Performance of HIRLAM in a Semiarid Heterogeneous Region: Evaluation of the Land Surface and Boundary Layer Description Using EFEDA Observations

ANNE M. JOCHUM

ALFAclima Asesoramiento Medioambiental, Albacete, Spain, and Wageningen University, Wageningen, Netherlands

ERNESTO RODRÍGUEZ CAMINO

Instituto Nacional de Meteorología, Madrid, Spain

HENDRIK A. R. DE BRUIN AND ALBERT A. M. HOLTSLAG

Wageningen University, Wageningen, Netherlands

(Manuscript received 27 January 2003, in final form 17 February 2004)

ABSTRACT

Observations from the European Field Experiment in a Desertification-threatened Area (EFEDA) are used to evaluate the performance of the radiation, land surface, and boundary layer description of the numerical weather prediction (NWP) system High-Resolution Limited Area Model (HIRLAM) in semiarid conditions. Model analysis and 6-h forecast data of the fully coupled three-dimensional model are compared with the comprehensive dataset of a case study representing a sample of 22 days of anticyclonic conditions. Distributed micrometeorological surface stations, radiosondes, flux aircraft, and airborne lidar provide a unique validation dataset of the diurnal cycle of surface and boundary layer processes.

The model surface, soil, and boundary layer are found to be too moist and slightly too cold during most of the diurnal cycle. The model radiation and surface energy budgets are biased toward more humid conditions.

Model shortcomings are identified essentially in four areas. These are the moisture data assimilation, the land-use and soil classification with its associated physiographic database, the aerosol parameterization in the radiation code, and the boundary layer vertical resolution and entrainment description.

Practical steps for immediate improvement of the model performance are proposed. They focus on the use of a land-use and soil classification and physiographic database adapted to Mediterranean landscapes, in combination with the inclusion of aerosol parameters in the radiation scheme, that account for the typically higher aerosol load of arid and semiarid environments.

1. Introduction

The parameterization of land surface and atmospheric boundary layer (ABL) processes has made considerable progress during the last decade. Advanced surface schemes have recently been incorporated in the major numerical weather prediction (NWP) models. A series of comparison and validation studies (Betts et al. 1993, 1998, 2001; Holtslag et al. 1995; Holtslag and Ek 1996; Yucel et al. 1998; Bringfelt et al. 1999; and others) have contributed significantly to this progress. These previous comparisons have all one feature in common in that the observational sites are located in fairly homogeneous areas, with only small variations in topography and gross vegetation type. All but the Arizona site of Yucel et al. (1998) (who do not provide a detailed study of boundary layer development) and the Sahelian site of van den Hurk et al. (2000) (who perform an offline validation of the land surface scheme) are furthermore located in semihumid or humid climate zones. Yet they all identify problems related to the transport of moisture. An interesting question is, therefore, how the models behave in arid or semiarid climate zones.

It is the objective of this work to evaluate the surface and ABL performance of the NWP High-Resolution Limited Area Model (HIRLAM, version 5.2; Unden et al. 2002) for a semiarid environment with heterogeneous land use. A consolidated dataset from the European Field Experiment in a Desertification-threatened Area (EFEDA; Bolle et al. 1993) is used for this purpose. The EFEDA area is located at an elevation of 700 m and is surrounded by mountains at distances of several tens of kilometers. The surface evapotranspiration is inhomogeneous on scales of several tens of kilometers because of partly irrigated land use. EFEDA is one of

Corresponding author address: Dr. Anne M. Jochum, ALFAclima Asesoramiento Medioambiental, Avenida de España, 9, E-02002 Albacete, Spain.
E-mail: ajochum@terra.es

a series of land surface experiments that were conducted in a wide range of ecosystem and climate zones (Jochum et al. 2000). These field experiments represent a concentrated effort to document a general circulation model (GCM) grid-scale volume and to study the aggregation of surface exchanges from subgrid scale to landscape units.

Each land surface experiment involves detailed ground-based measurements (micrometeorology, soils, vegetation) at selected sites, research aircraft with in situ and remote sensing instrumentation, aerological soundings, and a variety of satellite observations, and thus provides a complete dataset for model evaluation. The First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment (FIFE; Sellers et al. 1992) was the first land surface experiment to be used for NWP model evaluation. For that purpose, a single mean time series of surface and near-surface parameters was generated by Betts and Ball (1998). Betts et al. (1993) compared this time series with a 24-h forecast time series obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) operational model at the grid point next to the FIFE area. They found a number of systematic errors associated with overpredicted incoming solar radiation, too slow ground-surface interaction, surface and boundary layer drying out due to deficiencies in the formulations for soil moisture and surface evaporation, the roughness length for heat, and the description of entrainment at the boundary layer top.

Based on these results and further studies, improved land surface parameterizations were developed (Viterbo and Beljaars 1995) and introduced in the ECMWF operational model and used in the ECMWF 15-yr ReAnalysis Project (ERA-15). Betts et al. (1998) evaluated the performance of ERA-15 by comparing the corresponding model output time series with the same FIFE dataset. They found that some of the earlier model biases were removed, but still identified model errors and areas for further model improvement, such as the longwave radiation code and the incorporation of a seasonal cycle of vegetation and ground storage. An area of concern remains also the diurnal cycle of boundary layer development and the transport of moisture.

In a similar way and using the same FIFE dataset, Betts et al. (1997) evaluated the performances of the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis model and the NCEP Eta Model and contributed to improving the land surface parameterization in those models. Yucel et al. (1998) compared data from the NCEP Eta Model with longer time series of surface data from two observational networks, located in a semihumid, continental climate (Oklahoma and Kansas) and in a semiarid climate (Arizona). They again find systematic biases in the model solar radiation and in parts of the surface formulation.

Betts et al. (2001) used data from the old black spruce site of the Boreal Ecosystem–Atmosphere Study (BOREAS; Sellers et al. 1997) in Manitoba, Canada, for comparison with the ECMWF and NCEP–NCAR reanalysis models. Here they find again shortcomings of the surface and boundary layer parameterizations related to radiation, soil processes, evaporation, boundary layer development, and moisture. The work led to improved land surface parameterization for high-latitude forests.

Bringfelt et al. (1999) evaluated the tile approach (Avissar and Pielke 1989; Koster and Suarez 1992) as implemented in an experimental high-resolution version of HIRLAM. Their reference observations are from three sites of different land use (agricultural and forest) in the Northern Hemisphere Climate Processes Land-Surface Experiment (NOPEX; Halldin et al. 1999). They find that model results depend critically on a realistic description of the vegetation physiology and the initial soil moisture and that the tile approach works well for NOPEX.

We evaluate here the performance of the operational HIRLAM version 5.2 radiation, land surface, and boundary layer description for a semiarid environment. The output of the fully coupled 3D model is compared with a comprehensive observational dataset from EFEDA (Bolle et al. 1993). Distributed micrometeorological surface stations, radiosondes, flux aircraft, and airborne lidar provide a unique validation dataset of the diurnal cycle of surface and boundary layer processes. The purpose of this comparison is twofold. First, the results will provide indications for further model improvement. This corresponds to the classical model validation. Second, the information on model output errors enables us to assess uncertainties arising from the use of four-dimensional data assimilation (4DDA) and forecast data in practical applications, such as impact studies and the derivation of biogeophysical parameters from satellite data. The observational dataset, the model versions, and the evaluation strategy are described in section 2. The comparison starts with the atmospheric boundary layer structure and profiles (section 3). The surface thermodynamic cycle is investigated in section 4, and the surface radiation and energy budgets are compared in section 5. Section 6 summarizes the conclusions.

2. Validation layout

The HIRLAM system is a complete NWP system including 4DDA with analysis of observations and a limited-area short-range forecasting model (Unden et al. 2002). It is currently used operationally by seven European Weather Services (Denmark, Finland, Iceland, Ireland, the Netherlands, Norway, Spain, and Sweden). HIRLAM uses the Interaction Soil–Biosphere–Atmosphere (ISBA) surface scheme (Noilhan and Planton 1989; Mahfouf and Noilhan 1996; Rodríguez et al. 2003) and the tile approach to represent fractions of land use classes present within a single grid element (Avissar and Pielke 1989; Koster and Suarez 1992). The

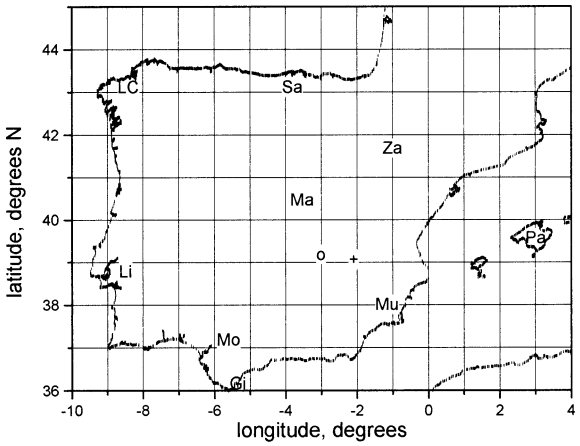

FIG. 1. Model grids and locations of radiosonde stations for the Iberian Peninsula. The solid grid lines mark a regular 1° by 1° grid, which holds four HIRLAM 0.5° grid elements or 16 HIRLAM 0.25° grid elements, respectively. Research radiosondes were launched during EFEDA from the center of the Tomelloso site (o) and from the center of the Barrax site (+). The letter codes indicate the operational radiosonde stations.

boundary layer scheme is based on a prognostic turbulent kinetic energy model together with a diagnostic length scale according to Cuxart et al. (2000). The turbulent surface fluxes are calculated from the standard bulk formulation using differences between the surface and lowest model layer following Louis (1979) and Louis et al. (1982). The fast radiation scheme is derived from Savijärvi (1990).

The operational version 5.2 of HIRLAM is used for the validation. Its horizontal grid resolution is 0.5°. The analysis of upper-air variables is based on an optimum interpolation method (Unden et al. 2002). A research version of HIRLAM [as currently used by the Spanish Weather Service Instituto Nacional de Meteorologia (INM)] serves to analyze details at higher spatial resolution (0.25°). Unless explicitly mentioned, the results shown in this paper refer to the 0.5° version. The comparison focuses on the 6-h forecast (fc06). Differences in prognostic variables introduced by the analysis increment are discussed whenever they are significant.

