Characteristics of global and regional drought, 1950–2000: Analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle

Justin Sheffield and Eric F. Wood

Received 28 November 2006; revised 16 May 2007; accepted 31 May 2007; published 15 September 2007.

Drought occurrence is analyzed over global land areas for 1950-2000 using soil moisture data from a simulation of the terrestrial water cycle with the Variable Infiltration Capacity (VIC) land surface model, which is forced by an observation based meteorological data set. A monthly drought index based on percentile soil moisture values relative to the 50-year climatology is analyzed in terms of duration, intensity and severity at global and regional scales. Short-term droughts (<= 6 months) are prevalent in the Tropics and midlatitudes, where inter-annual climate variability is highest. Medium term droughts (7–12 months) are more frequent in mid- to high-latitudes. Long term (12+ months) droughts are generally restricted to sub-Saharan Africa and higher northern latitudes. The Sahel region stands out for having experienced long-term and severe drought conditions. Severe regional drought events are systematically identified in terms of spatial coverage, based on different thresholds of duration and intensity. For example, in northern Europe, 1996 and 1975 were the years of most extensive 3- and 12-month duration drought, respectively. In northern Asia, severe drought events are characterized by persistent soil moisture anomalies over the wintertime. The drought index identifies several well-known events, including the 1988 US, 1982/83 Australian, 1983/4 Sahel and 1965/66 Indian droughts which are generally ranked as the severest and most spatially extensive in the record. Comparison with the PDSI shows general agreement at global scales and for these major events but they diverge considerably in cooler regions and seasons, and especially in latter years when the PDSI shows a larger drying trend.


1. Introduction

Drought is a pervasive climate phenomenon that is considered to be one of the most damaging natural hazards in terms of economic cost [Wilhite, 2000]. It can cover extensive areas and last from months to multiple years and may have major impacts on agriculture, water supply and the environment. Historically drought has persistently affected human activity [e.g., Hodell et al., 1995, Stine, 1994] and impacts in every part of the globe in which habitation is possible. Despite its omnipresent nature our knowledge of the onset, development and recession of drought is deficient. This hampers our ability to predict its occurrence at seasonal and longer timescales. Part of the reason for this is the dearth of detailed data about its spatial and temporal variability across large scales and the impact on various environmental and social sectors.

[3] To understand how drought varies, long-term observations of relevant variables, such as precipitation, streamflow, and soil moisture, are required. Global data sets of these variables are lacking at high spatial resolution or are available only for limited time periods. Alternatively, models can provide spatially and temporally consistent fields of these variables at large scales when forced with observed boundary conditions and can be used for prediction at seasonal to decadal timescales when run in forecast mode. Furthermore, atmosphere-ocean general circulation models (AOGCM) can be used to study decadal variability in drought when run in coupled mode, and may provide insights into the forcing mechanisms of historic drought events, such as the influence of sea surface temperature patterns [Hoerling and Kumar, 2003].

[4] Historically, the Palmer Drought Severity Index (PDSI) [Palmer, 1965] has been the tool of choice when monitoring and analyzing drought occurrence. At continental to global scales its simplicity makes it an attractive choice for reconstructing drought records [Cook et al., 1999; Dai et al., 2004]. However, it has been shown to be unsuitable for widespread application and suffers from simplifications in its physical basis [Alley, 1984; Heim,
2002]. With the emergence of physically based models over the last decade that simulate the detailed processes of energy and water transfer at the Earth’s surface, including detailed soil moisture transport and snow processes, the potential for more accurate drought monitoring is evident [Sheffield et al., 2004a]. Coupled with the growing availability of remote sensing products and detailed meteorological data at fine time and space scales to force these models with, both retrospectively and in real time, there is the potential for analyzing drought variability historically and for monitoring at regional and global scales.

[5] Drought in its various forms has been analyzed at large scales by many authors in recent years. Sheffield et al. [2004a] used an approach based on simulated soil moisture data to analyze the spatial and temporal extent of national and regional drought over the US. Similarly, Andreadis et al. [2005] used severity-area-duration curves to investigate major US droughts over the 20th century. At similar scales, van der Schrier et al. [2006a, 2006b] calculated summer PDSI over North America and Europe for 1901–2000 and identified the 1930s and 1950s as the driest periods of the record over North America and the late 1940s to early 1950s over Europe. Globally, Dai et al. [2004] calculated PDSI data from 1870 to 2002 and concluded that precipitation and temperature trends modulated by ENSO activity as the leading cause of variability. Peel et al. [2004, 2005] looked at 3863 precipitation station and 1236 streamflowgauge records globally to analyze the distribution of drought duration, magnitude, and severity and found that both precipitation and streamflow showed similar distributions of drought duration globally except for the Sahel. McCabe and Palecki [2006] analyzed decadal variability in global PDSI and sea surface temperatures.

[6] In this study, the spatial and temporal characteristics of global drought during the second half of the 20th century are analyzed using soil moisture data from an off-line land surface model simulation. Drought is defined conceptually as a sequence of soil moisture deficits relative to climatology. The simulation is driven by a hybrid forcing data set derived by combining global atmospheric reanalysis with a suite of observational data sets to remove biases and spurious trends in the reanalysis [Sheffield et al., 2006]. The output fields have been validated against observations of the terrestrial hydrologic budget, where available (Sheffield and Wood, Evaluation of retrospective off-line simulation of the global terrestrial water budget, 1950–2000, in preparation). This study focuses on the statistical properties of drought occurrence as derived from this simulation, in terms of duration, magnitude, and severity, and how these vary globally. This definition does not explicitly take into account the impacts of drought but these are generally small for short term droughts and increase with duration and magnitude. By providing a globally consistent picture of drought occurrence that is based on modeled physical processes and observation based boundary conditions, these analyses can help characterize historic droughts and thus form a basis for real time monitoring and prediction, including estimates of drought recovery. An analysis of the long-term variability and trends in drought characteristics over the latter half of the 20th century and the relationship with local and remote forcings will be reported elsewhere by the authors.

[7] The paper is laid out as follows. First, the soil moisture data set is described, including the land surface model and the meteorological forcings. Then methods for deriving the drought index are presented along with statistics for describing the attributes of drought and their temporal and spatial variation. These include the duration, magnitude (deviation from a threshold value), intensity (mean magnitude over the duration) and severity (intensity times duration). A general overview of the variation of soil moisture globally is given next, followed by a global and regional analysis of drought, its characteristics and their inter-relationships. Finally, the data set is compared to the PDSI and the methods presented are used to identify major drought periods over the last 50 years and some examples of these are evaluated within the context of the statistical framework.

