Florida State University Libraries

Electronic Theses, Treatises and Dissertations

The Graduate School

2013

Feng Yun-3B Microwave Humidity Sounder (MWHs) Data Noise Characterization and Filtering Using Principle Component Analysis

Yuan Ma

THE FLORIDA STATE UNIVERSITY
COLLEGE OF ARTS AND SCIENCES

FENGYUN-3B MICROWAVE HUMIDITY SOUNDER (MWHS) DATA NOISE
CHARACTERIZATION AND FILTERING USING PRINCIPLE COMPONENT ANALYSIS

By
YUAN MA

A Thesis submitted to the
Department of Earth, Oceanic and Atmospheric Sciences
in partial fulfillment of the
requirements for the degree of
Master of Science

Degree Awarded:
Spring Semester, 2013

Yuan Ma defended this thesis on June 26, 2012.

The members of the supervisory committee are:

Xiaolei Zou
Professor Directing Thesis

Guosheng Liu
Committee Member

Peter S. Ray
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the thesis has been approved in accordance with university requirements.

I humbly dedicate this manuscript to my parents, Jin Ma and Jianping Yu.

ACKNOWLEDGEMENTS

First and foremost, I want to thank Dr. Xiaolei Zou for her guidance, advice and help. She has provided me much more than I ever could have expected from an advisor and I owe her great deal of gratitude. I also want to thank Dr. Guosheng Liu and Dr. Peter S. Ray for serving as my committee members.

The work would have been impossible without the generous help from Dr. Zhengkun Qin, Mr. Xiang Wang and the rest of Prof. Zou's Data Assimilation lab, who have made my stay here incredible.

Last but not least important, I really appreciate encouragements from my family, especially my parents. You have made me who I am and your love is what keeps me fighting.

Funding was provided by Chinese Ministry of Science and Technology project 2010CB951600, and Chinese Ministry of Finance project GYHY200906006.

TABLE OF CONTENTS

List of Tables	vi
List of Figures	vii
Abstract	ix
1. INTRODUCTION	1
1.1 Background of FY3B MWHS	1
1.2 Researches of Microwave Humidity Sensors	2
1.3 Use of PCA in Remote Sensing Data	3
1.4 Thesis Organization	3
2. DATA	5
2.1 MWHS Data.....	5
2.2 GFS Data.....	9
2.3 MSPPS Data.....	9
3. METHODOLOGY	11
3.1 The Community Radiative Transfer Model	11
3.2 The PCA Approach.....	11
4. RESULTS	14
4.1 Biases and Standard Deviations.....	14
4.2 Characterization of MWHS Sensor Noise Using PCA.....	20
4.3 Elimination of Line-shape Noise in MWHS Data	26
5. CONCLUSION	39
5.1 Summary	39
5.2 Future Work	39
REFERENCES	40
BIOGRAPHICAL SKETCH	42

