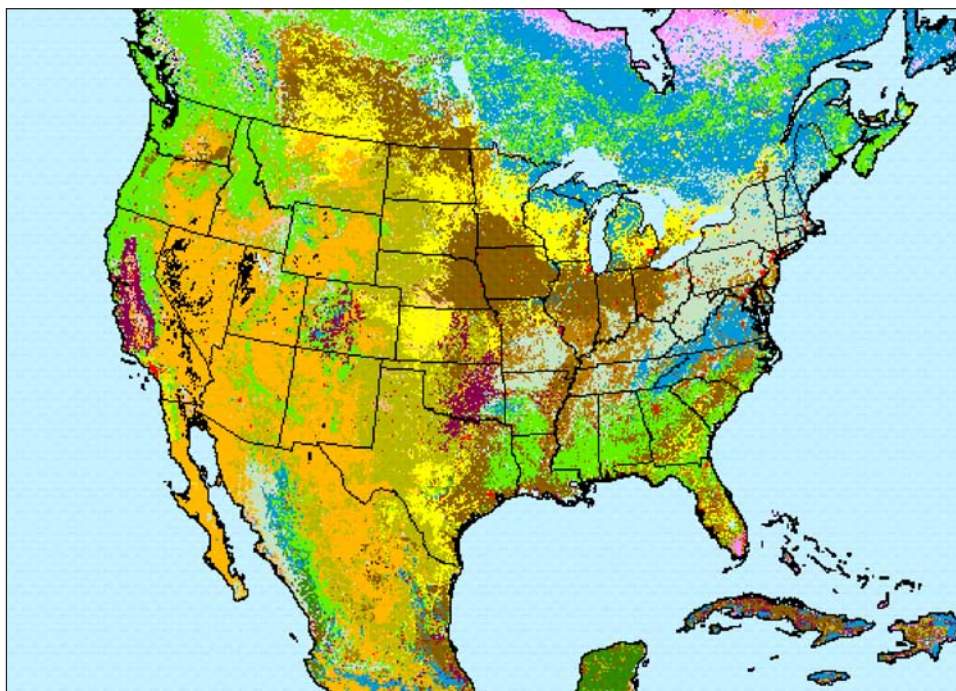**Scientific Report 2005-01**

Development of the Regional Climate-Weather Research and Forecasting (CWRF) Model: Surface Boundary Conditions

by

Xin-Zhong Liang, Hyan I. Choi, and Kenneth E. Kunkel
Illinois State Water Survey

Yongjiu Dai
Beijing Normal University

Everette Joseph
Howard University

Julian X.L. Wang
National Oceanic and Atmospheric Administration

Praveen Kumar
University of Illinois at Champaign-Urbana

Illinois State Water Survey
A Division of the Illinois Department of Natural Resources

2005

Front Cover: The figure illustrates the geographic distribution of the USGS land-cover categories based on the AVHRR satellite 1-km resolution data.

Contents

Acknowledgements

Abstract

The Climate-Weather Research and Forecasting (CWRF) is the climate *extension* of the Weather Research and Forecasting (WRF) model, incorporating all WRF functionalities for numerical weather predictions while enhancing the capability for climate applications. This report focuses on the construction and implementation of surface boundary conditions (SBCs) specifically designed for CWRF mesoscale modeling applications. The primary SBCs include surface topography (mean elevation, slope, curvature, and their standard deviations); bedrock, lakebed, or seafloor depth; soil sand, and clay fraction profiles; surface albedo localization factor; bottom soil temperature; surface characteristic identification; land cover category; fractional vegetation cover; leaf and stem area index; sea surface temperature, salinity, and current; and sea temperature and salinity profiles. They are currently presented in a CWRF domain suitable for the U.S applications at 30-kilometer spacing. The raw data sources and processing procedures, however, are elaborated in detail, by which the SBCs can be readily constructed for any specific CWRF domain anywhere in the world. For a specific field, alternative data sources, if available, were compared to quantify uncertainties and suggest the choice or improvement.

Introduction

Mesoscale regional climate models (RCMs) are recognized as an increasingly important tool to address scientific issues associated with climate variability, changes, and impacts at local and regional scales (Giorgi and Mearns 1999; IPCC 2001; Leung et al. 2003). Numerous RCMs have been developed, applied, and intercompared, demonstrating important downscaling skills achieved and model deficiencies yet to be resolved (Takle et al. 1999; Leung et al. 1999; Roads et al. 2003; and references therein). The most widely used RCM was the RegCM2, developed by Giorgi et al. (1993a, b) based on the Penn State/National Center for Atmospheric Research (NCAR) Mesoscale Model version 4 or MM4 (Anthes et al. 1987). This was later integrated with sea ice and ocean models to establish the ARCSyM (Lynch et al. 1995), the single existing fully coupled RCM. Over the years, the hydrostatic MM4 has been improved significantly and eventually succeeded by the nonhydrostatic MM5 (Grell et al. 1994; Dudhia et al. 2000). Several RCMs then built upon the MM5 as the basic framework and gained a wide range of applications (e.g., Leung and Ghan 1999; Liang et al. 2001, 2004a; Wei et al. 2002).

Meanwhile, the next-generation Weather Research and Forecasting (WRF) model (http://www.wrf-model.org/) is being developed by a broad community of government and university researchers (Klemp et al. 2000; Michalakes 2000; Chen and Dudhia 2000). Because the WRF was based upon the most advanced supercomputing technologies and promises greater efficiency in computation and flexibility in new module incorporation, it eventually will supercede the MM5. An initial study demonstrated WRF capability and limitation for regional climate applications (Liang et al. 2002). In particular, the WRF exhibits an encouraging advance in reproducing North American summer monsoon rainfall, which is poorly simulated in the MM5-based RCM or CMM5 (Liang et al. 2004a) and remains a challenging issue in the global and regional climate modeling community (NAME 2003). Since then, the authors have committed, in collaboration with the WRF Working Groups, to develop the *climate extension* of the WRF (CWRF) to replace the CMM5 (Liang et al. 2004a) and, after sufficient validation, for general public use.

The WRF originally was designed for numerical weather prediction (NWP) and not expressly for climate studies. To extend its capability for applications on regional climate scales, the authors developed the CWRF with four crucial characteristics to improve: [1] *planetary-mesoscale interaction* by including an optimal buffer zone treatment that integrates realistic energy and mass fluxes across the lateral boundaries of the RCM domains (Liang et al. 2001); [2] *surface-atmosphere interaction* by incorporating new physics modules for planetary boundary layer, and land surface and terrestrial hydrology, as well as observed variations or dynamic predictions of vegetation, ocean, and sea ice; [3] *convection-cloud-radiation interaction* by implementing fully coupled, new physical parameterizations for cumulus clouds, cloud microphysics, cloud formation, and radiative transfer; and [4] *system consistency throughout all process modules* by using unified water vapor saturation and solar zenith angle functions, common physical constants, and coherent tunable parameters. These improvements were accomplished through iterative, extensive model refinements, sensitivity experiments, and rigorous validations over the past three years. As a result, the CWRF has demonstrated greater capability and better performance in simulating the U.S. regional climate than the CMM5 (Liang et al. 2001, 2004a).

The concept of the CWRF is in its emphasis on the extension of the WRF. This extension incorporates inclusively all WRF functionalities for NWP while enhancing the capability for climate applications. As such, the CWRF has applications for both weather forecasts and climate predictions. This unification offers an unprecedented opportunity to develop, test, and verify new physical parameterizations of unresolved processes, identify their systematic errors, and eventually improve them over a wide range of frequencies from weather to climate scales. Given that systematic climate biases result from nonlinear interactions among many dynamic and physic processes, it is impossible to unravel RCM deficiencies in specific physical parameterizations solely by diagnosing climate simulations. Incorporation of the WRF data assimilation system, which ingests all available observations and applies variational methods to produce an optimal reality analysis, enables the CWRF to produce short-range weather forecasts from realistic initial conditions. Because the model dynamics evolve freely and interact fully with the physical parameterizations during the forecasts, the CWRF consistently generates all forcings and feedbacks. Thus high-frequency NWP analyses and unassimilated observations of parameterized variables can be used to identify parameterization deficiencies and gain insights on improvements manifested initially in short-range weather forecasts, but ultimately that persisted in climate simulations (Phillips et al. 2004). The parameterizations then can be modified to reduce the perceived high-frequency deficiencies and further evaluated beyond the deterministic forecast range (~15 days) to determine whether they ameliorate low-frequency biases in progressively longer (seasonal, interannual, or decadal) climate simulations. Some systematic climate biases that develop slowly probably cannot be identified and reduced by the NWP-based approach. For example, WRF mass and height dynamic cores reveal little sensitivity in weather forecasts while generating substantial systematic differences in climate simulations. The authors later found, via CWRF sensitivity experiments, that these differences were caused by erroneous lateral boundary conditions (LBCs) due to inappropriate use of surface rather than sea-level pressures by WRF Standard Initialization or SI (http://www.wrf-model.org/si/) in mapping the coarse-resolution global reanalysis into the RCM grid. Clearly, the CWRF provides a unique tool to develop improved schemes for realistic weather and climate applications.

A series of papers being prepared document details of CWRF formulations and skills for weather forecasts and climate predictions. This report is the first of the series to depict the construction and implementation of surface boundary conditions (SBCs). These SBCs also will be amenable to the requirements of existing WRF modules that simulate surface processes. A brief description of the current CWRF representation of these processes precedes general considerations introducing the SBCs wanted or required for normal model operations. There is also a discussion of how to

construct these SBCs, their important characteristics, and how to implement them into the WRF input/output application program interface (I/O API). This effort is unique in that, for the first time, the science community will have access to the most comprehensive SBCs specifically designed for mesoscale modeling applications, which is especially beneficial for the CWRF users.

CWRF Representation of Surface Processes

Most relevant to the SBCs is how the CWRF simulates local surface processes over land and oceans, for which a more detailed description follows. The WRF release version 2 included the 5-layer SLAB thermal diffusion model (Dudhia 1996); the 6-layer Rapid Update Cycle (RUC) land surface model or LSM (Smirnova et al. 2000); and the 4-layer NOAH LSM, jointly developed by National Centers for Environmental Prediction (NCEP), Oregon State University, U.S. Air Force Weather Agency, and NOAA Office of Hydrology (Chen and Dudhia 2001; Ek et al. 2003). The Common Land Model or CLM (Dai et al. 2003), a state-of-the-art model for Soil-Vegetation-Atmosphere Transfer (SVAT), was incorporated into the CWRF. Major CLM characteristics include: a 10-layer prediction of soil temperature and moisture; a 5-layer prediction of snow processes (mass, cover, and age); an explicit treatment of liquid and ice water mass and their phase change in soil and snow; a runoff parameterization based on the Topmodel concept (Beven and Kirkby 1979); a canopy photosynthesis-conductance scheme that describes the simultaneous transfer of carbon dioxide and water vapor to and from vegetation; a tiled treatment of subgrid fraction of energy and water balance; and high-resolution geographic distributions of land cover, vegetation, and root and soil properties. The CLM has been evaluated extensively in stand-alone mode with field measurements (Dai et al. 2003), indicating realistic simulations of state variables (soil moisture, soil temperature, and snow water equivalent) and flux terms (net radiation, latent and sensible heat fluxes, and runoff). Preliminary CWRF results showed that the CLM, when compared with the existing WRF modules, significantly improves representation of land surface processes, and also facilitates the consistent coupling with the radiation transfer, where surface albedo and emissivity are predicted by the CLM.

The latest version CLM incorporates several important updates and new modules. *Major updates* include a two-leaf (sunlit and shaded) canopy treatment for temperature, radiation, and photosynthesis-stomatal resistance; a two-stream approximation for canopy albedo; the bedrock depth effect on soil thermal and hydrological processes; a new canopy interception treatment accounting for partitioning between convective and stable precipitation; turbulent transfer under the canopy; an efficient iterative solution for leaf temperature; separation of surface and subsurface runoff; rooting fraction and water stress on transpiration; and perfect energy and water balance within every time-step. *A dynamic vegetation module* (DVM) that integrates the interactive canopies with the full carbon and nitrogen cycling mechanism (Dickinson et al. 1998, 2002) has been added to represent two-way interactions between climate and biosphere processes over a wide range of temporal scales from minutes to decades. The DVM combines process-based representations of terrestrial vegetation dynamics and land-atmosphere carbon, nitrogen, and water exchanges in a modular framework. Features include feedbacks through canopy conductance between photosynthesis and transpiration and between these "fast" and other "slow" ecosystem processes, such as tissue turnover, and soil organic matter and litter dynamics.

Most recently, the CLM was extended to include 11 layers with the bottom below 5 meters (m) to contain all water tables. In addition, the CLM was coupled with *a terrestrial hydrology module* (HYD) based on the three-dimensional averaged, localized Richard's equation to better predict runoff due to saturation and infiltration excess, base flow and snow melt, and surface energy fluxes (see Kumar 2004 for the initial concept and full implementation to be documented in an upcoming paper). The HYD predicts the averaged directional lateral flow from the bulk of each side of a grid box at individual soil layers. The prediction, completed in a soil column fully underneath each CWRF dynamic grid, thus makes obsolete the Topmodel concept, which operates at individual basins (Kumar and Chen 1999; Chen and Kumar 2001). The resultant four-directional lateral flows can be integrated readily into a comprehensive routing module to predict streamflow geographic distributions. Along with the DVM, the authors are coupling the most popular agroecosystem models, GOSSYM (Reddy et al. 2002) and DSSAT (Jones et al. 2003), to predict the life cycle of major crops, including cotton, corn, soybeans, wheat, and rice. These advances provide the foundation for modeling water quantity and quality at local, regional, and national scales over the United States.

The WRF release version 2 prescribes surface temperatures of water bodies, including inland lakes and oceans, and fixes them at the initial conditions from the WRFSI. This is one of the factors impeding WRF application in climate studies. In this regard, observed daily sea surface temperature (SST) variations have been incorporated into the CWRF following the approach for the CMM5 (Liang et al. 2004a). The SST data also define sea ice cover changes using a temperature threshold of 271.36 K. This noninteractive approach, however, excludes the observed negative radiative feedback that results from increased (decreased) convective clouds in response to positive (negative) SST anomalies. Feedback relaxes the SST anomalies and has impacts on other processes, including those responsible for the development and maintenance of convection. To facilitate the full interaction, the Geophysical Fluid Dynamics Laboratory Modular Ocean Model or MOM (Griffies et al. 2003) and the Los Alamos Sea Ice (CICE) model (Hunke and Lipscomb 2001; Liang et al. 2004b) are being coupled to predict ocean temperature, salinity, current, and sea ice distributions. The CLM then predicts water temperature profiles for inland shallow and deep lakes (Bonan 1995). The CLM also determines air-sea exchanges of heat, moisture, and momentum over all ocean and sea ice grids consistently for land surfaces.

The input parameter requirement depends on the formulation complexity of the surface modules. The SLAB

needs only the bulk surface moisture availability to calculate evaporation. The RUC and NOAH LSMs require the soil texture category to define static soil properties at each geographic grid; the land vegetation category to specify static canopy properties; and the green vegetation fraction to compute canopy versus bare soil contributions. Although soil texture categories are provided separately for top (0-30 centimeters or cm) and bottom (30-100 cm) layers, only the top values are used to provide constant soil properties in each entire column. The green vegetation fraction distribution is given by monthly means based on the NOAA Advanced Very High Resolution Radiometer (AVHRR) satellite-derived five-year climatology (Gutman and Ignatov 1998), but presently kept fixed at the initial conditions. The NOAH incorporates four additional input parameters: minimum areal coverage of annual green vegetation, background snow-free and maximum snow albedos to calculate the actual surface albedo with weighting of green vegetation and snow cover fractions, and LAI to account for the leaf density effect on canopy resistance. Due to limited data availability, the LAI, however, is presently set to a globally universal value of 4 across all vegetation categories.

On the other hand, the CWRF incorporating the most comprehensive representation of surface processes, when fully developed, includes the CLM, DVM, HYD, CICE, and MOM modules. The CLM needs soil, sand, and clay fraction profiles to specify eight static soil properties in individual layers of each geographic grid; bedrock depth to set the soil bottom level impervious to water; regression parameters to determine local characteristic dependence of surface albedos for direct and diffuse solar radiation at visible (0.3-0.7 micrometers or μm) and near-infrared (0.7-5.0 μm) spectral bands; the vegetation category to define 32 static canopy (morphological, optical, and physiological) properties; fractional vegetation cover to compute canopy versus bare soil contributions; and time-variant vegetation greenness, and leaf and stem area indices to describe canopy dynamic variations. The canopy dynamic variables can be predicted by the DVM, which, in turn, demands more input fields, such as plant phenology, stress thresholds, and crop cultural or management practices. The HYD requires terrain geographic distribution at the finest possible resolution to derive structure information for watersheds, including directional slopes and curvatures of elevation, to produce more realistic water redistribution (e.g., river routing, and surface and subsurface runoff). Without the interactive MOM and CICE, the CWRF entails specification of SST and sea ice distributions. The CICE predicts sea ice variations, but calls for input of sea surface current and salinity distributions. The MOM predicts ocean temperature, salinity, and current at the surface and throughout the water column above the seafloor. For a regional application, however, the MOM necessitates LBCs, especially for water temperature and salinity profiles, throughout the oceanic columns within the CWRF buffer zones. Unless a fully coupled global general circulation model (GCM) is driving the CWRF, these oceanic temperature and salinity profiles typically are supplied by coarse-resolution observational analyses of the monthly mean climatology. To force realistic upper ocean variations and constrain the unavoidable climate draft of the fully coupled system, the MOM dynamically can be relaxed toward the observed SST distributions.

General Considerations for Comprehensive SBCs

Like the WRF, the CWRF is targeted for community use. A comprehensive set of SBCs based on best observational data is desired for CWRF general applications for all effective, dynamically coupled or uncoupled, combinations of the surface modules, as well as for any specific region of the world. There is no universal, complete set of SBCs, however, because different modules may require specification of more or less surface parameters (see Chen et al. 1997 for an intercomparison of various LSM schemes). Some fields are necessary, in general, such as surface elevation. Others no longer require input when an interactive component is coupled to predict them. For example, if the MOM and CICE are fully coupled, input for oceanic conditions is not needed. Data for these fields still could be useful as initialization or relaxation to constrain the coupled system drift away from the observed climatology. Many other fields usually are presumed static and empirically derived from the basic parameters due to lack of global observations. For example, the CLM prescribes numerous canopy properties as static tabular values depending only on the vegetation category, while parameterizing several soil profiles in terms of clay and sand fractions. These parameters are not presented in this study, but rather treated as derivative static fields.

