Improved Atmospheric Modelling of the Oasis-Desert System in Central Asia Using WRF with Actual Satellite Products

Miao Zhang, Geping Luo, Philippe De Maeyer, Peng Cai and Alishir Kurban

Abstract: Because of the use of outdated terrestrial datasets, regional climate models (RCMs) have a limited ability to accurately simulate weather and climate conditions over heterogeneous oasis-desert systems, especially near large mountains. Using actual terrestrial datasets from satellite products for RCMs is the only possible solution to the limitation; however, it is impractical for long-period simulations due to the limited satellite products available before 2000 and the extremely time- and labor-consuming processes involved. In this study, we used the Weather Research and Forecasting (WRF) model with observed estimates of land use (LU), albedo, Leaf Area Index (LAI), and green Vegetation Fraction (VF) datasets from satellite products to examine which terrestrial datasets have a great impact on simulating water and heat conditions over heterogeneous oasis-desert systems in the northern Tianshan Mountains. Five simulations were conducted for 1–31 July in both 2010 and 2012. The decrease in the root mean squared error and increase in the coefficient of determination for the 2 m temperature (T2), humidity (RH), latent heat flux (LE), and wind speed (WS) suggest that these datasets improve the performance of WRF in both years; in particular, oasis effects are more realistically simulated. Using actual satellite-derived fractional vegetation coverage data has a much greater effect on the simulation of T2, RH, and LE than the other parameters, resulting in mean error correction values of 62%, 87%, and 92%, respectively. LU data is the primary parameter because it strongly influences other secondary land surface parameters, such as LAI and albedo. We conclude that actual LU and VF data should be used in the WRF for both weather and climate simulations.

Keywords: MODIS; Weather Research and Forecasting model; oasis-desert system; oasis effects; Northern Tianshan Mountains; Central Asia

1. Introduction

The arid region of Central Asia (CA), which includes Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, Uzbekistan, and the Xinjiang Province of China [1,2], is located deep inside the continent and has unique geomorphological characteristics, including mountain–basin systems. In this area,
Elevations can increase dramatically, from a few hundred metres above sea level in the basin areas to over 5000 m above sea level in the mountainous areas, over a horizontal distance of less than 200 km; thus, there is high heterogeneity in land cover types [3]. Water is scarce and is valuable for both human livelihoods and ecosystems in CA [4]; water resources are largely derived from the mountainous areas, whose rivers are fed by hydrologic processes of snow and glacial melt and precipitation [5]. These rivers flow into artificial lakes and then disappear into the desert areas in the basin [6]. Given the limited amount of runoff [7] and unrestricted groundwater exploitation in the area [8], oases form at the foothills of large mountains [6,9–12]. The geographical and ecological characteristics differ significantly between these oases and the surrounding deserts, causing significant differences in energy budgets, the exchange rate of momentum, and water vapor levels. These differences produce typical oasis effects [13] such as the “cold–wet” island effects of oases (an oasis is a wet, cold island capped by warm–dry air), and the thermal differences between oases and the surrounding deserts result in oasis breeze circulation (OBC). Such oasis effects increase in complexity both in and near mountain ranges. Although oases account for only a small proportion of the land surface (e.g., a proportion of 4–5% in Xinjiang, a typical region of the hinterland of the CA), more than 90% of the population and 95% of the socioeconomic wealth are concentrated there [14]. Therefore, CA, because of its large elevation differences and the importance of oases, can be divided into mountainous region, oases, and desert areas, often named the Mountain–Oasis–Desert System (MODS).

The northern Tianshan Mountains (NTM), the core section of the Silk Road, is a typical geomorphological part of CA; it is also sensitive to climate change [15]. Recent studies have indicated that annual mean air temperature in the NTM has been increasing at an average rate of 0.8 °C decade⁻¹ [16], which is greater than the average rate in CA (0.39 °C decade⁻¹ from 1979 to 2011) and the global land surface (0.27–0.31 °C decade⁻¹ from 1979 to 2005) [1]. Precipitation and the frequency of extreme precipitation show a rate of 11.3% in the NTM [16] amid a longer-term drying trend [17,18]. Other areas in CA generally show a slight decrease in average annual precipitation [4,19]. Additionally, the region has been experiencing distinct intense oases expansion since the 1950s [20–22]. Oases have expanded more than 400% in the past 60 years (from 121.0 × 10⁴ ha in 1949 to 512.5 × 10⁴ ha in 2010). A series of ecological problems have appeared as a result, including soil salinization, oasis degradation, and desertification [23–26]. Horton [27–29] found that regional climate change was largely independent or potentially related to land cover change processes. The abnormal regional temperature and precipitation changes in the NTM may be due to the rapid oases expansion. Therefore, understanding the mechanisms of oasis effects and quantitatively investigating the climate effects of oases expansion on the regional climate are important for ensuring the sustainable development and ecological stability of oases, and will also provide useful information for regional climate change assessments [30].

Numerical simulation using regional climate models (RCMs) is the most effective method to explore both oasis effects and climatic effects of the oases expansion in the complex mountain–basin systems of CA because RCMs can account for climatic mechanisms not included in field measurements [31,32] and Global Circulation Models (GCMs) [33,34]. GCMs are unable to adequately resolve many important meso-microscale processes, like wind patterns and precipitation due to orographic effects based on large-scale convective parameterization schemes, and simpler land surface processes [35]. The Weather Research and Forecasting model (WRF) is an RCM that has been widely used to simulate regional climatic patterns, particularly over the past 10 years [36,37]. Because the default terrestrial datasets in RCMs are generally derived from Advanced Very High Resolution Radiometer (AVHRR) data from 1992–1993, the ability of RCMs to accurately simulate weather and climate conditions is limited by the use of these outdated terrestrial datasets [38,39]. Integrating actual terrestrial datasets from satellite products or observation in the model simulations is a novel way to overcome these limitations. Many numerical simulations have used MODerate resolution Imaging Spectroradiometer (MODIS) products, including land use (LU), albedo, leaf area index (LAI), and green vegetation fraction (VF), to improve the boundary layer meteorology simulation and to
explore the climatic effects of land use cover change (LUCC) [14,37,40–43]. The results from the simulations using actual albedo, LAI, and VF indicated that LUCC led to local cooling of 1 °C in the summer and local warming exceeding 0.8 °C in the winter. By contrast, simulations using default terrestrial datasets showed random changes in temperature. However, these actual datasets mainly came online in the early 2000s; most numerical simulations, especially long-period simulations (many years, even hundreds of years) [40] and downscaled GCM runs, have to be performed using [44] default terrestrial datasets provided by RCMs. This choice is motivated by the fact that terrestrial datasets from satellite products are scarce and field measurements are temporally and spatially limited, especially in complex terrain. In addition, simulations using real-time, even monthly, actual various satellite terrestrial datasets in RCMs are very time- and labor-consuming processes. The question remains as to whether simulations using RCMs that update several key observed datasets can meet expected results in various applications, especially for land surface modelling or climate modelling, while reducing the time and labor cost, and also partly overcome the limitation of scarce observations and satellite products in such a complex region.