The observational datasets were collected during the EFEDA field phase in June 1991 (Bolle et al. 1993). A large sample of 22 case studies of anticyclonic conditions typically encountered in the area during the summer months is available. The 23 June 1991 is most densely covered by observations (Michels and Jochum 1995). The sample composite is almost identical to the individual 23 June observations (Jochum et al. 2004, manuscript submitted to *J. Appl. Meteor.*). Therefore, the individual case study approach is adopted in this work. The diurnal cycles of surface and ABL processes are compared in detail with observations of 23 June 1991. The performance evaluation at seasonal and annual time scales is the subject of a follow-up study that uses data from two permanent observing stations in the same area.

TABLE 1. Model grid characteristics. Grid size (km) gives average values for the Iberian Peninsula.

	Resolution latitude	Resolution longitude	E–W grid size (km)	N–S grid size (km)
HIRLAM-0.25	0.25°	0.25°	24	31
HIRLAM-0.5	0.5°	0.5°	48	62

The EFEDA experimental area is located in the region called La Mancha, part of the Castilian high plateau in the southeast of Spain, at an average elevation of 700 m above mean sea level. It extends about 60 km in north–south direction and 80 km in east–west direction, respectively. The area is generally flat, with elevation variations of up to 100 m, but surrounded and influenced by mountain ranges. The distance to the mountains is 35 km from the south, 45 km toward the northeast, and 40 km toward the northwest. Three intensive observation sites (supersites) were selected as to represent the major Mediterranean ecosystems found in the area.

Figure 1 shows the location of the two main supersites, Tomelloso (o) and Barrax (+), along with the locations of operational radiosonde stations. The symbols mark the center points of a 10 km by 10 km square of each supersite. These center points coincide with the high-resolution radiosonde launch sites in either case. A regular 1° by 1° grid (solid lines) serves to give a rough orientation about the HIRLAM grids. The two different resolutions of HIRLAM used here, 0.5° and 0.25°, would subdivide each 1° by 1° grid element into 4 and 16 smaller elements, respectively. Table 1 summarizes the model grid dimensions.

This comparison focuses on the Tomelloso site, which is representative of semiarid conditions and thus would be expected to be well reproduced by the model. The Barrax site is characterized by two additional sources of moisture, one at the ground from irrigation and one aloft from the Mediterranean sea breeze penetrating inland in the afternoon. The Barrax comparison is used occasionally, whenever it provides additional insight. Both represent typical Mediterranean land use. For each site, the comparison is performed in column mode, that is, aggregated site observations against model output at the corresponding grid points. The site-aggregated surface dataset for Tomelloso is taken from Linder et al. (1996). The site-aggregated dataset for Barrax was developed on the basis of Linder et al. (1996) and the vegetation classification of Calera Belmonte (2000). The procedure and results from aggregating surface fluxes to various NWP model grid elements are described by Jochum et al. (2004, manuscript submitted to *J. Appl. Meteor.*). The radiosonde and aircraft datasets are described in Michels and Jochum (1995) and Jochum (1993).

Because of the presence of mountainous terrain at

TABLE 2. Location and characteristics of HIRLAM grid points within the EFEDA area and fixed model parameters for these cells (and tiles when applicable). Values of albedo, leaf area index (LAI), and roughness length are given for Jun. Data for the EFEDA supersites from Linder et al. (1996). Model land-use classes after Manzi and Planton (1994). Supersite land-use classes after Martínez and Calera (2001).

	Tomelloso	HIRLAM T	HIRLAM TB	Barrax
Lon	3°01′W	3°W	2°30′W	2°06′W
Lat	39°10′N	39°N	39°N	39°03′N
Elev (m)	655	755	891	700
Vegetation percentage	0.05–0.15 (1–30 Jun)	0.5	0.4/0.9	0.07 (23 Jun)
Land-use class	Summer dry (vine)	17 (100%) deciduous shrub	1 (87%) crop, 5 (13%) deciduous broadleaf	30% irrigated, 70% bare or summer dry
Soil texture type	3	5	11/5	4
Emissivity	0.98	0.95	0.95	0.98
Albedo	0.28 (23 Jun)	0.17	0.18/0.16	0.15 irrigated, 0.23 bare/dry
LAI	0.1–0.4 (1–30 Jun)	1.5	2.0/4.0	0.26–0.36 (11–29 Jun)
r_{smin} (s m^{-1})	~ 140	150	40/250	(irrigated) 180
Roughness length (m)	0.01–0.05 (1–30 Jun)	0.1351*	0.0908/0.8488*	0.1 bare/dry 0.001–0.06 irrigated
Roughness length for heat (mm)	~ z_0/200	0.096	0.042/0.8	0.001 bare

* Includes orography.

model scales, the selection of model grid elements for comparison needs great care. The principal and obvious idea is to have the observational site inside the model grid cell, which in practice matches the nearest-grid-point concept of Betts et al. (1993, 1998). Both FIFE and BOREAS, however, are located within extended areas of similar characteristics and of small topographic variations. In contrast, the EFEDA area itself is rather flat, but is surrounded and influenced by mountain ridges, which come closest to the area from the south (35 km) and east. In a pragmatic way, we have still identified the nearest grid point to the location of each observational set of profiles (see Table 2 for all relevant coordinates), while monitoring the differences to a set of adjacent grid points, in order to assess the influence of the model orography. All points included in the comparison are located within the same climate regime. The model grid elements, which include the Tomelloso and Barrax sites, also contain part of the southern mountains. The model grid elements covering the northern part of the EFEDA area include more mountain ranges toward the north and east. The model uses a mean orography, which is consistent with the current mountain drag formulation. This results in a notable difference in elevation between grid element and site, depending on the model type and resolution. For Tomelloso, for example, the HIRLAM 0.5° grid surface is at 100 m above ground level.

3. Boundary layer structure and profiles

The daytime diurnal cycle of 23 June 1991 was densely covered with aircraft observations. Radiosondes were launched roughly every 2 h at Tomelloso and Barrax. The day was almost cloud free with weak winds. Michels and Jochum (1995) performed a detailed analysis of the diurnal cycle of the ABL structure at both sites, as obtained from all available observations (radiosoundings, aircraft ascents, airborne lidar).

At Tomelloso (Fig. 2), the stable nocturnal boundary layer (NBL) is about 500 m deep and remains until very late in the morning (1100 UTC). The mixed layer (ML) grows rather slowly until then. Its rapid growth phase starts only after complete erosion of the NBL and is furthermore slowed down by the slightly stable stratification of the very deep (3 km) residual layer (RL). The ML reaches its full depth by late afternoon (1730 UTC) only, when finally the RL has been eroded as well. The maintenance of the RL for most of the day is rather unusual, but was observed quite often during EFEDA (Michels and Jochum 1995). A possible reason is the occurrence of cirrus cloud reducing the incoming solar radiation. Also, the slightly stable stratification of the RL is relevant, since more energy is needed to remove a stable RL than a neutral one. The RL would be slightly stable for the combined effect of two processes. On one hand, entrainment of warm air at the top of the ML will continue for a short while after the ML has decoupled from the surface in the evening. On the other hand, nighttime cooling due to aerosol and water vapor occurs in the NBL and lower RL.

At Barrax (not shown; see Michels and Jochum 1995), the NBL is twice as deep and takes until noon to be eroded. The ML is growing more slowly and reaches a lower final depth. Both features are due to the larger amounts of water vapor available from the irrigated fields. Moisture advected by the sea breeze penetrating inland in the afternoon inhibits further ML growth here. The RL is similar to the one at Tomelloso.

The comparison of the vertical profiles of potential temperature at Tomelloso (Fig. 2) shows small differences between model forecast and observations. The profiles from analysis (not shown) and forecast are very similar. Given the individual sonde measurement errors,

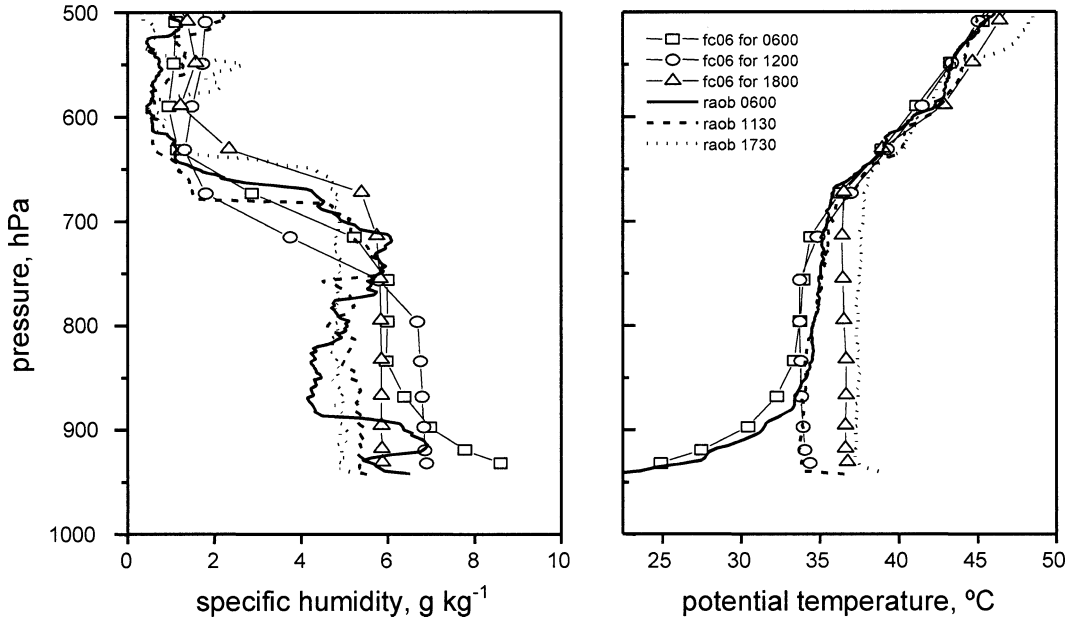

FIG. 2. Comparison of vertical profiles of (right) potential temperature and (left) specific humidity from fc06 (symbols) and Tomelloso radiosonde (no symbols) on 23 Jun 1991.

differences in temperature and specific humidity exceeding 0.8 K and 1 g kg^{-1}, respectively, are considered to be significant. At 0600 UTC, the model NBL temperature is about right, whereas the model RL is slightly (1 K) too cold. At 1200 noon, the ML temperature is reproduced very well. However, the model does not capture the weakly stable RL and, therefore, produces an ML top at about 740 hPa, which is almost 100 hPa too high. By 1800 UTC, the model ML top is about right, but the ML is 1 K too cold. There are no observations available at midnight. These differences in potential temperature are just above the limit of significance. Taking into account that each comparison involves a single radiosonde profile (which gives a ''snapshot'' only) and that the results for other fair-weather days (not shown here) are similar, it is concluded that the model performs well on daytime potential temperature.

