EFFECT OF CLOUD MACROSCALE PROPERTIES ON CLIMATE FORECASTING IN THE TROPICS

by

MICHAEL J. FOSTER

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Atmospheric Science

written under the direction of

Dr. Dana E. Veron

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2008

ABSTRACT OF THE DISSERTATION

Effect of Cloud Macroscale Properties on Climate Forecasting in the Tropics

By Michael J. Foster

Dissertation Director:

Dana E. Veron

This work represents an effort to improve climate model treatment of shortwave radiative transfer through cloudy atmosphere. Results from a stochastic model, which statistically represents radiative transfer through broken cloud fields using a Markovian distribution of observed cloud chord lengths, are compared against those of a plane-parallel model for a variety of cloudy scenes observed in the tropics over a four year period in an effort to reduce the error in radiative transfer calculations due to neglected cloud field morphology. A *k*-means clustering algorithm is applied to observed cloud optical and dynamical properties in order to identify the presence of specific cloud regimes, which in turn are categorized using large-scale simulated climate variables. The results from the model comparison and cluster analysis are used to develop criteria by which to identify situations where the stochastic model outperforms a traditional plane-parallel model. These criteria are applied to output from a climate model resulting in the identification of some of the issues involved in applying a parameterization developed with observations to model-generated cloud fields.

Acknowledgements

The path I've taken in attaining this degree has had its fair share of obstacles, and there are many people who have helped me along the way. For this I am extremely grateful, and though there are many colleagues, friends and loved ones to whom I owe a debt of gratitude, there are a handful I would like to mention here.

I would like to thank my advisor, Dana Veron, for her continued scientific, moral and financial support throughout this process. I have learned from her that bugs in code, times of slow progress, and unexpected results are poor reasons for becoming discouraged. This lesson, among many others, has greatly improved both my own skills as a scientist and the quality of the work presented here.

I would also like to thank Jennifer Francis and Tony Broccoli who, besides providing support and insight as members of my dissertation committee, helped make a steep learning curve a little easier in their classes and took the time to let me know I was making progress.

I am grateful to George Tselioudis for agreeing to serve as my outside committee member and for taking measures to fulfill that obligation when travel and other commitments made it difficult to do so.

I would like to thank my family, and especially my parents. For many years our family motto has jokingly been "Okay is good enough for us", but the examples that have been set for me and the support I have received have never been mediocre.

I should also mention the contribution of my puppy Luna, who kept me smiling during the long hours and late nights required to finish this dissertation.

Finally I would like to thank my wife Lauren, who has provided support of every conceivable kind: sometimes acting as a sounding board, sometimes by letting me vent frustrations, though mostly as a source of unwavering love and frequently a hot meal. I could not have accomplished this without her.

The material of Chapter 2 has been accepted in an article format and will appear in a forthcoming issue of *Journal of Geophysical Research*, under the title "Evaluating the Stochastic approach to Shortwave Radiative Transfer in the Tropical Western Pacific".

This research could not have been performed without the help of the many participants and resources of the ARM Program. This research was supported by the Office of Biological and Environmental Research of the U.S. Department of Energy under grant number DE-FG02-06ER64246 as part of the Atmospheric Radiation Measurement Program.

Table of Contents

List of Acronyms

AGCM	Atmospheric Global Climate Model
ARM	Atmospheric Radiation Measurement Program
ARSCL	Active Remotely-Sensed Clouds Locations
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation
CAM3	Community Atmosphere Model Version 3.0
CAPE	Convective Available Potential Energy
CART	Cloud and Radiation Testbed
CCSM	Community Climate System Model
CCSM3	Community Climate System Model Version 3.0
CINH	Convective Inhibition
CRM	Cloud Resolving Model / Column Radiation Model
DSTOC	Stochastic Shortwave Radiative Transfer Model
GCM	Global Climate Model
IPCC	Intergovernmental Panel on Climate Change
LWP	Liquid Water Path
MFRSR	Multifilter Rotating Shadowband Radiometer
MMCR	Millimeter Wavelength Cloud Radar
MPL	Micropulse Lidar
MWR	Microwave Radiometer
MWRRET	Microwave Radiometer Retrieval Value-Added Product
NSA	North Slope of Alaska
RFO	Relative Frequency of Occurrence

SGP	Southern Great Plains
SRES	Special Report on Emission Scenarios
SSI	Downward Solar Surface Irradiance
SZA	Solar Zenith Angle
TSI	Total Sky Imager
TWP	Tropical Western Pacific
VAP	Value-Added Product

List of Tables

List of Figures

CHAPTER 1 – INTRODUCTION

Clouds play an important role in the Earth's climate system. They reflect incoming solar radiation back into space, while absorbing and emitting infrared radiation. Clouds affect atmospheric heating rates through latent heating from phase changes of cloud water and ice. They transport moisture, heat, and momentum through convective processes, and are a primary component of the Earth's hydrological cycle (e.g., Houghton et al., 2001; Ackerman and Stokes, 2003). The process of cloud formation is composed of numerous interconnected complex processes, including large-scale and sub-grid-scale dynamics, microphysics of ice and water precipitation, formation, and removal, and atmospheric thermodynamics. Cloud-radiation interactions are important at both the microscale and cloud field scale, and are frequently cited as the largest reason for the large range in climate sensitivity estimates in atmospheric global climate models (AGCMs) (e.g., Cess et al., 1990; Senior, 1999; Potter and Cess, 2004; Randall et al., 2007). This is especially true in the tropics, where shortwave radiation interactions with boundary layer and mid-level clouds have a significant impact on the energy budgets that are major drivers of global atmospheric circulation. Consequently, it is important to improve our ability to realistically simulate cloud radiative effects in climate models. As an example of this, Figure 1.1 taken from the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4) relates how varied the treatment of shortwave radiation is in climate models. This figure illustrates the zonally-averaged mean difference between model-generated and observed shortwave radiation reflected into space between 1985 and 1989. Depending on the model being used, the magnitude

of reflected shortwave radiation may differ from that observed by as much as 20 Wm^{-2}. The multi-model mean difference from observations in the tropics is around 5 Wm^{-2}. Regionally these errors can be even more pronounced, though some recent studies have shown that in some instances errors are being offset by coinciding problems with unrealistic cloud optical properties, fraction, and vertical structure, thus masking the extent of the uncertainty (Webb et al., 2001; Potter and Cess, 2004).

Cloud treatment in climate forecasting

Current AGCM grid cells are typically much larger than the scale of individual clouds, necessitating the parameterization of cloud sub-grid scale impacts on the radiation fields. Many AGCMs take a simplified approach to this endeavor by diagnosing cloud amount in each vertical layer based on mean values of relative humidity averaged over the grid cell. In-cloud microphysical properties are often treated as homogeneous throughout the cloud, and the spatial relationship amongst clouds in different vertical layers is specified with a cloud overlap assumption. Recent improvements in the treatment of clouds in climate models include the separate treatment of cloud water and ice condensate, advection of these variables in large-scale circulations, improvement in convective parameterizations, and consistent treatment of cloud particles (Boville et al., 2006). The simulation of marine low-level clouds has also improved for some AGCMs (Randall et al., 2007). Many AGCMs separately parameterize convective cloud with large vertical velocities and stratiform cloud with weaker vertical motion, allowing for discrete treatment of microphysical properties. However, the treatment of cloud

properties through the use of mean values ultimately neglects the effects of sub-grid scale variability. When these sub-grid scale processes are non-linear, as is often the case, errors occur within the model simulations. Examples of processes where this can occur include cloud reflectance and microphysical properties (Cahalan et al., 1994; Pincus and Klein, 2000). These errors can in turn propagate and adversely affect atmospheric heating rates, cloud formation and dissipation processes, and precipitation rates; the end result being biases in the cloud feedback processes. Many studies have addressed these issues, and most can be grouped into those that address the effects of microscale or in-cloud inhomogeneity (Cahalan, 1994; Barker et al., 1996; Davis et al., 1997; Marshak et al., 1997; Pincus and Klein, 2000; among others), and those that address macroscale or cloud field configuration inhomogeneity. This work relates specifically to bias in shortwave radiative transfer calculations due to macroscale cloud field inhomogeneity.

The variance associated with cloud field geometry is a result of many factors, including the number and distribution of individual clouds, cloud shape and size (Zuev and Titov, 1995). Early studies examined isolated non-planar homogenous clouds and found that solar fluxes could change significantly compared to those of plane-parallel clouds due to cloud side illumination from direct solar beams and diffusion of photons (Davies, 1978; Barker and Davies, 1992; Zuev and Titov, 1995). Other studies looked at arrays of identical homogenous clouds with differing shapes and spacing and found that areas of enhanced illumination, interactions of photons between clouds, and cloud shadowing all play important roles for flux calculations (Welch and Wielicki, 1984; Kobayashi, 1988). These studies also showed that nonplanar clouds reflect more radiation than plane-parallel cloud when the sun is at large zenith angles and less when

the sun is overhead (Barker and Davies, 1992; Marshak et al., 1998). In addition, the radiative field has a nonlinear dependence on cloud parameters, as do related processes such as atmospheric sensible and latent heating. This further complicates the problem, and indicates that mean cloud properties are insufficient for calculating planetary albedo accurately (Cahalan, 1994; Zuev and Titov, 1995). To put this in perspective, a decrease in global albedo of 10 percent, assuming all other variables are held constant, would increase the global mean surface temperature by approximately five degrees Kelvin (Cahalan, 1994). This is roughly the same as doubling the atmospheric carbon dioxide concentration. From Figure 1.1, the 5 Wm^{-2} difference between multi-model mean reflected shortwave radiation and that observed over the tropics translates to a 5 percent difference in shortwave albedo.

Approaches to improving cloud simulation in climate models

One effort to resolve these issues is through the improvement of spatial resolution in models. While we are likely still decades away from achieving the computer processing power necessary to resolve individual clouds in a global climate model, regional or mesoscale models may currently benefit from this approach. Cloud resolving models (CRMs) have the necessary spatial and temporal resolution to simulate individual clouds and by extension fields throughout a cloud system lifetime. CRMs are often used for sensitivity and model intercomparison studies (Klein et al., 2008; Morrison et al., 2008). CRMs are also being used in the development of “superparameterizations”, where a CRM is embedded into an AGCM grid cell where it may explicitly simulate sub-grid

scale cloud processes that are in turn used as a statistical sampling for cloud properties throughout the entire grid cell (Grabowski, 2001; Randall et al., 2003a,b). While this approach is undoubtedly promising, there are issues that need to be addressed. Embedding single 2-dimensional CRMs in each grid cell allow for sampling errors: Grabowski (2001) embedded a 2D CRM in each grid cell of a simplified AGCM and found that his results changed based on the orientation of the CRM. A simple solution might be to embed multiple 2D or a single 3D CRM into each grid cell, but the problem of computational power quickly arises: Khairoutdinov and Randall (2001) found that after embedding a 2D CSRM into the Community Atmosphere Model (CAM) the running time was increased by a factor of 180, meaning it took one hour on a 64-processor server array to simulate one day.

Another approach to improving simulation of cloud-radiation interactions is through the development of statistical models, which use distributions of cloud properties, such as water content or (in our case) cloud chord lengths, to simulate the effects of sub-grid scale variability. There has been a number of studies that use statistical models to examine the bias created by ignoring the effects of cloud field microscale inhomogeneity (Cahalan, 1994; Marshak et al., 1998; Fu et al., 2000; Di Giuseppe and Tompkins, 2003; Di Giuseppe and Tompkins, 2005). This bias is often known as the "plane-parallel bias" and is defined as the difference between Independent Column Approximation (ICA) and plane-parallel calculations. The ICA method essentially breaks the grid cell into many smaller cells and calculates cloud microphysical properties in each one. The "ICA bias" is then defined as the difference between the ICA calculations and a fully 3D interactive radiative transfer calculation. This bias takes into account morphology of the cloud

fields. As discussed earlier, this includes horizontal photon flow, or "leakage", through cloud sides, cloud shadowing and areas of enhanced illumination, and has a dependence on solar zenith angle (SZA). Most of these studies use some form of a statistical Monte Carlo code to calculate the 3D radiative transfer. The Monte Carlo method simulates the physically realistic processes that occur as radiation traverses the atmosphere. That is, the flow of the radiation is calculated by simulating the paths individual photons travel. The length of these paths, and the events that occur when photons interact with cloud (absorption and scattering) are determined by probability functions. These functions vary depending on the set of parameters used in the atmosphere, such as the cloud microphysics, presence of aerosol, and the wavelength of the incident radiation (O'Hirok and Gautier, 1998).

One requirement shared by all these approaches is the realization of specific cloud fields for use by the Monte Carlo code. These cloud fields are taken from observations, or simulated by CRMs or Large Eddy Simulation (LES) models. While these individual cases provide insight into the potential for radiation calculation bias due to microscale and macroscale cloud inhomogeneity, the application to model simulations running on climate time scales is not immediately realized. Another type of statistical model, such as that used in this study, treats clear sky and cloud as two immiscible fluids in a mixture, and while the microphysical properties of the fluid are known deterministically, the type of fluid at any given location within the mixture is determined stochastically (Stephens et al., 1991; Malvagi et al., 1993; Lane-Veron and Somerville, 2004). One advantage to this method is that specific cloud fields do not need to be realized, though information about cloud size and spacing must still be provided. Lane et al. (2002) showed that

distributions of cloud chord lengths could be derived from ground-based observational data and used in the stochastic model to represent radiative fields averaged over a temporal or spatial domain. This makes it possible to incorporate realistic treatment of 3D radiative transfer into models running on climate time-scales without necessitating the formulation of specific cloud fields.

1.2 THEORY AND DATA

Stochastic theory

As a photon travels from the top of the atmosphere to the surface, the path it follows can be envisioned as a line. This is not to say that the three-dimensional atmosphere is not taken into account; as the photon is scattered in different directions the length of its path adjusts accordingly, thereby increasing (or reducing) its probability of being absorbed or reflected before reaching the surface. Let us assume the photon travels through an atmosphere composed of two non-mixing substances: clear sky and cloud. The optical properties in each substance, such as optical depth, single-scattering albedo, and asymmetry parameter, are known deterministically but the presence of either cloud or clear sky at any given point is determined stochastically. If we were to solve the equation of radiative transfer for an infinite number of clear-sky and cloud realizations, each sharing the same statistical properties, then the ensemble-averaged intensity, $\langle I \rangle$, of these solutions would represent the robustly known intensity averaged over time or space. Of course, solving for an infinite number of cases is unrealistic. Therefore a set of closures,

or approximations, must be used to solve for <*I*> directly using a set of simple equations (Malvagi et al., 1993).

Extending this idea of a line being the path a photon takes from the top of the atmosphere to the surface, statistics on a line can be used to determine the conditional probabilities of a photon being located in either clear sky or cloud at any given point on its path. For further reading on line statistics, there are good general descriptions available: (e.g., Marshak and Davis, 2005; Lane-Veron and Somerville, 2004), as well as more in-depth discussions: (e.g., Pomraning, 1991; Sanchez et al., 1994). With the conditional line probabilities calculated, an infinite number of possible photon paths can be determined. Figure 1.2 illustrates a number of possible paths as an example. The advantage of the stochastic method is that it allows us to formulate four situations: 1) the photon is in clear sky, 2) the photon is in cloud, 3) the photon is transitioning from clear sky into cloud, or 4) the photon is transitioning from cloud into clear sky. Traditional plane-parallel radiative transfer models only allow for situations where a photon is either in clear sky or cloud (Lane-Veron and Somerville, 2004). These transitional situations allow for more complex radiative interactions amongst clouds in broken fields.

Atmospheric Radiation Measurement Program

The Atmospheric Radiation Measurement Program (ARM) (Stokes and Schwartz, 1994) was created by the U.S. Department of Energy for the purpose of improving our treatment of clouds and radiative processes in Global Climate Models (GCMs). The method by which this is accomplished has been to develop Cloud and Radiation Testbed

(CART) sites where high resolution measurements of radiative, cloud, and general atmospheric properties are taken over time scales appropriate for climate modeling. There are three primary sites: The Southern Great Plains (SGP), The Tropical Western Pacific (TWP), and the North Slope of Alaska (NSA). These locations were chosen for their ability to provide a wide range of atmospheric and climatic conditions to study. In addition, there are ARM Mobile Facilities that are transported to various locations around the world and take measurements for shorter time periods, often one year. The focus of this dissertation is in the tropics, and as such the observational data are taken from the TWP CART site. This site is actually composed of three research facilities: the Manus facility located on Los Negros Island in Manus, Papua New Guinea, the Darwin facility located in Darwin, Northern Territory, Australia, and the Nauru facility located on Nauru Island, Republic of Nauru. Figure 1.3 displays a map of the three facilities and their relative locations. Model simulations are performed using data strictly from the Nauru facility, while the cluster analysis portion utilizes data from all three TWP facilities.

Cluster theory

The shortwave radiative properties of clouds vary considerably with height, amount, and particle shape, size, and water/ice composition. However, AGCMs don't predict cloud type with enough consistency and in enough detail to capture the sub-grid scale thermodynamic and dynamic processes that affect cloud radiative properties. For this reason it is helpful to use high-resolution surface measurements of cloud properties to separate observed cloud fields into different regimes, and then to look at the radiative

properties for each of these regimes. The tool chosen for this purpose is cluster analysis. More specifically, the k-means clustering algorithm is used to objectively distinguish amongst cloud regimes (Anderberg, 1973; Jakob and Tselioudis, 2003). The purpose of the k-means algorithm is to assign objects to a predetermined, k, number of groups (or clusters) based on similarity of the objects. This is accomplished by mapping objects onto a multi-dimensional attribute space, the number of dimensions being the number of object attributes being examined, and defining similarity as the Euclidean distance between these objects. Figure 1.4 shows a simple example of the clustering algorithm. The algorithm works using an iterative method, maximizing the Euclidean distance between clusters while minimizing the Euclidean distance between objects within each cluster. The cluster centroids are initially collocated with randomly chosen objects amongst the set, and then their locations are adjusted with each iteration of the algorithm. The process is completed after a specified number of iterations, in this case 50, or when the cluster memberships no longer change, whichever occurs first. Since the number of clusters is chosen before-hand, an objective set of criteria by which to choose a value of k may be applied.

1.3 RECENT WORK

Stochastic model

Lane et al. (2002) took data from the sensors located at the ARM SGP CART site in Oklahoma, and generated the cloud properties required to run the stochastic model. The statistical distribution of cloud size and spacing was determined using a threshold

technique and data on cloud-base height, the wind speed at that height, and radiance measurements from the Multifilter Rotating Shadowband Radiometer (MFRSR). The assumption was made that sudden changes in the direct beam measured by the MFRSR indicated the passage of cloud between the surface and the sun. The combination of these data allowed them to calculate the horizontal extent of the clouds and the spacing between them. These data were compiled to create a statistical distribution of cloud and clear-sky chord lengths. The study was constrained to times when broken cloud fields were present. In order for the observations to represent a statistical average of cloud properties over the area being examined, spatial and temporal averaging techniques were applied. The spatial averaging consisted of using data from multiple Oklahoma Mesonet stations, as well as multiple MFRSRs located within one-hundred kilometers of the central facility. The second technique involved taking the highly resolved temporal measurements of cloud properties and averaging them over one-hour intervals. In the end forty-five hours of data were input into the stochastic model and the results were compared to those of a plane-parallel model. Figures 1.5 and 1.6, taken from Lane et al. (2002), show some results from this study.

It was found that the stochastic model performed well in some cases and not so well in others. However, the results were ultimately inconclusive because it was found that the stochastic model was highly sensitive to certain input, such as cloud liquid water path (LWP) and cloud fraction. The most promising results of this study were that it was shown to be possible to derive the information necessary to build a statistical description of cloud field geometry from observations.

A more recent study by Veron et al. (2008) examines stochastic model results for the year 2000 at the ARM SGP, TWP, and NSA CART sites. The cloud information required as input to the stochastic model is extracted from a variety of instruments from these sites, including the Micropulse Lidar, 915-MHz Radar Wind Profiler and Radio Acoustic Sounding System, Vaisala Ceilometer, microwave radiometer, Active Remotely-Sensed Clouds Locations (ARSCL) products, Sky Radiation platform and Oklahoma Mesonet data. The cloud properties extracted include the base height, fraction, thickness, droplet effective radius, LWP, and the characteristic horizontal extent and spacing. Notably the horizontal extent and spacing translate into the Markovian distribution of chord lengths required for the stochastic portion of the model. The resulting radiation fields generated by the stochastic model are compared with those of a plane-parallel model and observational data. The stochastic model generates the best results at the NSA site and at the TWP site during times of deep convection. The work for this study is finished and a paper is in preparation.

Cluster analysis

Some recent studies have made use of the cluster analysis technique to identify cloud regimes (Jakob and Tselioudis, 2003; Jakob et al. 2005; Gordon et al. 2005; Williams et al. 2005). In the Jakob et al. (2005) study the primary data used for the clustering algorithm are cloud shortwave optical thickness and cloud-top pressure, processed by the International Satellite Cloud Climatology Project (ISCCP) from radiances measured by geostationary weather satellites (Rossow and Schiffer, 1999). They create histograms by

calculating the frequency-of-occurrence of specific cloud optical thickness and cloud-top pressure pairings for specific cloud coverage ranges over three-hour intervals. The study is run over a spatial domain of 280 x 280 km, located near the ARM TWP site. The data are subsampled within this domain every thirty kilometers, at five kilometer cell sizes, which is the resolution of the satellite pixels. The study is performed for the years of 1999 and 2000. Figure 1.7, from Jakob et al. (2005), shows the four distinct clusters the k-means algorithm generates, as well as a subset of the histograms located at the Manus TWP site within the domain being examined.

The authors describe the regimes represented by these clusters as consisting of two suppressed regimes: one containing primarily shallow clouds (SSC) and the other composed of thin cirrus clouds (STC), and two convectively active regimes: one made up of optically thin cirrus with large coverage (CC), and the other composed of optically thick clouds with large coverage (CD). The ISCCP data alone are insufficient to provide insight into the properties of these regimes, so instrumental and radiosonde data from the ARM TWP site are used to describe the regime characteristics, such as TOA and surface radiative fluxes, cloud properties, and temperature and water vapor profiles.

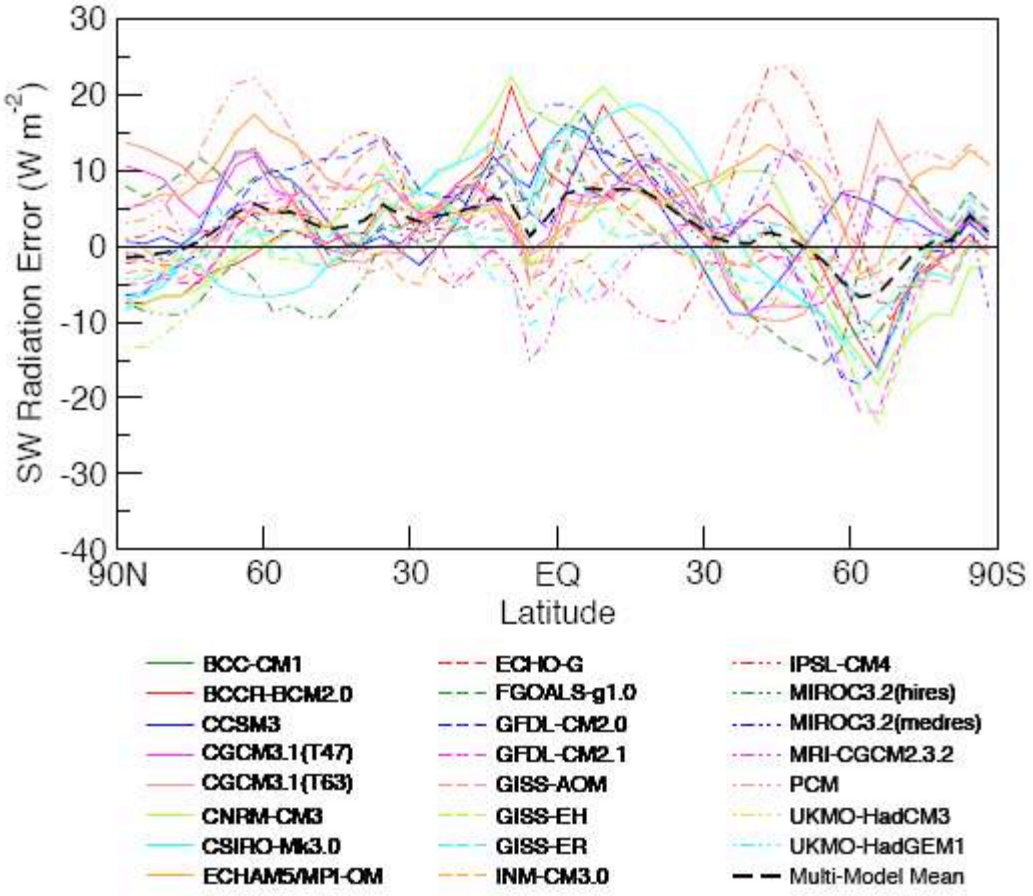


Figure 1.1 The ERBE (Barkstrom et al., 1989) observational estimates shown here are from 1985–1989 satellite-based radiometers, and the model results are for the same period of the CMIP3 20th Century simulations. Taken from Randall et al. (2007).

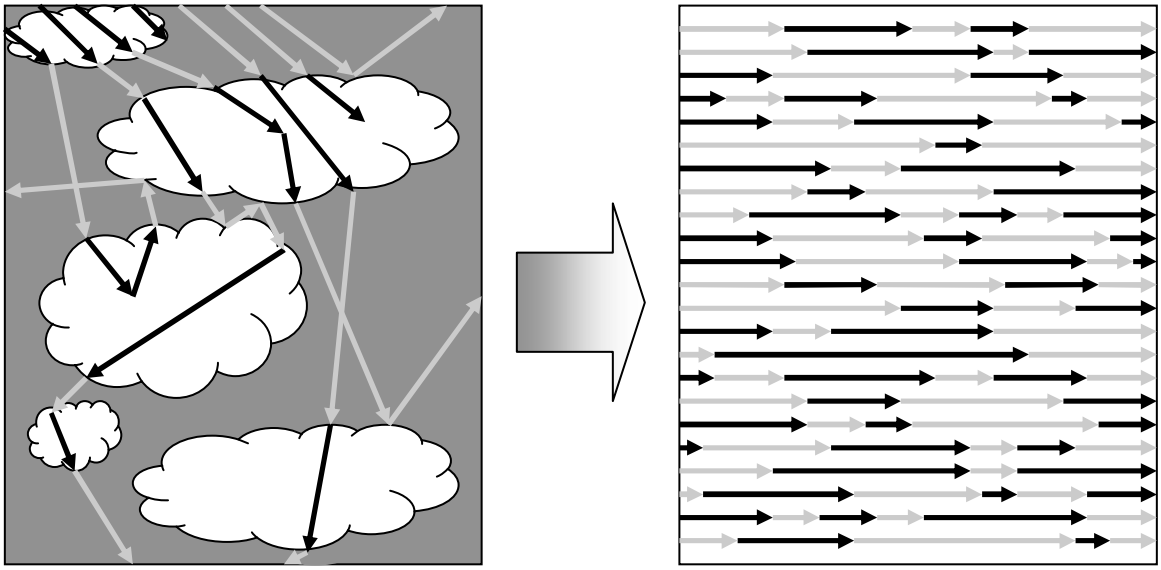
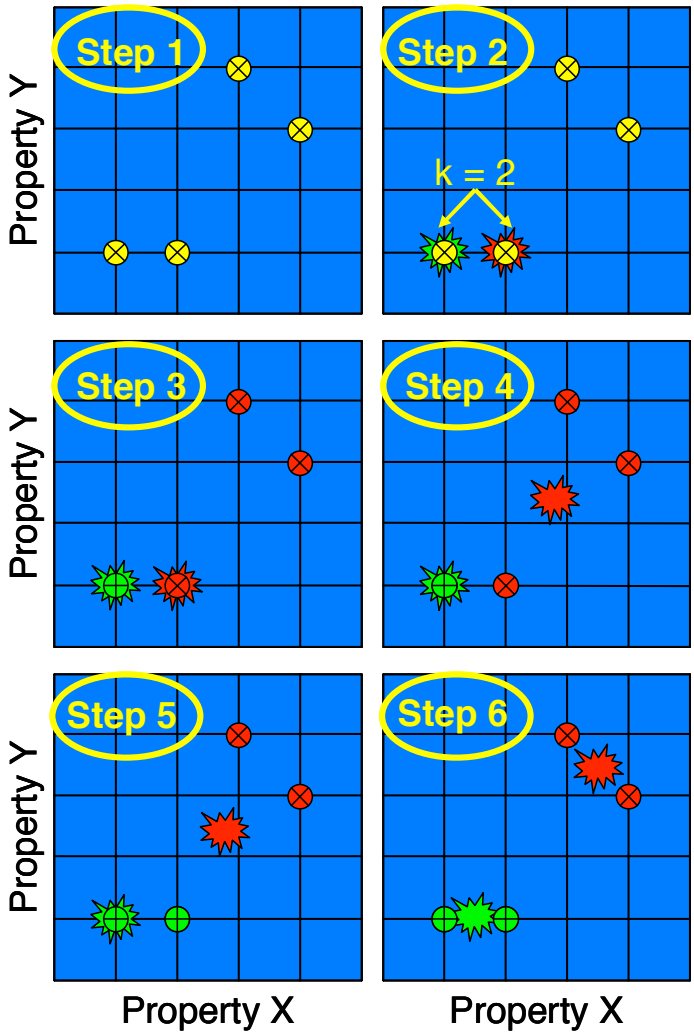


Figure 1.2 Illustration of several possible photon paths through a fractional cloud scene and grouping of several of these paths, all generated using the same statistical distribution of clear sky and cloud. Each path is composed of several arrows whose color is dependent on whether the photon is passing through clear sky or cloud.