2. Data Sets and Methods

[8] The analysis of drought is based on simulated soil moisture data derived from an off-line land surface hydrological model simulation. Soil moisture in the top meter or so provides a useful index of drought in that it provides an aggregate estimate of water availability from the balance of precipitation, evaporation, and runoff processes and takes into account the delays caused by infiltration and drainage, snow accumulation and melt, and the impacts of anomalies in meteorological forcings such as temperature and radiation. In drought terminology, soil moisture falls somewhere in between meteorological drought (a period of precipitation deficit) and hydrological drought (a shortfall in streamflow, reservoir and lake levels, groundwater, etc.) and may be representative of agricultural drought (deficient soil moisture relative to crop and plant demand) through its control on transpiration and thus vegetative vigor. The drought index is calculated as the deficit of soil moisture with respect to the model’s seasonal climatology at a given location [Sheffield et al., 2004a]. This allows us to quantify and compare drought characteristics between locations in a consistent manner. A drought is then defined as a period with a percentile soil moisture value less than a chosen level which represents the drought magnitude. This level reflects the rarity or extremeness of the event. The off-line simulation and the derivation of the drought index and related statistics are described in detail next.

2.1. Off-Line Land Surface Simulation

[9] The Variable Infiltration Capacity (VIC) land surface model [Liang et al., 1994, 1996; Cherkauer et al., 2002] was used to generate spatially and temporally consistent fields of soil moisture and other water budget flux and state variables. The VIC model simulates the terrestrial water and energy balances and distinguishes itself from other land surface schemes through the representation of sub-grid variability in soil storage capacity as a spatial probability distribution, to which surface runoff is related [Zhao et al., 1980], and by modeling base flow from a lower soil moisture zone as a nonlinear recession [Demenil and Todini, 1992]. The VIC model has been applied extensively at regional [Abdulla et al., 1996; Maurer et al., 2002] and global scales [Nijssen et al., 2001; Sheffield et al., 2004b],
including snow and ice dominated regions [Bowling et al., 2003; Su et al., 2006].

Horizontally, VIC represents the land surface by a number of tiled land cover classes. The land cover (vegetation) classes are specified by the fraction of the grid cell which they occupy, with their leaf area index, canopy resistance, and relative fraction of roots in each of the soil layers. Evapotranspiration is calculated using a Penman-Monteith formulation with adjustments to canopy conductance to account for environmental factors. The subsurface is discretized into multiple soil layers. Movement of moisture between the soil layers is modeled as gravity drainage, with the unsaturated hydraulic conductivity a function of the degree of saturation of the soil. Cold land processes in the form of canopy and ground snowpack storage, seasonally and permanently frozen soils and sub-grid distribution of snow based on elevation banding are represented in the model. Seasonally and permanently frozen soils are represented in the VIC model according to the algorithm of Cherkauer and Lettenmaier [1999]. Soil temperatures are calculated using a finite difference solution of the heat diffusion equation for a user-specified number of nodes that are independent of the soil moisture layers. Soil ice content is estimated based on node temperatures and infiltration and base flow are restricted based on the reduced liquid soil moisture capacity.

Previously, soil moisture fields from a retrospective simulation with the VIC model for the USA [Maurer et al., 2002] have been analyzed in terms of drought occurrence by Sheffield et al. [2004a], who found that the simulated soil moisture values were able to represent historic drought events, display coherency and sufficient detail at small space scales, and compare well with standard drought indices such as the PDSI. In snow dominated regions the VIC based data set and the PDSI data set were found to diverge, likely due to inadequate representation of cold season processes in the calculation of the PDSI.

For this study, the VIC model was run globally at 1.0 degree spatial resolution and 3-hourly time step, for the period 1950–2000. Three soil layers were used in the simulation with the top layer being 30 cm thick. The second layer, the main storage layer, was between 0.5 to 1.5 m and the lower layer, which provides moisture for subsurface runoff, was between 0.1m and 0.25m. These two layers are adjusted during the calibration process to result in routed streamflow that satisfactorily match observations at the large basin scale. The values of soil and vegetation parameters and their spatial distribution were specified following Nijssen et al. [2001]. Soil textural information and bulk densities were derived by combining the 5-min Food and Agricultural Organization–United Nations Educational, Scientific, and Cultural Organization (FAO–UNESCO) digital soil map of the world [FAO, 1995] with the World Inventory of Soil Emission Potentials (WISE) pedon database [Batjes, 1995]. The remaining soil characteristics, such as porosity, saturated hydraulic conductivity, and the exponent for the unsaturated hydraulic conductivity equation were based on Cosby et al. [1984]. Vegetation types were taken from the Advanced Very High Resolution Radiometer (AVHRR)-based, 1-km, global land classification of Hansen et al. [2000]. Vegetation parameters such as height, and minimum stomatal resistance were assigned to each vegetation class based on a variety of sources described by Nijssen et al. [2001]. Monthly values of leaf area index were based on Myneni et al. [1997] and were kept constant from year to year.

This simulation was forced by a hybrid data set of meteorological data derived from the National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) reanalysis [Kalnay et al., 1996] and a suite of global observation based products. In effect, the sub-daily variations in the reanalysis are used to downscale the monthly observations. These observations, which are generally available at higher spatial resolution, are concurrently used to downscale the reanalysis in space. Known biases in the reanalysis precipitation and near-surface meteorology were corrected at the monthly scale using observation-based data sets of precipitation, air temperature and radiation. Corrections were also made to the rain day statistics of the reanalysis precipitation which have been found to exhibit a spurious wave-like pattern in high-latitude wintertime. Wind-induced undercatch of solid precipitation was removed using the results from the World Meteorological Organization (WMO) Solid Precipitation Measurement Intercomparison project [Adam and Lettenmaier, 2003]. Other meteorological variables (downward short- and long-wave, specific humidity, surface air pressure and wind speed) were downscaled in space with account for changes in elevation. The forcing data set is described in detail by Sheffield et al. [2006]. The simulation has been validated against available observations of snow accumulation and melt that act at various timescales. For each model grid cell and month, beta distributions are fitted to the 51 monthly values (1 value for each year in 1950–2000) by finding distribution shape parameters that minimize the error between the statistical moments of the simulation sample and that of the fitted theoretical distribution. The current level of drought or wetness for a particular month and point in space can then be gauged relative to this fitted distribution or climatology. A detailed description of these methods is given by Sheffield et al. [2004a] and a summary is given next.

2.2. Analysis of Soil Moisture and Derivation of Drought Index

The drought index is calculated using the method of Sheffield et al. [2004a] and is briefly described here. Simulated soil moisture data at multiple model soil layers are aggregated over the total soil column, converted to volumetric values and averaged to monthly values. The moisture in the total soil column is used as it reflects the totality of modeled processes, including plant transpiration, soil evaporation, infiltration, runoff, base flow and snow accumulation and melt that act at various timescales. For each model grid cell and month, beta distributions are fitted to the 51 monthly values (1 value for each year in 1950–2000) by finding distribution shape parameters that minimize the error between the statistical moments of the simulation sample and that of the fitted theoretical distribution. The current level of drought or wetness for a particular month and point in space can then be gauged relative to this fitted distribution or climatology. A detailed description of these methods is given by Sheffield et al. [2004a] and a summary is given next.

