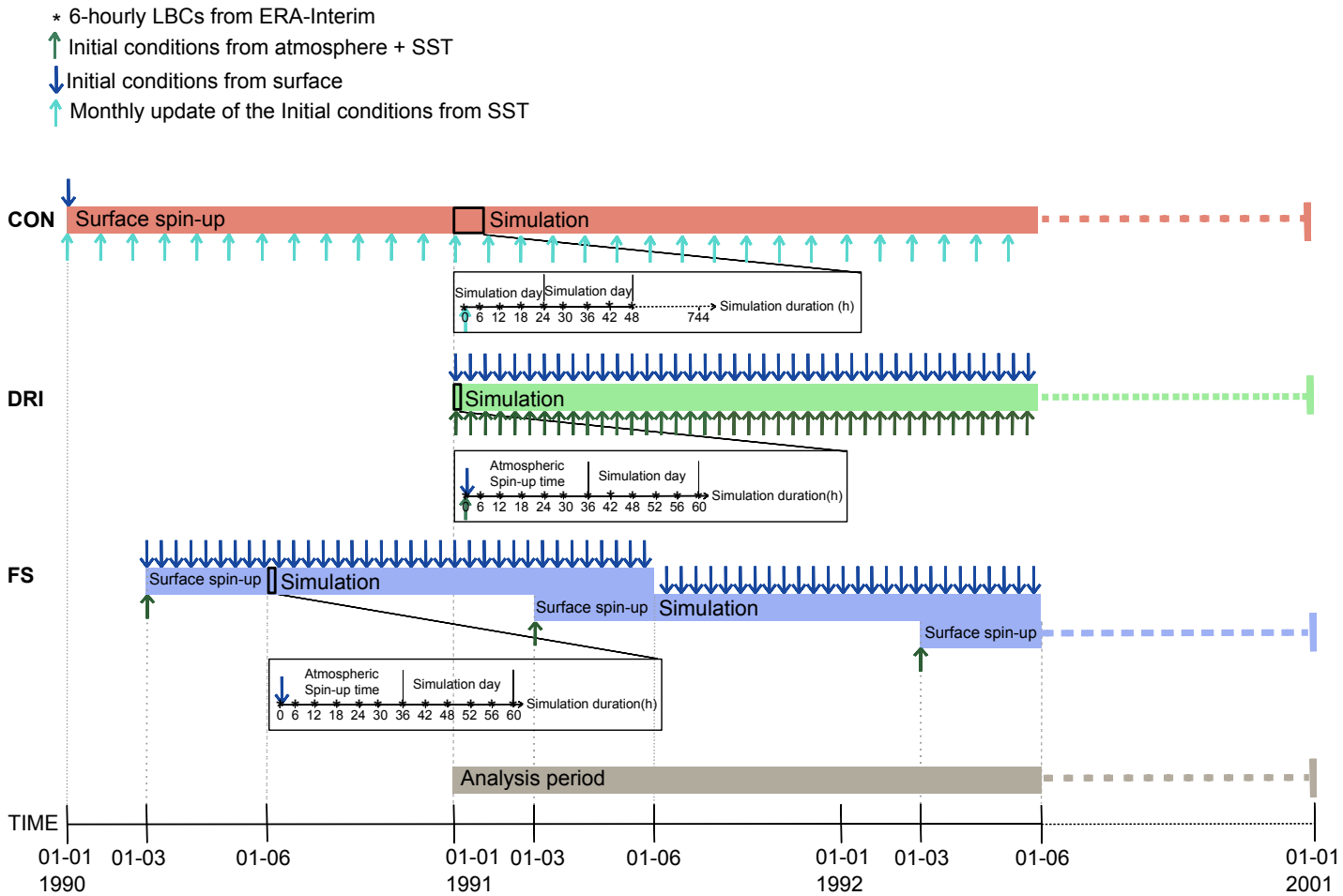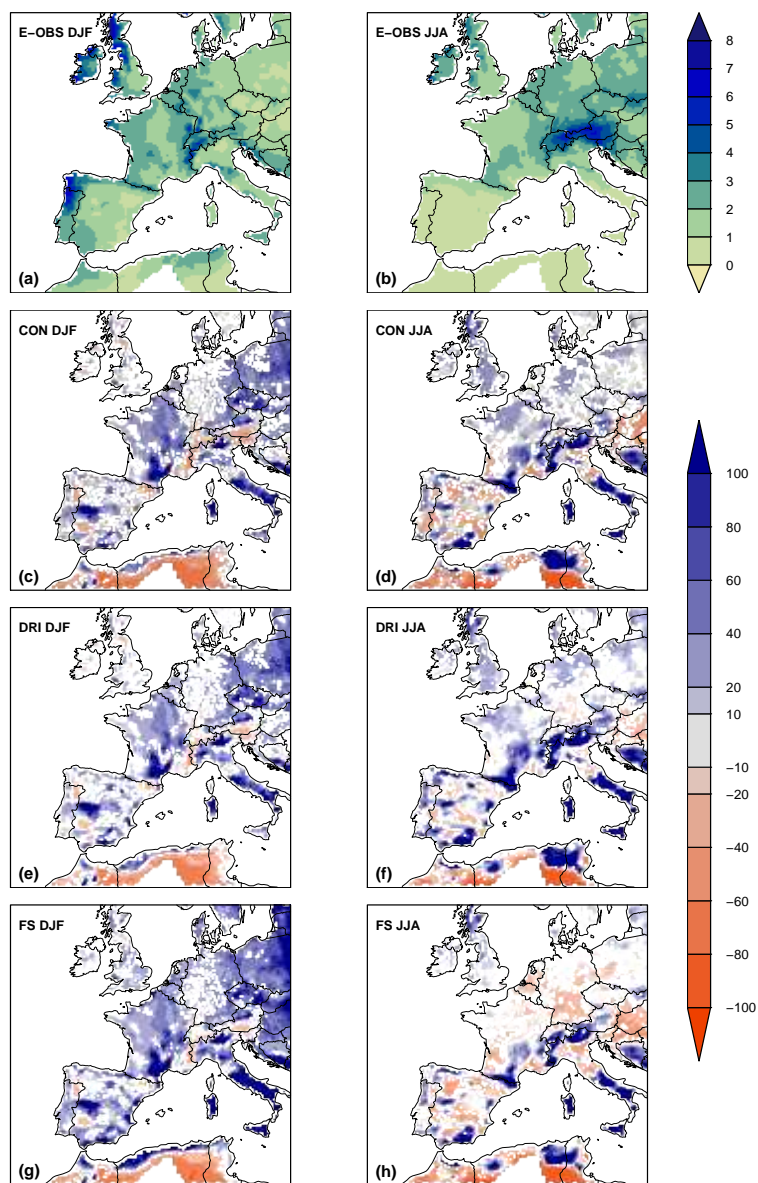
Figure 2. The setup of the three downscaling approaches CON, DRI and FS used in this study. It represents the spin-up time for the different simulations, the analysis period of the total experiment and the update frequency of the lateral and initial boundary conditions.

25
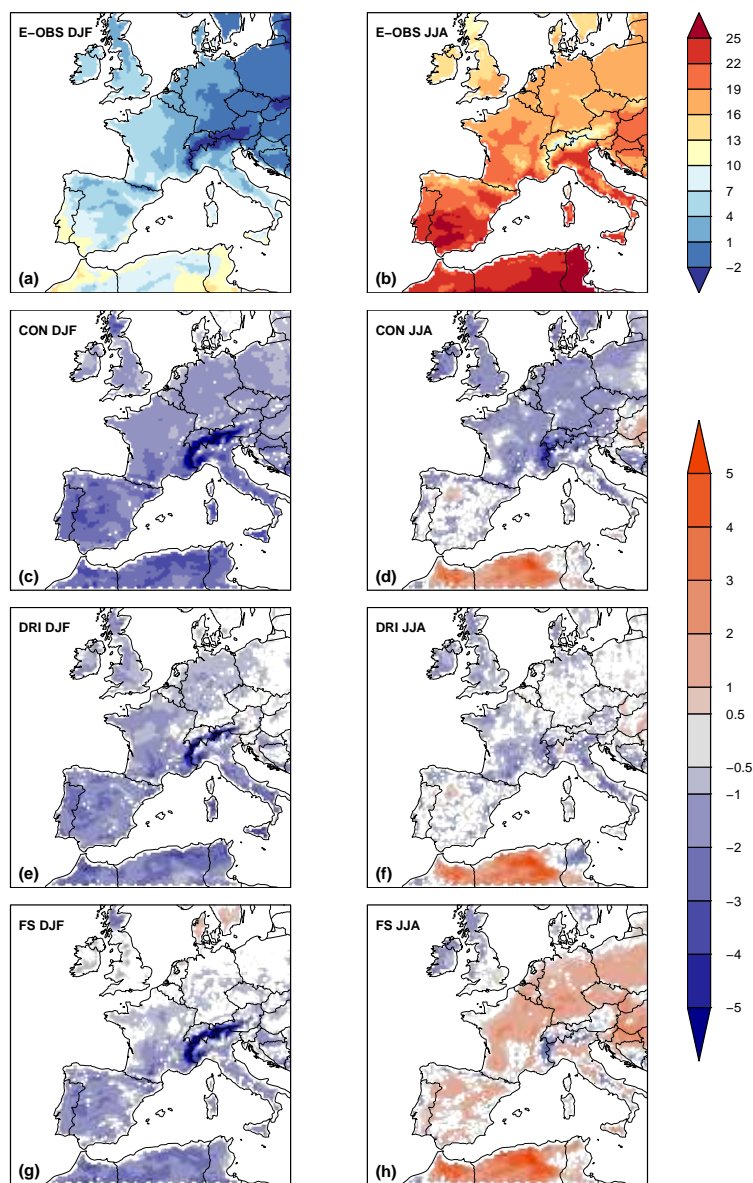
Figure 3. Daily mean 2 m temperature (°C) for E-OBS DJF (a) and JJA (b), and absolute bias (°C) of the model with E-OBS for CON DJF (c) and JJA (d), for DRI DJF (e) and JJA (f) and for FS DJF (g) and JJA (h), all at a 20 km horizontal resolution for the 10-year