In the 0600–1200 UTC period, the model-diagnosed warming (Fig. 3a) of the lower ABL is too high, especially at the lowest grid point (by 0.3 K h^{-1}). This period is characterized by erosion of the NBL and incipient growth of the ML, so a combination of factors could contribute to the overprediction. Part of the rapid ML growth phase occurs in the afternoon period 1200–1800 UTC. Therefore, the overall observed ML warming during that period is stronger in the lower layers. This feature is not captured by the model, which gives rise to a slight underprediction (0.1 K h^{-1}) of lower ML warming. The RL warming/cooling is reproduced fairly well in both periods.

For moisture, the situation is different. At 0600 UTC, the radiosonde shows a complex layered structure (Fig. 2), with a deep relatively dry (4 g kg^{-1}) RL and a moist NBL (6–7 g kg^{-1}) on top of the shallow drier incipient ML. The model forecast (at all times again similar to analysis) comes close to the NBL moisture (0.5 g kg^{-1} higher), but is unable to resolve the rest of the layers. This means also that the incipient ML is too moist. At 1200 UTC, the model gives an ML top between 800 and 760 hPa, with humidity slightly decreasing upward. The radiosonde-derived ML extends only up to 850 hPa, while the rest is RL. The model ML average of specific humidity (7 g kg^{-1}) is about 1.5 g kg^{-1} (27%) too moist. At 1800 UTC, the model ML is again about 1.5 g kg^{-1} too moist. The ML top around 660 hPa agrees reasonably well with observations.

The observed drying or moistening varies vertically according to the layered structure (Fig. 3b). During the morning period (0600–1200 UTC) the erosion of the NBL leads to drying of the lowest layers. The lower RL is moistening by entrainment, while the upper RL is strongly drying by entrainment of very dry free atmosphere (FA) air. The model is able to reproduce the general shape of the vertical profile in this case, even though the RL top is too low. During the afternoon period (1200–1800 UTC), the model ML dries out more than observed, which helps to partially offset the low bias in the 1200 UTC humidity profile (such that the 1800 UTC profile is less biased). The seemingly good agreement of the upper-RL moistening is due to another compensating effect rather than to the model's ability to reproduce the corresponding physical processes. This moistening peak results from the final ML growth going on in the early afternoon, when the remnants of moist RL air are entrained down into the ML. As explained above, the RL erosion takes a large part of the afternoon.

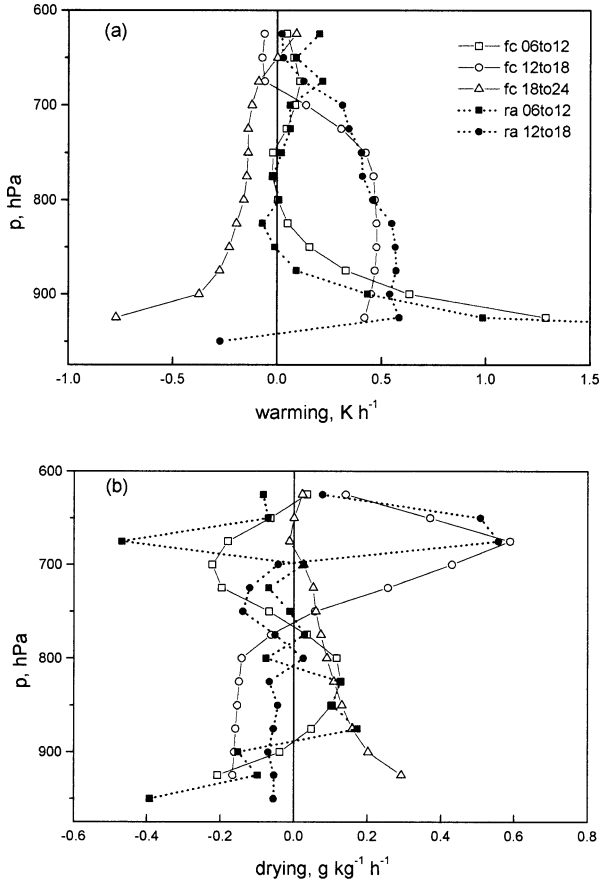

FIG. 3. Comparison of vertical profiles of (a) warming/cooling and (b) drying/moistening rates from fc06 (open symbols) and Tomelloso radiosonde (solid symbols) on 23 Jun 1991. Model warming from 0600 to 1200 UTC is denoted by fc6to12, ra6to12 denotes warming observed from radiosonde 0600–1200 UTC; other time periods are analogous.

The model does not resolve the moist RL, neither at 1200 nor at 1800 UTC. The model-produced moistening peak results rather from the underpredicted RL moisture at 1200 UTC.

Figure 4 shows the comparison of model forecast profiles at adjacent grid points. The Tomelloso gridpoint elevation is 755 m above mean sea level (MSL), the Barrax grid point is at 891 m MSL. Other than altitude effects, there are small differences in moisture (0.3–0.6 g kg^{-1}), but not in temperature. The comparison to observations for the Barrax grid point is slightly more favorable because of the enhanced moisture observed there. However, this cannot be attributed to the model performance, since the major moisture source (irrigation) is not represented in the model.

The general conclusion from the ABL comparison is that two types of differences occur. One is the general high bias in ML moisture. The other is in the vertical structure itself, since the model does not resolve the RL properly. The roots for the latter might be found in the vertical model resolution, which is insufficient for this type of landscape, and in the vertical diffusion and entrainment parameterization (see also Lenderink and Holtslag 2000). Part of the layered moisture structure observed in the RL is probably due to mesoscale circulations of topographic origin (Miao et al. 2003), which obviously are not resolved by the model.

The high bias in ML moisture could originate from various sources. The ML moisture is locally determined by the state and fluxes at the surface and at ML top, but is also conditioned by mesoscale dynamics and the large-scale flow field. The surface flux bias identified in section 5 below would produce a low bias in ML moisture, rather than the observed high bias. The entrainment fluxes cannot be directly validated. The unresolved vertical moisture layers suggest a potential error source there. The remaining candidate error source is the moisture data assimilation, where the range of relevant scales is not resolved in the model, but also not in the operational observations used in the assimilation. Problems in the analysis of moisture in regions of complex topography and close to coastlines have been reported on various occasions (e.g., Trenberth and Guillemot 1998) and remain unresolved to date (Dee and da Silva 2003).

In the present case, the only operational observations available for the moisture analysis are the radiosonde data at the locations shown in Fig. 1. First, standard radiosonde observation levels cannot capture the observed complex ABL vertical structure. Second, the two nearest stations to the EFEDA area are located in very different climate regimes (Madrid in dry continental; Murcia in semihumid Mediterranean climate). Figure 5 shows the analysis moisture profiles at the observation points Madrid (Ma) and Murcia (Mu; where the analysis scheme successfully corrects an observation error at 850 hPa).

The example provides evidence of a similar (albeit smaller) high bias as observed at the Tomelloso grid point, which lies in between these two stations and which is in the same climate and elevation regime as Madrid. This suggests shortcomings of the analysis scheme (which we do not investigate here). On the other hand, the spatial (horizontal and vertical) sampling of the observing system alone would explain part of the deficiencies in the analyzed moisture fields in this area. Along similar lines, control runs with a 3D variational assimilation scheme (instead of optimum interpolation as used in our reference version) do not improve the results.

4. Surface thermodynamic cycle

The analysis of screen-level variables uses univariate statistical interpolation (Rodríguez et al. 2003). Input data are the 2-m temperature and relative humidity from SYNOP reports. The model diagnostic package calculates the 2-m air temperature and humidity by means of interpolation between the surface (temperature and wet-

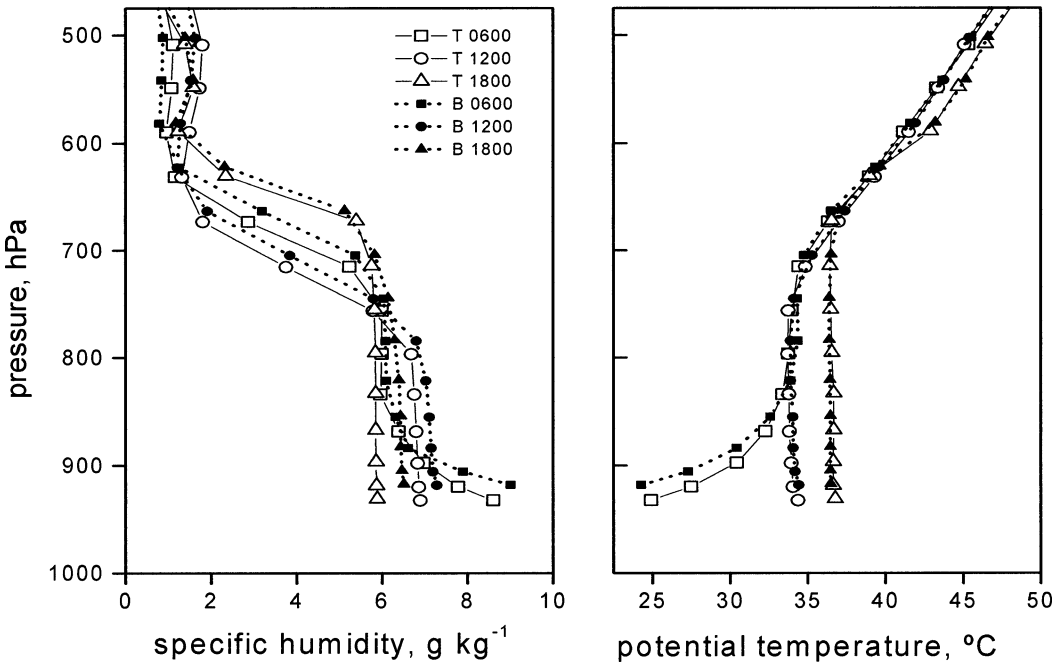

FIG. 4. Comparison of 6-h forecast profiles at Tomelloso (T; open symbols) and Barrax (B; solid symbols) grid points.