2.2.1. Empirical Moments of Soil Moisture

The statistical characteristics of hydrologic variables, including soil moisture can be best described by L-moments [Stedinger et al., 1993]. Hydrologic variables are generally non-Gaussian and often possess extreme values or outliers which hinder conventional statistical description. The
advantage of L-moments is that they are more robust to the presence of outliers and are able to characterize a wider range of distributions [Hosking, 1990]. However, autocorrelation or trends in the monthly soil moisture data will invalidate the application of L-moments that assumes the variable to be random. The areal extent of statistically significant autocorrelation (0.01 level) in the data is between 12 and 17% depending on the month with about 50% of this area in drier regions (precipitation < 0.5 mm/day) and the majority of the remainder in very high latitudes. Therefore the area that potentially invalidates the assumption of randomness is small and generally restricted to drier regions, such as the Sahara, which we ignore in the analysis. L-moments can be written in terms of linear combinations of probability-weighted moments (PWMs). For values of a random variable $X_j$ ($X_1, X_2, \ldots, X_n$) sorted in decreasing order, unbiased estimators for the first three PWMs are:

\[
\begin{align*}
b_0 &= X \\
b_1 &= \frac{1}{n} \sum_{j=1}^{n-1} (n-j)X_j/n(n-1) \\
b_2 &= \frac{1}{n} \sum_{j=1}^{n-2} (n-j)(n-j-1)X_j/n(n-1)(n-2)
\end{align*}
\]

The L-moments are calculated in terms of PWMs and are defined as

\[
\begin{align*}
\lambda_1 &= b_0 \\
\lambda_2 &= 2b_1 - b_0 \\
\lambda_3 &= 6b_2 - 6b_1 + b_0
\end{align*}
\]

The sample estimates using the L-moments are defined for the first few moments as

\[
\begin{align*}
L_{\text{mean}}(\mu_s) &= \lambda_1 \\
L_{\text{CV}}(\sigma_s/\mu_s) &= \frac{\lambda_2}{\lambda_1} \\
L_{\text{skew}}(\gamma_s) &= \frac{\lambda_3}{\lambda_2}
\end{align*}
\]

where $\mu_s, \sigma_s$ and $\gamma_s$ are the mean, standard deviation and skew, respectively and subscript $s$ indicates the sample estimates of these statistics.

2.2.2. Soil Moisture Probability Distributions

To simulate the continuous variation of soil moisture, beta probability distribution functions (PDF) are fitted to the simulated monthly soil moisture values. The beta distribution can represent a wide variety of shapes and is flexible enough to account for positive and negative skew values, which is necessary given the variation in soil moisture distributions globally. A generalized form of the beta distribution, defined on limits $a$ and $b$, with $a < b$, is:

\[
f(\theta) = \frac{1}{B(b-a)}(\theta-a)^{r-1}(b-\theta)^{s-1}, \quad a \leq \theta \leq b
\]

where $\theta$ is the volumetric soil moisture content; the distribution shape parameters, $r$ and $t$ are constrained by $r > 0$ and $t > r$ and $B = \Gamma(r) \Gamma(t-r) \Gamma(t)$, where $\Gamma(t)$ is the gamma function. For soil moisture, the parameters $a$ and $b$ represent the lower and upper limit on soil moisture, respectively, which are dependent on soil type and climate. The parameters $r$ and $t$ do not have any direct physical significance but determine the shape of the distribution and its moments.

The parameters $a, b, r$ and $t$ were estimated for each grid location and month. Parameter $a$ was estimated by assuming that the first (last) 10% of the sorted soil moisture values are linearly related to its empirical cumulative distribution function. We also investigated setting $a$ to zero and $b$ to soil saturation and this gave similar results for the fitted distributions. To determine $r$ and $t$, once $a$ and $b$ were estimated, the L-moment sample statistics were equated to the corresponding beta distribution moments and a “best fit” solution for $r$ and $t$ was found by minimizing the objective function:

\[
\text{error} = \frac{(\mu - \mu_s)^2}{\mu_s^2} + \frac{(\sigma - \sigma_s^2)}{(\sigma_s)^2} + \frac{(\gamma - \gamma_s)^2}{\gamma_s^2}
\]

using the shuffled complex evolution global optimization algorithm (SCE-UA) of Duan et al. [1993].

2.2.3. Calculation of the Drought Index

Equation 8 was used to estimate the PDF of monthly soil moisture, for each month and grid cell. The VIC drought index is then represented by the quantile, $q(\theta)$, corresponding to a soil moisture value $\theta$, and is determined by integrating the PDF over $(a, \theta)$. The integral of the PDF can be approximated as follows and is used to derive spatial fields of the drought index.

\[
q(\theta) = \int_a^\theta f(\theta)d\theta \approx \sum_{i=1}^{M} f(a + (i-1)\Delta\theta + \Delta\theta/2)\Delta\theta
\]

where $a \leq \theta \leq b$, $M$ is a large integer (1000 in this study), and $\Delta\theta = (\theta - a)/M$.

2.3. Temporal and Spatial Drought Statistics

To characterize the spatial and temporal variation of drought, a number of statistics are developed. First, a drought is defined in general terms as a period of anomalously low soil moisture. In engineering design or water resources management, the anomaly is often described in terms of the deficit below a critical or demand level. For the soil moisture drought index, this is a threshold quantile,
The magnitude of drought is the deficit from this threshold level:

\[ M = q_0(\theta) - q(\theta) \quad (11) \]

[20] In line with other drought indices (e.g., the PDSI), the level of drought can be categorized based on different arbitrary threshold values. These categories are usually referred to as moderate, severe and extreme or similar. In this study, a threshold value of 10% soil moisture quantile is used for the majority of the analysis to discern between a drought and a non-drought. The study of Sheffield et al. [2004a] and sensitivity tests of the threshold value indicate that this value satisfactorily characterizes major drought conditions globally. This value is also comparable to that used by the US National Drought Monitor (http://drought.unl.edu/dm) to denote “severe” drought conditions. Other analyses carried out in this study investigate the continuous variation of the soil moisture and drought magnitude as a function of the threshold value. Recognizing that the meaning and impacts of drought are regionally and sector specific, we also investigate the occurrence of short-term, severe drought and long-term, moderate drought (section 4) as characterized by different values of \( q_0(\theta) \) and duration.

[21] Following Yevjevich [1972], the occurrence of drought can be analyzed using the theory of runs, which has been applied frequently to time series of anomalous hydrologic events, most often in streamflow analysis [e.g., Peel et al., 2004a]. A run is defined as a consecutive sequence of \( D \) data values, in this case soil moisture values \( \theta \), below the threshold \( q_0(\theta) \) that is preceded and followed by at least one value \( q(\theta) > q_0(\theta) \). The cumulative deficit or severity of a run of duration \( D \) starting at time \( t_1 \) is:

\[ S = \sum_{i=t_1}^{t_1+D-1} q_0(\theta) - q(\theta)_i \quad (12) \]

which may also be written as

\[ S = I \times D \quad (13) \]

where \( I \) is the intensity or average magnitude of the run. Thus a run may be termed “severe” if the magnitude is large and/or deficits of any magnitude last for multiple months. This recognizes that the impact of drought is a subtle balance between these two related factors. It should be noted that because the drought index is defined relative to the local climatology, the magnitude of a drought at one location may be higher than at another even though the absolute soil moisture value is higher. The distribution of run durations, within the time series of soil moisture quantiles can give insight into the stochastic or deterministic nature of drought occurrence. Further it can be used to understand the frequency of droughts and the probability of future occurrence, in terms of magnitude or severity. The mean run duration is given by

\[ \bar{D} = \frac{1}{N} \sum_{i=1}^{N} D_i \quad (14) \]

where \( N \) is the total number of runs in the time series. Similarly, higher order statistics can be calculated that describe the spread in the distribution of runs (variance) and the bias toward longer or shorter runs (skew).