LIST OF TABLES

1.1	Platforms carrying microwave humidity sounders onboard.....	1
2.1	FY-3B MWHS instrument characteristics.	6

LIST OF FIGURES

1.1 Local Equatorial Crossing Time (LCT) of different NOAA satellites. (http://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh_avhrr_ect.php)	2
2.1 Weighting function calculated using US standard atmospheric profile for FY-3B MWHS (solid) and NOAA-18 MHS (dashed).	7
2.2 A single scanline of MWHS (green and orange) in ascending node (FY-3B travels along the track as indicated in blue) on 0417UTC, April 6, 2011.	8
4.1 (a)-(b) Observed, (c)-(d) simulated and (e)-(f) O-B differences of brightness temperatures at channel 3 of MWHS (left panels) and MHS (right panels) on 0300-1500 UTC April 2, 2011. ..	15
4.2 Global distribution of the total number of observations within each $1^{\circ} \times 1^{\circ}$ grid box during April 2-30, 2011 for (a) FY-3B MWHS and (b) NOAA-18 MHS.	16
4.3 Global bias (left panel) and standard deviation (right panel) of brightness temperature differences between observations and NCEP GFS during April 2-30, 2011 at nadir for channel 3 (red), channel 4 (blue) and channel 5 (green).	17
4.4 Monthly-mean scan-angle dependence of MWHS sounding (a)-(b) channel 3 (red), (c)-(d) channel 4 (blue) and (e)-(f) channel 5 (green) from FY-3B MWHS (left panels) and NOAA-18 MHS (right panels) observations (solid) and model simulations (dashed).....	18
4.5 Scan-angle dependence of O-B biases for (a) MWHS and (b) MHS sounding channel 3 (red), channel 4 (blue) and channel 5 (green). Nadir biases are subtracted.	19
4.6 Globally-averaged, monthly-mean O-B power spectral density (solid) and 95% confidence level (dashed) for MWHS channel 3 (red), channel 4 (blue) and channel 5 (green) during April 2011.	20
4.7 The explained variances by each of the 98 PCs (dashed) of MWHS channel 3 (red), channel 4 (blue) and channel 5 (green). The accumulated explained variances are shown as solid line and the y-axis on the right. A single swath data during 0114-0255 UTC April 2, 2011 is used in the calculation.	21
4.8 The PC coefficients (left panels) and the matrix of the vector products of the PC coefficients with PCs (right panels) for (a)-(b) the 1 st , (c)-(d) the 2 nd , and (e)-(f) the 3 rd modes calculated for MWHS channel 3 from the same swath data as Figure 4.7.	22

4.9 Distributions of power spectral density of each PC exceeding 95% confidence level for different recognized periods for MWHS (a)-(b) channel 3, (c)-(d) channel 4 and (e)-(f) channel 5 for all 98 PCs (left panels) or only the first 10 PCs (right panels) calculated for the same swath as in Figure 4.7 .	24
4.10 Spectra of the first five PCs for MWHS channels 3-5 extracted by Fourier analysis using the same swath as in Figure 4.7.	25
4.11 The O-B variations averaged over swath as in Figure 4.7 with respect to scan position of the MWHS data (black) and the PCA-reconstructed data (colored) for (a) channel 3, (b) channel 4 and (c) channel 5 observations or the PCA-reconstructed “observations” (black curve).	26
4.12 Variations of the standard deviation of the differences between MWHS observations and the PCA reconstructed fields with the number of PC modes employed for the data reconstruction for (a) channels 3, (b) channel 4, and (c) channel 5.	28
4.13 Differences between MWHS observations and the PC- reconstructed fields for (a) channel 3, (b) channel 4 and (c) channel 5 on 0114-0255 UTC, April 2, 2011	29
4.14 Same as Figure 4.13 but for an enlarged region.	30
4.15 (a) Cloud liquid water path (LWP) retrieved from surface-sensitive channels of AMSU-A on board NOAA-18 during 0000-1200 UTC April 2, 2011. (b) Cloudy data points of FY-3B MWHS (cyan) identified by their collocation with cloudy points in (a) under the criteria that $CLW > 0.01 \text{ kg/m}^2$ and spatial separation being less than 30 km.	31
4.16 (a) Scan-angle dependence of O-B bias in cloudy conditions during the entire month of April 2011 for channels 3 (red), channel 4 (blue) and channel 5 (green). (b) Power spectral density (solid) and 95% confidence level (dashed) for O-B bias in (a)	32
4.17 Same as Figure 4.16 but for clear-sky conditions.	33
4.18 Differences between MWHS observations and the five point smoothed fields for (a) channel 3, (b) channel 4 and (c) channel 5 on 0114-0255 UTC, April 2, 2011. (d) Same as (c) except for the entire swath.	35
4.19 Power spectral density (color shaded) obtained for data along all the complete scanlines in (a) Fig 4.14c and (b) Fig 4.18c. The contour in (a)-(b) indicates the 95 % confidence level.	37
4.20 Same as Figure 4.5 except for PCA-reconstructed fields.	38