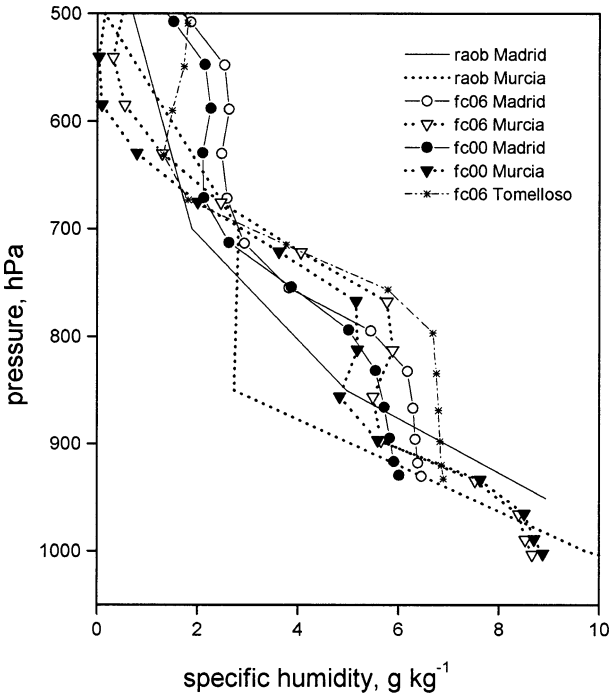

FIG. 5. Comparison of vertical profiles of specific humidity from analysis (solid symbols) and fc06 (open symbols) with operational radiosondes (no symbols) at observation points Madrid and Murcia on 23 Jun 1991 at 1200 UTC. Stars show fc06 at Tomelloso.

ness) and the lowest model level (about 30 m above ground). The special interpolation technique used for this purpose preserves the surface-layer similarity profiles (Geleyn 1988). Model data are compared with the site-aggregated screen-level surface observations. Based on the assessment of the individual measurement errors and the intrasite variability, differences in temperature and specific humidity exceeding 0.8 K and 1 g kg^{-1}, respectively, are considered to be significant and, thus, indicative of model biases.

The observations of 2-m temperature at Tomelloso (Fig. 6a) show a typical diurnal cycle of a fair-weather summer day in a dry continental climate. We find a narrow (2 h) minimum in the very early morning before sunrise (0445 UTC), followed by 6 h of rapid warming until almost noon. The warming then slows down to lead into a broad (4 h) afternoon temperature maximum. The evening cooling sets on about 1 h before sunset (1949 UTC). The Barrax 2-m temperature is generally lower, except for the early part of the rapid warming phase. After 1800 UTC it is further reduced by the onset of irrigation. The model analysis output at the four forecast intervals cannot be expected to resolve details of the diurnal cycle. At both sites (Barrax not shown), the model forecast reproduces well the midnight and 0600 UTC temperatures, although the model obviously misses the minimum. The latter is thought to be connected to the land surface and boundary layer scheme (see also Vogelezang and Holtslag 1996). The model performs well for the rapid warming phase, but then stops short of reaching the observed maximum. The 1800 UTC

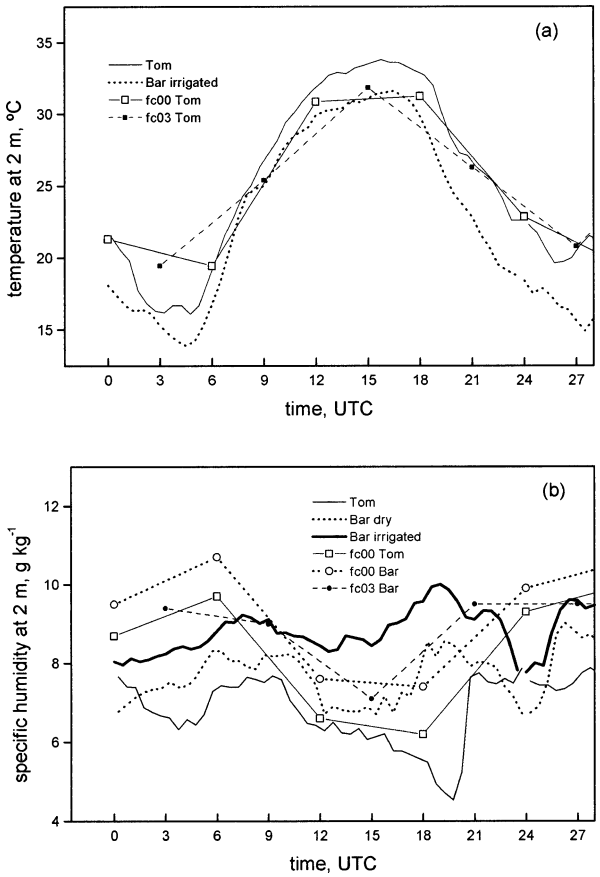

FIG. 6. Comparison of (a) 2-m temperature and (b) 2-m specific humidity from model analysis (squares) and 3-h forecast (circles) with surface observations at Tomelloso (Tom) and Barrax (Bar). For Barrax, the specific humidity observations for the major land-use classes are shown.

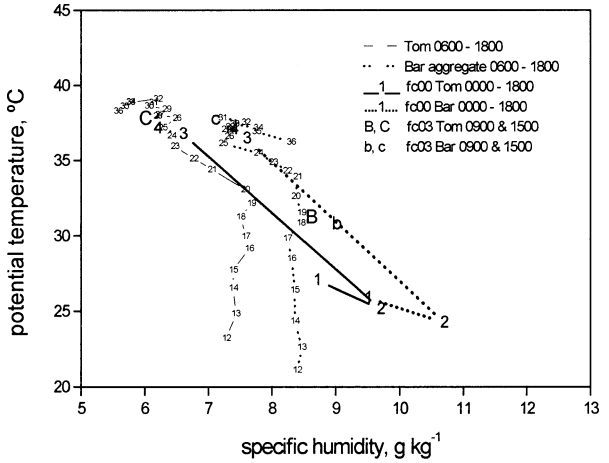

FIG. 7. Screen-level (θ, q) mixing lines from model analysis (thick lines) and surface observations (thin lines) at Tomelloso and Barrax. Observations are half-hourly values numbered consecutively from 12 (0600 UTC) to 36 (1800 UTC). Model values (marked 1–4) correspond to forecast hours (0000 to 1800 UTC). Letters B, C (b, c) show 3-h forecast at Tomelloso (Barrax) for 0900 and 1500 UTC, respectively.

temperature is slightly too low. This means that the morning surface warming slows down and stops earlier in the model than in reality. The observed high temperature amplitude of the day (17 K) cannot be fully reproduced by the model output (which does not provide daily minimum and maximum temperatures), but is obviously strongly underestimated. Intermediate 3-h forecast data (circles in Fig. 6a; not part of the routine output) confirm this conclusion.

The 2-m specific humidity diurnal cycle (Fig. 6b) exhibits a more complex behavior. We observe two maxima, one in the early-morning hours and one in the late evening till midnight. The daytime drying is strongest in the late morning, during the rapid ML growth phase, and again in the late afternoon, coinciding with the moment when the plants have closed their stomata. Because of the presence of moisture from irrigated fields, the surface at Barrax keeps moistening or stays at almost constant specific humidity until late afternoon, when some drying sets on. The model reproduces fairly well the daytime dry state and the weak drying from noon to 1800 UTC. At night, however, the model overestimates the moistening between 1800 and 2400 UTC and then keeps moistening until 0600 UTC, where it reaches a maximum that is 2.5 g kg^{-1} above the observed morning maximum.

The physical processes contributing to the daytime thermodynamic cycle can be understood from the (θ, q) mixing lines shown in Fig. 7, where θ is 2-m potential temperature. The observations show a decoupling of surface warming and drying processes during the early-morning warming period (0600–1000 UTC). To a lesser extent, this applies also to the late-morning drying period (1000–1200 UTC) (where slight warming does occur) and to the afternoon warming period (where some drying occurs). The model output obviously does not resolve these individual periods. Model warming and drying appear as coupled processes with strong drying and strong warming between 0600 and 1200 UTC (points 2 to 3 in Fig. 7). Less drying and very little warming occurs between 1200 and 1800 UTC (points 3 to 4). Recurring again to the (nonroutine) 3-h forecast data (points B at 0900 and C at 1500 UTC for Tomelloso in Fig. 7), we see evidence of two distinct, but still fully coupled, regimes in the morning phase: the first with stronger warming, and the second with stronger drying. In contrast, we see one continuous coupled regime at Barrax (points b and c).

The observed decoupling of warming and drying processes is essentially due to the complex and inhomogeneous moisture fields in the ABL and above (Michels and Jochum 1995). During the early-morning warming period, there is very little surface evaporation at Tomelloso. With the slowly growing ML, moist air is entrained from the overlying NBL. This contributes to a slight moistening of the surface until the shallow NBL

TABLE 3. Differences between surface state of model analysis and observations at Tomelloso (T) and Barrax (B).

	0000 UTC	0600 UTC	1200 UTC	1800 UTC
2-m specific humidity	Slightly too moist	Much too moist	About right	T too moist, B too dry
2-m temperature	OK	OK	Too cold	T too cold, B OK
Warming/cooling	Cooling OK	Overall cooling OK (misses minimum*)	OK, warming slows down earlier	OK
Drying/moistening	Too much moistening	Moistening instead of drying, misses minimum	Too much drying → 1200 UTC OK	OK

* The routine output does not provide daily minimum or maximum temperatures.

is eroded. This is then the start of the rapid ML growth phase, where entrainment of very dry air from the RL contributes to strong surface drying. Entrainment of RL air does not enhance warming, because the RL and ML have about the same potential temperature at that time. When the ML has grown to its maximum top (equal to the RL top), entrainment brings in warm, but not very dry, air from the free atmosphere aloft. This again leads to a weak decoupling of warming and drying.

The surface warming/cooling and drying/moistening is generally driven by the heat and moisture exchange at the surface and by entrainment processes reaching downward from the ML top to the surface (e.g., Mahrt 1991). Airborne lidar observations and bivariate conditional sampling of airborne fluxes show that entrained air almost reaches the surface layer (Michels and Jochum 1995).