The study considers an SBC that: (a) requires a geographic distribution, static or varying over time; (b) for which global observations are freely available; (c) is a fundamental input field, independent or defining other derivatives; and (d) plays an important role in surface-atmosphere interaction. Table 1 lists all CWRF SBCs that meet these criteria and are documented. They can be distinguished into three time-dependet groups. The first group contains static fields that need to be defined at the first step of an initial or restart run. The only field that may change between sea ice and open water over oceans is SCI. The second group includes SST, LAI, and SAI daily variations. These data, however, are derived from weekly or monthly means changing from year to year. The third group fields (SSS, SSC, STP, and SSP) vary relatively slowly and consist of only the monthly mean climatology due to limited observations. Among the static SBCs, fundamental inputs include surface elevation parameters defining the CWRF geographic boundaries, and soil characteristic fields determining terrestrial water movement (Webb et al. 1993). There exist rich studies for weather and climatic importance of SST (Gao et al. 2003; Thiébaux et al. 2003; Liang et al. 2004a), surface albedo (Charney 1975), and various vegetation parameters (Xue and Shukla 1993; Copeland et al. 1996; Chase et al. 1996; Betts et al. 1997; Hahmann and Dickinson 1997; Pielke et al. 1997; Hoffmann and Jackson 2000; Buermann et al. 2001; Lu and Shuttleworth 2002; Liang et al. 2003).

The list of the SBCs in Table 1 is designed for normal operation of the CWRF surface representation using the CLM and optional HYD, CICE, and MOM modules. The surface elevation fields developed show large differences from the

WRFSI result due to improved data processing techniques. The soil fields (DBED, SAND, CLAY, and SALF) consistently were introduced into the SVAT modeling for the first time ever. The use of a single SCI simplifies the CWRF control for various surface modules, where multiple indices were necessary in the WRF. Incorporation of daily SST variations common to other choices (SLAB, RUC, or NOAH) is the minimal requirement enabling the CWRF for climate applications. The vegetation fields (LCC, FVC, LAI, and SAI) also can be integrated with these choices, but important modifications are required. This would entail replacing the green vegetation fraction in the RUC with the new static FVC to disable canopy variability or replacing the NOAH combination of a variable green vegetation fraction and a universal LAI with the new static FVC and time-variant LAI plus SAI, assuming acceptable switching time dependence between the two fields. One problem arises in calculating surface albedo from prescribed snow-free and maximum snow values weighted by a green vegetation fraction that differs from the FVC. An alternative is to adopt the new parameterization based on the SALF. For convenience, in addition to the FVC, the old green vegetation fraction was retained but the model was improved to incorporate its temporal variation.

It may still be arguable whether the FVC, LAI, or both should carry the information for time variations of terrestrial vegetation phenology. The global distributions at fine spatial and temporal scales can be described only by means of remote sensing, such as the satellite product of Normalized Difference Vegetation Index (NDVI). Sellers et al. (1996) incorporated all NDVI geographic and seasonal variations into LAI distributions. In contrast, Gutman and Ignatov (1998) revealed that the limited information contained in NDVI precludes construction of seasonal variations for both FVC and LAI, and thus derived a time-variant FVC while prescribing a constant LAI. Recently, Zeng et al. (2000) argued that assuming a static FVC and varying LAI is more realistic from modeling perspective and also made feasible by current data availability. As such, FVC is determined by distinct vegetation categories and long-term edaphic and climatic controls, whereas LAI includes all canopy dynamic variability. This study follows the approach of Zeng et al. (2000). The three-dimensional canopy effects are parameterized by the combination of FVC for the fractional area of vegetation covering a model grid (horizontal extent), LAI, and SAI for the abundance of green leaves and stems of the vegetated area (vertical density).

A critical requirement in constructing the SBCs for CWRF use is that each field must be defined globally with no missing values and with physical consistency across all relevant parameters. Missing data, if any, must be appropriately filled. For mesoscale weather and climate modeling, the raw data should be available at the finest possible resolution to facilitate a more realistic representation

Table 1. Primary SBCs Incorporated into the CWRF

SBC	Description	Units	Level	Time	Usage	I/O
HSFC	Surface elevation	m	1	Static	All	ir
HSDV	Surface elevation standard deviation	m	1	Static	CLM	ir
HSLD	Surface elevation slope and deviation	m/m	4	Static	HYD	ir
HCVD	Surface elevation curvature and deviation	1/m	4	Static	HYD	ir
DBED	Bedrock, lakebed, or seafloor depth	m	1	Static	CLM, MOM	ir
SAND	Soil sand fraction profile		nl_soil	Static	CLM	ir
CLAY	Soil clay fraction profile		nl_soil	Static	CLM	ir
SALF	Soil albedo localization factor		4	Static	CLM	ir
TBS	Bottom soil temperature	K	1	Static	CLM	ir
SCI	Surface characteristic identification		1	Static	All	ir
LCC	Land cover category		1	Static	CLM	ir
FVC	Fractional vegetation cover		1	Static	CLM	ir
LAI	Leaf area index	m^2/m^2	1	Daily	CLM	i2r
SAI	Stem area index	m^2/m^2	1	Daily	CLM	i2r
SST	Sea surface temperature	K	1	Daily	All	i1r
SSS	Sea surface salinity	‰	1	Monthly	CICE	i3r
SSC	Sea surface current	m/s	1	Monthly	CICE	i3r
STP	Sea temperature profile	K	nl_ocean	Monthly	MOM	i3r
SSP	Sea salinity profile	‰	nl_ocean	Monthly	MOM	i3r

Note:
I/O indices used in Registry for DATASET=INPUT (i), AUXINPUT? (i?), and RESTART (r).

of surface heterogeneity effects. When the data resolution is sufficiently finer than the RCM grid, the subgrid effects can be incorporated further using composite, mosaic, or statistical-dynamical approaches (Avissar and Pielke 1989; Koster and Suarez 1992; Dickinson et al. 1993; Giorgi 1997; Leung and Ghan 1998). Although many raw datasets collected in this study have adequate resolutions (as fine as 1 km) to account for the subgrid effects, only the SBCs on a given RCM grid with a single dominant surface type are assumed. Note that a global database developed by Masson et al. (2003) at 1-km resolution for several land surface parameters is still insufficient for the CWRF applications.

Existing observational databases have various resolutions, finer or coarser than the RCM grid, a wide range of map projections and data formats, and often contain missing values or inconsistencies between variables. This presents significant challenges and requires labor-intensive efforts to process the data onto the RCM-specific grid mesh and input data format. Horizontal data remapping uses Geographic Information System (GIS) software application tools, Arc/Info and Arc/Map, from Environmental Systems Research Institute, Inc. In particular, the GIS tools are used to determine the geographic conversion information from a specific map projection of raw data to the identical RCM grid system. The information includes location indices, geometric distances, or fractional areas of all input cells contributing to each RCM grid. It then can be applied to remap all variables of the same projection. Remapping is completed by a bilinear interpolation method in terms of the geometric distances if the raw data resolution is low or otherwise a mass conservative approach as weighted by the fractional areas. For a categorical field such as SCI and LCC, the dominant field that occupies the largest fraction of the grid is chosen after calculating the total fractional area of each distinct surface category contributing to a given RCM grid.

In the vertical, soil profiles at each gridpoint are processed onto the same RCM standard soil layers. The CLM soil layers (Table 2) were used to construct SAND and CLAY profiles. Note that the CLM defines a soil layer at the node depth z_n (m):

$$z_n = \frac{1}{400}(e^{\frac{n}{2} - \frac{1}{4}} - 1), \quad n = 1, 2, ..., N \qquad (1)$$

where N is the total number of soil layers (i.e., nl_soil in Table 1). The interface is located between two layers (Dai et al. 2001). As such, the soil layers are thinner near the top of the profile to better resolve details of near-surface processes and have progressively larger thicknesses as an exponential function of the depth. Because the water tables are observed below a 4-m depth in some U.S. regions (Chen and Kumar 2001), N was increased from 10 to 11 to extend the model's soil bottom to 5.68 m. The number (nl_ocean in Table 1) and bounds of vertical layers for oceanic fields are set to be the same as the raw data. The authors anticipate that a smaller number of layers can provide a reasonable representation of upper ocean interaction with the atmosphere. This will be changed after completion of the MOM coupling.

The SBCs' data quality, value representation, and visual display largely depend on the RCM computational domain and grid resolution. The CWRF incorporates an optimal buffer zone treatment that integrates realistic energy and mass fluxes across the boundaries between the GCM and RCM domains, following that used for the CMM5 (Liang et al. 2001). This treatment consists of an advanced buffer zone positioning and LBC assimilation technique. A fundamental requirement is correct representation of key physical processes near buffer zones that govern the GCM resolvable circulation in the RCM domain interior.

Table 2. CLM Soil Layer Depth and Thickness (cm)

Layer n	Node depth (z_n)	Thickness (Δz_n)	Interface depth ($z_n^{/}$)
1	0.71	1.75	1.75
2	2.79	2.76	4.51
3	6.23	4.55	9.06
4	11.89	7.50	16.55
5	21.22	12.36	28.91
6	36.61	20.38	49.29
7	61.98	33.60	82.89
8	103.80	55.39	138.28
9	172.76	91.33	229.61
10	286.46	150.58	380.19
11	473.92	187.45	567.64

As such, the buffer zones are positioned over areas of small observational uncertainties or GCM biases but high GCM climate predictability. The assimilation technique incorporates improved dynamic relaxation coefficients within widened buffer zones such that robust GCM signals are integrated realistically into the RCM domain while LBC errors are effectively absorbed. For U.S. climate applications, the domain is centered (37.5ºN, 95.5ºW) using the Lambert Conformal Conic map projection and 30-km horizontal grid spacing, with total gridpoints of 196 (west-east) $\times$ 139 (south-north). The domain covers the continental United States and represents the regional climate that results from interactions between the planetary circulation (as forced by the LBCs) and North American surface processes, including orography, vegetation, soil, and coastal oceans. Buffer zones are located across 14 gridpoints along all four domain edges, where LBCs are specified throughout the entire integration period using a dynamic relaxation technique (see Eqs. [1-2] in Liang et al. 2001 with $L = 15 - 1$). This configuration enables the CMM5 to produce skillful simulations of temporal and spatial variations of precipitation over North America during 1982-2002 (Liang et al. 2001, 2004a). In this study, all SBCs constructed and displayed on this CWRF domain are suitable for U.S. applications.

Construction of the CWRF SBCs

For convenience, the geographic location of a point is hereafter referred as a "pixel" for raw data and a "grid" for the CWRF result. A given pixel or grid value represents the area surrounding the point as defined by its respective horizontal spacing. The following section elaborates in detail on raw data sources and processing procedures used to construct any specific CWRF domain over the globe. Most procedures use ArcInfo commands. In particular, IMAGEGRID and GRIDPOLY *convert* input data from the image to the ArcGIS raster grid and to the polygon coverage formats, respectively; PROJECT *remaps* the raw input data onto the CWRF grid projection; UNION and CLIP geometrically *intersect* polygon features of input data with the CWRF grid mesh and *extract* the fractional area of each pixel contributing to the grid; GRID DOCELL and IF statements conditionally *merge, replace,* or *adjust* different input datasets for an improved product. On the other hand, some raw input data are available only at coarse resolutions (e.g., 1°). They are *interpolated* into the CWRF grid using a bilinear approximation in longitude and latitude. These data, especially those for land or ocean only, also may contain gaps along coastal regions and over islands that are resolved by the CWRF. They are *extrapolated* with adjacent values using bilinear approximation. Even relatively finer resolution (1-km and 8-km) input data such as soil fraction and LAI have missing value pixels *filled* by the average over the nearby data pixels having the same land cover category within a certain radius around a missing point. The number of pixels and the range of radius used for filling depend on the resolution of the raw input data.

For brevity, italic commands appear throughout this section without repeating their corresponding ArcInfo commands or interpolation/extrapolation procedures.

[A] Surface Characteristic Identification (SCI)

The CWRF incorporates the SCI to distinguish broad surface categories that invoke distinct surface modules. Currently, SCI consists of eight categories: *urban and built-up, soil, wetland, glacier, shallow lake, deep lake, sea ice*, and *ocean*. The first four categories are dealt with differently in CLM calculations of soil properties, while the remaining categories are treated individually for specific processes in water bodies. The current SCI can be modified or expanded in the future as needed. For example, *ocean* can be further divided into *coastal ocean* and *deep ocean*.

Figure 1 depicts the SCI distribution over the CWRF domain, overlaid with the level-1 basin boundaries and main stream network (see [B] for details). The SCI follows the LCC to identify an *urban and built-up, soil, wetland, glacier,* or *ocean* grid, while defining a *shallow lake* or *deep lake* grid if the inland water body depth (DBED) is less or greater than 10 m. Note that no *glacier* grid exists in the current CWRF domain. Nor shown is *sea ice*, which is not static but specified by the changing SST or predicted by the CICE. The *sea ice* and *ocean* grids are interchangeable at their edges, where sea ice can form or melt out completely depending on whether the SST is cooler or warmer than 271.35 K. They may coexist in a grid, with their partitions predicted by the interactive CICE.

[B] Surface Elevation and Derivatives (HSFC, HSDV, HSLD, and HCVD)

These fields are constructed from the U.S. Geological Survey (USGS) HYDRO1k Digital Elevation Model (DEM) with a 1-km nominal cell size (http://edcdaac.usgs.gov/gtopo30/hydro/) based on the Global 30-arc-second elevation dataset (GTOPO30). The DEM data is available in the band interleaved by line (BIL) image format on the Lambert azimuthal equal area projection. The raw data are *converted* into ArcGIS raster grid format and then *remapped* onto the CWRF projection. Subsequently, the ArcInfo/GRID commands, ZONEALMEAN and ZONALSTD, are used to calculate the mean and standard deviation of the elevations within each CWRF grid. The mean and deviation from the centroid of each grid are picked up by ArcInfo's Arc Macro Language (AML) program Gridspot70 for HSFC and HSDV. Figure 2 illustrates the geographic distributions of HSFC, HSDV, and HSFC difference from the WRFSI product. Note that the SI mean elevation constrains the arithmetic averaging to no more than 10 $\times$ 10 raw data pixels surrounding each grid (John Smart, NOAA, personal communication, 2004), while the CWRF mean is the area-weighted averaging over all pixels within the grid. The local HSFC differences between the CWRF and SI can be large, up to ±1 km, especially over mountainous regions.

The HYD requires additional surface elevation derivatives, including slopes and curvatures, and their deviations along both x and y directions (HSLD and HCVD) to consider the topographic effect on soil moisture transport. These fields also can be constructed from the same DEM data. A more

accurate representation can be achieved by using the most recent Shuttle Radar Topography Mission (SRTM) elevation data, available at a 90-m (3 arc-second) resolution worldwide between 60ºS and 60ºN, and at a 30-m (1 arc-second) resolution for the U.S. domain south of 60ºN (ftp://edcftp.cr.usgs. gov /pub/data/srtm/).

Stream network lines and drainage basin boundaries of individual watersheds must be specified for the HYD to determine the source and sink terms closing the Richard's equation. Vector streamlines and the derived basin boundaries along with the flow directions also are provided in the HYDRO1k data (http://edcdaac.usgs.gov/ gtopo30/hydro/namerica.asp). Upstream watershed contributing areas greater than 1000 km^2 are selected and processed using the ArcInfo STREAMLINK function. Drainage basins are consecutively divided from the coarse to fine levels using procedures first articulated by Otto Pfafstetter (Verdin 2003). For a given level, each polygon in a basin has a unique Pfafstetter code depicting a sub-basin. The highest level (i.e., the most comprehensive structure representation) currently available is 6, which indicates identification of a total of 3612 basins in the CWRF domain. Interceptions of stream network lines and drainage basin boundaries with the CWRF grids are presently stored for the HYD use. Figure 1 depicts the level-1 basin boundaries and main stream network.
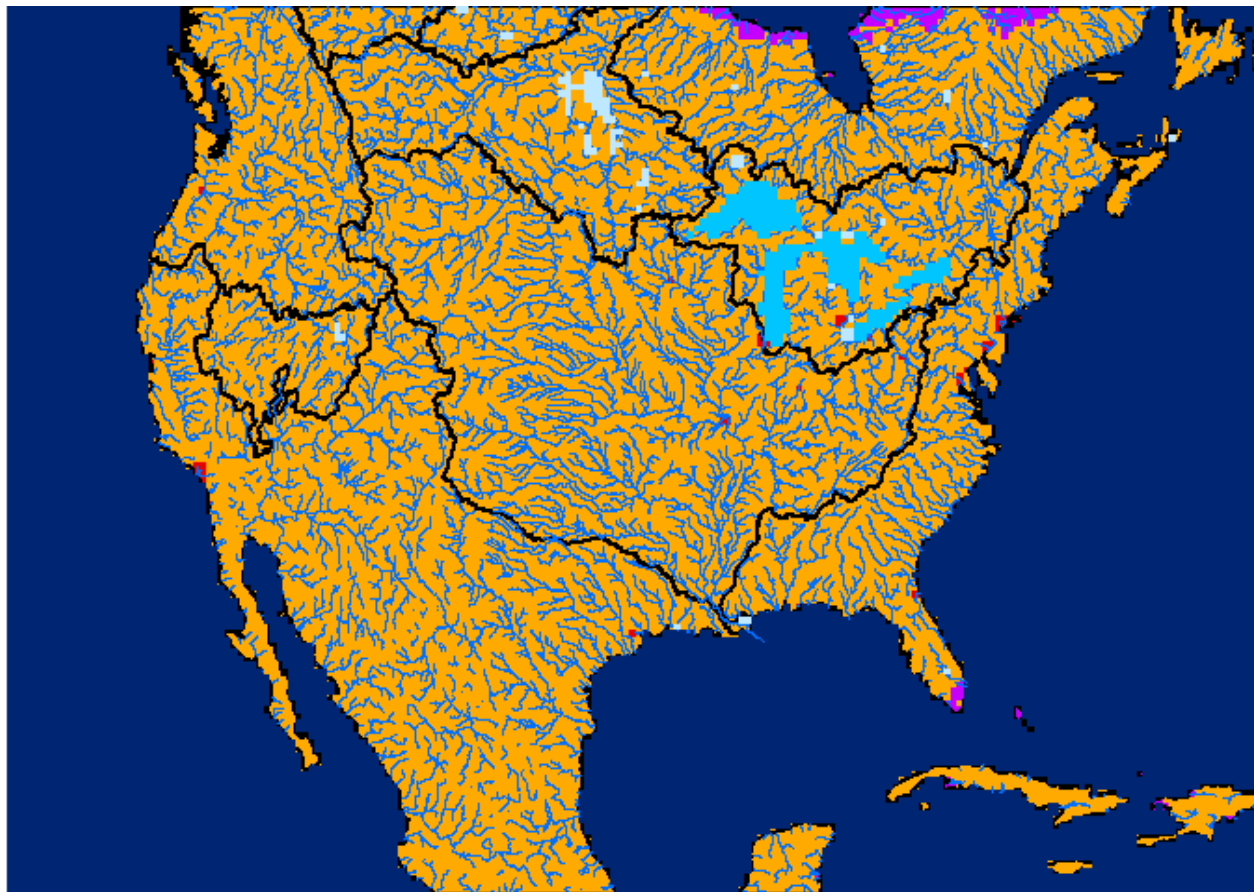