ABSTRACT

MicroWave Humidity Sounder (MWHS) onboard both FY-3A and FY-3B satellites have three channels (channels 3-5), near the 183 GHz water vapor absorption line. These channel frequencies are also used in other instruments such as Advanced Microwave Sounding Unit-B (AMSU-B) and Microwave Humidity Sounder (MHS) onboard MetOp and NOAA satellites. Both MWHS and MHS are cross-track scanners. In this study a comparison between the simulated brightness temperatures with MWHS measurements clearly shows that MWHS observations from the three sounding channels contain a scan angle dependent **coherent noise along the instrument scanline. This noise does not** cancel out when a large amount of data over a sufficiently long period of time is averaged. The noise is found around 0.3, 0.2, and 0.2 K for channels 3-5, respectively. A principle component analysis is used for the characterization of this **cohesive noise** using one-month of FY-3B MWHS data. It is shown that the MWHS cohesive noise is contained primarily in the first PC mode, which mainly describes a scan angle dependent brightness temperature variation, i.e., a unique feature of cross-tracking instrument. The 1st PC accounts for more than 99.91% of the total variance in the three MWHS sounding channels. A five-point smoother is then applied to the first PC, which effectively removes such a data noise in MWHS data. The reconstruction of the MWHS radiance spectra using the noise-filtered first PC component is essentially free of noise. The removal of the scan angle dependent bias produces a reconstructed MWHS data set that is more uniform and is consistent with NOAA-18 MHS data.

CHAPTER 1

INTRODUCTION

1.1 Background of FY-3B MWHS

The first satellite of the second series of Chinese polar orbiting satellites, Fengyun-3A (FY-3A), was successfully launched into orbit on May 27, 2008. The performance of the atmospheric sounding instruments, in particular, met the specifications (Dong et al. 2009 and Zhang et al. 2009). Two years later, the second satellite in the new series, FY-3B, was successfully launched on November 5, 2010.

The mission aims to enhance NWP and climate research with observation data; monitor large-scale natural disasters, the ecological environment, and provide meteorological information for aviation and navigation.

Both of FY-3A and FY-3B carry 11 sensors onboard: VIRR, IRAS, MWTS, MWHS, MERSI, MWRI, ERM, SIM, SBUS, TOU and SEM. FY-3B is the latest satellite to have a microwave humidity sounder on aboard, as shown in Table 1.1. The observation data for all sensors listed in the table is available till present.

Table 1.1: Satellites carrying microwave humidity sounders.

Satellite	Launch Date	Microwave Humidity Sounder
NOAA-15	13 May 1998	AMSU-B
NOAA-16	21 September 2000	
NOAA-18	20 May 2005	MHS
MetOp-A	19 October 2006	
NOAA-19	6 February 2009	
FY-3A	27 May 2007	MWHS
FY-3B	5 November 2010	

Figure 1.1 indicates the most updated Local Equatorial Crossing Time (LCT) of different NOAA satellites. LCTs are not constant since there are always differences

between pre-calculated and post-launched orbits for satellites. The LCTs of NOAA satellites, plus FY3A's 1015 am (Lu, 2011) and FY3B's 0130 pm, make it possible to observe the atmosphere several times during a single day at any location.

Figure 1.1: Local Equatorial Crossing Time (LCT) of different NOAA satellites.
(http://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh_avhrr_ect.php)

1.2 Researches of Microwave Humidity Sensors

Because of their capability to see through thin clouds, microwave humidity sounder data has long been used worldwide to generate products for weather prediction, climate forecasts, and hydrological studies.