Therefore, this study aims both to quantitatively examine which actual terrestrial datasets (including LU, albedo, LAI, and VF) have a great impact on WRF performance, and to improve the simulation of weather and climate conditions over complex and heterogeneous oasis–desert systems near to large mountains. Our specific research objectives are as follows: (1) to compare the differences between the actual LU, albedo, LAI, and VF datasets and the corresponding default terrestrial datasets over MODS; (2) to quantitatively examine the impacts of using each actual terrestrial dataset on WRF performance and to determine which is key for the WRF simulations with a complex underlying surface; and (3) to comprehensively assess oasis effects including temperature, humidity, energy flux, and circulation patterns.

2. Materials and Methods

2.1. Study Area

CA is characterized by typical mountain–basin systems. Due to its unique topography, runoff generated from snow- and glacier-melt and precipitation processes [7] in mountainous areas flows into the basin and, by the time these surface waters reach oasis and desert areas, has completely evaporated in the basin. CA is influenced by the westerly circulation at the middle–high latitudes and the polar air masses [7], and experiences an arid continental climate with scarce and concentrated rainfall (less than 250 mm in the basin regions and 900 mm in the mountains). The NTM is representative of the microcosm of the terrain and climate of CA and it includes the southern part of the Tianshan Mountains and the northern part of the Gurvan Tonggut Desert (Figure 1). The oasis area is in the groundwater overflow zone and the surrounding deserts in the basin. Two nested grid systems are used in this study (Figure 1). The coarse outer domain (D01) covers the entire area of the NTM and spans a total area of 890 km × 975 km with a grid spacing of 18 km in both horizontal directions. The inner domain (D02), our main region of interest, covers a total area of 530 km × 465 km with a grid spacing of 6 km (Figure 1).
2.2. Datasets

2.2.1. Forcing Data and In Situ Measurements

The latest global atmospheric reanalysis product, ERA-Interim, provided by the European Centre for Medium-Range Weather Forecasting Reanalysis [45], was used for the initial and lateral boundary conditions for WRF simulations in this study; it was chosen because it matches well with most of the local climate records, especially in the low-lying plain areas [1]. We used geopotential, relative humidity, temperature, and U and V wind component data at 30 pressure levels and surface forcing datasets including 10 m U wind, 10 m V wind, 2 m dewpoint temperature, 2 m temperature, mean sea level pressure, sea surface temperature, sea-ice cover, skin temperature, snow density, snow depth, 4-layer soil temperature, and soil water. The dataset has a spatial resolution of \(0.75^\circ\times0.75^\circ\) and is based on 6 h intervals.

Six meteorological stations were used to validate the simulation results (Table 1). These are distributed across the oasis and the surrounding desert areas (Figure 1). Qualified T2, RH at 2 m, wind speed (WS) and wind direction (WD) at 10 m, and precipitation on an hourly scale were retrieved. In addition, we validated simulations of latent heat flux (LE) over the oases using observations from one eddy covariance system installed at station S2 (no effective observations of sensible heat flux were available because the radiation sensor was damaged). Because no surface energy observations were available for the desert areas, we only validated temperature and relative humidity over the desert areas.

Table 1. Meteorological stations’ names, locations, elevations, and available elements.

<table>
<thead>
<tr>
<th>ID</th>
<th>Longitude/°E</th>
<th>Latitude/°N</th>
<th>Altitude/m</th>
<th>LU</th>
<th>Measurements</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>86.20</td>
<td>44.32</td>
<td>473.10</td>
<td>Crop/Urban</td>
<td>T2, P, RH, WS, WD</td>
<td>2012</td>
</tr>
<tr>
<td>S2</td>
<td>85.82</td>
<td>44.28</td>
<td>469.30</td>
<td>Crop</td>
<td>T2, RH, SW, LW, LE</td>
<td>2010, 2012</td>
</tr>
<tr>
<td>S3</td>
<td>85.25</td>
<td>44.85</td>
<td>338.10</td>
<td>Crop</td>
<td>T2, P, RH, WS, WD</td>
<td>2012</td>
</tr>
<tr>
<td>S4</td>
<td>86.10</td>
<td>45.02</td>
<td>347.80</td>
<td>Crop</td>
<td>T2, P, RH, WS, WD</td>
<td>2012</td>
</tr>
<tr>
<td>S5</td>
<td>87.93</td>
<td>44.29</td>
<td>476</td>
<td>Desert</td>
<td>T2, RH</td>
<td>2010, 2012</td>
</tr>
<tr>
<td>S6</td>
<td>87.92</td>
<td>44.48</td>
<td>448</td>
<td>Desert</td>
<td>T2, RH</td>
<td>2010, 2012</td>
</tr>
</tbody>
</table>

Note: LU represents land use; T2 represents 2 m air temperature; RH represents 2 m relative humidity; LE represents latent heat flux; WS and WD represent the wind speed and direction at 10 m, respectively; SW and LW represent the downward shortwave and longwave radiation, respectively.
2.2.2. LU

The default LU data included in WRF model are originally from the U.S. Geological Survey (USGS), which classifies LU into 24 categories (Table S2) [42]. A high-resolution LU image was produced for 2012 (2012LU) using visual interpretations based on Landsat images and a 1:1,000,000 scale topographic map. This image was generated by the Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences [46]. The 2012LU has a spatial resolution of 30 m and adopts a hierarchical classification system with a spatial resolution of 30 m, including 6 categories and 25 subcategories (Table S1). We converted it into the USGS classification system according to the relationships shown in Table S2 (please see the Supplementary Materials).