Figure 1.3 The location of the three TWP ARM CART sites: Nauru, Darwin and Manus. The red line indicates the equator. Taken from the ARM webite: http://www.arm.gov/sites/twp.stm.



Step 1: Suppose we have four objects, each with two properties, X and Y. We can map these objects, using their properties as coordinates, onto a two-dimensional histogram.

Step 2: Now we select a number of clusters, k. In this case, k = 2. We initially place the centroids for these clusters randomly at object locations.

Step 3: Next we determine cluster membership by measuring the Euclidean distance between each object and the cluster centroids, and placing the objects into the cluster with the closest centroid.

Step 4: Now we adjust the centroid positions to account for cluster membership, minimizing the distance between the cluster's centroid and its members.

Step 5: Next we reiterate Step 3 using the new centroid locations.

Step 6: Finally we reiterate Step 4 using the new cluster memberships. In this case, after this step the centroid locations and cluster membership will no longer change, so the algorithm is complete.

Figure 1.4 Example of k-means cluster analysis algorithm.

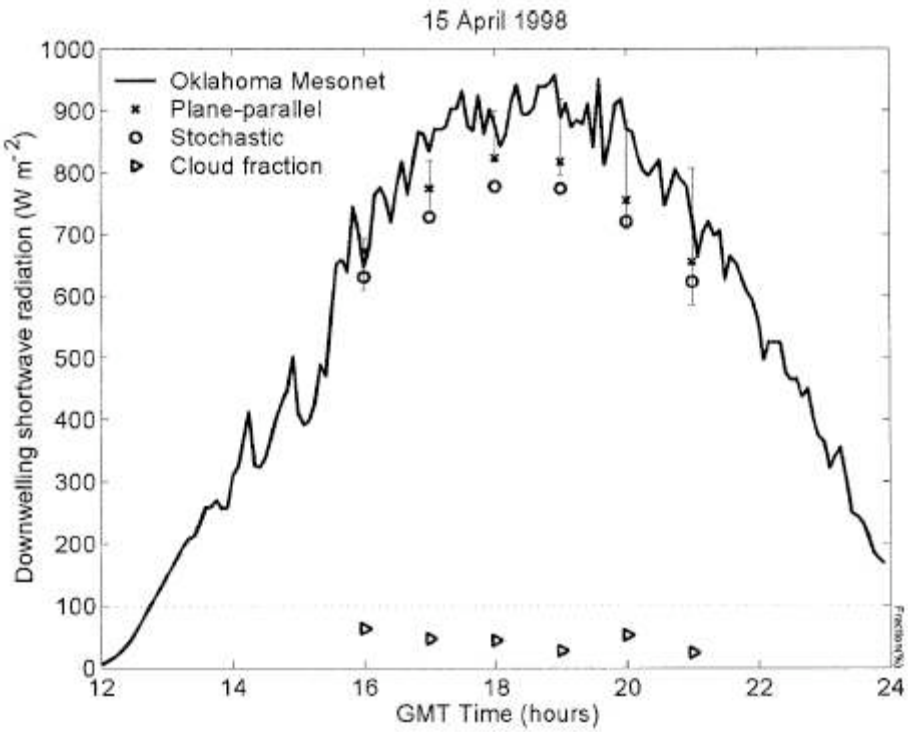


Figure 1.5 Model results of downwelling shortwave radiation for 15 Apr 1998 compared with averaged observations from the Oklahoma Mesonet. Cloud-fraction observations are from the MFRSR. Taken from Lane et al. (2002).

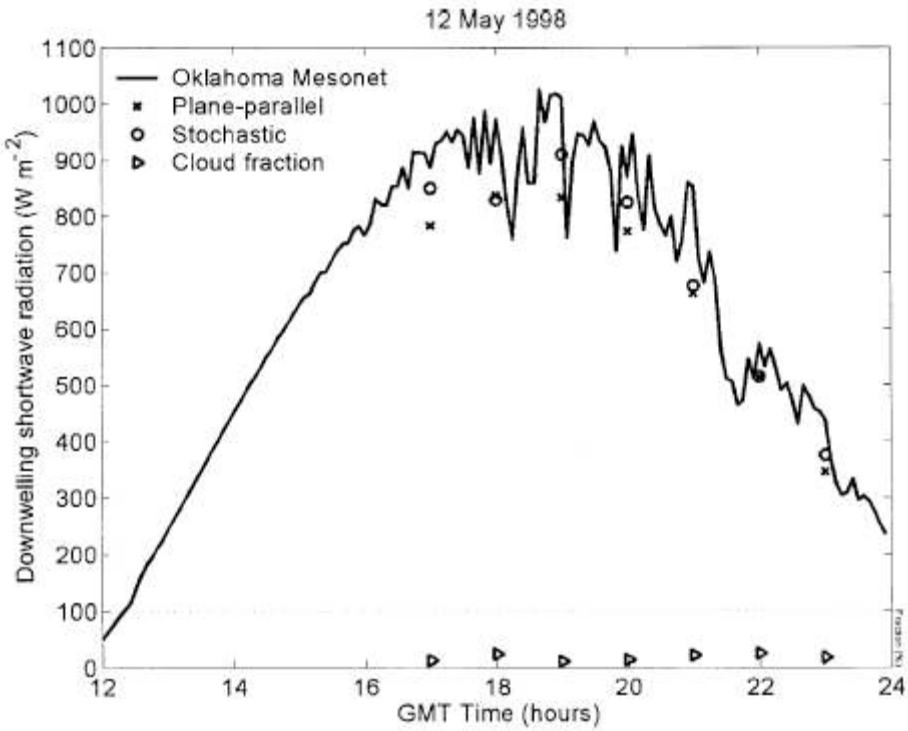


Figure 1.6 Model results of downwelling shortwave radiation for 12 May 1998 compared with averaged observations from the Oklahoma Mesonet. Cloud-fraction observations are from the MFRSR. Taken from Lane et al. (2002).

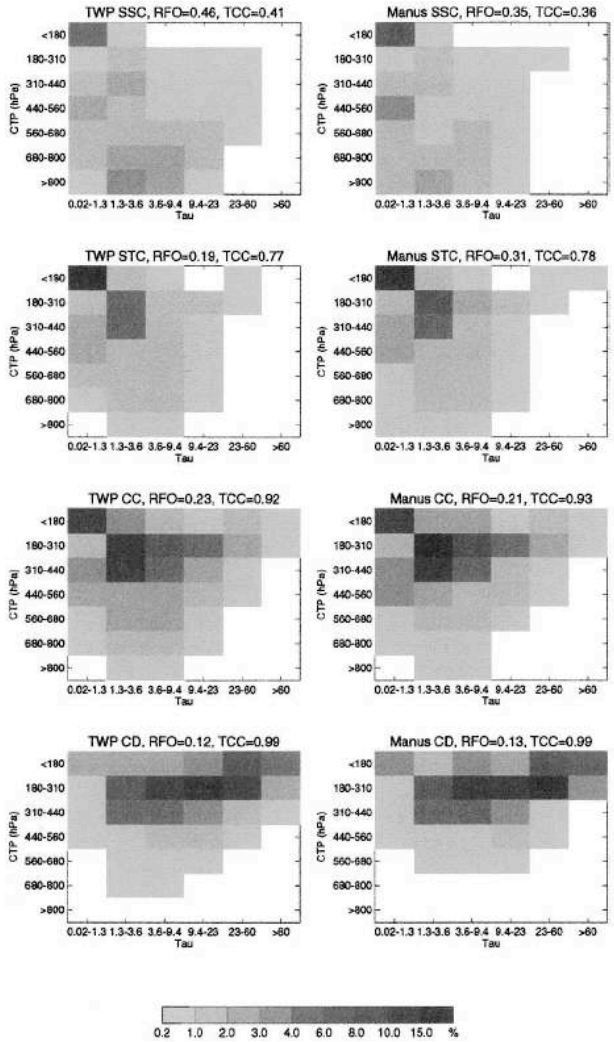


Figure 1.7 Cluster mean cloud-top pressure (CTP)–Tau histograms for 1999–2000 as identified (left) by cluster analysis for the TWP and (right) by projection for Manus. The regimes are suppressed shallow clouds (SSC), suppressed thin cirrus (STC), convectively active cirrus (CC), and convectively active deep cloud (CD). Also shown is the relative frequency-of-occurrence (RFO) and the total cloud cover (TCC). Taken from Jakob et al. (2005).

CHAPTER 2 – EVALUATION OF THE STOCHASTIC MODEL

In the last decade numerous advances have been made in parameterizing cloud-radiation interactions. One recent area of improvement involves statistical representations of the influence of three-dimensional cloud fields on the domain-averaged radiation. In this chapter, an off-line comparison of a traditional plane-parallel shortwave radiative transfer model against a stochastic model and observations is performed using four years of data from the ARM TWP CART site at Nauru Island. Statistical cloud properties are derived from observed cloud chord lengths and input into the stochastic model. A simple parameterization based on criteria developed from these results is applied to the plane-parallel modeled downwelling shortwave flux to represent the impact of complex cloud field geometry on the domain-averaged radiation.

A key component to the stochastic model simulations described in this chapter is the distribution of cloud size and spacing over the observed domain. The nature of cloud variability over differing spatial scales has been the subject of much research over the past few decades. Numerous studies have investigated cloud size distribution using aircraft observations (e.g. Planck, 1969; Hozumi et al., 1982; Sachs et al., 2002, Rodts et al., 2003), satellites (Wielicki and Welch, 1986; Cahalan and Joseph, 1989; Lovejoy and Schertzer, 2006, among others), radars (Kassianov et al., 2005a) and ground-based radiometers (Lane et al., 2002). From these studies several methods of representing cloud size distribution over varying spatial scales have been proposed, including the use of a decaying exponential (Markovian) function, a single or double power law, or a multifractal approach (Lovejoy and Schertzer, 2006). Complicating the debate is the

question of whether scale breaks exist where the description of cloud distributions shift and where they may occur (von Savigny et al., 2002; Zhao and Di Girolamo, 2007). For this study, previous work performed by Veron et al. (2008) found that most of the chord lengths measured at the ARM TWP sites were significantly less than 1 kilometer. For cases such as this, a Markovian distribution of cloud chord lengths has been found to be a reasonable approximation (Plank, 1969; Hozumi et al., 1982; Wielicki and Welch, 1986; Cahalan and Joseph, 1989). It is important to note that this approximation has significant effects on the performance of the stochastic model, and as a result, the appropriateness of the stochastic approach is dependent on many factors, including geographic location, time of year, and synoptic conditions. Much of the focus of this study then is to determine under which atmospheric and radiative conditions the stochastic model performs well and conversely under which situations it does not perform well, with a view to adapting the stochastic approach for use in an AGCM.

In section 2.1 the stochastic and plane-parallel models are described briefly as are the observed cloud statistical data used to force them. Section 2.2 contains the results and analysis of the model runs, and an evaluation of them performed using independent observations. In section 2.3 situations are identified where 3D radiative transfer is important and a simple parameterization is applied to the plane-parallel model to account for cloud-field morphology. The implications and conclusions from this study are discussed in section 2.4.

2.1 MODEL AND DATA DESCRIPTION

Stochastic model

The stochastic model used in this study, known as DSTOC, is based upon a model originally developed by Malvagi et al. (1993), though it is a more generalized form and has been modified for use in subsequent studies (Byrne et al., 1996; Lane et al., 2002; Lane-Veron and Somerville, 2004). Cloud fraction, liquid water content, SZA, fractional length of daylight, effective liquid droplet radius, surface albedo, and a measure of cloud scale and spacing are required input for the model at each time step. Model output includes ensemble-averaged radiative fluxes for each spectral band and model layer as well as in-cloud and clear-sky pathlengths.

DSTOC utilizes profiles of temperature, carbon dioxide, and ozone at thirty-two unequally spaced vertical layers taken from the McClatchey et al. (1972) climatology. The McClatchey moisture profile in the tropics was found to be significantly drier than that observed at Nauru so NCEP model-derived specific humidity was used instead. The horizontal domain of the model is related to the scale of the inhomogeneity in the cloud field; in this case the domain size is comparable to a relatively high resolution AGCM grid cell, approximately 20 kilometers per side. LWP is allocated to each liquid cloud layer using hydrometeor reflectivity measurements taken from the TWP ARM CART site at Nauru Island and then converted into cloud water content using cloud fraction and layer thickness, from which volume extinction and absorption coefficients are derived (Byrne et al., 1996; Lane-Veron and Somerville, 2004). Climatological values for surface albedo and effective ice particle radii are used. Surface albedo is set to 0.2,

effective ice particle radius is set to 23 microns, and liquid droplet radius uses a temperature-dependent equation with a minimum value of 6 microns. A single-moment temperature dependent function is used to partition cloud water and ice, and a random overlap assumption is used for adjacent cloudy layers.

The shortwave radiative transfer scheme is based upon the exponential sum-fitting method from Wiscombe and Evans (1977). It uses a discrete ordinate method with an approximate iterative technique for solving the radiative transfer equation (Byrne et al., 1996). The incoming solar radiation is separated into thirty-eight unequally spaced spectral bands, each of which may contain a maximum of two absorbing gases. Water vapor, ozone, carbon dioxide, and molecular oxygen are the gases currently included in the model: water vapor and ozone are the most commonly used absorbers, as they are the most important for many of the bands. Isotropic scattering is assumed. The solution is an ensemble mean of multiple clear and cloudy-sky scenarios sharing the same statistics. Solving for an ensemble of stochastic realizations generates the statistical variances of cloud-field properties required to calculate the non-linear reflectance of clouds. The standard time-independent radiative transfer equations are modified to contain two additional terms that describe the cloud-field geometry using conditional linear probabilities (Byrne et al., 1996) as shown below

$$\Omega \cdot \nabla\left(p_i I_i\right) + \sigma_i p_i I_i = \sigma_{si} \int_{4\pi} f_i\left(\Omega \cdot \Omega'\right) p_i I_i\left(\Omega'\right) d\Omega' + \frac{p_j \bar{I}_j}{\lambda_j} - \frac{p_i \bar{I}_i}{\lambda_i}, \; i = 0, 1, \; j \neq i \quad (2.1)$$

where $i = 0$ denotes clear sky and $i = 1$ denotes cloud. In this equation $p_i(r)$ is the probability of the *ith* material (clear sky or cloud) being at position r (the isotropic

assumption allows us to leave r out of the equation). Ω is the angular photon path direction, σ is the macroscopic total cross section, σ_s is the macroscopic scattering cross section, and f is the single-scatter angular redistribution function. $\bar{I}_i(r,\Omega)$ and $I_i(r,\Omega)$ are conditional ensemble-averaged intensities. $I_i(r,\Omega)$ is conditioned upon position r being located in fluid i, while $\bar{I}_i(r,\Omega)$ is conditioned upon position r being located at an interface between fluid i and j, with fluid i to the left (Malvagi et al., 1993). The statistical line theory needed to calculate the stochastic distribution of clear sky and cloud requires a probability distribution function of cloud chord lengths. Currently this function is governed by Markovian statistics, which simplifies the calculations and acts as one of the closure methods mentioned earlier. In addition, equation 2.1 actually represents two equations with four unknowns; $\bar{I}_1, I_1, \bar{I}_0$ and I_0. A simple closure method suggested by Adams et al. (1989) is:

$$\bar{I}_i = I_i \,. \tag{2.2}$$

Using Markovian statistics and removing photon scattering, these two closures still calculate an exact solution. With scattering, it has been shown the solution is approximate but still very good (Adams et al., 1989). The input required by the stochastic component is the sum of observed chord lengths. This information is then converted into a distribution of cloud chord lengths in the model, assuming Markovian statistics. The solution is the unconditional ensemble average intensity.

$$\langle I \rangle = p_0 I_0 + p_1 I_1 . \tag{2.3}$$

It should be noted that the stochastic model is too computationally costly to be directly coupled to an AGCM, therefore any adjustment to shortwave terms in an AGCM resulting from work done with this model would likely take the form of a look-up table, or possibly a fitted non-linear polynomial.

Column Radiation Model

The Column Radiation Model (CRM) is a standalone version of the plane-parallel radiative transfer code employed in the NCAR Community Climate Model (CCM3; Kiehl et al., 1998). The CRM is representative of shortwave radiative transfer codes used in many present-day AGCMs. The CRM utilizes the Delta-Eddington approximation described in Briegleb (1992) to solve the radiative transfer equation. The shortwave spectrum is divided into eighteen unequally-spaced bands with wavelengths ranging from 0.2 to 5.0 microns. Absorption sources include ozone, carbon dioxide, water vapor, and oxygen as well as cloud water and ice. Scattering sources include molecular, cloud, and aerosol, with isotropic scattering assumed between vertical layers. Profiles of temperature, carbon dioxide, and ozone at thirty-two unequally spaced vertical layers match those used by DSTOC and are taken from the McClatchey et al. (1972) climatology. The moisture profile also matches that used in DSTOC and is derived from NCEP model results. A double-moment microphysical function is normally used to partition cloud water and ice, but was replaced in this study by a temperature-dependent

single moment function that assumes clouds below -15°C are ice clouds and above are water clouds, in order to more closely match the stochastic model microphysics. A series of CRM runs were performed utilizing the double moment and single moment functions, and the resulting difference was negligible. The in-cloud microphysical properties for each layer are homogenous with adjacent cloudy layers overlapped randomly. For cloudy layers below -15°C LWP is allocated among the layers based on radar reflectivities. For temperatures below -15°C, the clouds in both models are treated as ice with an ice content of 0.015 gm^{-3}.

Model data

The forcing data used in this study are taken from observed data and Value-Added Products (VAPs) from the ARM TWP CART Nauru Island site. All measurements are averaged hourly, and then interpolated to thirty-two vertical layers. Cloud amounts, base and top heights are taken from the ARSCL VAP. The ARSCL VAP combines measurements from the Vaisala Ceilometer, Micropulse Lidar (MPL) and Millimeter Wavelength Cloud Radar (MMCR), and provides an estimate of layered cloud amount with a vertical resolution of 45 meters and a temporal resolution of 10 seconds (Clothiaux et al., 2000; Kollias et al., 2005). The ARSCL VAP also contains an estimate of hydrometeor reflectivity provided by MMCR measurements, which is used to allocate LWP in the vertical layers containing liquid cloud.

LWP is derived from the ARM two-channel (23.8 and 31.4 GHz) microwave radiometer (MWR) located at the Nauru Island facility using the Turner et al. (2007)

algorithm. The algorithm uses both a statistical and physical-iterative retrieval to derive LWP and precipitable water vapor from MWR measurements. Brightness temperature corrections are applied before the retrievals and the "monoRTM" radiative transfer model (Boukabara et al., 2001) is used to determine LWP. As mentioned earlier, LWP is allocated in each vertical layer containing liquid water clouds utilizing a scaling factor proportional to the reflectivity of the layer.

Downwelling shortwave broadband radiation data are taken from the ARM Shortwave Flux Analysis VAP, and are based on measurements taken from an ARM Sky Radiation (SKYRAD) radiometer. Global hemispheric shortwave irradiance is measured with an unshaded pyranometer with a hemispheric field of view, while diffuse shortwave irradiance is measured with a shaded pyranometer. The measurements are taken at one-minute intervals and averaged over each daytime hour. The Shortwave Flux Analysis VAP includes an algorithm that takes measurements of direct, diffuse, and total shortwave surface irradiance, detects periods of clear-sky, and fits a curve estimating the clear-sky shortwave irradiance over all time periods using SZA as the independent variable (Long, 2001). The fit-estimated surface clear-sky irradiance from this VAP is used for comparison against model results during clear-sky conditions. The total effective cloud amount is also calculated with an empirically derived formula that uses the measured and estimated diffuse and direct shortwave irradiance (Long et al, 1999; Long and Ackerman, 2000; Long et al., 2006).

The stochastic model requires a statistical description of the cloud size and spacing, given in the form of a probability distribution function (pdf) of cloud chord lengths. The cloud chord lengths are calculated using NCEP model-derived wind data and ARSCL

cloud fraction. In previous studies, a distribution function of individual cloud chord lengths is created using a narrow field-of-view upward pointing surface instrument such as a lidar or MFRSR to determine when clouds are overhead, and then converting the time a cloud is detected into a distance (chord length) using the wind speed at the detected cloud base height (Lane et al., 2002). In these studies it was found that for the ARM TWP sites the preponderance of cloud chords lengths calculated from ground-based remote sensors are at scales of 1 km and less, with greater than 50 percent of the chords having size of 200 m or less. The compilation of these chords provides a pdf of the cloud chord lengths, which along with the distribution of clear-sky chord lengths is used to produce the volume cloud optical depth necessary for the stochastic model to generate photon paths. Lane et al. (2002) determined that the observed Markovian distribution of chord lengths produced realistic results under low-level broken or scattered sky conditions. The Veron at al. (2008) study also found that the distributions at the ARM TWP sites for all cloudy chord lengths were well represented by a Markovian distribution. For this paper, the product of the hourly mean wind speed and the ARSCL cloud fraction at height is treated as the integration of the pdf of chord lengths. The stochastic model then solves for the pdf inversely assuming a Markovian distribution, similar to the observed chord length pdfs.

In order to complete the simulations described below, all input data from these diverse sources must be available and be able to meet certain criteria. Of the 35064 hours spanning from the beginning of 2001 to the end of 2004, 8408 of those hours were used in the simulations. Naturally, the largest constraint is that only daytime hours were used. 16,403 hours or 47 percent of the total hours were available from the Shortwave Flux

Analysis VAP. Most of the missing hours occurred at nighttime, while some were due to missing data, or a value of the cosine of the SZA below 0.02. LWP derived from the MWR instrumentation was available for 26,845 hours, or 77 percent of the time. The other hours were due to missing data or times with precipitation occurring (which accounted for a little under 3000 hours or 9 percent of the time). Removing the hours when precipitation is occurring could have a small effect on the model results, however as the models we are using do not currently have a suitable radiative treatment for precipitating hydrometeors, this is an issue that must be addressed in a later study. 12,180 hours (35 percent) were available for both data sources. Cloud fraction data taken from the ARSCL VAP were available for 21,934 hours (63 percent). Shortwave and cloud fraction data were available concurrently for 9831 hours, and LWP and cloud fraction data were available concurrently for 18,380 hours. The times when shortwave, LWP, and cloud fraction data were all available totaled 8408 hours, or 24 percent of the total hours from 2001 to 2004, approximately one-half of the total daylight hours. Figure 2.1 shows the concurrent availability of these datasets in a Venn diagram.

2.2 MODEL RESULTS

Clear sky conditions

The initial comparison of DSTOC and CRM involves simulations during clear-sky conditions. To this end both models were run for 241 hours between 2001 and 2004, during which times the ARSCL VAP detected no clouds. For downward solar surface irradiance (SSI), a mean difference of 6 Wm^{-2} with a standard deviation of 8 Wm^{-2}

represents agreement between the CRM and DSTOC models of 1.2 ± 1.5 percent, which gives confidence that the models are performing similarly in a control situation. Part of this difference may be because the spectral weighting differs slightly: the CRM solar spectral width spans from 0.2 to 5.0 microns and is broken into 18 bands, compared to that of DSTOC which spans from 0.25 to 5.0 microns and is broken into 38 bands. In addition a climatological mean value of aerosol optical depth used in the CRM (0.07) was derived from measurements taken from the ARM Cimel Sun Photometer, while DSTOC does not currently account for aerosol optical depth.

A comparison of the SSI produced by the models and measurements taken during times when the ARSCL VAP detected no clouds are shown in Figure 2.2. The CRM and DSTOC models generated mean surface irradiances of 22 ± 42 and 16 ± 43 Wm^{-2} higher than observations, respectively. When comparing the clear-sky model results against the algorithm-generated clear-sky estimations taken from the Shortwave Flux Analysis VAP, a separate analysis of clear-sky, there is no discernible positive or negative bias between model and clear-sky estimations (Figure 2.3). The clustering occurring in both figures is a result of averaging over hourly intervals for a relatively small set of data. The mean algorithm-generated clear-sky estimated SSI differed from the simulated CRM and DSTOC SSI by 7 ± 31 and 1 ± 33 Wm^{-2} respectively. This implies that the observed mean surface irradiance is 15 Wm^{-2} lower than the algorithm-generated clear-sky estimate. The most likely explanation for this discrepancy is that the ARSCL VAP is dependent on upward-pointing narrow-beam instruments for cloud detection and as such it is possible for undetected clouds to be present in the wider field-of-view. To test this data from the Total Sky Imager (TSI) are examined during the 241 hours the ARSCL

VAP detected no clouds. Out of the 241 hours, 177 coincided with times when TSI data are available. The TSI detected some amount of thin or opaque clouds during most of these hours. The mean opaque cloud fraction detected by the TSI over this period was 0.15 and the mean thin cloud fraction was 0.04. In addition, during the hours the ARSCL product detected no clouds in the atmosphere, the Microwave Radiometer did detect liquid water, however as there was no cloud in which to allocate this liquid water it was not input into the models.

Monthly and diurnal cycles

The median percent difference between the modeled and observed SSI is shown as a function of month of the year in Figure 2.4. One thing to note is that both models perform reasonably well from January to April and then diverge from observations with a positive bias from April to September. At the end of the year both models exhibit a negative bias with respect to observations. During the April to September period the mean DSTOC results are closer to observations than the CRM. This may be attributed to increased convective activity at Nauru during these months, due to warm SSTs and lighter surface winds. Increased occurrences of cumulus towers and broken cloud fields coincide with an increase in the number and importance of inter-cloud radiative interactions, making a stochastic approach preferable. Also included in Figure 2.4 is the median cloud coverage for these months, calculated using two different methods. One takes the largest value for ARSCL cloud mask from all cloudy layers for each hour and uses it as an estimate of the total cloud coverage, while the other method utilizes the

algorithm developed by Long et al. (2006), which uses measurements of diffuse and total shortwave radiation at the surface along with estimates of clear-sky flux. The correlation coefficient between mean ARSCL cloud coverage and the difference between the model divergences from observed SSI is 0.67, with the discrepancy between model results generally increasing as cloud coverage decreases. This solidifies the point that cloud amount is an important component in determining when the stochastic approach is most appropriate. It is interesting to note that the largest discrepancy between the Long et al. and ARSCL cloud coverage occur during roughly the same period the models diverge farthest from observed SSI. During this time the cloud coverage calculated by the Long et al. algorithm is higher than that of ARSCL, which coincides with the model divergence from observations, suggesting that less solar radiation is reaching the surface then the estimation of cloud coverage from ARSCL data implies. There are a couple of possible explanations for this: First, our use of the cloudiest layer of ARSCL data as an estimate for cloud coverage is tantamount to the use of a maximum overlap assumption. A study by Mace and Benson-Troth (2002) concluded that an assumption of random overlap could cause a positive bias in cloud coverage at Nauru for small separations between vertical cloudy layers, and that a maximum overlap assumption would be more appropriate in these instances. The study also found that for large separations between cloudy layers minimum overlap predominated at Nauru. Since the months at the beginning and end of the year are dominated by low thick boundary clouds, a maximum overlap assumption would be most appropriate and could cause the models to overestimate cloud coverage and underestimate SSI when compared to observations, as can be seen in Figure 2.4. From April to September increases in convective activity

could generate greater spacing between cloudy layers, which in turn could make a minimum overlap assumption most appropriate causing the models to underestimate cloud coverage and overestimate SSI compared to observations. Another explanation closely tied to the first is that thick low-level boundary layer clouds are likely to dominate direct and diffuse shortwave measurements and corresponding cloud coverage estimates whenever they are present. Finally, it should be noted that these two methods of estimating cloud coverage are using different instrumentation. The shortwave radiometers used in the Long et al. (2006) measurements have a nearly hemispheric field of view, while the instruments used in the ARSCL VAP are upward-pointing beams. While it has been shown that averaging nadir-view surface measurements over time can produce reasonable estimates of cloud coverage (Kassianov et al., 2005b), it should be recognized that there will undoubtedly be some discrepancies when comparing these measurements.