To better use the observation data and assimilate it into models, it is necessary to carefully calibrate, monitor and validate the microwave humidity sensors. Atkinson (2002) pointed out the problem of radio-frequency interference for NOAA-15 and AMSU-B. The comparison of observations between different sensors are often used to verify the data from some certain source. MHS onboard NOAA-19 and MetOp-A are verified by Bonsignori (2007), and the data acquired are regarded to be of better quality than their heritage AMSU-B. FY-3A MWHS observations of the whole month of January, 2010 are assessed by comparing them to the data from NOAA-18 MHS, with the help of a

radiative transfer model (RTM), and a Quality Control procedure, which made it possible to identify outliers and non-outliers (Guan et al. , 2011).

Satellite radiance's impact on NWP is highlighted since this data is the largest volume of input. Bormann (2010) presented methods of estimating observation errors for MHS ,when assimilating the data into European Centre for Medium-Range Weather Forecasts (ECMWF) system.

As atmospheric water vapor is an important green-house gas, the humidity observations have been employed to enhance climate studies. Since the middle and upper troposphere plays an important roles in climatic water vapor feedback, Moradi et al. (2010) compared the upper tropospheric humidity retrieved from NOAA-15, NOAA-16 AMSU-B as well NOAA-18, NOAA-19, MetOp-A MHS against radiosonde data.

The Microwave Surface and Precipitation Products System (MSPPS) provides the community with hydrological products retrieved from AMSU-A and AMSU-B (Ferraro et al. 2005), and MHS afterwards, which helps to promote the research on hydrology.

1.3 Use of PCA in Remote Sensing Data

In this study, we present a comparison study of FY-3B MWHS data with brightness temperatures simulated by a radiative transfer model using NWP fields as its input. A FOV-fixed, line-pattern cohesive noise is present in MWHS data. A principle component analysis (PCA) is employed for filtering the noise in MWHS data without altering other weather-related features in the MWHS observations. Using PCA to characterize sensor noise in the Earth scene data is not new. Antonelli et al. (2004) applied a PCA-based noise filter, which eliminates the higher-order principal components to reduce the random noise present in the simulated and real hyperspectral infrared observations. Tobin et al. (2007) showed that PCA is a powerful technique for diagnosing and filtering Atmospheric Infrared Sounder (AIRS) noise and other variable artifacts in hyperspectral data. Tobin et al. (2009) also used the PCA approach to remove the subtle distortions in the AIRS instrument line shape introduced by non-uniform scene effects.

1.4 Thesis Organization

This thesis is organized as follows. In section 2, the description of all data used in the work is presented. Model and numerical results, including global biases, standard deviations and scan-angle dependence of biases; the line-pattern noise in MWHS data, the PCA approach for characterizing it, a five-point smoother for filtering it, and the reconstructed MWHS data obtained by applying a PCA-based noise filtering to real MWHS observations are illustrated in section 4. Section 5 provides a brief summary.

CHAPTER 2

DATA

2.1 MWHS Data

Radiance data in Level-1B format from both FY-3B MWHS and NOAA-18 MHS during April 2011 are employed for this study. FY-3B MWHS data is obtained directly from Chinese Satellite Meteorological Center and NOAA-18 MHS data is downloaded at http://www.nsof.class.noaa.gov/saa/products/search?sub_id=0&datatype_family=TOVS&submit.x=29&submit.y=8

Both FY-3B and NOAA-18 have afternoon-configured orbits and the same equator-crossing local time (2:00pm). Major sensor characteristics are summarized in Table 1.

Table 2.1: FY-3B MWHS instrument characteristics.

Channel number		Frequency (GHz)		Bandwidth (MHz)		NE Δ T (K)	
MHS	MWHS	MHS	MWHS	MHS	MWHS	MHS	MWHS
	1	89(V)	150(V)	1000 $\times$ 2		0.23	0.90
	2	150 (V)	150(H)	1000 $\times$ 2		0.37	0.90
	3	183.31 $\pm$ 1 (H)	183.31 $\pm$ 1 (V)	500 $\times$ 2		0.55	1.10
	4	183.31 $\pm$ 3 (H)	183.31 $\pm$ 3 (V)	1000 $\times$ 2		0.42	0.90
	5	190.31 (V)	183.31 $\pm$ 7 (V)	2000 $\times$ 2		0.35	0.90
Channel number		Nadir Res. (km)		WF (hPa)		Swath width (km)	
MHS	MWHS	MHS	MWHS	MHS	MWHS	MHS	MWHS
	1		15		surface	2250	2700
	2		15		surface	2250	2700
	3		15		400	2250	2700
	4		15		600	2250	2700
	5		15		800	2250	2700