[22] To facilitate general comparisons, a number of run duration classes, \( D_i \), are defined as follows:

\[ D_{1-3}, \text{very short-term:} \quad 1 \leq D \leq 3, \quad q(\theta) < q_0(\theta) \quad (15a) \]

\[ D_{4-6}, \text{short-term:} \quad 4 \leq D \leq 6, \quad q(\theta) < q_0(\theta) \quad (15b) \]

\[ D_{7-12}, \text{medium-term:} \quad 7 \leq D \leq 12, \quad q(\theta) < q_0(\theta) \quad (15c) \]

\[ D_{12+}, \text{long-term:} \quad D > 12, q(\theta) < q_0(\theta) \quad (15d) \]

The total number of runs of a particular duration class can then be calculated over the time series and the frequency of occurrence over a defined period (e.g., number of medium-term runs per 30 years) used to compare with other locations.

[23] All the above statistics can be calculated for time series of soil moisture at a model grid cell. It is of interest to know the global variation of these statistics across climate and land cover zones and their spatial correlation structure. The latter is important for understanding how the extent of drought develops and decays over space. The spatial extent of deficits (for a particular value of \( q_0(\theta) \)) over a region is defined as the ratio of the area in deficit to the total area of the region:

\[ A = \frac{\sum_{i=1}^{N_{\text{grids}}} A(i)}{\sum_{i=1}^{N_{\text{grids}}} A(i)} \quad (16) \]

where \( A(i) \) is the area of grid cell \( i \) weighted by the cosine of the grid cell latitude and \( N_{\text{grids}} \) is the total number of grid cells in the region of interest. We are also interested in the spatial extent of contiguous drought, \( A_{C} \), which is calculated using a simple clustering algorithm based on Andreas et al. [2005]. Drought clusters of less than 10 grid cells (approximately 100,000 km²) are filtered out.

### 2.4. Some Caveats and Uncertainties

[24] Sheffield et al. [2004a] lists some of the potential deficiencies in the approach applied here. These include errors in the model forcing (unknown systematic biases and spurious trends), simplifications and biases in the model physics, uncertainties in the soil and vegetation parameter data, and errors in the fitted soil moisture distributions. Although some of the uncertainties will be masked by spatial and temporal averaging and the use of summary statistics, the compound effect of errors in all stages of the modeling process will inevitably lead to errors in final analysis. Furthermore, the full representation of drought variability over global scales is a difficult task given that knowledge of the variability of even basic climate variables, such as precipitation and temperature, is lacking over much...
of the world at any scale. However, the use of a carefully constructed forcing data set to drive a state of the art and physically based land surface model instills confidence in the soil moisture data (and other water budget variables), which are validated where observations are available. It should also be noted that the analysis does not take into account the direct influence of anthropogenic activities on the water budget such as irrigation and the potential response to climate variability and drought in terms of vegetation dynamics and land use change.

3. Results

3.1. Global Variation of Soil Moisture

The sample estimates of soil moisture statistics are derived using L-moments. Figure 1 shows the global variation in $L_{\text{mean}}$, $L_{\text{cv}}$ and $L_{\text{skew}}$ for January and July, and the seasonal range for each statistic. High mean values of soil moisture occur in the Tropics and follow the seasonal undulation of the terrestrial Inter Tropical Convergence Zone (ITCZ), most notably in central Africa and in Amazonia where the mean soil moisture approaches saturation. The southeast Asian Monsoon is reflected in the soil moisture values, as is wetting up during the North American Monsoon.

In mid-latitudes, soil moisture wets up in the Boreal winter and spring with extensive wetting across Europe and into Russia caused by low evaporation and melting snowpack. At higher northern latitudes, notable wet regions include near annually snow-covered regions of northern Quebec and Newfoundland in Canada and the northern parts of Ob-Yenisey basins in Siberia. Relatively low soil moisture values are spatially extensive and are found in perennially dry regions of the Sahara, the Middle East, central Asia, Australia and southern South America, among others.

The distribution of $L_{\text{cv}}$ values is more interesting as it represents the inter-annual variability and thus gives an indication of the range in soil moisture values and the potential for higher drought frequency. Regions of high variability tend to be collocated or located near to regions of high mean soil moisture. For example, in the Amazon basin the region of high mean soil moisture is surrounded by a band of higher variability likely a result of the annually variable spatial extent of precipitation over the region. Similarly, in Africa, bands of high and low variability coincide with the central part and edges of the ITCZ seasonal footprint. The thin band of high variability over the Sahel shows the vulnerability of this region to wet and dry extremes. In the US, the high wintertime variability...
Figure 2. Time series of regionally averaged soil moisture quantile and aerial extent of drought (total and contiguous) ($q(t) < 10\%$) for 1950–2000. The data are smoothed from the original monthly data with a 13-month moving average.
coincides with the region of high soil moisture in the southeast but is shifted slightly to the west.  

The skewness of the distribution of soil moisture ($L_{skew}$) is highly variable across the world and shows seasonality as indicated in Figure 1. Soil moisture is bounded between wilting point and porosity, and so its distribution will typically become skewed as the mean approaches these boundaries [Western et al., 2002] with positive (negative) skew for dry (wet) soils. In general, northern mid- and high-latitudes show little or no skew in the winter but positive skew in the summer as the soil dries. The latter is indicative of the presence of a few extreme wet years that punctuate the time series, which is confirmed by inspection of individual grid cell histograms. Values tend to be higher in Tropical and sub-Tropical regions, most notably in the Amazon and central Africa, where high seasonality in precipitation forces soil moisture toward dry or wet conditions. Within the Amazon, regions of high and low variability tend to be collocated with regions of extreme skew values. For example, in January low variability is associated with negative skew and high variability with positive skew. In July, the dipole of negative and positive skew to the northeast and southwest, respectively, are some of the highest skew values globally. However, these regions also have low variability indicating that the skewed monthly values are close to the mean and thus relatively unimportant. In Africa, positive skew is also associated with low variability.

