Atmospheric Research 203 (2018) 68–82

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

What model resolution is required in climatological downscaling over complex terrain?

Renalda El-Samra[a,b], Elie Bou-Zeid[c], Mutasem El-Fadel[a,*]

[a] Department of Civil and Environmental Engineering, American University of Beirut, Beirut, Lebanon
[b] Department of Civil and Environmental Engineering, Rakik Hariri University, Meshref, Lebanon
[c] Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ, USA

ARTICLE INFO

Keywords:
High resolution dynamical downscaling
Mesoscale modeling over complex terrain
Regional climate model
Weather research and forecasting model

ABSTRACT

This study presents results from the Weather Research and Forecasting (WRF) model applied for climatological downscaling simulations over highly complex terrain along the Eastern Mediterranean. We sequentially downscale general circulation model results, for a mild and wet year (2003) and a hot and dry year (2010), to three local horizontal resolutions of 9, 3 and 1 km. Simulated near-surface hydrometeorological variables are compared at different time scales against data from an observational network over the study area comprising rain gauges, anemometers, and thermometers. The overall performance of WRF at 1 and 3 km horizontal resolution was satisfactory, with significant improvement over the 9 km downscaling simulation. The total yearly precipitation from WRF's 1 km and 3 km domains exhibited < 10% bias with respect to observational data. The errors in minimum and maximum temperatures were reduced by the downscaling, along with a high-quality delineation of temperature variability and extremes for both the 1 and 3 km resolution runs. Wind speeds, on the other hand, are generally overestimated for all model resolutions, in comparison with observational data, particularly on the coast (up to 50%) compared to inland stations (up to 40%). The findings therefore indicate that a 3 km resolution is sufficient for the downscaling, especially that it would allow more years and scenarios to be investigated compared to the higher 1 km resolution at the same computational effort. In addition, the results provide a quantitative measure of the potential errors for various hydrometeorological variables.

1. Introduction

The climate is warming due to increased radiative forcing produced by anthropogenic emissions of greenhouse gases (GHGs). Regardless of the future emission scenario adopted, this warming is likely to exacerbate water scarcity in many regions of the world and lead to various other adverse impacts on human socio-economic activities and well-being. Since preparing for these variations is the wise course of action, an understanding of climate change impacts at small temporal and spatial scales is imperative to guide policy and management decisions. This motive fuelled a dramatic progress in weather forecasting and climate modeling over the past 50 years, involving a multidisciplinary top-down approach whereby an emission scenario is postulated to force a Global Climate Model (GCM) simulation, also known as General Circulation Model, which is then downscaled using dynamical (via a Regional Climate Model (RCM)) or statistical methods (Lynch, 2008; Quintana Seguí et al., 2010). While GCMs simulate several facets of the climate system and corresponding interactions (Murphy et al., 2004), they are constrained by computational limitations to horizontal grid spacing on the order of hundreds of kilometers. This coarse resolution produces serious truncation errors in the numerical solution of the governing differential equations (Lynch, 2008). Other factors that downgrade the results of GCMs include processes that are not explicitly simulated at these resolutions, like convection, clouds and precipitation, heterogeneity of surface fluxes, and planetary boundary layer (PBL) turbulence. These subgrid-scale processes are "parameterized" instead, using a physical understanding of the underlying physics, or semi-empirical relations (Giorgi and Mearns, 1991). Hence, the coarse grids of GCMs preclude the accurate representation of a range of processes that are important to the hydrometeorological projections used in climate-change impact studies at the regional to local levels. Examples of these projections are extreme weather events (floods, droughts, heat waves) that have significantly influenced the global community in the last decade. In 2015, 10 weather and climate catastrophic events occurred with damages exceeding $1 billion each in the United States alone (National Climatic Data Center (NCDC), 2016). More than 30,000 deaths were attributed to the 2003 European heat wave (Robine et al., 2008), while nearly 55,000 deaths were caused by

* Corresponding author.
E-mail address: mfadel@aub.edu.lb (M. El-Fadel).

https://doi.org/10.1016/j.atmosres.2017.11.030
Received 7 November 2016; Received in revised form 7 May 2017; Accepted 29 November 2017
Available online 01 December 2017
0169-8095/

the Russian heat wave in 2010 (Katsafados et al., 2014). Since these extremes are of utmost importance when climate change impacts are being assessed, coarse resolution GCM results are downscaled by RCMs to (1) add high resolution element to resolve regional scale constraints (e.g. topography, coastal lines, and land use/land cover) that interact with the larger-scale atmospheric circulations (Giorgi, 2006), and (2) resolve some small scale processes that are parameterized in GCMs (such as convection). Downscaling using RCMs, referred to as dynamical downscaling, has been reported to represent regional climate characteristics more accurately than GCM simulations (Giorgi and Mearns, 1991; Wang et al., 2004; Argüeso et al., 2012; Berg et al., 2013; Mohan and Sati, 2016), and therefore is more reliable to force impact models (for example, crop and hydrological models) and develop adaptation and mitigation strategies. However, various gaps remain in our understanding of the forecasting skill and optimal setup of such downscaling simulations, especially over complex terrain (Salameh et al., 2009; Di Luca et al., 2015). Two particularly critical open questions are: What model resolution is required in climatological downscaling over a complex terrain? And how does the downscaling skill vary among the main hydrometeorological variables of interest: temperature, precipitation, and wind speed?

In this study, we aim to address these questions using the Weather Research and Forecasting (WRF) model during a hot and dry year, which represents the most adverse conditions under future climate change scenarios, as well as a mild and wet year. While the performance of the WRF model as a regional climate model has been widely tested (see Supplementary material Table S1), assessments of its skill in hydrometeorological downscaling during extreme years and over a complex terrain remain very limited and insufficient to answer our motivating questions. However, establishing this skill is needed to support downscaling simulations for future years.

2. Material and methods

2.1. Model domain: pilot study area

The country of Lebanon, which is located in the northern temperate zone along the eastern Mediterranean (Fig. 1a), represents a compelling example of the need for high-resolution climatology due to its geomorphological complexity, as well as its large climate gradients (Fig. 1b). It was selected as the testing domain, with several observational stations (Fig. 1c) for validation. Downscaling simulations for future years becomes possible once the model is evaluated using historic periods, and the same set-up was used as the corner stone for future climate change downscaling simulations (El-Samra et al., 2017).

2.2. Model description

The Weather Research and Forecasting (WRF) version 3.4.1 is a three-dimensional mesoscale meteorology model, which is suitable for operational forecasting and atmospheric research needs (Heikkilä et al., 2011). It has a wide range of options for parameterizing physical processes, as well as various schemes for the numerical discretization of the governing equations (Skamarock et al., 2008). The model can directly extract, from several databases, information and data on terrain elevation, land cover and land use from the United States Geological Survey (USGS) and the Moderate Resolution Imaging Spectroradiometer (MODIS) (Friedl et al., 2001) at various resolutions that cover the entire globe. The finest resolution of these global data is 30 s in both latitudinal and longitudinal directions, which corresponds to about 1 km in length at mid-latitudes.

2.3. Model configuration

A regional high resolution is achieved in the current simulations by using three one-way nested domains (Fig. 1a), with 9, 3 and 1 km horizontal resolutions. The outer integration domain covers 1350 km × 1700 km to guarantee that synoptic-to-mesoscale systems that affect the coast are resolved in WRF. The inner domain extends over 154 km × 193 km (Fig. 1c), covering the entire study area. MODIS (Friedl et al., 2001) land use data was adopted with 21 land categories and Lambert Conformal projection, which is the most convenient for mid-latitude regions and provides homogeneous grid spacing, unlike latitude-longitude coordinates. A 30 second time step was used for the smallest domain. For the base cases, all domains had 35 vertical levels (with a vertically-stretched grid) arranged according to terrain-following hydrostatic pressure coordinates.

The National Center for Atmospheric Research-National Center for Environmental Prediction Final Analysis Data (National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce, 2000), which are global atmospheric fields with a 1-degree resolution and 26 pressure levels (1000–10 hPa), were used for initial, and boundary conditions. The time interval of the boundary data is 6 h, while the sea surface temperature (SST) was

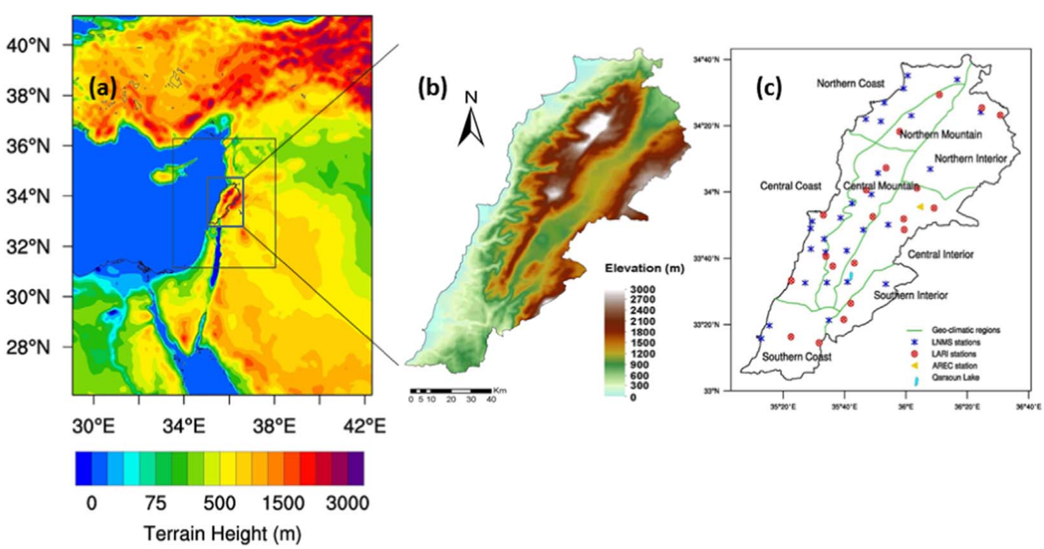

Fig. 1. (a) Eastern Mediterranean basin showing WRF's 3 domains (9:3:1 km) configuration with terrain elevation, (b) topographic features of the study area, (c) smallest WRF domain d03 (1 km) configuration depicting the observational stations with available data for 2003 and 2010 by geo-climatic region (43 rain, 2 wind and 31 temperature stations with records of both daily average and/or daily maximum and minimum temperatures).

updated once monthly based on previous studies that conclude that 6-hourly and daily SST does not vary significantly in the Eastern Mediterranean basin (Abudaya, 2013; Tonbol and El-Geziry, 2015). Daily updates of the SST will result in errors < 1 °C; hence it will not impact WRF results significantly. No nudging was adopted during the simulation to avoid biasing the validation (since in future downscaling, nudging data are not available) by "pulling" the model towards the observed conditions (Pielke et al., 2012; Xu and Yang, 2015), and making it difficult to assess whether WRF can describe features like air-mass formation that build within the domain (Paimazumder et al., 2012). All the WRF one-month-simulations were initialized on the first of each month of years 2003 and 2010 to prevent the regional model from drifting far from the observed weather by reinitializing it from the driving reanalysis data on a monthly basis. This method has the additional benefit of allowing simultaneous simulations for several months, significantly reducing the wall-clock simulation time (Pan et al., 1999; Caldwell et al., 2009). The best practice for model spin up under such conditions is still debated (Pan et al., 1999; Ferreira et al., 2014): spin up could under various conditions produce surface states that have more discrepancies compared to actual states if the modeling system does not assimilate observed data while the initialization product does. Longer spin up does not necessarily improve agreement with observations and in fact for specific humidity and temperature it seems that the shortest spin up works better as shown in our previous studies (Li et al., 2013). Moreover, for the one month simulations, initialization and spin up will only impact the first few days and thus the monthly averaged errors will not be affected significantly. In this study, we opted not to include spin-up periods in the monthly simulations because (i) the National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce (2000) surface initialization contains data assimilation for soil moisture and temperature and spin up is not required to set the surface state, (ii) analyses of various spinup periods generally show that they have a limited impact on meteorological fields (Ryu et al., 2016), and (iii) the effect of spin up would only affect the first few days of our monthly runs and will thus have a limited impact on the model evaluation we present later. We should also note that we initialize, run, and evaluate each month separately. Therefore, abrupt jumps in sea surface temperature or soil moisture and temperature from last day of a month to the first day of the subsequent month are inconsequential for the simulations.