In Figure 2.5, the hourly model-generated SSI is bin-averaged by 10 Wm^{-2} increments and plotted against that observed over the 4-year period. DSTOC tends to overestimate when values of SSI are low, such as times when SZA is large or cloud amount very high, and improves as values of SSI increase. For small values of SZA DSTOC tends to underestimate. This is illustrated with a line fitted to the scatter-plotted DSTOC results, which has a slope of 0.98. The opposite appears to be true for the CRM results. The CRM begins by doing a good job for low values of SSI, and then overestimates SSI when the sun travels closer to the zenith, as illustrated by the slope of the CRM data fit being 1.09. The two models converge around the point where SZA is equal to 60 degrees. For SZA larger than 60 degrees the mean CRM-generated minus observed SSI is 2 Wm^{-2}

while for DSTOC it is 11 Wm^{-2}. For SZA smaller than 60 degrees the mean CRM-generated minus observed SSI is 22 Wm^{-2} while for DSTOC it is -10 Wm^{-2}. This result is not necessarily intuitive, as one might expect the closest approximation of a plane-parallel atmosphere to occur when the sun is directly overhead, or nearly so. For large SZA there is more likely to be shading effects where clear sky is inter-dispersed among cloud, while for the case of small SZA the shadowing effects between broken clouds would presumably be diminished. However, when cumulus towers are frequently present, as is the case here, the large SZA case could more closely approximate a 'plane-parallel' situation. Also other processes such as rough cloud tops, side illumination or leakage (photons scattering horizontally out of a cloud), can cause radiative interactions in broken cloud fields that can significantly affect the shortwave radiative fields (Varnai and Davies, 1999).

Liquid water path

Cloud field morphology with radiatively significant 3D effects is more likely to occur when certain atmospheric conditions are present. Broken cumulus fields and dissipating stratus clouds are potential examples of cloud regimes prone to developing 3D cloud geometry that has an effect on the shortwave radiative fields. Figure 2.6 shows the median LWP for times when each model generates SSI closer to observed values, bin-averaged by the summed vertical extent of all layers containing liquid cloud. For this study we have selected a 5 percent threshold to determine when one model is performing better than another, meaning if the observed value of SSI is 100 Wm^{-2}, a model must

generate SSI closer to 100 Wm^{-2} by more than 5 Wm^{-2} in order for us to determine it is outperforming the other model. When both models generate SSIs that are within the 5 percent threshold they are considered to be performing equally well. For liquid cloud vertical extents below three kilometers both models have nearly identical median liquid waters paths. Above three kilometers, however, DSTOC performs better when LWP is near the median for all hours, while the CRM performs better when LWP is above the median. The LWP difference between the CRM and DSTOC plots are 10, 19, and 33 gm^{-2} for bin-averaged vertical extents of 3.5, 4.5, and 5.5 km respectively. For hours when it is determined one model is outperforming the other, the median percent by which each model is closer to observations is 12 for the CRM and 13 for DSTOC. One implication of this is that while LWP does seem to be an important factor in determining how the models will perform as vertical liquid cloud extent becomes large, (i.e. greater than 3 kilometers) it does not seem to be a good indicator when liquid cloud extent is small. Identifying thresholds such as these allows development of a set of criteria by which to determine when cloud field geometry is complex enough to require an approach other than a plane-parallel approximation. In this case liquid water clouds with low LWP and a large vertical extent may indicate the presence of broken cumulus clouds, where a stochastic approach to radiative transfer is more appropriate than the plane-parallel one.

Cloud fraction

Figure 2.7 plots the top of the atmosphere (TOA) upwelling shortwave irradiance due to cloud scattering divided by the TOA downwelling SW irradiance and the difference

ratio between model-generated and observed SSI against cloud fraction for SZA above and below 60 degrees. Figure 2.7a shows the CRM performing better than DSTOC for low cloud fractions with SZA greater than 60 degrees. For higher cloud fractions both models underestimate SSI and consequently DSTOC produces SSI closer to observations for fractions above around 0.6. The magnitude of the difference between the two models is relatively low since for large SZA the amount of incoming solar radiation is small. Much of this difference can be attributed to the larger TOA upwelling cloud-reflected shortwave irradiance simulated by the CRM when SZA is greater than 60 degrees.

Figure 2.7b shows DSTOC with more TOA upwelling cloud-reflected shortwave irradiance for almost all values of cloud fraction. Though the percent difference between the models is smaller, the magnitude is larger than Figure 2.7a due to the small SZA. DSTOC underestimates SSI for low cloud fractions and the CRM generates results close to observations. For cloud fraction greater than 0.5 the CRM overestimates SSI and DSTOC generates results close to observations. The implication of this is that DSTOC performs well for mostly cloudy to nearly overcast conditions when SZA is small.

Diffuse and direct solar radiation

The amount of diffuse and direct radiation reaching the surface can be a useful metric in determining the amount and radiative effects of cloud in the atmosphere (Long et al., 1999; Long and Ackerman, 2000; Long et al., 2006). As the ratio of observed diffuse to total surface irradiance is currently being used to determine when clear-skies are present, it is not unreasonable to hypothesize that it may be useful in determining when cloud

morphology is having a significant effect on radiation fields. Figure 2.8 illustrates the median of this ratio plotted as a function of SZA for times when the models outperform one another and for all hours. For large SZA the ratio of diffuse to total radiation is quite high, which is not surprising as the amount of diffuse radiation scattered in the atmosphere will naturally increase as its pathlength from the top of the atmosphere to the surface increases; consequently as SZA decreases so too does this ratio. What is interesting is that the median ratio for DSTOC drops quickly and then flattens out, remaining roughly between 0.5 and 0.6 for SZA smaller than 80 degrees. The CRM ratio meanwhile stays at a very high value until around 60 degrees, where it too quickly drops and flattens out, but to much lower values generally below 0.4. This suggests that the stochastic model performs best when the ratio of diffuse to direct radiation is roughly equal, while the CRM tends to perform best when either direct or diffuse radiation dominates the surface irradiance. The ratio for all hours may be used as an indicator of how often one model is outperforming the other. For example, the ratio for all hours is closer to times when DSTOC outperforms the CRM for large SZA and times when the CRM outperforms DSTOC for small SZA. Long et al. (2006) noted that it was not unusual for partly cloudy skies to enhance the amount of diffuse radiation reaching the surface. From this we hypothesize that complex cloud field geometries in partly cloudy skies increase the radiative cloud interactions like those simulated in DSTOC, thereby increasing the amount of diffuse radiation reaching the surface.

2.3 PARAMETERIZATION DEVELOPMENT

Criteria selection

A preliminary set of criteria is developed that can be used to indicate when a stochastic approach to parameterizing SW radiative transfer through the cloudy atmosphere is more appropriate than the plane-parallel approach. In section 2.2, it is shown that DSTOC performs well when the average LWP is relatively low and the liquid cloud has large geometric vertical extent (i.e., greater than 3 kilometers). These criteria are consistent with a description of broken cumulus clouds with high vertical extent, providing space for inter-cloud radiative interaction in both the horizontal and vertical dimensions. The LWP threshold used for this set of criteria is the midpoint between the mean for times when DSTOC outperforms the CRM and the mean LWP for the reverse situation. Any LWP value found to be below this threshold during times when liquid cloud has a vertical extent greater than 3 kilometers is considered to be meeting the criteria necessary for use of the stochastic approach.

The cloud fraction and SZA analysis described in section 2.2 indicate that DSTOC performs well during times with moderate to high cloud fraction, regardless of SZA. In addition the diffuse to total SSI analysis indicates that DSTOC performs well when this ratio is between 0.5 and 0.6. These times are consistent with the presence of broken cloud fields with complex top shapes and spacing conducive to horizontal scattering between clouds. Therefore a second set of criteria for using the stochastic model over the plane-parallel model will be when cloud coverage is between 0.5 and 0.9 and the diffuse to total SSI ratio is between 0.5 and 0.6. It should be noted that when either the LWP-

vertical extent criteria or the cloud coverage-diffuse to total ratio criteria is met the stochastic approach will be used.

These criteria are used as an indicator of when the stochastic approach to resolving radiative effects due to cloud field geometry provides a significant benefit to surface shortwave flux calculations, and as such when they are met the CRM SSI results are replaced with the stochastic SSI. In addition, they are used to apply a simplistic parameterization of the stochastic approach described below.

Parameterization

A parameterization of the stochastic approach to SW radiative transfer was developed using a multiple regression technique applied to the following variables: LWP divided by vertical extent of liquid cloud, cloud fraction, and SZA. Our goal is to identify key atmospheric processes that have a discernible relationship to cloud field configuration, and as such a simple multiple linear regression seems an appropriate initial choice. The correlation coefficients for the difference between DSTOC and CRM SSI and these properties are 0.5 for cloud fraction, 0.12 for LWP divided by vertical extent of liquid cloud, and -0.37 for SZA. The parameterization is as follows:

$$\text{PARAM} = \text{DSTOC SSI - CRM SSI} = a + b * CF + c\left(\frac{LWP}{LCV}\right)\frac{1}{q_w} + d * \cos(SZA)$$

Where CF is cloud fraction, LWP is in gm^{-2}, LCV is liquid cloud vertical extent in kilometers and q_w is the density of water in gm^{-3}. The values for coefficients a, b, c and d are -12.8, 92.7, 267.7 and -81.5 respectively. The R^2 value for this parameterization is 0.6, meaning this parameterization explains 60 percent of the original variability.

Results

Figure 2.9 shows a comparison of the model results along with the CRM with DSTOC and CRM with parameterization results. Also included is a plot which, for each hour, uses the results from the model that generates SSI closest to observation. This plot is labeled 'Best Performing', and can be thought of as a best case scenario for this simple version of the parameterization criteria. The CRM with DSTOC and CRM with parameterization results use the CRM calculations unless the above-mentioned criteria are met, in which case the DSTOC calculations or parameterization are used, respectively. The plot shows all five model results generally overestimating observed SSI for large values of SZA and then underestimating SSI as SZA becomes small. The CRM overestimates SSI for almost all values of SZA, only underestimating when SZA is less than 15 degrees. DSTOC experiences the largest variability, ranging from overestimating by 12 percent for SZA above 80 degrees to underestimating by 9 percent for SZA less than 10 degrees. As expected, the 'Best Performing" results are the closest to observations, staying within 5 percent for almost all values of SZA and very close to zero for much of the time. Both the CRM and CRM w/ DSTOC simulations do well for large values of SZA and then tend to underestimate as SZA decreases. Perhaps the most encouraging results are for value of SZA between 80 and 60 degrees and between 35 and 15 degrees. During these intervals the CRM with DSTOC and CRM with parameterization results outperform both the CRM and DSTOC, and in some cases even the 'Best Performing' simulation.

Figure 2.10 shows a monthly comparison of the model results, with the same simulations found in Figure 2.9. Again the "Best Performing" plot is closest to observations throughout the year. The CRM with DSTOC results remain very close to those of DSTOC for much of the year, which is interesting as the criteria used for substituting DSTOC results for those of the CRM are met a little less than half of all hours. In general, CRM with DSTOC and CRM with parameterization results show improvement over the CRM from May to September, but then largely underestimate at the end of the year. One interesting point to note is that the CRM with parameterization results do extremely well from May to August, better even than those of "Best Performing".

For all hours, the median difference from observed SSI is 10 Wm^{-2} for the CRM, -3 Wm^{-2} for DSTOC, 2 Wm^{-2} for the "Best Performing" simulation, -9 Wm^{-2} for the CRM with parameterization, and -5 Wm^{-2} for the CRM with DSTOC combination. The parameterization criteria are met 46.9 percent of the time in the CRM model, 9.4 percent of the time due to the LWP criteria and 38.8 percent of the time due to the cloud coverage and diffuse to total SSI ratio criteria. 1.3 percent of the time both criteria are met.

2.4 CONCLUSIONS

These results indicate that a) the stochastic approach does in fact have the ability to significantly improve SW radiative calculations through cloudy atmosphere and b) it should be possible to develop a set of objective criteria by which to know when bias due to complex cloud field geometry is significant. The stochastic model is currently limited

in its ability to use different descriptions of variability in the cloud field, using only a Markovian description of the cloud sizes in the field. However, this statistical approach outperforms current binary cloud-radiation models that do not account for 3D cloud effects.

The comparison between off-line stochastic and traditional plane-parallel model runs suggest that the relationships between cloud coverage and SZA, LWP and vertical liquid cloud extent, and diffuse and direct SSI all have identifiable characteristics during times when the two models perform well. There are also seasonal differences in model performance, with the stochastic model showing the greatest improvement over plane-parallel results from April to September. These features can be used to derive objective criteria indicating when cloud-field geometry impacts the domain-average shortwave radiative fields. These criteria are representative of the clouds present at the ARM TWP CART site and may not be appropriate for the entire tropics, let alone the rest of the globe. The following chapter expands the scope of the areas being observed to include the other two ARM sites in the Tropical Western Pacific.

The stochastic approach is too computationally expensive to run directly as a parameterization in an AGCM. A simple parameterization that captures the impact of the stochastic approach was developed from stand-alone stochastic and traditional plane-parallel model runs using ground-based observations as a reference. The new parameterization provides an improvement in the CRM downwelling shortwave surface irradiance calculations under certain conditions. The next step is to develop a more sophisticated approach to parameterizing the stochastic approach to modeling shortwave radiative transfer that will account for more variability in the modeled SSI differences.

This may include a sinusoidal dependence on SZA or inclusion of a dynamical term and may result in an adjusted cloud fraction as opposed to an irradiance correction.

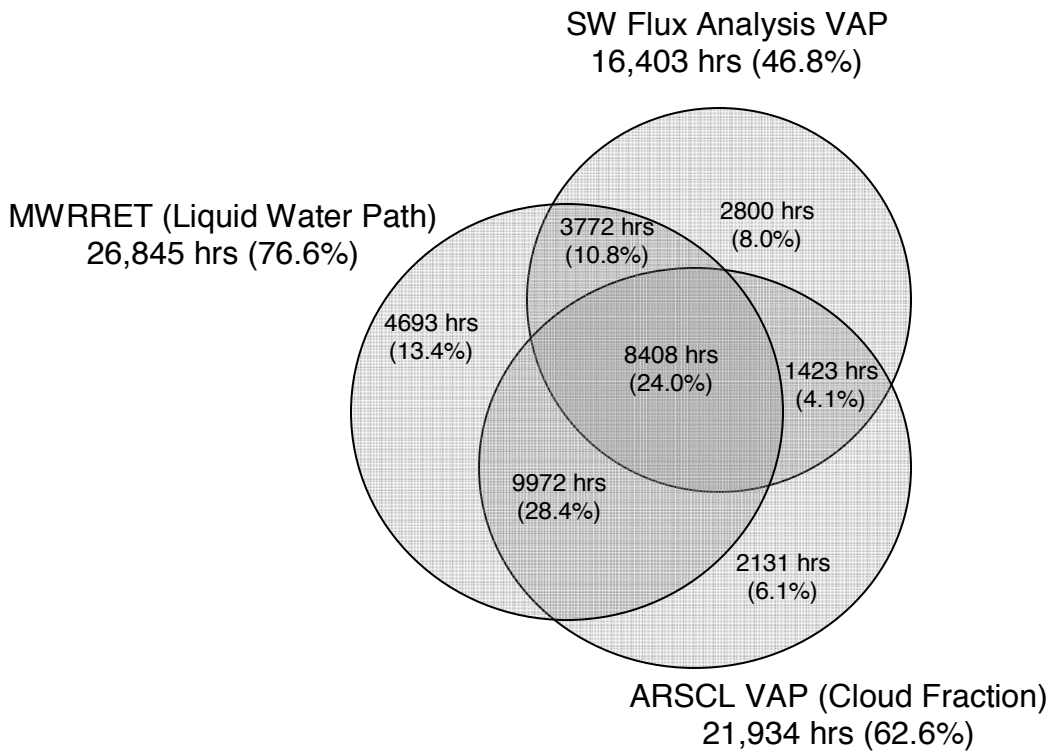


Figure 2.1 Venn diagram showing the concurrent availability of ARM data at the Nauru facility from the beginning of 2001 to the end of 2004.

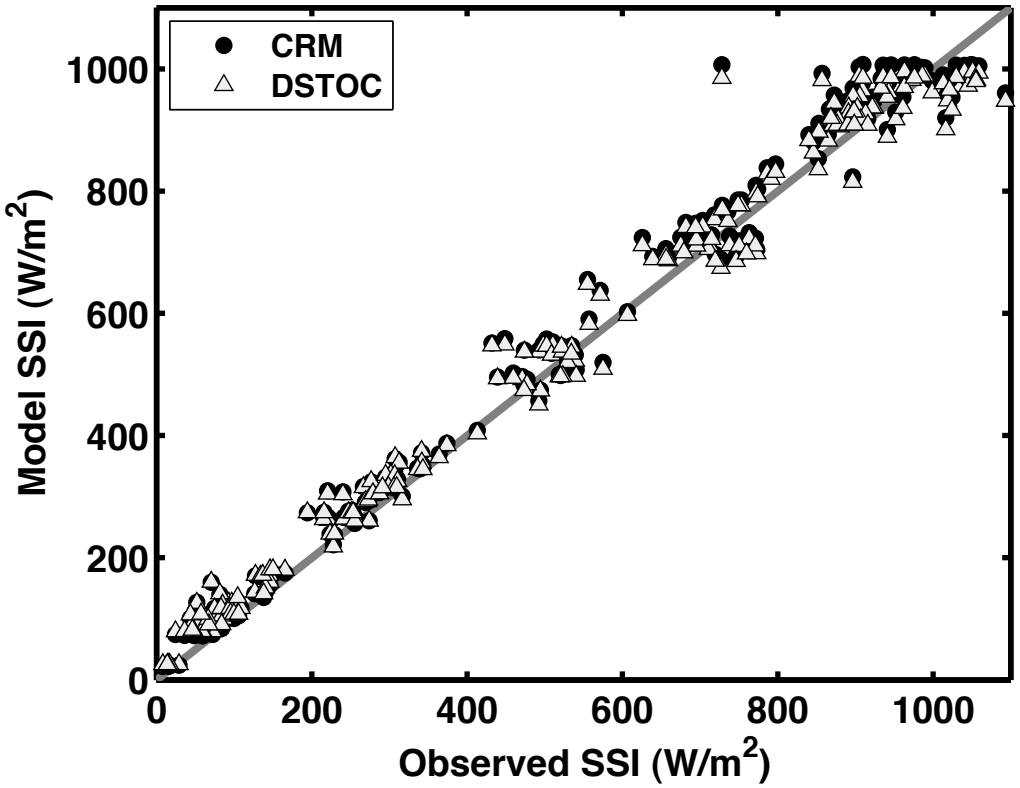


Figure 2.2 Scatterplot comparing observed solar surface irradiance to modeled for all clear-sky hours detected in the four-year 2001-2004 period. DSTOC results are indicated by solid circles and CRM results are indicated by triangles. The grey diagonal line indicates a one-to-one match between model and observation or fit-estimate.

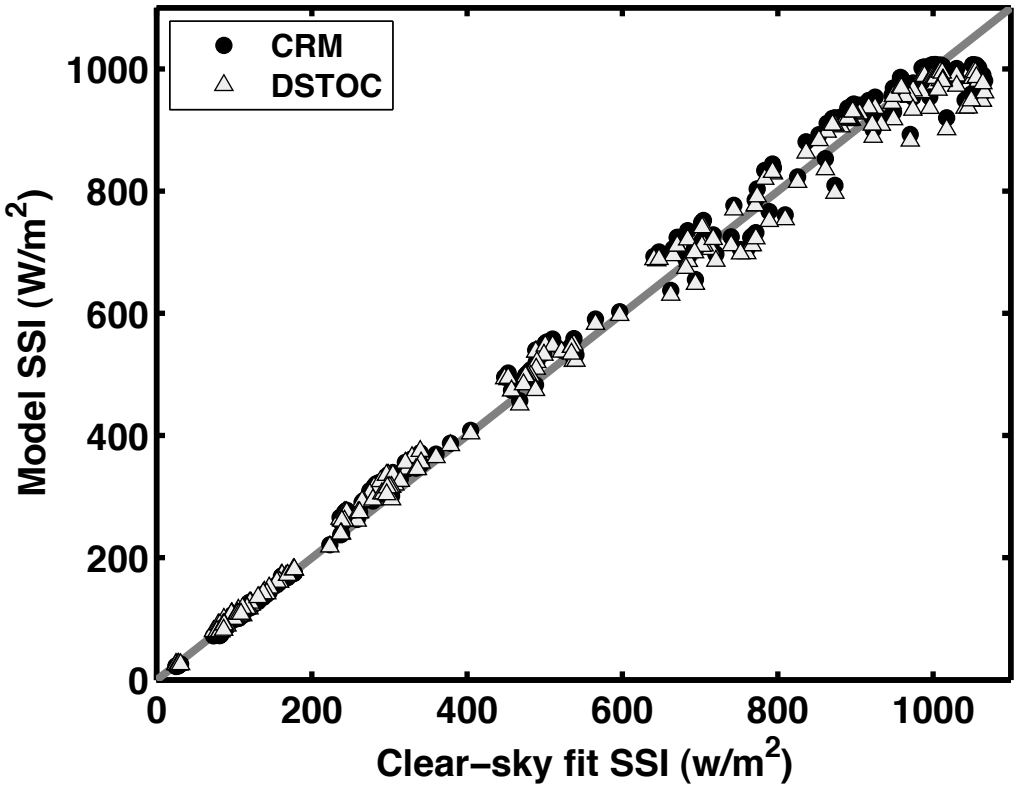


Figure 2.3 Scatterplot comparing clear-sky fit estimated solar surface irradiance to modeled for all clear-sky hours detected in the four-year 2001-2004 period. DSTOC results are indicated by solid circles and CRM results are indicated by triangles. The grey diagonal line indicates a one-to-one match between model and observation or fit-estimate.

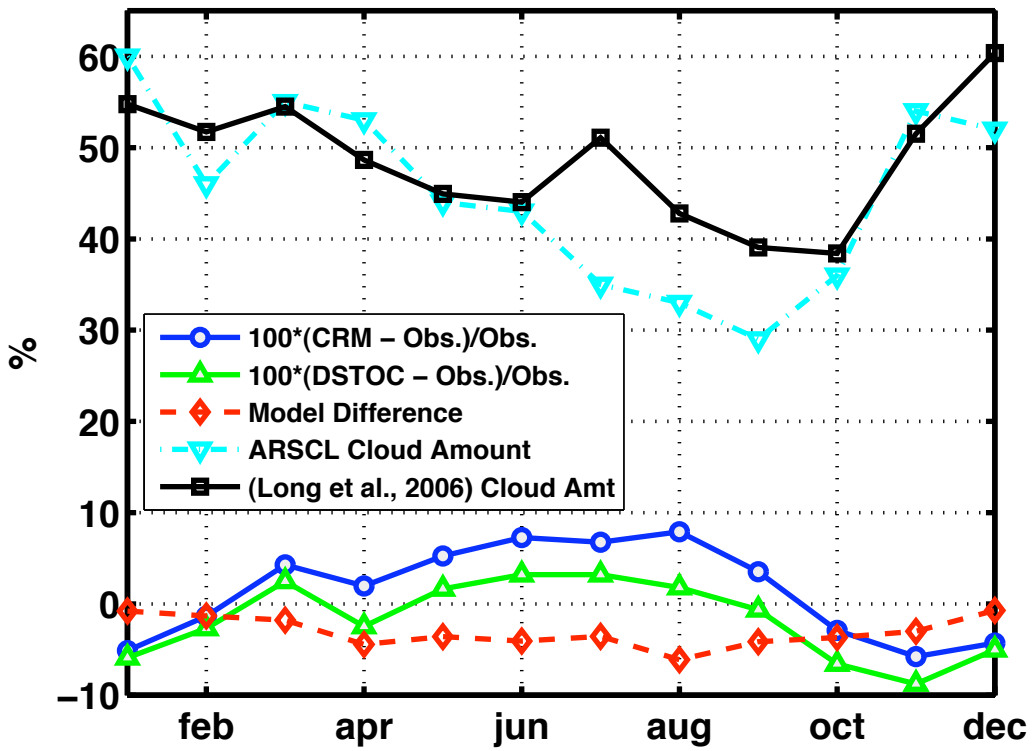


Figure 2.4 Monthly comparison of model-generated median SSI shown as a percent-difference from observations. CRM results are shown as circles on a solid blue line. DSTOC results are shown as upward pointing triangles on a solid green line, and the difference between the two models is shown as diamonds on a dashed red line. Monthly median cloud coverage calculated from the ARSCL layer with the highest cloud fraction is displayed on a cyan dashed line with downward pointing triangles, while cloud coverage as generated using the Long et al. (2006) algorithm is displayed with squares on a black solid line.

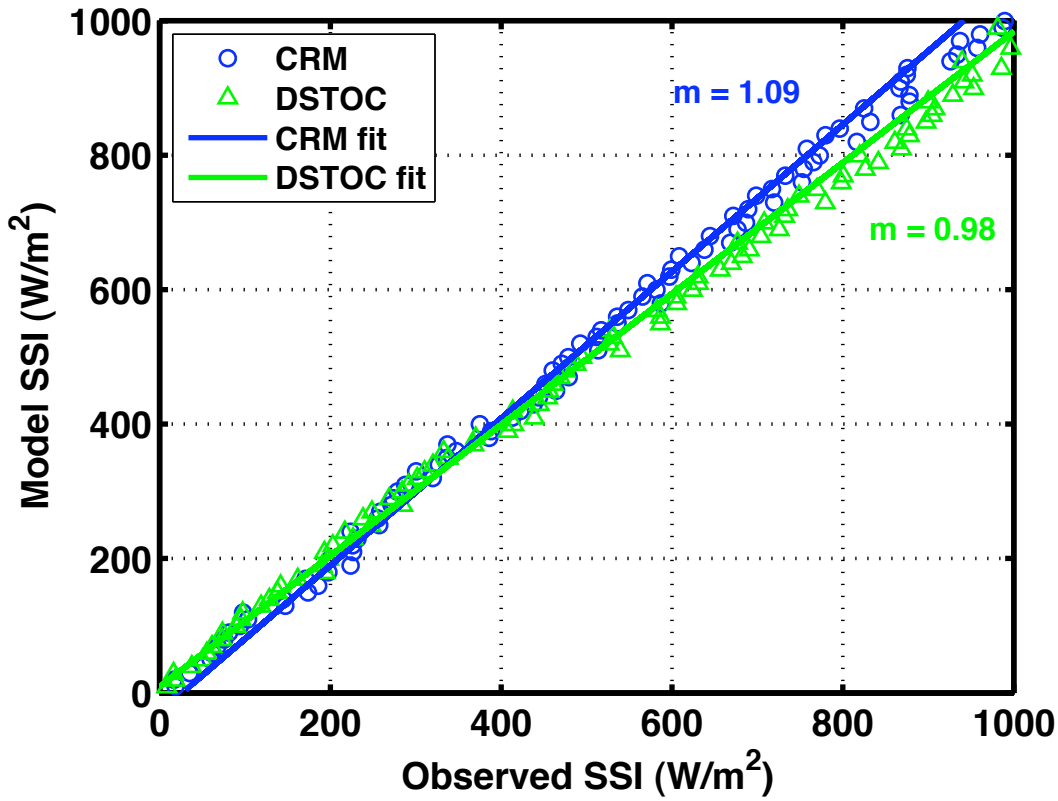


Figure 2.5 Comparison of model-generated versus observed solar surface irradiance. Results are bin-averaged at 10 Wm^{-2} observed SSI intervals using median values of model-generated SSI and plotted against one another. Blue circles indicate CRM results while green triangles indicate DSTOC results. The solid blue line represents the linear fit of the CRM results generated using the least-squares method while the green line represents the linear fit of DSTOC results. The numbers next to the linear fit indicate the slope of the line.