The MWHS and MHS channels 1-2 are window channels. The center frequency of MWHS channel 1 is 150 GHz (vertical polarization), which is different from the MHS channel 1 at 89 GHz (vertical polarization). The center frequencies of MWHS channels 2-5 are the same as MHS channels 2-5, located at 150 GHz (horizontal polarization), 183.31 $\pm$ 1 GHz (vertical polarization), 183.31 $\pm$ 3 GHz (vertical polarization), and 183.31 $\pm$ 7 GHz (vertical polarization), respectively. The three sounding channels 3-5 have their weighting function peaks located near 400hPa, 600hPa and 800hPa, respectively, and weighting functions of channel 1 and channel 2 of MWHS are very closed, as shown in Figure 2.1.

Figure 2.1: Weighting function calculated using US standard atmospheric profile for FY-3B MWHS (solid) and NOAA-18 MHS (dashed).

MWHS provides a total of 98 field-of-views (FOVs) along each scanline in $8/3$ seconds, while MHS observes 90 FOVs on each scanline in $8/3$ seconds. A sketch of how FY-3B MWHS takes its scans is shown. When on ascending node of FY-3B, MWHS scan from the first scan position at the right (152.72 E, 0.10S) to the 98th position at the left (129.54E, 3.55S), which makes a single scanline. The subsequent scanline would be toward the North and still from right to left.

Figure 2.2: A single scanline of MWHS (green and orange) in ascending node (FY-3B travels along the track as indicated in blue) on 0417UTC, April 6, 2011

The observation resolution at nadir is 15 km for both MWHS and MHS. The MWHS swath width is 2700 km, which is wider than the swath width of MHS (e.g., 2250 km), since MWHS has 98 scan positions while MHS 90. The Noise Equivalent Delta Temperature (NEDT) values are 1.10, 0.90, and 0.90 K for MWHS channel 3-5, which are comparable to the NEDT values of 0.55, 0.42 and 0.35 K for MHS (Boukabara et al. 2010).

Data comparisons between MWHS and MHS are made only for channels 3-5. Measurements from these three sounding channels are sensitive to temperature, water vapor and cloud information in the troposphere (Weng et al., 2003). NOAA-18 MHS data are routinely ingested and assimilated at major NWP operational centers around the world. NOAA, ECMWF and China are planning to incorporate FY-3A/B MWHS data into operational forecasts. Comparisons between MWHS channel 1 and MHS channel 1 are not possible because of their frequency differences. Comparisons for MWHS channel 2 with MHS channel 2 will be made in a future study since the comparison results need to be interpreted and analyzed with many other auxiliary data (e.g., surface parameters) and

more advanced radiative transfer capability.

2.2 GFS Data

GFS is short for National Center for Environmental Prediction (NCEP)'s Global Forecast Systems 6-hour forecast fields. It is available at <http://nomads.ncdc.noaa.gov/data/gfsanl/>, the temporal coverage is from March 2004 to present. The NCEP GFS 6-hour forecast fields have a horizontal resolution 1°x1° and 26 vertical levels at 10.0, 20.0, 30.0, 50.0, 70.0, 100.0, 150.0, 200.0, 250.0, 300.0, 350.0, 400.0, 450.0, 500.0, 550.0, 600.0, 650.0, 700.0, 750.0, 800.0, 850.0, 900.0, 925.0, 950.0, 975.0, 1000.0 hPa.

