Windowed and Wavelet Analysis of Marine Stratocumulus Cloud Inhomogeneity

STEVEN M. GOLLMER
Dept. Science and Mathematics, Cedarville College, Cedarville, OH 45314

HARSHVARDHAN
Dept. of Earth and Atmospheric Sciences, Purdue University, West Lafayette, IN 47907

ROBERT F. CAHALAN
NASA/Goddard Space Flight Center, Laboratory for Atmospheres, Greenbelt, MD 20771

JACK B. SNIDER
NOAA/ERL, ETL, Boulder, CO 80303

Manuscript Submitted 1994

Corresponding author address: Steven M. Gollmer, Dept. Science and Mathematics, Cedarville College, Box 601, Cedarville, OH 45314-0601

ABSTRACT

In order to improve radiative transfer calculations for inhomogeneous clouds, a consistent means of modeling inhomogeneity is needed. One current method of modeling cloud inhomogeneity is through the use of fractal parameters. This method is based on the supposition that cloud inhomogeneity over a large range of scales is related. An analysis technique named wavelet analysis provides a means of studying the multiscale nature of cloud inhomogeneity. In this paper, we discuss the analysis and modeling of cloud inhomogeneity through the use of wavelet analysis.

Wavelet analysis as well as other windowed analysis techniques are used to study liquid water path (LWP) measurements obtained during the marine stratocumulus phase of the First ISCCP Regional Experiment (FIRE). Statistics obtained using analysis windows, which are translated to span the LWP data set, are used to study the local (small scale) properties of the cloud field as well as their time dependence. The LWP data is transformed onto an orthogonal wavelet basis which represents the data as a number of times series. Each of these time series lie within a frequency band and have a mean frequency which is half the frequency of the previous band. Wavelet analysis combined with translated analysis windows reveals that the local standard deviation of each frequency band is correlated with the local standard deviation of the other frequency bands. The ratio between the standard deviation of adjacent frequency bands is 0.9 and remains constant with respect time. This ratio defined as the variance coupling parameter is applicable to all of the frequency bands studied and appears to be related to the slope of the data's power spectrum.

Similar analyses are performed on two cloud inhomogeneity models, which use fractal based concepts to introduce inhomogeneity into a uniform cloud field. The bounded cascade model does this by iteratively redistributing LWP at each scale using the value of the local mean. This model is reformulated into a wavelet multiresolution framework, thereby, presenting a number of variants of the bounded cascade model. One variant introduced in this paper is the "variance coupled model" which redistributes LWP using the local standard deviation and the variance coupling parameter. While the bounded cascade model provides an elegant two-parameter model for generating cloud inhomogeneity, the multiresolution framework provides more flexibility at the expense of model complexity. Comparisons are made with the results from the LWP data analysis to demonstrate both the strengths and weaknesses of these models.

1. Introduction

The effect of clouds on radiative transfer has been shown to have a significant influence on model results. The magnitude of the radiative forcing effect of clouds has been studied during the Earth Radiation Budget Experiment (ERBE) and the International Satellite Cloud Climatology Project (ISCCP). Comparisons between satellite data and general circulation model results have shown discrepancies which are due in part to how clouds are represented in these models. Currently a cloud layer is modeled as a homogeneous plane parallel feature that occupies a certain percentage of a model grid box. Although this provides a reasonable first-order approximation for the purpose of radiative transfer, significant errors can be introduced by not incorporating cloud inhomogeneity. As shown by Harshvardhan and Randall (1985), cloud layers with identical mean liquid water path (LWP) can have dramatically different albedos. Methods of calculating radiative transfer in inhomogeneous media have been developed by Davies (1978), Stephens (1988), Gabriel et al. (1993), Evans (1993) and others, and have been used to study the magnitude of bias due to cloud inhomogeneity (eg. Cahalan 1989; Lovejoy et al. 1990; Kobayashi 1991). Data collected during the marine stratocumulus phase of FIRE indicate that cloud inhomogeneity reduces cloud albedo by 10% or more (Cahalan et al, 1994a). Though observational errors in cloud albedo often exceed 10%, such a bias can have a noticeable effect on climate model predictions (Cahalan and Wiscombe 1992). Having an even larger impact on remote sensing, the effect of cloud inhomogeneity on radiation thus merits careful study.

The purpose of this work is to model cloud inhomogeneity based on LWP data derived from the surface-based microwave radiometer during the marine stratocumulus phase of FIRE (Snider 1988) with eventual application to radiative transfer calculations. Although analysis of this LWP data has been done by Cahalan et al. (1994a), we extend their work by applying the technique of wavelet analysis. This approach not only provides information about the amount of inhomogeneity present at different temporal and spatial scales, but also provides a multiresolution framework for modeling cloud inhomogeneity. This wavelet multiresolution framework is a generalization of Cahalan's bounded cascade (BC) model and provides a means of changing the statistical properties of the model yet maintaining the multiscale nature of the LWP inhomogeneity.

Recent approaches to the study of cloud inhomogeneity have drawn upon the popular topic of fractals and nonlinear dynamics. The connection of cloud fields to fractal concepts was made by Lovejoy (1982), who observed that cloud fields have a constant perimeter fractal dimension of 1.35 over areas ranging from 1 km^2 to 1000 km^2. Hentschel and Procaccia (1984) linked the fractal dimension to the turbulent diffusion involved in cloud development. Within this framework it is assumed that certain properties are invariant with respect to scale and that it is possible to derive features at one scale knowing the properties of the cloud at a larger scale. The connection between scales does not provide a deterministic representation of the smaller scale, but rather a possible realization of the smaller scale which is statistically consistent with observed large-scale variations. This connection has also been observed in spectra of wind speed and potential temperature analyzed by Gage and Nastrom (1986), who also observed a change in scaling at a few hundred kilometers. The relationship between scales also changes for areas smaller than 1 km^2 (Cahalan and Joseph 1989) where the scaling exponents depend on cloud type and brightness threshold.

Simple macrophysical cloud models have been developed which implement scaling properties. These models are phenomenological rather than based on the equations of fluid flow and thermodynamics, and are designed to give a simple representation of the inhomogeneity within the cloud for purposes of calculating the cloud's radiative properties. Schertzer and Lovejoy (1987) considered a general class of multiplicative cascades. Cahalan (1989) studied the radiative properties of a one-parameter multifractal generated by such a cascade. Davis et al. (1990) focused on the thick cloud limit of a similar scaling model having a single fractal dimension. Using LWP data obtained during the marine stratocumulus phase of FIRE, Cahalan developed a two-parameter fractal model for studying inhomogeneity effects (Cahalan and Snider 1990). This model redistributes LWP within a uniform cloud through an iterative cascade process which generates inhomogeneity at all spatial scales. Here we reformulate this model within a multiresolution framework. This provides a broader framework for modeling cloud inhomogeneity and leads to a number of variants, such as the variance-coupled (VC) model introduced in this paper. These multiscale statistical models generate features at different spatial scales and then combine these features using an inverse wavelet transform.

In order to accurately model the cloud inhomogeneity present at different scales, we focus on correlations between features of different scales as a function of time and frequency (or the analogous function of position and wave number). We first present a conventional "windowed" analysis in which the time-dependent relationship is studied by calculating statistics for data within a window which is translated across the length of the entire data set. The time dependence of the frequency component is studied by performing a Fourier transform on the data located within each analysis window. Although the windowed analysis provides useful information about the temporal and frequency characteristics of the LWP data set, it does not optimize the window size to the frequency being analyzed. An optimized window function balances the localization of temporal information with the localization of frequency information.

Wavelet analysis was designed to optimize the trade-off between temporal and frequency localization. This is done by defining an oscillatory function of finite length.