While the year 2010 was a relatively hot and dry year that is used here as an example of an extreme period, we also compare the modeling results to a mild and wet year, 2003. The years 2003 and 2010 were selected precisely to have extreme hydrometeorological events such as heat waves, strong storms, or drought periods that pose particular challenges in dynamic modeling. The selection was based on continuous daily data spanning the period from 2001 till 2010. WRF simulations of historically wet and dry years also allow "measuring" the sensitivity of complex topographical regions to climate extremes, which are good proxies for the conditions expected in the future as the climate means shift.

The parameterization schemes adopted include the Single-Moment 6-Class Microphysics Scheme (WSM6) (Hong and Lim, 2006), Monin-Obukhov and Mellor-Yamada Janjic (Eta) for surface layer and PBL physics (Mellor and Yamada, 1974; Janjic, 2002) (we also evaluated the YSU scheme, as we will detail later), Dudhia Short Wave (SW) for radiative processes (Dudhia, 1989), Rapid Radiative Transfer Model (RRTM) (Mlawer et al., 1997) for Long Wave (LW) for radiative processes and the Noah Land Surface Model for surface processes (Chen and Dudhia, 2001). The choices are motivated by previous WRF tests detailed in Talbot et al. (2012) and Li et al. (2013), and various other studies that examined the influence of the choice of parameterizations on WRF's performance (Bukovsky and Karoly, 2009; Ikeda et al., 2010; Ruiz et al., 2010; Argüeso et al., 2011; Liu et al., 2011, Remesan et al., 2014).

2.4. Observational data

The duration, quality, and exhaustiveness of several climatic data sources (Atlas Climatique du Liban; NOAA's National Climatic Data Center, Lebanese National Meteorological Services (LNMS), Lebanese Agricultural Research Institute (LARI), American University of Beirut Advancing Research Enabling Communities Center (AREC)) were assessed to identify the spatial and temporal climatic data that can be relied upon. Based on long-term trends of weather parameters including temperature, relative humidity, precipitation, and wind, the Lebanese National Meteorological Services (LNMS) divides the study area into three broad climatic regions: the coastal, the mountainous and the inland. These are further subdivided into sub-regions making up a total of eight geoclimatic regions (Fig. 1c):

– The coastal strip, which includes the northern (NC), central (CS) and southern coasts (SC);
– The mountainous area, which is divided into the northern (NM) and central mountains (CM);
– The inland region, which is divided into the northern (NI), central (CI) and southern (SI) areas.

The analysis of climatic records from weather stations in the study area (Fig. 1c) indicated variation in span and quality. The availability of continuous observations was a hindrance since we could get complete daily data for a limited number of stations for the various geoclimatic regions. We selected only the stations from which 70% or more of the daily and/or monthly precipitation, temperature and wind measurements were available during 2003 and/or 2010 (see Supplementary material Table S2), which required the exclusion from WRF's output analysis of the periods with missing data for each selected station before model evaluation statistics were carried out. Comparisons between simulated and observed 2 m temperatures, precipitation, and 10 m wind, were then performed to answer the driving question of this study.

2.5. Data handling and analysis

Model simulations from WRF in its three resolutions (9, 3 and 1 km) were compared with observations at various stations throughout the inner domain. While model validation is imperative, there is no agreement on an ideal evaluation technique. Traditionally, RCMs evaluation procedures rely on gridded data or reanalysis (Argüeso et al., 2011; Berg et al., 2013). However, these datasets are sometimes created using a rather small number of stations, and the resulting interpolation can lead to excessively smoothed precipitation and temperature values that do not capture the extremes (Hofstra et al., 2009).

In this study, several statistics were used to evaluate model performance using observational datasets such as the mean bias error (MBE), mean absolute error (MAE), percentage bias (PBIAS), root mean square error (RMSE), percentage RMSE, and coefficient of determination R2 (Willmott, 1982; Kobayashi and Salam, 2000). Assessment of the contribution to the RMSE of (i) the bias of the mean (referred to as BIAS), (ii) the variance or standard deviation (referred to as VAR) and (iii) the covariance or dispersion error (referred to as COVAR) are also conducted following the approaches of Murphy (1988) and Horvath et al. (2012) as expressed in Eqs. (1a) and (1b).

$$MSE = \frac{1}{MN}\sum_{k=1}^{M}\sum_{i=1}^{N}(x_{i,k} - y_{i,k})^2 \tag{1a}$$

$$MSE = \frac{1}{M}\sum_{k=1}^{M}(\overline{x}_k - \overline{y}_k)^2 + [\sigma_k(x) - \sigma_k(y)]^2 + 2\sigma_k(x)\sigma_k(y)[1 - r_k(x,y)] \tag{1b}$$

where MSE is the mean square error, x and y are respectively the simulated and observed data, k and i are indices denoting various points in space and time, M is the number of stations per geo-climatic region, N is the number of points in the time series being compared, σ is the standard deviation, r is the correlation coefficient between simulated and observed data, and bars indicate time-averages. In this equation, the first term on the right hand side is the BIAS that denotes errors related to differences in the means between observations and model outputs, the second is the VAR that reflects errors resulting from the differences in variability between the two, while the third is the COVAR that is related to the synchronization errors between observation and simulated time series.

A regional statistical analysis was performed on the limited number of stations corresponding to every geo-climatic zone. Depiction of extreme events is through the probability density functions (PDFs) of the temperature and precipitation. Comparisons concentrated on daily average, maximum and minimum temperatures (Tavg, Tmax and Tmin) and accumulated total precipitation (daily, monthly and yearly). The daily average, maximum and minimum statistics are derived from the WRF simulations (hourly data) to assess the value of WRF downscaling. The simulated values are taken at the center of the grid cell for temperature and precipitation. Interpolation to match the exact observational location would result in minor differences in the simulated values derived from the smallest grids (WRF 1 km and 3 km), which are the focus of this study. For the same reason, no lapse rate correction was applied to temperature at the fine grid resolutions. Furthermore, since the lapse rate over complex topographical terrain does not depend only on altitude but is also a function of other aspects such as buoyancy, mountain width, moisture (Barstad and Smith, 2005) and wind (Esteban and Chen, 2008), estimating an altitude correction confidently is challenging. The use of a uniform and constant lapse rate of 6–6.5 °C/km is reportedly not illustrative of real surface environments over complex topographic terrains based on the use of temperature sensors (Bolstad et al., 1998; Rolland, 2003; Tang and Fang, 2006; Blandford et al., 2008; Gardner et al., 2009; Minder et al., 2010). Such observational studies report that the average surface lapse rate varies considerably from the 6–6.5 °C/km values often used to correct for altitude discrepancies between grid centers and observational points, displaying noticeable seasonal changes beyond 2 °C/km, diurnal erraticism and spatial fluctuations linked to the topography or position with respect to valleys (Minder et al., 2010). In light of the above, we opted not to apply a lapse rate correction to the 9 km domain either since it is highly uncertain whether such a correction improves or degrades the results.

In the case of wind, the model evaluation is more complex since altitude as well as small local topographic features around the station can influence the model validation. Moreover, only 1 station recorded hourly wind speed and direction in the central coast sub-region, while 2 stations on the northern coast and the central interior zone recorded these data on a 3-hour interval. After testing whether the wind data at these stations follow a Weibull probability distribution (Di Piazza et al., 2009), which is expected if the stations are not influenced by local obstacles or very location-specific features, the central coast station failed the test and only the last 2 stations were retained for wind evaluation for 2003 and 2010. While it is possible to have wind statistics at some locations that do not follow the Weibull probability distribution, the fact that two stations in the study region followed it while one did not, casts doubts on the wind data quality at that central coast station, and thus it was excluded from the analysis.

3. Results and discussion

3.1. Temperature

3.1.1. Average temperatures

The simulated impact of topography and altitude on the spatial patterns of yearly-averaged 2 m temperature, and the clear distinction between Tavg in wet and cool (2003) versus dry and hot (2010) years in all three WRF resolutions, are depicted in Fig. 2. The coast is the warmest region, followed by a small zone of the mid-mountain ranges, and the majority of the area of the inland central and northern regions. WRF reproduced the expected average temperature drop at high elevations of the mountain ranges and in the southern inland regions with higher spatial detail as the resolution is made finer from 9 to 1 km. The average temperature of the coast varied between 18 and 22 °C, which are consistent with reported values of 20 °C (Atlas Climatique du Liban, 1977; Lebanon SNC, 2011). Over the coast and in interior regions, average temperatures in 2010 (18 °C) are only slightly higher than in 2003 (17 °C).

Since the study area has an extended and narrow form in addition to its complex topographical features, local variations tend to be large. Therefore, the model performance over individual sub-regions is more informative than its countrywide performance. The number of weather stations measuring Tavg in each of the eight geo-climatic regions varies from 0 to 4, since most stations for 2003 have records only for Tmax and Tmin (refer to supplementary material Table S2). In order to compute the regional averages, the stations that report Tavg in each climate zone were grouped and then averaged (WRF results are also taken at the station locations only, and then averaged). The MAE and the MBE of the daily-averaged 2 m temperature for the different geo-climatic regions, categorized during various seasons of 2003 and 2010, are improved as WRF's grid is refined down to 1 km (Tables 1 and 2 for MAE and Tables 3 and 4 for MBE) (DJF: December, January, February; MAM: March, April, May; JJA: June, July, August; SON: September, October, December). They confirm that higher resolution leads to higher coefficients of determination and smaller errors.

Fig. 3 illustrates that the WRF simulation with 1 km resolution also outperformed the 9 km resolution in all regions, resulting in an average reduction in the RMSE of 0.47 °C for 2003 and 0.41 °C for 2010. WRF 1 km errors were also smaller than those of WRF 3 km, but the differences were less significant with only 0.16 and 0.12 °C improvement in the RMSE for 2003 and 2010, respectively. In most regions and seasons, the largest part of the RMSE error was due to the covariance term in all WRF resolutions, with few exceptions in the CC during MAM, CI during MAM, JJA and SON. This indicates that weather systems in WRF are arriving too early or too late compared to observations (note that since these are daily values, timing errors of less than one day cannot be detected). The BIAS was the second largest contributor to the error overall. On the other hand, the smallest part of the error is due to the variance (standard deviation) in all regions during all seasons, implying that the observed time variability (variance) is reproduced well by WRF, which is a direct representation of WRF's skill in reproducing the annual cycle. We are showing the decomposition of RMSE errors of WRF-1 only (Fig. 3) since the remaining resolutions follow the same trend. Fig. S3-1 to 3-4 in the supplementary material illustrates the results using Taylor diagrams (Taylor, 2001).

The PDFs of the daily-averaged temperatures from the observations, and from the 9, 3 and 1 km WRF simulations, for all four seasons of 2003 and 2010, were computed for all stations combined, and for conciseness only the best and worst seasonal performance of WRF (based on differences between WRF-1 simulated and observed Tavg) are presented in Fig. 4. The general trends of the distribution of WRF in its 3 resolutions match observations reasonably well, with WRF's finest resolution in general yielding PDFs closer to observations for all seasons. The satisfactory reproduction of these PDFs, particularly their tails, is a positive indication of WRF's ability to reproduce extremes.