Figure 1. Geographic distribution of SCI with overlays of basin boundaries and main stream network.
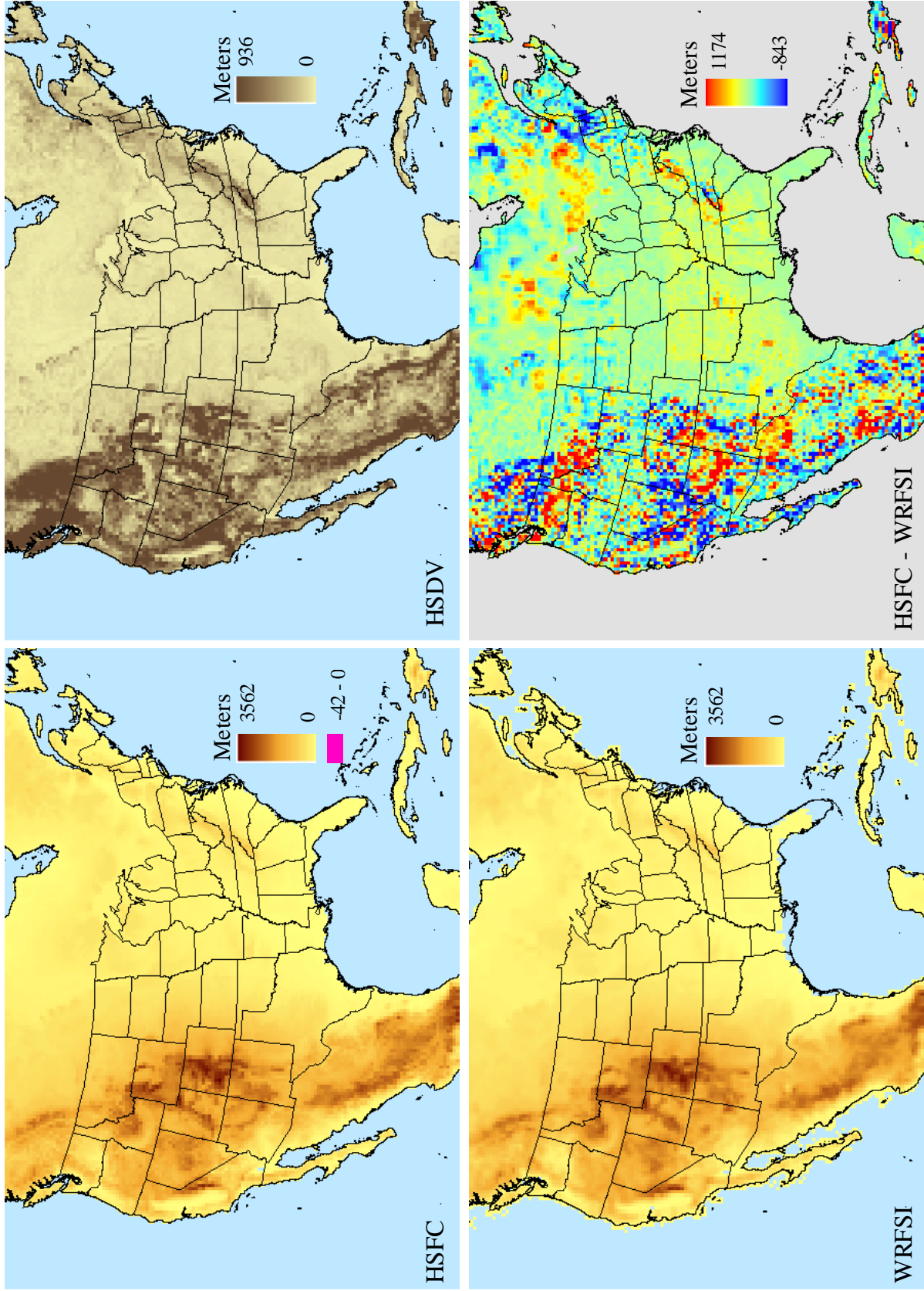
Figure 2. Geographic distributions of HSFC, HSDV, and HSFC difference from the WRFSI product.

[C] Bedrock, Lakebed, or Seafloor Depth (DBED)

The DBED includes bedrock depth of the land, lakebed depth of the lakes, and seafloor depth of the oceans. The CLM uses bedrock depth to determine thermal and hydraulic properties in terms of the SAND and CLAY profiles (see [D]) for the soil layers above the bedrock. The CLM also predicts separate water temperature profiles for shallow and deep lakes, as distinguished by the SCI via lakebed depth (see [A]). The MOM requires specification of the seafloor depth to define the lower boundary of the water circulation. Eventually, the MOM will be integrated with the HYD and a comprehensive routing model to predict the water level (HSFC) of major inland water bodies, including the Mississippi River and the Great Lakes. The DBED defines the bottom of all surface modules impermeable to water over the entire CWRF domain.

The bedrock depth is based on the combination of the Continental U.S. Multi-Layer Soil Characteristics (CONUS-SOIL) Dataset and FAO-UNESCO Soil Map of the World. The FAO-UNESCO data (FAO 1996) include the geographic distribution at 5-minute longitude-latitude spacing of nearly 5000 mapping units over the globe (also available from http://www.lib.berkeley.edu/EART/fao.html). Each mapping unit contains a maximum of eight soil units among the 106 categories of the FAO 1974 soil classification system (FAO, 1974). Each soil unit corresponds to one of five bedrock depth classes: 0-10, 10-50, 50-100, 100-150, and 150-300 cm (Curt Reynolds, USDA Foreign Agricultural Service, personal communication, 2003). For each mapping unit, all soil units are assigned with respective depth of their upper bounds (i.e., 10, 50, 100, 150, or 300 cm) and then integrated with their occurrence rates to estimate the mean bedrock depth. A global soil bedrock depth distribution at a 5-minute resolution then is constructed.

The CONUS-SOIL, developed at Penn State University (Miller and White 1998) from the USDA State Soil Geographic Database (STATSGO), has a finer resolution over the conterminous 48 states with 1-km spacing on the geographic coordinate system (http://www.essc.psu.edu/soil_info/index.cgi?soil_data&conus&data_cov). About one third of the data pixels were coded as 152 cm, which generally indicates the maximum depth of soil data from an area where bedrock was not encountered. Most CONUS-SOIL regions with bedrock deeper than 152 cm are overlaid with certain FAO-UNESCO mapping units having soil depths of 150 cm. Given large uncertainties involved in these estimates, a uniform bedrock depth of 600 cm (deeper than the bottom of the last CLM soil layer) is assigned to all CONUS-SOIL pixels with values of 152 cm and the corresponding FAO-UNESCO mapping units. The FAO-UNESCO data are then *replaced* by CONUS-SOIL data to better resolve the U.S. bedrock distribution. These hybrid data are *adjusted* further to be confined by the USGS land cover classification (see [F]) for a consistent representation of water bodies.

Lakebed depth is calculated by subtracting the lake topographic data from the DEM surface elevations (see [B]), consistent with the long-term mean water levels of each lake (http://www.glerl.noaa.gov/data/now/wlevels/). Currently, only the Great Lakes topographic data are available at 2.56-km spacing on the Mercator projection from the NOAA Great Lakes Environmental Research Laboratory (http://www.glerl.noaa.gov/data/bathy/bathy.html). For all other lakes within the CWRF domain, including the Great Salt Lake in Utah, Lake Okeechobee in Florida, and lakes in Canada, no digital data were available so it was assumed that these water bodies were 10 m deep (*shallow lake*).
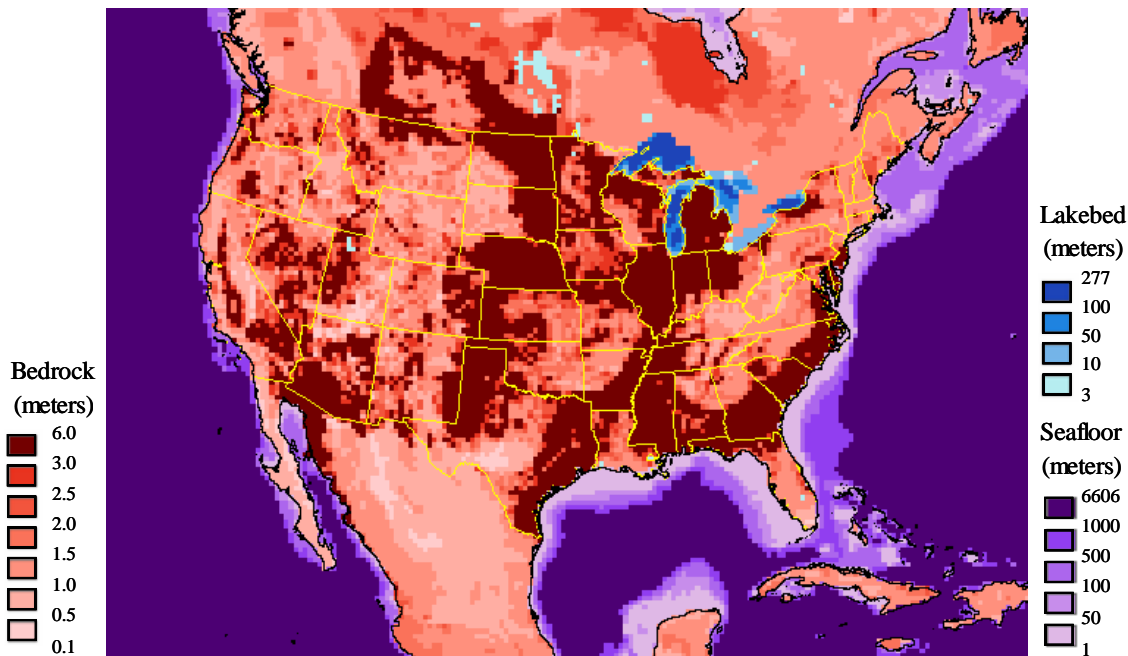

Figure 3. Geographic distribution of DBED.

Seafloor depth is based on the global 2-minute bathymetry data (ETOPO2) available from the National Geophysical Data Center (http://www.ngdc.noaa.gov/mgg/fliers/01mgg04.html). Between 64°N and 72°S, it is derived from satellite altimetry observations in combination with shipboard echo-sounding measurements, version 8.2 (Smith and Sandwell 1997). South of 72°S, the U.S. Naval Oceanographic Office's Digital Bathymetric Data Base Variable Resolution, version 4.1, is used. North of 64°N, the new International Bathymetric Chart of the Arctic Ocean, version 1 (Jakobsson et al. 2001), is used.

For each of the above three datasets with various spatial resolutions, the pre-processed depth distribution is *converted* from the raster grid into the polygon coverage, *remapped* onto the CWRF projection, and later *intersected* with the CWRF grid mesh. The fractional area of each pixel contributing to the grid is *extracted*. The final depth is obtained by the area-weighted averaging of all pixels within each CWRF grid. The resulting three depth distributions are *merged* conditionally with SCI guidance. The depth of the dominant SCI category (soil, lakes, or ocean) within each CWRF grid is chosen to represent the DBED for that grid. Figure 3 depicts the merged DBED geographic distribution over the CWRF domain.

[D] Soil Sand and Clay Fraction Profiles (SAND and CLAY)

The CLM requires SAND and CLAY profiles to parameterize soil thermal and hydraulic properties (Dai et al. 2003) following Cosby et al. (1984). The WRFSI currently provides global 1-km distribution of 16 soil texture categories for top (0-30 cm) and bottom (30-100 cm) layers (http://www.rap.ucar.edu/projects/land/LSM/). In addition, Webb et al. (1993) produced global 1° distributions of sand and clay for the two layers by combining the FAO-UNESCO Soil Map of the World with the World Soil Data File (Zobler, 1986). Both datasets are insufficient to define the required SAND and CLAY profiles over all 11 CLM layers. Consist with the bedrock depth (see [C]), these profiles are determined by a combination of the CONUS-SOIL and FAO-UNESCO data.

Reynolds et al. (2000) reproduced the FAO-UNESCO global 5-minute distributions of sand and clay fractions for 0-30 and 30-100 cm (http://hydrolab.arsusda.gov/soils/start.htm). Top layer data are assigned uniformly for the upper five CLM layers above 28.91 cm, while bottom layer values are used for the remaining layers. Over the conterminous 48 states, they are *replaced* by the CONUS-SOIL 1-km distributions of sand and clay fractions at 11 standard layers, divided at 5, 10, 20, 30, 40, 60, 80, 100, 150, 200, and 250 cm (Miller and White 1998). As previously discussed, raw data below the bedrock depth of 152 cm likely were not measured, and thus those for standard layers 10 and 11 (150-250 cm) are discarded. Data in the top 9 standard layers are interpolated with thickness-weighting to the upper 8 CLM layers above 138.28 cm, while those of the standard layer 9 are extended uniformly down to the remaining CLM layers. In addition, the CONUS-SOIL contains points with other soil texture classifications without providing sand and clay fractions. Each missing point is *filled* by averaging over all nearby data pixels having the same USGS land cover

classification (see [F]) within a certain radius starting at 10 km (440 pixels) around the point and increasing until a minimum of 50 data pixels are obtained.

As with the bedrock depth, the pre-processed hybrid soil fraction data are *adjusted* to be confined by the USGS land cover classification (see [F]) to provide consistent representation of water bodies. The resultant sand and clay fractions in each of the 11 CLM layers at the varying horizontal resolutions of the raw data (1 km or 5 minutes) are *converted* from the raster grid into the polygon coverage, *remapped* onto the CWRF projection, and then *intersected* with the CWRF grid mesh. The fractional area of each pixel contributing to the grid is *extracted*. Final SAND and CLAY profiles are obtained by area-weighted averaging of all pixels within each CWRF grid. When bedrock occurs within a CLM layer, averaging applies an additional thickness weight for the portion of the layer above the bedrock depth. Note that there exist some regions with soil texture classified as organic material for which neither sand nor clay data are given. These regions are mainly in Florida, Minnesota, and several western states. Corresponding sand and clay fractions are assigned a negative unit as an indicator of organic material properties in the CLM. Figure 4 shows the geographic distributions of SAND and CLAY profiles for CLM layers 1 and 8 in the CWRF domain.

[E] Bottom Soil Temperature (TBS)

To specify lower boundary or initial conditions, some LSMs (e.g., SLAB and NOAH) require the soil temperature for the bottom layer. A similar need may arise for lakebed or seafloor temperature when coupling the interactive HYD or MOM. These three temperatures are the TBS. Unfortunately, there is no global observation for this field. As a proxy, TBS is defined by combining the annual mean climatology of surface air temperature over land, SST over lakes (see [J]), and STP at the seafloor over oceans (see [L]). The 1971-2000 land air temperature data (http://www.cru.uea.ac.uk) are available at 0.5° longitude-latitude spacing (New et al. 2002). Given the coarse resolution, these proxy data are *extrapolated* beyond the coastal boundaries and then *interpolated* into the CWRF grid. They are conditionally *merged* to be confined by the LCC for the final TBS. Figure 5 depicts the geographic distribution of TBS over the CWRF domain.

[F] Land Cover Category (LCC)