Convolution of this function with the data set provides a time-dependent analysis of the data over a given band of frequencies. By stretching or compressing this oscillatory function, the mean frequency of the frequency band extracted by the wavelet analysis is decreased or increased respectively. Windowed Fourier analysis can be used to extract similar information; however, the size of the analysis window is independent of the frequency being analyzed. Therefore, the low frequency component, which may consist of only one oscillation within the sampling window, is not sampled to the same degree as the high frequency component, which consists of many oscillations within the sampling window. For wavelet analysis the effective window size changes with the mean frequency of the band being analyzed so that a similar number of oscillations are sampled. The Daubechies wavelets used in this study (Daubechies 1988) have an additional computational advantage over windowed Fourier analysis in that an orthogonal decomposition of the data set is made using an efficient iterative algorithm based on multiresolution analysis. This is not true of a windowed Fourier analysis since the efficiency of the fast Fourier transform is reduced by including a tapered window to the analysis. Reconstruction of the original data is also difficult in a windowed Fourier analysis since overlapping windows are needed to account for the information lost while using the tapered window.

This paper is presented in five sections. Section 2 provides a description of the windowed and wavelet analyses used in this work. Section 3 describes the LWP data set and presents the results of the analyses on this data set. In Section 4 the BC model is recast within the multiresolution framework. Within this context the VC model is introduced and compared to the BC model. Finally, Section 5 discusses the primary observations made during the LWP analyses and their relationship to the modeling of cloud inhomogeneity.

2. Windowed analysis

In this section we describe two approaches to determine the variability within different frequency bands as a function of time and frequency. Since the data are not statistically stationary, the global statistics of the data set are not representative of the statistics at any particular time within the data set. Therefore, it is necessary to select subsets of the data over which the stationarity assumption is approximately satisfied. In the first approach, the time dependence of the data set is studied by translating a window of fixed width over the entire data set and computing statistics as a function of the window location. We will find that these windowed statistics provide useful information about the time dependence of the LWP data set. By performing a Fourier transform at each window location information about the frequency content is also obtained. The second approach uses a different windowing technique named wavelet analysis, which partitions the data into a series of frequency bands. Since the wavelet transform retains time information, the time dependence of variability within each frequency band can be studied. To facilitate comparison between the two approaches, statistics from the result of the wavelet transform are calculated using translated windows matching those of the first approach. This provides common ground for discussing the nature of inhomogeneity present within different frequency ranges.

a. Windowed statistics

The windowed statistics calculated from the LWP data set include the first three moments of the data – the mean, standard deviation, and skewness. Although these statistics do not completely determine the probability density function (PDF) of the data within the analysis window, they do constrain the shape of the PDF and provide useful statistics for modeling inhomogeneity. Since these statistics are calculated for a data window of limited width, they represent a local statistic rather than a global statistic of the data set. Use of the term "mean" for example will imply a "local mean" unless otherwise noted.

The local statistics calculated for the LWP data set use 64-minute windows which are translated by 32-minute increments. Since the sampling rate of the LWP data set is one minute, a window of this width provides a sufficient statistical sample. This window size also restricts the segment of analyzed data to a length which nearly satisfies the stationarity assumption. Changes in the statistics of the LWP data are primarily due to a diurnal cycle, although a synoptic trend is also present. It is assumed that the solar forcing of the cloud layer does not change significantly over the length of one hour and, therefore, a 64-minute window is small enough to ensure stationarity. The validity of the stationarity assumption is studied by adjusting the window width from one half hour to eight hours where the later width clearly violates the stationarity assumption.

The local statistics are also studied using a composite PDF which uses information from all of the analysis windows. A composite PDF is used since binning noise for a 64 point windowed analysis may obscure trends within the PDF. A composite PDF is generated by standardizing the data within each analysis window and then combining all of the standardized data points into a single PDF. A standardized data point, W', is calculated using the equation

$$W' = \frac{W - \mu_i}{\sigma_i} \tag{1}$$

where μ_i and σ_i are respectively the mean and standard deviation of the ith analysis window and W is the original value of the data point. This composite PDF indicates how the local PDF compares to a normal distribution. Since the mean is removed and the standard deviation is set to one, only the skewness and higher order moments affect the shape of this PDF. The usefulness of the composite PDF is limited to the assumption that the higher order moments do not significantly change in sign over the extent of the entire data set. For example, if the data has local positive skewness, the predominant weighting of the composite PDF is to the left of zero (left of the mean). However, if a similar number of data windows having both positively and negatively skewed statistics are combined, the composite PDF may be symmetric about zero indicating a non-skewed distribution. The composite PDFs used in this study are generated using 200 equally spaced bins ranging from –4 to +4 standardized deviations about the mean.

Frequency information is calculated for each window by performing a Fourier transform. In order to avoid problems due to boundary conditions at the edges of the analysis window a Parzen's window is used. Neglecting to use a windowing function affects the relationship between frequencies since there is a false high frequency component introduced into the analysis (Press et al. 1992). The plot of the power spectrum provides a convenient means of studying the relationship between energy and frequency. Since the plot of the power spectrum is log-log, the energy as well as the frequency is normalized to a reference. Therefore, the slope of this plot yields a dimensionless quantity. From turbulence theory a constant negative slope indicates that the energy contained in the large scale features cascades into the smaller scale features. At the smallest scale, the kinetic energy is dissipated away as heat. In our analysis we assume that the inhomogeneity present at each time scale is related to the energy of the power spectrum at its corresponding frequency range. Since the LWP data used in this analysis is a vertically integrated quantity, this inhomogeneity represents the horizontal non-uniformity of such processes as entrainment and vertical motion. For modeling purposes the slope of the power spectrum provides an important parameter for applying scaling properties.

b. Wavelet analysis

Wavelet analysis is used to separate features into different frequency bands while maintaining temporal information. This makes it possible not only to study the relationship between the relative amounts of energy present in each frequency band, but also to study any temporal dependence that might exist in this relationship. This is done by performing an inner product between the LWP data and each member of a wavelet family. Each member of a wavelet family, $\psi_{m,n}(t_n)$, is related to the others by a combination of translations n and scalings (stretching of the function) m. Since neither of these operations affect the relative shape of the wavelet, this analysis extracts features which are similar in shape. This is well illustrated in Weng and Lau (1994) where examples are given for both continuous and orthogonal wavelets.

The wavelet family used for analysis in this paper is based on the D8 wavelet (Figure 1a). This orthogonal wavelet family was developed by Daubechies (1988) and is defined by eight "taps" (Table 1). These taps are used as weighting coefficients to extract the highest frequency content from the data. The transform

$$d_1\left(t_{n/2}\right) = \left\langle \psi_{1,n}(t_n), s_0(t_n) \right\rangle = \sum_{i=0}^{7} g_i s_0(t_n + i) \qquad (n \textit{ is even}) \tag{2}$$

multiplies the eight taps g_i to the data series $s_0(t_n)$ in order to generate the highest frequency detail $d_1(t_{n/2})$ of the data series. The value of n is restricted to even values since the family of orthogonal wavelet transforms are translated two data points with respect to each other. As a result, the series generated by this transform is only half the length of the original data. This transform only analyzes the data at the smallest scale which corresponds to the highest frequency range. When the transform in Equation 2 is applied iteratively to multiple scales of the data, analysis by the discrete values of the taps are equivalent to analysis done by the nearly continuous function seen in Figure 1a (Daubechies 1988). Although the transform behaves as a high-pass filter at the smallest time scale, at any larger time scale the transform performs a band-pass filtering of the data as shown by the Fourier transform of the D8 wavelet in Figure 1b.

In general the taps for Daubechies wavelets are defined such that members of the wavelet family are orthogonal with each other. This is only true if the functions are translated by increments of two data points and the scale is changed by a factor of two. These restrictions allow the data to be expressed as a series of dyadic scales (Weng and Lau 1994). Mahrt (1991) discusses the statistics related to analysis with the Haar function which is the simplest orthogonal wavelet. As with the Haar function, the D8 wavelet preserves energy and, therefore, allows a comparison of statistics between scales. In this study the D8 wavelet is chosen over the Haar wavelet so that leakage between frequency bands is minimized. Since the D8 wavelet is not a perfect filter function, there still exists some frequency leakage represented by the lobes at 3 and 5 Hz in Figure 1b. The improved frequency characteristic is at the expense of some loss of temporal localization.