3.1.2. Maximum temperatures

Seasonal and regional MAE analyses for the summer season for maximum 2 m temperatures, along with the correlation coefficients are depicted in Table 5. MAEs for WRF 1 and 3 km are lowest in all regions, except in CM during 2003 and NM during 2010. The coefficients of determination are closely comparable among the 3 WRF resolutions. In addition, the RMSEs of Tmax for years 2003 and 2010 were computed from differences in the daily maximum values between WRF and station

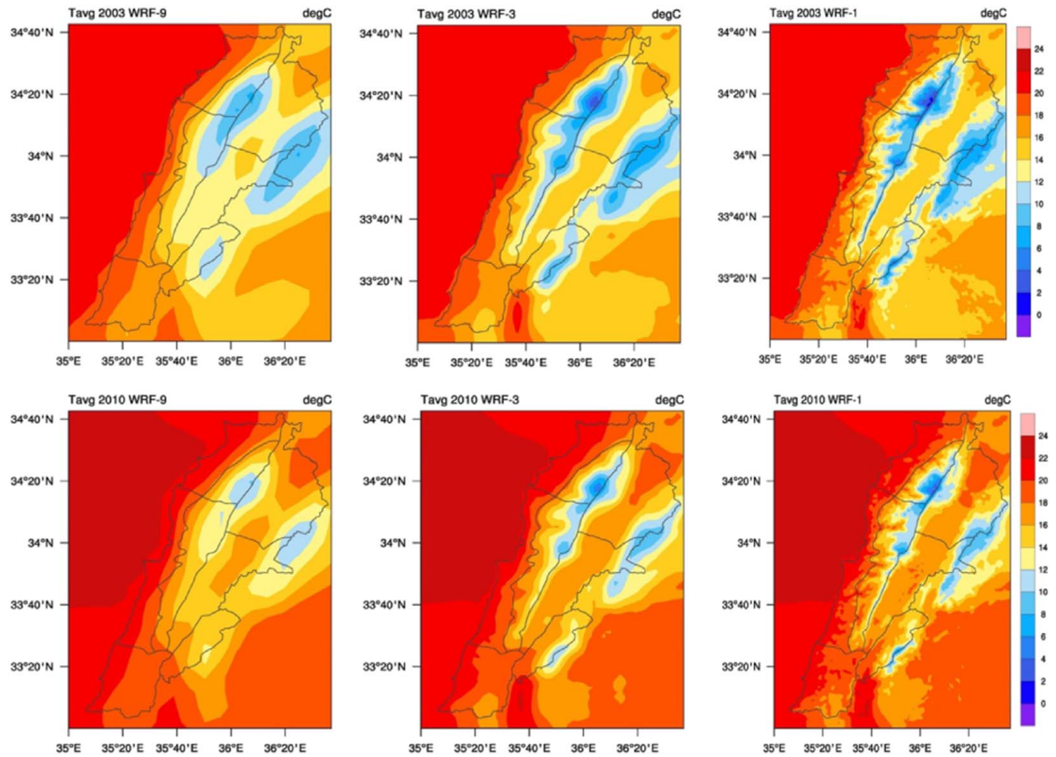

Fig. 2. Average 2 m temperature from WRF-9, 3 & 1 simulations for 2003 and 2010.

observations (Fig. 5), but we will only consider the warm summer month since this is the critical periods for high temperatures. In both years, WRF-1 is yielding smaller errors in all regions except the central mountains in 2003. WRF-3 is also markedly better than WRF-9 in all regions for both years. Interestingly, WRF-3 outperforms WRF-1 biases in the northern mountain and northern inland regions in 2010, and the differences between these two finest resolutions remain limited. Regarding the decomposition of the RMSE in time and space, the 3 model resolutions exhibited similar trends. For clarity, we are showing only the decomposition of RMSE related to WRF 1 in Fig. 5. Once again, the errors related to the variance (standard deviation) are minimal in all regions during both years. Interestingly, and unlike the mean temperature, the largest contribution to the RMSE came from the bias of the mean during 2003 and most of 2010, with the exception of the mountainous regions (NM and CM) and northern and southern inland regions (NI and SI) where the contribution of the covariance error is the greatest. This implies that WRF errors in underpredicting or overpredicting the maximum temperatures are overall more significant than the errors in the timing of these extremes (again the minimal time error that can be detected is 1 day). An alternative way of showing the above results using Taylor diagrams can be found in the Supplementary material Fig. S3-5 and S3-6.

The PDFs of daily maximum temperatures for the 2003 and 2010 summer season, comparing WRF simulations to observations, are presented in Fig. 6. The summer season is normally defined as June, July and August (JJA), but for the computation of these PDFs we opted to include the month of May (M) in the summer season: early in May the Khamsin wind blows up the Levantine coast from the Sahara Desert and could yield high Tmax values. The temperature of the wind might exceed 40 °C, conveying a heat wave over the study area (Blanchet, 1965). The added value of finer resolutions (WRF-1 and WRF-3) can be detected in the PDFs of MJJA of 2003 and 2010, although WRF tends to overestimate the frequency of Tmax < 30 °C, and underestimate the frequency of Tmax > 30 °C (particularly in 2003). A significant underestimation is noted in bins between 30 °C and 35 °C. Higher WRF resolutions can only mitigate some of this misrepresentation: WRF-1 and WRF-3 captured the observed hot season PDF of 2010 well, while WRF-9 continued to exhibit an underestimation between 35 °C and 40 °C bins. In general, the match between the simulated and observed PDFs of Tmax is inferior to the match for Tavg.

3.1.3. Minimum temperatures

A summary of the regional MAE analysis is shown in Table 6 for Tmin. MAE values indicate a good representation of Tmin by the WRF-1 during both years. In most regions, the error is improved by WRF-1 by ~0.4 °C during 2003 and by 1.0 °C during 2010, when compared with

Table 1
MAE (°C) and R^2 for seasonally Tavg in 2003, lowest errors are in bold*.

Region	WRF-1				WRF-3				WRF-9			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
NC	**1.1**	**1.5**	**1.3**	**1.2**	1.2	1.7	1.7	1.3	1.4	1.7	1.4	1.4
CC	**0.7**	2.0	**2.5**	**0.7**	**0.7**	**1.8**	2.6	0.8	0.9	2.0	2.8	0.9
CI	**1.4**	**1.4**	**1.3**	**1.4**	1.8	1.7	1.5	1.8	2.0	2.3	2.7	2.5
Study area	**1.1**	**1.6**	**1.7**	**1.1**	1.2	1.7	1.9	1.3	1.4	2.0	2.3	1.6
R^2	**0.75**	**0.91**	0.75	**0.90**	0.74	0.90	0.73	0.88	0.71	0.89	**0.83**	0.86

No stations recorded Tavg in SC, NM, CM, NI and SI regions.

Table 2
MAE (°C) and R^2 for seasonally Tavg in 2010, lowest errors are in bold.

Region	WRF-1				WRF-3				WRF-9			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
NC	**1.7**	**1.4**	**2.0**	**1.6**	1.7	1.6	2.3	1.6	1.9	1.5	2.2	1.7
CC	**0.8**	**1.5**	**3.1**	**1.0**	**0.8**	1.6	**3.1**	1.1	1.2	1.8	3.4	1.4
SC	**1.3**	1.0	1.7	**1.3**	1.3	**0.8**	**1.5**	1.3	1.5	1.1	1.7	1.4
NM	**0.7**	**0.7**	**0.9**	**0.8**	1.2	0.8	1.0	1.3	1.7	1.3	1.2	1.9
CM	**1.1**	**1.0**	**1.2**	**1.2**	1.8	1.4	1.3	1.7	1.6	1.5	1.3	1.7
NI	**1.4**	**1.3**	**1.2**	**1.2**	1.4	1.4	1.2	1.3	2.4	1.6	1.4	1.7
CI	1.7	1.6	1.6	1.5	**1.7**	**1.5**	**1.5**	**1.3**	1.8	1.9	2.1	1.5
SI	**0.7**	**0.7**	1.2	**0.8**	0.7	0.8	**1.2**	0.9	1.0	1.3	1.8	1.3
Study area	**1.2**	**1.1**	**1.6**	**1.2**	1.3	1.2	1.7	1.3	1.6	1.5	1.9	1.6
R^2	**0.82**	**0.87**	**0.77**	**0.88**	0.81	0.86	0.76	0.86	0.78	0.85	0.75	0.84

WRF-9; WRF-3 is also better than WRF-9 in most regions by ~0.2 °C during 2003 and by 0.7 °C during 2010. The RMSE of winter daily Tmin at a regional and a seasonal scale are presented in Fig. 7 and as a Taylor diagram in Fig. S3-6 in the Supplementary material. The good skill of WRF-1 in capturing Tmin is noticeable in most regions, except in the central mountain (CM) in 2003 and central inland (CI) in 2010 where WRF-1 yields the largest errors of all three resolutions. WRF-3 also yields lower errors than WRF-1 in the central coast (CC) 2003 and southern inland (SI) in 2010. Regarding the decomposition of the RMSE, Tmin follows the same trend as Tavg (rather than Tmax) in all of WRF 3 resolutions (only WRF 1 km is shown in Fig. 7). The errors related to the variance (standard deviation) are minimal (except in NI during 2003) and the highest contributions to the RMSE were from the covariance error in most regions during both years with a few exceptions.

The coefficients of determination of Tmin between observations and WRF simulations are always lower than those of Tmax (Tables 5 and 6). A possible cause that could account for these differences is that night-time temperature minima in the diurnal cycle are difficult to capture: this is a limitation with atmospheric models such as WRF that has often been reported before (Paimazumder et al., 2012; Talbot et al., 2012). It is related to the difficulty in parameterizing PBL turbulence under stable night-time conditions (Huang and Bou-Zeid, 2013; Huang et al., 2013; Shah and Bou-Zeid, 2014), compared to unstable/mixed daytime conditions. In an attempt to reduce the nighttime errors, a change from Mellor-Yamada Janjic (Eta) to Yonsei University (YSU) PBL scheme was tested for a 4-months simulation period (February and November 2003, August and December 2010). The YSU scheme led to even larger errors in nighttime temperatures as well as to an increase in precipitation bias (Supplementary material Figs. S3-7 and S3-8). Subsequently, the original Mellor-Yamada Janjic (Eta) PBL parameterization scheme was retained in the WRF model set-up.

The frequency distributions of Tmin, shown in Fig. 8, illustrate in a different way the seasonal dependence of the model's performance in capturing daily variability and extremes. Generally, the agreement between WRF-1 and observed minimums is satisfactory in the winter of 2003. In the winter of 2010, WRF-1 is close to observations in more than half of the bins. The WRF-1 simulated distribution of minimum temperatures is skewed to the right in DFJ 2010, where the model is predicting higher Tmin, and hence underpredicting the probability of occurrence of very cold minima. Along with Tmax simulations, these results point to the fact that WRF tends to underestimate the severity of extreme temperatures in general (it produces milder temperatures), although not excessive, this bias should be taken into account when downscaling future climate scenarios.

3.2. Precipitation

The various climatic references for the study area report an average annual rainfall along the coastal zones between 700 and 1000 mm, with an increase from South to North due to the increase in the mountain heights producing the orographic precipitation as shown in the supplementary material Fig. S3-9 (Ministry of Agriculture (MOA)/United Nations Development Program (UNDP)/GTZ, 2003), which depicts a multi-year average precipitation map. The West Mountains range forms a barrier against inland moisture movement, and the precipitation it generates on the upwind western slopes can reach > 1400 mm per annum (there are no snowfall measurements in the study area, hence the term precipitation used here is indicative of rainfall only for both observed and simulated values). This dries up the air masses flowing east, and rainfall decreases rapidly on the eastern slopes of the west mountain range and registers only 600 mm inland. Rainfall inland varies between 800 mm (central inland, leeward of the lowest mountains) and 200 mm (northern inland, leeward of the highest mountains). As for the East Mountains chain along the eastern border of the study area, rainfall is about 600 mm and increases up to > 1000 mm at high elevations. Precipitation in dry years can be as low as 50% of the average. Coastal regions receive between 50 and 80 days of precipitation (Atlas Climatique du Liban, 1977; Ministry of Agriculture (MOA)/United Nations Development Program (UNDP)/GTZ, 2003, UNDP/GEF and MPWT/DGU, 2005; Lebanon SNC, 2011).

The cumulative annual precipitation from WRF 9, 3 and 1 km for 2003 and 2010 is illustrated in Fig. 9, depicting a coherent pattern corresponding to the topography and matching the sub-regional divisions remarkably well (Atlas Climatique du Liban, 1977, Ministry of Agriculture (MOA)/United Nations Development Program (UNDP)/

Table 3
MBE (°C) for seasonally Tavg in 2003, lowest errors are in bold[a].

Region	WRF-1				WRF-3				WRF-9			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
NC	**1.1**	**1.5**	**1.3**	**1.2**	1.2	1.7	1.7	1.3	1.4	1.7	1.4	1.4
CC	**0.7**	2.0	**2.5**	**0.7**	**0.7**	**1.8**	2.6	0.8	0.9	2.0	2.8	0.9
CI	**1.4**	**1.4**	**1.3**	**1.4**	1.8	1.7	1.5	1.8	2.0	2.3	2.7	2.5
Study area	**1.1**	**1.6**	**1.7**	**1.1**	1.2	1.7	1.9	1.3	1.4	2.0	2.3	1.6

[a] No stations recorded Tavg in SC, NM, CM, NI and SI regions.

Table 4
MBE (°C) for seasonally Tavg in 2010, lowest errors are in bold.