2.2.3. Albedo Product (MCD43A4)

The MODIS Bidirectional Distribution Reflectance Model (BRDF) 16 Day surface albedo standard products have been validated by comparison to in situ measurements [47,48]. The high-quality primary algorithm for the MODIS standard albedo product (MCD43) has also been shown to produce consistent global quantities over a variety of land surface types and snow-covered conditions [49]. We used the nadir BRDF-adjusted reflectance MCD43A4 (MODIS Terra + Aqua Nadir BRDF-Adjusted Reflectance 16 Day L3 Global 500 m SIN Grid V005), which is computed for each MODIS spectral band (1–7) at the mean solar zenith angle. MCD43A4 images, with the strip numbers h23v04 and h24v04 for 3 July 2010 and the same day in 2012, were downloaded from the MODIS website. We reprocessed them using the same coordinate systems and resolutions via numerical simulations.

2.2.4. LAI Product (MYD15A2)

The MODIS global LAI product has been validated into stage 2 by the Committee on Earth Observation Satellites (CEOS) [50,51], and has been determined to have high continuity and consistency for all biome types. On global and regional scales, earth observation (EO)-based estimates of LAI serve as valuable inputs for climate and hydrologic modelling [52]. In this study, the level 4 MODIS global LAI MYD15A2 (MODIS/Terra + Aqua LAI/FPAR 8 Day L3 Global 1 km SIN Grid V005) was used. The data were downloaded from the website listed in Section 2.2.3 and reprocessed using the same coordinate system and resolution as in the numerical simulation.

2.2.5. VF Data from MODIS Vegetation Indices (VI) (MOD13A2)

Currently, validation to stage 3 has been achieved for MODIS VI data (MOD13), and analyses produced by various airborne and field validation campaigns demonstrate that, over most biomes, MODIS near-nadir satellite VI shows strong agreement with top-of-canopy nadir VI and land surface biophysical properties [53,54]. Using this qualified MODIS Normalized Difference Vegetation Index (NDVI), the VF can be calculated as follows [33,41,55]:

\[ VF = \frac{NDVI - NDVI_s}{NDVI_V - NDVI_s} \]  

where NDVI denotes the NDVI value for each pixel from the MODIS NDVI; NDVI_s is the NDVI value for a sparsely vegetated or barren vegetation area; and NDVI_V is the NDVI value corresponding to a full vegetation cover type. Both NDVI_V and NDVI_s are constant, allowing the pixel-level VF to reach theoretical values of 0.0 to 1.0 for any LU. Previous studies [56–59] have empirically determined NDVI_s and NDVI_V values of 0.05 and 0.87, respectively. These two parameters serve as global bounds to ensure that the derived VFs vary from 0.0 to 1.0 (i.e., VF = 1.0 when NDVI > 0.87 and VF = 0.0 when NDVI < 0.05 in Equation (1)). The MOD13A2 [60] Version 6 product (MODIS/Terra VI 16 Day L3 Global 1 km Grid SIN V006) was downloaded from the MODIS website [60].
2.3. Model Configuration and Experimental Design

WRF is an advanced mesoscale numerical weather prediction system designed for both atmospheric research and operational forecasting needs. It is jointly administered by the National Center for Atmospheric Research and the National Centers for Environmental Prediction. In this study, simulations used WRF version 3.6 coupled with the Noah land surface model.

WRF was configured for fine-scale simulation with two nested domains (D01 and D02 in Figure 1). In the vertical direction, 35 unevenly spaced full eta levels were defined, and the model top was fixed at 50 hPa. The WRF model was forced by ERA-Interim reanalysis data and was updated every 6 h. Qiu et al. [44] performed a series of analyses examining the model’s sensitivity to different parameterizations of the physical atmospheric processes operating over the study region. In this study, we used the optimal WRF configuration. Planetary boundary layer processes were resolved with the Yonsei University (YSU) scheme [61], microphysics were elucidated via WRF Single Moment-3 (WSM3) [62], cumulus clouds were simulated using the Kain–Fritsch Scheme [63], and the Community Atmospheric Model (CAM) scheme was used to calculate longwave and shortwave radiation [64].

We designed five sets of numerical experiments to investigate using the impact of each of actual LU, albedo, LAI, and VF on model results for 2010 and 2012. The experiments were as follows: the def simulation used default LU, albedo, LAI, and VF provided by the WRF itself; the LU simulation using only actual 2012LU [46]; the Alb simulation used actual LU and albedo datasets; the LAI simulation used actual LU, albedo, and LAI datasets; and the VF simulation used all of actual LU, albedo, LAI, and VF data. The model simulation was initialized from 00:00 UTC on 1 July to 18:00 UTC on 31 July in each year. During this period, the interaction of water and energy between oases and deserts are often the strongest, and oasis crops are at their growth peak. The simulation results were stored hourly with a 60 s time step for integration. Generally, in the absence of accurate, gridded initial soil moisture conditions, a spin-up period is needed to allow the soil moisture within Noah to approach equilibrium within the hydrological cycle [36]. The optimal spin-up period for any particular application is uncertain and may require years to reach equilibrium [65]. In this study, the soil moisture values of oasis and desert areas were initialized via interpolation from observed soil moisture data from similar oasis and desert regions referenced in a previous paper [42] (Table 2). In addition, following previous simulations that were similar for mesoscale water, surface energy, and circulation [14,66,67], the simulation results for the first 21 days were discarded as spin-up, and only simulations for 19:00 UTC on 22 July to 18:00 UTC on 31 July were used for the analysis. According to observation and simulations, 22–31 July were with anticyclonic and clear-sky conditions (Figure S1); thus, the effects of cloud distribution on results were excluded.

Table 2. Soil moisture values for the oasis and desert areas in the four Noah soil layers [42].