As seen in Figure 1c, the D8 scale function $\phi_{m,n}(t_n)$ is defined by reversing the order of the D8 wavelet taps and negating each even tap (Table 1). This function performs a weighted sum of the data points by the formula

$$s_1\left(t_{n/2}\right) = \left\langle \phi_{1,n}(t_n), s_0(t_n) \right\rangle = \sum_{i=0}^{7} h_i s_0(t_n + i) \quad (n \text{ } is \text{ } even) \tag{3}$$

which generates the smoothed data $s_1(t_{n/2})$. This data has a resolution which is reduced by a factor of two and, therefore, the transform acts as a low-pass filter as shown by the Fourier transform of the D8 scale function (Figure 1d). As with the orthogonal wavelet transform, the translations are incremented by two, thereby, generating a series which is half the length of the original data set.

Due to the orthogonal nature of the wavelet transform any information lost by the wavelet transform is maintained by its associated scaling transform. Therefore, the two series generated by these two complementary transforms can be used to perfectly reconstruct the original data set through the use of the inverse transform

$$s_0(t_n + k) = \sum_{i=0}^{4} g_{2i+k} d_1(t_{n/2} - i) + h_{2i+k} s_1(t_{n/2} - i) \qquad (k = 0,1) \quad (n \text{ } is \text{ } even). \tag{4}$$

From this notation it is seen that the same taps used in the forward transform are used to determine the weighing of particular values of the smoothed, $s_1(t_{n/2})$, and detail, $d_1(t_{n/2})$, data to perform the inverse transform. It is also evident that the even taps are used for reconstructing the even locations of the data series and the odd taps for the odd locations.

Although a D8 wavelet transform can be defined for scales other than the smallest scale, multiresolution analysis (Mallat 1989) combines the wavelet and scale transforms to iteratively generate the wavelet transforms for larger scales. Since the D8 scale transform decreases the number of data points by a factor of two, this low-pass or averaging transform effectively increases the scale of the resultant coefficients by a factor of two. Therefore, a wavelet transform performed on the resultant components of the scale transform $s_1(t_{n/2})$ is the same as performing the wavelet transform of the next largest scale on the original data $s_0(t_n)$.

Notation for the multiresolution analysis with the D8 wavelet is a generalization of Equations 2, 3, and 4. The scale transform,

$$s_{m+1}\left(t_{n/2}\right) = \left\langle \phi_{m,n}(t_n), s_m(t_n) \right\rangle = \sum_{i=0}^{7} h_i s_m(t_n + i) \quad (n \text{ } is \text{ } even), \tag{5}$$

reduces the resolution of the smoothed data $s_m(t_n)$ obtained from the previous level of the scale transform. Once again, notice that the length of the *(m+1)*th level scale transform is half the length of the *m*th level scale transform. The *m*th level scale transform spans the original data series with $N/2^m$ points since translations are by increments of two and each change in scale increases the translation distance by an additional factor of two. In this case N is the number of points present in the original data series. The wavelet transform,

$$d_{m+1}\left(t_{n/2}\right) = \left\langle \psi_{m,n}(t_n), s_m(t_n) \right\rangle = \sum_{i=0}^{7} g_i s_m(t_n + i) \qquad (n \text{ is even}), \tag{6}$$

removes the detail from the smoothed data obtained from the previous level of the scale transform. This iterative definition for the wavelet and scale transform makes computation very efficient since larger levels of the transform are performed on a reduced length of the data set. The inverse function,

$$s_m(t_n + k) = \sum_{i=0}^{4} g_{2i+k} d_{m+1}(t_{n/2} - i) + h_{2i+k} s_{m+1}(t_{n/2} - i) \qquad (k = 0,1) \qquad (n \text{ is even}), \tag{7}$$

is also defined iteratively and provides perfect reconstruction given that infinite precision is maintained in the transform.

This iterative transform process is illustrated in Figure 2 where the multiresolution analysis is performed on data sampled at one minute intervals to generate smoothed and detail information at the two, four and eight minute scales. The original LWP data set, which has a resolution of one minute, is represented by the rectangle labeled $s_0(t_n)$. Performing the scale transform with the scaling function $\phi_{1,n}$ reduces the resolution (increases the scale) of the data by a factor of two and generates the smoothed data $s_1(t_{n/2})$. The index $t_{n/2}$ is used since the transformed data has only half the number of points of the LWP data while spanning the same time period. Therefore, each data point in $s_1(t_{n/2})$ represents a local two-minute average of the LWP data set. Although an arbitrary position index of n could be used, Figure 2 uses the index $t_{n/2^m}$ to simplify the collocation of features from the scale m, with a location within the original LWP data set. Since the wavelet and scale function complement each other, the detail lost by this reduction in resolution is retained by the detail $d_1(t_{n/2})$ obtained through the wavelet transform. To obtain the next lowest resolution (next largest scale) the multiresolution process is repeated on the transformed data set $s_1(t_{n/2})$ to obtain the smoothed data $s_2(t_{n/4})$, and the detail data $d_2(t_{n/4})$. At this scale features are representative of the four-minute time scale and each transformed data series has only a quarter of the number of points of the LWP data set. This process is repeated once again on $s_2(t_{n/4})$ to obtain both the smoothed, $s_3(t_{n/8})$, and the detail, $d_3(t_{n/8})$, data sets at the eight-minute time scale.

Continuing this processes to the largest scale allows the original data set to be represented by a series of wavelet transforms which span all possible scales. Since the scale transform can be represented by the scale and wavelet transform at the next largest scale, the concept of the scale transform only serves as an intermediate step for iteratively generating the wavelet transform at all scales. At the largest scale a single component represents the scale transform, and its value corresponds to the global average of the LWP data set. Since this analysis is based on an orthogonal decomposition, all information from the LWP data set is retained and it is possible to reconstruct the original data set from the components of the wavelet transform at all scales.

The multiresolution analysis represented in frequency space reveals the trade-off between frequency and temporal localization. Figure 3 illustrates the frequency ranges represented by each scale of the wavelet analysis. At the smallest scale the wavelet transform acts as a high-pass filter where the cut-off frequency is at half of the Nyquist frequency. Since the sample rate of the LWP data is one minute, the Nyquist frequency is 30 hr $^{-1}$ and the cut-off frequency of the first level detail $d_1(t_{n/2})$ is 15 hr $^{-1}$. The length of time represented by each component of the first level detail is two minutes. At the next largest scale each component of the wavelet analysis represents four minutes, while in frequency space these components represent information that ranges between 7.5 and 15 hr $^{-1}$. In a time-frequency phase space representation the product of the time localization and the frequency localization remains constant. Therefore, as the scale of the wavelet analysis increases, lower frequencies are analyzed, the range of frequencies contributing to the analysis decreases, and the number of original data points contributing to the analysis increases.

From this presentation of the multiresolution analysis an obvious limitation exists which affects data analysis. Since the change in scale is limited to factors of two, a translation of the data set by one position will change the representation of the wavelet transform of the data. In addition, if any one frequency has a strong influence on the data set, the wavelet analysis will attribute it to a range of frequencies or, worse yet, to two frequency ranges when the frequency lies near a transition between two scales. Since the scales analyzed by the wavelet transform are powers of two of the resolution of the original data, the analyzed scales cannot be shifted apart from remapping the LWP data into a different sampling rate. Although many of these problems can be resolved by using nonorthogonal wavelets, these problems are not prohibitive to the analysis used in this research. Since statistics are calculated from the wavelet transform, any change in the specific representation does not significantly change the statistical results. The features present in the LWP data used in this research are the result of turbulent processes which do not favor a specific frequency over another. The only predominant frequency present in the data is the diurnal cycle, and any analysis at that scale violates the stationarity assumption. The advantage of retaining the orthogonal condition is that features present in the data cannot be over-represented in the wavelet analysis. This provides grounds for comparing statistics contributed by different scales without a bias due to one frequency having a stronger contribution than originally present in the data.