Region	WRF-1				WRF-3				WRF-9			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
NC	**1.7**	**1.4**	**2.0**	**1.6**	1.7	1.6	2.3	1.6	1.9	1.5	2.2	1.7
CC	**0.8**	**1.5**	**3.1**	**1.0**	**0.8**	1.6	**3.1**	1.1	1.2	1.8	3.4	1.4
SC	**1.3**	1.0	1.7	**1.3**	1.3	**0.8**	**1.5**	1.3	1.5	1.1	1.7	1.4
NM	**0.7**	**0.7**	**0.9**	**0.8**	1.2	0.8	1.0	1.3	1.7	1.3	1.2	1.9
CM	**1.1**	**1.0**	**1.2**	**1.2**	1.8	1.4	1.3	1.7	1.6	1.5	1.3	1.7
NI	**1.4**	**1.3**	**1.2**	**1.2**	1.4	1.4	1.2	1.3	2.4	1.6	1.4	1.7
CI	1.7	1.6	1.6	1.5	**1.7**	**1.5**	**1.5**	**1.3**	1.8	1.9	2.1	1.5
SI	**0.7**	**0.7**	1.2	**0.8**	0.7	0.8	**1.2**	0.9	1.0	1.3	1.8	1.3
Study area	**1.2**	**1.1**	**1.6**	**1.2**	1.3	1.2	1.7	1.3	1.6	1.5	1.9	1.6

GTZ, 2003). The spatial distribution in WRF-1 clearly captures the broad tendency towards decreased precipitation at the northern coast (which matches observations) in comparison to the other coastal zones (partially due to the wider coastal plains upwind of the west mountain range as depicted in Fig. 9). Additionally, WRF-1 captures the maxima along the west mountain range and the minima further inland, as well as the peak in the coastal zone and the southern inland region. WRF-3 captures most of these spatial patterns and tendencies, while WRF-9 seems to blur the precipitation contours significantly. Interestingly, the 3 km resolution was adopted in similar topographies by Trapero et al. (2013) to examine the effect of the orography on heavy precipitation events in the Eastern Pyrenees. The difference in the rainfall between the wet 2003 and the drier 2010 is significant, and is also captured by WRF.

For every station, total monthly precipitation was calculated and the process was repeated for the closest WRF grid point, then the bias was computed as a percentage difference between the simulated and observed total yearly precipitation for every station independently, and then averaged for each region, and then over the whole study area (Table 7). The study area bias and coefficient of determination between WRF and the observational data were calculated by pooling all individual stations together without regionalization. This approach was adopted because the interest is primarily in the overall water budget of the region and due to the fact that the temporal and spatial locations of precipitation events are exceedingly difficult to reproduce accurately (Li et al., 2013; Ryu et al., 2016). In individual sub-regions, particularly the Northern Mountains and Northern Inland regions, biases can reach about 50%. However, with respect to the overall study area, which is of most importance from a water resources management perspective, the biases of WRF 1 km and 3 km are very acceptable, and they perform better than WRF 9 km, especially in the dry year 2010.

Note that in an attempt to reduce the bias, a change in model configuration from 35 to 56 vertical levels was tested for a 2-months period (February 2003, August 2010). In parallel, the same period was also configured based on 26 vertical levels. An increase in the precipitation bias was noticeable as the number of levels increased with a concurrent decrease in the temperature bias (Supplementary material Figs. S3-10 and S3-11 and Tables S3-1 and S3-2). Subsequently, the original 35 vertical levels configuration was retained in the WRF model set-up since precipitation is more critical for our future applications. However, for applications where temperature is the more important metric, a higher number of vertical levels might be useful since the change in vertical resolution tends to cause a change in surface energy flux and consequently in the simulated temperature field (Aligo et al., 2009; Zeng et al., 2016).

Fig. 10 shows the RMSE for annual accumulated precipitation of WRF 9, 3 and 1 km along with the decomposition of the RMSE of WRF 1 km. Taylor diagrams portraying the results is also included in the supplementary material (Fig. S3-12). The performance of WRF 1 and 3 km was inferior to that of WRF 9 km during 2003, confirming the results obtained for the MBE for that year (Table 5). Results for 2010 yield an improvement of WRF 1 over WRF 3 and 9 km. Interestingly, the major source of error for precipitation seems to vary regionally and between years, with no strong dominance of the BIAS, VAR, or COVAR

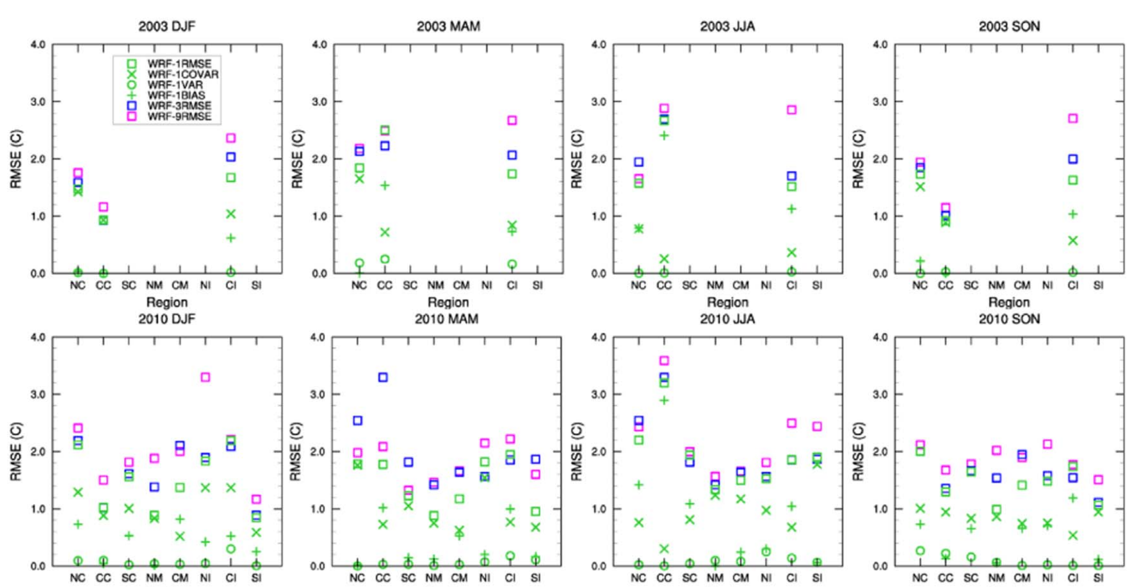

Fig. 3. Regional and seasonal RMSE for daily average 2 m temperatures for years 2003 and 2010, with decomposition of RMSE for WRF 1.

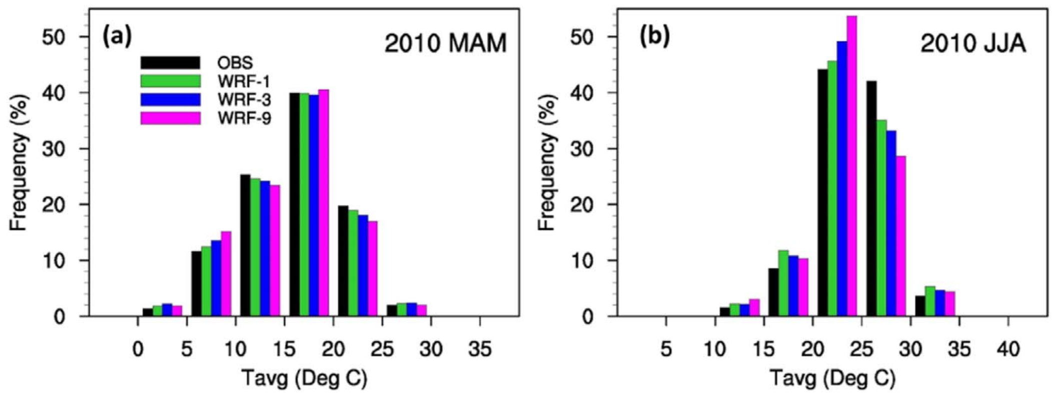

Fig. 4. PDFs of the daily average 2 m temperature (a) best WRF fit (b) worst WRF fit.

Table 5
MAE (°C) and R^2 for the summer season Tmax in 2003 & 2010, lowest errors are in bold.

Region	2003			2010		
	WRF-1	WRF-3	WRF-9	WRF-1	WRF-3	WRF-9
NC	**2.1**	2.5	2.7	**2.7**	3.1	3.4
CC	3.2	**2.9**	3.9	**4.0**	**4.0**	5.6
SC	**5.0**	5.6	5.6	**2.2**	2.4	2.7
NM	**3.4**	5.8	5.8	1.9	**1.7**	**1.7**
CM	2.8	2.9	**2.7**	**2.1**	2.4	2.5
NI	**3.1**	3.6	4.4	1.3	**1.2**	2.2
CI	**1.9**	2.9	4.4	**1.6**	2.6	4.0
SI	**3.6**	4.7	4.8	**1.8**	2.1	3.3
Study area	**3.1**	3.9	4.3	**2.2**	2.4	3.2
R^2	0.66	**0.68**	**0.68**	**0.75**	**0.75**	0.74

terms.

One important aspect in the evaluation of a model is its competence in reproducing the strength and occurrence of separate precipitation events. Precipitation in the study area happens mostly as short time rainfall events and displays a clear annual cycle with very dry summers. The high precipitation events in the study area are often connected with adverse hydrological consequences such as road and channel floods and landslides. Consequently, simulating the higher ends of the precipitations PDFs is a necessary skill for WRF. To account for the rainy season extreme events, we applied a filter so that dry days from May to September at all localities are removed, keeping only the days with precipitation strictly > 0 mm/day. Fig. 11 shows a representative selection of the resulting simulated PDFs of the mean daily precipitation simulated compared with the observations for the eight regions of the rainy season (October to May), and for both 2003 and 2010, the complete PDFs for all regions are shown in supplementary material Figs. 2–3. Given the challenges related to the simulation of precipitation the match between observed and simulated PDFs is in fact surprisingly good. During both years, WRF overestimates the occurrence of rainy days with < 0.2 mm of accumulation in all regions (Fig. 11 and Fig. S3-13 in the Supplementary material). This might be due either to light rainfall formation in WRF (e.g. as dew) or to the low resolution of rainfall gages (typically tipping buckets are used) that prevents them from detecting light rainfall events. It also overestimates the prevalence of days with rainfall accumulation between 0.2 and 1 mm along the coast (Fig. 11a and Fig. S3-13a, b & c) while it underestimates the occurrence of days with rainfall between 5 and 10 mm in the mountains (Fig. 11b and Fig. S3-13d & e) as well as along the northern and central coast (Fig. 11a and Fig. S3-13a & b). The PDF is better reproduced for the higher rainfall rates, which are more relevant. The overall distribution of rainfall is simulated particularly well in the central and southern inland (Fig. 11c and Fig. S3-13g & h), albeit with a tendency to overestimate the probability of moderate precipitation (days of rainfall between 1 and 5 mm); the overestimation is reduced by WRF-1. Likewise, WRF-1 is generally capable of simulating topographic impacts on extremes (days of rainfall beyond 5 mm), especially over mountainous regions (Fig. 11b and Fig. S3-13d & e) during the wet year 2003, with WRF-9 overestimating the probability of intense rainfall more than the finer resolutions. The change in the results with increasing resolution is in general smaller when going from WRF-1 to WRF-3, compared

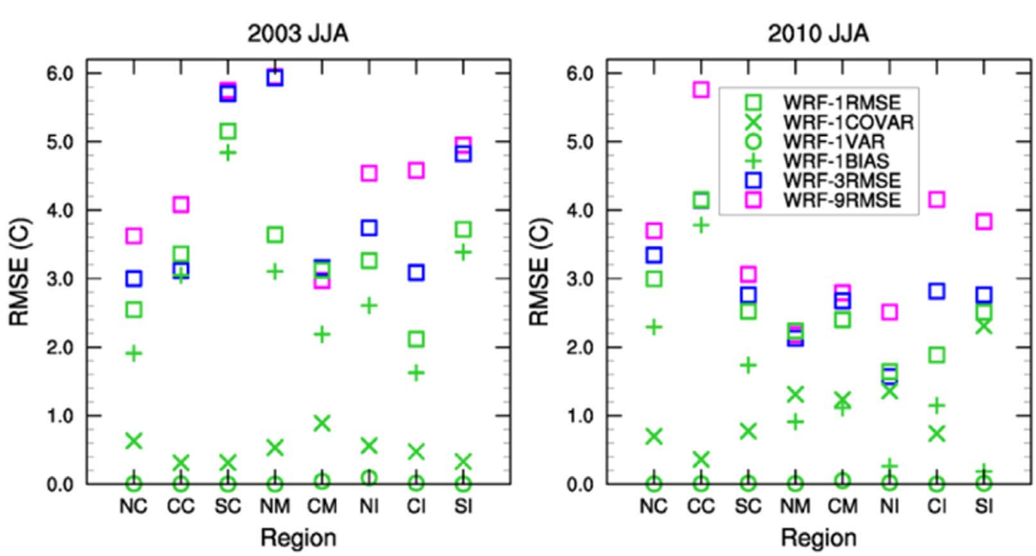

Fig. 5. Regional and seasonal RMSE for the summer season daily maximum 2 m temperatures, with decomposition of RMSE for WRF-1.