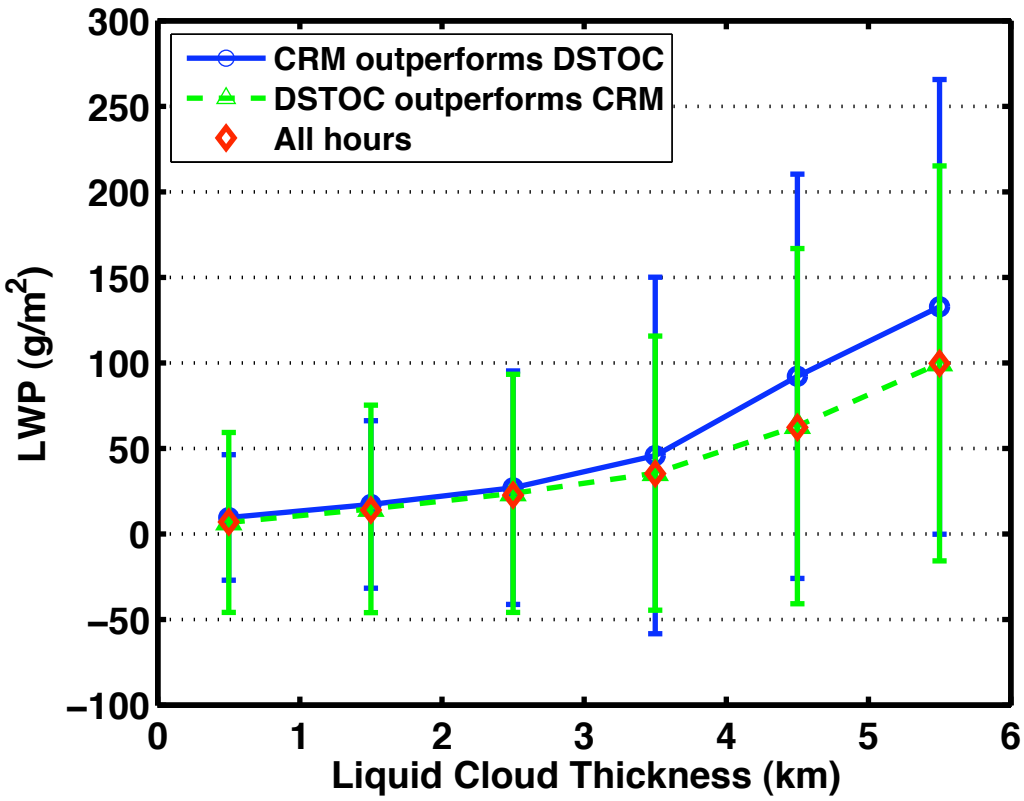


Figure 2.6 Median liquid water paths for both models during times when they generate SSI closer to observations as a function of the bin-averaged vertical extent of liquid cloud. CRM outperforming DSTOC is indicated by circles with a solid blue line and DSTOC outperforming CRM is indicated by triangles with a dashed green line. Red diamonds indicate the median liquid water path for all hours. Error bars extend to one standard deviation.

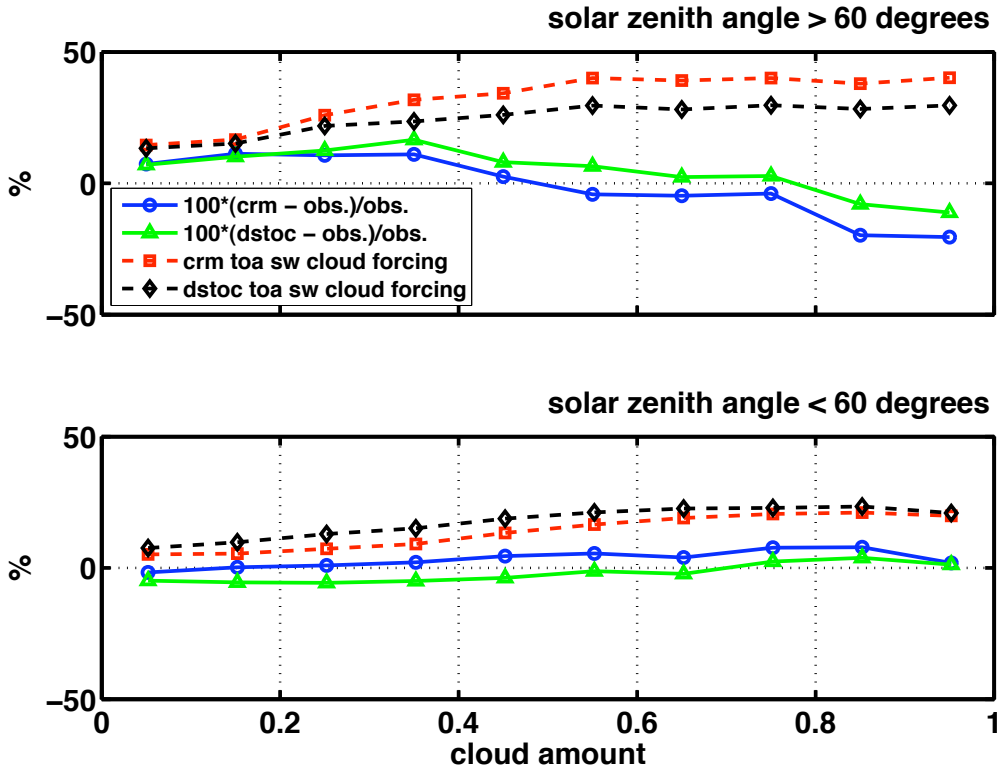


Figure 2.7 Normalized difference between model-generated and observed solar surface irradiance along with model-generated TOA upwelling shortwave irradiance due to cloud scattering divided by downwelling TOA shortwave irradiance plotted against cloud fraction. 2.6a (top) represents times when SZA is above 60 degrees, while 2.6b (bottom) represents times when SZA is below 60 degrees. CRM results are shown by the solid blue line with circles, DSTOC results are shown as a solid green line with triangles, CRM albedo is shown as a dashed red line with squares, and DSTOC albedo is shown as a dashed black line with diamonds. Results are bin-averaged at 0.1 cloud amount intervals using median values.

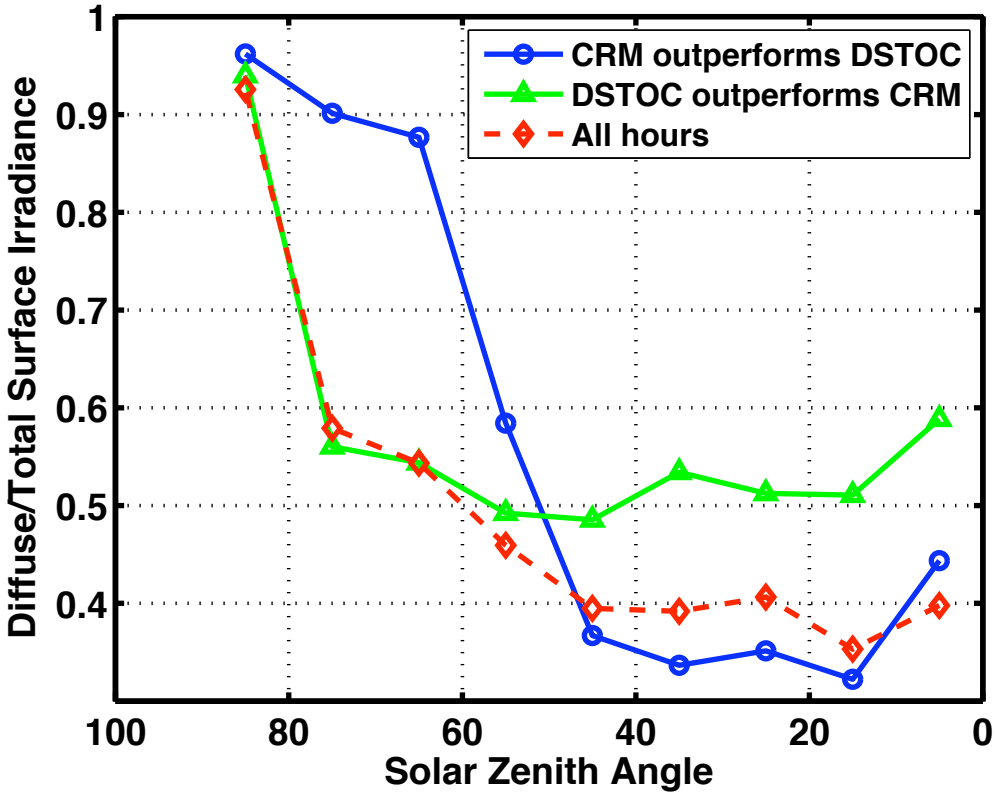


Figure 2.8 Comparison of surface diffuse divided by total solar radiation. Circles on a solid blue line indicate times when CRM generates SSI results 5 percent closer to observations than did DSTOC, while triangles on a solid green line indicate times when the opposite is true. Diamonds on a dashed red line indicate the ratio for all hours. Results are bin-averaged by SZA at intervals of 10 degrees using median values of diffuse to total surface irradiance.

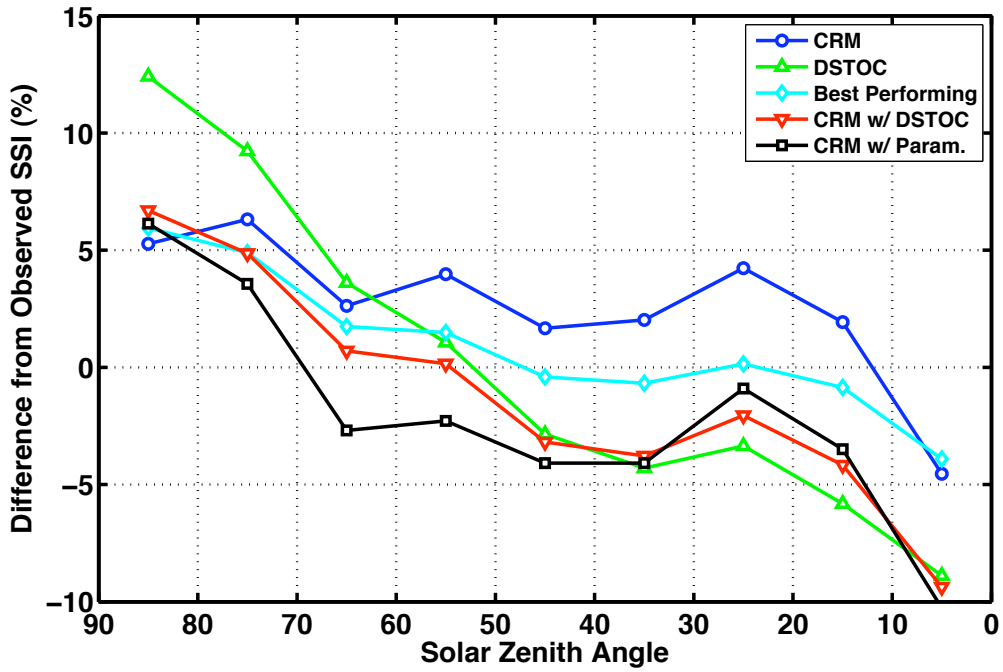


Figure 2.9 Comparison of model-generated solar surface irradiance with respect to SZA. Each model-generated solar surface irradiance is treated as follows: 100*(Model SSI – Observed SSI)/(Observed SSI) to obtain a percent difference from observations. CRM results are displayed on a blue line with circles, DSTOC results are displayed on a green line with upward pointing triangles, the cyan line with diamonds represents whichever model has the generated SSI closer to that observed for each hour, the red line with downward pointing triangles displays CRM results with DSTOC results inserted whenever the parameterization criteria are met, and the black line with squares represents CRM results adjusted with the results from the multiple linear repression whenever the parameterization criteria are met. The bin-averaging interval is 10 degrees and median values are used.

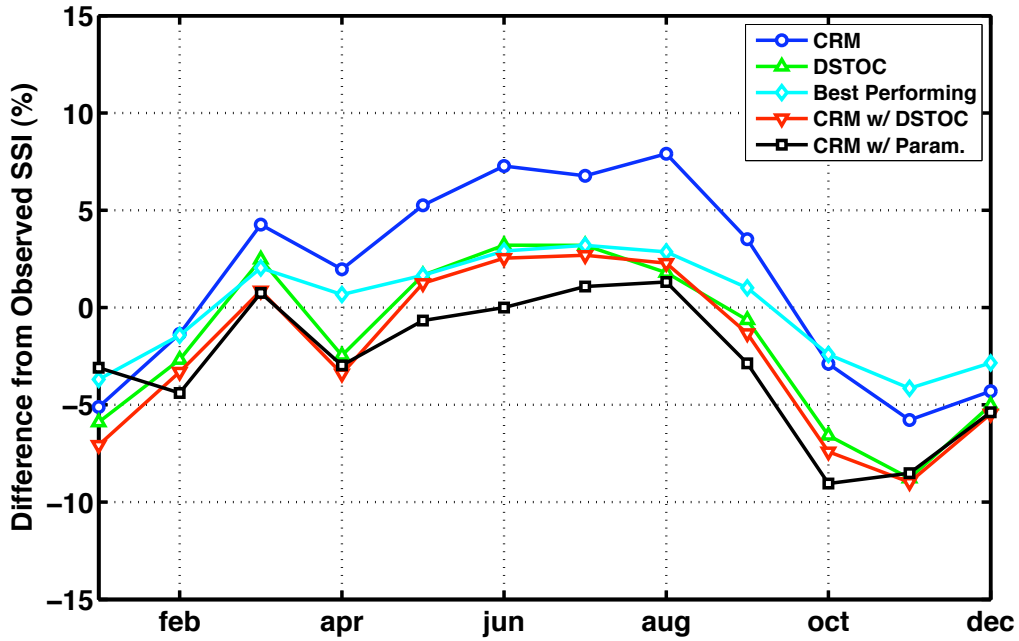


Figure 2.10 Monthly comparison of model-generated median SSI shown as a percent-difference from observations. CRM results are displayed on a blue line with circles, DSTOC results are displayed on a green line with upward pointing triangles, the cyan line with diamonds represents whichever model has the generated SSI closer to that observed for each hour, the red line with downward pointing triangles displays CRM results with DSTOC results inserted whenever the parameterization criteria are met, and the black line with squares represents CRM results adjusted with the results from the multiple linear repression whenever the parameterization criteria are met.

CHAPTER 3 – IDENTIFICATION OF MAJOR CLOUD REGIMES IN THE TROPICAL WESTERN PACIFIC

Recent studies have indicated that one way to improve global climate models is to improve their treatment of three-dimensional cloud field geometry. There are several challenges to realizing this goal, one of which is relating large-scale (GCM-scale) fields to unresolved cloud scale variability in the cloud and radiation properties (Potter and Cess, 2004; Randall et al., 2007). One way to address this issue is to identify specific cloud regimes using large-scale atmospheric properties that GCMs are able to resolve, and then to relate the dynamic and radiative properties of each regime individually using high resolution observations. To this end a clustering algorithm is applied to ground-based cloud and radiation measurements in order to identify said cloud regimes.

Following work by Jakob et al. (2005) and Gordon et al. (2005), a *k*-means clustering algorithm (Anderberg, 1973; Jakob and Tselioudis, 2003; Jakob et al., 2005) is applied to cloud and radiation data measured at the ARM CART facility on Nauru Island from 2001 to 2004. The radiative and dynamic characteristics of the resulting clusters are analyzed using additional data from the ARM archive, as well as atmospheric profiles of temperature, humidity and wind taken from ECMWF data. Four distinct cloud regimes are identified: a convectively active optically thick cirrus regime with large coverage, a convectively active optically thin cirrus regime with low coverage, a suppressed regime composed primarily of boundary layer clouds, and a convectively active optically thick regime composed of multiple cloud types. Stochastic and plane-parallel shortwave radiative transfer simulations, as described in the last chapter, are compared against

observations and evaluated for each of the clusters. As one of the goals of this work is to develop a parameterization applicable throughout the tropics, the clustering algorithm is applied to data from the ARM CART Manus and Darwin facilities. Using the cluster analysis results from all three sites as well as the model results, a set of 3D cloud-effect criteria is developed.

Section 3.1 details the data used for the clustering algorithm and additional analysis of the cloud regimes. In section 3.2 the results and the ensuing examination of radiative and dynamic properties for each cluster are described. Model performance for each cluster is evaluated in section 3.3. Section 3.4 relates the inclusion of the Manus and Darwin data in the cluster analysis and the development of criteria for application of a statistical cloud-radiation scheme. The implications and conclusions from this study are discussed in section 3.5.

3.1 METHODOLOGY

Cluster analysis

The Jakob et al. (2005) study described in the first chapter and subsequent work by Gordon et al. (2005) and Williams et al. (2005) provide the framework for the cluster analysis performed here, with a few key differences. While these studies used satellite data from the ISCCP climatology (Rossow and Schiffer, 1999) we are using surface-based measurements from a smaller spatial domain. We are also using LWP in lieu of optical depth and geometric cloud-top height instead of cloud-top pressure. Cloud coverage is used as a third variable to generate the histograms used in the cluster analysis,

similar to the method employed by Gordon et al. (2005). Averages of the data are taken at 5-minute intervals, and 36 of these intervals are used to develop 3-hour histograms. The data are grouped into ten bins each, meaning each histogram contains 1000 elements. Analysis is restricted to histograms that contain ten or more 5-minute intervals, and times with clear sky are not included. Also not included are times when rainfall is occurring, identified when the 31.4 GHz brightness temperature as measured by the microwave radiometer exceeds 100 K, or when LWP exceeds a predetermined threshold of 500 gm^{-2}. While this may have some effect on the final clusters, the radiative transfer models with which these results are being compared do not currently have a method of dealing with precipitating hydrometeors.

Since the number of clusters is chosen beforehand, an objective set of criteria must be used to determine a value for k. Rossow et al. (2005) used a set of four criteria to objectively determine a value for k: (1) the centroid histogram patterns should not change when the initial conditions are varied, (2) the centroid patterns should significantly differ from one another, (3) the spatial-temporal correlations of the cluster members should be low, and (4) the distances between cluster centroids should be greater than the distance between the cluster members and their centroid. Beginning with k=2 and successively increasing its value by one we found that k=4 provides the optimal number of clusters following these criteria. For values of k less than 4, changes in the initial centroids lead to significant changes in the mean histogram patterns thereby violating the first criterion, while for values of k greater than 4 we produce centroids patterns that do not significantly differ from one another, violating the second criterion. Even for k=4 changes in the initial centroid results in small deviations in the final mean cluster patterns. To take this

into account, the clustering algorithm was run several times using randomly chosen initial centroids and the variance around each centroid was calculated. The final cluster set chosen was that with the least sum of the variance. This procedure follows that used in Gordon et al. (2005).

Data

All of the data presented in this chapter are from the ARM Program's TWP site, composed of the Manus, Darwin and Nauru Island observational facilities. The initial cluster analysis is performed with data from the Nauru Island facility that span from the start of 2001 to the end of 2004. The variables used to develop the clusters are LWP derived from the ARM two-channel (23.8 and 31.4 GHz) microwave radiometer using the Turner et al. (2007) algorithm, cloud amount calculated with the (Long et al., 2006) algorithm using data from the Shortwave Flux Analysis VAP which is derived from pyranometer measurements, and cloud-top height taken from the ARSCL VAP, which combines measurements from the Vaisala ceilometer, micropulse lidar and millimeter wavelength cloud radar (Clothiaux et al., 2000; Kollias et al., 2005). It is important to note that many of these instruments produce nadir narrow-beam single-point measurements, so that the size and source of the spatial domain being examined is strongly dependent on the magnitude and direction of the wind. The measurements are averaged over five-minute intervals and then used to generate three-hourly relative frequency-of-occurrence histograms. Diagnostic data derived from ECMWF model runs specially generated for TWP ARM sites provide profiles of wind speed and direction

which, along with cloud boundary data taken from the ARSCL VAP, are used to calculate wind shear between cloudy layers. The ECMWF diagnostic data also provide profiles of temperature and moisture that are used to calculate Convective Available Potential Temperature (CAPE). In addition to providing cloud profiles for the model simulations, the ARSCL product also contains cloud boundaries used for calculating spacing between vertical cloudy layers, Doppler velocity that is tracked for each cluster, and vertical profiles of reflectivity measurements used to allocate cloud liquid water in each cloudy vertical layer. Details of the data sources, from which input for the radiative transfer models is derived, and descriptions of the models themselves can be found in chapter 2 but Table 3.1 provides a quick comparison of the model configurations.

Once the analysis of the Nauru Island clustering results is performed, the method is expanded to include the Manus and Darwin facilities. Though the cloud-top height and total cloud coverage measurements are still taken from the ARSCL and Shortwave Flux Analysis VAPs, respectively, the LWP measurement is taken from a different source. The LWP measurements in the initial cluster analysis are processed using the Turner et al. (2007) algorithm to derive LWP from microwave radiometer measurements. The algorithm applies both a statistical and physical-iterative retrieval to derive LWP and precipitable water vapor from MWR measurements. Brightness temperature corrections are applied before the retrievals and the "monoRTM" radiative transfer model (Boukabara et al., 2001) is used to determine LWP. This algorithm-processed data is not available during the time periods in which we are interested at Manus and Darwin, and so LWP is derived directly from the line-of-site (LOS) microwave radiometer measurements. To account for this, the cluster analysis is run a second time at the Nauru

Island site using the LOS measurements and a comparison is done between the two results.

3.2 CLUSTER ANALYSIS RESULTS

The *k*-means clustering algorithm is applied to the histograms described above using *k*=4 for the number of clusters. Table 3.2 contains information about the cluster characteristics. Figure 3.1 shows the first cluster, which represents a regime composed primarily of high cirrus with median LWP of 22.7 gm^{-2}, high total cloud coverage and the presence of some deep convective activity (indicated by a large median Doppler velocity and cloud geometric thickness). The large cloud coverage suggests this regime may also include outflow from neighboring convection. The second regime, shown in Figure 3.2, is also dominated by high cirrus, but has lower median LWP (12.4 gm^{-2}) with lower total cloud coverage and less coincident convective activity. The third regime represented in Figure 3.3 is dominated by low boundary layer clouds with a median LWP of 22 gm^{-2} and small total cloud coverage. This is a stable regime with little to no convective activity. Figure 3.4 shows the fourth and largest regime, which often contains multiple coincident cloud types ranging from boundary layer to cirrus. Because of the presence of multiple cloudy layers, this regime has the most variance in cloud coverage, cloud-top height and cloud geometric thickness, but also has the highest median value of LWP with 37.1 gm^{-2}. This cluster also has a positive (upward) median Doppler velocity indicating it may be a convectively active regime, and seems to be the only cluster that contains

significant amounts of mid-level cloud. The relative frequencies of occurrence of the regimes are 15%, 13%, 18% and 54%, respectively.

Spatial domain

Before additional analysis of these clusters takes place a look at the magnitude and direction of the wind at the Nauru facility is warranted, as this will determine the size and location of the spatial domain being observed by the nadir-pointing surface instruments. The mean magnitudes of wind speed throughout the portion of the vertical column being observed, that being between a couple of hundred meters and twenty kilometers above the surface, are 5.0, 4.9, 5.4 and 4.9 m/s for clusters 1, 2, 3 and 4, respectively. Figure 3.5 shows the relative frequency-of-occurrence of wind direction; where for all hours it is dominated by the easterly trade winds. This indicates that the size and source of the spatial domain is pretty consistent for all four identified cloud regimes, and that the spatial domain from which hourly data is being pulled for the DSTOC and CRM models is close to twenty kilometers in size and approaching from the east of Nauru.

Convective Available Potential Energy

One of the goals of this study is to determine the dynamic properties of the cloud regimes represented by the four identified clusters. Vertical velocity may serve as one indicator of how much convective activity is present in the clusters, while another metric that may be used for this purpose is Convective Available Potential Energy (CAPE).

CAPE is essentially an indicator of how much buoyant energy is theoretically available for a parcel of air to vertically rise and accelerate. Outside factors such as entrainment of dry air and horizontal pressure gradients can inhibit parcels from realizing their maximum potential height. In terms of a skew-T diagram, positive CAPE is the area between the parcel temperature and the environmental temperature where the parcel temperature is warmer than that of the environment. The lower boundary limit of CAPE generally occurs at the level of free convection, though inversions near the surface may also be included, and the upper boundary limit is the equilibrium or anvil level. There are also different types of CAPE; for example in this study we calculate CAPE both for the entire atmospheric column and for only the area below 5 km where clouds are composed primarily of water. Figure 3.6 shows a scatterplot of CAPE against cloud-top height. The CAPE used in Figure 3.6 is that calculated below 5 km and has been bin-averaged by intervals of 5 J/kg. After restricting the dataset to daytime hours, as we are focused on shortwave fluxes, the data shown here represents nearly 13,000 hours spanning the four year period. Though there is no statistically significant correlation between the raw hourly CAPE and cloud-top height measurements, once these data have been bin-averaged the correlation coefficient between the two becomes 0.62. The gray least squares fit bar in Figure 3.6 shows this linear relationship between cloud-top height and the magnitude of CAPE, though the error bars indicates that the signal is within the noise of the data. When this comparison is made using CAPE calculated throughout the atmospheric column this signal becomes weaker (not shown). These results are similar to those of Jenson and Del Genio (2006), though their study was limited to primarily cumulus congestus clouds and covered a shorter time span.

In order to further examine the relationship between CAPE and cloud-top height, we separate CAPE by cloud regime. Using the same set of low-level CAPE and cloud-top height data, Figure 3.7 shows relative frequency-of-occurrence histograms based on cluster membership. Membership is determined by matching the hourly cloud-top height and CAPE measurements with the coincident histograms used to generate the initial four cloud regime clusters. Clusters 1 and 2 display similar patterns of high cloud top with CAPE median values of 32 and 38 J/kg (See Table 3.3), respectively, and a relatively stable distribution across the range of values. Interestingly, though cluster 3 seems primarily composed of stable low-level boundary layer clouds, it contains a relatively high concentration of CAPE with a median value of 52 J/kg. Cluster 4 displays a wide array of cloud-top heights with a CAPE median value of 33 J/kg. To put these numbers into better context we calculate convective inhibition (CINH) which is essentially the opposite of CAPE, or the area in a skew-T diagram where parcel temperature is cooler than that of the environment. CINH may be thought of as the energy that must be overcome before an air parcel may develop into deep convection. The median CINH values below 5 km are 290, 289, 250 and 279 J/kg for clusters 1, 2, 3 and 4 respectively. This further illustrates the complex thermodynamic processes that are occurring during the third cloud regime, as these clouds coincide with the highest values of CAPE and the lowest values of CINH. Possible explanations for limiting the cloud-top heights of cluster 3, at least for the case of cumulus congestus clouds, include entrainment of dry air and the presence of weak stable layers near the freezing level (Redelsperger et al., 2002).

From this analysis it appears that there is no easily discernible linear relationship between CAPE and cloud-top height, though given the correlation of the bin-averaged

data and the differing values of CAPE for each cluster this topic may warrant further study. In particular, it is unclear why cluster 3, which appears to be composed primarily of stable boundary layer clouds and has a negative median value for Doppler velocity, would have the largest median value of CAPE and smallest median value for CINH.

Wind shear and vertical spacing

Wind shear and vertical spacing between cloudy layers are potentially significant factors in determining cloud macroscale effects on shortwave flux calculations as they are both related to cloud overlap (Naud et al., 2008). Vertical spacing between cloudy layers, for this study, is defined as the geometric distance between the top height of a cloud detected by the ARSCL VAP and the bottom height of a cloud located above it. Wind shear is calculated between these same heights. Table 3.3 contains wind shear and vertical cloud spacing information for each cluster, and Figure 3.5 shows relative frequency-of-occurrence histograms for wind shear and cloud spacing based on cluster membership. The histograms for clusters 1 and 2 display similar patterns, with these clusters containing relatively high wind shear and small spacing between cloudy layers when compared to those of clusters 3 and 4. Cluster 1 has wind shear greater than 0.003 s^{-1}, which may be considered of at least moderate strength, 32 percent of the time while cluster 2 has moderate or higher wind shear strength 38 percent of the time, compared to 14 percent for cluster 3 and 23 percent for cluster 4. The small values for spacing between cloudy layers could indicate that a maximum overlap assumption would be appropriate for these clusters, but Naud et al. (2008) found that large values for wind

shear often prescribe a minimum overlap assumption. Cluster 3 has the lowest values for wind shear and the highest concentration of both very low and very high (8 km and greater) spacing between cloudy layers. These numbers should be taken in context, however, as 36 percent of the time there is only a single cloud layer present in cluster 3, where cloud spacing is zero and not included in these histograms. That compares to only 3, 5 and 9 percent of the time for clusters 1, 2 and 4 respectively. The times with very high cloud spacing are due to the coincident presence of low boundary layer clouds and high cirrus, and account for only 9 percent of cluster 3. Cluster 4 contains a wide array of different cloud spacing and wind shear, though the median wind shear value is relatively low (1.6E-03 s^{-1}). The large size and wide range of values in cluster 4 make it difficult to characterize cloud field morphology or overlap type.

From this analysis it seems clusters 1 and 2, which represent cloud regimes with high cloud-top heights, may be characterized as having relatively large wind shear and small spacing between noncontiguous cloudy layers. Cluster 3 has small median values for wind shear and the largest values for vertical spacing, but this is a bit misleading as cluster 3 often has only single or contiguous cloudy layers, which are not represented in these calculations. Cluster 4 contains the mid-range values of wind shear and vertical spacing, likely due in part to its large relative frequency-of-occurrence and corresponding variety of coincident cloud types.