The correspondence between windowed analysis and wavelet analysis is illustrated in Figure 4. The window Γ_n is chosen to be ~1 hour wide (64 pts) and is translated by

~1/2 hour (32 pts) increments. Although there is overlap between adjacent windows, this permits a larger number of windowed statistics to be generated. Of comparable size to the data windows the fourth level D8 wavelets $\psi_{4,n}$ are translated by only 16 pts. In addition to the difference in translation, the weighting of the data is equal over the window Γ_n while transforms with the $\psi_{4,n}$ wavelets have its largest weighting based on the middle half of the window. Analyzing inhomogeneity of the data over a window Γ_n includes all scales smaller than the window size. However, by using the $\psi_{5,n}$, $\psi_{4,n}$, and $\psi_{3,n}$ wavelet functions inhomogeneity is partitioned into a series of scales which lie in the 32, 16, and 8 minute scales respectively. Although the size of the windowed analysis can be adjusted to these different sizes, it does not provide an orthogonal representation between scales and, therefore, does not provide a simple method for reconstructing the original data set.

c. Wavelet statistics

The important statistics calculated from the wavelet analysis are the standard deviation and the Markov parameter for each level of detail. The wavelet transform partitions the inhomogeneity of the data into a number of time series each associated with a different frequency band. The windowed analysis discussed in Section 2a is applied to each of the wavelet generated time series to calculate a standard deviation. This statistic is equivalent to the square root of Gamage and Hagelberg's (1993) wavelet variance. This scale-dependent standard deviation is calculated for the first level detail $d_1(t_{n/2})$ through the third level detail $d_3(t_{n/8})$. Since each level of detail has half the number of points as the preceding level, the number of points used for calculating the detail standard deviation is halved. For a 64-point window, the standard deviation of the first level detail $d_1(t_{n/2})$ is calculated using 32 points, the second level using 16 points, and the third level using eight points. Due to the limited statistics for calculating the standard deviation of the third level detail, analysis focuses on the first two levels and the relationship between them.

The ratio of *m*th level standard deviation to *(m+1)*th level standard deviation is related to the slope of the power spectrum. Since the wavelet transform separates data into different frequency ranges, it is possible to relate a particular level of the detail to a particular portion of the Fourier spectrum. For data that has a linear power spectrum, the ratio of standard deviation between adjacent levels of the wavelet transform should be constant. This in turn is related to the self-similar or self-affine nature of the data. The relationship between the ratio of standard deviation between levels and slope of the power spectrum is based on observations while conducting our research and still needs to be quantified by a mathematical relationship.

An additional statistic for the *m*th level detail is the Markov parameter. This parameter is equal to the lag one autocorrelation of the *m*th level detail. Since each level of the detail spans a limited frequency band, the Markov parameter measures the correlation of inhomogeneity at a particular scale. The wavelet transform is not a perfect filter and, therefore, it is possible for structures at the next largest scale to generate autocorrelation at the level being analyzed. The presence of autocorrelation at all scales effects the slope of the power spectrum which complicates the mathematical relationship between ratio of variability and slope of the power spectrum. The presence of negative autocorrelation at all scales causes a time series to be more intermittent than one with positive autocorrelation given that the power spectrum is forced to be the same. The LWP data set used in this analysis has intermittent features as reported by Davis et al. (1994) and, therefore, makes the Markov parameter important for modeling applications. The term autocorrelation or Markov parameter is used interchangeably, since only the lag one autocorrelation is used in this study.

3. Vertically integrated liquid water data

a. Description

The data used for this study were obtained during the marine stratocumulus phase of FIRE. The intensive field operation was conducted from 1 July through 19 July 1987, and consisted of ground-based, airborne, and satellite measurements. The ground-based instruments were stationed on San Nicolas Island, which is located 100 km southwest of Los Angeles. One goal of the experiment was to improve the characterization of clouds for radiative transfer studies (Cox et al. 1987). The LWP data are well suited for this since they provide a lengthy record of cloud inhomogeneity and contain changes therein due to time development.

The LWP data used in this study were obtained from a ground-based, upward-looking microwave radiometer. Derived from an algorithm using the 20.6 and 31.65 GHz channels, the LWP data have an absolute accuracy of 20% (Albrecht et al. 1990). Since the retrieval method may contain a bias which affects the absolute amount of LWP, retrieval of variations in the LWP can be known to a greater accuracy of 10% (Cahalan and Snider 1989). Consisting of one minute averages of 10 Hz data, this LWP data set runs nearly continuously for 19 days. Missing data gaps of several hours are present in the data along with two minute gaps approximately every four hours which are used for calibration of the equipment.

Since the LWP measurements are from a zenith view of the sky, the data series records not only time development during the course of the 19 days, but also spatial development as the clouds advect over the instrument. The instantaneous sampled volume for the microwave radiometer is a truncated cone which has an average diameter of 21 m and depth of 100 m assuming an average cloud base of 425 m during the experimental period (Blaskovic et al. 1991). Assuming an average wind speed at cloud height of 5 m/s, the one minute measurements of the microwave radiometer are an average of LWP over a 300 m by 21 m by 100 m section of cloud. Using 64-minute analysis windows, the spatial extent for each window corresponds to a 19.2 km section of cloud. On this time scale the data are assumed to be stationary as discussed in Section 2a and, therefore, variability within the window is assumed to be the result of spatial structure rather than time development of the cloud field.

b. Analysis

The LWP data contain unusable portions due to missing data and clear skies. Using only data from 9 July through 17 July eliminates most of these unusable sections. The bias due to the remaining missing data is reduced by linearly interpolating over these short gaps. Since the greatest variability is found on time scales long compared to the gaps, the interpolation introduces little error. Observations of LWP below a threshold of 10 g/m^2 are assumed to be unreliable due to the bias present in the microwave retrieval algorithm for thin clouds and due to the presence of clear skies. To avoid problems due to poor sampling, any window containing more than 25% clear skies, thin clouds, or missing data is rejected. When calculating the mean, standard deviation, skewness, and PDF, only cloudy points are used within each window. However, when calculating the Fourier transform and any statistic related to the wavelet transform, all data points are used since continuity between adjacent points is important. For these last two calculations, the original values are used since it is expected that thresholding the data will introduce a false high-frequency component to the data.

The mean of the LWP has a strong diurnal cycle as seen in Figure 5a. The mean is greatest just before sunrise and reaches a minimum in the late afternoon. This variation in LWP is related to changes in both the cloud base and cloud thickness (Albrecht et al. 1990). A synoptic trend is also present in the data as demonstrated by the differences in the daily maximums of the LWP.

Although the mean of the LWP has a strong diurnal dependence, the slope of the power spectrum (Figure 5b) is nearly time independent. As discussed in Section 2a, the slope of the power spectrum is related to the frequency dependence of energy and, therefore, is related to the scale dependence of inhomogeneity. There are fluctuations in the value of the power spectral slope; however, there is no obvious trend in this statistic. Using one-hour analysis windows the slope has a mean value of -1.75. This value remains at -1.75 for windows which extend to eight hours in length, although a window of this size violates the assumption of stationarity. For window sizes beyond eight hours the slope of the power spectrum is affected by the diurnal cycle; however, the value of the slope fluctuates about the same global value of -1.75.

The most important result of the LWP's wavelet analysis is the relationship between inhomogeneity at different scales. The degree of inhomogeneity is related to the fluctuations in the LWP and is quantified through the statistic of standard deviation. The standard deviation for the LWP data (Figure 6a) consists of a diurnal cycle with a daily peak value near 1200 UTC which corresponds to dawn. The standard deviation decreases during the day to reach a minimum in late afternoon. Due to the sampling rate of the microwave radiometer the frequencies present in this data set range from 0 to 30 hr^{-1}. The first three levels of the wavelet analysis partition the original data set into the frequency ranges of 15 – 30, 7.5 – 15 and 3.75 – 7.5 hr^{-1}. The standard deviation of these three levels (Figures 6b, 6c and 6d, respectively) corresponds well with the standard deviation of the original data set and with each other. The windowed standard deviation also contains a synoptic feature which has a time period of approximately 7.5 days. This synoptic feature also exists in the standard deviation of the wavelet transform for levels 1, 2, and 3.