<table>
<thead>
<tr>
<th>Land Use Type</th>
<th>Noah Soil Layer</th>
<th>Soil Moisture (cm³ cm⁻³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oasis</td>
<td>0–10 cm</td>
<td>0.38 (at 5 cm)</td>
</tr>
<tr>
<td></td>
<td>10–40 cm</td>
<td>0.47 (at 25 cm)</td>
</tr>
<tr>
<td></td>
<td>40–100 cm</td>
<td>0.33 (at 70 cm)</td>
</tr>
<tr>
<td></td>
<td>100–200 cm</td>
<td>0.26 (at 150 cm)</td>
</tr>
<tr>
<td>Desert</td>
<td>0–10 cm</td>
<td>0.07 (at 5 cm)</td>
</tr>
<tr>
<td></td>
<td>10–40 cm</td>
<td>0.10 (at 25 cm)</td>
</tr>
<tr>
<td></td>
<td>40–100 cm</td>
<td>0.05 (at 70 cm)</td>
</tr>
<tr>
<td></td>
<td>100–200 cm</td>
<td>0.06 (at 150 cm)</td>
</tr>
</tbody>
</table>

3. Results

3.1. Differences between Actual Terrestrial Datasets and the Default Datasets

We first examined the differences between the actual LU, albedo, LAI, and VF data and the corresponding default datasets one by one. Both the defaults and the actual satellite images showed
generally correct land surface information for MODS. However, there were significant differences, especially in oasis and desert areas, that were strongly impacted by human activities.

Large areas of cropland and barren desert in oases and north-eastern desert areas are apparent in the actual LU data (Figure 2a), while the default data show grassland and shrubland in these oases and desert areas. The default LU data in WRF is based on AVHRR satellite data for 1992–1993 [42], which represents the original oasis and desert land cover. There has been a large expansion in urban areas and irrigated cropland in the NTM at the expense of sparse shrubland during the last 20 years [68]. In addition, the default LU data shows a large area of forest in the Ili River basin and did not indicate that ice was found on the mountaintops. These areas are misclassified, because dense grassland and cultivated lands have accounted for the largest areal proportion in this river basin over the past 40 years [69], and glaciers are common at the mountaintops in CA. The spatial consistency was only 38.42% between the actual and default LU data using a rough pixel-by-pixel comparison (Figure 2i) [70].

The default albedo, LAI, and VF are based on inter-annual averages of monthly climatology for 1986–1991 [71]. Figure 2b,f,j respectively show the spatial distributions of the actual and default albedo data and the difference between the two. The actual albedo level is overall higher than the default data, ranging from 0.05–0.25 (Figure 2j); the difference between the two datasets mainly occurs in the oasis areas and the northern desert areas (Figure 2b, black rectangle). There is also a slight difference in the values of approximately 0.10 occurring at mountaintops. It is difficult to explain

Figure 2. Comparison between actual LU (a), albedo (b), leaf area index (LAI) (c), and green vegetation fraction (VF) data (d) and the corresponding default data (e–h), with differences shown in (i–l). The red line indicates the border of the key oasis areas, and the black rectangle indicates the region strongly influenced by human activities in oases and desert areas over the past 20 years.
why albedo would be greater in the actual image than in the default dataset; we would expect albedo to decrease with an increase in crop cover. Albedo is influenced by multiple factors, including LU, VF, dynamic roughness lengths, solar elevation angle, soil color, and humidity [72]. The difference in albedo in the oasis area could be related to the severe salinization caused by irrigation [73,74] as well as the expansion of plastic-mulched areas [75] over the past 20 years. The greater albedo in the northern desert area (black rectangle) in Figure 2 can be attributed to the degradation of the desert flora following drawdown of groundwater levels in the oasis–desert transition zone [23,24], which implies that land reclamation and groundwater extraction have led to serious ecological problems in CA oases. In addition, the albedo over crops in the actual image is higher than over the surrounding northern desert area (black rectangle in Figure 2b), but this is not the case for the default data. This indirectly confirms our speculation that oasis salinization and larger areas with plastic mulching could increase the actual albedo levels. The possible reason for slight differences at the mountaintops between the actual and default albedo is that the surface reflectance estimation from different satellite images has large uncertainties over rugged terrain [42,76–80]. The spatial distributions of the actual versus default LAI and VF data, as well as the differences between them, are shown in Figure 2c,d,g,h,k,l, respectively. The differences between actual LAI and VF data and the corresponding default data range from 0 to 3 and from 0 to 85%, respectively. Major differences are evident across the basin, especially near the Ili River basin and the northern oasis border (black rectangle in Figure 2k,l), consistent with the expanded oasis region. There are few differences in the desert and mountainous areas between the actual and default LAI data. In contrast, there are noticeable differences between the actual and default VF data in the desert area, indicating that the default data do not realistically represent VF conditions in this region. Field verification shows that there are sparse desert plants (e.g., *Haloxylon* and *Tamarix ramosissima*) with a coverage of approximately 20%. The differences between the actual and default terrestrial datasets confirm that the default datasets are outdated and are less representative of land surface information.

3.2. Validation and Impacts of Actual LU, Albedo, LAI, and VF Data on Atmospheric Modelling

The validation of simulated results is one focus of this paper. We use several statistical measures, including the mean bias error (MBE), root mean squared error (RMSE) and coefficient of determination \( R^2 \), to comprehensively evaluate the simulation results [81]. These measures describe the direction of the error bias, and indicate the average error magnitude. We also assess spatial patterns of temperature, humidity, energy, and circulation and determine the difference in their daytime and night-time values by averaging daytime simulations from 19:00 to 2:00 UTC and nocturnal simulations from 08:00 to 14:00 UTC.