As discussed in the context of the ABL profiles, the model is obviously not able to resolve the observed vertically layered moisture fields and probably does not produce strong enough entrainment. Its daytime surface thermodynamic cycle starts at 0600 UTC, potentially too warm and much too moist. The warming rate is less than observed, which offsets the wrong start and helps the model to catch up with the observed thermodynamic state at noon. During the afternoon forecast period, it seems that the model again does not reproduce the entrainment contribution to warming and drying.

Table 3 gives a summary overview of the major findings of this section.

5. Surface radiation and energy budgets

The model output surface fluxes are time integrated over the past 6-h forecast period at each forecast time (0000, 0600, 1200, and 1800 UTC). The observations consist of time series of half-hourly flux averages. Figure 8 shows the observed diurnal cycles of the surface radiation and energy budgets, respectively. For comparison with model output data, 6-h averages were derived from each time series of surface flux observations. Figure 9 and Tables 4 and 5 show the comparison of the resulting 6-h average fluxes from model and observations. The reference point for the comparison is again Tomelloso. Results for Barrax are shown and used whenever they help to gain additional insight.

a. Incoming solar radiation and aerosol load

The model overestimates the incoming solar radiation by 10% in the morning and 17% in the afternoon (Table 5 and Fig. 9a). This is surprising for the sign of the bias and even more for its magnitude. The model atmosphere is too moist and slightly too cold (section 3; Fig. 2). Therefore, one would expect a reduced shortwave (SW) transmission in the model. With a more realistic atmospheric moisture profile, the bias would even be larger. Betts et al. (1993) found a similar (10%) high bias in the T-106 L-19 cycle 39 version of the ECMWF model as compared to FIFE data. With the introduction of aerosols and the prognostic cloud scheme in the ERA-15 reanalysis scheme, the incoming shortwave radiation performed rather well for the FIFE conditions (Betts et al. 1998). Comparison of the ERA-15 data with the same EFEDA–Tomelloso observations (not shown here) does indeed show a better agreement of the incoming solar radiation (5% low in the morning; 10% high in the afternoon).

About 1–2/10 of thin cirrus cloud was observed in the area on most days. The model grid point Tomelloso does not show any cloud cover on 23 June, whereas at the Barrax grid point it gives 0.11 cirrus cloud cover at noon. The lack of cloud cover at the model grid point would explain about 2% of the bias.

In order to explore further potential sources of the SW high bias, details of the HIRLAM SW radiation scheme (Sass et al. 1994) were analyzed. The scheme is based on the parameterization of Savijärvi (1990), whose terminology is used here. For cloud-free grid elements, the incoming solar radiation at the surface (watts per meter squared) is expressed as

$$R_{sd} = S \sin h\{\underbrace{1 - 0.024(\sin h)^{-0.5}}_{\text{I}} - \underbrace{\text{aa}0.11(u/\sin h)^{0.25}}_{\text{II}} - \underbrace{\text{as}[0.28/(1 + 6.43\sin h)}_{\text{III}} - \underbrace{0.07\alpha]}_{\text{IV}}\}, \quad (1)$$

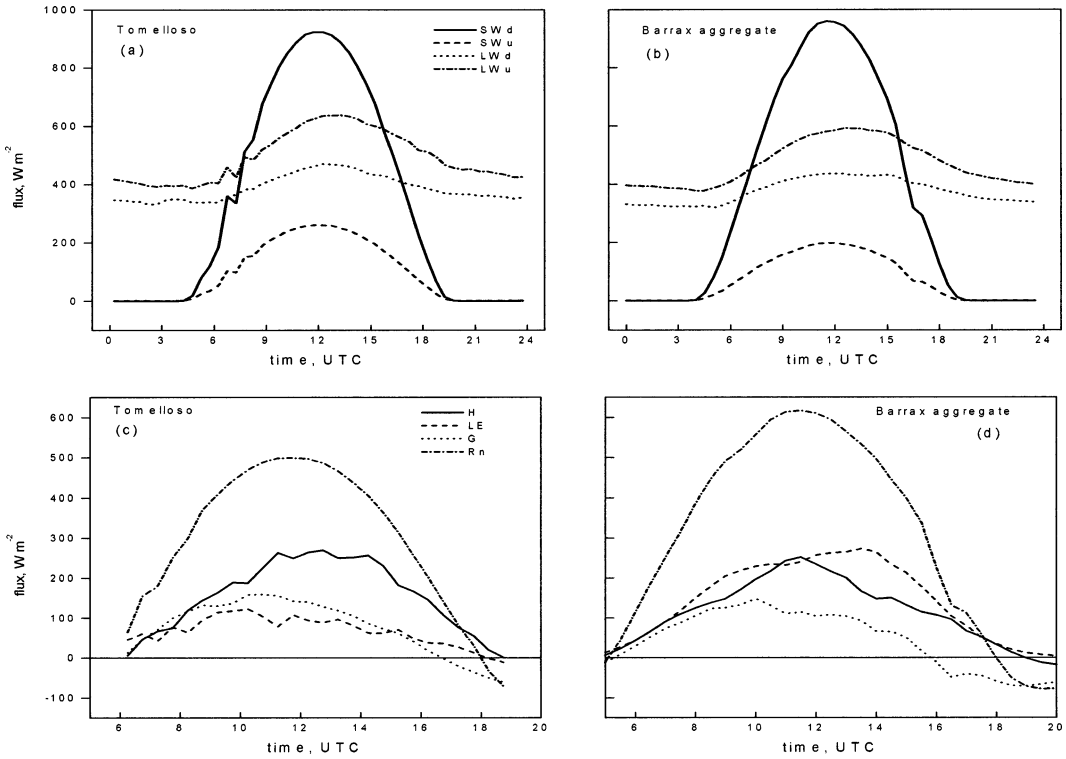

FIG. 8. (a), (b) Surface energy and (c), (d) radiation budget components at Tomelloso in (a) and (c) and Barrax in (b) and (d) on 23 Jun 1991. Observations are site-aggregated values (see text). Key: swd = SW down; swu = SW up, and for LW accordingly.

where *aa* and *as* are constants (see below), *u* is the column-integrated (linearly) pressure-scaled water vapor (centimeters), and α is the SW albedo of the ground. The ''solar constant'' *S* (watts per meter squared) is approximated as

$$S = 1365[1 + 0.034\,22\cos(2\pi d/365) + 0.0013\sin(2\pi d/365) + 0.000\,719(4\pi d/365)], \quad (2)$$

where *d* is the running date from 1 January, and *h* the local sun-height angle.

The terms in Eq. (1) represent the reduction of the top-of-the-atmosphere solar radiation ($S\sin h$) by ozone UV and visible absorption in the stratosphere (term I), water and tropospheric ozone and CO_2 absorption (II), Rayleigh scattering (III), and backscattering from the reflected beams (IV). The constants aa = 1.2 and as = 1.25 serve to include the effects of aerosol absorption and scattering, with values selected from empirical fits to data in ''continental industrialized areas during normal visibility'' (Savijärvi 1990).

EFEDA field visibility records and airborne lidar observations (Jochum 1993; Kiemle et al. 1995) suggest that the atmosphere was hazy and that visibility was low on most days. Therefore, the aerosol parameterization constants are not expected to be adequate for the area. We have calculated the individual terms of Eq. (1) from the corresponding model data. Terms I and IV are very small (below 3%). Terms II and III contribute about 20% and 7%, respectively. Enhancing the aerosol absorption coefficient *aa* from 1.2 to 1.5 increases the contribution of the corresponding term to 26%, which would result in reducing the high bias by about 6%.

Some of the aerosol may also be of marine origin, especially in the afternoon (brought into the area by the sea breeze penetrating inland from the Mediterranean coast in the afternoon). This would contribute to enhanced scattering as well, increasing the aerosol scattering coefficient as to 1.9 (Paltridge 1973) and, thus, augmenting term III by 2%–3%.

b. Radiation budget

The site-aggregated observations at Tomelloso show an essentially clear-sky net radiation, slightly reduced in amplitude by light cirrus cloud, peaking at 500 W m^{-2} shortly before local noon (Fig. 8c). The intrasite variability is around 20 W m^{-2}, which lies within the range of individual measurement errors of 5%.

The model overestimates net radiation considerably, by 28% in the morning period and 47% in the afternoon period. Part of this bias results from the incoming solar radiation bias analyzed above. In order to track further error sources, we now compare the individual terms in the surface radiation budget:

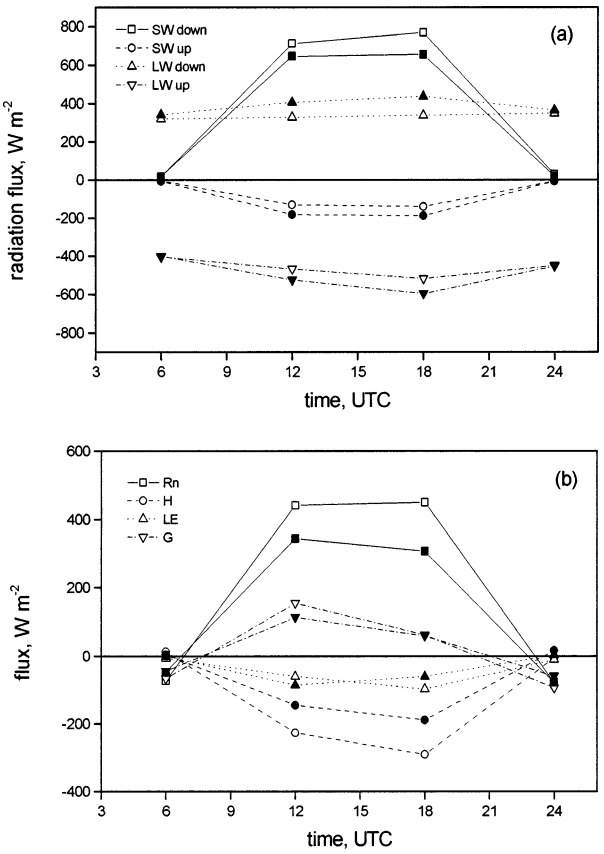

FIG. 9. Comparison of (a) surface radiation and (b) energy budget components at Tomelloso on 23 Jun 1991. Open symbols denote model output (fluxes averaged over 6-h forecast period). Solid symbols denote site-aggregated observations (mean values of 6-h period).

$$R_n = R_{sd} - R_{su} + R_{ld} - R_{lu} \quad (3)$$

where R_n is net radiation, R_{sd} downwelling SW or solar radiation, R_{su} upwelling (reflected) SW radiation, R_{ld} downward longwave (LW) radiation, and R_{lu} upward LW radiation.