The variables included in GFS fields are: pressure, geopotential height, total ozone, temperature, potential temperature, geopotential height anomaly, u-component of wind, v-component of wind, pressure vertical velocity, absolute vorticity, specific humidity, relative humidity, precipitable water, water equivalent accumulative snow depth, total cloud cover, cloud water, land-sea mask, ice concentration, surface lifted index, vertical speed shear, volumetric soil moisture content, ozone mixing ratio, convective inhibition, potential energy, planetary boundary layer height, 5-wave geopotential height, etc. GFS data of the entire month of April 2011 is used to simulate brightness temperature.

2.3 MSPPS Data

Microwave Surface and Precipitation Products System (MSPPS) is available on http://www.nsof.class.noaa.gov/saa/products/search?sub_id=0&datatype_family=MSP_PS_ORB&submit.x=19&submit. The products MSPPS provides including Total Precipitable Water, Cloud Liquid Water, Ice Water Path, etc, among which Cloud Liquid Water retrieved from surface channels of NOAA-18 AMSU- during April 2011 are employed to identify cloudy and clear-sky data points. The data points for Cloud Liquid Water are consistent with NOAA-18 AMSU-A observations. There are 30 beam positions in a single scanline, which takes 8 seconds to collect. The resolution at nadir is around 60

km and swath width is 2300 km. Cloud Liquid Water is the cumulative water content in units of kg/m^2 , and the value should be divided by 100.

CHAPTER 3

METHODOLOGY

3.1 The Community Radiative Transfer Model

The Community Radiative Transfer Model (CRTM) was developed by the US Joint Center for Satellite Data Assimilation (JCSDA) for rapid calculations of satellite radiances and their derivatives under various atmospheric and surface conditions (F. Weng 2007). The model (Y. Han et al. 2007) is used for producing global simulations of brightness temperatures that are measured by either MWHS or MHS from GFS 6-hour forecast fields.

The input required by the CRTM includes: number of profile, latitude, longitude, scan angle, zenith angle; surface type, surface temperature, 10 m wind speed U, 10 m wind speed V; vertical profile of pressure, temperature and mixing ratio. In our case, the vertical atmospheric profiles and surface parameters are from GFS data, the geometry is extracted from MWHS or MHS data. The output of CRTM includes brightness temperature, optical length, etc. CRTM is employed in the study to simulate brightness temperature as ‘truth’ in order to calculate the bias of observation.

3.2 The PCA Approach

The principal component analysis (PCA) is carried out on a swath-by-swath basis. The data matrix (A) is constructed from MWHS brightness temperature observations:

$$A_{98 \times N} = \begin{pmatrix} TB_{1,1} & \cdots & TB_{1,N} \\ \vdots & \ddots & \vdots \\ TB_{98,1} & \cdots & TB_{98,N} \end{pmatrix} \quad (1)$$

Where $TB(k, j)$ ($k=1, 2, \dots, 98, j=1, 2, \dots, N$) indicates brightness temperature at the k^{th}

FOV and the j^{th} scanline of a swath, N is the total number of scanlines involved in the calculation.

The covariance matrix is then constructed from the data matrix A:

$$\mathbf{S} = \mathbf{A}\mathbf{A}^T,$$

whose eigenvalues and eigenvectors are then calculated:

$$\mathbf{S}\bar{e}_i = \lambda_i \bar{e}_i \quad (\lambda_i, i=1, \dots, 98) \quad (2)$$

where

$$\bar{e}_i = \begin{pmatrix} e_{1,i} \\ e_{2,i} \\ \vdots \\ e_{98,i} \end{pmatrix}$$

is called the principal component for the i^{th} mode, the i^{th} eigenvalue represents the contribution of the i^{th} mode to the total variance of data in matrix A. Equation (2) can be written in matrix form:

$$\mathbf{S}\mathbf{E} = \mathbf{E}\mathbf{\Lambda} \quad (3)$$

where

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_{98} \end{pmatrix}, \quad \mathbf{E} = (\bar{e}_1 \quad \bar{e}_2 \quad \dots \quad \bar{e}_{98})$$