3.2. Spatial Extent of Drought

The regionally averaged time series of soil moisture quantile and spatial extent of drought for $q(\theta) < 10\%$ (total and contiguous area) is illustrated in Figure 2. The regions are defined by Giorgi and Francisco [2000] with the addition of the northeast Canada region (NEC), and are shown in Figure 3 and defined in Table 1 in terms of area. These regions are chosen as they cover all land areas, are

Table 1. List of Regions Used in This Study, Including Their Total Area and Maximum Spatial Extent of Drought, $A (q_d(\theta) = 10.0\%)^a$

<table>
<thead>
<tr>
<th>Region</th>
<th>Total Area, $(km^2 \times 10^{12})$</th>
<th>$A_{max}$, (%)</th>
<th>$A_{C,max}$, (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>World</td>
<td>124.9</td>
<td>15.9</td>
<td>12.9</td>
</tr>
<tr>
<td>Northern Europe (NEU)</td>
<td>5.1</td>
<td>46.6</td>
<td>44.4</td>
</tr>
<tr>
<td>Mediterranean (MED)</td>
<td>6.3</td>
<td>38.4</td>
<td>36.6</td>
</tr>
<tr>
<td>Western Africa (WAF)</td>
<td>8.2</td>
<td>46.0</td>
<td>43.0</td>
</tr>
<tr>
<td>Eastern Africa (EAF)</td>
<td>8.6</td>
<td>32.6</td>
<td>28.8</td>
</tr>
<tr>
<td>Southern Africa (SAF)</td>
<td>6.2</td>
<td>50.7</td>
<td>49.8</td>
</tr>
<tr>
<td>Northern Asia (NAS)</td>
<td>14.7</td>
<td>26.3</td>
<td>24.6</td>
</tr>
<tr>
<td>Central Asia (CAS)</td>
<td>6.3</td>
<td>57.5</td>
<td>55.0</td>
</tr>
<tr>
<td>Tibetan Plateau (TIB)</td>
<td>4.7</td>
<td>33.3</td>
<td>30.1</td>
</tr>
<tr>
<td>East Asia (EAS)</td>
<td>8.5</td>
<td>34.6</td>
<td>31.9</td>
</tr>
<tr>
<td>Southeast Asia (SEA)</td>
<td>6.5</td>
<td>47.7</td>
<td>35.3</td>
</tr>
<tr>
<td>Southern Asia (SAS)</td>
<td>5.4</td>
<td>35.1</td>
<td>31.6</td>
</tr>
<tr>
<td>Australia (AUS)</td>
<td>8.6</td>
<td>61.9</td>
<td>59.8</td>
</tr>
<tr>
<td>Alaska (ALA)</td>
<td>3.5</td>
<td>40.5</td>
<td>36.8</td>
</tr>
<tr>
<td>Northeastern Canada (NEC)</td>
<td>4.5</td>
<td>42.9</td>
<td>41.1</td>
</tr>
<tr>
<td>Western North America (WNA)</td>
<td>5.6</td>
<td>39.7</td>
<td>38.9</td>
</tr>
<tr>
<td>Central North America (CNA)</td>
<td>3.4</td>
<td>74.7</td>
<td>74.4</td>
</tr>
<tr>
<td>Eastern North America (ENA)</td>
<td>2.7</td>
<td>53.5</td>
<td>52.3</td>
</tr>
<tr>
<td>Central America (CAM)</td>
<td>3.0</td>
<td>42.9</td>
<td>39.9</td>
</tr>
<tr>
<td>Amazon (AMZ)</td>
<td>12.9</td>
<td>48.2</td>
<td>46.2</td>
</tr>
<tr>
<td>Southern South America (SSA)</td>
<td>5.9</td>
<td>35.2</td>
<td>31.6</td>
</tr>
</tbody>
</table>

$^a$The regions are taken from Giorgi and Francisco [2000] but exclude the Greenland region which has been replaced by the northeastern Canada (NEC) region.
simple in shape, and have been used extensively in previous climate variability and change studies [e.g., Lopez et al., 2006; Kharin and Zwiers, 2005; Giorgi, 2002]. Note that the time series in Figure 2 are smoothed to aid visualization, which will tend to diminish the extreme values. The maximum drought extent values from the original unsmoothed monthly time series are given in Table 1.

Globally, there is little variation in the extent of drought due to the spatial averaging, with slightly less extensive drought in the 1970s. NEU shows higher variability than MED and dry conditions in NEU are at a maximum in the 1950s and the mid 1990s. WAF is dominated by the long-term drought period in the Sahel from the late 1960s to mid 1980s, also reflected but to a lesser extent in EAF. In SAF, there are a number of peaks that cover up to 40% of the region and a particularly extensive wet period in the 1970s. The size of the NAS region tends to dampen the variation in soil moisture and spatial extent although an underlying decadal variability is apparent. Spatially extensive drought conditions in the late 1950s to early 1960s over TIB are followed by increasingly wet conditions until the 1990s. SEA has experienced several periods of large drought extent in the early 1970s and 1980s. The AUS data are dominated by an extensive dry period in the 1950s and 1960s followed by an upward jump in soil moisture around 1973, also seen in the precipitation data. The extent of drought in WNA is relatively uniform but there is greater variability in CNA and a number of events that surpass 60% coverage. The data for the Amazon region are damped by the large area but show peak extents in the 1960s, 1980s and 1990s.

3.3. Run Frequency and Duration

The geographic variation of run duration can be characterized by a summary statistic such as the mean, median or higher moment. Figure 5 shows the total number of runs of any duration and the statistics (mean, standard deviation and skew) of run durations calculated using \( q_0(\theta) = 10\% \). Note that the total time in deficit is equivalent to the total number of runs multiplied by the mean run length, with an expected value of 10% as defined by the threshold value. The total number of runs is minimal in dry regions. The highest values (>25 runs) occur in more humid regions such as eastern North America, the Amazon

---

**Figure 4.** Frequency of occurrence of very short \((D_{1-3})\), short \((D_{4-6})\), medium \((D_{7-12})\), and long-term \((D_{12+})\) runs \((q_0(\theta) = 10\%)\) for 1950–2000. The units are the number of runs over the whole time period.
and Paraná River basins, Tropical Africa, Europe and southeast Asia. The Sahel, scattered regions across central Asia and high northern latitudes stand out as having high mean run lengths, a result of the high frequency of long-term drought conditions. Regions of higher mean durations are generally collocated with regions of higher variability (standard deviation). Tropical latitudes tend to have the lowest standard deviations. The skew values are generally positive, indicating that the frequency of long duration runs is low.

### 3.4. Run Intensity and Severity

Figure 6 shows the distribution of run intensity and severity as calculated by equations 11 and 12. The data were calculated for a threshold value $q_0(\theta) = 50\%$ to capture the statistics of all events, but the results using other threshold values show similar patterns but different amplitudes. Mean run intensity, $I_{\text{mean}}$, is $\sim 15\text{–}25\%$ (departure below the threshold) for humid regions and $5\text{–}15\%$ for drier regions, explained by the greater frequency of short (long) term droughts in more humid (dry) regions. The eastern United States, central and eastern Europe, Southeast Asia, China and Tropical Africa exhibit the largest $I_{\text{mean}}$ values. The pattern for maximum run intensity, $I_{\text{max}}$, is similar in terms of the general distribution of high and low values. The majority of values are clustered within the range 33–50%, with only the drier regions dropping below about 30%. For mean run severity, $S_{\text{mean}}$, the variation is much more distinct and several regions stand out as having experienced relatively severe conditions, driven more by long durations than by high intensities. $S_{\text{mean}}$ is greater than 200% (cumulative departure below the threshold) over the Sahel, northern Canada, northern Siberia and the Taklamakan desert north of the Himalayas, with $S_{\text{mean}}$ exceeding 500% for parts of the Sahel. Over much of the Tropics, $S_{\text{max}} < 400\%$, although in the northern half of Amazonia, $S_{\text{max}} \approx 800\%$. In the majority of regions elsewhere $S_{\text{max}} \approx 400\text{–}1200\%$. The regions of highest $S_{\text{mean}}$ also have the highest $S_{\text{max}}$. In the Sahel, $S_{\text{max}} \approx 1000$ and reaches 4000% in parts.