3.2.1. Radiation and Surface Energy Fluxes

Figure 3 shows the WRF performance in simulating LE from five simulations over cropland at S2 in 2010 and 2012. The daily average LE from five simulations correctly reproduces the overall shape of the observations (Figure 3a,c), and a strong linear relationship is obtained from all of simulations with coefficients of determination \( R^2 \) larger than 0.73 \( (p < 0.05) \) (Figure 3b,d). The difference between the five simulations and observations mainly occurs during daytime from 17:00 to 6:00 UTC. The observed daily average maximum value in LE is 244.67 W/m\(^2\) in 2010, while the peak values of LE from the def, LU, Alb, LAI, and VF simulations are 158.78 W/m\(^2\), 161.81 W/m\(^2\), 150.63 W/m\(^2\), 176.70 W/m\(^2\), and 355.06 W/m\(^2\), respectively (Figure 3a). In 2012, the daily average maximum value of LE from observation and the def, LU, Alb, LAI, and VF simulations are 371.50 W/m\(^2\), 192.50 W/m\(^2\), 191.51 W/m\(^2\), 182.74 W/m\(^2\), 211.46 W/m\(^2\), and 427.27 W/m\(^2\), respectively (Figure 3c). The VF simulation slightly overestimates LE and the other four simulations underestimate LE during the daytime in both of the two years. However, the RMSE value of the simulations decreases and \( R^2 \) increases in the following order: def, LU, Alb, LAI, and VF simulations. The \( R^2 \) increased from 0.73 to 0.74 in 2010 (Figure 3b) and from 0.92 to 0.94 in 2012 (Figure 3d). The RMSE reduced from
75.81 to 69.05 W/m² in 2010 (Figure 3b) and from 87.64 to 53.52 W/m² in 2012 (Figure 3d). These results indicate that the performance of WRF is improved in simulating the surface energy budget by the inclusion of actual LU, albedo, LAI, and VF data in the model. The VF simulation in particular shows considerable improvements in both years; the daily maximum LE value has a much closer resemblance to observations after approximately 1:00 UTC. The VF simulation may have overestimated LE because plastic mulching resulted in lower evaporation [82,83]. This process is not considered in the simulations [14,42].

Using the actual LU, albedo, LAI, and VF datasets during the night-time results in similar spatial patterns of average sensible heat (H) flux. Thus, Figures 4 and 5 only show the daytime spatial patterns of average H and LE from the def, LU, Alb, LAI, and VF simulations and the differences of H and LE resulting from using each of the actual LU, albedo, LAI, and VF datasets. The VF simulation (Figures 4 and 5c) indicates an obvious difference in spatial patterns of H and LE in the basin, compared with that from the other four simulations (Figures 4 and 5a–d). Since evident differences with values of 0–85% are across the Ili River basin and the oases areas between the actual and default VF (Figure 2f), the VF simulation considerably decreases simulation of H by a value of approximately 50–150 W/m² (Figure 4i) and considerably increases LE by approximately 90–270 W/m² (Figure 5i) over these areas.
Although the def, LU, Alb, and LAI simulations present relatively similar overall spatial patterns of H and LE, some detailed differences have to be noted. Since there is no urban representation in the default LU (Figure 2e), the value of H (LE) is obviously underestimated (overestimated) by approximately 100–150 W/m$^2$ (Figures 4 and 5f) over corresponding grids in the def simulation. Since the default LU data includes a large area of shrubland over the oasis region, and forest in the Ili River valley rather than cropland in the actual LU data, the misclassification results in the def simulation overestimating (underestimating) LE (H) by approximately 50 W/m$^2$ over the corresponding grids (Figures 4 and 5f). In addition, since there is no ice found in the default LU compared with the actual LU, the def simulation overestimates (underestimates) H (LE) in the corresponding areas. These results indicate that the outdated default LU results in an incorrect energy response, especially over the oasis area, the Ili River valley basin, and the glacier region. Realistic representation of LU is important for energy budget simulation, since it determines secondary parameters such as LAI, albedo, emissivity, and surface roughness length. Given that the actual albedo values are slightly greater than the default values, ranging from 0.05–0.25 (Figure 2j), this decreases H and LE in the Alb simulation compared with the LU simulation (Figures 4 and 5c), especially in mountainous areas. Slight differences in H and LE result from using the actual LAI data, with the most obvious differences in the Ili River basin and the northern oasis border (black rectangle in Figure 2k,l).

3.2.2. Air Temperature, Humidity at 2 m

Figures 6 and 7 show aspects of the WRF performance in simulating T2 and RH, respectively, from the five simulations. All of the def, LU, Alb, LAI, and VF simulations reproduce the shape and peak of T2 and RH, and produce a strong linear relationship of T2 and a relatively moderate relationship of RH with the observations at six stations in the two years. The $R^2$ of T2 ranges between 0.71 and 0.95 ($p < 0.05$), and that of RH ranges between 0.44 and 0.75 ($p < 0.05$) obtained from all of the five simulations. Although the T2 (RH) from all of the simulations, compared with the observations, is overestimated (underestimated) over both cropland sites (S1, S2, S3, and S4) and over desert sites (S5 and S6) throughout all times of the day, a stronger relationship (increasing $R^2$ progressively) and similar magnitudes of T2 and RH (decreasing RMSE progressively) are observed when each of actual LU, albedo, LAI, and VF datasets was used in the simulations. In particular, at S2, S3, and S4, the bias of temperature was corrected by up to 0.35–2.25 °C and that of relative humidity was corrected by up to 8.85%. Thus, using actual terrestrial datasets improves the WRF performance.

Note that the improvements are relatively smaller at station S1 compared with those at stations S2, S3, and S4. This can be attributed to the fact that the S1 station is located in an urban area, so using actual vegetation parameters such as the LAI or VF does not affect the performance of the WRF model in these areas. The overestimations of T2 and underestimations of RH over oasis areas could be attributed to the cooling or wetting effects of soil evaporation from irrigation; these cannot be simulated by lake irrigation schemes in WRF. All five simulations (the def, LU, Alb, LAI, and VF) captured the rain event that occurred on 28 July 2012 at S1, S3, and S4 (not shown). However, it is difficult to determine whether the use of actual LU, albedo, LAI, and VF data improved the simulation of precipitation due to the limited statistics.
Figure 4. Daytime spatial patterns of sensible heat flux (H) from the (a) def, (b) LU, (c) Alb, (d) LAI and (e) VF simulations, and their differences (f) b–a, (g) c–b, (h) d–c and (i) e–d (these difference pixels are statistically significant at \( p < 0.05 \)). The def simulation used default LU, albedo, LAI, and VF provided by WRF itself; the LU simulation used only actual LU data; the Alb simulation used only actual LU and albedo data; the LAI simulation used actual LU, albedo, and LAI data, and the VF simulation used all of the actual terrestrial datasets. The red line represents the border of the key oasis area.
Figure 5. Daytime spatial patterns of latent heat flux (LE) from the (a) def, (b) LU, (c) Alb, (d) LAI and (e) VF simulations, and their differences (f) b–a, (g) c–b, (h) d–c and (i) e–d (these difference pixels are statistically significant at $p < 0.05$). The def simulation used default LU, albedo, LAI, and VF provided by WRF itself; the LU simulation used only actual LU data; the Alb simulation used only actual LU and albedo data; the LAI simulation used actual LU, albedo, and LAI data, and the VF simulation used all of the actual terrestrial datasets. The red line represents the border of the key oasis area.
Figure 6. Comparisons of hourly averaged 2 m air temperature (T2) between observations and five simulations (a,c,e,g,i,k), and corresponding scatter diagram (b,d,f,h,j,l) at six stations (S1–S6) in 2010 and 2012. The five simulations are as follows: def (using default LU, albedo, LAI, and VF data); LU (using actual LU data), Alb (using actual LU and albedo data), LAI (using actual LU, albedo, and LAI), and VF (using all of actual LU, albedo, LAI, and VF data).