$$\mathbf{E}^{-1} = \mathbf{E}^T$$

$$\mathbf{S} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^T \quad (4)$$

Finally, the PC coefficient matrix (U) is calculated:

$$\mathbf{U} = \mathbf{E}^T \mathbf{A} = \begin{pmatrix} u_{1,1} & u_{1,2} & \dots & u_{1,N} \\ u_{2,1} & u_{2,2} & \dots & u_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ u_{98,1} & u_{98,2} & \dots & u_{98,N} \end{pmatrix} = \begin{pmatrix} \bar{u}_1 \\ \bar{u}_2 \\ \vdots \\ \bar{u}_{98} \end{pmatrix} \quad (5)$$

where u-vector is the so-called PC coefficients for the i^{th} mode.

The correction algorithm simply involves the reconstruction of the spectra with the first “noisy” principal component being smoothed, i.e.,

$$\mathbf{A}^{reconstructed} = \overleftarrow{e_1^{sm}} \overleftarrow{u_1} + \sum_{i=2}^{98} \overleftarrow{e_i} \overleftarrow{u_i}, \overleftarrow{e_1^{sm}} : \text{five-point smoothed } \overleftarrow{e_1} \quad (6)$$

PCA is used to identify contribution from each PC of the observations and reconstruct the observed fields.

CHAPTER 4

RESULTS

4.1 Biases and Standard Deviations

Figure 4.1 presents a 12-h global distribution of MWHS and MHS channel 3 observations (Figure 4.1a-b), model simulations (Figure 4.1c-d), and differences between observations and simulations (Figure 4.1e-f) of brightness temperatures during 0300-1500 UTC April 2, 2011.

The overall distributions of brightness temperatures provided by MWHS are very similar to those of MHS. Brightness temperatures from MWHS have a larger distribution of values. And MWHS has a wider swath.

Figure 4.1: (a)-(b) Observed, (c)-(d) simulated and (e)-(f) O-B differences of brightness temperatures at channel 3 of MWHS (left panels) and MHS (right panels) on 0300-1500 UTC April 2, 2011.

The distributions of the numbers of observations for FY-3B MWHS and NOAA-18 MHS during April 2-30 are as shown in Figure 4.2.

Figure 4.2: Global distribution of the total number of observations within each $1^{\circ} \times 1^{\circ}$ grid box during April 2-30, 2011 for (a) FY-3B MWHS and (b) NOAA-18 MHS.

Compared with mid-latitude, more observations are within high latitudes of 70S-80S and 70N-80N while near-polar regions have much fewer observations. This fact is consistent with the spatial coverage of polar-orbit satellites. A noticeable difference between MWHS and MHS is MHS observations are more evenly distributed 60S-60N.

The monthly mean global nadir biases and the standard deviations of brightness

temperature differences between observations and model simulations during April 2011 for FY-3B MWHS and NOAA-18 MHS data are provided in Figure 4.3. The biases for MWHS data are comparable to those of MHS, with channel 3 having the largest global bias (between 3-6K). The standard deviations of the MWHS data are consistently larger than the MHS data, suggesting a larger variability in the MWHS data. The standard deviations increase with channel number.

Figure 4.3: Global bias (upper panel) and standard deviation (lower panel) of brightness temperature differences between observations and NCEP GFS during April 2-30, 2011 at nadir for channel 3 (red), channel 4 (blue) and channel 5 (green).

The scan-angle dependences of the MWHS and the MHS biases are shown in Figure 4.4a and 4.4b, respectively. The scan-angle dependence of the MWHS and the MHS biases are asymmetric for all sounding channels 3-5. A data noise characterized by a scan angle dependence, high-frequency oscillatory signal, occurred in the MWHS observations (Figure 4.4 left panels), resulting a similar high-frequency oscillatory signal noise in the O-B biases for all three MWHS sounding channels (Figure 4.4a). The O-B

biases of the MHS data (Figure 4.4b) are much smoother than those of MWHS data.