The CWRF uses the 24-category USGS land cover classification (Table 3) developed from the April 1992-March 1993 AVHRR satellite-derived NDVI composites. The raw data are available at 1-km spacing on the geographic coordinate system in BIL image format (http://edcdaac.usgs.gov/glcc/globe_int.html), *converted* into the ArcGIS raster grid and polygon coverage, and *remapped* onto the CWRF projection. The fractional area of each pixel contributing to the grid is *extracted* after the result is *intersected* with the CWRF grid mesh. The contributing area for each of the 24 LCCs is summed over all pixels of the same category within each CWRF grid. The category contributing the largest area is chosen as the LCC for the grid. When the fractional area of water bodies (*shallow* or *deep lakes*, *sea ice* or *ocean*) is less than 0.5 but dominates the grid, the category chosen is the one contributing the second largest area.
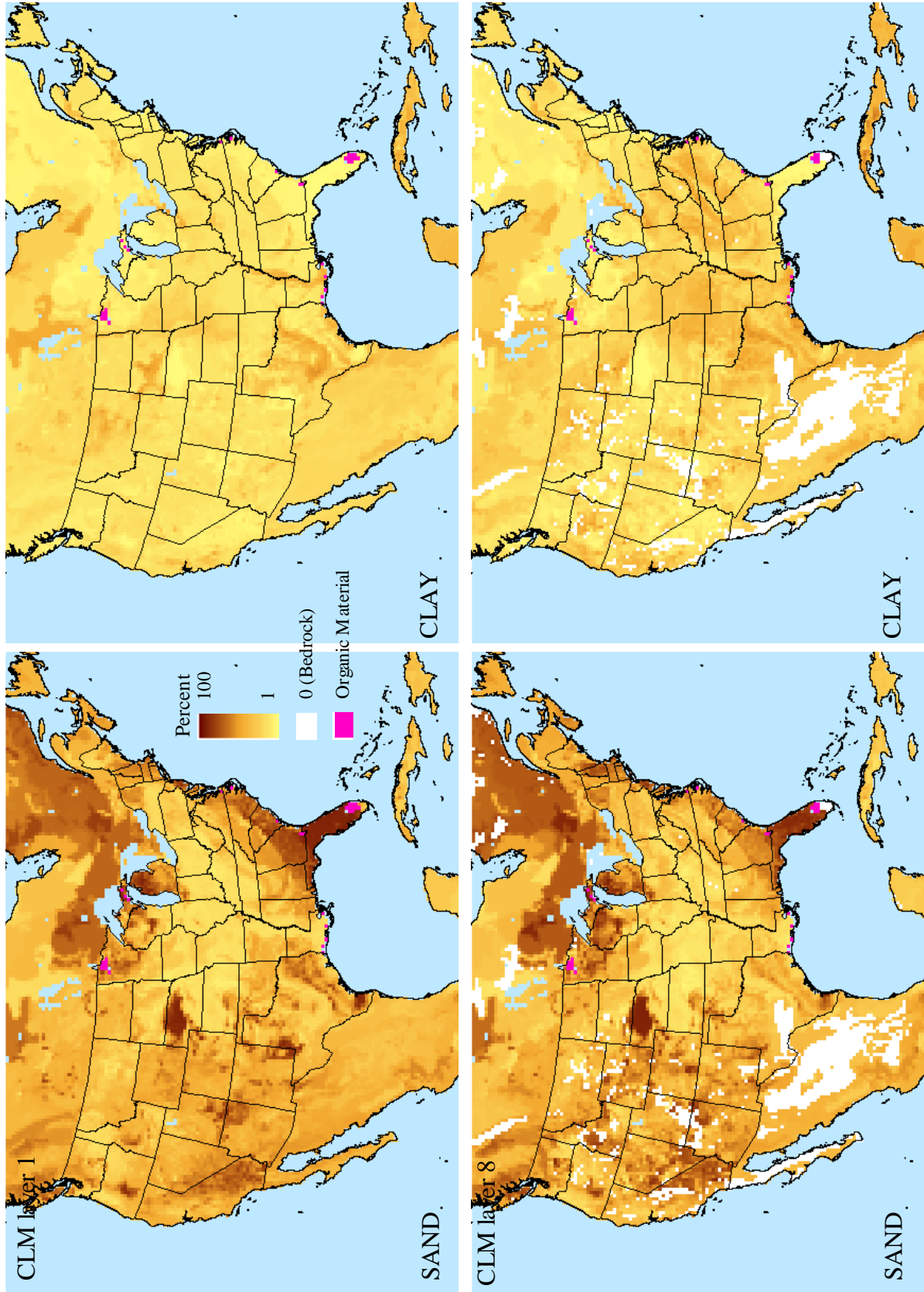
Figure 4. Geographic distributions of percent SAND and CLAY for top (layer 1: 0-1.75 cm) and bottom (layer 8: 82.89-138.28 cm) CLM layers.
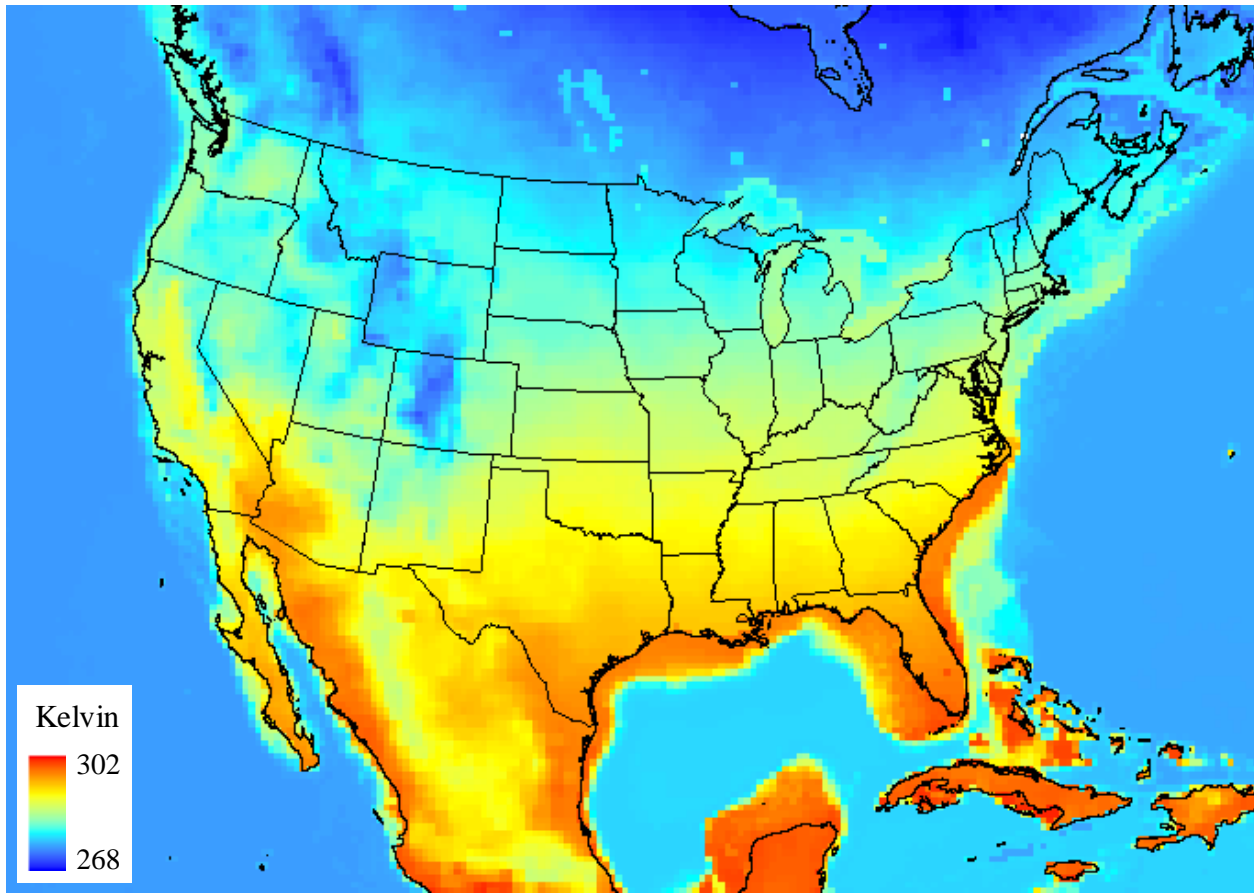
Figure 5. Geographic distribution of TBS.

Table 3. Comparison of $N_{c,v}$ between USGS and IGBP Land Cover Legends

USGS						IGBP				
Type	Description	$N_{c,v}$		Distribution ratio (%)		Type	Description	$N_{c,v}$	Contributing ratio for corresponding USGS Legend (%)	
		RCM	Global	RCM	Global				RCM	Global
1	Urban and Built-Up Land	0.62	0.62	0.34	0.12	13	Urban and Built-Up	0.62	100	100
2	Dryland,Cropland,and Pasture	0.61	0.61	5.49	5.64	12	Croplands	0.61	100	100
3	Irrigated Cropland and Pasture	0.61	0.61	0.48	1.52	12	Croplands	0.61	100	94.41
						14	Cropland/Natural Vegetation Mosaic	0.65	0	5.59
4	Mixed Dryland/Irrigated Cropland and Pasture**									
5	Cropland/Grassland Mosaic	0.65	0.65	4.42	2.05	12	Croplands	0.61	0	1.33
						14	Cropland/Natural Vegetation Mosaic	0.65	100	98.67
6	Cropland/Woodland Mosaic	0.65	0.65	2.26	3.27	14	Cropland/Natural Vegetation Mosaic	0.65	100	100
7	Grassland	0.49	0.49	6.10	4.82	10	Grasslands	0.49	100	100
8	Shrubland	0.60	0.60	8.23	7.23	6	Closed Shrublands	0.60	6.67	14.81
						7	Open Shrublands	0.60	79.52	81.00
						8	Woody Savannas	0.62	13.81	4.19
9	Mixed Shrubland/Grassland	0.60	0.59	0.11	1.02	6	Closed Shrublands	0.60	100	15.28
						7	Open Shrublands	0.60	0	76.53
						10	Grasslands	0.49	0	8.19
10	Savanna	0.61	0.60	0.99	7.13	8	Woody Savannas	0.62	86.88	50.48
						9	Savanna	0.58	13.12	49.52
11	Deciduous Broadleaf Forest	0.69	0.70	5.46	2.55	2	Evergreen Broadleaf Forest	0.69	0	20.41
						4	Deciduous Broadleaf Forest	0.70	68.78	68.26
						5	Mixed Forest	0.68	31.22	11.33
12	Deciduous Needleleaf Forest*		0.63	0.00	0.91	4	Deciduous Needleleaf Forest	0.63		100
13	Evergreen Broadleaf Forest	0.69	0.69	0.08	5.75	2	Evergreen Broadleaf Forest	0.69	100	100
14	Evergreen Needleleaf Forest	0.63	0.63	10.32	2.26	1	Evergreen Needleleaf Forest	0.63	100	100
15	Mixed Forest	0.68	0.68	7.59	3.59	5	Mixed Forest	0.68	100	100
16	Water Bodies			43.09	38.90	17	Water Bodies		100	100
17	Herbaceous Wetland*		0.56	0.00	0.03	11	Permanent Wetlands	0.56		100
18	Wooded Wetland	0.56	0.56	0.84	0.43	11	Permanent Wetlands	0.56	100	100
19	Barren or Sparsely Vegetated	0.60	0.60	0.46	7.56	16	Barren or Sparsely Vegetated	0.60	100	100
20	Herbaceous Tundra**									
21	Wooded Tundra	0.60	0.61	3.35	2.99	7	Open Shrublands	0.60	100	68.11
						8	Woody Savannas	0.62	0	31.89
22	Mixed Tundra	0.60	0.60	0.37	0.97	16	Barren or Sparsely Vegetated	0.60	100	100
23	Bare Ground Tundra*		0.60	0.00	0.02	16	Barren or Sparsely Vegetated	0.60		100
24	Snow or Ice			0.03	1.23	15	Snow and Ice		100	100

Notes:
*Land cover type does not exist in the CWRF domain.
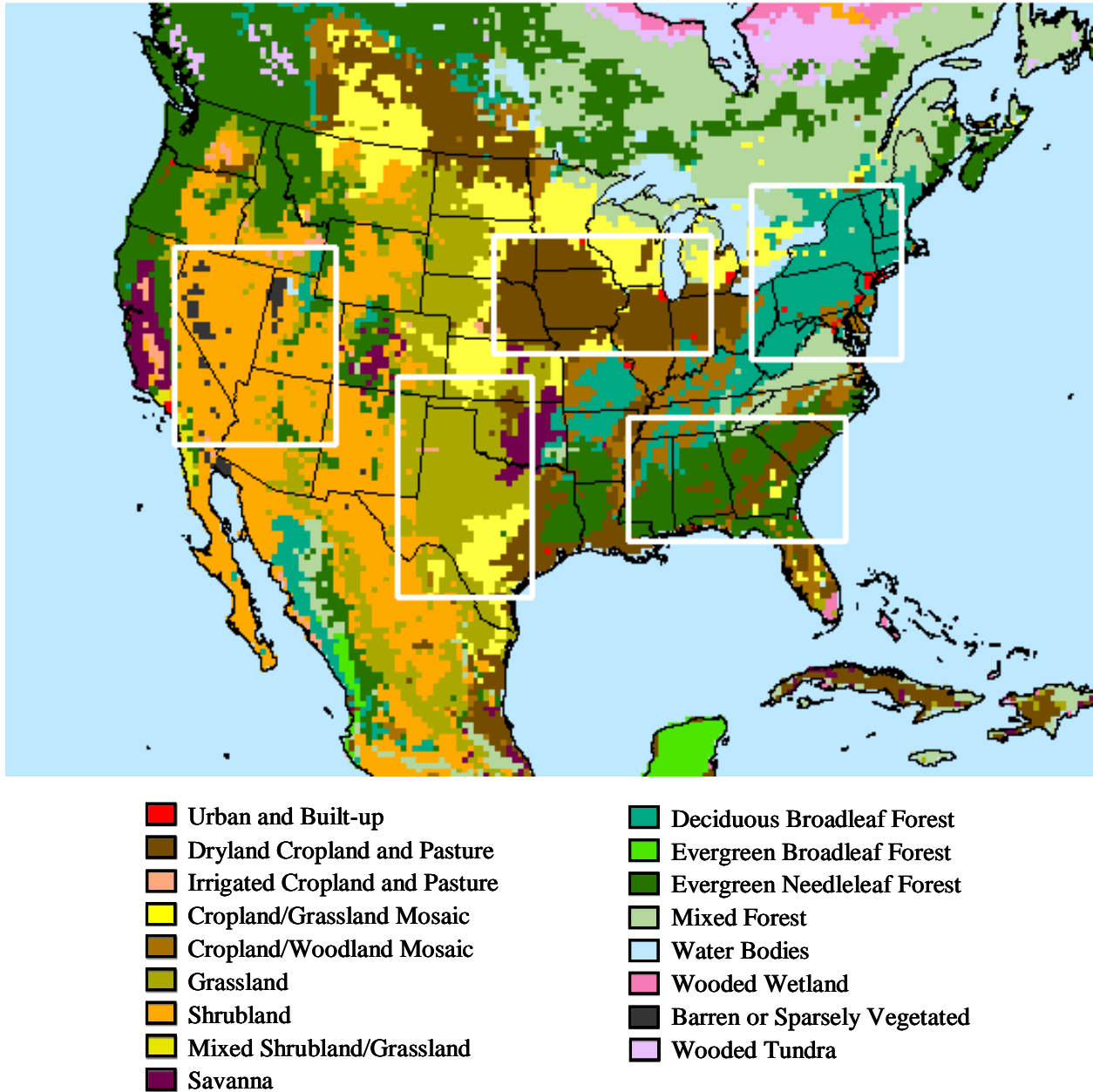**Land cover type does not exist in the global dataset.

Figure 6. Geographic distribution of 17 LCC categories occurred over the CWRF domain and 5 U.S. key regions of interest, each with a predominant category: Texas (grassland), Southwest (shrubland), Midwest (dryland cropland and pasture), Southeast (evergreen needleleaf forest), and Northeast (deciduous broadleaf forest).

Figure 6 illustrates the LCC geographic distribution over the CWRF domain. Note that the USGS raw data do not contain LCCs 4 and 20 over the globe, and additionally LCCs 12, 17, and 23 within the present CWRF domain. Moreover, LCCs 22 and 24 are not LCC majority categories. Therefore, the final LCC includes only 17 LCCs over the CWRF domain.

[G] Fractional Vegetation Cover (FVC)

The FVC is one ecological parameter that determines contribution partitioning between bare soil and vegetation for surface evapotranspiration, photosynthesis, albedo, and other fluxes crucial to land-atmosphere interactions. It is assumed to be time-invariant or static, and derived following Zeng et al. (2000, 2002), from the same global 1-km AVHRR satellite product as for LCC (see [F]). The 10-day April 1992-March

1993 composites were used to determine the annual maximum NDVI $\left(N_{p,max}\right)$ for each LCC, minimizing the effect of cloud contamination on data quality. For each pixel, the vegetation cover is computed by:

$$C_v = \frac{N_{p,max} - N_s}{N_{c,v} - N_s} \qquad (2)$$

where $N_{c,v}$ is the NDVI value for a complete coverage of a specific USGS LCC over the pixel and N_s for bare soil. Using a commercial imagery database, Zeng et al. (2000) determined $N_{c,v}$ by examining percentiles of the $N_{p,max}$ histogram for each LCC of the International Geosphere Biosphere Programme (IGBP) classification (Belward 1996; Loveland et al. 2000). To avoid redundant data processing, the $N_{c,v}$ values for the 24 USGS LCCs (Table 3) are calculated from those of the 17 IGBP categories by *intersecting* the USGS and IGBP land cover maps to determine the fractional areas of individual IGBP categories contributing to each USGS category. The final $N_{c,v}$ is the average of all contributing IGBP values weighted by their corresponding fractional areas. Corresponding $N_{c,v}$ values and contributing areas for the USGS and IGBP categories are listed in Table 3. In addition, the lower percentiles of the $N_{p,max}$ histograms for most categories that define N_s occur mainly in winter and have larger uncertainties (than in summer) due to cloud contamination and atmospheric effects. After Zeng et al. (2000), a uniform value of 0.05 is assigned to N_s for all USGS LCCs.

Note that there exist significant differences between the NDVI from the AVHRR and the most recent Moderate Resolution Imaging Spectroradiometer (MODIS) sensors (http://edcimswww.cr.usgs.gov/pub/imswelcome/). Gallo et al. (2004) compared the concurrent 16-day composite data during 2001 and found a linear relationship between the two methods. The regression intercept and slope values change with LCCs, but all are significantly different than 0 and 1, respectively. The MODIS has generally larger values than the AVHRR, causing Equation (2) to produce greater C_v values. On the other hand, Zeng et al. (2002, 2003) demonstrated that, using the same method with Equation (2), the C_v derived from 8-km AVHRR NDVI during 1982-2000 (James and Kalluri 1994) is consistent with that derived from the 1-km data for April 1992-March 1993. Given the good agreement with field surveys and observational studies and the small interannual variability over areas expecting small anthropogenic impacts, the FVC derived from the AVHRR NDVI was believed to be robust.

The MODIS C_v is therefore scaled toward the AVHRR. For each USGS LCC, a scaling factor $f_{p,v}$ is defined to remove the systematic MODIS difference from the AVHRR in $N_{p,max}$ averaged over all pixels. Assuming the same N_s and multiplying $N_{p,max}$ by $f_{p,v}$ in Equation (2), the corresponding $N_{c,v}$ is estimated to minimize the C_v difference between MODIS and AVHRR. Table 4 lists the resultant $f_{p,v}$ and $N_{c,v}$ values, as well as the correlation coefficients and root mean square (RMS) differences between the C_v based on the AVHRR and MODIS after scaling. The $f_{p,v}$ ranges from 0.50 to 0.81, while the $N_{c,v}$ remains close to the respective AVHRR value except for category 19. The correlations are generally excellent and above 0.5, except for categories 18 and 22 (~0.4) and quite low for categories 6 and 19 (~0.3). Nonetheless, the RMS differences are small for all categories.

The resultant C_v point data at 1-km spacing are *converted* to polygon coverage data, *remapped* onto the CWRF projection, and *intersected* with the CWRF grid mesh. The fractional area of each pixel contributing to the grid is *extracted*. The final FVC is obtained by the area-weighted averaging of C_v values for all pixels within each CWRF grid. Figure 7 compares the FVC geographic distributions derived from the AVHRR and scaled MODIS data over the CWRF domain.
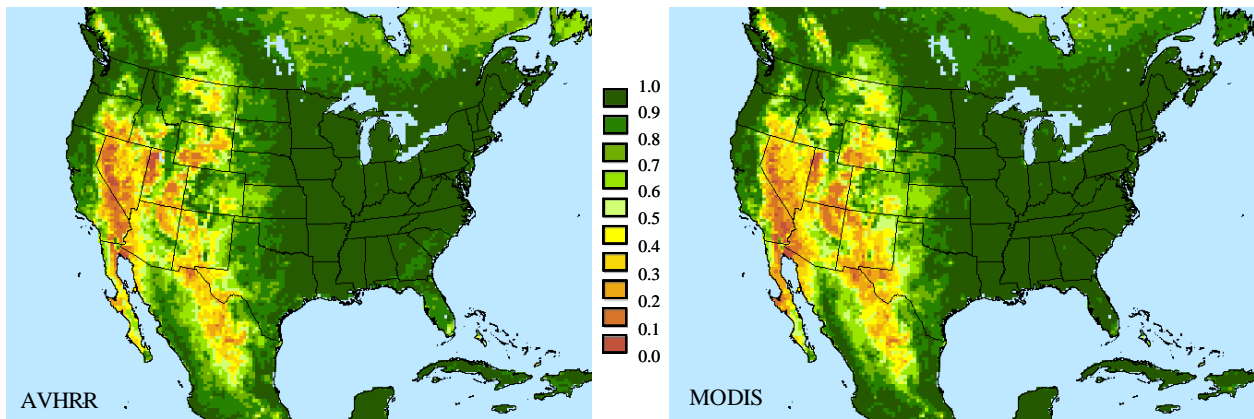

Figure 7. Geographic distributions of FVC derived from the April 1992-March 1993 AVHRR (left) and scaled January 2000-December 2003 MODIS (right) NDVI data.