period 1991-2000. ~~The dots represent~~ Only the ~~grid points with a~~ significant ~~difference at 5%~~ biases are shown, using the Student's t-test ~~with~~ at a ~~null hypothesis stating that the means of the model~~ 5% level, and ~~observations~~ non-signicant biases are ~~equal~~ shown in white.

26

Figure 4. Daily accumulated precipitation (mm day^{-1}) for E-OBS DJF (a) and JJA (b), and relative bias (%) of the model with E-OBS for CON DJF (c) and JJA (d), for DRI DJF (e) and JJA (f) and for FS DJF (g) and JJA (h), all at a 20 km horizontal resolution for a 10-year period 1991-2000. ~~The dots represent~~ Only the ~~grid points with~~ significant ~~different variations at 5%~~biases are shown, using the F-test ~~with at~~ a ~~null hypothesis stating that the variances of the model~~ 5% level, and ~~observations~~ non-significant biases are ~~equal~~shown in white.
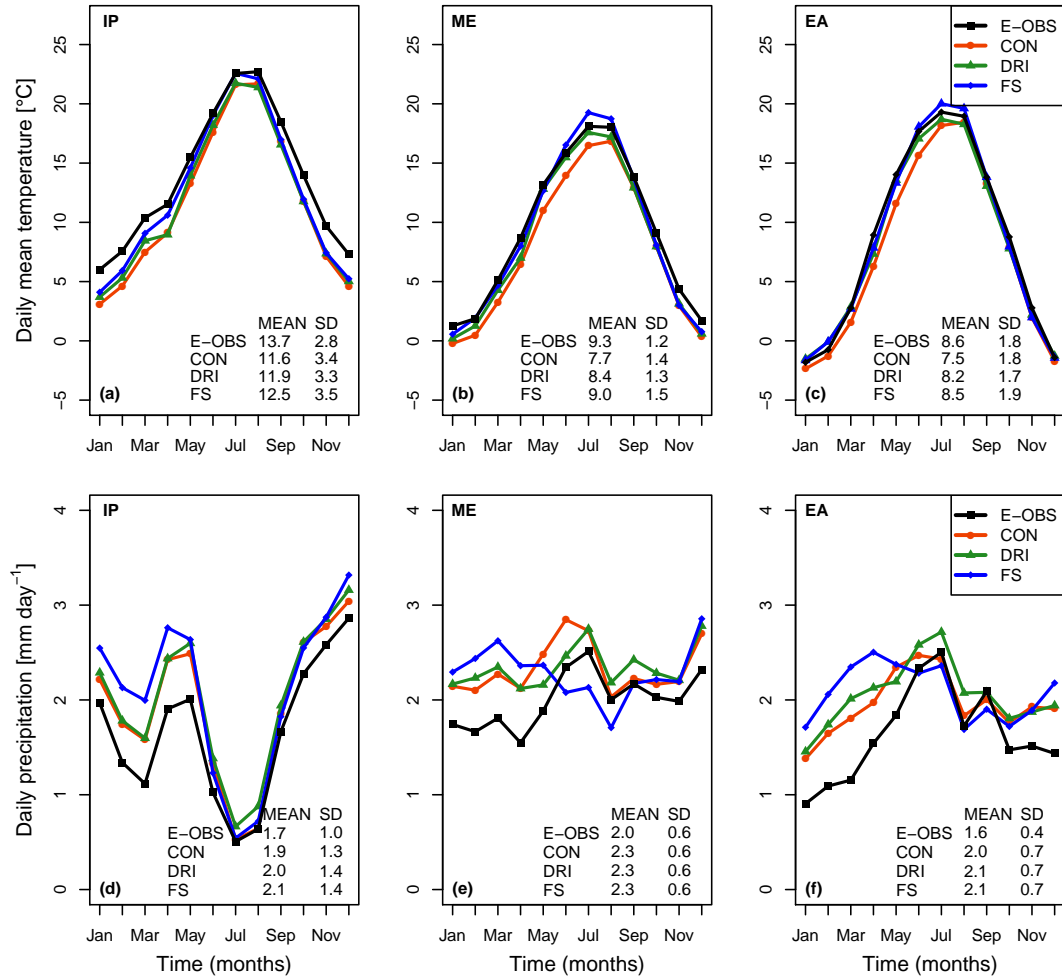
Figure 5. Mean annual cycle of the daily 2 m temperature (°C) with E-OBS, CON and FS for (a) the Iberian Peninsula, (b) Mid-Europe, and (c) Eastern Europe, and daily accumulated precipitation (mm day^{-1}) for (d) the Iberian Peninsula, (e) Mid-Europe, and (f) Eastern Europe, averaged over the 10-year period 1991-2000. Both the mean and standard deviation (SD) are displayed as text.