### 3.5. Relationship Between Run Duration and Intensity

The relationship between $D$ and $I$ and thus the distribution of $S$ is explored further in Figure 7 for three example regions, WNA, WAF and SAS. Neighboring regions show similar behavior. The figure shows scatter-plots of coincident values of $D$ and $I$ for various values of $q_0(\theta)$. By definition, $I$ cannot exceed the threshold value (i.e., $I < q_0(\theta)$) and this then forms an upper bound on the cloud of points. All plots show a wide range of values of $I$ for short duration runs ($\sim D < 10$), but as $D$ increases, $I$ converges to approximately 50–75% of $q_0(\theta)$, most notably for WNA and for higher $q_0(\theta)$ values, for which runs are more numerous and longer durations ($D > 30$ months) are more probable. At lower thresholds ($q_0(\theta) = 10.0\%$), such a relationship may be equally valid, at least empirically, but is less tractable because of the smaller sample size.

For WAF, some very different behavior is apparent, in addition to the larger range in duration. First, there is a
clustering of points at regular intervals of duration, most notable for higher threshold values. The interval is approximately 12 months and is a result of the strong seasonal cycle in precipitation over this region. The seasonality ensures that long-term deficit conditions \((D > 12)\) can be extended until at least the next rainy season as this is the only time of the year that the deficit can be dissipated. Secondly, the convergence to a small range of \(I\) at higher values of \(D\) is not as obvious for the WAF data. In fact there is considerable variability in \(I\) at high \(D\), especially at higher \(q_0(\theta)\) values. The reasons for this are unclear, although the greater number of long runs over WAF will tend to increase the variability in \(I\).

4. Discussion

4.1. Comparison With the PDSI

[35] The PDSI has been extensively analyzed and applied in research studies [e.g., Dai et al., 2004, van der Schrier et al., 2006a; 2006b] and is arguably the most prevalent drought index in operational use. It is a proxy for soil moisture that correlates well with soil moisture and streamflow observations [Dai et al., 2004], but has been criticized for its hydrologic simplicity and lack of spatially consistency [Alley, 1984]. In this section we compare the soil moisture index with a PDSI data set driven by the same precipitation and temperature forcings and the PDSI data set of Dai et al. [2004]. To calculate the PDSI we use the self-calibrating algorithm of Wells et al. [2004] that removes the spatial inconsistency. At global scales the soil moisture index correlates well with both PDSI data sets (Figure 8) and all show decreasing tendencies since the mid-1970s. The soil moisture index shows greater month to month variability compared to the PDSI data sets. This has been attributed, in part, to the PDSI ignoring the daily variation of precipitation and the effects of snowmelt, and the use of time invariant vegetative cover, which will tend to dampen the index over seasonal scales [Sheffield et al., 2004a]. At the grid scale, the data sets are generally well correlated but tend to diverge in cooler seasons and high latitudes and substantially so in dry regions (Figures 9 and 10), which is consistent with Sheffield et al. [2004a] who analyzed a similar soil moisture data set generated at high resolution over the contiguous USA.

4.2. Identification of Severe Drought Events

[36] Figure 2 gives a general overview of the monthly spatial extent of drought on a regional basis but does not take into account the severity (combined intensity and duration) which are relevant for evaluating the impacts. Figure 9 shows regional time series of the spatial extent of drought in a similar manner to Figure 2, but filtered for high severity to reveal drought events that were either of long duration or high intensity. The filtering is applied at two scales: using a moving window of 3 month duration and 10% threshold \((D = 3, q_0(\theta) = 10.0\%); \) short duration, high intensity droughts and using a 12 month duration and 50% threshold \((D = 12, q_0(\theta) = 50.0\%); \) long duration, low intensity droughts. The filtering process leaves only the severe events and could have been implemented in any
number of ways using different combinations of duration and intensity. Nevertheless, Figure 9 reveals those events that are both severe and spatially extensive as derived from this data set. In the following discussion, short duration, high intensity events are referred to as short-term. Long duration, low intensity events are referred to as long term. It should be noted that the spatial extent calculated here is not necessarily contiguous and this is more likely the case for larger regions as drought can occur in multiple disconnected locations at the same time.

Over North America, the 1950s were the decade of most spatially extensive and prolonged drought in the west and central regions, as expected, and would only be surpassed by the 1930s drought during this century [Andreadis and Lettenmaier, 2006]. In terms of individual years, the winter drought of 1976/77 was considerably more extensive than the drought of 1988 that was purported to be the worst natural disaster in US history [Trenberth and Branstator, 1992].

The most extensive droughts in CAM occurred in the 1950s, which is consistent with reported conditions [Liverman, 1999] and with gauge based precipitation records that show an increasing trend since the early 1960s [Aguilar et al., 2005]. In the Amazon, maximum extent coincides quite satisfactorily with El Nino events, including 1957/8, 1972/3, 1992, and 1997, which are associated with dry and warm conditions, especially in the northern part of the region [Foley et al., 2002]. In southern South America, 1962 and 1988 stand out as years of spatially extensive long term drought. These years coincide with La Nina events which are known to cause dry conditions in the east of this region [Boulanger et al., 2005], although curiously other La Nina events are not reflected in the drought record which requires further investigation.

The series for NAS is interesting for two reasons. First, the extent of severe drought is much greater since the mid-1970s for short term droughts, possibly a result, in part,
of the switch to a positive NAO phase [Visbeck et al., 2001]. Secondly, the peaks tend to occur at regular intervals. Closer inspection of the data shows that these occur in the autumn and winter. As the soil column freezes, the soil moisture at any location tends to remain fairly constant over the winter period, which will tend to prolong any preceding drought conditions. CAS shows major events in the early 1950s and at the end of the record, the latter analyzed by Barlow et al. [2002], with a spate of less extensive events in the 1970s. For TIB, the series of short-term drought is dominated by a prolonged event during the late 1950s to early 1960s, although peak spatial extent occurs in 1997 (short term) and 1994 (long term). In EAS, which covers most of eastern China and Japan, spatially extensive severe drought is limited at the short term but peak coverage for long term drought is evident in 1951, 1968, and 1992. SAS shows a few major events which tend to occur in the middle part of the period. In SEA there are distinct major events in 1972, 1982/3, and 1997 at both timescales (short and long term) that coincide with major El Nino episodes. For AUS, drought events are more extensive in the earlier part of the period most notably in 1952 and 1965 for short term droughts and multiple years in the 1950s and 1960s for long term droughts, which may be due to changes in the influence on ENSO on Australian climate since the 1970s [Nicholls et al., 1996].