The correlation between inhomogeneity at different scales is very strong and is not influenced by the diurnal cycle. The correlation coefficient between level 1 and level 2 standard deviation is 0.81. This implies that 66% of the changes in local standard deviation at the two minute scale is explained by the changes in local standard deviation of the four minute scale. At higher levels of the wavelet transform the correlation coefficient between adjacent levels of standard deviation is also near 0.81. The ratio between level 1 and level 2 standard deviation (Figure 7a) is also time independent, although there may be a dependence on the synoptic scale as indicated by the slight increase in the ratio around days 13 and 14. This lack of time dependence is also observed in the ratio of standard deviation between levels 2 and 3. Another important feature is that the ratio of standard deviation is independent of which two adjacent scales are being compared.

The ratio of standard deviation between scales is approximately 0.9, and is valid up to the ratio of standard deviation between levels 5 and 6. The ratio of standard deviation between levels 6 and 7 drops to 0.7 and is most likely due to the size of the analyzing window. Although the ratio of standard deviation between levels 1 and 2 uses a one-hour window, the window size is doubled at each increase of scale to ensure sufficient statistics. The ratio of standard deviation between levels 6 and 7 depends on a window size of 32 hours which is influenced by the synoptic trend of the data. Each wavelet coefficient within this window is representative of one hour for level 6 and two hours for level 7. Therefore, the reduction in the ratio of standard deviation between scales may be a combined effect of the diurnal and synoptic trends of the data. It is interesting that the ratio of standard deviation between scales remains unaffected by window sizes up to 32 hours, which is well beyond a reasonable stationarity assumption.

As discussed previously, the windowed mean (Figure 5a) has a strong diurnal dependence and this dependence is positively correlated with the windowed standard deviation (Figure 6a) by a correlation coefficient of 0.67. However, on the 2-minute scale the correlation coefficient between the mean and the level 1 standard deviation is 0.56. This means that only 30% of the change in standard deviation at the two minute scale is explained by the mean. This is made evident by taking the ratio between the level 1 standard deviation and the mean (Figure 7b). This ratio is affected by the diurnal cycle and has a minimum of 0.03 at dawn and a maximum of 0.1 in late afternoon. This ratio does not appear to contain a synoptic trend and, therefore, is useful for modeling as long as the diurnal dependence is taken into account. Although the ratio of standard deviation between scales is a better predictor of small scale inhomogeneity, the ratio of standard deviation to mean is useful when information about the inhomogeneity of the LWP is unavailable.

The windowed Markov parameter for the first level of the wavelet transform does not contain a diurnal cycle. Having a mean value of –0.28, this statistic becomes less negative around 13 July where the mean LWP is small. This statistic is also time independent at the second and third levels of the wavelet transform and has mean values of –0.28 and –0.29 respectively. Beyond these levels the statistic becomes less negative and may be affected by diurnal influences. The autocorrelation at each level of the wavelet transform indicates that correlated structures exist at all the analyzed scales. This observation corresponds well with the fractal concepts of self-similarity and self-affinity. Since wavelet analysis restricts the inhomogeneity to a given time scale, the autocorrelation is only indirectly related to larger scale structures within the cloud field. Due to the existence of the Markov parameter it is important that any inhomogeneity model manifest this time independent statistic.

Unlike the statistics discussed previously, the skewness of the LWP is both time and scale dependent without any obvious connection to the time dependence of the mean.

From the one-hour analysis windows the skewness fluctuates between + and −1 and has a mean skewness of 0.1. When the analysis window size increases to two hours, the skewness increases to 0.28. Skewness at this scale is less uniform and appears to have some time dependence to it. Extending the analysis window to four hours gives a global skewness of 0.41 and begins to reveal the diurnal dependence of the skewness. The skewness over this larger window has a weak diurnal dependence which is also related to the mean LWP. When the mean LWP remains relatively constant, the skewness is zero or even negative. However, when the local LWP begins to change, the skewness is positive. This appears to be related to the nonlinear way in which cloud thickness develops. When the cloud LWP increases, it does so at a faster than linear rate. As a result the PDF of LWP over a set interval develops a long, positively skewed tail. Likewise as the cloud decreases in LWP, the dissipation is faster than linear and the distribution of LWP for a set interval is again weighted towards a thinner section of cloud with a few thick sections of cloud.

The dependence of local skewness on analysis window size is best illustrated using the composite PDFs for different window sizes. The composite PDF for a one-hour window (Figure 8a) appears nearly Gaussian when compared to a normal distribution. However, as the window size increases, the largest weighting of the PDF shifts to the left and begins to generate a longer positive tail. This shift in the PDF is the result of violating the stationarity assumption since the statistics of the cloud change significantly over 8 hours due to diurnal effects. For window sizes of eight hours (Figure 8b) and larger, the PDF becomes more log-normal as reported by Cahalan and Snider (1990). Although the contribution of positive skewness is present over the entire data set, a large amount of skewness exists on 17 July where the mean LWP becomes very large over a short period of time.

Since it can be argued that the PDF of the one-hour window is dominated by instrument noise and, thereby, predominantly Gaussian, an alternate composite PDF is generated. The alternate PDF uses only windows which have a standard deviation at least four times as large as the expected standard deviation due only to noise. The skewness dependence of the one, two , four, and eight hour alternate PDFs is the same as the composite PDFs and, therefore, indicates that this observation is not due to noise contamination.

The last statistic to be discussed is the composite PDF for the first level wavelet transform (Figure 9). The wavelet transform is sensitive to inhomogeneity at a particular scale and, therefore, the PDF of the wavelet transform is related to the distribution of inhomogeneity at the corresponding scale. Using a one hour analysis window the composite PDF (Figure 9a) is symmetric about the mean and has a shape similar to a normal distribution. Since the PDF is peaked near the mean, this distribution has positive kurtosis which increases with window size. We interpret this to mean that the contribution of inhomogeneity at a particular time scale is not evenly distributed over the window, but is concentrated to smaller regions which contribute a disproportionate amount of inhomogeneity to the window. This description corresponds to the concept of an intermittent geophysical signal. As the window size increases to eight hours, the localized contributions of inhomogeneity are better sampled and the PDF becomes more kurtose (Figure 9b). An eight hour window violates the stationarity assumption and, therefore, raises the question whether the increase in kurtosis is an indication of cloud development or of better sampled intermittency.

Results from the microwave radiometer LWP analysis appear to be reasonable and provide a foundation for comparison to inhomogeneity model results. Both the ratio of variability and the slope of the power spectrum are time independent and appear to be related to each other, although the relationship has yet to be developed. The correlation of standard deviation between adjacent levels is strong and provides motivation for the development of a wavelet-based cloud inhomogeneity model. This wavelet-based model must also include the Markov parameter since it indicates the presence of correlated structures at all time scales. On the smallest scale the distribution of the LWP inhomogeneity is similar to a normal distribution. Only as the window size includes a significant amount of cloud development does a strong positive skewness appear. This provides a basis for determine the suitability of an inhomogeneity model for a particular application.