3.3 MODEL PERFORMANCE

The DSTOC and CRM models described in chapter 2 are run, using the same data sources as the clustering algorithm: the ARSCL VAP, the MWRRET VAP, and the Shortwave Flux Analysis VAP, as forcing data from the beginning of 2001 to the end of 2004. Limiting the runs by the availability of instrument measurements and sunlit hours reduces the number of hours run from the 35064 possible to just under 8000, or a little less than half the daytime hours. A thorough description of these simulations can be found in chapter 2. For the purposes of this study we are concerned with the models' ability to simulate downwelling solar surface irradiance as compared to measurements taken from the Shortwave Flux Analysis VAP. Figure 3.9 displays the performance of the models when compared to one another and against observations, separated by cluster membership. In this context the term "outperform" is defined the same way is in chapter 2; as one model generating SSI that is at least 5 percent closer to that observed than the other model. For example, if the observed SSI is 100 Wm^{-2} then a model must generate SSI at least 5 Wm^{-2} closer to 100 Wm^{-2} than the other model in order to be considered outperforming it. When both models are within this 5 percent range they are considered to be performing equally well. It should be noted that the choice of 5 percent is somewhat arbitrary, and as such this analysis was also performed using 2, 3 and 4 percent as the criteria for one model outperforming the other (not shown). The time when the models performed equally well decreased with a narrower definition of "outperformance", but otherwise the ratio of one model outperforming the other remained close for all cases. The distribution of one model outperforming the other varies

considerably amongst the clusters, with cluster 3 showing similar performances by both models and cluster 2 showing a considerable difference between DSTOC and CRM performance, with DSTOC outperforming the CRM 46 percent of the time compared to the CRM "outperforming" DSTOC 21 percent of the time. Cluster 1 shows a smaller difference of this kind, with DSTOC outperforming the CRM 41 percent of the time compared to 27 percent of the time when the opposite occurs. Cluster 4 shows a difference between the models of only about 5 percent, with the models performing equally most often by a small margin.

Table 3.4 contains information on how the models perform in terms of mean, median and standard deviations of differences between model-generated and observed SSI normalized by the incoming solar radiation at the top of the atmosphere. This can be thought of as the percent difference between modeled and observed atmospheric transmissivity for solar radiation. The data are grouped by cluster and for times when each model "outperforms" the other. As can be expected, during times when one model "outperforms" the other the difference from observations is smaller. For all hours, DSTOC has smaller mean and median differences than the CRM for clusters 2, 3 and 4, while the CRM has smaller mean and median differences than DSTOC for cluster 1. The largest discrepancy between model performances for all hours occurs for cluster 2, where DSTOC atmospheric transmissivity is 3.8 and 2.6 percent (mean and median) closer to observations than that simulated by the CRM. Perhaps the data that stand out the most in Table 3.4 are the standard deviations, which are all large compared to the difference between model performances. This is because there is a large amount of variability in the hour-by-hour model-simulated SSI as compared to observations. There could be many

reasons for this variability, but likely suspects include the use of single point nadir-pointing measurements for many of the observations, the use of climatological mean values for temperature, carbon dioxide, ozone and aerosol profiles, and the assumption of homogenous in-cloud microphysical properties. Regardless, this variability makes it difficult to determine if one model's improved performance as compared to the other's is actually significant, and impossible to pronounce it statistically so.

In an effort to address this issue and provide an additional tool for evaluating model performance, the ability of the models to capture the observed variability in solar radiative cloud forcing at the surface is compared in Table 3.5 for each cluster. A correlation coefficient is calculated for model-generated and observed solar radiative cloud forcing at the surface, which is defined as the clear-sky fit estimated downwelling surface solar irradiance minus the model-generated or observed SSI. Similar to the results in Table 3.4, when one model "outperforms" the other the correlation coefficient is higher, and in some cases significantly so. For example, during hours when DSTOC "outperforms" the CRM for the second cluster, the CRM cloud forcing is poorly correlated with that observed, only 0.18, while the DSTOC correlation coefficient is 0.48, which is still not very highly correlated. One point to take from Table 3.5 is that DSTOC has higher correlation coefficients for all hours and clusters than the CRM. This is not unexpected, as one of the primary purposes of the stochastic approach is to be able to capture some of the large-scale geometric variability in the cloud fields not accounted for by plane-parallel calculations, but it does serve as some validation. Another point to take from Table 3.5 is that there is quite a bit of variability in the models' performance among

clusters, ranging from correlation coefficients for all hours of 0.44 and 0.33 for cluster 2 to 0.69 and 0.65 for cluster 4 for DSTOC and the CRM, respectively.

3.4 EXPANSION OF CLUSTER ANALYSIS SPATIAL DOMAIN

For these cloud regimes to be useful they must occur in more locations than just a single small island in the Tropical Western Pacific. For this reason it is helpful to expand the spatial domain of the cluster analysis to include the Darwin and Manus facilities. There are a couple of caveats that should be mentioned regarding these sites. First, the time period examined at Manus coincides with that of Nauru, from the beginning of 2001 to the end of 2004, however the Darwin site is newer than its counterparts and does not have coincident measurements of LWP, cloud coverage and cloud top during this period. Therefore the cluster analysis at Darwin is performed for the period from the beginning of 2006 through July of 2007. This may have some effect on the resulting clusters at Darwin, as there is bound to be inter-annual variability due to events such as the El Nino Southern Oscillation. Also, as mentioned earlier the MWRRET VAP used at the Nauru site for LWP was not available at Darwin and Manus, so LWP was derived from the line-of-site microwave radiometer (MWRLOS) measurements. To account for this the analysis was rerun at Nauru using the MWRLOS measurements.

Table 3.6 contains information about the mean centroids generated at Nauru using the MWRLOS measurements of LWP. Clusters 1, 2, 3 and 4 correspond with those generated in Table 3.2. The difference in median LWP is considerable, with the MWRLOS clusters all sharing similar median values of LWP near 50 gm^{-2} and the

MWRRET clusters having median values of 22.7, 12.4, 22.0, and 37.1 gm^{-2} for the four clusters, respectively. One reason for this difference may be that the MWRRET VAP uses infrared measurements at 11 microns as well as the microwave measurements to determine LWP. The advantage to this approach is that when LWP is relatively small, say under 60 gm^{-2}, the infrared measurements are much more sensitive than the microwave measurements and are able to reduce uncertainty in LWP (Turner, 2007). Since many of the LWP measurements at the TWP facilities fall into this under 60 gm^{-2} category, it is not surprising that there are marked differences between the MWRLOS and MWRRET LWP measurements. Regardless, the basic structure of the cloud regimes remains similar for both sets of clusters. Due to the lack of sensitivity in LWP, clusters 1 and 3, which in the first set of clusters had LWP median values of 22 gm^{-2}, lose some members to cluster 4. The result is that cluster 4's relative frequency-of-occurrence increases by 0.1 to 0.64 at the expense of clusters 1 and 3. Since clusters 1 and 3 are composed primarily of very different cloud types, namely boundary layer low coverage clouds and nearly overcast cirrus clouds, it leads to the question of whether cluster 4 can be broken up into smaller clusters. Ultimately, after running the *k*-means algorithm with *k* values up to 12 it was found that cluster 4 consistently existed and contained a large number of member histograms. For $k = 12$ this cluster still has a relative frequency-of-occurrence of 0.44 while the next largest cluster has a relative frequency-of-occurrence of 0.11. This mixed-cluster is primarily made up of the multiple cloud layers, oftentimes high cirrus and low boundary layer clouds. This generates a rather unique patterned histogram, as can be seen in Figure 3.4, which may explain why the clustering algorithm does not break it apart even with large values for *k*.

The same set of criteria for selecting a value for k is used at the Manus and Darwin facilities, as well as with the clustering of all three sites combined, as was used for the Nauru facility. However, even with the use of a set of criteria for which to select k, there is still a subjective component to determining how many clusters are appropriate for each observed area. The goal, at least for this study, is to find the minimum number of clusters with distinct cloud properties and dynamical conditions. For example, when increasing k from 4 to 5 at the Manus site the third cluster, composed primarily of low clouds, split into two clusters with median cloud fractions of 0.2 and 0.4 that otherwise shared similar properties. It was determined that this distinction did not produce a significant amount of additional insight and therefore the number of clusters produced at Manus remains at four. Similarly at the Darwin site when k was increased from four to five the first cluster, composed primarily of high cirrus, split into two clusters with one of the clusters composed primarily of nearly overcast optically thick cirrus and the other composed primarily of partly cloudy cirrus with mid-range optical thicknesses. In this case it was determined that the two new clusters represented regimes with significantly different optical and dynamical properties and should remain separate.

Manus

Table 3.7 contains information about the four clusters generated at the Manus facility. Three of the four clusters share similar properties with those generated at Nauru, and the rows have been color-coded in each table to reflect this similarity. Specifically, cluster 2 at Manus is composed primarily of high cirrus clouds with medium coverage

(approximately between 0.4 and 0.7) and a positive median Doppler velocity indicating the possibility of convective activity, which corresponds with cluster 2 at Nauru. Cluster 3 at Manus is composed primarily of low boundary layer clouds with relatively low cloud coverage (between 0.2 and 0.4), very low geometric thickness, and very small or negative (downward) Doppler velocities indicating more stable conditions, which corresponds with the 3rd cluster at Nauru. Finally, cluster 4 at Manus is the often-occurring mixed cluster discussed earlier which contains many cloud types that exist simultaneously, corresponding with cluster 4 at Nauru. Cluster 1 at Manus is somewhat similar to cluster 1 at Nauru in that both clusters contain primarily clouds with nearly overcast to completely overcast coverage that are optically thick. The major difference between the clusters are that the median value of cloud-top height for Manus is around 5 kilometers while the median cloud-top height at Nauru is close to 13 kilometers. Also the median geometric thickness of the Manus cluster is much smaller than that of Nauru, 0.41 kilometers compared to 1.82 kilometers

Darwin

Table 3.8 contains information about the five clusters generated at the Darwin facility. Four of the clusters generated at Darwin share similar characteristics with the Nauru clusters. As mentioned in the description of the Manus clusters, clusters 2, 3 and 4 at Darwin correspond with clusters 2, 3 and 4 at Manus and Nauru respectively. The 1st cluster at Darwin is composed primarily of high cirrus clouds with nearly overcast to overcast coverage, large median Doppler velocity, and high LWPs, corresponding to

cluster 1 at Nauru. The 5th cluster at Darwin has a small relative frequency-of-occurrence, 0.05, very small cloud amount with a median of 0.08, small Doppler velocity, geometric thickness, and LWP, and a median top height of 5.26 kilometers. It does not correspond with any cluster from either Nauru or Manus.

All ARM TWP facilities

Table 3.9 contains information about the five clusters generated using the histograms from all three ARM TWP facilities: Manus, Darwin and Nauru. When the clustering algorithm is run with k=4, the clusters generated match those at the Nauru site. With higher values of k smaller clusters appear that are generally just small variations on these four primary cloud regimes. For example, depending on the selection of the initial centroids, several iterations split the low level boundary layer regime into two clusters containing median values of cloud fraction of 0.2 and 0.4, instead of a single cluster with a median value of 0.3. This is not totally unlike the final selection of k=5 where it can be seen by the color coding in Table 3.9 that clusters 1, 3 and 4 all correspond with those found at the other sites, while clusters 2a and 2b have both been labeled as being part of the same cirrus medium coverage convective regime. The reason these two clusters remain separate is that not only are there differences in the median cloud fraction, but one cluster is composed primarily of geometrically thick cirrus clouds with low median LWP (29.9 gm^{-2}) while the other is composed primarily of cirrus that are geometrically thinner but have higher median LWP (53.1 gm^{-2}).

The results of this analysis indicate that regional variations exist in the composition and frequency of specific cloud regimes, but that these differences tend to be small. Certain cloud regimes, such as convectively active high coverage cirrus, low-level boundary layer stratus, and medium coverage cirrus exist at several locations with large enough frequency to suggest that improving the radiative treatment of cloud-field geometry for these regimes could prove beneficial for many parts of the tropics. The multi-layer regime, which exists with very large frequency at all sites, is more difficult to characterize as it has the largest standard deviations for its cloud properties.

3.5 CONCLUSIONS

The application of the *k*-means clustering algorithm to the surface-based measurements of atmospheric state taken from Nauru Island has produced four cloud regimes with distinct characteristics. As the primary goal of this chapter is to identify criteria that could be used as indicators of complex cloud field geometry, it has been helpful to look at the dynamical and thermodynamical properties of each of these cloud regimes individually. Three of the four regimes show signs of being convectively active, while the other regime (cluster 3) appears to be composed primarily of stable boundary layer clouds. Though the convectively active regimes have on average higher cloud tops then the stable regime, we find only a weak linear relationship between cloud-top height and the magnitude of CAPE, both for the entire set of data and for each cluster individually. In fact, the cloud regime found to be most stable also had the highest median values of CAPE and the lowest median values of CINH. Examination of wind

shear and vertical spacing between cloudy layers suggest that the convectively active regimes tend toward large values of wind shear and smaller spacing between cloudy layers than does the stable regime.

Another focus of this study is to compare the performance of a stochastic radiative transfer model against that of a plane-parallel RT model for a variety of cloud field configurations, specifically those affiliated with the four cloud regimes identified by the cluster analysis. This task is complicated by the large amount of variability in the hour-by-hour comparisons of modeled and observed surface radiative fluxes. Large standard deviations for mean and median differences in model-generated versus observed SSI make it difficult to determine whether one model "outperforming" the other for specific clusters is statistically significant. However, it is found that the stochastic model is able to capture more of the variability of the observed solar radiative cloud forcing at the surface than the CRM for all four clusters, and that this difference is particularly noticeable for clusters 1 and 2. These clusters represent regimes that often contain cirrus clouds with large ranges of cloud coverage and LWP. Clusters 1 and 2 also have the highest median values for wind shear between cloudy layers and lowest median values for spacing between cloudy layers. The combination of cloud-top height, wind shear, and spacing between cloudy layers directly affect the depth of the cloud field and the vertical and horizontal spacing between clouds, making them good indicators of significant macroscale inhomogeneity in the cloud fields. From chapter 1 it is also known that cloud coverage is a key factor in determining the utility of a stochastic approach to radiative transfer, as is the ratio of LWP to cloud geometric thickness.

Other variables may require additional interpretation to be most useful. For example, cluster 1 has high cloud tops, the strongest vertical velocity and the smallest spacing between cloudy layers of all the clusters, yet the stochastic model outperforms the CRM more frequently under cluster 2 conditions. The most likely explanation for this is the large total cloud coverage and relatively high optical depth of the first cluster, since optically thick overcast or nearly overcast skies are likely to minimize the importance of radiative effects due to interactions between clouds. Cluster 4 is also difficult to characterize, as its high occurrence and variety of different cloud types generate a large range of values for all cloud properties, making it easy to mask the strengths of the stochastic model.

The cluster analysis is expanded to include the ARM Manus and Darwin facilities. It is found that although there are small changes in the relative frequency-of-occurrence and the number and structure of the clusters, the primary cloud regimes found at Nauru may also be found at other locations. This suggests it should be possible to apply a parameterization developed at Nauru to other areas in the tropics.

	CRM	DSTOC
# vertical layers	32	32
q profile	NCEP-derived	NCEP-derived
O3,T,CO2 profiles	McClatchey et al. (1972)	McClatchey et al. (1972)
RT solver	Delta-Eddington method (Briegleb 1992)	Discrete ordinate with approximate iterative technique (Wiscombe and Evans, 1977)
SW spectrum	18 unequally spaced bands	38 unequally spaced bands
Cloud ice/water partitioning	Single-moment T-dependent	Single-moment T-dependent
Cloud overlap	Random	Random
Droplet effective radius	Usually 6 microns	Usually 6 microns
Ice crystal effective radius	23 microns	23 microns
Aerosol	0.07 visible extinction optical depth	None

Table 3.1 Description of the CRM and DSTOC model configurations. Droplet effective radius is determined by a temperature dependent function with a minimum value of 6 microns, but in reality almost always uses the minimum value.

cluster	rfo	cloud amount		doppler (m/s)		diffuse/total		topheight (km)		thickness (km)		lwp (g/m^2)	
		median	*std*	*median*	*std*	*median*	*std*	*median*	*std*	*median*	*std*	*median*	*std*
1	0.15	0.74	0.15	0.37	0.35	0.73	0.18	12.74	2.23	1.76	1.71	22.70	28.87
2	0.13	0.39	0.15	0.16	0.38	0.42	0.27	13.52	2.79	0.91	1.94	12.40	21.38
3	0.18	0.30	0.11	-0.03	0.22	0.28	0.16	1.43	1.82	0.08	0.23	22.02	22.73
4	0.54	0.50	0.20	0.29	0.64	0.56	0.27	7.10	3.45	0.47	2.25	37.08	60.71

Table 3.2 Median and one standard deviation of atmospheric variables for the four clusters generated by the k-means clustering algorithm at Nauru.

cluster	CAPE (J/kg)		CINH (J/kg)		cloud spacing (m)		wind shear (s^{-1})	
	median	*std*	*median*	*std*	*median*	*std*	*median*	*std*
1	32.2	38.9	289.7	122.6	1696	1151	2.10E-03	2.10E-03
2	37.9	44.4	288.6	120.4	1542	1384	2.30E-03	2.30E-03
3	52.1	42.9	249.9	111.0	2893	3958	1.40E-03	1.40E-03
4	33.1	33.8	279.2	115.2	2247	2179	1.60E-03	2.00E-03

Table 3.3 Median and one standard deviation values of CAPE, CINH, spacing between cloudy layers and wind shear between cloudy layers for the four clusters generated at Nauru.

Difference from Observed SSI as a % of incoming SW		Cluster 1			Cluster 2			Cluster 3			Cluster 4		
		mean	*median*	*std*	*mean*	*median*	*std*	*mean*	*median*	*std*	*mean*	*median*	*std*
DSTOC outperforms CRM	**DSTOC**	0.2	-1.0	12.8	-7.1	-9.0	12.1	5.3	3.5	9.1	5.3	4.1	13.8
	CRM	-1.8	-8.3	19.2	-14.1	-17.3	16.2	11.5	10.4	10.9	8.2	10.9	18.8
CRM outperforms DSTOC	**DSTOC**	6.9	10.0	14.6	7.1	8.7	16.8	-9.6	-9.9	8.7	-1.3	-7.2	17.3
	CRM	3.6	3.2	10.4	4.4	2.5	12.4	-2.8	-2.5	6.7	1.0	-0.9	11.8
All hours	**DSTOC**	2.4	2.8	13.8	-3.3	-5.9	14.1	-1.7	-2.5	10.9	1.1	-0.3	14.8
	CRM	0.6	1.2	15.7	-7.1	-8.5	15.8	3.7	3.0	10.9	3.3	2.5	15.2

Table 3.4 Comparison of DSTOC and CRM performance for each of the clusters developed at the ARM Nauru Island facility. Data is calculated as follows: Model-generated minus observed SSI normalized by the incoming solar radiation at the top of the atmosphere. The *std* column under each cluster represents one standard deviation.

Correlation between model and observed solar cloud forcing					
Cluster #		**1**	**2**	**3**	**4**
DSTOC outperforms CRM	**DSTOC**	0.74	0.48	0.73	0.78
	CRM	0.35	0.18	0.55	0.54
CRM outperforms DSTOC	**DSTOC**	0.47	0.38	0.70	0.53
	CRM	0.73	0.72	0.80	0.79
All hours	**DSTOC**	0.61	0.44	0.57	0.69
	CRM	0.49	0.33	0.50	0.65

Table 3.5 Comparison of DSTOC and CRM ability to capture the observed changes in solar radiative cloud forcing for each of the clusters developed at the ARM Nauru Island facility. A linear correlation coefficient is calculated based on clear sky fit estimated downwelling surface irradiance minus the model-generated and observed SSI.

Nauru

cluster	rfo	cloud amount		doppler (m/s)		diffuse/total		topheight (km)		thickness (km)		lwp (g/m^2)	
		median	*std*	*median*	*std*	*median*	*std*	*median*	*std*	*median*	*std*	*median*	*std*
1	0.08	0.76	0.16	0.37	0.35	0.76	0.17	12.86	2.39	1.82	1.70	55.6	29.4
2	0.13	0.46	0.16	0.18	0.34	0.49	0.25	12.21	2.99	0.87	1.60	49.3	23.0
3	0.15	0.30	0.11	-0.03	0.27	0.28	0.14	1.45	1.75	0.08	0.23	49.6	20.8
4	0.64	0.48	0.21	0.25	0.57	0.55	0.28	7.61	4.14	0.50	2.06	50.9	57.6

	nearly overcast cirrus regime
	cirrus medium coverage regime
	boundary layer low coverage stable regime
	multiple cloud-type mixed regime

Table 3.6 Median and one standard deviation of atmospheric variables for the four clusters generated by the k-means clustering algorithm at the ARM TWP Nauru facility from the beginning of 2001 to the end of 2004. The LWP used for this analysis is derived from MWRLOS measurements from a microwave radiometer as opposed to the MWRRET VAP used in the previous clusters. The colors in each row correspond with specific primary cloud regimes: blue represents a nearly overcast cirrus regime, pink represents a cirrus medium coverage convective regime, green represents a boundary layer low coverage stable regime, and purple represents a multiple cloud-type mixed regime.

Manus

cluster	rfo	cloud amount		doppler (m/s)		diffuse/total		topheight (km)		thickness (km)		lwp (g/m^2)	
		median	*std*	*median*	*std*	*median*	*std*	*median*	*std*	*median*	*std*	*median*	*std*
1	0.10	0.75	0.16	0.20	0.83	0.75	0.18	5.20	1.80	0.41	1.05	64.2	42.8
2	0.19	0.54	0.18	0.28	0.51	0.60	0.23	12.56	2.39	2.17	2.20	54.2	30.9
3	0.12	0.31	0.14	-0.29	0.72	0.30	0.16	1.34	1.58	0.09	0.42	48.3	22.3
4	0.60	0.47	0.22	0.30	0.84	0.63	0.27	6.69	4.16	0.76	2.45	56.9	62.3

Table 3.7 Same as Table 3.6 but for the Manus Facility.

Darwin

cluster	rfo	cloud amount		doppler (m/s)		diffuse/total		topheight (km)		thickness (km)		lwp (g/m^2)	
		median	*std*	*median*	*std*	*median*	*std*	*median*	*std*	*median*	*std*	*median*	*std*
1	0.07	0.81	0.16	0.32	0.25	0.79	0.14	13.05	1.93	2.27	1.89	75.7	33.3
2	0.05	0.57	0.15	0.21	0.27	0.60	0.21	12.11	2.52	1.49	1.60	58.3	29.8
3	0.11	0.33	0.14	0.06	0.15	0.30	0.11	1.31	1.61	0.16	0.30	54.0	24.2
4	0.72	0.45	0.24	0.18	0.40	0.56	0.30	9.08	4.28	0.86	2.46	67.2	64.3
5	0.05	0.08	0.11	0.05	0.14	0.17	0.08	5.26	5.01	0.13	0.86	16.3	8.3

Table 3.8 Same as Table 3.6 but for the Darwin Facility.

TWP

cluster	rfo	cloud amount		doppler (m/s)		diffuse/total		topheight (km)		thickness (km)		lwp (g/m^2)	
		median	*std*	*median*	*std*	*median*	*std*	*median*	*std*	*median*	*std*	*median*	*std*
1	0.08	0.74	0.16	0.34	0.37	0.73	0.17	12.72	2.25	2.10	1.83	60.2	30.3
2a	0.11	0.43	0.15	0.17	0.49	0.46	0.26	12.08	2.94	1.04	2.15	53.1	28.6
2b	0.09	0.59	0.18	0.30	0.54	0.58	0.24	12.32	2.61	1.63	2.02	29.9	22.6
3	0.12	0.30	0.12	-0.04	0.49	0.28	0.14	1.37	1.68	0.10	0.30	49.1	20.9
4	0.61	0.48	0.23	0.21	0.66	0.59	0.28	6.56	4.10	0.53	2.24	60.0	61.3

Table 3.9 Same as Table 3.6 but for all three ARM TWP facilities.

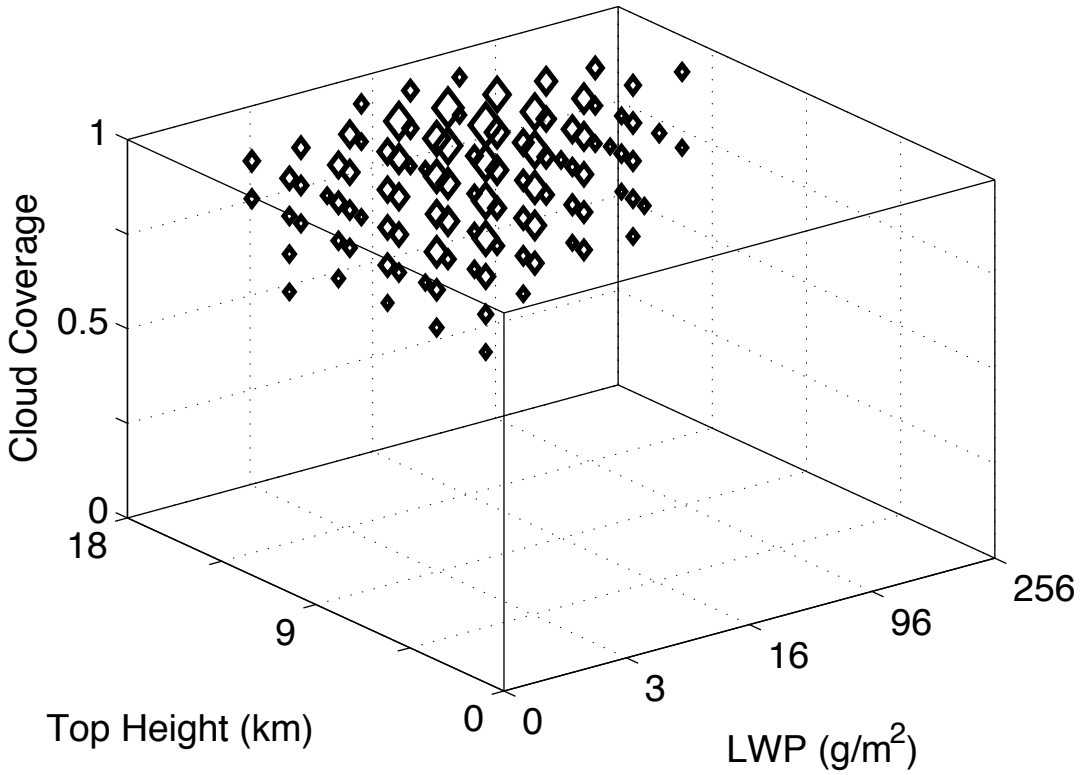


Figure 3.1 First three-dimensional centroid resulting from the k-means clustering of LWP, cloud-top height and total cloud coverage as measured by surface instrumentation on Nauru Island from the beginning of 2001 to the end of 2004. The larger the diamond, the greater the relative frequency-of-occurrence.

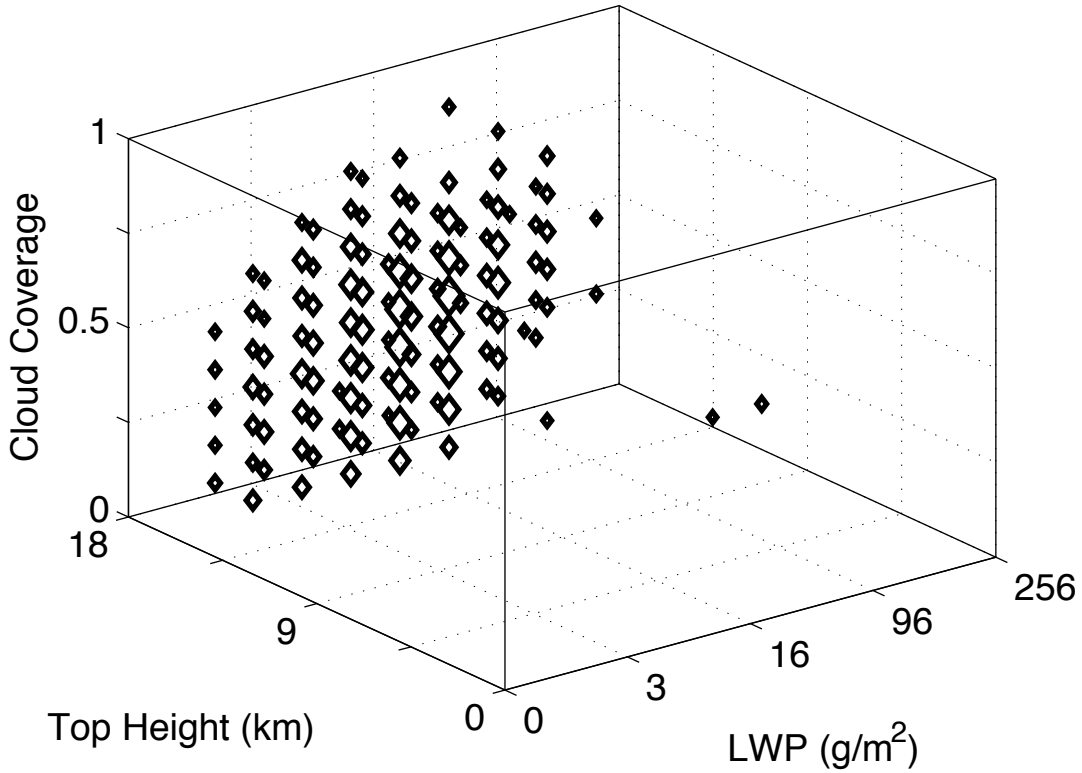


Figure 3.2 Same as Figure 3.1 but for the second three-dimensional centroid.