The upwelling (reflected) SW radiation is underestimated by the model. This is clearly due to the low model albedo. Observations show an albedo of 0.28–0.29, whereas the model physiographic database gives a value as low as 0.17 (Table 2). This represents an underestimation of 35%–37% (morning–afternoon), which cannot even be offset by the high model incoming solar radiation. The low model albedo is close to that observed at the irrigated fields of Barrax, which is consistent with the moist and cold model surface and atmosphere.

Both LW radiation components are too low in the model, but the resulting total LW radiation is too high (18% in the morning; 12% in the afternoon), which may be responsible for part of the excessive cooling. The downward LW radiation depends very much on the vertical moisture and temperature profiles. The model ABL is too moist and slightly too cold, which might explain most of the 19%–22% low bias of downwelling LW radiation. The 11%–13% low bias in outgoing LW radiation is closely linked to the surface temperature, which is greatly underestimated by the model (Fig. 9a).

The high bias in net radiation can now be explained by means of Table 5. For the morning period, we see that the total SW bias (-117 W m^{-2}) is much higher than the total LW bias ($+21$ W m^{-2}). The relative contribution of the total SW bias to the net radiation bias is -34% (in percent of observed net radiation). This overestimate is slightly offset by the LW bias, which contributes $+6\%$ of the observed net radiation, resulting in a total high bias of 28% for net radiation. The overestimation of incoming SW radiation and the low bias in albedo contribute about equally to the total SW bias.

For the afternoon period, the total LW bias contribution is the same as in the morning ($+6\%$). The total SW bias, however, contributes much more to the overall error (-162 W m^{-2}, corresponding to -53% of observed net radiation) and, thus, explains almost entirely the gross overestimation of net radiation (47%). The incoming SW radiation bias has the largest share in it.

c. Heat and moisture fluxes

Figure 9b shows the comparison of the surface energy budget. The observational uncertainty is estimated to be up to 20% for sensible heat flux and 10%–70% for moisture flux, depending on the method used. Model and observations agree that the daytime available energy is dominated by the sensible heat flux (H). The bias in

TABLE 4. Components of the surface energy budget at Tomelloso on 23 Jun 1991, where ob = site-aggregated surface observation (mean values of 6-h period), fc = HIRLAM 6-h forecast (fluxes integrated over 6-h forecast period), dev = difference (ob − fc); reldev = difference in percent (=dev/ob), efra = evaporative fraction [$LE/(H + LE)$]. All flux units are W m^{-2}.

	Mean value 0600–1200 UTC				Mean value 1200–1800 UTC			
	ob	fc	dev	reldev	ob	fc	dev	reldev
R_n	344	441	−97	−28	307	450	−143	−47
$H + LE$	231	286	−55	−24	248	387	−140	−56
G	133	155	−41	−36	59	62	−3	−6
LE	86	60	26	30	60	97	−38	−63
H	145	226	−81	−56	188	290	−102	−54
Bowen ratio	1.6	3.8	−2.2		3.3	3.0	0.3	
Efra	0.43	0.21	0.22		0.24	0.25	−0.01	

TABLE 5. Same as in Table 4, but for surface radiation budget, where SWd = SW down, SWu = SW up, and for LW accordingly.

	ob12	fc12	dev12	reldev12	ob18	fc18	dev18	reldev18
SWd	646	712	−66	−10	655	769	−114	−17
SWu	183	132	51	28	188	141	48	25
albedo	0.28	0.18	0.10	35	0.29	0.18	0.11	37
SW total	463	580	−117	−25	467	629	−162	−35
LW total	−119	−140	21	−18	−160	−179	19	−12
LWd	406	328	78	19	437	338	98	22
LWu	525	467	57	11	597	517	79	13

latent heat flux (*LE*) is smaller in absolute magnitude and changes sign from morning to afternoon. The sensible heat flux is overestimated by 56% in the morning and by 54% in the afternoon.

It is difficult and may seem futile to perform a detailed comparison of the surface energy budget in view of the fact that the model surface scheme acts on a stage very different from the actual site setting. The model grid element is entirely covered by rather dense (50% vegetated) deciduous shrub (Table 2). Consequently, the key vegetation parameters are very different from observations (columns 2 and 5 in Table 2). The low albedo, high vegetation fraction, and low surface resistance jointly explain the biases in sensible and latent heat fluxes. We have already seen above that the low albedo explains much of the net radiation bias. The ISBA model has been extended for Mediterranean vegetation and thoroughly validated offline for the EFEDA area (Braud et al. 1993; Giordani et al. 1996; Noilhan et al. 1997). So there is no question that it works well in this environment, provided it is given the adequate input parameters. Tests with a modified land-use classification are under way at INM, in order to assess the feedbacks in the fully coupled 3D model. Yet these feedbacks are not assumed to completely alter the performance of the land surface scheme.

The potential impact of the high bias in *H* would be an increased warming of the ML, but the model ML is slightly too cold. Similarly, the potential impact of the low bias in *LE* at noon would be a too dry ML, but the model ML is very moist. This indicates that the underestimation of ML temperature and especially the overestimation of ML moisture originate from sources other than the surface forcing. As discussed above, the moisture assimilation is the prime candidate here.

d. Soil parameters

The ISBA scheme uses force-restore models to predict soil temperature and moisture (Mahfouf and Noilhan 1996). The soil is divided in two layers. The upper layer (''surface,'' with a typical depth of 1 cm) follows the diurnal cycle. The second layer (''total'') is considered to extend down to about 1 m and to act on time scales of days. The soil-surface-layer temperature T_s evolves from diurnal forcing by the net soil heat flux G and restoring of the 24-h mean temperature T_d. The soil moisture force-restore equations are based on the water balance (Rodríguez et al. 2003).

Surface layer, T_s, and 24-h mean temperature, T_d, are separately assimilated for each tile, following Giard and Bazile (2000). The method is a simple correction using the calculated 2-m temperature analysis increment at each assimilation step. Surface temperature is recalculated in the process of 2-m temperature analysis, in order to preserve the surface-layer potential temperature lapse rate (Rodríguez et al. 2003).

Surface, w_s, and total layer, w_d, water contents are also assimilated separately at each tile. Water intercepted by vegetation is not analyzed, but simply copied from the first guess to the analysis. The method to initialize w_s and w_d is based on the sequential assimilation developed by Mahfouf (1991), with optimum coefficients approximated analytically by Bouttier et al. (1993a,b), and further modified for operational implementation in the Action de Recherche Petite Echelle Grande Echelle (ARPEGE) model of the French Weather Service by Giard and Bazile (2000). The analyzed 2-m temperature and relative humidity are used as observations for the analysis of soil water contents, which is based on optimum interpolation and soil moisture nudging as described by Rodríguez et al. (2003).

The model soil heat flux G is calculated from the surface energy budget ($R_n = H + LE + G$). Thus, the soil heat flux bias (Fig. 9b; Table 4) is a direct result of the errors in the other three terms of the surface energy budget. In the afternoon period, the errors in the three components R_n, H, and LE offset each other, such that the resulting G is almost unbiased (−6%). The situation is different in the morning period, where the remaining bias in G is −36%; G is also too high at night (during both periods), basically as a result of the R_n bias.

The observed radiometric surface temperature (Fig. 10) has a wide amplitude diurnal cycle, reaching from an 18°C minimum at sunrise to a noon maximum of 55°C. Soil temperature was measured at various levels (at a depth of 3, 5, 10, 25, and 50 cm at some and at 0, 5, 10, 16, 22, and 30 cm at other locations within the Tomelloso site). These observations were vertically averaged to match the ISBA soil temperature definitions. Obviously, the results depend on the choice of the vertical averaging bins. On the other hand, the two layers are not quantitatively defined in ISBA. Figure 10 shows

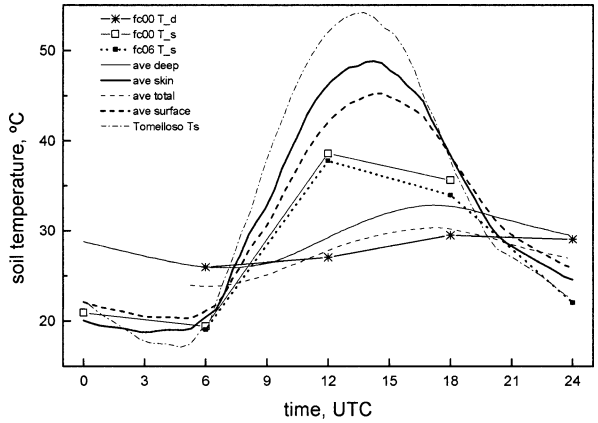

FIG. 10. Surface and soil temperature comparison at Tomelloso: Model analysis 24-h mean temperature (asterisks) and analysis, and 6-h forecast surface temperature (open and solid squares, respectively). Two pairs of observational averages correspond to two different depth range combinations (see text for details). Pair 1: ''total'' (averaged over levels 10–50 cm) and ''surface'' (averaged over levels 0–5 cm). Pair 2: ''deep'' (levels 5–50 cm) and ''skin'' (0–3 cm). Dashed–dotted curve shows radiometric skin temperature.