4. Inhomogeneity models

The LWP inhomogeneity models used in this paper are multiscale statistical models. Operating within the framework of multiresolution analysis, values for the detail information at each scale are statistically generated and then combined using an inverse wavelet transform. The detail information, $d_m(t_{n/2^m})$, at scale *m*, is dependent on the smoothed information, $s_m(t_{n/2^m})$, of the same scale and the detail information of lower resolutions, $d_{m+1}(t_{n/2^{m+1}})$, $d_{m+2}(t_{n/2^{m+2}})$, ... The values used to initialize the lowest resolution and the relationship between each scale provide the degrees of freedom for fitting the model. Cahalan's BC model uses the values in $s_m(t_{n/2^m})$ to generate $d_m(t_{n/2^m})$ while the VC model uses $d_{m+1}(t_{n/2^{m+1}})$ to generate $d_m(t_{n/2^m})$. Although each model uses a different approach, they still fit within the multiscale statistical framework.

a. Bounded cascade model

The BC model is a two-parameter model that uses an iterative cascading process to generate inhomogeneity. Although the model is more thoroughly described in Cahalan et al. (1994a), a brief discussion is provided here. The BC model begins as a uniform cloud of thickness *W*, which is equal to the mean of the LWP. The first step of the cascade model moves a fraction of LWP, f_0, from one side of the cloud to the other. The direction that the LWP fraction is moved is arbitrary. Upon completion of the first level of the cascade the cloud consists of two regions. The first has a uniform thickness of $W(1-f_0)$, while the second has a uniform thickness of $W(1+f_0)$. The cascade process is repeated by moving a fraction, f_1, from one side to the other for each of the two new regions of cloud. This results in four regions which have the following liquid water amounts: $W(1+f_0)(1+f_1)$, $W(1+f_0)(1-f_1)$, $W(1-f_0)(1+f_1)$ and $W(1-f_0)(1-f_1)$. Each time this process is repeated the number of regions is doubled, thus adding inhomogeneity to the cloud. Since the LWP is being redistributed, the mean LWP is conserved, and only the distribution of LWP is changed at each cascade. If the fraction moved at each cascade is the same, $f_0 = f_1 = f_2 = \ldots$, after an infinite number of cascades the model will consist of an infinite number of singularities with varying intensities. Although this singularity model is truly a self-similar model, it generates results which are nonphysical. This model is made more physically sound by adding a second parameter, *c*, which reduces the fraction moved at each level of the cascade. In this case

$$f_0 = fc^0;\ \ f_1 = fc^1;\ \ f_2 = fc^2;\ \ \ldots;\ \ f_n = fc^n \ \text{ where } \ c < 1. \qquad (8)$$

No singularities are generated for this model since an insignificant amount of LWP is redistributed after a large number of cascades. By adding this damping factor the model is no longer self-similar, but multiaffine since there is a rescaling of the amount of LWP moved at each step of the cascade (Marshak et al. 1994). This damping factor directly affects the amount of variability at each scale and, therefore, affects the slope of the power spectrum. The PDF of LWP generated by the BC model is log-normal in shape.

The multiscale relationship of the BC model is illustrated in Figure 10 using the multiresolution framework. The initial uniform cloud LWP, *W*, corresponds to $s_i(t_{n/2^i})$, which is the lowest resolution of the model. The detail, $d_i(t_{n/2^i})$, is generated by multiplying $s_i(t_{n/2^i})$ with f_0 which is the fraction moved by the first cascade. The smoothed and detail information are combined by performing an inverse transform with the Haar wavelet. This process is equivalent to generating a single level of the cascade model. The detail information for the second cascade is generated by multiplying values of the smoothed information, $s_{i-1}(t_{n/2^{i-1}})$, with f_1 and repeating the inverse transform. Any level of inhomogeneity can be generated by repeating this process through the desired number of iterations. The sign of each coefficient in the detail information is randomly generated and determines whether the liquid water fraction is moved from right to left (positive) or from left to right (negative). This model is classified as a multiscale statistical model on the basis of this random element. Although the PDF, spectra, and moments of the model are deterministic, each realization of the model is randomly generated.

Within the multiresolution framework a number of variants of the BC model become apparent. The simplest variant comes with the choice of wavelet used in the inverse transform step of the cascade. Since the BC model redistributes LWP based on a Haar wavelet function, discontinuities between adjacent regions are generated as seen in Figure 11a. By using a wavelet with a higher order approximation, such as the D8 wavelet, the discontinuities at the boundary between adjacent regions is removed and the smoothness of the resultant LWP values is increased (Figure 11b). As a result of this change, the maximum value is reduced and the skewness of the generated LWP is also reduced. A disadvantage of this change is the presence of negative LWP. This represents a nonphysical condition that must be modified in order to apply the results to radiative transfer problems.

b. Variance coupled model

Two additional changes are made to the BC model in order to define the VC model. Unlike the BC model, where the inhomogeneity at one level is dependent on the local mean, the VC model uses the inhomogeneity at the previous level to generate the detail information at the next level. As indicated in the analysis of the LWP data, the standard deviation at any particular level is well correlated with the standard deviation at the previous level. Although the mean is related to the standard deviation at a particular level, it is not as well correlated and has a time-dependent element. When initializing the VC model with a uniform cloud layer, the mean LWP is used to generate inhomogeneity at the lowest level. Once this is done, the remaining levels of inhomogeneity are generated using only the inhomogeneity at the previous level.

The second change in the BC model is related to the method of redistributing the LWP. The BC model redistributes LWP using a deterministic fraction of the local LWP. The random element of the model is a binomial value of either −1 or +1 which determines the direction that the LWP is moved. The VC model uses a Gaussian distribution of mean zero to redistribute the LWP. The standard deviation of this distribution is determined by using the ratio of standard deviation between levels and the standard deviation of the previous level.

The VC model was developed using multiresolution analysis and is best presented within this framework. The model is initialized by prescribing a series of values to $s_i(t_{n/2^i})$. As illustrated in Figure 12, the detail at the lowest resolution is generated from this initialization data. The detail is assumed to have a Gaussian distribution of mean, μ, and standard deviation, σ. Since the mean of the wavelet transform is near zero, only the standard deviation needs to be determined. The standard deviation of the detail, $d_i(t_{n/2^i})$, is not constant and is calculated by multiplying a coupling factor, c_μ, with the corresponding value in $s_i(t_{n/2^i})$. The detail at the lowest resolution is then generated using the equation

$$d_i(t_n) = G(0, C_\mu \cdot s_i(t_n)) + \alpha \cdot d_i(t_{n-1}) \qquad (9)$$

where $G(\mu,\sigma)$ is a random number with Gaussian distribution and α is an autocorrelation factor. The autocorrelation term is included because wavelet analysis of the LWP data indicates that a Markov process is present at each level of the wavelet transform.

To generate the variability at the next highest level, the standard deviation is calculated using the variability of the current level. The standard deviation of the current level is multiplied by the variance coupling parameter, C_σ, which is the coupling between standard deviations of adjacent levels of the wavelet transform. The detail is then generated with the equation

$$d_{i-1}(t_{2n+k}) = G(0, C_\sigma \cdot \sigma_i(t_n)) + \alpha \cdot d_{i-1}(t_{2n+k-1}) \quad (k = 0,1) \qquad (10)$$

where $\sigma_i(t_n)$ is defined as the standard deviation of a window of width, 2λ, centered about t_n in the series $d_i(t_n)$. $\sigma_i(t_n)$ is calculated using the equations

$$\sigma_i(t_n) = \frac{1}{2\lambda - 1} \sum_{j=-\lambda}^{\lambda-1} (d_i(t_{n+j}) - \mu_i(t_n))^2 \qquad (11)$$

and

$$\mu_i(t_n) = \frac{1}{2\lambda} \sum_{j=-\lambda}^{\lambda-1} d_i(t_{n+j}) \qquad (12)$$

where $\mu_i(t_n)$ is the mean over this same window. It is necessary to calculate the standard deviation over a window of $d_i(t_n)$ in order to provide a sufficient sampling and yet allow for local fluctuations in the standard deviation. Additional levels of detail are generated using Equation 10 since the coupling with respect to standard deviation appears to be independent of scale. As in multiresolution analysis, the detail information at each level is combined with the initialization data to generate the full resolution of the model. The choice of wavelet for the inverse transform is an additional variable; however, to be consistent with the analysis of the LWP data, the D8 wavelet is used.

c. Analysis

Analysis of the BC and VC models corresponds with that of the LWP data. The primary parameters used to generate the models' results include the mean, standard deviation, and slope of the power spectrum. Values for these parameters are recorded in Table 2 along with other relevant parameters. An autocorrelation parameter is also included for the VC model since this is not inherently generated as in the cascade process of the BC model.