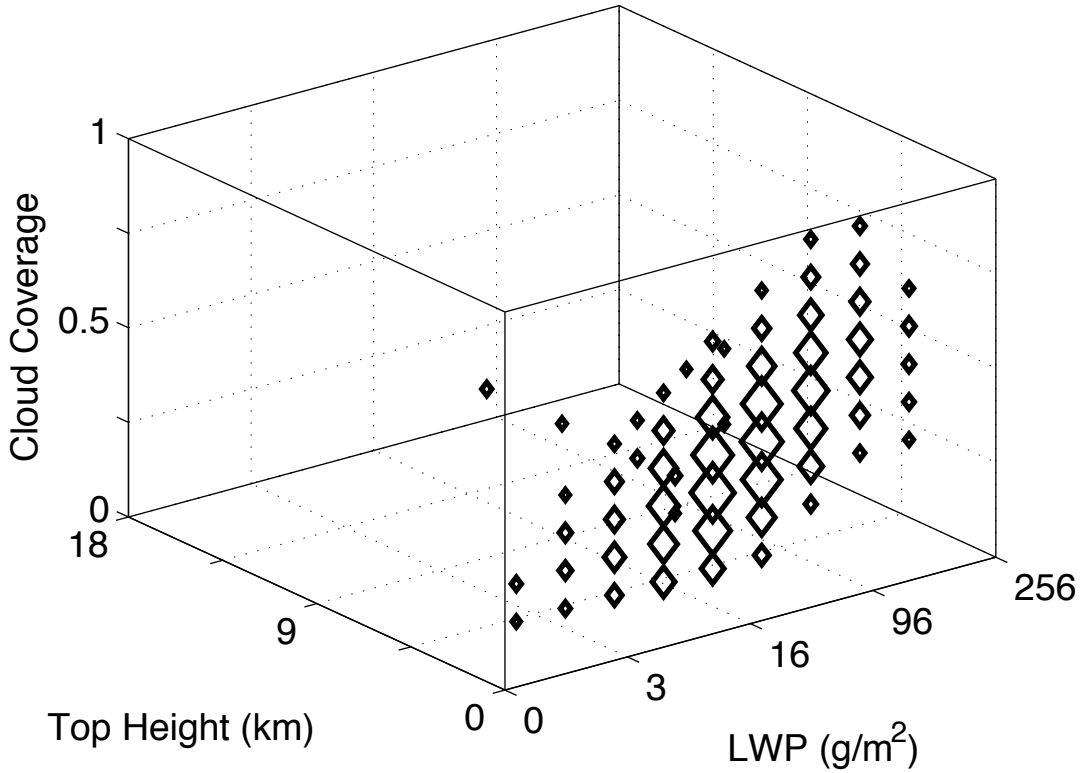


Figure 3.3 Same as Figure 3.1 but for the third three-dimensional centroid.

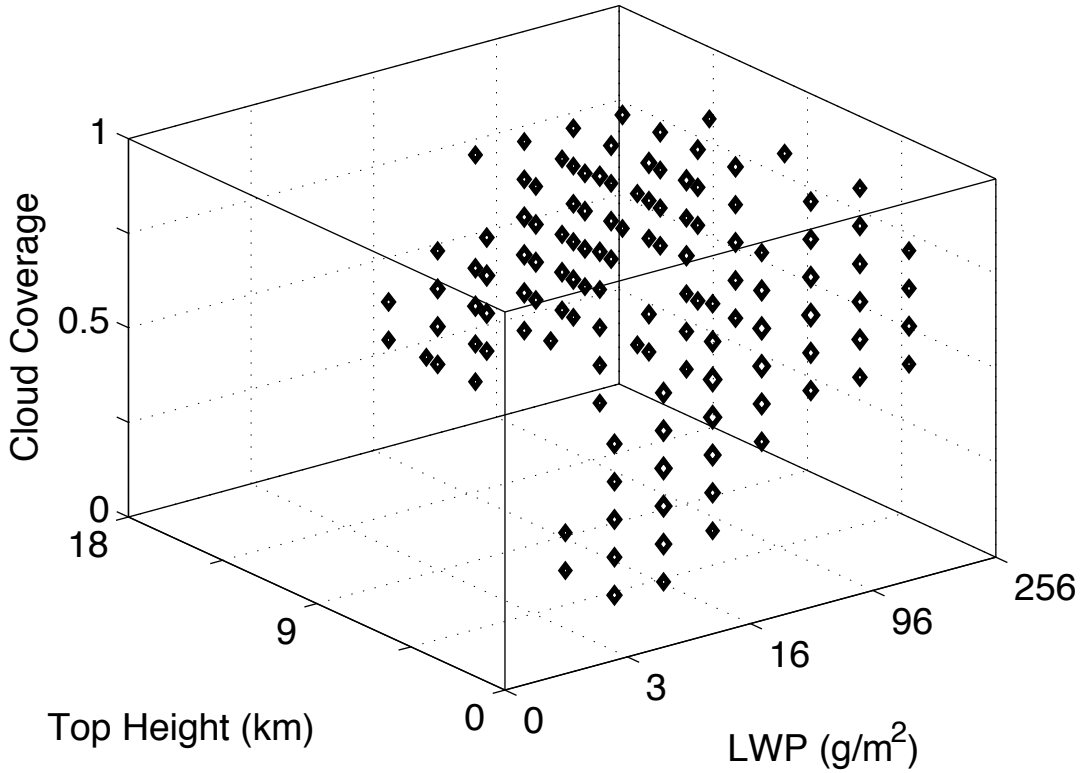


Figure 3.4 Same as Figure 3.1 but for the fourth three-dimensional centroid.

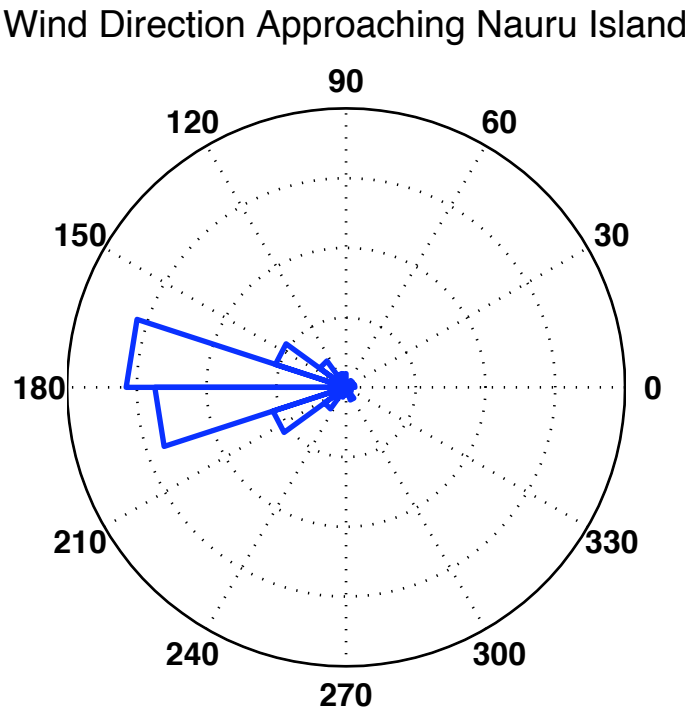


Figure 3.5 Relative frequency-of-occurrence of wind direction approaching Nauru Island between 2001 to 2004. Occurrences between 150 and 210 degrees indicate the wind is approaching from the east.

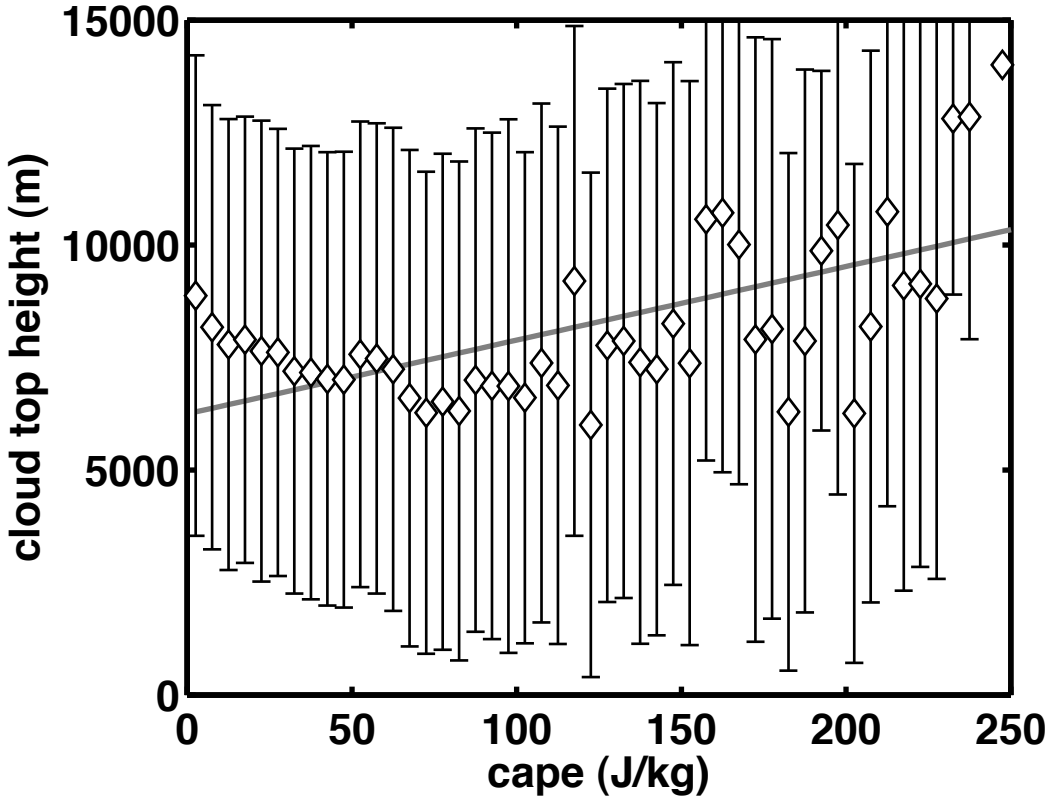


Figure 3.6 Scatterplot of CAPE versus cloud-top height bin-averaged by 5 J/kg intervals. Diamonds represent the average value and the error bars represent one standard deviation. The solid grey line represents a least mean squares fit of the data.

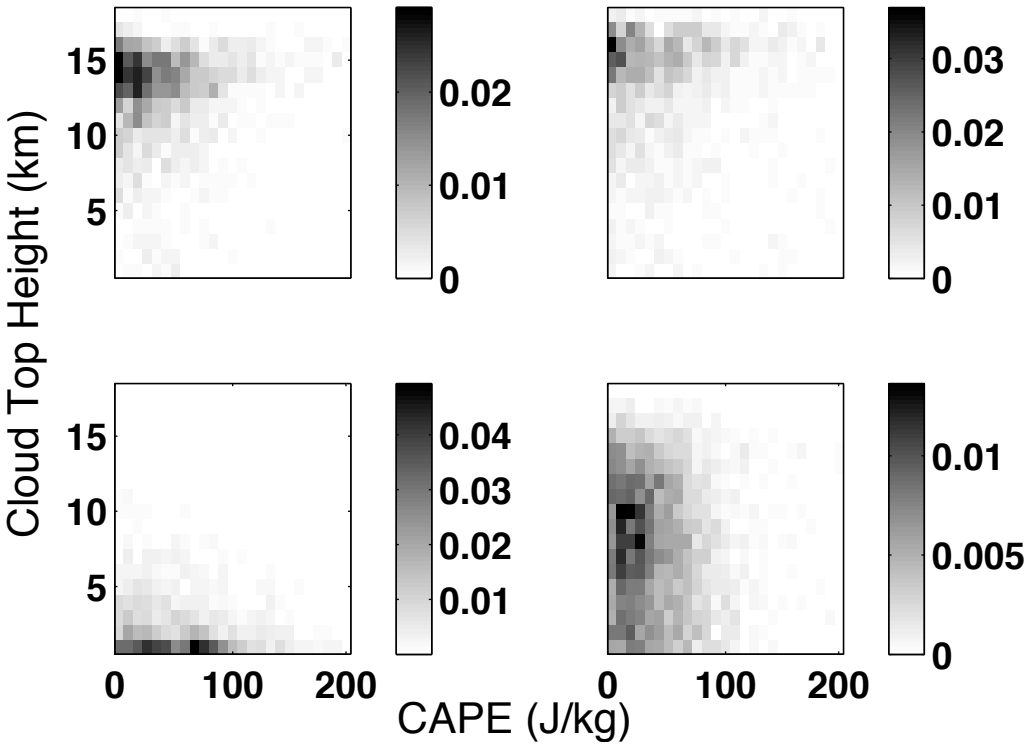


Figure 3.7 CAPE and cloud-top height relative frequency-of-occurrence histograms for each of the four clusters. Cluster 1 is upper left, cluster 2 is upper right, cluster 3 is lower left and cluster 4 is lower right.

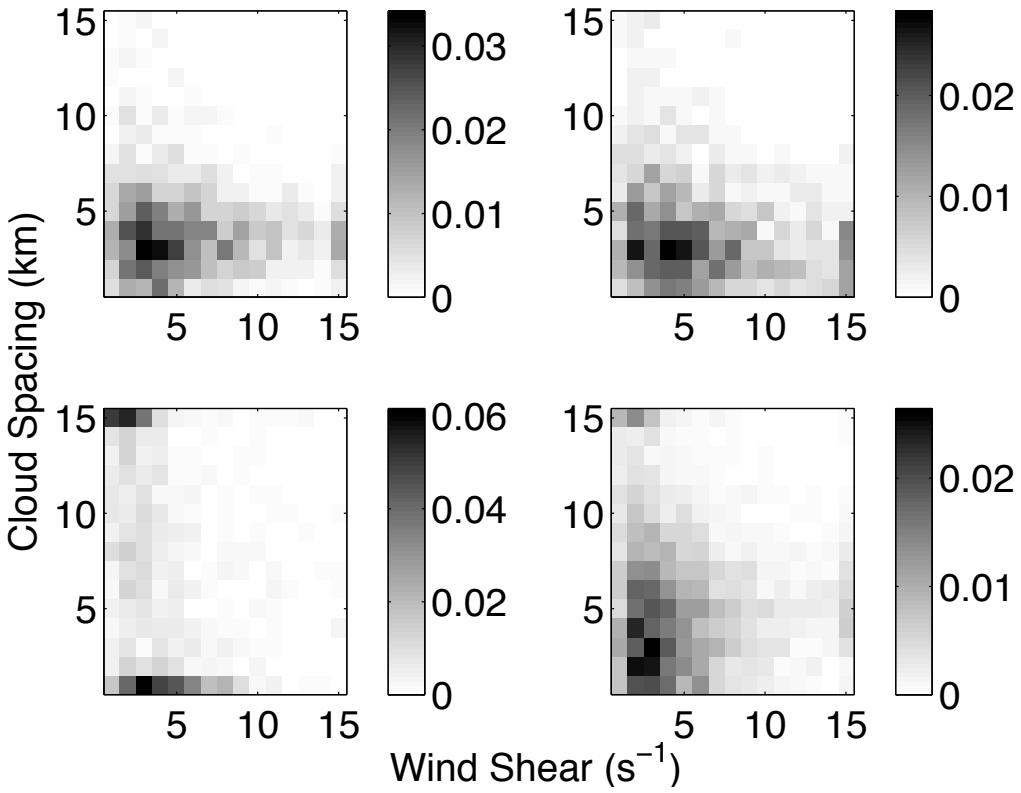


Figure 3.8 Mean wind shear between cloudy layers and mean spacing between cloudy layers relative frequency-of-occurrence histograms for each of the four clusters. Cluster 1 is upper left, cluster 2 is upper right, cluster 3 is lower left and cluster 4 is lower right.

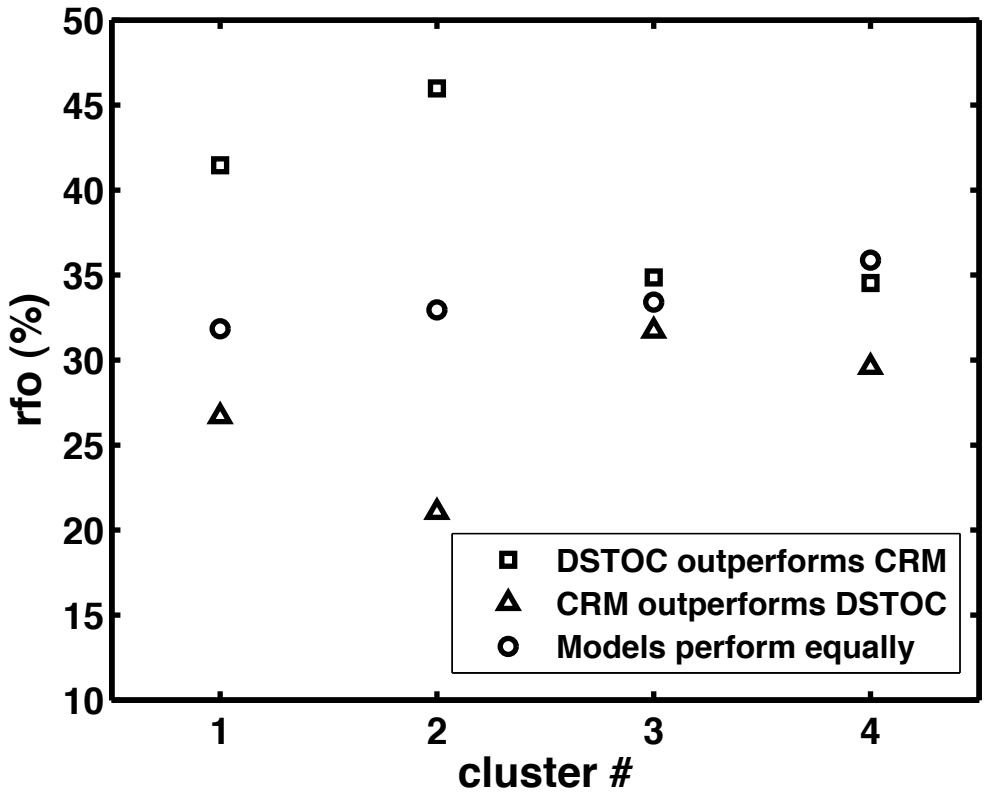


Figure 3.9 Relative frequency-of-occurrence (rfo) plot for each of the four clusters comparing the DSTOC and CRM model abilities to generate downwelling solar surface irradiance close to that observed. One model outperforms another when results are closer to observations by at least 5 percent of the observed value.

CHAPTER 4 – IDENTIFICATION OF OBSERVED CLOUD REGIMES IN CAM3

One of the primary goals of this work is to develop a method to improve model treatment of solar radiative transfer through geometrically complex cloud scenes. The focus of chapters 2 and 3 has been the development of an objective set of criteria, resolvable on a scale appropriate to GCM simulations, which may be used to determine when cloud field geometry is having significant effects on solar radiative fields. This has been done through the use of high-resolution, surface-based measurements from a handful of ARM TWP facilities. The next step, and the goal of this chapter, is to compare the results obtained from these high-resolution, single-point observations to those obtained from relatively low-resolution GCM simulations that span the entire tropics. Of course a comparison of this type is challenging; direct comparisons may only be performed at the ARM TWP sites, spatial resolution of the GCM and single-point measurements are difficult to compare, and over long time periods it is unrealistic to expect GCM-simulated variables to directly match observations. With these limitations in mind this chapter will address the following questions: 1) Can the GCM used in this study simulate the cloud regimes identified using the clustering techniques detailed in chapter 3; 2) How do distributions of observed versus GCM-simulated LWP, cloud-top height, and cloud coverage compare over time at the ARM TWP facilities; and 3) What can be concluded about differences between observed and GCM-simulated variables and how will these differences affect attempts to improve solar radiative transfer through cloud? The GCM used for this purpose is the Community Atmosphere Model (CAM3).

Section 4.1 describes the configuration of the CAM3 model and the simulations performed. Section 4.2 examines the distribution of cloud properties at two of the ARM facilities and the corresponding CAM3 results. Section 4.3 examines CAM3's ability to identify the cloud regimes discussed in chapter 3, and the implications and conclusions of this analysis are discussed in section 4.4.

4.1 METHODOLOGY

Model description

The Community Climate System Model (CCSM) has been developed through collaboration by researchers at Universities and government labs, and is housed, distributed, and maintained by the National Center for Atmospheric Research (NCAR), in Boulder, Colorado. The CCSM is a fully-coupled climate model composed of four primary components simulating the Earth's atmosphere, ocean, land-surface, and sea-ice. There are multiple versions of CCSM and different horizontal resolutions available for each version. CCSM3 is the most current version of the CCSM, and CAM3 represents the sixth generation of the atmospheric component of the CCSM. CAM3 is a three-dimensional global spectral model capable of being run either in a stand-alone or coupled mode, and uses 26 vertical layers (Collins et al., 2006). The version run for this study utilizes the T85 horizontal resolution, which equates approximately to 1.4 by 1.4 degree grid cell size or 256 by 128 horizontal gridpoints globally.

There are significant improvements in the treatment of cloud microphysical and condensation processes introduced in CAM3. These include the separate treatment of

cloud water and ice condensate, advection of these variables in large-scale circulations, improvement in convective parameterizations, and consistent treatment of cloud particles including sedimentation and radiative properties (Boville et al., 2006). Further information about the performance of CAM3 with respect to specific aspects of the global climate system can be found in a special issue of *Journal of Climate* (Vol. 19, No. 11; e.g., Hack et al., 2006a,b; Rasch et al., 2006). The cloud overlap in CAM3 is decoupled from the radiative parameterization and can vary among grid cells as well as time steps. The overlap assumption is capable of being set to any combination of maximum and random cloud overlap, but in general practice employs the standard maximum-random overlap assumption (Collins et al., 2006).

Model simulations

The model simulations used in this study are taken from those generated for the IPCC AR4 Special Report on Emission Scenarios (SRES). The CCSM3 IPCC SRES scenarios were run from 2000 to 2100 to estimate the likelihood of various aspects of future climate change. The simulations cover a range of possible futures including idealized emission or concentration assumptions for greenhouse gases and aerosol. Since the time period examined in this work is from the beginning of 2001 to the end of 2004, which is at the very beginning of the simulations and not yet projecting future changes, many of the SRES scenarios may be considered suitable for the purposes of this comparison. The SRES results used are those from the A1F1 scenario, which is a scenario that represents a continued heavy reliance on the burning of fossil fuels throughout the 21st century. This

scenario was chosen for a number of reasons: 1) The aerosol and greenhouse gas emissions are not drastically changed for the purposes of an idealized scenario (e.g., a complete freeze at 2000 levels); 2) The output is written at 6-hour intervals, and uses instantaneous as opposed to mean values; 3) Additional cloud-related output is available for this scenario, such as cloud fraction at each vertical level. The output from this CCSM3 simulation is available for download at the Earth System Grid website (http://www.earthsystemgrid.org).

4.2 CLOUD PROPERTIES

Liquid water path, cloud-top height, and cloud coverage

Since the CAM3 model results used in this study are generated for future climate scenarios and thus not forced with observations, it is difficult to directly compare hour-by-hour observed versus model-generated atmospheric properties. Another method, the one applied here, is to generate a distribution of these properties. In this case cloud-top height, total cloud coverage and LWP are used for the clustering algorithm in the previous chapter, and should therefore not only measure the CAM3 ability to match observed atmospheric conditions but also its potential to identify the cloud regimes generated from the ARM TWP facilities. Figures 4.1 and 4.2 show distributions of cloud-top height, LWP and cloud coverage generated by CAM3 and observed at the ARM Nauru and Manus facilities. A comparison to Darwin is not shown here as the CAM3 simulations span from 2001 to 2004 and do not overlap the 2006-2007 observations from Darwin. CAM3 output is written every six hours, so observations

taken at matching times are used for this comparison. Excluding times when observations are not available, either from missing data or quality control problems, a total of approximately 3000 hours are being examined. While the distributions of LWP are close for CAM3 and ARM observations at both sites, there are marked differences in cloud coverage and top height. The largest difference by far is for cloud-top height, where 69 percent of the time CAM3 simulates cirrus with a top height between 17 and 18 kilometers above Nauru Island and 98 percent of the time the cloud-top height is above 14 kilometers. Manus shares a similar situation with CAM3 simulating a cloud-top height between 17 and 18 kilometers 71 percent of the time and above 14 kilometers 99 percent of the time. This suggests that the clusters developed in chapter 3 with low- to mid-level cloud tops will be virtually nonexistent, based solely on these cloud-top heights.

There are a number of possible reasons for the persistent high cirrus in the CAM3 simulations. The transition from CCM3 to CAM2 introduced a large cold bias near the tropical tropopause resulting in a dry bias for stratospheric water vapor (Boville et al., 2006). In CAM3 the treatment of subvisible cirrus clouds was improved by separating the treatment of cloud ice and liquid particles and including additional sources and sinks, such as large-scale advection of cloud and gravitational settling of cloud particles. The result was that the radiative imbalance causing the cold bias in CAM2 was largely removed. It is possible that an increase in the stratospheric water vapor could generate more subvisible cirrus, causing the issues seen here. The question then remains of whether this high cirrus is supposed to exist or if it is an artifact of CAM3. It is certainly possible that the surface instruments used for the ARSCL VAP could miss some optically

thin high cirrus cloud, as the heights being discussed are near the upper limits of lidar range, and cirrus clouds are common in these regions. However, the cloud regimes generated in the Jakob et al. (2005) study, which uses satellite measurements of cloud-top pressure, suggest that the occurrence of cloud tops with pressures below 640 hPa is not uncommon. Specifically, at Manus a suppressed regime dominated by low clouds (cloud-top pressure > 680 hPa) was found to occur about 36 percent of the time, while the preponderance of other regimes contained cloud tops between 680 and 180 hPa. This is not definitive proof, as the retrieval of optical depths requires the use of a visible channel, leaving open the possibility of undetected subvisible cirrus. The next step in answering this question is to examine Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) data, which use active lidar and passive infrared measurements to probe thin cloud properties globally.

The distribution of cloud coverage also varies between CAM3 and ARM observations at both sites with the largest difference being that CAM3 generates more times with overcast conditions. The location of these overcast conditions is examined in more depth in the following section.

High, mid-level, and low cloud occurrence and cloud thickness

Figures 4.3 and 4.4 contain histograms of cloud fraction and geometric vertical thickness for high, mid-level and low cloud for Nauru and Manus, respectively. High cloud is defined as being located at levels less than 400 hPa, mid-level cloud is located between 400 and 740 hPa, and low cloud is located at levels greater than 740 hPa. These

pressure levels were chosen to correspond with level interfaces used by CAM3. These figures also include a relative frequency-of-occurrence of high, mid-level and low cloud. The data used in these histograms are the same as those used in the previous figures, with observed 5-minute averages matched to the CAM3 6-hourly instantaneous output. The cloud geometric thickness and fraction are derived from the ARSCL VAP at Nauru and Manus, which uses 512 vertical layers. This greater resolution is the reason that the CAM3 histograms have a blocky appearance for geometric thickness, as it uses only 26 vertical layers giving it a smaller range of possible thickness combinations. If cloud is detected in a layer it is assumed that it exists throughout the vertical extent of the layer.

Figure 4.3 shows the results at the Nauru facility. The first thing to note is that CAM3 generates high cloud 98 percent of the hours being examined. This indicates a persistent presence of high cirrus, as indicated by the large proportion of high cloud tops in the previous figures. This compares to 68 percent of the time when the ARSCL VAP identifies the presence of high cloud. One possible explanation for this is that since the spatial domain of CAM3 is approximately 1.4° by 1.4°, which is considerably larger than that of the single point measurements taken at the Nauru facility, there may often be small cirrus clouds present even when the spatial domain in general is clear. However, the distribution of cloud fraction and vertical extent for CAM3 shows a relatively high concentration of overcast conditions with geometrically thick clouds, suggesting these are not simply small individual cirrus clouds. This also explains the difference in the distribution of cloud coverage in figure 4.1, as the ARSCL data identifies far fewer times with overcast high clouds. The histogram pattern for high cloud for the ARSCL VAP indicate a concentration of geometrically thin clouds with low coverage along with a

general increase in the distribution of cloud thickness as cloud fraction increases. The ARSCL VAP does not identify times with overcast conditions with very thick clouds (> 6 kilometers), while CAM3 does not generate times with overcast clouds that are geometrically thin (< 3 kilometers). At least part of reason for the difference in thickness may be explained by the difference in vertical resolution. As mentioned earlier, the ARSCL VAP uses 512 vertical layers while CAM3 uses only 26, meaning a cloud generated in a single CAM3 layer is necessarily much thicker than one detected in a single ARSCL layer.