For Europe, NEU has far more spatially extensive droughts than the MED region. Spatially extensive short term droughts in NEU occur in 1954 and 1996. Other drought events, such as that in 1975/6, are less extensive but span multiple 3-month periods, and 1975 is the most extensive long term drought. 1954 is the most extensive short term drought year in the MED region, yet this stands alone in the first half of the period and drought appears to be much more extensive in the second half for both timescales, especially in the late 1980s and 1990s. This is generally consistent with Lloyd-Hughes and Saunders [2002] and van der Schrier et al. [2006b] who found that the 1950s and 1990s were the most drought prone periods across the whole of Europe in terms of PDSI and 3 and 12 month SPI.

WAF is dominated by prolonged drought in the 1970s and 1980s [Hulme, 1992; L’Hôte et al., 2002] that reached peak extent in 1984 and 1972. This is reflected in EAF but the spatial coverage is considerably lower. In SAF, peak events occur in 1970 and 1992 for short term drought and 1992 for long term drought which corresponds well with the SPI analysis of Rouault and Richard [2005] who determined that the majority of major drought events were coincident with El Nino episodes.

### 4.3. Analysis of Selected Major Drought Events

Several major drought events are selected for further analysis. These are chosen because of their documented socio-economic impacts and the large number of studies carried out into their origins and dynamics. We are interested in whether the hydrologic conditions, as represented by the soil moisture index, are good indicators of such events and how this compares with the PDSI. Other major droughts have been identified in the previous section but may not have impacted as greatly because of a sparse population or low agricultural activity, or may be less well reported. Snapshots of the spatial distribution of soil moisture quantiles for four selected major historic regional drought events are shown in Figure 11: USA, 1988; the Sahel, 1983-84; India, 1965–66; and Australia, 1982–83. These four droughts are put in the context of the statistical analysis presented in section 2 by determining the regional extent of drought for \( q(\theta) = 10.0\% \) and the mean regional intensity and duration for \( q(\theta) = 50.0\% \) on a seasonal basis. We compare this with the other years in the record and the
results from the two PDSI data sets in Figure 12. The maximum values and years in which they occur are summarized in Table 2.

4.3.1. USA 1988

[44] The drought of 1988 over the central United States is estimated to have cost $30 billion in agricultural losses alone and has been considered to be the worst natural disaster in U.S. history [Trenberth and Branstator, 1992]. It is generally agreed that a combination of sea surface temperature (SST) anomalies, persistent stationary atmospheric circulation anomalies and soil moisture-precipitation feedbacks were key factors in the development and longevity of the drought [Sud et al., 2003]. Drought conditions, as derived from the soil moisture data, developed through the spring and summer of 1988, reaching a peak spatial extent in June 1988 (65.5% at \( q_{d}(θ) = 10.0\% \), region CNA) that was only exceeded by conditions in October 1952 (74.7%). In the context of the full time period (Figure 12), 1988 is ranked 1st, in terms of summer (JJA) drought extent, with 1980 and years of the early to mid 1950s ranked next. In terms of regional severity, defined as the regional intensity multiplied by the regional duration, 1988 is ranked 2nd with 1980 being the exceptional year by this definition. 1977 is ranked highly in annual terms, but dry conditions were manifested in the winter months (peak extent in January 1977 of 59.9%). The PDSI data sets exhibit similar spatial

---

**Figure 9.** Correlation between the VIC soil moisture index and the PDSI\textsubscript{VIC} data set at monthly, annual and seasonal timescales.
Figure 10. Time series of regional spatial extent of short (black bars) and long timescale (red bars) drought events. Short timescale droughts are defined as soil moisture quantile $q_0 < 10\%$ over a 3 month moving window $\{q_0(\theta) = 10.0, D = 3\}$. Long timescale droughts are defined as soil moisture quantile $q_0(\theta) < 50\%$ over a 12 month moving window $\{q_0(\theta) = 50.0, D = 12\}$. Note that the $y$ axis scales differ between regions.
extent to the VIC index for 1988, but show 1956 to have a much greater spatial extent, also found by van der Schrier et al. [2006a] (see their Figure 6).

4.3.2. Sahel 1983-84

Long-term drought conditions in the Sahel region of Africa during the 1970s and 1980s had devastating social and environmental consequences [Mortimore and Adams, 2001; Tarhule and Lamb, 2003]. Decreasing precipitation trends over the region have been well documented [Hulme, 1992; L'Hôte et al., 2002] but the forcing mechanisms have been the subject of debate. Previously, it was thought that overuse, in the form of overgrazing, deforestation and poor land management, was responsible. However, recent studies have shown that a combination of land-atmosphere interactions [Nicholson, 2000], ocean temperatures [Giannini et al., 2003] and anthropogenic forcing [Held et al., 2005] are likely causes. From the soil moisture data, the highest monthly spatial extent of drought over the Sahel region (10–20N, 20W–20E) occurred during September 1984 (68.7%). For the growing season average (May–October), 1984 also has the highest spatial extent (38.9%) closely followed by 1987 (38.5%). The PDSI spatial extent values are similarly ranked, although they are much higher (69.5–74.2%). The scatterplot of duration versus intensity (Figure 12) shows that 1983 was by far the severest drought year, with 1984 ranked 6th and showing particularly high intensities but relatively lower durations. For the encompassing WAF region, the maximum extent of drought conditions occurred during October 1983 (46.0%) and 18 out of the top 20 ranked months of spatial extent occurred in the 1980s.

4.3.3. India 1965-66

India has experienced a multitude of severe droughts over the last century, which are generally forced by inter annual variability in the Indian monsoon during June to September [Krishnamurthy and Shukla, 2000]. In the last 50 years, the droughts of 1965/66, 1972 and 1987 have been the most widespread and damaging [OFDA/CRED, 2006].
Figure 12. Regional time series of spatial extent of drought and scatterplots of average run duration versus intensity for the CNA, Sahel, India and east AUS regions. The spatial extent data calculated for $q_0(\theta) = 10\%$ and are shown for the VIC index, PDSI$_{VIC}$ and PDSI$_{DAI}$ data sets. The PDSI data sets are first transformed into quantile space. Each point in the scatterplots represents the average duration and intensity of runs for within each season averaged over all grid cells within the region. Selected historic drought event years are highlighted.
Conditions in 1965–66 were particularly devastating because of two consecutive years of drought. The soil moisture data reveal that the maximum monthly spatial extent of drought over India at \( q_{0}(0) = 10.0\% \) was 35.3\% in November 1965, although 1987 had five months with higher values up to 46.7\%. The PDSI data sets are generally in agreement but show slightly larger values. Notably the PDSI\(_{VIC}\) data set ranks the 1965 event highest, which may be due, in part, to the coarser spatial resolution and different source for the forcings. Figure 12 shows that, in terms of regional severity, 1965 and 1966 are not ranked highly within the full time period. However, this does not take into account multiyear droughts that spanned 1965–66 and it is conceivable that the overall impacts were compounded by consecutive years of severe drought conditions.