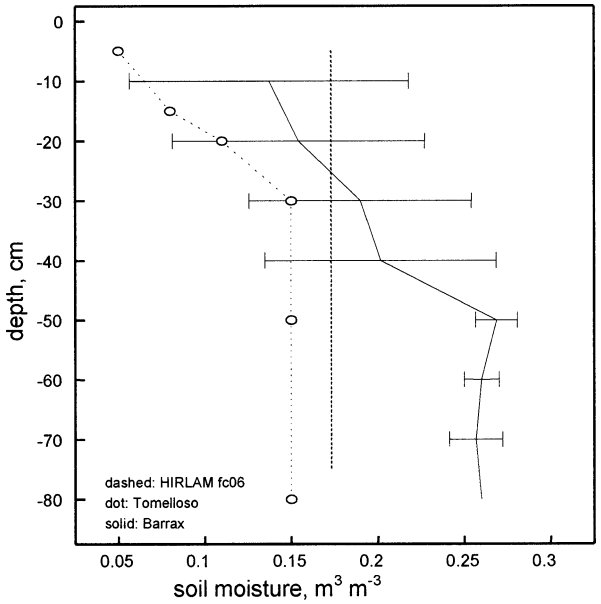

FIG. 11. Soil moisture comparison of model (dashed vertical line) and observations at Tomelloso (dotted, o) and Barrax (solid curve, error bars give standard deviation of layer mean).

two alternative groupings of depth ranges. The first pair consists of a ''surface'' layer (with observations averaged over a depth of 0–5 cm) and a ''total'' layer (10–50 cm). The second pair holds a ''skin'' layer (0–3 cm) and a ''deep'' layer (5–50 cm). Both the ''skin'' and ''surface'' layer temperatures have also a large diurnal amplitude, albeit smaller than the radiometric skin temperature. Fractional ground cover plays an important role here, since midday differences between bare soil and foliage temperatures are large (e.g., Giordani et al. 1996). The model reproduces well the nighttime low values and the 1800 UTC values, but fails by far to reach the daytime maximum of either surface temperature.

The sources of the surface temperature bias are basically twofold. First, the surface analysis scheme ties the soil surface temperature closely to the 2-m temperature, which is slightly underestimated. Second, the soil heat flux bias propagates directly into the soil temperatures. For the afternoon period (1200–1800 UTC), the modeled G agrees very well with observations (-6%). This leads to the observed good agreement of surface temperature at 1800 UTC. The morning G is 36% too high. In addition to other potentially relevant soil parameterization effects, this results in the significant underestimation of surface temperature at noon, which in turn affects both the longwave radiation and the heat and moisture flux. Note that the modeled surface temperature and related variables are also very sensitive to the setting of surface parameters, in particular the roughness length of heat (e.g., Holtslag and Ek 1996).

The comparison of the model 24-h mean soil temperature, T_d, against the observational averages over certain vertical bins is not straightforward, as T_d has not an exact correspondance with a certain depth level. However, the model T_d matches fairly well the daytime values of the total layer, while overestimating them at nighttime (0600 and 2400 UTC). Conversely, it comes close to the observational deep layer at night, but shows a low bias during daytime, which is probably again due to the soil heat flux bias. The slight high bias of total-layer soil moisture (Fig. 11) might be simply due to the assignment of soil type (class 5 has a higher clay percentage than class 3 found at the site.). However, we have not considered the possible effects of soil moisture assimilation, so the bias origin might be more complex.

e. Summary discussion

In conclusion, we have identified a high bias in daytime incoming solar radiation (originating from the aerosol description in the radiation code) and a low bias in albedo (rooted in the vegetation class assignment). Both (in descending order of magnitude) cause a large overestimation of net radiation (28% in the morning; 47% in the afternoon). Model errors in LW radiation components are of smaller magnitude and have less impact on net radiation.

We also have identified deficiencies in all other terms of the surface energy budget. The sensible heat flux has a high bias of 56% (morning; 54% afternoon), while the moisture flux bias changes sign from 30% low (morning) to 63% high (afternoon). The discrepancy essentially results from the vegetation class and subsequent parameter attribution in the model physiographic database, where the Tomelloso grid element is much more densely vegetated than the sparse vine canopy

actually growing there. Even at the irrigated site Barrax, the situation is similar.

The soil heat flux bias is a direct result of the errors in the other three terms of the surface energy budget, R_n, H, and LE. In the afternoon period, these errors offset each other, such that the resulting G is almost unbiased. This leads to a very good agreement of modeled surface temperature with observations at that time. The situation is different in the morning period, where the remaining bias in G is -36%. This overestimation contributes significantly to the large low bias in surface temperature at noon; G is also too high at night (during both periods), basically as a result of the R_n bias.

Based on these conclusions and knowing that the ISBA scheme has performed well in offline validations for the same environment, we have not entered in a detailed comparison of the surface energy budget. This will be an interesting task once the model's land-use classification and physiographic database includes an adequate representation of Mediterranean ecosystems. Based on the same EFEDA dataset, Jochum et al. (2004, manuscript submitted to *J. Appl. Meteor.*) show that the EFEDA grid-scale aggregated fluxes depend critically on the fraction of irrigated crops in a given grid cell. Consequently, they propose a simple and efficient way to incorporate the corresponding plant phenological parameters into the physiographic database of a given NWP model, by distinguishing two classes of irrigated crops. Spring-irrigated and summer-irrigated crops have different phenological cycles; their discrimination in multitemporal remote sensing–based land-use classification is straightforward (Martínez and Calera 2001). The classification of Calera Belmonte (2000) is a regional-scale example of the approach, and Jochum et al. (2004, manuscript submitted to *J. Appl. Meteor.*) show its usefulness in scale aggregation for Mediterranean landscapes.

6. Conclusions

Observations from the European Field Experiment in a Desertification-threatened Area (EFEDA) were used to evaluate the performance of the radiation, land surface, and boundary layer description of HIRLAM in semiarid conditions. Analysis and 6-h forecast data of the fully coupled three-dimensional model were compared with the comprehensive dataset of a case study representative for a sample of 22 days of anticyclonic conditions.

The model surface, soil, and boundary layer are found to be too moist and slightly too cold during most of the diurnal cycle. The model radiation and surface energy budgets are biased toward more humid conditions.

Summarizing the overall results of the comparison and assigning priorities (based on bias magnitudes and impacts) to the observed discrepancies, we obtain the following picture:

1) Assignment of too dense vegetation leads to biased surface fluxes and net radiation. With adequate land-use class assignment, the ISBA surface scheme performs well.
2) The incoming solar radiation is overestimated, which results from the aerosol parameterization in the radiation scheme. Modification of the aerosol parameters to account for lower visibility reduces the bias significantly.
3) Large errors (up to 50%) in the ABL humidity profiles result from deficiencies in the moisture assimilation.
4) The residual layer is not resolved properly and the entrainment contribution to surface and ML warming and drying is not reproduced adequately because of the predominant effect of vertical model resolution, which is insufficient for this type of landscape, and associated deficiencies in the vertical diffusion and entrainment parameterization.
5) A soil type with too high clay percentage might generate too much soil moisture, although the soil moisture assimilation effects would need to be assessed here as well.
6) The high bias in net longwave radiation is probably related to the excess atmospheric moisture and the underestimated surface temperature.

Immediate improvement of the model performance can be expected from the use of a land-use and soil classification (with its associated physiographic database) adapted to Mediterranean landscapes, in combination with the use of aerosol parameters in the radiation scheme, that account for the typically higher aerosol load of arid and semiarid environments.

An adequate land-use classification for Mediterranean landscapes would need to account for sparse dry canopies as well as for two classes of irrigated crops (spring-irrigated and summer-irrigated), with a physiographic database reflecting their different phenological cycles. The classification of Calera Belmonte (2000) is a regional-scale example of the approach. The classification of Champeaux et al. (2000) follows a similar approach, based on the seasonality of the Normalized Difference Vegetation Index (NDVI) in more general terms and covering all western Europe. The use of this classification in HIRLAM is currently being tested at the Spanish Weather Service.

Most of these conclusions are expected to be valid for other NWP models as well. In fact, the ongoing comparison of ERA-15 and ERA-40 data with the same observational data gives similar results. Our results are in line with other NWP validation studies, which have often focused on the land surface scheme. Major NWP centers are currently upgrading their land-use classification and physiographic database (e.g., van den Hurk et al. 2000; Gustafsson et al. 2003; Masson et al. 2003) to account for a wider variety of land-use classes and for more realistic seasonal variations in key canopy pa-

rameters. Irrigated crops, however, have not yet been separated in seasonal classes.

The radiation scheme is more specific for HIRLAM. Other schemes might require a different analysis and solution, although the higher aerosol load prevailing in (semi-) arid environments would generally need to be accounted for. An improved aerosol parameterization for Mediterranean areas could draw on aerosol climatologies (e.g., Holben et al. 2001) or use satellite-derived aerosol datasets (e.g., Yu et al. 2003) that are recently becoming available.

Acknowledgments. Working with such a complete and high quality dataset was only possible thanks to the efforts of all EFEDA participants. In particular, we would like to thank Pierre Bessemoulin and Norbert Kalthoff for making available their radiosonde data and Joël Noilhan for providing us with the surface dataset of Linder et al. (1996). Thanks also go to Alfonso Calera for making available his land-use classification, to Beatriz Navascués for the fruitful discussions on the performance of the surface parameterization and assimilation modules, and to Bart van den Hurk for comments on an earlier manuscript version. The detailed comments of one of the reviewers were very helpful. The EFEDA 1991 field phase was cofunded by the European Commission under its Environment Program (Grant EV5V-CT93-0271).

REFERENCES

Avissar, R., and R. A. Pielke, 1989: A parameterization of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology. *Mon. Wea. Rev.,* **117,** 2113–2136.

Betts, A. K., and J. H. Ball, 1998: FIFE surface climate and site-averaged dataset 1987–89. *J. Atmos. Sci.,* **55,** 1091–1108.

——, ——, and A. C. M. Beljaars, 1993: Comparison between the land-surface response of the ECMWF model and the FIFE-1987 data. *Quart. J. Roy. Meteor. Soc.,* **119,** 975–1001.

——, F. Chen, K. E. Mitchell, and Z. I. Janjic, 1997: Assessment of the land surface and boundary layer models in two operational versions of the NCEP Eta Model using FIFE data. *Mon. Wea. Rev.,* **125,** 2896–2916.

——, P. Viterbo, and A. C. M. Beljaars, 1998: Comparison of the land-surface interaction in the ECMWF reanalysis model with the 1987 FIFE data. *Mon. Wea. Rev.,* **126,** 186–198.

——, ——, ——, and B. J. J. M. van den Hurk, 2001: Impact of BOREAS on the ECMWF forecast model. *J. Geophys. Res.,* **106,** 33 593–33 603.

Bolle, H.-J., and Coauthors, 1993: EFEDA: European Field Experiment in a Desertification-threatened Area. *Ann. Geophys.,* **11,** 173–189.