CAM3 generates mid-level cloud 36 percent of the time, compared to 42 percent of the time when it is detected by the ARSCL VAP. For low clouds, CAM3 generates them 85 percent of the time while the ARSCL VAP detects them 77 percent of the time. As with the high clouds CAM3 tends to have geometrically thicker clouds than the ARSCL data for both mid-level and low clouds. In addition, for both mid-level and low clouds the largest concentration of cloud occurrences for the ARSCL data have fractions between 0 and 0.1, while the highest concentration of cloud occurrences for the CAM3 data had fractions between 0.2 and 0.3. One interesting thing to note is that the ARSCL data show a pattern of very geometrically thick clouds with fractions between 0.3 and 0.4 for both the mid-level and low cloud histograms. These clouds are likely cumulus towers, and a similar pattern is not seen in the CAM3 data.

Figure 4.4 shows the same histogram comparison but for the ARM Manus facility. Many of the spatial patterns are similar to those at Nauru with a few key differences. First, similar to Nauru, the Manus CAM3 simulations generate persistent high cirrus cloud with a large concentration of overcast conditions. In fact, CAM3 generates high

cloud over Manus 99 percent of the time, compared to 67 percent of the time when the ARSCL VAP detects high cloud. Unlike Nauru, however, the CAM3 simulations generate quite a bit more cloud at the mid- and low levels as well; 13 and 12 percent more respectively. Both CAM3 and the ARSCL data detect mid-level cloud more often at Manus than at Nauru. One other point to make is that, while it was mentioned that the Nauru ARSCL VAP detected low clouds with large vertical extents with fractions between 0.3 and 0.4, the Manus ARSCL VAP detects low cloud with large geometric thickness across a much wider range of cloud fractions; between 0.2 and 0.7. This along with the greater occurrence of mid-level cloud suggests that the Manus facility may detect convective towers more often than Nauru.

4.3 IDENTIFYING CLUSTERS USING CAM3 OUTPUT

Figure 4.5 shows relative frequency-of-occurrence maps throughout the tropics for the 4 clusters generated at the ARM Nauru facility described in chapter 3. The maps are generated using output from the CAM3 simulations. The three variables used in the *k*-means clustering algorithm: cloud-top height, cloud coverage and LWP, are used as criteria for the CAM3 output. Values of cloud-top height, cloud coverage, and LWP that are within one standard deviation of the median values for each cluster are considered to have met the required criteria for that cluster. A cluster is considered present when all three variables fall within the specified range of values. Cluster 2, the moderate coverage cirrus regime, is the only cluster that maintains a consistent presence throughout the tropics, with an occurrence of between 8 and 12 percent over much of the ocean. There

is a drop in occurrence over the tropical warm pool east of New Zealand, which is coincidentally the approximate location of the ARM TWP facilities. In fact cluster 2 occurs only about 2 percent of the time at Nauru and Manus in the CAM3 simulations, with the other clusters occurring less than 1 percent of the time. Cluster 3, the low cloud-top boundary layer regime, occurs infrequently over most of the tropics but has a few areas of higher concentration: in the Indian Ocean, east of the Gulf of Mexico and off the eastern coast of Brazil in the Atlantic. Clusters 1 and 4, the overcast cirrus and mixed cloud-type regimes, both have some small pockets where their relative frequency-of-occurrence reached between 8 and 12 percent, but for the vast majority of the tropics they are virtually nonexistent.

Much of the reason for why the CAM3 cluster occurrence does not match that generated using the surface-based measurements at the ARM TWP facilities can be attributed to the CAM3 cloud-top height. As explained earlier, CAM3 frequently generates high cirrus clouds with cloud-top heights often above 17 kilometers, which does not fall within the range of cloud-top height values for any of the 4 clusters. This makes clustering based on cloud-top height problematic. There are a number of possible approaches to this problem. One would be to avoid using cloud-top height and instead opt with another variable, such as cloud-base height or thickness. This approach may have its own issues, however, as Figures 4.3 and 4.4 show distinct differences in the cloud thicknesses observed versus those generated by CAM3, and it is uncertain how well cloud base-height may determine the presence of specific cloud regimes. Another possibility would be to look for the presence of cloud within the range of values for cloud-top height for each cluster. One issue with this approach is that for a cluster

composed of multiple cloud types, such as cluster 4 at the Nauru facility, there could be considerable intersection among the clusters. Another approach, and the one used here, is to remove the subvisible cirrus generated by CAM3 when determining cloud-top height. Subvisible cirrus is defined as having an optical depth below 0.03 (Seifert et al., 2007). The optical depth for each cloudy layer is calculated for CAM3 output using the parameterization:

$$LWP = \tau * R_e * 5/9 \tag{4.1}$$

Where τ is visible optical depth and R_e is effective droplet radius. In this case the equation is rearranged to solve for τ. LWP is calculated for each layer using the layer thickness, cloud fraction and in-cloud water mixing ratio, while R_e is assumed to be 6 μm, which is consistent with the value used for the radiative transfer calculations from chapter 2.

Figure 4.6 shows relative frequency-of-occurrence maps for each of the clusters throughout the tropics with subvisible cirrus cloud removed. Cluster 2 has almost no occurrence in the tropics while cluster 3 exists 20 percent or more of the time over much of the tropical oceans. The range of values for cloud-top height in cluster 3 is from the surface to 3.25 kilometers. It would appear that the reliance on LWP to generate visible optical depth has removed most of the high cloud. It should be noted that cluster 4, which contains clouds with tops between 3.65 and 10.55 kilometers, has increased its frequency-of-occurrence in the equatorial region near New Zealand by 4 or 5 percent. Given that CAM3 generates mid-level cloud at Nauru and Manus 36 and 76 percent of the time respectively, we can surmise that the 0.03 optical depth restriction has also

removed some mid-level cloud, or else the cluster 4 cloud fraction requirement of being between 0.3 and 0.7 is not often being met. Regardless the extremely low occurrences of clusters 1 and 2 suggest the removal of subvisible cirrus using liquid water as a calculator of optical depth does not produce a highly accurate translation from CAM3 generated cloud tops to those observed at the ARM TWP sites.

A final point to consider is how often any of the clusters exists in the two scenarios given. Figure 4.7 shows frequency-of-occurrence maps for any of the clusters for the original CAM3 output and applying the 0.03 cloud optical depth requirement. For both maps there are areas close to the subtropics where the frequency-of-occurrence is between 25 and 35 percent. For the original model output much of the equatorial region between 60°E and 120°W has low occurrence, below 5 percent, with increased occurrences over the oceans closer to the subtropics. For the CAM3 output with the 0.03 cloud optical depth requirement, the occurrence ranges between 15 and 30 percent for most of the ocean areas, with discontinuities off the west coasts of Chile and Northern Africa, the east coasts of Australia, Madagascar and Brazil, and the Gulf of Mexico.

4.4 CONCLUSIONS

Model output from a CAM3 climate change scenario generated for use in IPCC AR4 is analyzed to determine how well a GCM can reproduce distributions of cloud variables such as cloud-top height, LWP and cloud coverage when compared with those observed at the ARM TWP Nauru and Manus facilities. The CAM3 output does a very good job of reproducing the distribution of LWP, but displays marked differences in the distribution

of cloud-top height and cloud coverage. Specifically, the CAM3 generates persistent, often overcast, cirrus cloud with very high top heights, while the surface-based observations at Nauru and Manus detect cirrus approximately 70 percent of the time with fewer occurrences of overcast conditions.

The frequency-of-occurrences of high, mid-level and low cloud is also compared along with pairings of geometric thickness and cloud fraction. It is found that CAM3 generates high cloud 98 and 99 percent of the time over Nauru and Manus, respectively, which is about 30 percent more often than it is observed by the ARSCL VAP at either site. It is also found that the CAM3 clouds tend to have higher vertical extent, which is likely due in part to the lower vertical resolution used in the model; 26 vertical layers in CAM3 versus 512 vertical layers in the ARSCL VAP. CAM3 cloud fractions have a higher concentration between 0.2 and 0.3 while the ARSCL VAP clouds have a concentration between 0.0 and 0.1. This is in addition to the high occurrence of overcast cirrus in CAM3 previously mentioned.

Relative frequency-of-occurrence maps are generated throughout the tropics using CAM3 output to determine how often the cloud regimes represented by the clusters developed in chapter 3 are present in the CAM3 simulations. The persistent high cirrus generated by CAM3 does not often fall within the cloud-top height ranges for the low- and mid-level cloud regimes. Cluster 2, which has the highest range of cloud-top height, is found to be present approximately 8 percent of the time over much of the tropical oceans, with a reduction around the equatorial region near New Zealand. In an attempt to bridge the gap between observed and model-generated cloud properties, a 0.03 cloud optical depth requirement is imposed upon CAM3 clouds when cloud-top height is

calculated. The result is a large increase in the frequency of cluster 3, which is composed primarily of low boundary-layer clouds, along with an accompanying slight increase in cluster 4, which is composed of multiple cloud types with heights in the mid-level range. Clusters 1 and 2 occur rarely under these conditions.

Finally the frequency-of-occurrence of any of the clusters is calculated for the original CAM3 output and with the 0.03 cloud optical depth requirement. In the original output there is little cluster presence in the equatorial region between 60°E and 120°W. This is the area where the ARM TWP sites are located and the persistent presence of very high cloud tops is an important reason for the rare presence of the clusters. For this scenario elsewhere over the tropical oceans, clusters are present between 15 and 20 percent of the time. In the modified cloud optical depth scenario there is a presence of clusters between 15 and 25 percent of the time throughout much of the ocean, most of which may be attributed to boundary-layer clouds from cluster 3.

With respect to the primary goal of this work of improving solar radiative transfer through geometrically complex cloud scenes, the translation of GCM-generated cloud properties to those observed is problematic. The clusters developed in chapter 3 may serve as a guide to determine when certain types of cloud regimes are present. This information, along with the data gathered from the comparison of the stochastic and plane-parallel radiative transfer models in chapter 2, may be used to determine when and how the stochastic approach to radiative transfer can best improve the treatment of 3D cloud field geometry. In this chapter, we examine one AGCM's ability to identify these clusters in the tropics. Though the translation from observed to model-generated cloud properties is clearly imperfect, a rough estimate of the potential of the stochastic

approach to improve solar radiative transfer may still be made. Averaged over the tropical oceans, both the original and removal of subvisible cirrus methodologies detect cluster presence approximately 20 percent of the time. In chapter 3 it was determined that DSTOC outperforms the CRM approximately 40 percent of the time averaged over all clusters. Using just these numbers, the stochastic approach has the potential to significantly improve shortwave radiative transfer calculations over the tropical oceans 8 percent of the time.

In this chapter we have identified some of the differences between model-generated and observed cloud properties, and have proposed one method, though imperfect, of bridging this gap. There are several possible next steps: 1) The large-scale variables used to develop the initial clusters could be reexamined and possibly replaced or modified with variables that are easier to interpret with the CAM3 output; 2) The method of identifying the current clusters in the CAM3 output may be refined, one readily apparent course of action is to include the effect of cloud ice when detecting subvisible cirrus with the hope of not removing so much of the observable mid-level and high cloud when determining cloud-top height; 3) The clusters could be initially developed using CAM3 output, and the stochastic radiative transfer model could likewise be forced with model output.

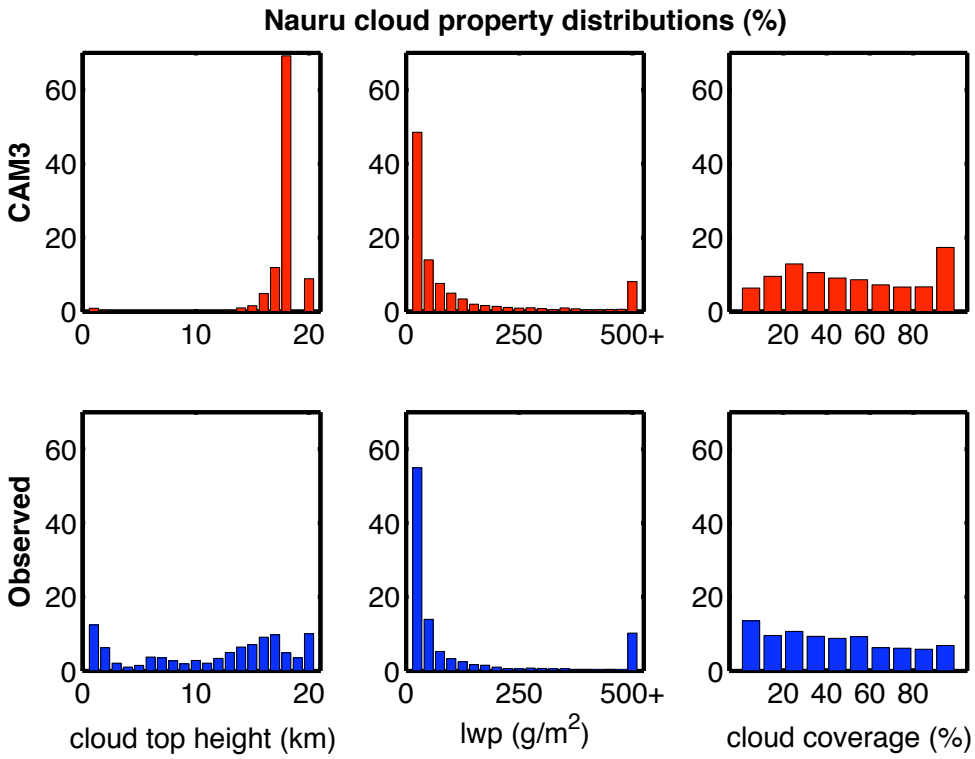


Figure 4.1 Distribution of cloud-top height, liquid water path and cloud coverage for observed versus model-generated 6-hourly data from the beginning of 2001 to the end of 2004 at the Nauru Island ARM facility. The top row represents the CAM3 simulations, and the bottom row represents data taken from the ARM ARSCL and MWRRET VAPs. The location of Nauru Island is 0.521° S, 130.891° E, while the closest CAM3 grid cell is centered at 0.7004° S, 167.3438° E.

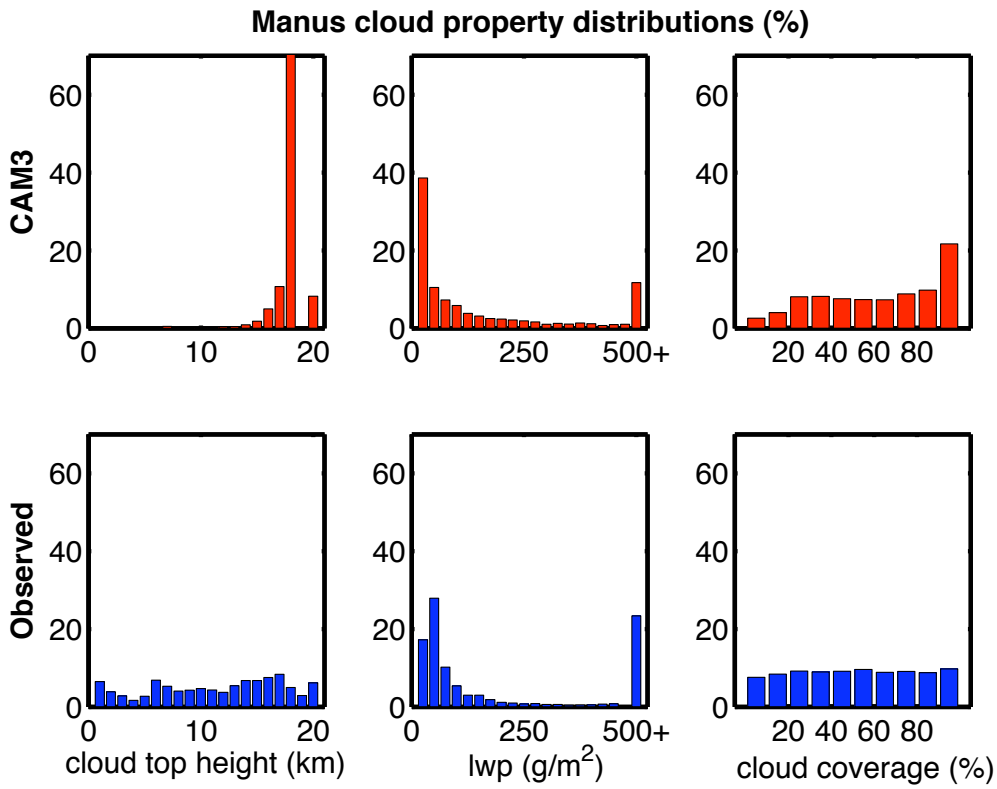


Figure 4.2 Distribution of cloud-top height, liquid water path and cloud coverage for observed versus model-generated 6-hourly data from the beginning of 2001 to the end of 2004 at the Manus ARM facility. The top row represents the CAM3 simulations, and the bottom row represents data taken from the ARM ARSCL and MWRRET VAPs. The location of Manus is 2.006° S, 147.425° E, while the closest CAM3 grid cell is centered at 2.1012° S, 147.6563° E.

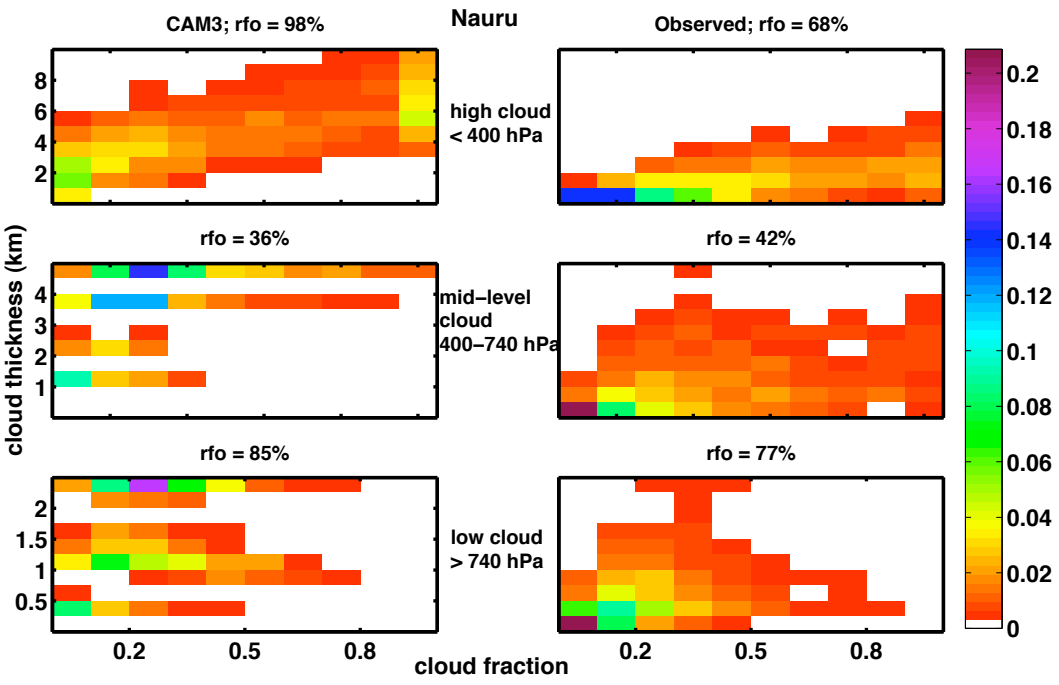


Figure 4.3 Paired histograms of geometric vertical cloud thickness and cloud fraction for CAM3 (left column) and observations (right column) for high, mid-level, and low clouds. The relative frequency-of-occurrence (rfo) of cloud type is displayed above each box. The observed clouds are derived from the ARSCL VAP at the ARM Nauru facility. The color bar represents the rfo of each pairing. Cloud fraction and geometric height are allocated into 10 equally-spaced bins.

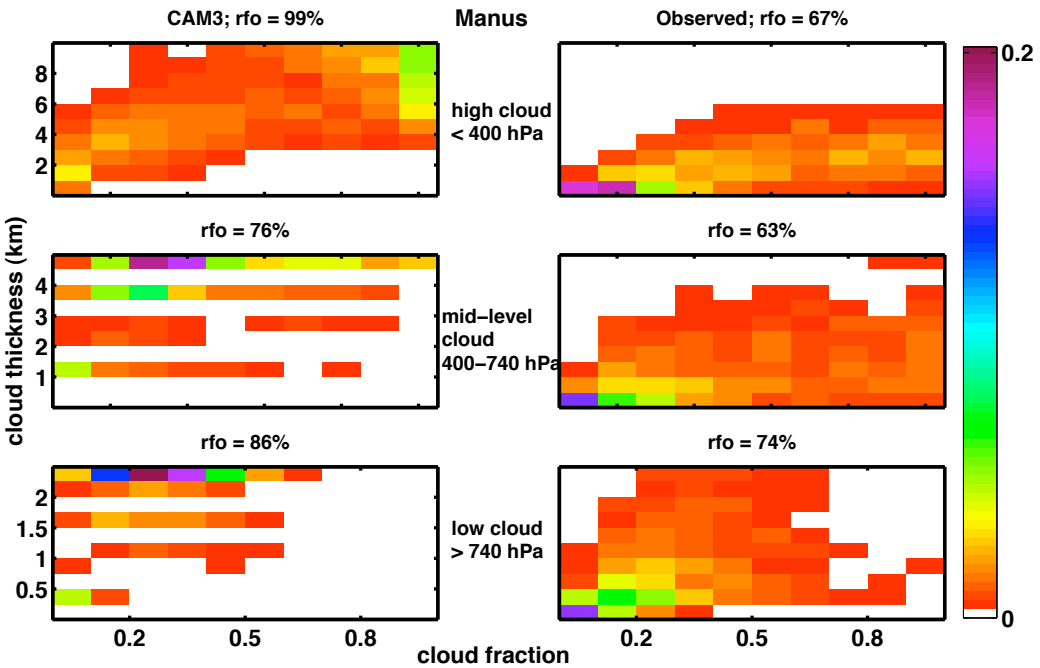


Figure 4.4 Same as Figure 4.3 but for Manus Facility.

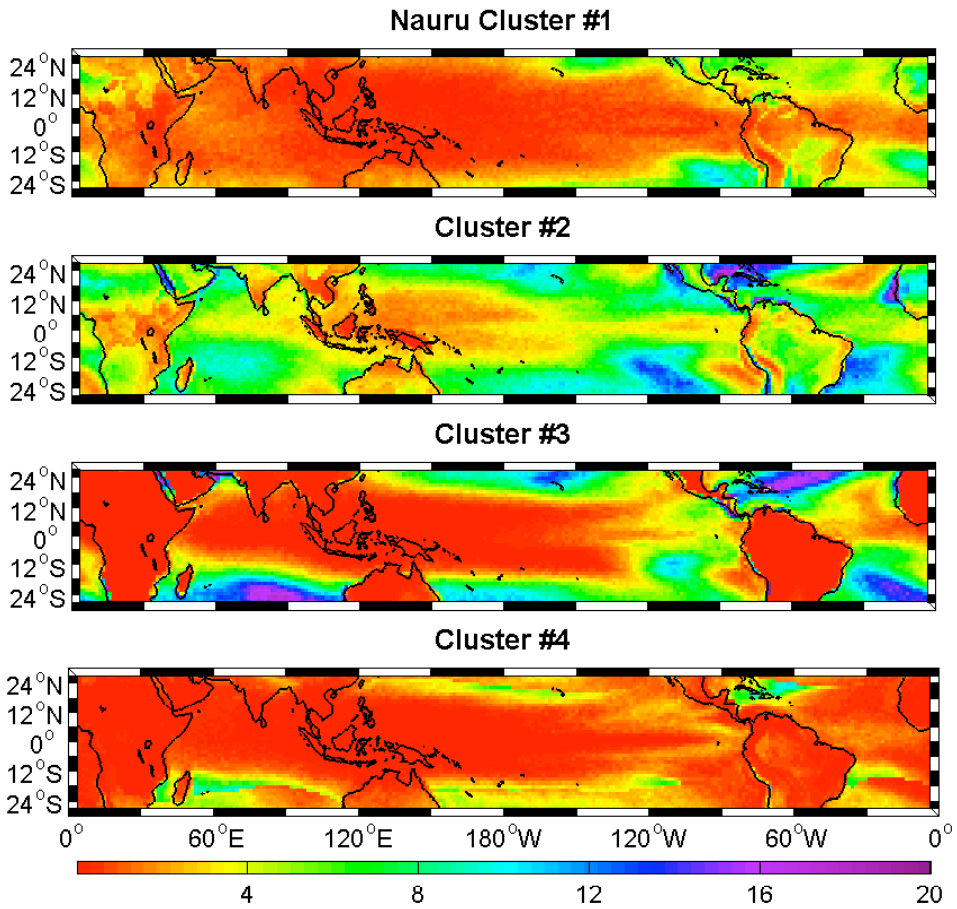


Figure 4.5 Relative frequency-of-occurrence maps generated using CAM3 simulations in the tropics from 2001 to 2004. The color bar represents the percent occurrence of the clusters described in chapter 3.

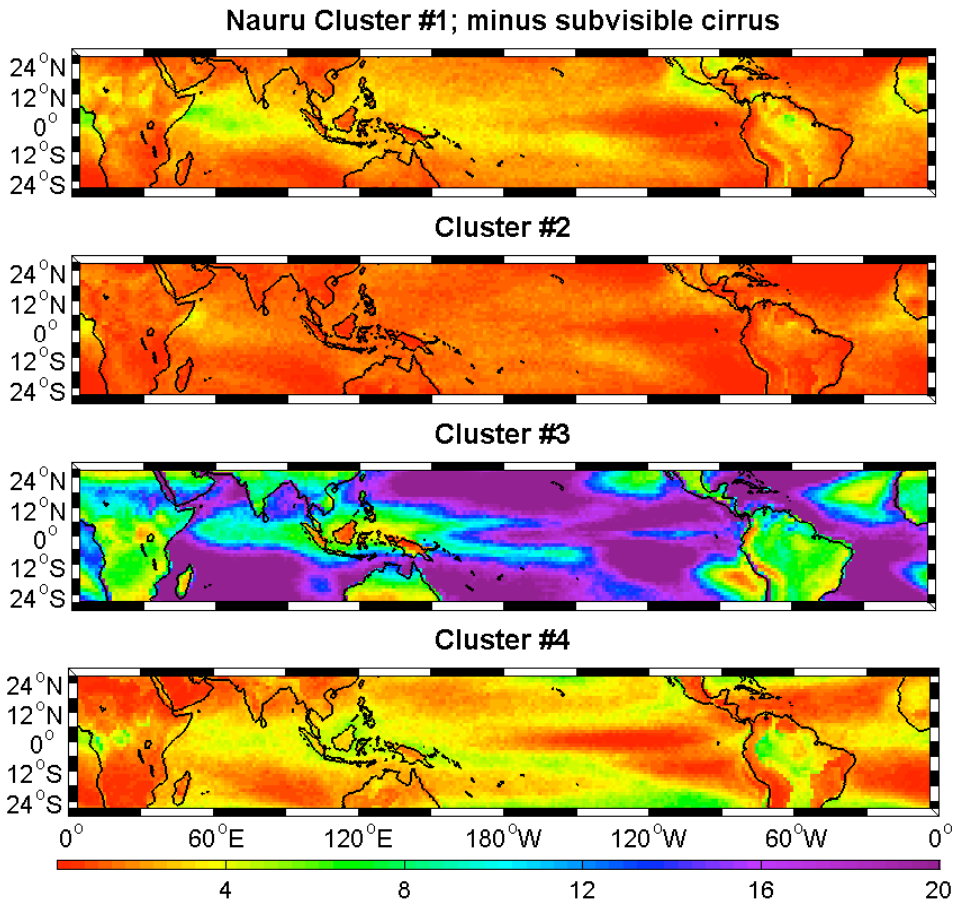


Figure 4.6 Relative frequency-of-occurrence maps generated using CAM3 simulations in the tropics from 2001 to 2004. The color bar represents the percent occurrence of the clusters described in chapter 3. In order for a cloud-top height to be used as a criterion for these clusters the cloud layer must possess a minimum optical depth of 0.03.

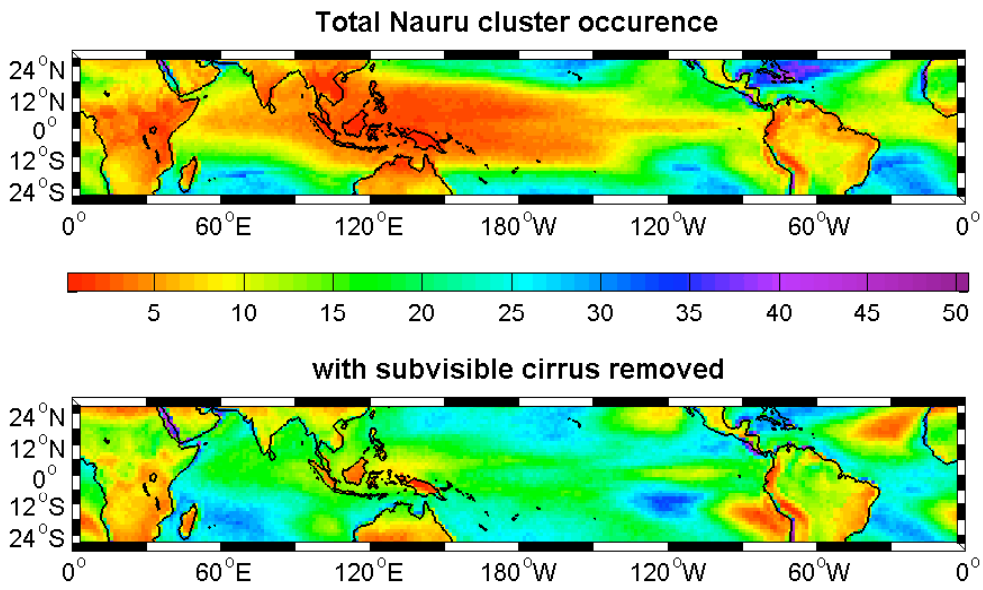


Figure 4.7 Relative frequency-of-occurrence maps generated using CAM3 simulations in the tropics from 2001 to 2004. The color bar represents the percent summed occurrence of all the clusters presented in chapter 3. For a cloud-top height to be used as a criterion for these clusters in the bottom map, the cloud layer must possess a minimum optical depth of 0.03.