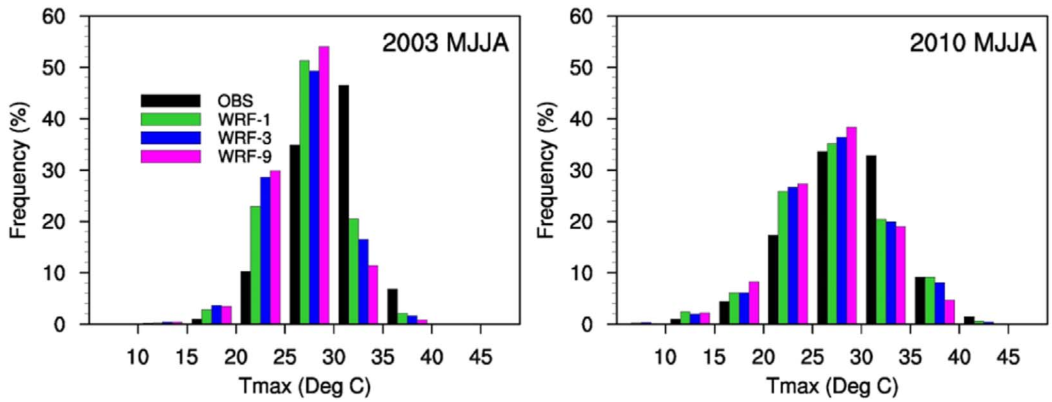

Fig. 6. Probability density plots of the daily maximum 2 m temperature.

Table 6
MAE (°C) and R^2 for the winter season Tmin in 2003 & 2010, lowest errors are in bold.

Region	2003			2010		
	WRF-1	WRF-3	WRF-9	WRF-1	WRF-3	WRF-9
NC	**1.4**	1.6	2.0	**3.0**	3.1	4.2
CC	1.5	**1.3**	1.8	**1.6**	1.7	3.0
SC	2.0	1.4	**1.0**	**2.3**	2.8	3.1
NM	**1.3**	1.6	1.6	**1.4**	2.2	2.7
CM	3.8	**3.3**	3.4	**1.1**	1.3	1.6
NI	**2.3**	2.6	2.6	**2.1**	2.3	3.7
CI	**1.7**	1.9	1.9	3.5	3.6	**3.4**
SI	**1.9**	2.4	2.8	1.3	**1.2**	1.3
Study area	**2.0**	2.0	2.1	**2.0**	2.3	2.9
R^2	**0.61**	0.60	0.59	**0.64**	0.63	0.57

to going from WRF-3 to WRF-9.

3.3. Wind

The wind data in the study area is not as comprehensive as the temperature or precipitation data. Humid maritime air is gusted from the sea during the rainy season (October to May) and produces snow at high altitudes when mixed with continental dry and cold air from the Caucasus and the Balkans entering the study area through the northeast and the southern inland region. During the summer season, very humid and warm air blows also from the littoral after passage over the sea, but the higher temperature reduces the relative humidity, which eliminates rainfall occurrence (UNDP/GEF and MPWT/DGU, 2005).

As stated in Section 2.5, two stations were used for the wind evaluation: one (TRP) located along the northern coast and the second (HAO) situated in the central inland. The predominant wind directions experienced during 2010 are displayed in Fig. 12 in a series of wind roses for the observed (at 3 hour interval) and corresponding simulated WRF-1 speeds and directions. The wind roses for 2003 follow the same direction and hence are shown in supplementary material Fig. S3-14. The number of periods with wind blowing from a particular direction is indicated by the radius of the wedge extending towards that direction, while the colors denote ranges of wind speeds. The average mean wind direction for both stations exhibited a consistent bias during both years of + 16° for HAO and − 22° for TRP, which is also apparent in the wind roses (the roses align better if bias correction is made). The persistence of the bias with similar values for both years at each station is more suggestive of errors related to the alignment of the wind vanes than model errors. Wind vanes should be oriented to the geographic North, which is often identified using a compass that indicates magnetic North. Accurate correction for this discrepancy is not always performed; in addition, potential human errors in reading the compass and aligning the vane correctly could also lead to a consistent bias. The magnetic declination angle in the project area is about 12° west, which would explain most of the inland station bias, but would exacerbate the northern coast station bias. This bias explains why at the inland station, while most of the observed wind blows from the West and South-West with a significantly lower frequency for the wind from the North-West direction, the occurrence of wind blowing from the North-West direction in WRF is more common. In addition, WRF overestimates the wind

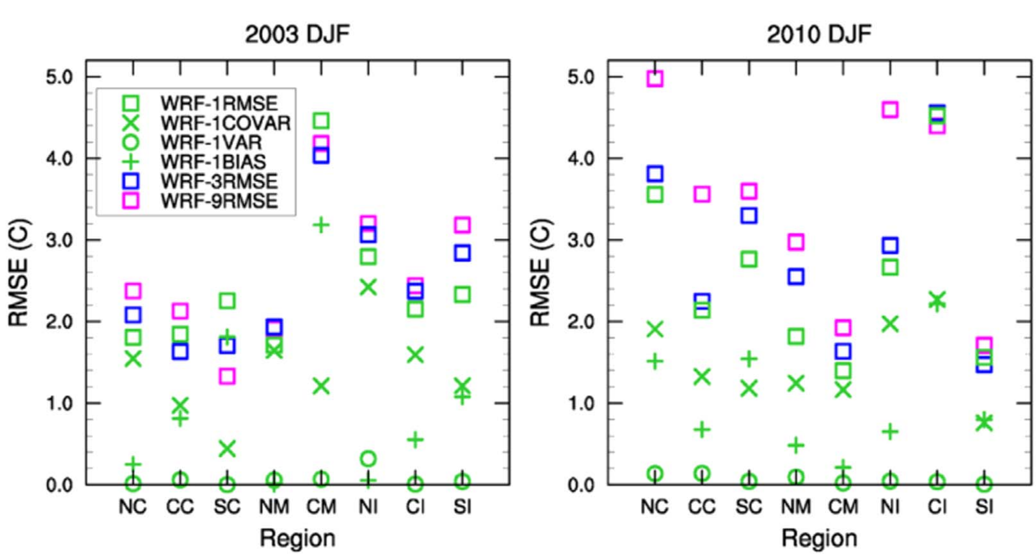

Fig. 7. Regional and seasonal RMSE for winter season daily minimum 2 m temperatures, with decomposition of RMSE for WRF-1.

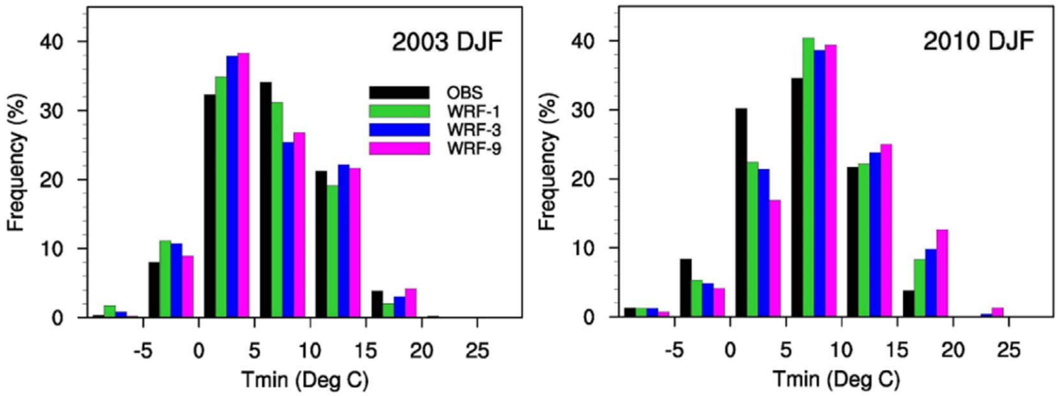

Fig. 8. Probability density plots of the winter season daily minimum 2 m temperature.

speed consistently from all directions. The same observed wind rose shows that the wind rarely blows from the South East direction at the inland station, a fact mirrored by WRF-1. At the coastal station, the dominant observed wind direction is from the West, followed by weak winds from the South-West and South-East directions, and a minor fraction of slightly higher speed winds blowing from the West-North-West direction. WRF-1 intensifies the speed of winds blowing in all directions (as in the inland station), and its bias of − 22° diminishes the occurrence of westerly winds and increases the occurrence of south-westerly winds.

The observed average daily wind speed at these stations varies from 0.5 to 6.6 m/s inland, while it increases in the simulations up to 12.6 m/s at the northern coast. The wind speeds are generally more overpredicted (about 50%) on the coast than inland (around 40%) in the three model resolutions (not shown). Concerning the statistical errors, Table 8 points out an important improvement in the MAE added by WRF 3 km, which in general performs better than the 1 km, and

Table 7
Percentage MBE and R^2 for yearly precipitation in 2003 and 2010, lowest errors are in bold.

Region	2003			2010		
	WRF-1	WRF-3	WRF 9	WRF-1	WRF-3	WRF 9
NC	− 7.2%	− 42.5%	**0.6%**	16.1%	**15.1%**	38.4%
CC	19.4%	**3.9%**	24.3%	14.1%	**6.0%**	15.4%
SC	**− 2.3%**	− 15.7%	− 9.3%	14.0%	− 3.4%	**0.2%**
NM	**37.8%**	46.2%	49.7%	**26.0%**	52.9%	47.7%
CM	10.9%	**− 3.9%**	− 11.1%	4.7%	**4.6%**	6.3%
NI	− 49.2%	− 54.5%	**− 43.1%**	− 37.4%	− 37.8%	**− 8.4%**
CI	− 18.3%	− 27.9%	**− 6.0%**	21.4%	**16.4%**	64.2%
SI	**19.7%**	37.1%	25.5%	9.3%	**5.4%**	40.0%
Study area	**− 2.01%**	− 9.75%	2.16%	6.8%	**3.8%**	17.6%
R^2	0.64	0.60	**0.66**	**0.70**	0.64	0.57

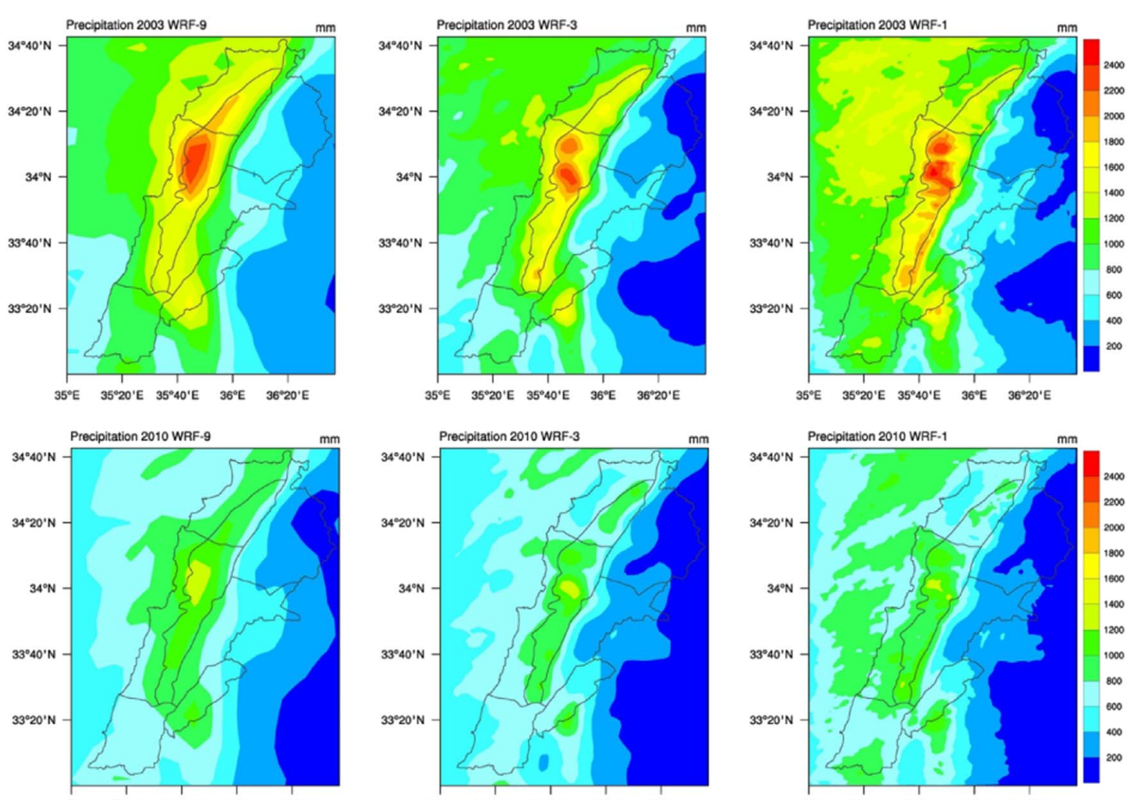

Fig. 9. Annual precipitation from WRF-9, 3 and 1 km simulation for 2003 and 2010.