Performing a windowed analysis on the BC model demonstrates the correlation between mean thickness and standard deviation. As seen in Figures 13a and 13b, the mean and standard deviation appear identical apart from a scaling factor. This is an inherent property of the BC model where the variance parameter, f, determines the ratio between these two statistics. Since inhomogeneity is generated at each scale of the model through the use of the local mean, correlation of standard deviation between scales is also quite strong. The ratio of standard deviation between levels 1 and 2 is fairly uniform apart from a certain amount of noise. A similar relationship exists between standard deviations at larger scales. Although the inhomogeneity at any one level is a nonlinear combination of the mean and inhomogeneity at all previous scales, the ratio of standard deviation between levels is strongly influenced by the spectral parameter.

Although the analysis of the VC model appears markedly different from the BC model, the relationship between scales is still present. The most prevalent difference between the two models is that the uniform initialization field of LWP in the VC model is not greatly redistributed (Figure 13c). This is due primarily to the model being initialized with 64 data points each having a value of 100 g/m². The standard deviation (Figure 13d) does show variability; however, the magnitude is not as great as for the BC model. Since this model does not force a redistribution of LWP, the fluctuations in standard deviation give an indication of the natural variability of standard deviation within the VC model.

The standard deviation at each scale of the VC model correlates well with the standard deviation at other scales. This is a natural outcome of the VC model's use of the variance coupling parameter. Although the parameter is set at a value of 0.9, the calculated ratio of variability between scales is 0.99. This indicates that the ratio of standard deviation between scales is not the best estimator of the variance parameter and a better estimator is needed. One advantage of the variance coupling parameter is that the ratio of standard deviation between each level remains constant. This does not exist in the BC model since the ratio decreases as you go to higher levels in the analysis. This difference exists because the D8 wavelet used in the VC model generates a much smoother field than the D2 wavelet of the BC model. The discontinuities generated in the BC model introduce higher frequencies at each level of the cascade process. This artificially enhances the high frequency component and, therefore, the ratio of standard deviation between scales approaches unity at small scales. With the VC model there is not as much frequency leakage and the ratio of standard deviation is maintained to the smallest scales.

The power spectrum of both models is time independent and its slope is determined by the spectral parameter of each model. For the spectral parameters used in Table 2 the slope of the power spectrum is −1.7 and −1.6 for the BC and VC models, respectively. The difference in the values used for the spectral parameter of the two models is related to the method of generating the inhomogeneity and the wavelet used in the inverse transform. The choice of wavelet has the most profound effect since it is not possible for a Haar wavelet to generate a power spectrum slope more negative than −2 since it is a step function (Barker and Davies 1992). For particular ranges of the spectral parameter, the VC model as well as a modified BC model using higher order Daubechies wavelets is able to generate power spectral slopes more negative than −2.

As with the slope of the power spectrum, the autocorrelation of the wavelet transform at each level is time independent. Results from each model have an autocorrelation of −0.31. This corresponds to the value of the Markov parameter used in the VC model. The BC model generates autocorrelation through the cascading process and, therefore, has an advantage over the VC model since an additional parameter is not needed to determine the autocorrelation.

The skewness of each model poses the greatest difference between the BC and VC models. As mentioned previously the BC model inherently generates skewness through the cascading process. Using the parameters in Table 2 the skewness is 0.45, although it can be decreased by decreasing the values of both the variance and spectral parameters. For the VC model the skewness remains near zero. Apart from conditions which generate negative LWP, the skewness of the VC model is unaffected by the choice of the variance and spectral parameters.

The effect of skewness is easily seen in the composite PDF of results from both models. As illustrated in Figure 14a, the composite PDF for the BC model is clearly non-Gaussian and is actually log-normal as a result of the cascading process. This shape is present at all scales since the fundamental cascading process does not change for smaller scales. On the other hand the composite PDF for the VC model is symmetric about the mean and is nearly Gaussian for all temporal scales (Figure 14b). This demonstrates a clear difference between the two models and is partially due to the choice of wavelet and mostly dependent on the redistribution process of the model.

5. Discussion and summary

a. Liquid water path data

The analysis of the LWP data reveals some relationships which are useful for evaluating the effectiveness of the two models. The most important is the relationship between inhomogeneity at different scales. From the windowed analysis it is evident that the inhomogeneity, quantified by the standard deviation, is correlated to the local mean of the LWP data. This inhomogeneity is distributed over many scales and, through the use of wavelet analysis, is seen to have a set relationship between scales. This relationship is quantified by defining the ratio of standard deviation between scales. Since this ratio has a value that is less than one, the amount of inhomogeneity decreases at smaller scales. This ratio is present in the LWP data as well as in the BC and VC models, and indicates the existence of scaling properties. Being time and scale independent the ratio of standard deviation between scales is closely related to the slope of the power spectrum.

Another important relationship is the presence of skewness in the LWP data and its dependence on window size. Using a half-hour window (9 km spatial scale) the skewness is very small and the composite PDF is nearly a Gaussian distribution. This Gaussian shape may not be indicative of the LWP data due to instrument noise and limitations in the retrieval algorithm. However, limiting measurements to those where the expected noise is less than 25% of the variability does not appreciably change the PDF. This implies that either the inherent skewness at the small scale is relatively small, or the noise in the data is underestimated. With the 2 hour window (36 km spatial scale) the skewness increases and noticeably affects the composite PDF. The source of this skewness can only be speculated; however, a few observations can be made when comparing the LWP data with model results. Unlike the BC model, the presence of skewness is not present at all scales. This implies that certain dynamic processes which introduce skewness only become important beyond the half-hour time scale (9 km length scale). By comparing the time series for skewness and the mean, the skewness appears to be largest when cloud thickness is changing. This may indicate that the thickening and thinning process of cloud dynamics is nonlinear in marine stratocumulus.

b. Inhomogeneity models

Evaluation of the BC model shows that it is useful in capturing some very important features of the data. This model is able to capture the correlation between standard deviation and mean as well as the correlation of standard deviation between scales. Using only two parameters, the BC model not only adds inhomogeneity consistent with the standard deviation and power spectrum, but also reproduces the autocorrelation and large scale skewness which leads to a log-normal distribution of LWP. The limitations of this model are the discontinuities generated by using a step function in the cascade process and the presence of skewness to the smallest scales of the model, which appears to conflict with results from the LWP analysis.

Results from the VC model demonstrate both strengths and weaknesses of this model. As with the BC model, the VC model is able to capture the correlation between variability at different scales. However, due to the redistribution process of the VC model, the standard deviation is not as closely correlated to the mean. The standard deviation, power spectrum, and autocorrelation can all be adjusted using the parameters of the model. The VC model is best suited for generating small-scale inhomogeneity within a known large scale cloud field. Although this model produces distributions which are nonskewed locally, it is not able to generate the skewness that is present at the larger scales of the LWP data. A possibility for future work is to make a time-dependent version of the VC model which is able to generate skewness through the use of a nonlinear cloud development scheme.

The difference between the BC and VC models demonstrates the flexibility of the multiresolution framework for modeling cloud inhomogeneity. A number of variants to the BC model are implied by this research, such as the change in the cascading wavelet and the random number generator function. Although a high correlation exists between inhomogeneity at adjacent scales, the method for generating this correlation can take a variety of forms. Finally, the multiresolution framework has an analogous two-dimensional form that provides a means of extending these inhomogeneity models into higher dimensions.

Both models involved in this study use fractal principles to generate cloud inhomogeneity. As seen in the analysis of the LWP data and the inhomogeneity models, one way in which this principle manifests itself is in the correlation of inhomogeneity at different scales. This provides a method not only of generating a consistent PDF, but also provides a constraint when adding inhomogeneity into a known cloud LWP field. This feature may not be important in general circulation models where knowledge of the PDF appears to be sufficient for calculating radiative fluxes (Cahalan 1994b). However, consistent spatial distribution becomes important when calculating the radiance of a cloud field for remote sensing applications. In this later application the BC and VC models may prove useful for incorporating sub-resolution cloud inhomogeneity.