Table 4. Estimated $f_{p,v}$ and $N_{c,v}$ for C_v Based on MODIS NDVI (2000-2003)

Type	USGS land cover legend	$f_{p,v}$	$N_{c,v}$	Correlation	RMS
1	Urban and Built-Up Land	0.77	0.63	0.81	0.14
2	Dryland Cropland and Pasture	0.78	0.61	0.63	0.12
3	Irrigated Cropland and Pasture	0.81	0.62	0.53	0.16
4	Mixed Dryland/Irrigated Cropland and Pasture**				
5	Cropland/Grassland Mosaic	0.78	0.65	0.80	0.11
6	Cropland/Woodland Mosaic	0.78	0.65	0.26	0.08
7	Grassland	0.77	0.51	0.69	0.18
8	Shrubland	0.75	0.64	0.82	0.13
9	Mixed Shrubland/Grassland	0.81	0.62	0.64	0.14
10	Savanna	0.77	0.62	0.59	0.12
11	Deciduous Broadleaf Forest	0.81	0.69	0.66	0.08
12	Deciduous Needleleaf Forest*				
13	Evergreen Broadleaf Forest	0.77	0.69	0.78	0.06
14	Evergreen Needleleaf Forest	0.75	0.64	0.64	0.11
15	Mixed Forest	0.74	0.68	0.58	0.12
16	Water Bodies				
17	Herbaceous Wetland*				
18	Wooded Wetland	0.59	0.56	0.42	0.13
19	Barren or Sparsely Vegetated	0.53	0.96	0.28	0.10
20	Herbaceous Tundra**				
21	Wooded Tundra	0.63	0.62	0.61	0.16
22	Mixed Tundra	0.50	0.66	0.36	0.19
23	Bare Ground Tundra*				
24	Snow or Ice				

Notes:
*Land cover type does not exist in the CWRF domain.
**Land cover type does not exist in the global dataset.

[H] Leaf and Stem Area Index (LAI and SAI)

The LAI and SAI are defined, respectively, as the total one-sided area of all green canopy elements and stems plus dead leaves over vegetated ground area. They are constructed from the global monthly mean distributions of green vegetation leaf area index, based on the July 1981-December 1999 AVHRR NDVI data at 8-km spacing on the Interrupted Goode Homolosine projection provided by Boston University (Zhou et al. 2001; Buermann et al. 2002). There exist missing data zones in some land cover regions: urban and built-up, permanent wetlands, marshes, tundra, barren, desert, or very sparsely vegetated area. These missing zones are *filled* by the average over nearby data pixels having the same LCC within a certain radius starting from 16 km (24 pixels) around a missing point and increasing until a 3-pixel minimum is obtained. Filled data are *converted* into the raster

grid, then the polygon coverage, and *remapped* onto the CWRF projection. The result is further *adjusted* to be confined by the USGS LCC (see [F]) for a consistent representation of water bodies. The product is denoted as L_{raw}.

Because L_{raw} is defined with respect to unit ground area, it is divided by local vegetation cover C_v to define L_{gv} representing the green leaf area index with respect to vegetated area only (Zeng et al. 2002). Due to inconsistency between C_v and L_{raw} data at individual pixels, some L_{gv} values are abnormally large, up to several hundreds. The inconsistency arises mainly because C_v was derived based on the 24 USGS LCCs at 1-km spacing, but L_{raw} in terms of six alternative biomes with distinct vegetation structures at an 8-km interval. Zeng et al. (2002) determined C_v at every point, while defining LAI for each IGBP LCC by a mean seasonal variation within a 10° latitude zone.

Table 5. Ecological Parameters in Deriving LAI and SAI for each USGS Land Cover

Type	USGS land cover legend	FVC (1 km) RCM	FVC (1 km) Global	Displacement height (m)	L_{gv} filter threshold	γ	SAI_{min}
1	Urban and Built-Up Land	0.767	0.735	0.667	7	0.00	0.1
2	Dryland Cropland and Pasture	0.941	0.875	0.667	7	0.00	0.1
3	Irrigated Cropland and Pasture	0.882	0.804	0.667	7	0.00	0.1
4	Mixed Dryland/Irrigated Cropland and Pasture**			0.667			
5	Cropland/Grassland Mosaic	0.823	0.729	0.667	7	0.25	0.5
6	Cropland/Woodland Mosaic	0.958	0.869	0.667	7	0.25	0.5
7	Grassland	0.805	0.711	0.667	6	0.50	1.0
8	Shrubland	0.417	0.381	0.333	5	0.50	1.0
9	Mixed Shrubland/Grassland	0.722	0.391	0.333	5	0.50	1.0
10	Savanna	0.899	0.848	0.667	7	0.50	1.0
11	Deciduous Broadleaf Forest	0.947	0.871	13.333	8	0.50	1.0
12	Deciduous Needleleaf Forest*		0.920	13.333	8	0.50	1.0
13	Evergreen Broadleaf Forest	0.955	0.953	23.333	8	0.50	1.0
14	Evergreen Needleleaf Forest	0.898	0.895	13.333	8	0.50	1.0
15	Mixed Forest	0.848	0.875	13.333	8	0.50	1.0
16	Water Bodies			0.667			
17	Herbaceous Wetland*		0.947	13.333	6	0.50	1.0
18	Wooded Wetland	0.729	0.835	0.667	8	0.50	1.0
19	Barren or Sparsely Vegetated	0.061	0.073	0.333	4	0.50	1.0
20	Herbaceous Tundra**			0.667			
21	Wooded Tundra	0.704	0.714	0.667	6	0.50	1.0
22	Mixed Tundra	0.396	0.323	0.333	6	0.50	1.0
23	Bare Ground Tundra*		0.018	0.333	6	0.50	1.0
24	Snow or Ice			0.667			

Notes:
*Land cover type does not exist in the CWRF domain.
**Land cover type does not exist in the global dataset.

For the CWRF, the 1-km C_v data are integrated onto the 8-km L_{raw} map to compute L_{gv} guess values and then a smoothing filter removes abnormal values. The filter was designed through trial and error by examining the frequency distribution of abnormal L_{gv} values and considering the canopy displacement height in the CLM for each USGS LCC. The point value that exceeds the filter threshold listed in Table 5 is *filled* by the average over nearby data pixels having the same LCC within a certain radius starting from 16 km (24 pixels) around the point and increasing until a 3-pixel minimum is obtained. In addition, L_{gv} data contain large uncertainties in winter due to cloud contamination, especially for the USGS LCCs 13 and 14 (evergreen broadleaf and needleleaf forests). Following Zeng et al. (2002), L_{gv} values in winter months for these two categories are adjusted by:

$$L_{gv} = max(L_{gv}, c\, L_{gv,max}) \quad (3)$$

where correction coefficient c is 0.8 and 0.7 for category 13 and 14 respectively, and $L_{gv,max}$ is the maximum L_{gv}. For the climatology, the maximum can be determined from all monthly values during the entire period, while for interannual variations it is taken in three consecutive years.

After extreme value removal and winter adjustment at each 8-km pixel, the new L_{gv} is multiplied by its respective C_v and then *intersected* with the CWRF grid mesh. The fractional area of each pixel contributing to the grid is *extracted*. The area-weighted averaging of all pixels within each CWRF grid results in the new LAI per unit ground, which will be divided by local FVC (see [G]) to produce the final LAI.
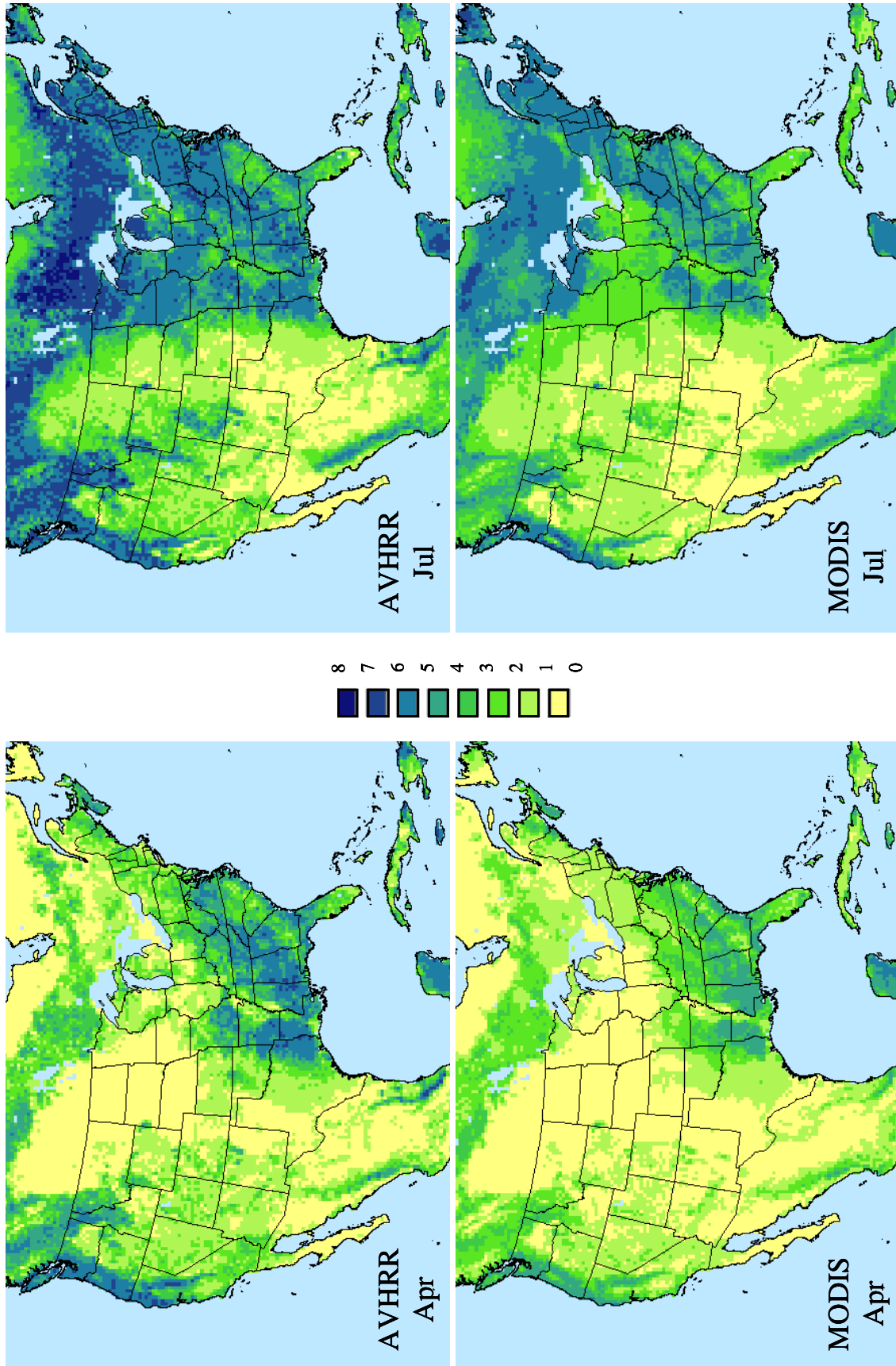
17

Figure 8. Geographic distributions of April (left) and July (right) mean LAI based on the original 1981-1999 AVHRR (top) and 2000-2003 MODIS (bottom) data,

For each USGS LCC, SAI is then approximated as in Zeng et al. (2002) by:

$$SAI^m = max\{ SAI_{min}, [\gamma\, SAI^{m-1} + max(\, LAI^{m-1} - LAI^m, 0\,)]\}$$

(4)

where m denotes month, SAI_{min} the prescribed minimum SAI, and γ the monthly remaining rate after dead leaves removed. Both γ and SAI_{min} are listed in Table 5. Most LCCs reach the minimum SAI in winter and the maximum SAI in fall (October or November). This seasonal trend may not be appropriate for certain categories, especially those with croplands where nothing may remain on the field after fall harvest.

A serious concern is the systematic difference in the L_{raw} products based on the AVHRR (Zhou et al. 2001) and the recent MODIS (Knyazikhin et al. 1998; Myneni et al. 2002) data. The MODIS L_{raw}, available from February 2000 onward, has a finer resolution at 1-km spacing (http://edcimswww.cr.usgs.gov/pub/imswelcome/). Following the same procedure described above, the corresponding LAI and SAI can be constructed from the MODIS L_{raw}. Figure 8 depicts April and July mean LAI distributions of the AVHRR and MODIS climatologies over the CWRF domain, while Figure 9 presents seasonal variations of the five key regions outlined in Figure 6. The MODIS values are clearly smaller, which is not a result of long-term trends. Figure 10 compares AVHRR and MODIS monthly mean LAI variations averaged over five U.S. key regions for the respective predominant LCC types. These include Texas (grassland), Southwest (shrubland), Midwest (dryland cropland and pasture), Southeast (evergreen needleleaf forest), and Northeast (deciduous broadleaf forest). Apparent discontinuities exist between the two datasets, when the MODIS values are systematically smaller, especially for the Midwest cropland. Analyses indicate that certain relationships exist, but vary greatly with regions. No physically sound and statistically robust adjustment can be made for consistency. A first-order correction is to obtain the same climatology while retaining the interannual variability at each CWRF grid. Figure 10 shows corrected regional mean time series using the AVHRR or MODIS climatology.

An open question is which climatology, AVHRR or MODIS, is more realistic. More than 1000 published estimates during 1932-2000 at nearly 400 field sites over the globe have been compiled (Scurlock et al. 2001) and validation is underway, but no direct intercomparison between field measurements and satellite products is currently available. Figure 11 compares monthly mean LAI variations for the Midwest cropland based on the January-December 1999 AVHRR 8-km and January 1999-May 2001 16-km data (ftp://crsa.bu.edu/pub/rmyneni/myneniproducts/datasets/), and February 2000-May 2001 MODIS 1-km data with June-September 1999-2000 field measurements at a central Illinois soybean/corn site (Steven Hollinger, ISWS, personal communication, 2003). The two AVHRR-based LAI estimates have good agreement and well capture the peak values of the field observations. Clearly, the MODIS-based estimates substantially underestimate the growing season LAI for the cropland.
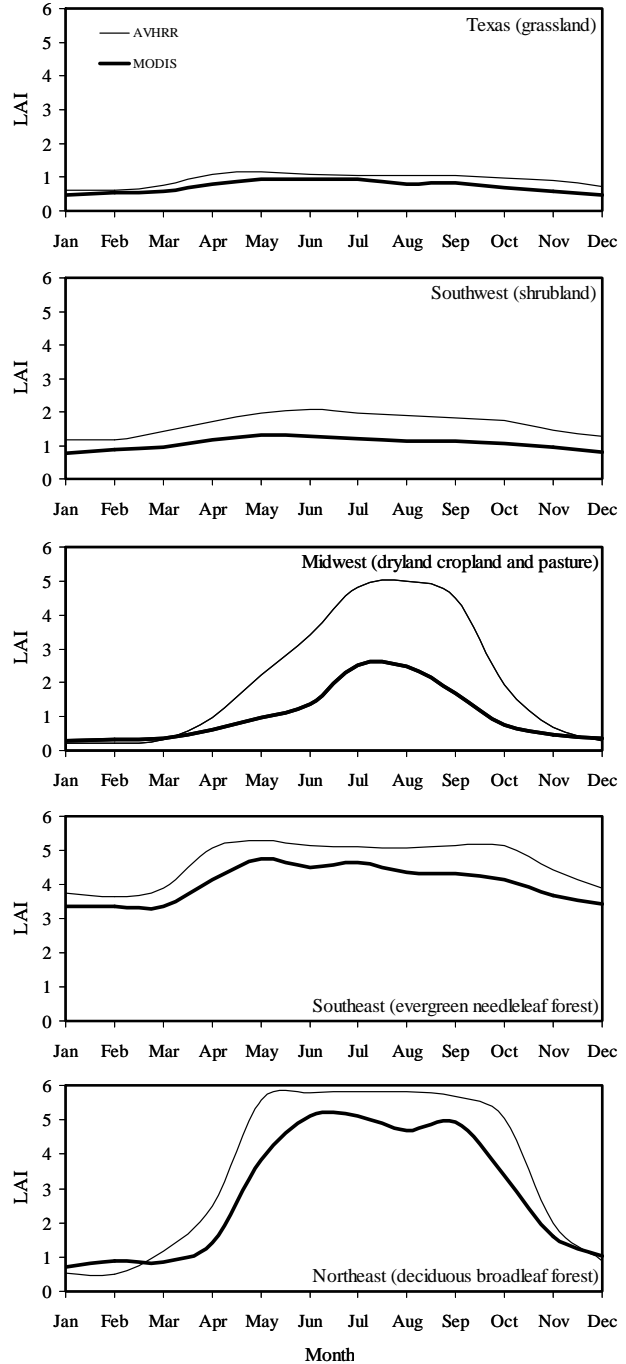

Figure 9. Annual cycle of LAI climatologies for predominant LCC types over 5 key regions, as derived from the original 1981-1999 AVHRR (thin solid) and 2000-2003 MODIS (thick solid).