### 4.3.4. Australia 1982-83

[47] Australia has experienced multiple drought events over the 20th century that are generally forced by variability at inter-annual and inter-decadal scales associated with El Nino [Chiew et al., 1998] and modified by Pacific inter-decadal variability [Power et al., 1999]. The strong El Nino of 1982/83 forced drought conditions that affected much of Australia during this time. Figure 11 shows a map of soil moisture quantiles in February 1983 indicating severe drought conditions in eastern Australia, the culmination of record low precipitation from July 1982 to February 1983. Drought conditions \( q_{0}(0) = 10.0\% \) covered more than 45\% of the east AUS region on average from September 1982 to February 1983, with the maximum spatial extent of 62.1\% during February 1983. Mean regional drought duration during the 1982/83 warm season (SONDIF) was 2.7 months and mean intensity was 27.5\% and is ranked highly relative to the whole period. In terms of spatial extent, 1982/83 is easily the highest ranked year according to the VIC index and both PDSI data sets. Like other regions, the PDSI data sets tend to give larger values of spatial extent.

### 5. Summary and Conclusions

[48] A monthly soil moisture based drought index is developed for global terrestrial areas, excluding Greenland and Antarctica, from an off-line land surface model simulation forced by an observation based meteorological data set. The index is used to investigate the occurrence and variability of drought globally over 1950–2000. Drought is described in terms of duration, intensity and severity, and various statistics that summarize their distributions in time and space. These variables are analyzed spatially, at global and regional scales, and temporally with respect to severity and spatial coverage. The inter-dependence of these correlated variables is also explored along with the sensitivity to the threshold soil moisture value that defines drought.

[49] An analysis of the statistics of drought events reveals considerable global variability and some interesting relationships between drought characteristics. Based on a soil moisture quantile threshold of \( q_{0}(0) = 10\% \), the frequencies of short-term droughts (6 months and less) and droughts of any length are highest in humid regions. Medium term droughts (6–12 months) are more prevalent in mid- to high-latitudes, which for the latter is a result in part of freezing temperatures causing static soil moisture conditions and forcing drought conditions to persist through the wintertime. Over the Sahel and parts of high northern latitudes, the frequency of long-term droughts is at a maximum.

[50] Drought intensity is defined as the mean departure below the threshold soil moisture quantile over the drought duration and tends to be higher over humid regions. This is likely a result of the higher inter annual variability in soil moisture that tends to prevail in humid regions, even if the range in soil moisture is small in absolute terms. Drought severity, calculated as intensity times duration, tends to be lowest in more humid regions and is highest in regions of high mean duration, such that drought duration is a more dominant factor in severity for longer duration droughts. The Sahel region stands out globally for having long-term and severe drought conditions.

[51] The relationship between duration and intensity, and thus the distribution of severity, is of particular interest as this governs the impacts of drought. A more detailed analysis of the joint and conditional distributions of duration, intensity and severity [e.g., Kim et al., 2003] would be required to quantify the relationship between these highly correlated drought attributes, but some general remarks can be made nevertheless. The data shows that for some regions,
the drought intensities will tend to converge to a small range of values at higher duration. This is consistent with the possibility that drought duration dominates severity at longer durations, as discussed previously. Elsewhere, strong seasonality in a region’s climate may result in a wider spread of intensity values and cluster the distribution of long-term drought durations into annual and multiannual lengths.

Severe drought events are systematically identified in terms of spatial coverage for various regions based on different thresholds of duration and intensity that relate to either high intensity, short duration droughts or low intensity, long duration droughts. For example, in northern Europe 1975 was the year of most spatially extensive drought at annual timescale and 1996 was the equivalent year at 3-month timescale. In northern Asia, severe drought events at short and long timescales are characterized by persistent soil moisture anomalies over the wintertime. Droughts in western and eastern Africa are dominated by events in the Sahel.

The drought index identifies several well-known drought events, including the 1988 USA, 1982/83 Australian, 1983/4 Sahel and 1965/66 Indian droughts, which are analyzed in more depth. These are generally ranked as the severest events in the record, although some are ranked relatively low and the severity of their reported impacts is likely compounded by socio-economic and other factors. Comparison of the results with those from two PDSI data sets shows general agreement, although the PDSI tends to give larger spatial extent values. Some events, however, (e.g., 1988 CNA and 1965 India) are ranked somewhat differently by each data set that may be due to differences in scale and forcings, but is also likely a result of fundamental differences in the modeling approach between the VIC index and the PDSI. At global scales the VIC index and the PDSI are reasonably well correlated but this breaks down in cooler regions and seasons, and especially in the latter half of the 20th century, when the PDSI shows a larger drying trend. Given these comparisons, the known deficiencies and simplifications in the PDSI and the history of evaluations of the VIC model, we conclude that the VIC index is a good indicator of major drought events that is applicable to a wider range of climate regimes than the PDSI.

The overall analysis is inevitably subject to errors in the representation of the actual variation in soil moisture and drought occurrence, which are listed in section 2.3 and Sheffield et al. [2004a]. However, the soil moisture based index can provide a useful indicator of drought given its physical basis and consistent global coverage. The utility of a drought index is nevertheless, measured by how well it describes the development and recession of drought and whether this information can be used to monitor and react to drought. As the impacts of a certain magnitude and duration of soil moisture deficit are particular to the region, time of year and sector it is impossible to specify whether drought is occurring as a simple yes or no. Rather it is more useful to present the actual deficit and duration and let the user decide the implications of this.

This data set forms a climatology that provides a useful benchmark against which current and potential future changes in drought can be assessed. An increase in the number of droughts and/or drought severity is a possible outcome of future global warming and intensification of the water cycle [Wetherald and Manabe, 1999]. Predicted temperature rise will lead to increased evaporation and thus reduced soil moisture, yet accompanying changes in precipitation, in terms of totals and statistics, may act to increase or decrease drought occurrence. Projections of precipitation changes are highly scenario and model dependent [Covey et al., 2003] and regional variability [Giorgi and Bi, 2005] compounds the uncertainty in future changes in drought regime. Recent increases in global temperatures may already have acted to change the global drought regime [Dai et al., 2004]. An analysis of changes and trends in various drought statistics over the second half of the 20th century using this data set is work in progress that will also examine the processes that force and modulate the temporal and spatial variation of drought at large scales.

Acknowledgments. This work has been supported by NOAA grants NA86GP0248 and NA0303AR410001, NASA grant NAG5-9486 and DOE grant 983338. The VIC soil moisture data will be made available from http://hydrology.princeton.edu along with other water and energy balance variables from the 50-yr global simulations. We thank the three anonymous reviewers for their comments and suggestions that helped to improve the manuscript.

References


J. Sheffield and E. F. Wood, Department of Civil Engineering, Princeton University, Princeton, NJ 08544, USA. (justin@princeton.edu)