Acknowledgments. The introductory help in learning wavelets by Praveen Kumar and the comments by the reviewers are greatly appreciated. The facilities at Purdue University and the faculty and staff of the Department of Earth and Atmospheric Sciences assisted in the completion of this work. This work was done with the support of NASA GSRP

Fellowship NGT-50963. Partial support for this work was provided by NSF Grant ATM-8909870 and NASA Grant NAGW-3150. This research was initiated while SG attended NASA/GSFC under the University Space Research Association's Graduate Student Program in the Earth System Sciences.

REFERENCES

Albrecht, B. A., C. W. Fairall, D. W. Thomson, A. B. White, J. B. Snider, and W. H. Schubert, 1990: Surface-based remote sensing of the observed and the adiabatic liquid water content of stratocumulus clouds. *Geophys. Res. Lett.* **17**, 89-92.

Argoul, F., A. Arneodo, G. Grasseau, Y. Gagne, E. J. Hopfinger and U. Frisch, 1989: Wavelet analysis of turbulence reveals the multifractal nature of the Richardson cascade. *Nature*, **338**, 51-53.

Barker, H. W. and J. A. Davies, 1992: Radiative transfer, cloud structure, and the characteristics of satellite radiance wavenumber spectra. *Remote Sens. Environ.*, **42**, 51-64.

Blaskovic, M., R. Davies and J. B. Snider, 1991: Diurnal variation of marine stratocumulus over San Nicolas Island during July 1987. *Mon. Wea. Rev.*, **119**, 1469-1478.

Cahalan, R. F., 1989: Overview of fractal clouds. Advances in Remote Sensing Retrieval Methods, Deepak Pub., 371-388.

_____, and J. H. Joseph, 1989: Fractal statistics of cloud fields. *Mon. Wea. Rev.* **117**, 261-272.

_____, and J. B. Snider, 1989: Marine stratocumulus structure. *Remote Sens. Environ.* **28**, 95-107.

_____, and _____, 1990: Turbulence in marine stratocumulus. *Proc. of the 9th Symp. on Turbulence and Diffusion*, Roskilda, Denmark, 11-13.

_____, and W. J. Wiscombe, 1992: Impact of cloud structure upon climate. *Current Problems in Atmospheric Radiation*, Deepak Pub., 120-124.

_____, W. Ridgway, W. J. Wiscombe, T. L. Bell, and J. B. Snider, 1994a: The albedo of fractal stratocumulus clouds, *J. Atmos. Sci.* **51**, in press.

_____, _____, _____, S. M. Gollmer, and Harshvardhan, 1994b: Independent pixel and Monte Carlo estimates of stratocumulus albedo, *J. Atmos. Sci.* **51**, in press.

Cox, S. K., D. S. McDougal, D. A. Randall, and R. A. Schiffer, 1987: FIRE - the first ISCCP regional experiment. *Bull. Amer. Meteor. Soc.,* **68**, 114-118.

Daubechies, I., 1988: Orthogonal bases of compactly supported wavelets. *Commun. on Pure Appl. Math.* **XLI**, 901-996.

_____, 1992: *Ten Lectures on Wavelets*, Soc. Indus. and Appl. Math., 357 pp.

Davies, R., 1978: The effect of finite geometry on the three-dimensional transfer of solar irradiance in clouds. *J. Atmos. Sci.* **35**, 1712-1725.

Davis, A., P. Gabriel, S. Lovejoy, D. Schertzer, and G. L. Austin, 1990: Discrete angle radiative transfer 3, Numerical results and meteorological applications. *J. Geophys. Res.* **95**, 11,729-11,742.

_____, A. Marshak, and W. Wiscombe, 1994: Wavelet-based multifractal analysis of non-stationary and/or intermittent geophysical signals. *Wavelet Transforms in Geophysics*, Academic Press (In publication).

Evans, K. F., 1993. Two-dimensional radiative transfer in cloudy atmospheres: The spherical harmonic spatial grid method. *J. Atmos. Sci.* **50**, 3111-3124.

Gabriel, P. M., S.-C. Tsay and G. L. Stephens, 1993. A Fourier-Riccati approach to radiative transfer. Part I: Foundations. *J. Atmos. Sci.* **50**, 3125-3147.

Gage, K. S. and G. D. Nastrom, 1986: Theoretical interpretation of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft during GASP. *J. Atmos. Sci.* **43**, 729-740.

Gamage, N. and C. Hagelberg, 1993: Detection and analysis of microfronts and associated coherent events using localized transforms. *J. Atmos. Sci.* **50**, 750-756.

Harshvardhan and D. A. Randall, 1985: Comments on "The parameterization of radiation for numerical weather prediction and climate models." *Mon. Wea. Rev.* **113**, 1832-1833.

Hentschel, H.G.E. and I. Procaccia, 1984: Relative diffusion in turbulent media: The fractal dimension of clouds. *Phys. Rev. A* **29**, 1461-1470.

Kobayashi, T., 1991: Reflected solar flux for horizontally inhomogeneous atmospheres. *J. Atmos. Sci.* **48**, 2436-2447.

Lovejoy, S., 1982: Area-perimeter relation for rain and cloud areas. *Science* **216**, 185-187.

_____, A. Davis, P. Gabriel, D. Schertzer and G. L. Austin, 1990: Discrete angle radiative transfer 1, Scaling and similarity, universality and diffusion. *J. Geophys. Res.* **95**, 11,699-11,715.

Mahrt, L., 1991: Eddy asymmetry in the sheared heated boundary layer. *J. Atmos. Sci.* **48**, 472-492.

Mallat, S. G., 1989: A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intel.* **11**, 674-693.

Marshak, A., A. Davis, R. Cahalan and W. Wiscombe, 1994: Bounded cascade models as nonstationary multifractals. *Phys. Rev.*, **E49**, 55-69.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1992: *Numerical Recipes in C*, Cambridge U. Press, 735 pp.

Schertzer, D. and S. Lovejoy, 1987: Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *J. Geophys. Res.*, **92**, 9693-9714.

Snider, J. B., 1988: Estimated accuracy of ground-based liquid water measurements during FIRE, *FIRE Science Results 1988*, NASA Conference Publication, 295-298.

Stephens, G. L., 1988: Radiative transfer through arbitrary shaped optical media. Part I: A general method of solution. *J. Atmos. Sci.* **45**, 1818-1836.

Weng, H., and K.-M. Lau, 1994: Wavelets, period doubling, and time-frequency localization with application to organization of convection over the tropical western pacific. *J. Atmos. Sci.* **51**, 2523-2541.

Figure Captions

Figure 1. Plot of the (a) D8 wavelet function, (b) its Fourier transform, (c) the corresponding scale function and its (d) Fourier transform.

Figure 2. Diagram of multiresolution analysis.

Figure 3. Frequency ranges of multiresolution analysis for the LWP data.

Figure 4. Correspondence between windowed analysis and wavelet analysis.

Figure 5. Time development of the LWP data's local (a) mean and (b) slope of the power spectrum.

Figure 6. Time development of the LWP data's local (a) standard deviation and standard deviation of the level (b) 1, (c) 2, and (d) 3 wavelet transform.

Figure 7. Time development of the LWP data's local (a) ratio of standard deviation between levels 1 and 2 and (b) ratio between level 1 standard deviation and mean.

Figure 8. LWP composite PDFs for (a) 1-hour and (b) 8-hour window sizes.

Figure 9. LWP composite PDFs for 1st level detail with (a) 1-hour and (b) 8-hour window sizes.

Figure 10. Diagram of the BC model cast into the multiresolution framework.

Figure 11. LWP field generated from the BC model using the (a) D2 (Haar) and (b) D8 wavelets.

Figure 12. Diagram of the VC model in the multiresolution framework.

Figure 13. Position dependence of the BC model's local (a) mean and (b) standard deviation and the VC model's local (c) mean and (d) standard deviation.

Figure 14. Composite PDFs of the LWP from the (a) BC and (b) VC models.

Table Captions

Table 1. Tap values, g_i and h_i for the Daubechies D8 wavelet and scale functions respectively.

Table 2. Parameters for the BC and VC models. The number of parameters specified gives an indication of the flexibility provided by the multiresolution framework for modeling.