Because the MODIS measurement is continuing and providing finer resolution and quality-controlled data with improved atmospheric correction and cloud screening (Justice et al. 1998) over the AVHRR measurement (Goward et al. 1991), its LAI product is preferred. Until a comprehensive

evaluation or validation is completed and improved products are made available, the authors suggest correcting the AVHRR LAI to reflect the same monthly mean climatology as the MODIS LAI except for the cropland-related LCC categories (2-6), where the opposite correction is applied due to the obvious MODIS underestimation. The result is long-term, continuous, and consistent LAI data. This is particularly important because the surface albedo parameterization developed from the MODIS data depends on LAI (see [I]).
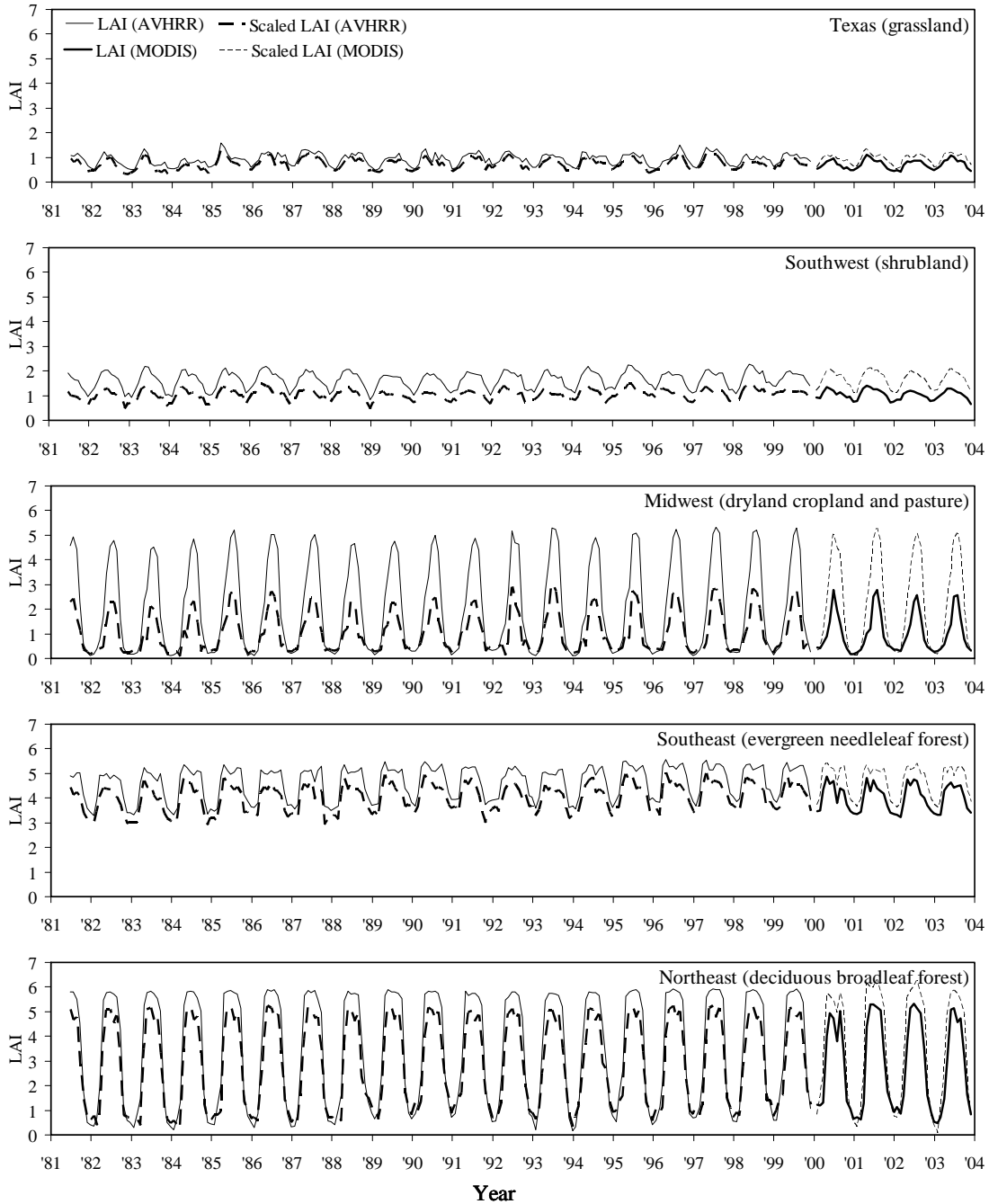

Figure 10. Interannual variations of LAI averaged over 5 key regions for the respective predominant LCC types 7 (Texas), 8 (Southwest), 2 (Midwest), 14 (Southeast), and 11 (Northeast) as derived from the original 1981-1999 AVHRR (thin solid) and 2000-2003 MODIS (thick solid) and bias-corrected correspondence (thin and thick dashed).
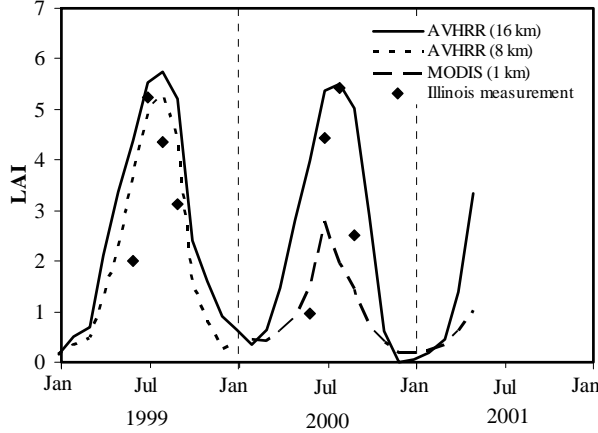
20

Figure 11. Comparison of LAI (Midwest dryland cropland and pasture) based on AVHRR 8-km (dot) and 16-km (solid), and MODIS 1-km (dash) data with Illinois soybean/corn field measurements (diamond), January 1999-May 2001.

[I] Soil Albedo Localization Factor (SALF)

The original CLM parameterization of surface albedo (Dai et al. 2001) was adopted from Dickinson et al. (1993) with improvements by Zeng et al. (2002). It consists of surface albedos for bare ground $\alpha_g(\lambda, w, \text{ and soil color})$, vegetation canopy $\alpha_c(\lambda, \eta, \mu, L_{sai}, \text{ and } \alpha_g)$, and snow cover $\alpha_s(\lambda, \eta, \mu, \text{ snow fraction, amount, and age})$, weighted by their effective fractional areas contributing to the grid. Listed in parentheses are the corresponding dependence parameters, where λ denotes visible and near-infrared wavebands (0.3-0.7 and 0.7-5.0 µm), η represents direct beam and diffuse radiation, μ represents cosine solar zenith angle (only in direct beam), w represents surface soil water volumetric content, and $L_{sai}=LAI+SAI$. Note that α_g, not a function of μ, ignores the difference between direct and diffuse radiation while depending on soil color classes that lack global observations. In addition, Zhou et al. (2003) showed that the CLM-parameterized albedos have substantial biases from the MODIS measurements (http://edcimswww.cr.usgs.gov; Schaaf et al. 2002). Correlation coefficients between the CLM and MODIS snow-free albedos during 2000-2003 range from -0.5 to 0.3 over the CWRF domain.

Therefore the authors developed an improved *snow-free* surface albedo parameterization based on the MODIS data. Although a separate paper being prepared documents the details of the algorithm development, a brief description of the parameterization directly relevant to this study is given here. In particular, the static parts that are geographically dependent and the dynamic components that can be predicted in terms of the CWRF variables are distinguished:

$$\alpha_{snow\,free} = SALF(\lambda, \eta) \cdot \begin{bmatrix} (1-FVC) \cdot \alpha_g(\lambda, \eta, \mu, w) + \\ FVC \cdot \alpha_c(\lambda, \eta, \mu, LCC, L_{sai}, \alpha_g) \end{bmatrix}$$

(5)

where SALF is the static localization factor to represent albedo dependence on local soil characteristics (e.g., soil color, and surface roughness) and canopy structures (e.g., mosaic distribution of multiple vegetation categories). Hence this factor, like FVC and LCC, is also a function of geographic locations and is included in the CWRF SBCs.

The new parameterization is a substantial improvement over the old one in which the correlation coefficients with the MODIS data significantly increased toward a 0.6-0.9 range over the CWRF domain. Figure 12 compares the geographic distributions of the old and new parameterizations with the MODIS-measured July surface albedos averaged during 2000-2003, along with SALF, for near-infrared direct beam.

[J] Sea Surface Temperature (SST)

One of the most important variables responsible for seasonal-interannual climate variability worldwide is SST. Given the lack of fine-resolution data, most mesoscale models have been using the weekly Optimum Interpolation (OI) SST analysis at 1° longitude-latitude spacing, a blend of multi-channel AVHRR infrared measurements with in-situ ship and buoy observations (Reynolds et al. 2002). The CWRF, following that for the CMM5 (Liang et al. 2004a), incorporates daily variations based on a conservative spline-fit procedure from the weekly OISST data, available over oceans globally from November 1981 onward (http://www.emc.ncep. noaa.gov/research/cmb/sst_analysis). Note that daily SSTs interpolated directly from weekly values (treated as if they were at mid-week) do not conserve weekly means nor preserve the extremes (Taylor et al. 2000). Thus, an iterative spline-fit procedure was used to interpolate daily SST variations from the weekly data while conserving the weekly means. These daily SST variations then were incorporated into the CWRF with daily updates. This procedure effectively preserves the extremes revealed in the original weekly data.

A new blended real-time global (RTG) SST analysis (Thiébaux et al. 2003) is now available with daily data at 0.5° longitude-latitude spacing since 11 February 2001 (ftp://polar. wwb.noaa.gov/pub). The analysis ingests the most current observations in the previous 24 hours, including NOAA-16 SEATEMP retrievals from the Naval Oceanographic Office Major Shared Resource Center, ship and buoy in-situ SST reports from the Global Telecommunications System, and Special Sensor Microwave Imager sea ice concentrations. Most recently, the MODIS level-3 mapped SST data are available twice daily at 4-km spacing, derived from infrared brightness temperature measurements by Terra since 19 July 2000 (http://daac.gsfc.nasa.gov/data/datapool/MODIS/ 03_Ocean_/03_Level3_Mapped/001_MO04MD_4KM_Daily/ index.html) or Aqua since 7 January 2003 (http://daac.gsfc. nasa.gov/data/datapool/MODIS-qua/03_Ocean/ 03_Level3_Mapped/001_MY04MD_4KM_Daily/index.html). The MODIS sensor was designed with higher sensitivity and lower signal-to-noise ratios than AVHRR radiometers on NOAA satellites. Despite the similarity in the analysis algorithm, the MODIS SST has potentially better accuracy and higher resolution than the OISST.
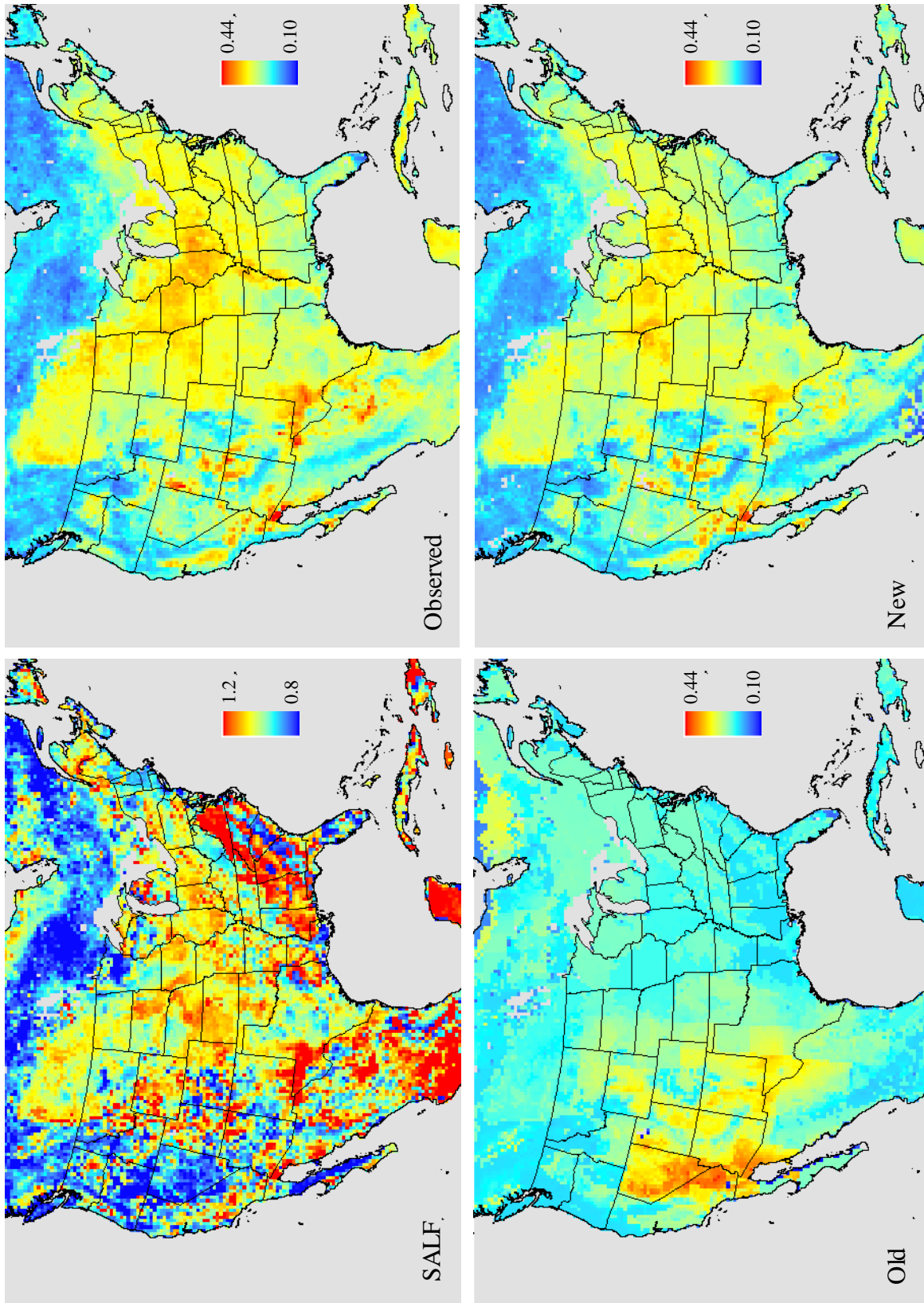
21

Figure 12. Geographic distributions of July mean surface albedo for near-infrared direct beam as observed by the MODIS (top right) and parameterized by the old (bottom left) and new (bottom right) schemes averaged during 2000-2003, along with SALF (top left).
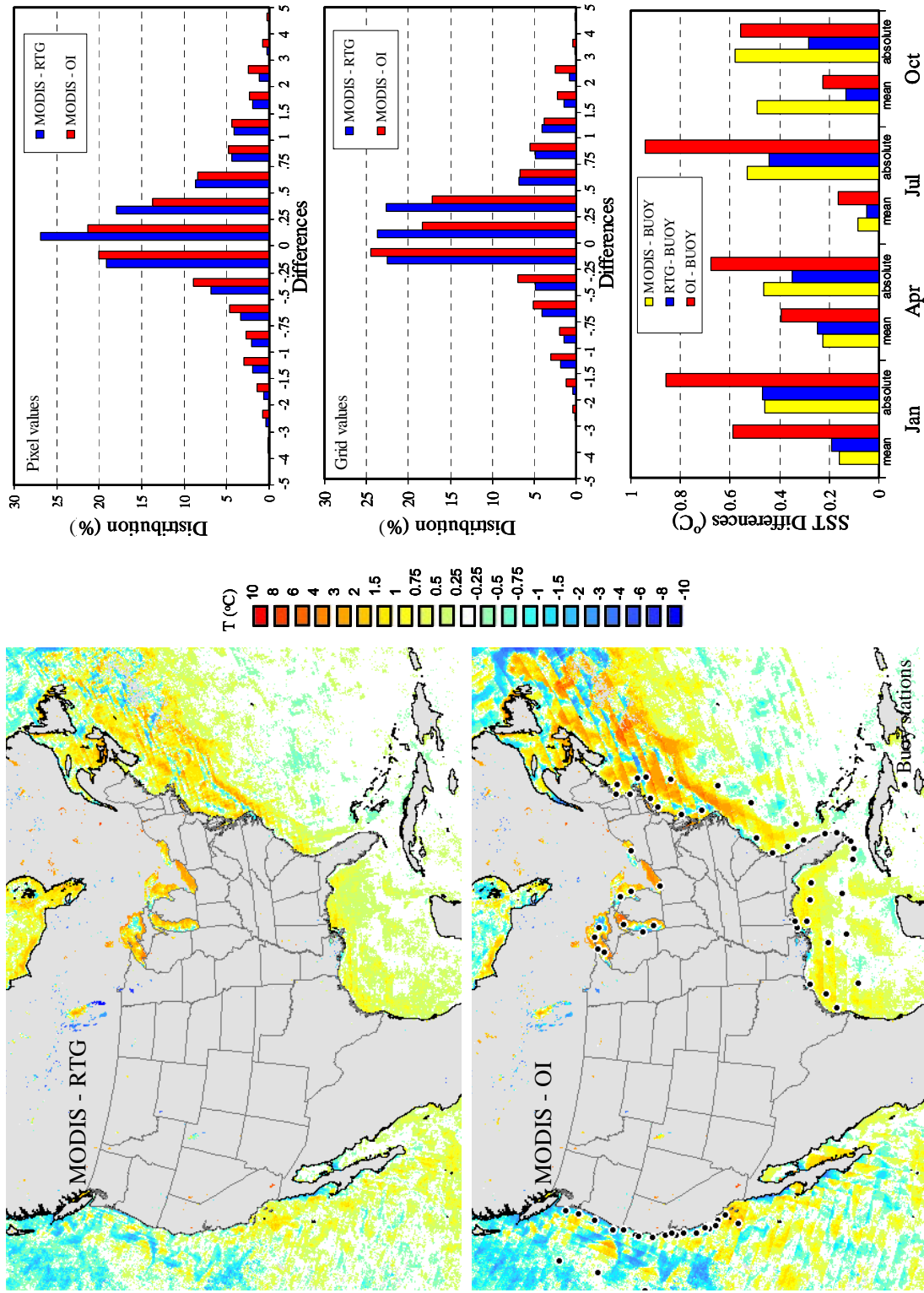
Figure 13. Geographic distributions (left) of the October 2002 mean MODIS SST differences (°C) from RTGSST (top) and OISST (bottom). The right panels depict the frequency distributions of the differences at raw data pixels (top) and 30-km CWRF grids (middle), and mean and absolute differences (bottom) of MODIS, RTG and OI SST from hourly observations at 70 buoy stations (dots in left bottom map) uniformly distributed over U.S. coastal oceans and Great Lakes in January, April, July, and October.

Figure 13 compares the OISST and RTGSST analyses with the October 2002 MODIS data in terms of geographic variations and frequency distributions of the differences. To depict the resolution enhancement effect, both OISST and RTGSST data are mapped onto the MODIS pixel mesh using longitude-latitude bilinear interpolation. Approximately 41 and 64 percentage or % (46 and 71%) of oceanic pixels in the current CWRF domain have SST differences from the MODIS within ±0.25 and 0.5 ºC for the OISST (RTGSST) analysis, respectively. About 15 and 4% (11 and 2%) of pixels contain absolute differences greater than 1 and 2 ºC, respectively. In particular, the MODIS data tend to be cooler than both OISST and RTGSST values in the northeast Pacific Ocean, but warmer over the Great Lakes, the Gulf of Mexico, and western coastal Atlantic Ocean. As shown in Figure 13, a comparison with the hourly observations at 70 buoy stations uniformly distributed over these water bodies indicates overall better agreement by the RTGSST and MODIS data, but a relatively larger error with the OISST data. In addition, MODIS data daytime and nighttime differences contain significant biases from the buoy observations, and thus more effort is needed before their real applications.