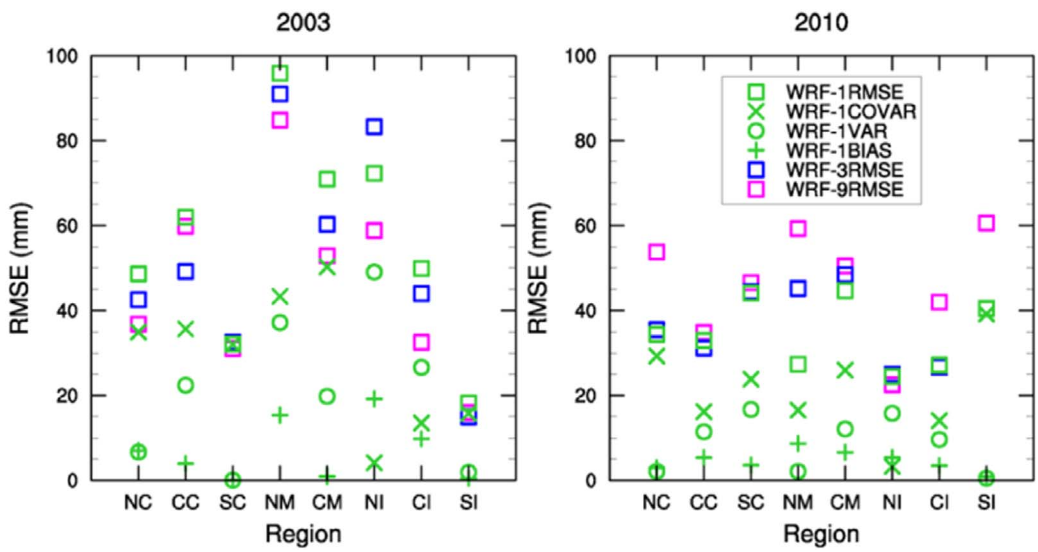

Fig. 10. Regional RMSE (mm) for rainy season precipitation, with decomposition of RMSE for WRF-1.

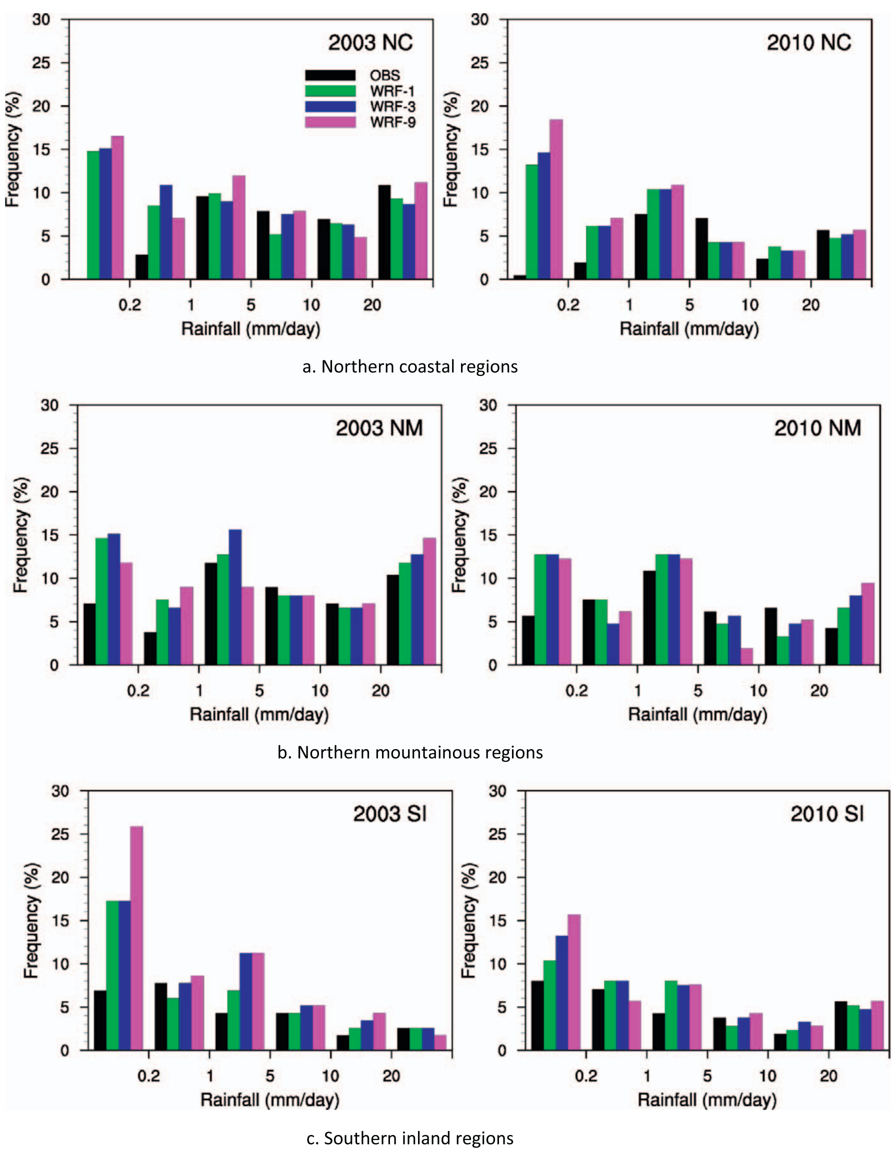

Fig. 11. Probability density of averaged daily precipitation for 2003 and 2010 in northern coast (NC), northern mountainous (NM), and southern inland (SI) regions.

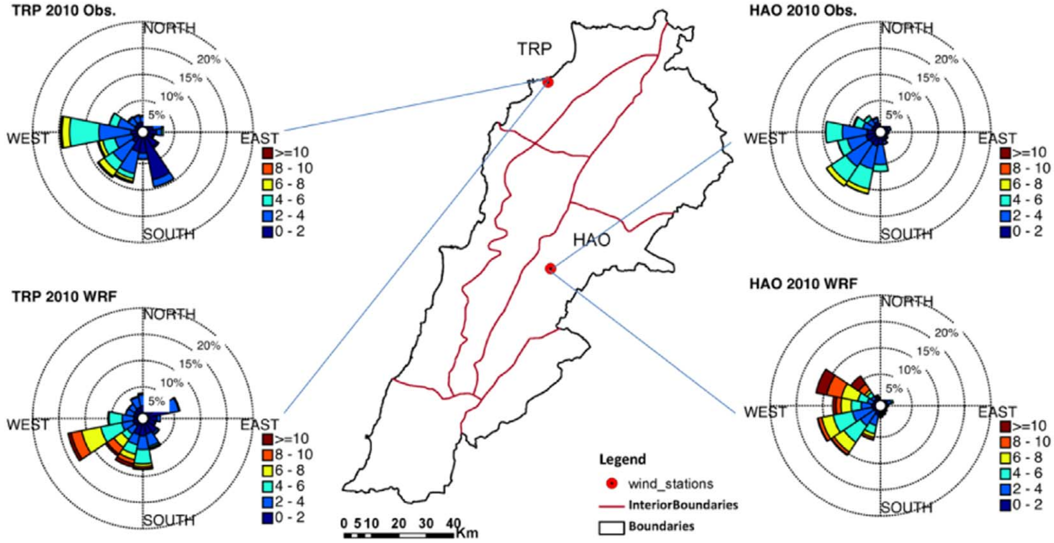

Fig. 12. Wind roses for available stations for 2010.

Table 8
MAE (m/s) and R^2 for the yearly average wind speed in 2003 & 2010, lowest errors are in bold.

Station	2003			2010		
	WRF-1	WRF-3	WRF 9	WRF-1	WRF-3	WRF 9
HAO	1.13	**0.85**	1.14	1.22	**0.94**	1.06
TRP	1.11	1.10	**1.10**	1.03	1.00	**0.94**
Study area	1.12	**0.98**	1.12	1.13	**0.97**	1.00
R^2	**0.53**	0.51	0.46	**0.46**	**0.46**	0.38

some deterioration for the WRF 9 km. However, the highest coefficient of determination achieved by WRF 1 km. Fig. 13 shows that the largest part of the RMSE for WRF 1 km was due to covariance error at both stations during both years. A Taylor diagram for the annual average wind speed at both stations is included in the supplementary material Fig. S3-15. The average statistics confirm that all resolutions have analogous results and that increasing the grid resolution to 1 km does not generally result in a significant improvement. These results clearly reflect what was expected: as the terrain complexity increases, the model struggles to simulate accurately the wind regime, especially the wind speed, which is consistent with the work of Jiménez et al. (2010) and Lorente-Plazas et al. (2016).

Regarding the wind speed extremes, a comparison from the seasonal PDFs of the two stations for 2003 in Fig. 14 illustrates that WRF underestimates the low winds (< 4 m/s) and overemphasizes the moderate winds (5–10 m/s) during all seasons in both years. The same trend is observed during 2010, as shown in supplementary material Fig. S3-16. The error for the wind speed may be aggravated by measurement errors if the cup anemometers, which have errors that are typically higher than other anemometers, do not start registering wind speeds until a minimum speed is reached (usually around 0.5 m/s, which are a common shortcoming of mechanical anemometers). If that is the case, wind speeds less than this cutoff will be registered as zero, erroneously decreasing the observed averages.

4. Summary and conclusion

This study presents an evaluation of WRF simulations over a complex topography domain along the Eastern Mediterranean for a wet year (2003) and a dry/hot year (2010) to examine its skill as a tool for climatological downscaling and its capability to capture the climatology of extreme events. The study area covers ≈ 30,000 km^2, but is characterized by strong spatial heterogeneity of the precipitation fields

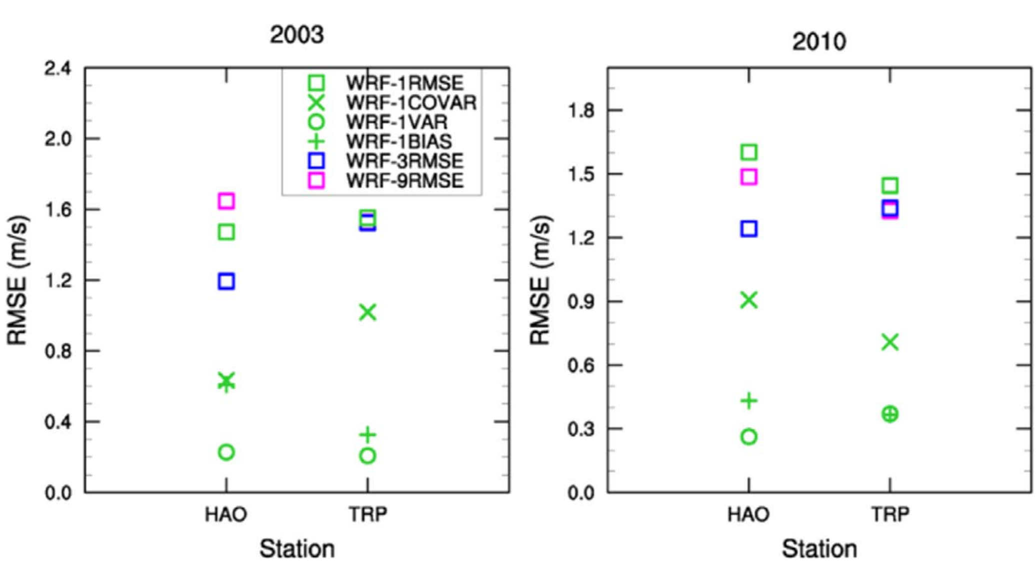

Fig. 13. Regional RMSE (m/s) for annual wind speed at the considered stations, with decomposition of RMSE for WRF-1.