Figure 7. Cont.
Figure 7. Comparisons of hourly averaged 2 m relative humidity (RH) between observations and five simulations (a,c,e,g,i,k), and corresponding scatter diagram (b,d,f,h,j,l) at six stations (S1–S6) in 2010 and 2012. The five simulations are as follows: def (using default LU, albedo, LAI, and VF data); LU (using actual LU data), Alb (using actual LU and albedo data), LAI (using actual LU, albedo, and LAI), and VF (using all of actual LU, albedo, LAI, and VF data).

Figures 8 and 9 present spatial patterns of T2 and Q2 from the def, LU, Alb, LAI, and VF simulations, and their differences (the differences of pixels are statistically significant at $p < 0.05$) using each of actual LU, albedo, LAI, and VF data, respectively. Overall, all of the simulations (the def, LU, Alb, LAI, and VF) generally suggest continuous stripelike T2 increases from the mountainous areas to the basin due to the lapse rate of temperature resulting from the elevation gradient difference. In addition, the simulations show lower Q2 in the mountainous regions as compared to basin areas due to the large difference of temperature in each. Focusing on the difference between oases and desert regions, in accordance with spatial patterns of H and LE, the spatial patterns of T2 and Q2 from the def (Figures 8 and 9a,a1) and VF (Figures 8 and 9e,e1) simulations differ from the LU, Alb, and LAI simulations during the daytime (Figures 8 and 9b–d) and night-time (Figures 8 and 9b1–d1).

Specifically, as indicated by the difference of T2 and Q2 between the LU and def simulations (Figures 8 and 9f,f1), the averaged T2 differences over the oasis area and Ili River valley from the LU simulation decrease up to approximately $-0.5$ °C in the daytime (Figure 8f) and $-1.2$ °C at night-time (Figure 8f1), and the averaged Q2 over these areas increases (decreases) 0.25 g kg$^{-1}$ during the daytime (in night-time) when actual LU data is used in the WRF model (Figure 9f,f1). The differences most likely result from the fact that the default LU over these areas includes sparse shrubland and forest, while the actual LU data show cropland. Since the default LU data do not include urban areas or ice (Figure 2b), there are also large differences in T2 and Q2 over these grids. In addition, there are also obvious differences in T2 and Q2 in the northeast desert when actual LU data was used. Although the LU, Alb, and LAI simulations can simulate the cold–wet island effects of the area during the daytime, this effect is more intense in the VF simulation than in the other three simulations. Using actual VF data results in a significant decrease in T2 of $-0.5$ to $-1.5$ °C during the day (Figure 8i) and $-0.5$ to $-4.5$ °C (Figure 8i1) at night; similarly, there is a large increase in Q2 by 0.5–2.5 g kg$^{-1}$ during the day (Figure 9i) and a decrease of 1.0 g kg$^{-1}$ at night (Figure 9i1) over the oasis area and Ili River Valley. Using actual albedo data results in a slight decrease in T2 of approximately $-0.45$ °C (Figure 8g1) and a decrease in Q2 of approximately 0.5 g kg$^{-1}$ (Figure 9g1) over the north-eastern desert at night, as shown by the difference between the Alb and LU simulations. Using actual LAI data results in a slight decrease in T2 by approximately $-0.15$ °C (Figure 8h,h1), and an increase in Q2 by approximately 0.25 g kg$^{-1}$ over the oasis region for the whole day (Figure 9h,h1). Overall, actual LU and VF data strongly influence the simulations of T2 and Q2 patterns in the oasis–desert system, while albedo and LAI have a lesser impact.
Figure 8. Daytime spatial patterns of 2 m air temperature (T2) from the (a) def, (b) LU, (c) Alb, (d) LAI and (e) VF simulations, and their differences (f) b–a, (g) c–b, (h) d–c and (i) e–d (these difference pixels are statistically significant at \( p < 0.05 \)). And corresponding patterns during the night-time, which are labelled with the corresponding daytime label and the number 1. The def simulation used default LU, albedo, LAI, and VF provided by WRF itself; the LU simulation used only actual LU data; the Alb simulation used only actual LU and albedo data; the LAI simulation used actual LU, albedo, and LAI data, and the VF simulation used all of the actual terrestrial datasets. The red line represents the border of the key oasis area.

Figure 9. Cont.
Figure 9. Daytime spatial patterns of 2 m specific humidity patterns (Q2, g kg$^{-1}$) from the (a) def, (b) LU, (c) Alb, (d) LAI and (e) VF simulations, and their differences (f) b–a, (g) c–b, (h) d–c and (i) e–d (these difference pixels are statistically significant at $p < 0.05$). And corresponding patterns during the night-time, which are labelled with the corresponding daytime label and the number 1. The def simulation used default LU, albedo, LAI, and VF provided by WRF itself; the LU simulation used only actual LU data; the Alb simulation used only actual LU and albedo data; the LAI simulation used actual LU, albedo, and LAI data, and the VF simulation used all of the actual terrestrial datasets. The red line represents the border of the key oasis area.

3.2.3. Surface Circulation

Figure 10 shows comparisons of the simulated 10 m horizontal WS and WD from the def, LU, Alb, LAI, and VF simulations with observations at three meteorological stations over cropland. Station S1 is located in the upper part of the oasis, near its southern border, and stations S3 and S4 are located in the lower part of the oasis, near its northern border. Although the RMSE ($R^2$) of the WS decreases slightly (increases) as actual data are added in the def, LU, Alb, LAI, and VF simulations, most of the simulated WS values are higher than the observed values by approximately 2 m/s; 60% of the simulated WS range from 2–6 m/s, whereas the observed values range from 2–4 m/s. The reason for this bias in WS is that the uncertainty of randomized turbulence processes results in difficulties in the accurate simulation of wind patterns [84]. The trends in WD are consistent with the observations. The dominant WD is WNW or NW during the daytime and WSW or SW during the night-time for all stations; these directions are observed in all simulations and in the observations, and reflect the circulation characteristics in a mountain–valley region.
Figure 10. Comparisons of hourly 10 m (a,c,e) wind speed (WS) and (b,d,f) wind direction (WD) between observations and five simulations at three stations (S1, S3, S4).