CHAPTER 5 – DISCUSSION AND CONCLUSIONS

The comparison of off-line plane-parallel and stochastic shortwave radiative transfer simulations for a variety of observed cloud scenes over a four-year period indicate that, given the proper atmospheric conditions, the stochastic approach does in fact have the ability to significantly improve calculations of radiative transfer through cloudy atmospheres. Further analysis of these results suggest that cloud coverage, LWP in relation to geometric vertical cloud thickness, and the ratio of diffuse to direct SSI exhibit identifiable characteristics during times when the stochastic model outperforms its plane-parallel counterpart. Monthly and diurnal patterns in model performance have also been identified. This information is used to develop a preliminary set of criteria to determine when the stochastic approach to radiative transfer is most appropriate. A simplistic parameterization is developed using a multiple linear regression technique and applied using the aforementioned criteria. The results generated using this parameterization are encouraging, often showing improvement over the plane-parallel radiative transfer model when compared against observed SSI.

Limitations in this comparison include the stochastic models' ability to use different descriptions of variability in the cloud field, as only a Markovian description of the cloud sizes in the field is utilized for these simulations. This distribution of cloud chord lengths is representative of the clouds present at the ARM TWP Nauru site and may not be appropriate for the entire tropics, let alone the rest of the globe. In addition, the large number of hours and atmospheric conditions creates significant variance in the model results when compared to observations, indicating that a further refinement of the criteria

and method of radiative flux adjustment is required before this parameterization is ready for general application to an AGCM.

In order to further differentiate the atmospheric conditions necessary for optimal use of the stochastic approach, as well as to expand the scope of the global area for which this approach is appropriate, a *k*-means clustering algorithm is applied to surface based measurements of cloud-top height, cloud coverage, and LWP for three ARM TWP facilities. Four cloud regimes are identified at the ARM TWP Nauru site, each with distinct characteristics. The dynamical and thermodynamical properties of each of these cloud regimes are examined individually. Three of the four regimes show signs of being convectively active, while the other regime appears to be composed primarily of stable boundary layer clouds. Though the convectively active regimes have on average higher cloud tops then the stable regime, only a relatively weak linear relationship is found between cloud-top height and the magnitude of CAPE, both for the entire set of data and for each cluster individually. Examination of wind shear and vertical spacing between cloudy layers suggest that the convectively active regimes tend toward large values of wind shear and smaller spacing between cloudy layers than does the stable regime. The combination of cloud-top height, wind shear, and spacing between cloudy layers directly affects the depth of the cloud field and the vertical and horizontal spacing between clouds, making it a potentially good indicator of macroscale cloud field inhomogeneity.

Comparison of model performance for each cluster involves examination of the models' ability to capture the variability in observed shortwave cloud forcing at the surface, the percent time when one model outperforms the other as previously defined, and the percent difference of model-generated from observed SSI. The stochastic model

captures more of the variability of the observed solar radiative cloud forcing at the surface than the CRM for all four clusters, and this difference is particularly noticeable for clusters 1 and 2. These clusters represent regimes that often contain cirrus clouds with large ranges of cloud coverage and LWP, and they also have the highest median values for wind shear between cloudy layers and lowest median values for spacing between cloudy layers. In addition they are the clusters where the stochastic model outperforms the CRM a significant portion of the time. In terms of difference from observed SSI, the stochastic model generates median SSI closer to that observed by 6 percent or greater for all clusters during times when it is considered to be outperforming the CRM. This suggests that, in order to maximize the potential for the stochastic approach to improve AGCM radiative transfer calculations through cloud, a secondary set of criteria must be developed for each cluster. Results from the initial comparison of the stochastic and the plane-parallel model against observations may serve as a starting point, but additional work still needs to be performed in this area.

The cluster analysis is expanded to include the ARM TWP Manus and Darwin facilities. It is found that although there are variations in the relative frequency-of-occurrence and the number and structure of the clusters, the primary cloud regimes found at Nauru may also be found at other locations. This suggests it should be possible to apply a parameterization developed at Nauru to other areas in the tropics.

Translating results obtained from observations at the ARM TWP facilities to a parameterization useable by an AGCM requires a comparison of observed to AGCM-generated cloud properties. To this end output from the CAM3 is analyzed to determine how well it can reproduce distributions of cloud-top height, LWP, and cloud coverage

when compared against those observed. The comparison is also expanded to include CAM3's ability to reproduce histograms of cloud coverage and vertical geometric thickness as well as frequency-of-occurrence of high, mid-level, and low clouds. It is found that CAM3 reproduces distributions similar to those observed for LWP over the ARM TWP sites, but generates persistent, often overcast, cirrus cloud with tops often reaching above 17 kilometers, while the surface-based observations at Nauru and Manus detect cirrus approximately 70 percent of the time with fewer occurrences of overcast conditions. An examination of passive infrared and active lidar measurements from CALIPSO is currently underway to determine if this discrepancy is an artifact of CAM3 or due to the surface instrumentation's inability to detect high subvisible cirrus clouds. Compared to observations, CAM3 also underestimates mid-level cloud and overestimates low cloud at Nauru by small margins, and overestimates mid-level and low cloud occurrence at Manus by 13 and 12 percent, respectively. CAM3 clouds tend to have higher geometric vertical thickness, which is likely due in part to the lower vertical resolution used in the model; 26 vertical layers in CAM3 versus 512 vertical layers in the ARSCL VAP. In addition to the higher frequency of overcast cirrus, CAM3 cloud fractions have higher values between 0.2 and 0.3 while the ARSCL VAP clouds have larger values between 0 and 0.1.

Two methodologies are tested for identifying occurrence of the clusters presented in chapter 3 for CAM3 output generated throughout the tropics. The first is a direct comparison of CAM3 cloud fields with those observed in the clusters; using a range of one standard deviation of the clusters' median values for cloud coverage, LWP, and cloud-top height as criteria for a cluster's presence. Cluster 2, which has the highest

cloud-top range, is present approximately 8 percent of the time over much of the tropical ocean, with a dip around the equatorial region near New Zealand. The other clusters are present less frequently. Overall there is little cluster presence in the equatorial region between 60°E and 120°W, which is the area where the ARM TWP sites are located and is likely largely due to persistent very high cloud tops. For this scenario elsewhere over the tropical oceans, clusters are present between 15 and 20 percent of the time.

The second methodology attempts to bridge the gap between observed and model-generated cloud properties by imposing a 0.03 cloud optical depth requirement upon CAM3 clouds when cloud-top height is calculated; the idea being that much of the high cirrus generated by CAM3 is subvisible, and is either not detected by the surface-based instruments or is an artifact of the model. The result is a large increase in the presence of cluster 3, which is composed primarily of low boundary layer clouds, with an accompanying slight increase in cluster 4, which is composed of multiple cloud types with heights in the mid-level range. Clusters 1 and 2 occur rarely in this scenario. Clusters are present between 15 and 25 percent of the time throughout much of the ocean, most of which may be attributed to boundary-layer clouds from cluster 3.

Possible next steps for improvement of the CAM3's ability to identify these clusters include changing the large-scale variables used to develop the clusters with ones that translate more smoothly to CAM3 output, refining the method of identifying the current clusters in CAM3, or developing clusters directly using CAM3 output. The issue with this third approach is that if CAM3, or AGCMs in general, contain biases when generating certain cloud properties, then tuning a parameterization to accommodate these biases is not necessarily a step forward.

Perhaps the most encouraging aspect of these results is the potential of the stochastic approach to improve AGCM treatment of complex cloud field geometry throughout the tropics. If we only examine the two methods put forth in this study, and assume the stochastic approach has the ability to improve shortwave fluxes 40 percent of the time, which is the average amount for the 4 clusters derived in chapter 3, a significant improvement in SSI can be achieved in CAM3 over the tropical oceans approximately 8 percent of the time. This translates to an average improvement of 2 Wm^{-2} for all hours in the tropics. These numbers will almost certainly increase with additional work in identifying the presence of specific cloud regimes for CAM3 output. This calculation is perhaps somewhat premature given the additional work that needs to be performed before a parameterization ready for general application is ready, but the potential for improvement is encouraging nonetheless.

References

Ackerman, T. P., and G. M. Stokes, 2003: The Atmospheric Radiation Measurement Program. *Phys. Today*, **56,** 38–44.

Adams, M. L., E. W. Larsen, and G. C. Pomraning, 1989: Benchmark results for particle transport in a binary Markov statistical medium, *J. Quant. Spectrosc. Radiat. Transfer*, **42**, 253-266.

Anderberg, M. R., 1973: *Cluster Analysis for Applications*. Academic Press, 359 pp.

Barker, H. W., and J. A. Davies, 1992: Solar radiative fluxes for stochastic, scale-invariant broken cloud fields. *J. Atmos. Sci.*, **49**, 1115-1126.

Barker, H. W., B. A. Wielicki, and L. Parker, 1996: A parameterization for computing grid-averaged solar fluxes for inhomogeneous marine boundary layer clouds. Part II: Validation using satellite data. *J. Atmos. Sci*, **53**, 2304- 2316.

Boukabara, S.A., Clough, S.A. and R.N. Hoffman, 2001: MonoRTM: A monochromatic radiative transfer model for microwave and laser calculation. *In Programs and Abstracts: Specialist Meeting on Microwave Remote Sensing*, **158**, November 2001.

Boville, B.A., P.J. Rasch, J.J. Hack, and J.R. McCaa, 2006: Representation of Clouds and Precipitation Processes in the Community Atmosphere Model Version 3 (CAM3). *J. Climate*, **19**, 2184–2198.

Briegleb, B.P., 1992: Delta-Eddington C Approximation for Solar Radiation in the NCAR Community Climate Model, *J. Geophy. Res.*, **97**, 7603-7612.

Byrne, R. N., R. C. J. Somerville, and B. Subasilar, 1996: Broken-cloud enhancement of solar radiation absorption, *J. Atmos. Sci.,* **53**, 878-886.

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

Cahalan, R. F., 1994: Bounded cascade clouds albedo and effective thickness. *Nonlinear Proc. Geophys.*, **1**, 156-167.

Cess, R.D., G.L. Potter, J.P. Blanchet, G.J. Boer, A.D. Del Genio, M. Deque, V. Dymnikov, V. Galin, W.L. Gates, S.J. Ghan, J.T. Kiehl, A.A. Lacis, H. Le Treut, Z.-X. Li, X.-Z. Liang, B.J. McAvaney, V.P. Meleshko, J.F.B. Mitchell, J.-J. Morcrette, D.A. Randall, L. Rikus, E. Roeckner, J.F. Royer, U. Schlese, D.A. Sheinin, A. Slingo, A.P. Sokolov, K.E. Taylor, W.M. Washington, R.T. Wetherald, Yagai, and M.-H. Zhang 1990: Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. *J. Geophys. Res.* **95**, 16601-16615.

Clothiaux, E.E., T.P. Ackerman, G.G. Mace, K.P. Moran, R.T. Marchand, M.A. Miller, and B.E. Martner, 2000: Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites. *J. Appl. Meteor.,* **39,** 645–665.

Collins, W. D., P. J. Rasch, B. A. Boville, J. J. Hack, J. R. McCaa, D. L . Williamson, B. P. Briegleb, C. M. Bitz, S.-J. Lin, and M. Zhang, 2006: The formulation and atmospheric simulation of the Community Atmospheric Model Version 3 (CAM3). *J. Climate*, **19**, 2144-2161, doi:10.1175/JCLI3760.1.

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., A. Marshak, R. F. Cahalan, and W. J. Wiscombe, 1997: The Landsat scale break in stratocumulus as a three-dimensional radiative transfer effect: Implications for cloud remote sensing. *J. Atmos. Sci.*, **54**, 241-260.

Di Giuseppe, F., and A. M. Tompkins, 2003: Effect of spatial organisation on solar radiative transfer in three dimensional idealised stratocumulus cloud fields. *J. Atmos. Sci.*, **60**, 1774-1794.

Di Giuseppe, F., and A. M. Tompkins, 2005: Impact of cloud cover on solar radiative biases in deep convective regimes. *J. Atmos. Sci.*, **62**, 1989-2000.

Fu, Q., M. C. Cribb, H. W. Barker, S. K. Krueger, and A. Grossman, 2000: Cloud geometry effects on atmospheric solar absorption. *J. Atmos. Sci.*, **57**, 1156-1168.

Gordon, N. D., J. R. Norris, C. P. Weaver, and S. A. Klein, 2005: Cluster analysis of cloud regimes and characteristic dynamics of midlatitude synoptic systems in observations and a model. *J. Geophys. Res.*, **110**, doi:10.1029/2004JD005027.

Grabowski, W. W., 2001: Coupling cloud processes with the large-scale dynamics using the cloud-resolving convection parameterization (CRCP). *J. Atmos. Sci.*, **58**, 978-997.

Hack J., J. Caron, G. Danabasoglu, K. Oleson, C. Bitz, and J. Truesdale, 2006a: CCSM–CAM3 climate simulation sensitivity to changes in horizontal resolution. *J. Climate*, **19**, 2267–2289.

Hack J., J. Caron, S. Yeager, K. Oleson, M. Holland, J. Truesdale, and P. Rasch, 2006b: Simulation of the global hydrological cycle in the CCSM Community Atmosphere Model (CAM3): Mean features. *J. Climate*, **19**, 2199–2221.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, Eds., 2001: *Climate Change 2001: The Scientific Basis*. Cambridge University Press, 881 pp.

Hozumi, K, T Harimaya, and C Magono, 1982: The size distribution of cumulus clouds as a function of cloud amount. *J. Meteor. Soc. Japan*, **60**, 691-699.

Jakob, C., G. Tselioudis, and T. Hume, 2005: The radiative, cloud and thermodynamic properties of the major tropical western Pacific cloud regimes. *J. Climate*, **18**, 1203-1215.

Jakob, C. and G. Tselioudis, 2003: Objective identification of cloud regimes in the Tropical Western Pacific. *Geophys. Res. Lett.*, **30**, 2082, doi:10.1092/2003GL018367.

Jensen, M.P., and A.D. Del Genio, 2006: Factors limiting convective cloud-top height at the ARM Nauru Island climate research facility. *J. Climate*, **19**, 2105-2117, doi:10.1175/JCLI3722.1.

Kassianov, E., T. Ackerman, and P. Kollias, 2005a: The role of cloud-scale resolution on radiative properties of oceanic cumulus clouds. *J. Quant. Spect. Rad. Trans.*, **91**, 211-226.

Kassianov, E., C.N. Long, and M Ovtchinnikov, 2005b: Cloud Sky Cover versus Cloud Fraction: Whole-Sky Simulations and Observations. *J. Appl. Meteorol.,* **44**, 86-98.

Khairoutdinov, M. F., and D. A. Randall, 2001: A Cloud Resolving Model as a Cloud Parameterization in the NCAR Community Climate System Model: Preliminary Results. *Geophys. Res. Lett.*, **28**, 3617-3620.

Kiehl, J.T., J.J. Hack, G.B. Bonan, B.A. Boville, D.L. Williamson and P.J. Rasch, 1998: The National Center for Atmospheric Research Community Climate Model: CCM3, *J. Clim.*, **11**, 1131-1149.

Klein, S.A., R.B. McCoy, H. Morrison, A.S. Ackerman, A. Avramov, G. de Boer, M. Chen, J.N.S. Cole, A.D. Del Genio, M. Falk, M.J. Foster, A. Fridlind, J.C. Golaz, T. Hashino, J.Y. Harrington, C. Hoose, M.F. Khairoutdinov, V.E. Larson, X. Liu, Y. Luo, G.M. McFarquhar, S. Menon, R.A.J. Neggers, S. Park, M.R. Poellot, J.M. Schmidt, I. Sednev, B.J. Shipway, M.D. Shupe, D.A. Spangenberg, Y.C. Sud, D.D. Turner, D.E. Veron, K. von Salzen, G.K. Walker, Z. Wang, A.B. Wolf, S. Xie, K. Xu, F. Yang, and G. Zhang, 2008: Intercomparison of model simulations of mixed-phase clouds observed during the ARM Mixed-Phase Arctic Cloud Experiment. Part I: Single layer cloud. Submitted to *Q. J. Roy. Meteor. Soc.*

Kobayashi, T., 1988: Parameterization of reflectivity for broken cloud fields. *J. Atmos. Sci.*, **45**, 3034-3045.

Kollias, P., E.E. Clothiaux, B.A. Albrecht, M.A. Miller, K.P. Moran, and K.L. Johnson, 2005: The Atmospheric Radiation Measurement program cloud profiling radars: An evaluation of signal processing and sampling strategies, *J. Atmos. Oceanic Technol.*, **22**(7), 930– 948.

Lane, D. E., K. Goris, and R. C. J. Somerville, 2002: Radiative transfer through broken cloud fields: Observations and model validation. *J. Climate*, **15**(20), 2921-2933.

Lane-Veron, D. E., and R. C. J. Somerville, 2004: Stochastic Theory of Radiative Transfer Through Generalized Cloud Fields, *J. Geophys. Res.*, **109**, doi:10.1029/2004JD004524.

Long, C.N., T.P. Ackerman, J.J. DeLuisi, and J. Augustine, 1999: Estimation of fractional sky cover from broadband SW radiometer measurements. *Proc. 10th Conf. on Atmos. Rad.*, June 28-July 2, 1999, Madison, Wisconsin.

Long, C.N. and T.P. Ackerman, 2000: Identification of clear skies from broadband pyranometer measurements and calculation of downwelling shortwave cloud effects. *J. Geophys. Res.*, **105** (D12), 15609-15626.

Long, C.N., 2001: The Shortwave (SW) Clear-Sky Detection and Fitting Algorithm: Algorithm Operational Details and Explanations, Atmospheric Radiation Measurement Program Technical Report, ARM TR-004.

Long, C.N., T.P. Ackerman, K.L. Gaustad, and J.N.S. Cole, 2006: Estimation of fractional sky cover from broadband shortwave radiometer measurements. *J. Geophys. Res.*, **111**, doi:10.1029/2005JD006475.

Lovejoy, S. and D. Schertzer, 2006: Multifractals, cloud radiances and rain, *J. Hyrol.*, **322**, 59-88.

Mace, G.G. and S. Benson-Troth, 2002: Cloud-Layer Overlap Characteristics Derived from Long-Term Cloud Radar Data. *J. Climate*, **15**, 2505-2515.

Malvagi, F., R. N. Byrne, G. C. Pomraning, and R. C. J. Somerville, 1993: Stochastic radiative transfer in a partially cloudy atmosphere, *J. Atmos. Sci.*, **50**, 2146-2158.

Marshak, A., A. B. Davis, W. J. Wiscombe, and R. F. Cahalan, 1997: Inhomogeneity effects on cloud shortwave absorption measurements: Two-aircraft simulations. *J. Geophys. Res.*, **102**, 16619-16637.

Marshak, A., A. Davis, W. Wiscombe, W. Ridgway, and R. Cahalan, 1998: Biases in shortwave column absorption in the presence of fractal clouds. *J. Climate*, **11**, 431-446.

Marshak A., and A.B. Davis (editors), 2005: *3D Radiative Transfer in Cloudy Atmospheres*. Springer, New York, 686 pp.

McClatchey, R. A., R. W. Fenn, J. E. A. Selby, F. E. Volz, and J. S. Garing, *Optical properties of the atmosphere*, 3rd ed., Environmental research papers, no. 411, Air Force Cambridge Research Laboratories, Bedford, MA, 1972.

Morrison, H., R.B. McCoy, S.A. Klein, S. Xie, Y. Luo, A. Avramov, M. Chen, J.N.S. Cole, M. Falk, M.J. Foster, A.D. Del Genio, J.Y. Harrington, C. Hoose, M.F. Khairoutdinov, V.E. Larson, X. Liu, G.M. McFarquhar, M.R. Poellot, K. von Salzen, B.J. Shipway, M.D. Shupe, Y.C. Sud, D.D. Turner, D.E. Veron, G.K. Walker, Z. Wang, A.B. Wolf, K. Xu, F. Yang, and G. Zhang, 2008: Intercomparison of model simulations of mixed-phase clouds observed during the ARM Mixed-Phase Arctic Cloud Experiment, Part II: Multi-layered cloud. Submitted to *Q. J. Roy. Meteor. Soc.*

Naud, C.M., A. Del Genio, G.G. Mace, S. Benson, E.E. Clothiaux, and P. Kollias, 2008: Impact of dynamics and atmospheric state on cloud vertical overlap. *J. Climate*, **21**, 1758-1770, doi:10.1175/2007JCLI1828.1.

O'Hirok, W., and C. Gautier, 1998: A three-dimensional radiative transfer model to investigate the solar radiation within a cloudy atmosphere. Part I: Spatial effects. *J. Atmos. Sci.,* **55**, 2162–2179.

Pincus, R. and S.A. Klein, 2000: Unresolved spatial variability and microphysical process rates in large-scale models. *J. Geophys. Res.* **105**, 27,059-27,065.

Planck, V. G., 1969: The size distribution of cumulus clouds in representative Florida populations. *J. Appl. Meteorol.*, **8**, 46-67.

Pomraning, G. C., 1991: Linear *Kinetic Theory and Particle Transport in Stochastic Mixtures*. World Scientific Publishing, Singapore, 234 pp.

Potter, P. L. and R. D. Cess, 2004: Testing the impact of clouds on the radiation budgets of 19 atmospheric general circulation models. *J. Geophys. Res.*, **109**, D02106.

Randall, D., S. Krueger, C. Bretherton, J. Curry, P. Duynkerke, M. Moncrieff, B. Ryan, D. Starr, M. Miller, W. Rossow, G. Tselioudis, and B. Wielicki, 2003: Confronting models with data: the GEWEX cloud systems study. *Bull. Am. Meteor. Soc.*, **84**, 455–469.

Randall, D., M. Khairoutdinov, A. Arakawa, and W. Grabowski, 2003: Breaking the Cloud Parameterization Deadlock. *Bull. Am. Meteor. Soc.*, **84**, 1547-1564.

Randall, D.A., R.A. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, R.J. Stouffer, A. Sumi and K.E. Taylor, 2007: Climate Models and Their Evaluation. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Rasch P., M. Stevens, L. Ricciardulli, A. Dai, R. Wood, B. Boville, B. Eaton, and J. Hack, 2006: A characterization of tropical transient activity in the CAM3 atmospheric hydrologic cycle. *J. Climate*, **19**, 2222–2242.

Redelsperger, J. L., D. B. Parsons, and F. Guichard, 2002: Recovery Processes and Factors Limiting Cloud-Top Height following the Arrival of a Dry Intrusion Observed during TOGA COARE. *J. Atmos. Sci.,* **59**, 2438-2457.

Rodts, S. M. A., P. G. Duynkerke, H. J. J. Jonker, 2003: Size distribution and dynamical properties of shallow cumulus clouds from aircraft observations and satellite data.*J. Atmos. Sci.,* **60**, 1895-1912.

Rossow, W. B., and R. A. Schiffer, 1999: Advances in understanding ISCCP. *Bull. Am. Meteorol. Soc.*, **80**, 2261– 2287.

Rossow, W.B., G. Tselioudis, A. Polak, and C. Jakob 2005: Tropical climate described as a distribution of weather states indicated by distinct mesoscale cloud property mixtures. *Geophys. Res. Lett.* **32**, L21812, doi:10.1029/2005GL024584.

Sachs, D., S. Lovejoy, and D. Schertzer, 2002: The multifractal scaling of cloud radiances from 1m to 1km, *Fractals*, **10**(3), 1-12.

Sanchez, R., O. Zuchuat, F. Malvagi and I. Zmijarevic, Symmetry and translations in multimaterial line statistics, *J. Quant. Spectrosc. Radiat. Transfer*, **51**, 801-812, 1993.

Seifert, P., Ansmann, A., Muller, D., Wandinger, U., Althausen, D., Heymsfield, A. J., Massie, S. T., and Schmitt C.: Cirrus optical properties observed with lidar, radiosonde, and satellite over the tropical Indian Ocean during the aerosol-polluted northeast and clean maritime southwest monsoon, *J. Geophys. Res.*, 112, D17205, doi:10.1029/2006JD008352, 2007.

Senior, C.A., 1999: Comparison of mechanisms of cloud-climate feedbacks in GCMs. *J. Clim.*, **12**, 1480-1489.

Stephens, G. L., P. M. Gabriel, and S.-C. Tsay, 1991: Statistical radiative transfer in one-dimensional media and its application to the terrestrial atmosphere, *Trans. Th. Stat. Phys.*, **20**, 139-175.

Stokes, G.M., and S.E. Schwartz, 1994: The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the Cloud and Radiation Testbed. *Bull. Amer. Meteor. Soc.*, **75**, 1201-1221.

Turner, D.D., S.A. Clough, J.C. Liljegren, E.E. Clothiaux, K. Cady-Pereira, and K.L. Gaustad, 2007: Retrieving liquid water path and precipitable water vapor from Atmospheric Radiation Measurement (ARM) microwave radiometers. *IEEE Trans. Geosci. Remote Sens.*, **45**, 3680-3690.

Varnai, T., and R. Davies, 1999: Effects of cloud heterogeneities on shortwave radiation: comparison of cloud-top variability and internal heterogeneity. *J. Atmos. Sci.*, **56**, 4206-4224.

Veron, D. E., J. Secora, and M. J. Foster, 2008: Assessing cloud field characteristics for radiative transfer at the ARM CART Sites, *J. Climate*, in preparation.

von Savigny, C., A. B. Davis, O. Funk and K. Pfeilsticker, 2002: Time-series of zenith radiance and surface flux under cloudy skies: Radiative smoothing, optical thickness retrievals and large-scale stationarity. *Geophys. Res. Lett.*, **29**(17), 1825-1831, doi:10.1029/2001GL014153.

Webb, M., C. Senior, S. Bony, and J.J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models, *Climate Dynamics*, **17**, 905 – 922.

Welch, R. M., and B. A. Wielicki, 1984: Stratocumulus cloud field reflected fluxes: The effect of cloud shape. *J. Atmos. Sci.*, **41**, 3085-3103.

Wielicki, B. A. and R. J. Welch, 1986: Cumulus cloud properties derived using Landsat satellite data.*J. Appl. Meteorol.*, **25**, 261-276.

Williams KD, Senior CA, Slingo A, Mitchell JFB, 2005: Towards evaluating cloud response to climate change using clustering technique identification of cloud regimes, *Climate Dynamics*, **24**, 701–719.

Wiscombe, W. J., and J. W. Evans, 1977: Exponential-sum fitting of radiative transmission functions, *J. Comp. Phys.,* **24**, 416-444.

Zhao, G., and L. Di Girolamo, 2007: Statistics on the macrophysical properties of trade wind cumuli over the tropical western Atlantic. *J. Geophys. Res.*, **112**(D10204), doi:10.1029/2006JD007371.

Zuev, V. E., and G. A. Titov, 1995: Radiative transfer in cloud fields with random geometry. *J. Atmos. Sci.*, **52**, 176-190.

Curriculum Vita

Michael J. Foster

1994 – 1998 B.A., College of Arts and Sciences
Majors in Mathematics and Political Science
Boston College
Chestnut Hill, Massachusetts

1998 – 2003 Co-founder of and IT Systems Consultant for *NetTech Solutions*, a firm dedicated to the design and implementation of various local-area and wide-area network architectures.

2003 – 2006 M.S., Environmental Sciences Department
Rutgers, The State University of New Jersey
Graduate School-New Brunswick, New Jersey

2006 – 2008 Ph.D., Atmospheric Sciences Program
Rutgers, The State University of New Jersey
Graduate School-New Brunswick, New Jersey

Publications

Foster, M. J. and D. E. Veron, 2008: Evaluating the Stochastic approach to Shortwave Radiative Transfer in the Tropical Western Pacific. *J. Geophys. Res.,* accepted

Veron, D. E., C. P. Weaver, F. Veron, and M. J. Foster, 2008: Stochastic Radiative Transfer on Modeled Cloud Fields. *IEEE Geoscience and Remote Sensing Letters*, accepted