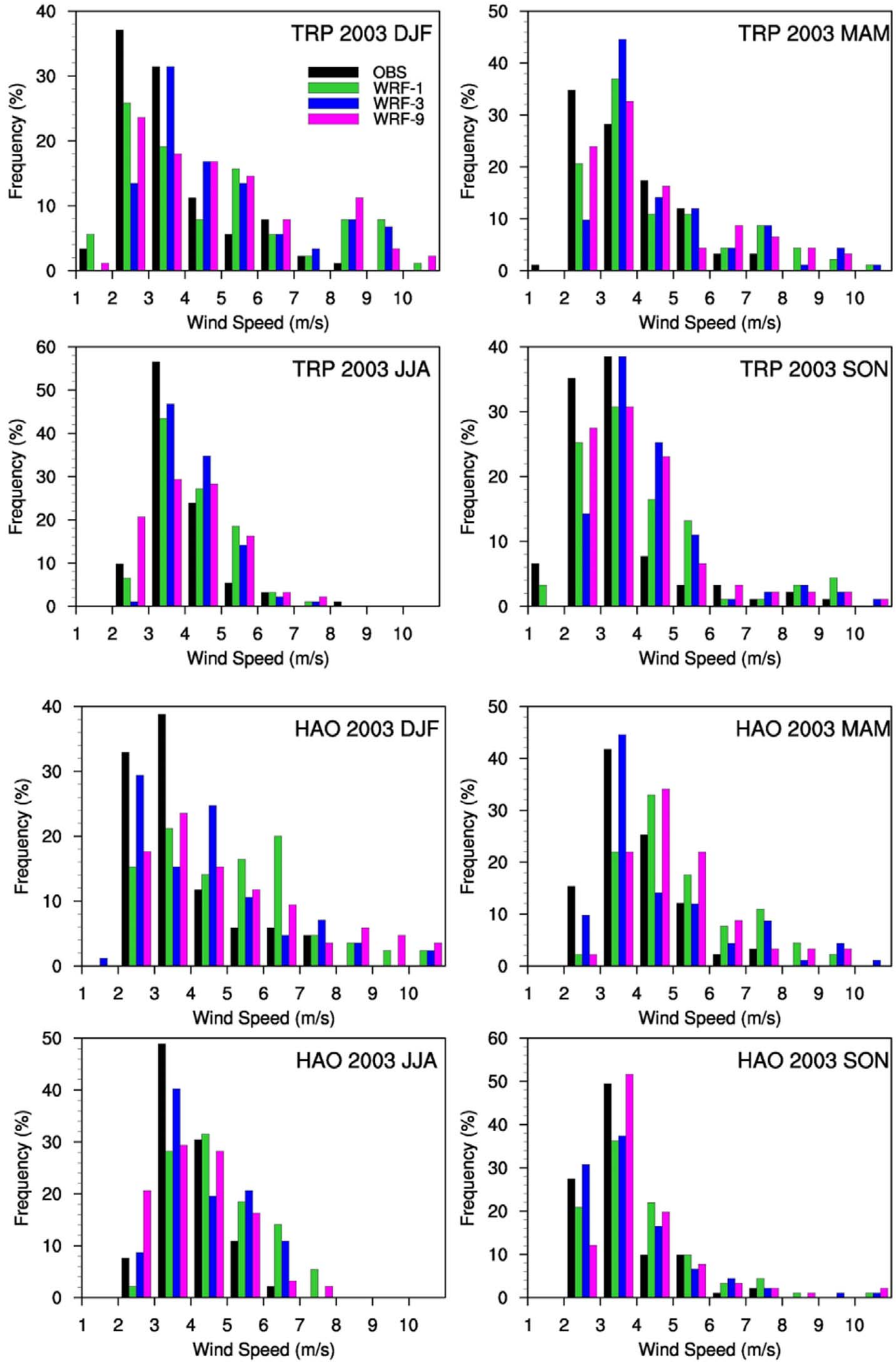

Fig. 14. Wind speed seasonal PDF of coastal (TRP) and inland stations (HAO) for 2003.

associated with its complex topography and coastal effects, and a large inter-annual variability typical of the Mediterranean climate. These attributes make it an excellent candidate for assessing the skill of WRF over complex terrain.

Overall, WRF performance was satisfactory. It captured the complex spatial and seasonal variability that characterizes the observed climate of the study area and reproduced the differences between the mild, wet year and the hot, dry year well. The reduction in errors gained from the increase in WRF resolution was noticeable for maximum temperatures, which are central to climate change impact studies. Accumulated yearly rainfall, another key parameter in impact studies, was reproduced very well by WRF. These conclusions hold for the 1 km resolution domain but also for the 3 km resolution run, the performance of which was overall quite similar to the finer domain. At a resolution of 9 km, deterioration in the simulated temperature and precipitation fields was clear, indicating that a 3 km resolution is necessary and sufficient, at least for the region considered here, for downscaling climate fields over complex terrain if year-long simulations are to be performed (1 km might be a better option for shorter periods). The wind speed and direction were more challenging for the model to reproduce at all resolutions.

Given that omitting the 1 km domain (which has to be run with a shorter time step) reduces the computational cost of the simulation by about 70%, allowing the downscaling of more future years to represent better the climatic extremes, limiting the finest resolution for downscaling to 3 km is advisable. In addition, the MYJ PBL scheme yielded better precipitation than the YSU PBL schemes in our tests, but the comparison to other schemes and over other domains would be required to develop better guidance regarding the role and influence of the land and PBL schemes, as well as other key schemes such as the one for microphysics.

Acknowledgments

The authors thank the United States Agency for International Development for providing support for this work through the USAID-NSF PEER initiative [Grant Number –AID-OAA-A_I1_00012] in conjunction with support from the US National Science Foundation under a supplement to grant [# CBET-1058027]. NCAR provided the supercomputing resources through project P36861020. We would also like to thank Dr. Dan Li for his technical support, and LARI, AREC, Dr. R. Al Khodari at LNMS, Dr. M. Traboulsi at the Lebanese University, and Dr. S. Katafago at the Litani Water Authority for their help in data acquisition.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosres.2017.11.030.

References

Abudaya, M., 2013. Seasonal and spatial variation in sea surface temperature in the South-East Mediterranean Sea. J. Environ. Earth Sci. 3 (2), 42–52. ISSN 2225-0948. http://dspace.up.edu.ps/xmlui/handle/123456789/153.

Aligo, E.A., Gallus Jr., W.A., Segal, M., 2009. On the impact of WRF model vertical grid resolution on midwest summer rainfall forecasts. Weather Forecast. 24 (2), 575–594. http://dx.doi.org/10.1175/2008WAF2007101.1.

Argüeso, D., Hidalgo-Muñoz, J.M., Gámiz-Fortis, S.R., Esteban-Parra, M.J., Dudhia, J., Castro-Díez, Y., 2011. Evaluation of WRF parameterizations for climate studies over southern Spain using a multistep regionalization. J. Clim. 24, 5633–5651. http://dx.doi.org/10.1175/JCLI-D-11-00073.1.

Argüeso, D., Hidalgo-Muñoz, J.M., Gámiz-Fortis, S.R., Esteban-Parra, M.J., Castro-Díez, Y., 2012. High-resolution projections of mean and extreme precipitation over Spain using the WRF model (2070–2099 versus 1970–1999). J. Geophys. Res. Atmos. 117 (12), D12108. http://dx.doi.org/10.1029/2011JD017399.

Atlas Climatique du Liban, 1977. Beyrouth: Republique libanaise, Ministere des travaux publics et des transports, Direction de l'aviation civile.

Barstad, I., Smith, R.B., 2005. Evaluation of an orographic precipitation model. J. Hydrometeorol. 6, 85–99. http://dx.doi.org/10.1175/JHM-404.1.

Berg, P., Wagner, S., Kunstmann, H., Schädler, G., 2013. High resolution regional climate model simulations for Germany: part I—validation. Clim. Dyn. 40, 401–414. http://dx.doi.org/10.1007/s00382-012-1508-8.

Blanchet, G., 1965. Le climat de Beyrouth. Revue de Géographie de Lyon 40, 131–158. http://dx.doi.org/10.3406/geoca.1965.1778.

Blandford, T.R., Humes, K.S., Harshburger, B.J., Moore, B.C., Walden, V.P., Ye, H., 2008. Seasonal and synoptic variations in near-surface air temperature lapse rates in a mountainous basin. J. Appl. Meteorol. Climatol. 47, 249–261. http://dx.doi.org/10.1175/2007JAMC1565.1.

Bolstad, P.V., Swift, L., Collins, F., Régnière, J., 1998. Measured and predicted air temperatures at basin to regional scales in the southern Appalachian mountains. Agric. For. Meteorol. 91, 161–176. http://dx.doi.org/10.1016/S0168-1923(98)00076-8.

Bukovsky, M.S., Karoly, D.J., 2009. Precipitation simulations using WRF as a nested regional climate model. J. Appl. Meteorol. Climatol. 48, 2152–2159. http://dx.doi.org/10.1175/2009JAMC2186.1.

Caldwell, P., Chin, H.-N.S., Bader, D.C., Bala, G., 2009. Evaluation of a WRF dynamical downscaling simulation over California. Clim. Chang. 95, 499–521. http://dx.doi.org/10.1007/s10584-009-9583-5.

Chen, F., Dudhia, J., 2001. Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: model implementation and sensitivity. Mon. Weather Rev. 129, 569–585. http://dx.doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Di Luca, A., de Elía, R., Laprise, R., 2015. Challenges in the quest for added value of regional climate dynamical downscaling. Curr. Clim. Chang. Reports 1, 10–21. http://dx.doi.org/10.1007/s40641-015-0003-9.

Di Piazza, A., Di Piazza, M., Heikkilä, U., Sandvik, A., Sorteberg, A., Giorgi, F., Mearns, L., Friedl, M.A., McIver, D.K., Zhang, X.Y., Hodges, J.C.F., Schnieder, A., Bacinni, A., Strahler, A.H., Cooper, A., Gao, F., Schaaf, Liu, W., Esteban, M.A., Chen, Y.-L., Di Luca, A., de Elía, R., Laprise, R., Chen, F., Dudhia, J., Caldwell, P., Chin, H.N.S., Bader, D.C., Bala, G., Bukovsky, M.S., Karoly, D.J., Berg, P., Wagner, S., Kunstmann, H., Schädler, G., Argüeso, D., Hidalgo-Muñoz, J.M., Gámiz-Fortis, S.R., Esteban-Parra, M.J., Dudhia, J., Castro-Díez, Y., 2009. Statistical processing of wind speed data for energy forecast and planning. Clim. Dyn. 33, 401–414. http://dx.doi.org/10.1029/2011JD017399.

Dudhia, J., 1989. Numerical study of convection observed during the winter monsoon experiment using a Mesoscale two-dimensional model. J. Atmos. Sci. http://dx.doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

El-Samra, R., Bou-Zeid, E., Bangalath, H.K., Stenchikov, G., El-Fadel, M., 2017. Future intensification of hydro-meteorological extremes: downscaling using the weather research and forecasting model. Clim. Dyn. 49 (11), 3765–3785. http://dx.doi.org/10.1007/s00382-017-3542-z.

Esteban, M.A., Chen, Y.-L., 2008. The impact of trade wind strength on precipitation over the windward side of the island of Hawaii. Mon. Weather Rev. 136, 913–928. http://dx.doi.org/10.1175/2007MWR2059.1.

Ferreira, J.A., Carvalho, A.C., Carvalheiro, L., Rocha, A., Castanheira, J.M., 2014. On the influence of physical parameterisations and domains configuration in the simulation of an extreme precipitation event. Dyn. Atmos. Oceans 68, 35–55. http://dx.doi.org/10.1016/j.dynatmoce.2014.08.001.

Friedl, M.A., McIver, D.K., Zhang, X.Y., Hodges, J.C.F., Schnieder, A., Bacinni, A., Strahler, A.H., Cooper, A., Gao, F., Schaaf, Liu, W., 2001. Global land cover classification results from MODIS. Geosci. Remote Sens. Symp. 733–735. http://dx.doi.org/10.1109/IGARSS.2001.976618. (2001. IGARSS '01. IEEE 2001 Int. 2).

Gardner, A.S., Sharp, M.J., Koerner, R.M., Labine, C., Boon, S., Marshall, S.J., Burgess, D.O., Lewis, D., 2009. Near-surface temperature lapse rates over arctic glaciers and their implications for temperature downscaling. J. Clim. 22, 4281–4298. http://dx.doi.org/10.1175/2009JCLI2845.1.

Giorgi, F., 2006. Climate change hot-spots. Geophys. Res. Lett. 33, 1–4. http://dx.doi.org/10.1029/2006GL025734.

Giorgi, F., Mearns, L., 1991. Approaches to the simulation of regional climate change: a review. Rev. Geophys. 191–216. http://dx.doi.org/10.1029/90RG02636.

Heikkilä, U., Sandvik, A., Sorteberg, A., 2011. Dynamical downscaling of ERA-40 in complex terrain using the WRF regional climate model. Clim. Dyn. 37, 1551–1564. http://dx.doi.org/10.1007/s00382-010-0928-6.

Hofstra, N., Haylock, M., New, M., Jones, P.D., 2009. Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature. J. Geophys. Res. Atmos. 114. http://dx.doi.org/10.1029/2009JD011799.