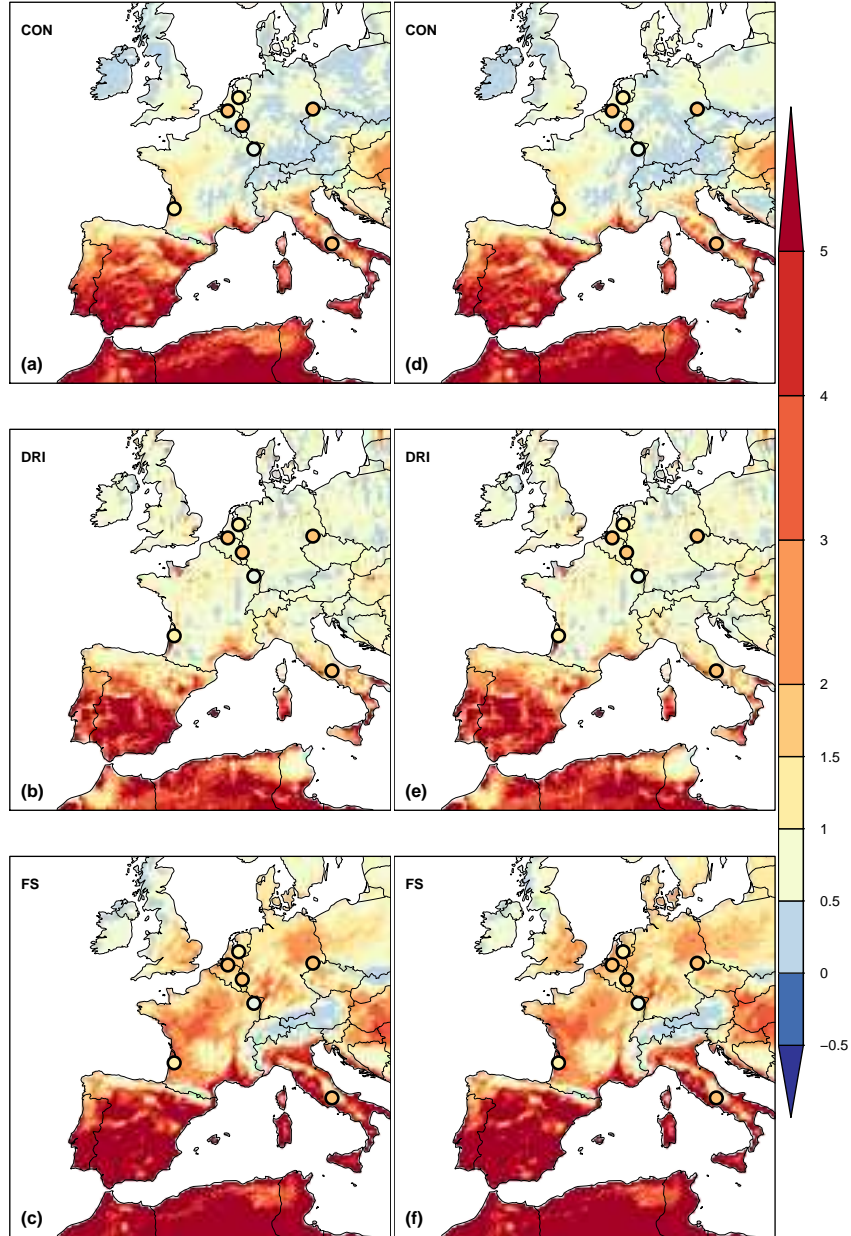
Figure 6. Daily maximum Bowen ratio averaged over the 5 year JJA period 1996-2000 for (a) CON, (c) DRI and (e) FS and averaged over the 10-year JJA period 1991-2000 for (b) CON, (d) DRI and (f) FS. The dots represent the values for the FLUXNET stations Vielsalm (Belgium) and Collelongo (Italy).