Bouttier, F., J.-F. Mahfouf, and J. Noilhan, 1993a: Sequential assimilation of soil moisture from atmospheric low-level parameters. Part I: Sensitivity and calibration studies. *J. Appl. Meteor.,* **32,** 1335–1351.

——, ——, and ——, 1993b: Sequential assimilation of soil moisture from atmospheric low-level parameters. Part II: Implementation in a mesoscale model. *J. Appl. Meteor.,* **32,** 1352–1364.

Braud, I., J. Noilhan, P. Bessemoulin, P. Masacart, R. Haverkamp, and M. Vauclin, 1993: Bare-ground surface heat and water exchanges under dry conditions: Observations and parameterization. *Bound.-Layer Meteor.,* **66,** 173–200.

Bringfelt, B., M. Heikinheimo, N. Gustafsson, V. Perov, and A. Lindroth, 1999: A new land surface treatment for HIRLAM—Comparisons with NOPEX measurements. *Agric. For. Meteor.,* **198–199,** 239–256.

Calera Belmonte, A., 2000: Seguimiento mediante teledetección de la cubierta vegetal de los cultivos de secano y su relación con variables climáticas en Castilla La Mancha. Ph.D. thesis, University of Valencia, 315 pp. [Available from A. Calera, IDR, E-02071 Albacete, Spain.]

Champeaux, J.-L., D. Arcos, E. Bazile, D. Giard, J.-P. Goutorbe, F. Habets, J. Noilhan, and J.-L. Roujean 2000: AVHRR-derived vegetation mapping over western Europe for use in numerical weather prediction models. *Int. J. Remote Sens.,* **21,** 1183–1199.

Cuxart, J., P. Bougeault, and J. L. Redelsperger, 2000: A turbulence scheme allowing for mesoscale and large-eddy simulations. *Quart J. Roy. Meteor. Soc.,* **126,** 1–30.

Dee, D. P., and A. M. da Silva, 2003: The choice of variable for atmospheric moisture analysis. *Mon. Wea. Rev.,* **131,** 155–171.

Geleyn, J. F., 1988: Interpolation of wind, temperature and humidity values from the model levels to the height of measurement. *Tellus,* **40A,** 347–351.

Giard, D., and E. Bazile, 2000: Implementation of a new assimilation scheme for soil and surface variables in a global NWP model. *Mon. Wea. Rev.,* **128,** 997–1015.

Giordani, H., J. Noilhan, P. Lacarrère, P. Bessemoulin, and P. Mascart, 1996: Modelling the surface processes and the atmospheric boundary layer for semi-arid conditions. *Agric. For. Meteor.,* **80,** 263–287.

Gustafsson, D., and Coauthors, 2003: Boreal forest surface parameterization in the ECMWF model–1D test with NOPEX long-term data. *J. Appl. Meteor.,* **42,** 95–112.

Halldin, S., S. E. Gryning, L. Gottschalk, A. M. Jochum, L.-C. Lundin, and A. A. Van de Griend, 1999: Energy, water and carbon exchange in a boreal forest landscape—NOPEX experiences. *Agric. For. Meteor.,* **98–99,** 5–29.

Holben, B. N., and Coauthors, 2001: An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET. *J. Geophys. Res.,* **106,** 12 067–12 097.

Holtslag, A. A. M., E. van Meijgaard, and W. C. de Rooy, 1995: A comparison of boundary-layer diffusion schemes in unstable conditions over land. *Bound.-Layer Meteor.,* **76,** 69–95.

——, and M. Ek, 1996: Simulation of surface fluxes and boundary layer development over the pine forest in HAPEX-MOBILHY. *J. Appl. Meteor.,* **35,** 202–213.

Jochum, A. M., 1993: Evaporation and energy fluxes during EFEDA: Horizontal variability and area averaging. *Exchange Processes at the Land Surface for a Range of Space and Time Scales*. H.-J. Bolle, R. A. Feddes, and J. D. Kalma, Eds., IAHS Press, 373–381.

——, P. Kabat, and R. Hutjes, 2000: The role of remote sensing in land-surface experiments within BAHC and ISLSCP. *Observing Land from Space: Science, Customers and Technology,* M. Verstraete, M. Menenti, and J. Peltoniemi, Eds., Kluwer Academic, 91–103.

Kiemle, C., M. Kästner, and G. Ehret, 1995: The convective boundary layer structure from lidar and radiosonde measurements during the EFEDA'91 campaign. *J. Atmos. Oceanic Technol.,* **12,** 771–782.

Koster, R. D., and M. J. Suarez, 1992: Modeling the land surface boundary in climate models as a composite of independent vegetation stands. *J. Geophys. Res.,* **97,** 2697–2715.

Lenderink, G., and A. A. M. Holtslag, 2000: Evaluation of the kinetic energy approach for modeling turbulent fluxes in stratocumulus. *Mon. Wea. Rev.,* **128,** 244–258.

Linder, W., J. Noilhan, M. Berger, K. Bluemel, E. Blyth, and G. Boulet, 1996: Intercomparison of surface schemes using EFEDA flux data. CNRM Rep. 39, Météo-France, 86 pp. [Available from J. Noilhan, CNRM, F-31057, Toulouse, France.]

Louis, J. F., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Bound.-Layer Meteor.,* **17,** 187–202.

——, M. Tiedtke, and J. F. Geleyn, 1982: A short history of the PBL parameterization at ECMWF. *Proc. ECMWF Workshop on Planetary Boundary Layer Parameterization,* Reading, United Kingdom, ECMWF, 59–80. [Available from ECMWF, Shinfield Park, Reading, Berkshire RG29AX, United Kingdom.]

Mahfouf, J.-F., 1991: Analysis of soil moisture from near-surface parameters: A feasibility study. *J. Appl. Meteor.,* **30,** 1534–1547.

——, and J. Noilhan, 1996: Inclusion of gravitational drainage in a land surface scheme based on the force-restore method. *J. Appl. Meteor.,* **35,** 987–992.

Mahrt, L., 1991: Boundary-layer moisture regimes. *Quart. J. Roy. Meteor. Soc.,* **117,** 151–176.

Manzi, A. O., and S. Planton, 1994: Implementation of the ISBA parameterization scheme for land surface processes in a GCM—An annual cycle experiment. *J. Hydrol.,* **155,** 353–387.

Martínez, C., and A. Calera, 2001: Irrigated crop area estimation from thematic map using Landsat TM imagery in La Mancha (Spain). *Photogramm. Eng. Remote Sens.,* **67,** 1177–1184.

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

Miao, J.-F., L. J. M. Kroon, J. Vilà-Guerau de Arellano, and A. A. M. Holtslag, 2003: Impacts of topography and land degradation on the sea breeze over eastern Spain. *Meteor. Atmos. Phys.,* **84,** 157–170.

Michels, B. I., and A. M. Jochum, 1995: Heat and moisture flux profiles in a region with inhomogeneous surface evaporation. *J. Hydrol.,* **166,** 383–407.

Noilhan, J., and S. Planton, 1989: A simple parameterization of land surface processes for meteorological models. *Mon. Wea. Rev.,* **117,** 536–549.

——, P. Lacarrère, A. J. Dolman, and E. M. Blyth, 1997: Defining area-average parameters in meteorological models for land surface with mesoscale heterogeneity. *J. Hydrol.,* **190,** 302–326.

Paltridge, G. W., 1973: Direct measurements of water vapor absorption of solar radiation in the free atmosphere. *J. Atmos. Sci.,* **30,** 156–160.

Rodríguez, E., B. Navascués, J. J. Ayuso, and S. Järvenoja, 2003: Analysis of surface variables and parameterization of surface processes in HIRLAM. Part I: Approach and verification by parallel runs. HIRLAM Tech. Rep. 58, SMHI, 52 pp. [Available from Per Unden, SMHI, S-601 76 Norrkoeping, Sweden.]

Sass, B. H., L. Rontu, and P. Räisänen, 1994: HIRLAM-2 radiation scheme: Documentation and tests. HIRLAM Tech. Rep. 16, SMHI, 42 pp. [Available from Per Unden, SMHI, S-601 76 Norrkoeping, Sweden.]

Savijärvi, H., 1990: Fast radiation parameterization schemes for mesoscale and short-range forecast models. *J. Appl. Meteor.,* **29,** 437–447.

Sellers, P. J., F. G. Hall, G. Asrar, D. E. Strebel, and R. E. Murphy, 1992: An overview of the First ISLSCP Field Experiment. *J. Geophys. Res.,* **97,** 18 345–18 372.

——, and Coauthors, 1997: BOREAS in 1997: Experiment overview, scientific results and future directions. *J. Geophys. Res.,* **102,** 28 731–28 769.

Trenberth, K. E., and C. J. Guillemot, 1998: Evaluation of the atmospheric moisture and hydrological cycles in the NCEP/NCAR reanalyses. *Climate Dyn.,* **14,** 213–231.

Unden, P., and Coauthors, 2002: The HIRLAM model (version 5.2). HIRLAM Scientific Rep., 144 pp. [Available online at http://hirlam.knmi.nl/open/publications/SciDoc_Dec2002.pdf.]

van den Hurk, B. J. J. M., P. Viterbo, A. C. M. Beljaars, and A. K. Betts, 2000: Offline validation of the ERA40 surface scheme. ECMWF Tech. Memo. 295, 42 pp. [Available from ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, United Kingdom.]

Viterbo, P., and A. C. M. Beljaars, 1995: A new land surface parameterization scheme in the ECMWF model and its validation. *J. Climate,* **8,** 2716–2748.

Vogelezang, D. H. P., and A. A. M. Holtslag, 1996: Evaluation and model impacts of alternative boundary-layer height formulations. *Bound.-Layer Meteor.,* **81,** 245–269.

Yu, H., R. E. Dickinson, M. Chin, Y. J. Kaufman, B. N. Holsen, I. V. Geogdzhayev, and M. I. Mischenko, 2003: Annual cycle of global distributions of aerosol optical depth from integration of MODIS retrievals and GOCART model simulations. *J. Geophys. Res.,* **108,** 4128, doi:10.1029/2002JD002717.

Yucel, I., W. J. Shuttleworth, and J. Washburne, 1998: Evaluating NCEP Eta Model–derived data against observations. *Mon. Wea. Rev.,* **126,** 1977–1991.