Hong, S., Lim, J., 2006. The WRF single-moment 6-class microphysics scheme (WSM6). J. Korean Meteorol. Soc. 42 (2), 129–151. http://box.mmm.ucar.edu/wrf/users/docs/WSM6-hong_and_lim_JKMS.pdf (accessed 11.12.2015).

Horvath, K., Koracin, D., Vellore, R., Jiang, J., Belu, R., 2012. Sub-kilometer dynamical downscaling of near-surface winds in complex terrain using WRF and MM5 mesoscale models. J. Geophys. Res. Atmos. 117, 1–19. http://dx.doi.org/10.1029/2012JD017432.

Huang, J., Bou-Zeid, E., 2013. Turbulence and vertical fluxes in the stable atmospheric boundary layer. Part I: a large-eddy simulation study. J. Atmos. Sci. 70, 1513–1527. http://dx.doi.org/10.1175/JAS-D-12-0167.1.

Huang, J., Bou-Zeid, E., Golaz, J.C., 2013. Turbulence and vertical fluxes in the stable atmospheric boundary-layer II: a novel mixing length model. J. Atmos. Sci. 70 (6), 1528–1542. http://dx.doi.org/10.1175/JAS-D-12-0168.1.

Ikeda, K., Rasmussen, R., Liu, C., Gochis, D., Yates, D., Chen, F., Tewari, M., Barlage, M., Dudhia, J., Miller, K., Arsenault, K., Grubišić, V., Thompson, G., Guttman, E., 2010. Simulation of seasonal snowfall over Colorado. Atmos. Res. 97 (4), 462–477. http://www.sciencedirect.com/science/article/pii/S0169809510001043.

Janjic, Z., 2002. Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP Meso model. NCEP off. Note 437, 61. http://www.lib.ncep.noaa.gov/ncepofficenotes/files/on437.pdf, Accessed date: 14 October 2014.

Jiménez, P.A., González-Rouco, J.F., García-Bustamante, E., Navarro, J., Montávez, J.P., Vilà-Guerau de Arellano, J., ... Muñoz-Roldan, A., 2010. Surface wind regionalization over complex terrain: evaluation and analysis of a high-resolution WRF simulation. J. Appl. Meteorol. Climatol. 49 (2), 268–287. http://dx.doi.org/10.1175/2009JAMC2175.1.

Katsafados, P., Papadopoulos, A., Varlas, G., Papadopoulou, E., Mavromatidis, E., 2014. Seasonal predictability of the 2010 Russian heat wave. Nat. Hazards Earth Syst. Sci. 14, 1531–1542. http://dx.doi.org/10.5194/nhess-14-1531-2014.

Kobayashi, K., Salam, M.U., 2000. Comparing simulated and measured values using mean squared desviation and its components. Agron. J. 92, 345–352. http://dx.doi.org/10.1007/s100870050043.

Lebanon SNC, 2011. Second National Communications to the UNFCCC. Republic of Lebanon, Ministry of Environment. http://unfccc.int/resource/docs/natc/lebanon_snc.pdf (accessed 13.06.2014).

Li, D., Bou-Zeid, E., Baeck, M.L., Jessup, S., Smith, J.A., 2013. Modeling land surface processes and heavy rainfall in urban environments: sensitivity to urban surface representations. J. Hydrometeorol. 14, 1098–1118. http://dx.doi.org/10.1175/JHM-D-12-0154.1.

Liu, C., Ikeda, K., Thompson, G., Rasmussen, R., Dudhia, J., 2011. High-resolution simulations of wintertime precipitation in the Colorado headwaters region: sensitivity to physics parameterizations. Mon. Weather Rev. 139 (11), 3533–3553. http://dx.doi.org/10.1175/MWR-D-11-00009.1.

Lorente-Plazas, R., Jiménez, P.A., Dudhia, J., Montávez, J.P., 2016. Evaluating and improving the impact of the atmospheric stability and orography on surface winds in the WRF model. Mon. Weather Rev. 144 (7), 2685–2693. http://dx.doi.org/10.1175/MWR-D-15-0449.1.

Lynch, P., 2008. The origins of computer weather prediction and climate modeling. J. Comput. Phys. 227, 3431–3444. http://dx.doi.org/10.1016/j.jcp.2007.02.034.

Mellor, G.L., Yamada, T., 1974. A hierarchy of turbulence closure models for planetary boundary layers. J. Atmos. Sci. http://dx.doi.org/10.1175/1520-0469(1974)031<1791:AHOTCM>2.0.CO;2.

Minder, J.R., Mote, P.W., Lundquist, J.D., 2010. Surface temperature lapse rates over complex terrain: lessons from the Cascade Mountains. J. Geophys. Res. Atmos. 115, 1–13. http://dx.doi.org/10.1029/2009JD013493.

Ministry of Agriculture (MOA)/United Nations Development Program (UNDP)/GTZ, 2003. National action plan to combat desertification. http://www.unccd.int/

ActionProgrammes/lebanon-eng2003.pdf, Accessed date: 7 January 2014.
Mlawer, E.J., Taubman, S.J., Brown, P.D., Iacono, M.J., Clough, S.A., 1997. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J. Geophys. Res. 102, 16663–16682. http://dx.doi.org/10.1029/97JD00237.
Mohan, M., Sati, A.P., 2016. WRF model performance analysis for a suite of simulation design. Atmos. Res. 169, 280–291. http://dx.doi.org/10.1016/j.atmosres.2015.10.013.
Murphy, A.H., 1988. Skill scores based on the mean square error and their relationships to the correlation coefficient. Mon. Weather Rev. http://dx.doi.org/10.1175/1520-0493(1988)116 < 2417:SSBOTM > 2.0.CO;2.
Murphy, J.M., Sexton, D.M., Barnett, D.N., Jones, G.S., Webb, M.J., Collins, M., Stainforth, D.A., 2004. Quantification of modelling uncertainties in a large ensemble of climate change simulations. Nature 430, 768–772. http://dx.doi.org/10.1038/nature02771.
National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce, 2000. NCEP FNL Operational Model Global Tropospheric Analyses, Continuing from July 1999. Research Data Archive at the National Center for Atmospheric Research. Boulder, CO., Computational and Information Systems Laboratory. https://doi.org/10.5065/D6M043C6 (accessed 23.02.2015).
National Climatic Data Center (NCDC), 2016. Billion-dollar weather/climate disasters. https://www.ncdc.noaa.gov/billions/summary-stats, Accessed date: 8 November 2016.
Paimazumder, D., Henderson, D., Mölders, N., 2012. Evaluation of WRF-forecasts over Siberia: air mass formation. Clouds Precipit. 93–110. http://dx.doi.org/10.2174/1874282301206010093.
Pan, Z., Takle, E., Gutowski, W., Turner, R., 1999. Long simulation of regional climate as a sequence of short segments. Mon. Weather Rev. 127, 308–321. http://dx.doi.org/10.1175/1520-0493(1999)127 < 0308:LSORCA > 2.0.CO;2.
Pielke Sr., R.A., Wilby, R., Niyogi, D., Hossain, F., Dairuku, K., Adegoke, J., Kallos, G., Seastedt, T., Suding, K., 2012. Dealing with complexity and extreme events using a bottom-up, resource-based vulnerability perspective. In: Sharma, A.S. (Ed.), Extreme Events and Natural Hazards: The Complexity Perspective. Geophys. Monogr. Ser. 196. AGU, Washington, D. C, pp. 345–359. http://dx.doi.org/10.1029/2011GM001086.
Quintana Seguí, P., Ribes, A., Martin, E., Habets, F., Boé, J., 2010. Comparison of three downscaling methods in simulating the impact of climate change on the hydrology of Mediterranean basins. J. Hydrol. 383, 111–124. http://dx.doi.org/10.1016/j.jhydrol.2009.09.050.
Remesan, R., Bellerby, T., Holman, I., Frostick, L., 2014. WRF model sensitivity to choice of parameterization: a study of the "York Flood 1999". Theor. Appl. Climatol. 122, 229–247. http://dx.doi.org/10.1007/s00704-014-1282-0.
Robine, J.M., Cheung, S.L.K., Le Roy, S., Van Oyen, H., Griffiths, C., Michel, J.P., Herrmann, F.R., 2008. Death toll exceeded 70,000 in Europe during the summer of 2003. C. R. Biol. 331, 171–178. http://dx.doi.org/10.1016/j.crvi.2007.12.001.
Rolland, C., 2003. Spatial and seasonal variations of air temperature lapse rates in alpine regions. J. Clim. 16, 1032–1046. http://dx.doi.org/10.1175/1520-0442(2003)016 < 1032:SASVOA > 2.0.CO;2.
Ruiz, J.J., Saulo, C., Nogués-Paegle, J., 2010. WRF model sensitivity to choice of parameterization over South America: validation against surface variables. Mon. Weather Rev. 138, 3342–3355. http://dx.doi.org/10.1175/2010MWR3358.1.
Ryu, Y.H., Smith, J.A., Bou-Zeid, E., Baeck, M.L., 2016. The influence of land-surface heterogeneities on heavy convective rainfall in the Baltimore-Washington metropolitan area. Mon. Weather Rev. 144, 553–573. http://dx.doi.org/10.1175/MWR-D-15-0192.1.
Salameh, T., Drobinski, P., Vrac, M., Naveau, P., 2009. Statistical downscaling of near-surface wind over complex terrain in southern France. Meteorog. Atmos. Phys. 103, 253–265. http://dx.doi.org/10.1007/s00703-008-0330-7.
Shah, S.K., Bou-Zeid, E., 2014. Direct numerical simulations of turbulent Ekman layers with increasing static stability: modifications to the bulk structure and second-order statistics. J. Fluid Mech. 760, 494–539. http://dx.doi.org/10.1017/jfm.2014.597.
Skamarock, W.C., Klemp, J.B., Dudhi, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.-Y., Wang, W., Powers, J.G., 2008. A description of the advanced research WRF version 3. Tech. Rep. 113. http://dx.doi.org/10.5065/D6DZ069T.
Talbot, C., Bou-Zeid, E., Smith, J., 2012. Nested mesoscale large-Eddy simulations with WRF: performance in real test cases. J. Hydrometeorol. 13, 1421–1441. http://dx.doi.org/10.1175/JHM-D-11-048.1.
Tang, Z., Fang, J., 2006. Temperature variation along the northern and southern slopes of Mt. Taibai, China. Agric. For. Meteorol. 139, 200–207. http://dx.doi.org/10.1016/j.agrformet.2006.07.001.
Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. 106, 7183–7192. http://dx.doi.org/10.1029/2000JD900719.
Tonbol, K.M., El-Geziry, T.M., 2015. The daily SST variations within the south eastern Mediterranean Sea. MEDCOAST 15, Varna, Bulgaria. 2, 1005–1016. https://www.researchgate.net/publication/282650084_The_Daily_SST_Variations_within_the_South_Eastern_Mediterranean_Sea.
Trapero, L., Bech, J., Lorente, J., 2013. Numerical modelling of heavy precipitation events over Eastern Pyrenees: analysis of orographic effects. Atmos. Res. 123, 368–383. http://dx.doi.org/10.1016/j.atmosres.2012.09.014.
UNDP/GEF and MPWT/DGU, 2005. Climatic Zoning for Building in Lebanon. Republic of Lebanon, Ministry of Public Works and Transport, General Directorate of Urban Planning. http://test.moe.gov.lb/pdf/Climate%20Change/Climatic%20Zonic%20for%20Buildings%20in%20Lebanon.pdf (accessed 14.06.2014).
Wang, Y.Q., Leung, L.R., McGregor, J.L., Lee, D.-K.K., Wang, W.-C.C., Ding, Y.H., Kimura, F., 2004. Regional climate modeling: progress, challenges, and prospects. J. Meteorol. Soc. Japan 82, 1599–1628. http://dx.doi.org/10.2151/jmsj.82.1599.
Willmott, C.J., 1982. Some comments on the evaluation of model performance. Bull. Am. Meteorol. Soc. 63, 1309–1313. http://dx.doi.org/10.1175/1520-0477(1982)063 < 1309:SCOTEO > 2.0.CO;2.
Xu, Z., Yang, Z., 2015. A new dynamical downscaling approach with GCM bias corrections and spectral nudging: a new dynamical downscaling approach. J. Geophys. Res. Atmos. 120 (8), 3063–3084. http://dx.doi.org/10.1002/2014JD022958.
Zeng, X.M., Wang, M., Zhang, Y., Wang, Y., Zheng, Y., 2016. Assessing the effects of spatial resolution on regional climate model simulated summer temperature and precipitation in China: a case study. Adv. Meteorol. 2016, 1–12. http://dx.doi.org/10.1155/2016/7639567.