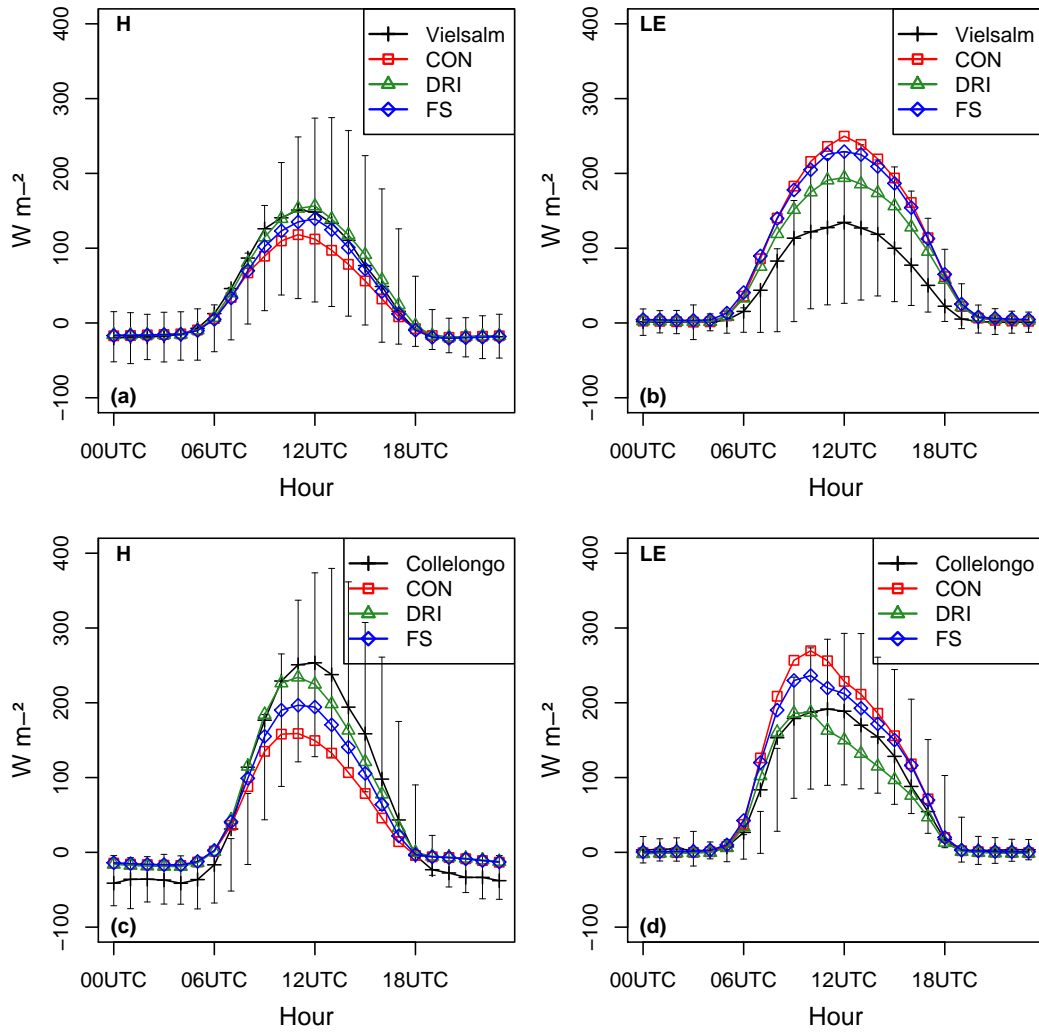
Figure 7. Daily cycle of the energy fluxes (W m^{-2}) in JJA 1996-2000 for Vielsalm in the top row and Collelongo in bottom row for (a,c) H, and (b,d) LE, for the FLUXNET observations and their corresponding model grid points by CON, DRI and FS. The error bars represent the ~~standard deviation~~ estimated uncertainties of the ~~observations~~observed turbulent fluxes.