Note that SSTs from infrared satellite retrievals are subject to varying degrees of residual cloud contamination, which is difficult to characterize, particularly in the absence of in-situ observations. The cloud masking results in cold biases for SST retrievals from single-view infrared sensor data and is difficult to correct using the OI analysis in regions with sparse in-situ data. In addition, the MODIS data, albeit twice daily, are available only over the scanning tracks along the satellite passes, and hence contain broad areas with missing data. Similar to the OISST and RTGSST analyses, the MODIS measurements must be blended with in-situ ship and buoy observations to fill all missing data objectively before their application in the CWRF. The CWRF incorporates daily SST variations from the weekly OISST data via the iterative, conservative spline-fit procedure or the daily RTGSST data, both of which are *interpolated* into the CWRF grid due to their coarse resolutions. In the future, the improved, blended MODIS data will *intersect* with the

CWRF grid mesh to determine SST by area-weighted averaging over all contributing pixels.

[K] Sea Surface Current and Salinity (SSC and SSS)

The monthly mean SSC climatology was produced from a ship drift database provided by the U.S. Coast Guard (Mariano et al. 1995). For each velocity component, *u* or *v*, a median filter was applied to remove gross outliers; a scalar objective analysis following Mariano and Brown (1992) then yielded the final estimates. The quality of the estimates is poor in southern oceans due to missing data, especially south of 50°S. The raw data are available at 1° longitude-latitude spacing over global oceans (http://oceancurrents.rsmas.miami.edu/). The monthly mean SSS climatology is adopted from that of SSP (see [L]) at sea level. Given the coarse resolution, both SSC and SSS data are *extrapolated* for the coastal boundaries and then *interpolated* into the CWRF grid. Over inland lakes with missing data, both SSC and SSS are assigned a value of zero. Figure 14 depicts the SSC and SSS distributions in January, April, July, and October.

[L] Sea Temperature and Salinity Profiles (STP and SSP)

Both STP and SSP are based on the Polar Science Center Hydrographic Climatology or PHC (Steele et al. 2001), available on 33 depth levels (0, 10, 20, 30, 50, 75, 100, 125, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1750, 2000, 2500, 3000, 3500, 4000, 4500, 5000, and 5500 m) at 1° longitude-latitude spacing over global oceans (http://psc.apl.washington.edu/Climatology.html). The PHC data were created via an optimal interpolation procedure that merges the 1998 version of the World Ocean Atlas (Antonov et al. 1998) with the recent Arctic Ocean Atlas (EWG 1997, 1998). The monthly data are provided for the upper 24 levels to a 1500-m depth, while seasonal (March-April-May and July-August-September) and annual means are available for all levels. Both STP and SSP data are *extrapolated* for the coastal boundaries at each depth level and then *interpolated* into the CWRF grid. Figure 15 depicts the STP and SSP distributions on the 50- and 300-m depth levels in July.

Implementation of SBCs into the WRF I/O API

As stated in Michalakes (2000), I/O presents the greatest stumbling block to a software and data architecture that strives for independence with respect to machine architecture, institutional conventions, and application needs, with issues in parallel I/O data, and metadata formats and conventions. The WRF incorporates an I/O API with a package-independent set of routines implemented using GRIdded Binary (GRIB) , Network Common Data Form (netCDF), or Hierarchical Data Format (HDF) packages. Implementation of these packages, parallel I/O, and vendor specifics are hidden by the I/O API and transparent to the WRF application. The CWRF complies with the same API standard to acquire I/O of all SBCs listed in Table 1.

The WRF inputs or outputs data on four distinct data streams: initial, restart, boundary (lateral), and history. They can be directed separately to different I/O packages or formats within the same model run. The static or time-

invariant SBCs (HSFC, HSDV, HSLD, HCVD, DBED, SAND, CLAY, SALF, TBS, SCI, LCC, and FVC) are inserted into the existing files *wrfinput* and *wrfrst*, using data stream INPUT and RESTART, respectively. The time-variant SBCs are incorporated through three auxiliary boundary (surface) data streams, using AUXINPUT1, 2, and 3, respectively, with *wrfsst* for daily SST, *wrfveg* for daily vegetation variables (LAI and SAI), and *wrfocn* for monthly oceanic fields (SSS, SSC, STP, and SSP). The WRF Registry registers a time-variant SBC as <field>B with double memory addresses that store two consecutive data records for temporal interpolation into its instantaneous value at each shorter step calling the surface module between the data period. This facilitates simple implementation with an easy distinction from <field> predicted by an interactive or dynamic surface module. Effective implementation requires consistent modifications to the WRF mediation modules and the WRFSI and REAL pre-
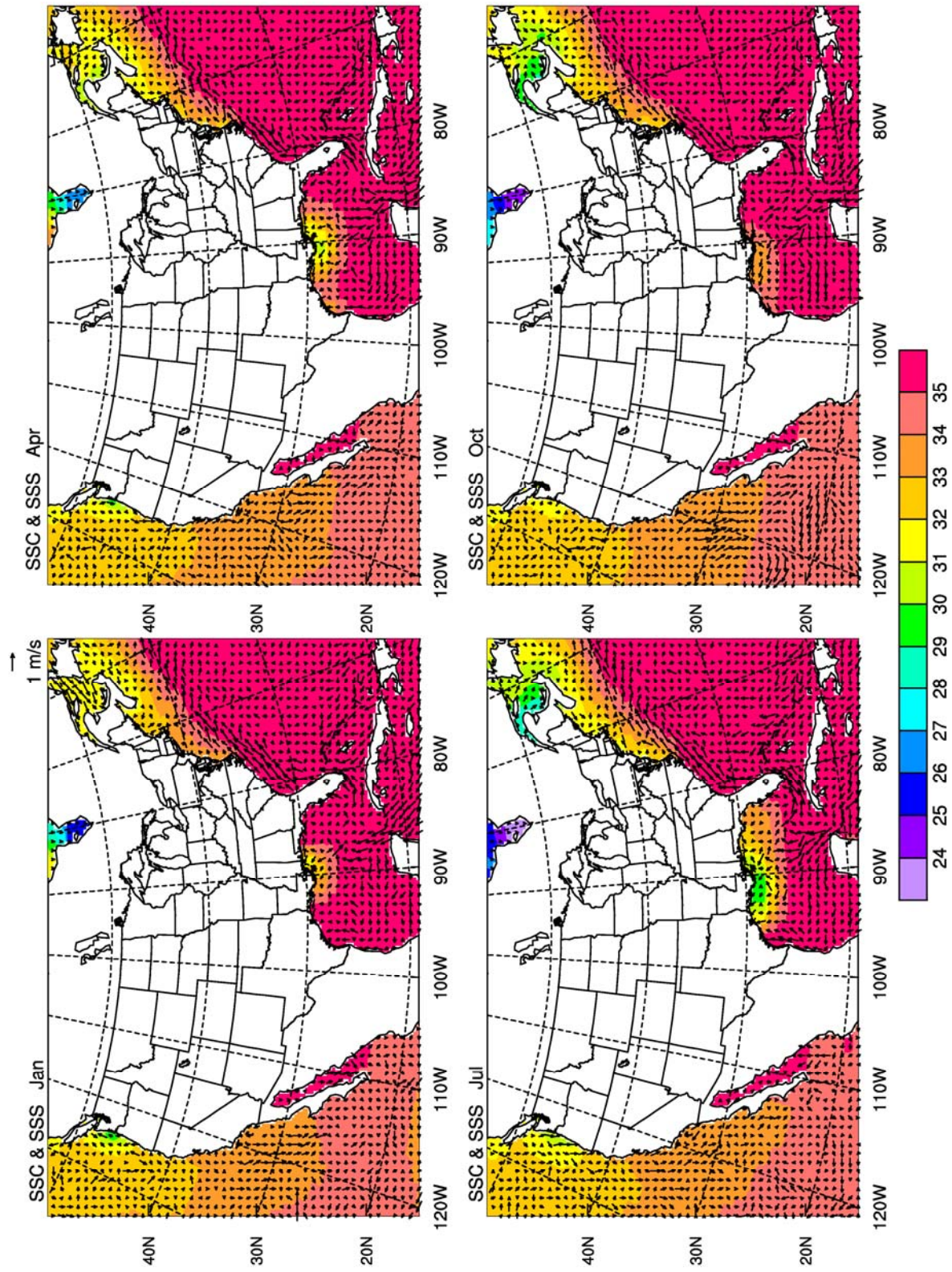
Figure 14. Geographic distributions of SSC (m s^{-1}; vector) and SSS (‰; color shading) in January, April, July, and October.
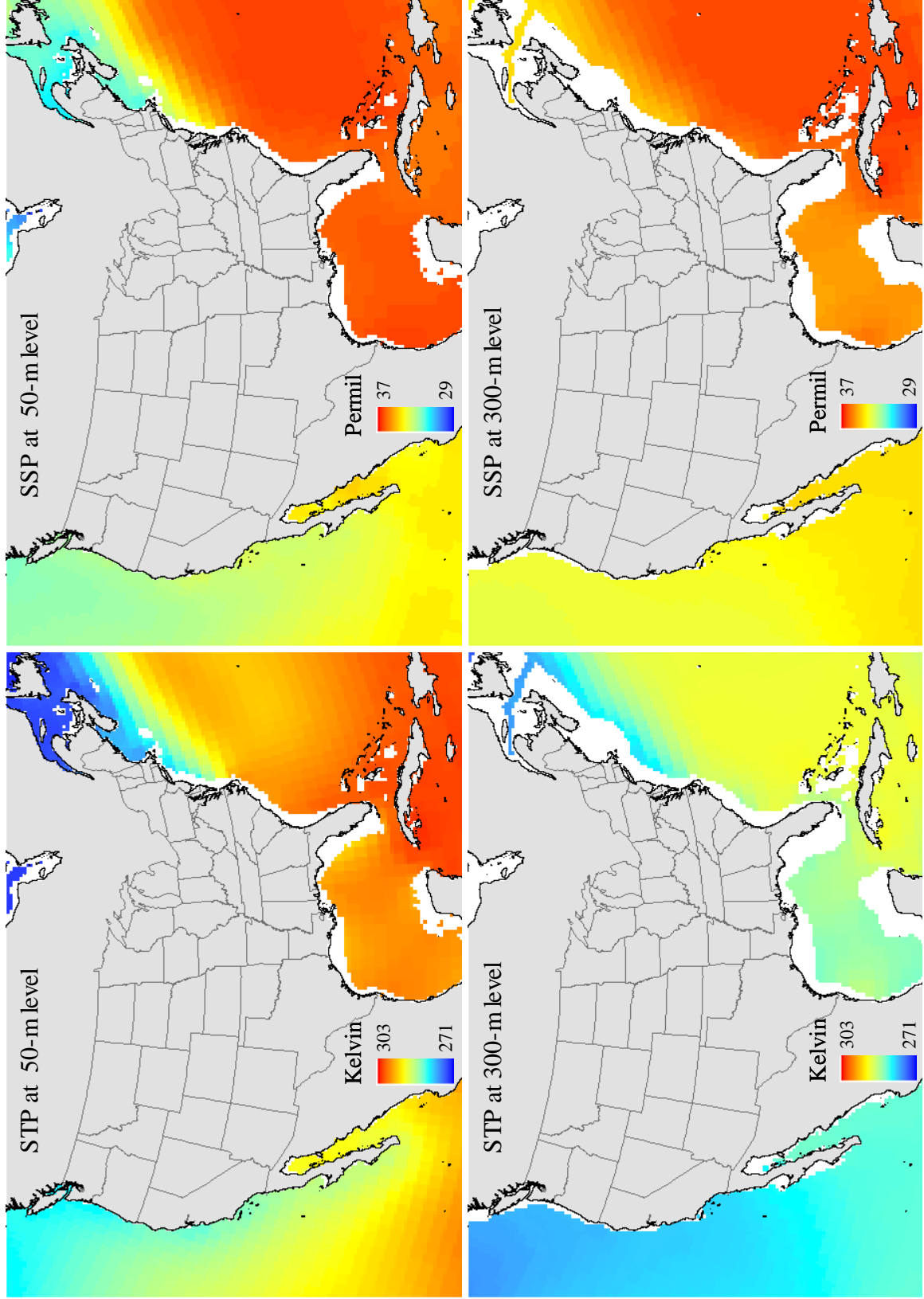
Figure 15. Geographic distributions of STP (K; left) and SSP (‰; right) at the 50-m (top) and 300-m (bottom) depth levels in July with white areas corresponding to level below seafloor.

processors, all following the WRF I/O API standard. Special care has been taken to enable the CWRF restart from any time step between the first and last data records in an SBC file. Individual SBC files may contain data records with different starting and ending dates.

Summary

The CWRF, the climate *extension* of the WRF, is being developed with key improvements on planetary-mesoscale, surface-atmosphere, and convection-cloud-radiation interactions and system consistency throughout all process modules. The extension inclusively incorporates all WRF functionalities for weather forecasts while enhancing the capability for climate predictions. As an initial effort of a series documenting CWRF formulation and performance, this report focuses on the construction and implementation of SBCs desired for general CWRF applications to all effective, dynamically coupled or uncoupled, combinations of the surface modules, as well as portability to any specific region of the world. The primary SBCs include surface topography (mean elevation, slope, curvature, and their standard deviations); bedrock, lakebed or seafloor depth; soil sand and clay fraction profiles; surface albedo localization factor; bottom soil temperature; surface characteristic identification; land cover category; fractional vegetation cover; leaf and stem area index; sea surface temperature, salinity, and current; and sea temperature and salinity profiles. This is by far the most comprehensive set of SBCs specifically designed for mesoscale modeling applications, and is especially beneficial for CWRF users. The climatic impacts of these new or improved SBCs, singly or in combination, are being investigated and will be documented in the future.

Given that data quality and value representation depend on the RCM computational domain and grid resolution, all the SBCs are constructed onto the 30-km CWRF domain suitable for U.S applications. Nonetheless, the authors elaborate, in detail, on raw data sources and processing procedures by which the SBCs can be readily constructed for any specific CWRF domain in the world. Although the authors have strived for the best-available data quality, comprehensive processing procedures, and consistency between alternatives, the SBCs so constructed carry over uncertainties inherent in the raw data. Evaluation of the overall data quality of the SBCs was beyond the scope of this study.

References

Anthes, R.A., E.Y. Hsie, and Y.H. Kuo. 1987. *Description of the Penn State/NCAR Mesoscale Model Version 4 (MM4).* NCAR Tech. Note, NCAR/TN-282+STR:66 pp, Boulder, CO.

Antonov, J.I., S. Levitus, T.P. Boyer, M.E. Conkright, T.D. O'Brien, and C. Stephens. 1998. *World Ocean Atlas 1998 Vol. 1: Temperature of the Atlantic Ocean.* NOAA Atlas NESDIS 27:166 pp, Washington, D.C.

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.

Belward, A. S. 1996. *The IGBP-DIS global 1 km land cover data set (DISCover): Proposal and implementation plans.* IGBP-DIS Working Paper No. 13:61 pp. [Available from IGBP-DIS Office, 42 Ave. Gustave Coriolis 31057 Toulouse Cedex, France.]

Betts, A.K., 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.

Beven, K.J., and M.J. Kirkby. 1979. A physically based variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.* **24**(1):4-69.

Bonan, G.B. 1995. Sensitivity of a GCM simulation to inclusion of inland water surfaces. *J. Climate* **8**:2691-2704.

Bonan, G.B. 1996. *A land surface model (LSM version 1.0) for ecological, hydrological, and atmospheric studies: Technical description and User's guide.* NCAR Tech. Note, NCAR/TN-417+STR:150 pp, Boulder, CO.

Buermann, W., J. Dong, X. Zeng, R.B. Myneni, and R.E. Dickinson. 2001. Evaluation of the utility of satellite-based vegetation leaf area index data for climate simulations. *J. Climate* **14**:3536-3551.

Buermann, W., Y. Wang, J. Dong, L. Zhou, X. Zeng, R.E. Dickinson, C.S. Potter, and R.B. Myneni. 2002. Analysis of a multiyear global vegetation leaf area index data set, *J. Geophys. Res.* **107**:4646, doi:10.1029/2001JD000975.

Charney, J.G. 1975. Dynamics of deserts and drought in the Sahel. *Quart. J. Roy. Meteor. Soc.* **101**:193-202.

Chase, T.N., R.A. Pielke, T.G.F. Kittel, R. Nemani, and S.W. Running. 1996. The sensitivity of a general circulation model to global changes in leaf area index. *J. Geophys. Res.* **101**:7393-7408.

Chen, F., and J. Dudhia. 2001. Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.* **129**:569-585.

Chen, J., and P. Kumar. 2001. Topographic influence of the seasonal and interannual variation of water and energy balance of basins in North America. *J. Climate* **14**:1989-2014.

Chen, S.-H., and J. Dudhia. 2000. *Annual Report: WRF Physics* (http://www.mmm.ucar.edu/wrf/users/wrf-doc-physics .pdf).

Chen, T.H., A. Henderson-Sellers, P.C.D. Milly, A.J. Pitman, A.C.M. Beljaars, J. Polcher, F. Abramopoulos, A. Boone, S. Chang, F. Chen, Y. Dai, C.E. Desborough, R.E. Dickinson, L. Dümenil, M. Ek, J.R. Garratt, N. Gedney, Y.M. Gusev, J. Kim, R. Koster, E.A. Kowalczyk, K. Laval, J. Lean, D. Lettenmaier, X. Liang, J.-F. Mahfouf, H.-T. Mengelkamp, K. Mitchell, O.N. Nasonova, J. Noilhan, A. Robock, C. Rosenzweig, J. Schaake, C.A. Schlosser, J.-P. Schulz, Y. Shao, A.B. Shmakin, D.L. Verseghy, P. Wetzel, E.F. Wood, Y. Xue, Z.-L. Yang, and Q. Zeng. 1997. Cabauw experimental results from the project for intercomparison of land-surface parameterization schemes. *J. Climate* **10**:1194-1215.

Cosby, B.J., G.M. Hornberger, R.B. Clapp, and T.R. Ginn. 1984. A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.* **20**:682-690.

Copeland, J.H., R.A. Pielke, and T.G.F. Kittel. 1996. Potential climatic impacts of vegetation change: A regional modeling study. *J. Geophys. Res.* **101**:7409-7418.

Dai, Y., X. Zeng, R.E. Dickinson, and co-authors. 2001. The Common Land Model (CLM): Technical Documentation and User's Guide (http://climate.eas.gatech.edu/dickinson/).

Dai, Y., X. Zeng, R.E. Dickinson, I. Baker, G.B. Bonan, M.G. Bosilovich, A.S. Denning, P.A. Dirmeyer, P.R. Houser, G. Niu, K.W. Oleson, C.A. Schlosser, and Z-L. Yang. 2003. The Common Land Model. *Bull. Amer. Meteor. Soc.* **84**:1013-1023.

Dickinson, R.E., J.A. Berry, G.B. Bonan, G.J. Collatz, C.B. Field, I.Y. Fung, M. Goulden, W.A. Hoffman, R.B. Jackson, R. Myneni, P.J. Sellers, and M. Shaikh. 2002. Nitrogen controls on climate model evapotranspiration. *J. Climate* **15**:278-295.