Figure 11. Daytime spatial patterns of 10 m wind speed (WS) and wind direction (WD) from the (a) def, (b) LU, (c) Alb, (d) LAI and (e) VF simulations, and their differences (f) b—a, (g) c—b, (h) d—c and (i) e—d (these difference pixels are statistically significant at $p < 0.05$). And corresponding patterns during the night-time, which are labelled with the corresponding daytime label and the number 1. The def simulation used default LU, albedo, LAI, and VF provided by WRF itself; the LU simulation used only actual LU data; the Alb simulation used only actual LU and albedo data; the LAI simulation used actual LU, albedo, and LAI data, and the VF simulation used all of the actual terrestrial datasets. The red line represents the border of the key oasis area.
Figure 11 shows the WS and WD patterns from the def, LU, Alb, LAI, and VF simulations and the differences caused by using each of the actual LU, albedo, LAI, and VF datasets. Overall, the simulations (def, LU, Alb, LAI, and VF) reflect the characteristics of mountain–valley winds, which have WDs to the WNW or NW during the day and to the WSW or SW at night. Using actual LU, albedo, and LAI data has very little impact on the spatial patterns of WD (Figure 11f–h), but using actual VF data causes slight differences in the oasis center and the surrounding desert (Figure 11i,i1). These results suggest that using actual VF data increases the intensity of oasis effects (cold–wet island effects, and OBC).

3.2.4. Impacts of Using Actual Datasets on Atmospheric Modelling

To quantitatively discern which terrestrial datasets have the strongest influence on the meteorological elements simulated in this region, the bias percentage was calculated as the general regional influence index, following the approach in [85]. After land surface parameters are replaced, the equation is as follows:

$$BP_i = \frac{1}{n} \sum_{j=1}^{n} \left( Y_{\text{sim}(j)} - Y_{\text{sim}(j-1)} \right) \frac{1}{n} \sum_{i=1}^{n} Y_{\text{obs}}.$$  

(2)

In the current study, $n$ is the number of stations, $Y_{\text{sim}(j)}$ is the simulated meteorological variables (e.g., temperature) with updated actual land surface parameters from experiment $j$, and $Y_{\text{obs}}$ is the observed value at each station. Four modelled predictors (T2, RH, WS, and LE) are analyzed.

Figure 12 presents the bias percentage of the simulated T2, RH, WS, and LE due to using each actual dataset. The $y$ axis shows the mean bias percentage for 2010 and 2012, which represents the impact of using each actual dataset on atmospheric simulations. Using actual LAI and VF data mainly affects the LE (Figure 12g,h), while using actual LU and albedo data affects the WS (Figure 12c,f). T2 (RH) decreases (increases) by a total of $-3.5\%$ ($10.2\%$) from using actual LU, albedo, LAI, and VF data. In total, $-2.26\%$ ($8.85\%$) of the change in T2 (RH) is contributed by using actual VF data, and the remainder comes from using actual data for albedo and the other two parameters (Figure 12a,b). The WS decreases by $-13.31\%$ when actual LU, albedo, LAI, and VF data are used; of that total, LU and albedo contribute $-5.51\%$ and $-4.47\%$, respectively—far more than the other two parameters (Figure 12c). The LE first decreases due to the use of actual LU and albedo data and then increases with the addition of actual LAI and VF data; the total change is $58.19\%$, of which VF contributes $54.94\%$, which is far more than the other three parameters (Figure 12d). In general, using actual land surface parameters alters the near-surface meteorology simulation in the lower atmospheric layer (Figure 12e–h). Using actual VF data has a large influence on the simulation of T2, RH, and LE in the oasis–desert system, which contributes to error correction values of $62\%$, $87\%$, and $92\%$. Thus, using actual VF data is very important for simulating near-surface meteorology. Using actual LU is the principal parameter for near-surface water and heat simulation, since it determines the value of secondary parameters such as LAI, albedo, emissivity, and surface roughness length.
4. Discussion

Outdated default terrestrial datasets in WRF, including LU, albedo, LAI, and VF, limit this model’s ability to accurately simulate the meteorological characteristics of the complex oasis–desert system in NTM. In this study, we examined the impact of using actual LU, albedo, LAI, and VF data from satellite products in WRF on the model performance. Five simulations were conducted with the same meteorological forcing data and model schemes: def (using default LU, albedo, LAI, and VF data), LU (using actual LU data only), Alb (using actual LU and albedo data only), LAI (using actual LU, albedo, and LAI), and VF (using all of actual LU, albedo, LAI, and VF together).

WRF simulations of temperature, humidity, energy, and WS are improved by the incorporation of actual LU, albedo, LAI, and VF into the model, as evidenced by the decrease in RMSE values and increase in $R^2$ in both 2010 and 2012. Using actual VF data greatly affects the simulation of $T_2$, RH, and LE in the oasis–desert system, contributing to error correction values of 62%, 87%, and 92%, respectively. LU data is a primary parameter and it determines the values of many secondary parameters. Although all of the simulations conducted in this study produce the characteristic of the “wet–cold” island effect over the oasis area, as reported previously [7,14,71,72], WRF can more accurately reflect the intensity of the oasis cold–wet effects when using all actual LU, albedo, LAI, and VF data. The results of this study contribute to a greater knowledge of the impacts of land surface parameters on the performance of WRF [42], which is beneficial for various applications, especially for land surface and climate modelling. Our results are also critical to accurately understanding the intensity of cold–wet effects of oases and the OBC.
We note that our simulations have several limitations. For example, overall, we found that the simulations overestimated (underestimated) T2 (RH), and the VF simulation overestimates the LE relative to the observations. These errors can be attributed to the fact that soil evaporation resulting from irrigation and plastic mulching effects are not considered in the simulations of WRF [14,42]. Adding irrigation and plastic mulching schemes may help to correct these errors.