Dickinson, R.E., A. Henderson-Sellers, P.J. Kennedy, and M.F. Wilson. 1993. Biosphere atmosphere transfer scheme (BATS) version 1e as coupled for Community Climate Model. NCAR Tech. Note NCAR/TN-378+STR:72 pp, Boulder, CO.

Dickinson, R.E., M. Shaikh, R. Bryant, and L. Graumlich. 1998. Interactive canopies for a climate model. *J. Climate* **11**:2823-2836.

Dudhia, J. 1996. A multi-layer soil temperature model for MM5. *Preprint from the Sixth PSU/NCAR Mesoscale Model Users' Workshop*, Boulder, CO (http://www.mmm.ucar.edu/mm5/lsm/ lsm-docs.html).

Dudhia, J., D. Gill, Y.-R. Guo, K. Manning, W. Wang, and J. Chriszar. 2000. *PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's Guide: MM5 Modeling System Version 3*, Boulder, CO (http://www.mmm.ucar.edu/mm5/doc.html).

Ek, M.B., K.E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J.D. Tarpley. 2003. Implementation of NOAH land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.* **108**:8851, doi:10.1029/2002JD003296.

EWG (Environmental Working Group). 1997. *Joint U.S.-Russian Atlas of the Arctic Ocean for the Winter Period.*

Natl. Snow and Ice Data Cent., CD-ROM,Washington, DC.

EWG (Environmental Working Group). 1998. *Joint U.S.-Russian Atlas of the Arctic Ocean for the Summer Period.* Natl. Snow and Ice Data Cent., CD-ROM,Washington, DC.

FAO (Food and Agriculture Organization). 1974. *Soil map of the world Vol. 1-10.* Paris, France.

FAO (Food and Agriculture Organization). 1996. *The Digitized Soil Map of the World Including Derived Soil Properities,* CD-ROM, Rime, France.

Gallo, K., L. Ji, B. Reed, J. Dwyer, and J. Eidenshink. 2004. Comparison of MODIS and AVHRR 16-day normalized difference vegetation index composite data. *Geophys. Res. Lett.* **31**:L07502, doi:10.1029/2003GL019385.

Gao, X., S. Sorooshian, J. Li, and J. Xu. 2003. SST data improve modeling of North American monsoon rainfall. *EOS* **84**(43):457 & 462.

Giorgi, F. 1997. An approach for the representation of surface heterogeneity in land surface models. Part I: Theoretical framework. *Mon. Wea. Rev.* **125**:1885-1899.

Giorgi, F., M.R. Marinucci, and G.T. Bates. 1993a. Development of a second generation regional climate model (RegCM2): Part I. Boundary layer and radiative transfer processes. *Mon. Wea. Rev.* **121**:2794-2813.

Giorgi, F., M.R. Marinucci, and G.T. Bates. 1993b. Development of a second generation regional climate model (RegCM2): Part II. Convective processes and assimilation of lateral boundary conditions. *Mon. Wea. Rev.* **121**:2814-2832.

Giorgi, F., and L.O. Mearns. 1999. Introduction to special section: Regional climate modeling revisited. *J. Geophys. Res.* **104**:6335-6352.

Goward, S.N., B. Markham, D.G. Dye, W. Dulaney, and A.J. Yang. 1991. Normalized difference vegetation index measurements from the Advanced Very High Resolution Radiometer. *Remote Sens. Environ.* **35**:257-277.

Grell, A.G., J. Dudhia, and D.R. Stauffer. 1994. *A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5).* NCAR Tech. Note, NCAR/TN-398+STR, National Center for Atmospheric Research, Boulder, CO:122 pp

Griffies, S.M., M.J. Harrison, R.C. Pacanowski, and A. Rosati. 2003. *A Technical Guide to MOM4.* GFDL Ocean Group Technical Report No. 5, NOAA/Geophysical Fluid Dynamics Laboratory:337 pp, Washington, DC (http://www.gfdl.noaa.gov/~fms/).

Gutman, G., and A. Ignatov. 1998. The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *Int. J. Remote Sens.* **19**:1533-1543.

Hahmann, A.N., and R.E. Dickinson. 1997. RCCM2-BATS model over tropical South America: Applications to tropical deforestation. *J. Climate* **10**:1944-1964.

Hoffmann, W.A., and R.B. Jackson. 2000. Vegetation-climate feedbacks in the conversion of tropical savanna to grassland. *J. Climate* **13**:1593-1602.

Hunke, E.C., and W.H. Lipscomb. 2001. *CICE: the Los Alamos Sea Ice Model Documentation and Software User's Manual.* T-3 Fluid Dynamics Group, Los Alamos National Laboratory, Los Alamos, NM (http://climate.lanl.gov/Models/ CICE/index.htm).

IPCC (Intergovernmental Panel on Climate Change). 2001. *Climate Change 2001: The Scientific Basis.* J.T. Houghton, Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson (eds.),Cambridge University Press, Cambridge and New York :881 pp. [Especially refer to Chapter 10: Regional Climate Information - Evaluation and Projections, pp. 538-638.]

Jakobsson, M., R. Macnab, and IBCAO member of the Editorial Board. 2001. International Bathymetric Chart of the Arctic Ocean (IBCAO) Version 1.0 Technical Reference and User's Guide (http://www.ngdc.noaa.gov/ mgg/bathymetry/arctic/IBCAO_TechnicalReference.PDF).

James, M.E., and S.N.V. Kalluri. 1994. The Pathfinder AVHRR land data set - An improved coarse resolution data set for terrestrial monitoring. *Int. J. Remote Sens.* **15**:3347-3363.

Jones, J.W., G. Hoogenboom, C.H. Porter, K.J. Boote, W.D. Batchelor, L.A. Hunt, P.W. Wilkens, U. Sinh, A.J. Gijsman, and J.T. Ritchie. 2003. The DSSAT cropping system model. *Europ. J. Agronomy* **18**:235-265.

Justice, D.H., V. Salomonson, J. Privette, G. Riggs, A. Strahler, R. Lucht, R. Myneni, Y. Knjazihhin, S. Running, R. Nemani, E. Vermote, J. Townshend, R. Defries, D. Roy, Z. Wan, A. Huete, W. van Leeuwen, R. Wolfe, L. Giglio, J. Muller, P. Lewis, and M. Barnsley. 1998. The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research. *IEEE Trans. Geosci. Remote Sens.* **36**:1228-1249.

Klemp, J., W. Skamarock, and J. Dudhia. 2000. Conservative split-explicit time integration methods for the compressible nonhydrostatic equations: WRF Eulerian prototype model equations on height and mass coordinates (http://www.mmm.ucar.edu/wrf/users/wrf-dyn-num.html).

Knyazikhin, Y., J.V. Martonchik, R.B. Myneni, D.J. Diner, and S.W. Running. 1998. Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *J. Geophys. Res.* **103**:32257-32276.

Koster, R.D., and M.J. Suarez. 1992. A comparative analysis of two land surface heterogeneity representations. *J. Climate* **5**:1379-1390.

Kumar, P. 2004. Layer averaged Richard's equation with lateral flow. *Advances in Water Resources* **27**(5):521-531.

Kumar P., and J. Chen. 1999. Validation of large-area basin-scale model using streamflow observations from the Mississippi River region. *EOS, Transactions, American Geophysical Union*, **80**(46):F388.

Leung, L.R., and S.J. Ghan. 1998. Parameterizing subgrid orographic precipitation and surface cover in climate models. *Mon. Wea. Rev.* **126**:3271-3291.

Leung, L.R., and S.J. Ghan. 1999. Pacific Northwest climate sensitivity simulated by a regional climate model driven by a GCM. Part I: control simulations. *J. Climate* **12**:2010-2030.

Leung, L.R., S.J. Ghan, Z.-C. Zhao, Y. Luo, W.-C. Wang, and H.-L. Wei. 1999. Intercomparison of regional climate

simulations of the 1991 summer monsoon in eastern Asia. *J. Geophys. Res.* **104**:6425-6454.

Leung, L.R., L.O. Mearns, F. Giorgi, and R.L. Wilby. 2003. Regional climate research. *Bull. Amer. Meteorol. Soc.* **84**:89-95.

Liang, X.-Z., W. Gao, K.E. Kunkel, J. Slusser, X. Pan, H. Liu, and Y. Ma. 2003. Sustainability of vegetation over northwest China. Part 1: Climate response to grassland. In *Ecosystems Dynamics, Ecosystem-Society Interactions, and Remote Sensing Applications for Semi-Arid and Arid Land.* X. Pan, W. Gao, M.H. Glantz, and Y. Honda (eds.), SPIE Press, Volume 4890:pp. 29-44.

Liang, X.-Z., K.E. Kunkel, and A.N. Samel. 2001. Development of a regional climate model for U.S. Midwest applications. Part 1: Sensitivity to buffer zone treatment. *J. Climate.* **14**:4363-4378.

Liang, X.-Z., K.E. Kunkel, R. Wilhelmson, J. Dudhia, and J.X.L. Wang. 2002. The WRF simulation of the 1993 central U.S. heavy rain: Sensitivity to cloud microphysics representation. *Proceedings of the 82nd AMS Annual Meeting: 16th Conference on Hydrology, Orlando, FL*, January 13-17:pp. 123-126.

Liang, X.-Z., L. Li, K.E. Kunkel, M. Ting, and J.X.L. Wang. 2004a. *Regional climate model simulation of U.S. precipitation during 1982-2002. Part 1: Annual cycle. J. Climate* **17**:3510-3528.

Liang, X.-Z., J. Pan, K.E. Kunkel, J.X.L. Wang, E.C. Hunke, and W.H. Lipscomb. 2004b. Coupling the CWRF with the CICE for Arctic climate applications. *Preprints of the 5th WRF/14th MM5 User's Workshop*, June 22-25:pp. 215-220, Boulder, CO.

Loveland, T.R., B.C. Reed, J.F. Brown, D.O. Ohlen, Z. Zhu, L. Yang, and J.W. Merchant. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *Int. J. Remote Sens.* **21**:1303-1330.

Lu, L., and W.J. Shuttleworth. 2002. Incorporating NDVI-derived LAI into the climate version of RAMS and its impact on regional climate. *J. Hydrometeorology* **3**:347-362.

Lynch, A.H., W.L. Chapman, J.E. Walsh, and G. Weller. 1995. Development of a regional climate model of the western Arctic. *J. Climate* **8**:1555-1570.

Mariano, A.J., and O.B. Brown. 1992. Efficient objective analysis of dynamically heterogeneous and nonstationary fields via the parameter matrix. *Deep-Sea Res.* **39**:1255-1271.

Mariano, A.J., E.H. Ryan, B.D. Perkins, and S. Smithers. 1995. *The Mariano Global Surface Velocity Analysis 1.0.* U.S. Coast Guard Technical Report, CG-D-34-95 (see latest updates at http://oceancurrents.rsmas.miami.edu/).

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.

Michalakes, J. 2000. Weather research and forecast model 1.0: Software design and implementation (http://www.mmm.ucar.edu/wrf/users/wrf-architect.html).

Miller, D.A., and R.A. White. 1998. A Conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. *Earth Interactions* **2**:1-26.

Myneni, R.B., S. Hoffman, Y. Knyazikhin, J.L. Privette, J. Glassy, Y. Tian, Y. Wang , X. Song , Y. Zhang, G.R. Smith, A. Lotsch, M. Friedl, J.T. Morisette, P. Votava, R.R. Nemani, S.W. Running. 2002. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sens. Environ.* **83**:214-231.

NAME. 2003. *The North American Monsoon Experiment (NAME): Science and Implementation Plan* (http://www.cpc.ncep.noaa.gov/products/precip/monsoon/NAME.html).

New, M., D. Lister, M. Hulme, and I. Makin. 2002. A high-resolution data set of surface climate over global land areas. *Climate Research* **21**:1-25.

Phillips, T.J., G.L. Potter, D.L. Williamson, R.T. Cederwall, J.S. Boyle, M. Fiorino, J.J. Hnilo, J.G. Olson, S.C. Xie, and J.J. Yio. 2004. Evaluation parameterizations in general circulation models: Climate simulation meets weather prediction. *Bull. Amer. Meteor. Soc.* (**in press**).

Pielke, R.A., T.J. Lee, J.H. Copeland, J.L. Eastman, C.L. Ziegler, and C.A. Finley. 1997. Use of USGS-provided data to improve weather and climate simulations. *Ecol. Appl.* **7**:3-21.

Reddy, K.R., V.G. Kakani, J.M. McKinion, and D.N. Baker. 2002. Applications of a cotton simulation model, GOSSYM, for crop management, economic and policy decisions. *Agricultural System Models in Field Research and Technology Transfer* (L.R. Ahuja, L. Ma, and T.A. Howell, eds.), CRC Press, LLC, Boca Raton, FL, pp.33-73.

Reynolds, C.A., T.J. Jackson, and W.J. Rawls. 2000. Estimating soil water-holding capacities by linking the Food and Agriculture Organization soil map of the world with global pedon databases and continuous pedotransfer functions. *Water Resources Research* **36**:3653-3662.

Reynolds, R.W., N.A. Rayner, T.M. Smith, D.C. Stokes, and W. Wang. 2002. An improved in situ and satellite SST analysis for climate. *J. Climate* **15**:1609-1625.

Roads, J., S. Chen, S. Cocke, L. Druyan, M. Fulakeza, T. LaRow, P. Lonergan, J.-H. Qian, and S. Zebiak. 2003. International Research Institute/Applied Research Centers (IRI/ARCs) regional model intercomparison over South America. *J. Geophys. Res.* **108**:4425, doi:10.1029/2002JD003201.

Schaaf, C.B., F. Gao, A.H. Strahler, W. Lucht, X. Li, T. Tsang, N.C. Strugnell, X. Zhang, Y. Jin, J.-P. Muller, P. Lewis, M. Barnsley, P. Hobson, M. Disney, G. Roberts, M. Dunderdale, C. Doll, R.P. d'Entremont, B. Hu, S. Liang, J.L. Privette, and D. Roy. 2002. First operational BRDF, albedo and nadir reflectance products from MODIS. *Remote Sens. Environ.* **83**:135-148.

Scurlock, J.M.O., G.P. Asner, and S.T. Gower. 2001. *Worldwide Historical Estimates of Leaf Area Index, 1932-2000.* ORNL Technical Memorandum ORNL/TM-2001/268, Oak Ridge National Laboratory, Oak Ridge, TN (http://www-eosdis.ornl.gov/VEGETATION/lai_des.html).

Sellers, P.J., S.O. Los, C.J. Tucker, C.O. Justice, D.A. Dazlich, G.J. Collatz, and D.A. Randall. 1996. A revised land surface parameterization (SiB2) for atmospheric GCMs. Part II: The generation of global fields of terrestrial biophysical parameters from satellite data. *J. Climate* **9**:706-737.

Smith, W.H.F., and D.T. Sandwell. 1997. Global sea floor topography from satellite altimetry and ship depth soundings. *Science* **277**(5334):1956-1962.

Smirnova, T.G., J.M. Brown, and D. Kim. 2000. Parameterization of cold-season processes in the MAPS land-surface scheme. *J. Geophys. Res.* **105**:4077-4086.

Steele, M., R. Morley, and W. Ermold. 2001. PHC: A global ocean hydrography with a high-quality Arctic Ocean. *J. Climate* **14**:2079-2087.

Takle, E.S., W.J. Gutowski Jr., R.W. Arritt, Z. Pan, C.J. Anderson, R.R. da Silva, D. Caya, S.-C. Chen, F. Giorgi, J.H. Christensen, S.-Y. Hong, H.-M.H. Juang, J. Katzfey, W.M. Lapenta, R. Laprise, G.E. Listen, P. Lopez, J. McGregor, R.A. Pielke Sr., and J.O. Roads. 1999. Project to intercompare regional climate simulations (PIRCS): Description and initial results. *J. Geophys. Res.* **104**:19443-19461.

Taylor, K.E., D. Williamson, and F. Zwiers. 2000. The sea surface temperature and sea ice concentration boundary conditions for AMIP II simulations. PCMDI Report No. 60, UCRL-JC-125597:24 pp. (http://www-pcmdi.llnl.gov/pcmdi/pubs/pdf/60.pdf).

Thiébaux, J., E. Rogers, W. Wang, and B. Katz. 2003. A new high-resolution blended real-time global sea surface temperature analysis. *Bull. Amer. Meteor. Soc.* **84**:645-656.

Verdin, K.L. 2003. A System for Topologically Coding Global Drainage Basins and Stream Networks (http://edcdaac.usgs.gov/gtopo30/hydro/P311.asp).

Webb, R.S., C.E. Rosenzweig, and E.R. Levine. 1993. Specifying land surface characteristics in General Circulation Models: Soil profile data set and derived water-holding capacities, *Global Biochemical Cycles* **7**:97-108.

Wei, H., W.J. Gutowski Jr., C.J. Vorosmarty, and B.M. Fekete. 2002. Calibration and validation of a regional climate model for Pan-Arctic hydrologic simulation. *J. Climate* **15**:3222-3236.

Xue, Y., and J. Shukla. 1993. The influence of land surface properties on Sahel climate. Part I: Desertification. *J. Climate* **6**:2232-2245.

Zeng, X., R.E. Dickinson, A. Walker, M. Shaikh, R.S. DeFries, and J. Qi. 2000. Derivation and evaluation of global 1-km fractional vegetation cover data for land modeling. *J. Appl. Meteor.* **39**:826-839.

Zeng, X., P. Rao, R.S. Defries, and M.C. Hansen. 2003. Interannual variability and decadal trend of global fractional vegetation cover from 1982 to 2000. *J. Appl. Meteorol.* **42**:1525-1530

Zeng, X., M. Shaikh, Y. Dai, R.E. Dickinson, and R. Myneni. 2002. Coupling of the Common Land Model to the NCAR Community Climate Model. *J. Climate* **15**:1832-1854.

Zhou, L., R.E. Dickinson, Y. Tian, X. Zeng, Y. Dai, Z.-L. Yang, C.B. Schaaf, F. Gao, Y. Jin, A. Strahler, R. B. Myneni, H. Yu, W. Wu, and M. Shaikh. 2003. Comparison of seasonal and spatial variations of albedos from Moderate-Resolution Imaging Spectroradiometer (MODIS) and Common Land Model. *J. Geophys. Res*. **108**(D15):4488, doi:10.1029/2002JD003326.

Zhou, L., C.J. Tucker, R.K. Kaufmann, D. Slayback, N.V. Shabanov, and R.B. Myneni. 2001. Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999. *J. Geophys. Res.* **106**:20069-20083.

Zobler, L. 1986. A world soil file for global climate modeling. *NASA Tech. Memo.* 87802:pp. 32-14, Washington, DC.

Illinois State WATER Survey (1895)