5. Conclusions

The current study used WRF with actual LU, albedo, LAI, and VF data derived from satellite products to improve the simulation of weather and climate conditions in the oasis–desert system of the NTM in 2010 and 2012. Model evaluations for temperature, humidity, and energy demonstrated that our simulations, which were performed using actual terrestrial datasets, improved the performance of WRF, as evidenced by the decrease in RMSE and the increase in $R^2$. All of the simulations exhibit the “wet–cold” island effects of the oases. However, the intensity of the wet–cold effect varies depending on the use of actual LU, albedo, LAI, and VF data. Using actual VF data results in error correction values of 62%, 87%, and 92%, respectively, for simulated T2, RH, and LE in the oasis–desert system. Using actual LU data is crucial for near-surface water and heat simulation, since it determines the values of additional secondary parameters. We conclude that it is important to use, at least, actual LU and VF data for weather and climate simulations in WRF.

Supplementary Materials: The following are available online at www.mdpi.com/2072-4292/9/12/1273/s1, Table S1: Shows LU types and codes of for the 2012LU; Table S2: Land use type and its categories for WRF and the 2012LU.

Acknowledgments: This work was supported financially by the International Partnership Program of the Chinese Academy of Science (Grant No. 131965KYSB20160004) and the National Natural Science Foundation of China (Grant No. U1303382, 41671108). The authors declare no conflicts of interest.

Author Contributions: Miao Zhang and Geping Luo conceived and designed the experiments; Miao Zhang and Peng Cai performed the experiments and analyzed the data; and Miao Zhang, Philippe De Maeyer and Alishir Kurban jointly revised the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVHRR</td>
<td>Advanced Very High Resolution Radiometer</td>
</tr>
<tr>
<td>BRDF</td>
<td>Bidirectional reflectance distribution function</td>
</tr>
<tr>
<td>CA</td>
<td>Central Asia</td>
</tr>
<tr>
<td>LAI</td>
<td>Leaf area index</td>
</tr>
<tr>
<td>LE</td>
<td>Latent heat flux</td>
</tr>
<tr>
<td>LU</td>
<td>Land use</td>
</tr>
<tr>
<td>MBE</td>
<td>Mean bias error</td>
</tr>
<tr>
<td>MM5</td>
<td>Fifth-generation Penn State/NCAR Mesoscale Model</td>
</tr>
<tr>
<td>MODIS</td>
<td>MODerate Resolution Imaging Spectroradiometer</td>
</tr>
<tr>
<td>NCAR</td>
<td>National Center for Atmospheric Research</td>
</tr>
<tr>
<td>NCEP</td>
<td>National Centers for Environmental Prediction</td>
</tr>
<tr>
<td>NTM</td>
<td>North Tianshan Mountains</td>
</tr>
<tr>
<td>OBC</td>
<td>Oasis breeze circulation</td>
</tr>
<tr>
<td>Q2</td>
<td>Specific humidity at 2 m</td>
</tr>
<tr>
<td>$R^2$</td>
<td>Coefficient of determination</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root mean squared error</td>
</tr>
<tr>
<td>RH</td>
<td>2-m relative humidity</td>
</tr>
<tr>
<td>RCMs</td>
<td>Regional climate models</td>
</tr>
<tr>
<td>T2</td>
<td>2-m air temperature</td>
</tr>
<tr>
<td>TIL</td>
<td>Temperature inversion layer</td>
</tr>
<tr>
<td>USGS</td>
<td>U.S. Geological Survey</td>
</tr>
<tr>
<td>VF</td>
<td>Vegetation fraction</td>
</tr>
<tr>
<td>WRF</td>
<td>Weather Research and Forecasting model</td>
</tr>
<tr>
<td>WS</td>
<td>Wind speed</td>
</tr>
<tr>
<td>WD</td>
<td>Wind direction</td>
</tr>
</tbody>
</table>
References


5. Immerzeel, W.W.; Van Beek, L.P.; Bierkens, M.F. Climate change will affect the Asian water towers. *Science*, 2010, 328, 1382–1385. [CrossRef] [PubMed]


23. Sun, D.; Zhao, C.; Wei, H.; Peng, D. Simulation of the relationship between land use and groundwater level in taihan river basin, Xinjiang, China. Quat. Int. 2011, 244, 254–263. [CrossRef]


30. Yan, J.W.; Liu, J.Y.; Chen, B.Z.; Feng, M.; Fang, S.F.; Xu, G.; Zhang, H.F.; Che, M.L.; Liang, W.; Hu, Y.F. Changes in the land surface energy budget in eastern China over the past three decades: Contributions of land-cover change and climate change. J. Clim. 2014, 27, 9233–9252. [CrossRef]


42. Wen, X.H.; Lu, S.H.; Jin, J.M. Integrating remote sensing data with WRF for improved simulations of oasis effects on local weather processes over an arid region in Northwestern China. J. Hydrol. 2012, 13, 573–587. [CrossRef]

43. Xu, Z.F.; Mahmood, R.; Yang, Z.L.; Fu, C.; Su, H. Investigating diurnal and seasonal climatic response to land use and land cover change over monsoon Asia with the community earth system model. Int. J. Climatol. 2015, 120, 1137–1152. [CrossRef]

45. Dee, D.P.; Uppala, S.M.; Simmons, A.J.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balsamo, G.; Bauer, P. The era-interim reanalysis: Configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 2011, 137, 553–597. [CrossRef]


70. Zhang, M.; Ma, M.; De Maeyer, P.; Kurban, A. Uncertainties in classification system conversion and an analysis of inconsistencies in global land cover products. *ISPRS Int. J. Geo Inf.* 2017, 6, 112. [CrossRef]


79. Wen, J.; Liu, Q.; Liu, Q.; Xiao, Q.; Li, X. Scale effect and scale correction of land-surface albedo in rugged terrain. *Int. J. Remote Sens.* 2009, 30, 5397–5420. [CrossRef]

80. Wen, J.; Liu, Q.; Xi, Q.; Xiao, Q.; Li, X. Parametrized brdf for atmospheric and topographic correction and albedo estimation in Jiangxi rugged terrain, China. *Int. J. Remote Sens.* 2009, 30, 2875–2896. [CrossRef]