Fig. 4.4: Monthly-mean scan-angle dependence of MWHS sounding (a)-(b) channel 3 (red), (c)-(d) channel 4 (blue) and (e)-(f) channel 5 (green) from FY-3B MWHS (left panels) and NOAA-18 MHS (right panels) observations (solid) and model simulations (dashed).

Figure 4.5: Scan-angle dependence of O-B biases for (a) MWHS and (b) MHS sounding channel 3 (red), channel 4 (blue) and channel 5 (green). Nadir biases are subtracted.

To find out which frequencies the O-B variations are strong, the power spectral density is calculated for all swaths and then averaged to obtain a global-average, monthly-mean O-B power spectral densities for the MWHS data in April 2011. Results

are shown in Figure 4.6.

Figure 4.6: Globally-averaged, monthly-mean O-B power spectral density (solid) and 95% confidence level (dashed) for the MWHs channel 3 (red), channel 4 (blue) and channel 5 (green) during April 2011.

The 95% confidence levels for all three MWHs channels are also indicated. The power spectral densities in Figure 4.6 provide a quantitative measure of the overall frequency content of the MWHs data along different scanlines. It is seen that the strength of data variation weakens as the frequency increases except for a strong high-frequency oscillatory noise signal with a periodicity of 2.6 FOVs in all three sounding channels. The strength of the signal noise is stronger for the lower tropospheric channel 5 than the upper-level channels 3 and 4.

4.2 Characterization of the MWHs Sensor Noise Using PCA

The PCA approach is employed to convert a set of observations of possibly correlated scan angle dependent brightness temperatures into a set of uncorrelated PCs, each revealing the internal structure of the MWHs data in a way that the 1st PC accounts for as much of the variability in the data as possible and explains most variance in the data, and each succeeding component in turn has the highest variance possible under the

constraint that it be orthogonal to (uncorrelated with) the preceding PC components. Figure 4.7 presents the explained variances by each of the 98 PCs of MWHS channels 3-5, as well as the accumulated explained variances for a single swath data at 0100 UTC April 2, 2011. It is seen that the 1st PC explains more than 99.91% of the total variance in the MWHS data. The first 10 PCs explain more than 99.99% of the total variance.

Figure 4.7: The explained variances by each of the 98 PCs (dashed) of MWHS channel 3 (red), channel 4 (blue) and channel 5 (green). The accumulated explained variances are shown as solid line and the y-axis on the right. A single swath data at 0114-0255 UTC April 2, 2011 is used in the calculation.

As examples, we show in Figure 4.8 the PC coefficients (as in equation (5)) and the matrix of the vector products of the PC coefficients with PCs (as in equation (6)) for the first three PC modes calculated for MWHS channel 3 from a single swath data at 0100 UTC April 2, 2011.

Figure 4.8: The PC coefficients (left panels) and the matrix of the vector products of the PC coefficients with PCs (right panels) for (a)-(b) the 1st, (c)-(d) the 2nd, and (e)-(f) the 3rd modes calculated for MWHS channel 3 from the same swath data as Figure 4.7.

It is seen that the latitudinal variations of brightness temperature of an orbit going

from the South Pole to the North Pole and back again are captured by the PC coefficients of the 1st PC mode (Figure 4.8a). The scan angle dependence of the brightness temperature of a cross-track radiometer instrument, with the brightness temperatures being warmer at smaller scan angles than larger scan angles, is captured by the vector products of the PC coefficients with PCs (Figure 4.8b). The 2nd PC mode mainly describes the asymmetric variations with scan angle, and the 3rd PC seems to represent the symmetric part that are of different magnitudes in high latitudes and low and middle latitudes.

The individual contributions of the 98 PC modes to the total power spectral density are provided in Figure 4.9, which shows the power spectral densities of each PC exceeding 95% confidence level for different MWHS channel 3 periods for (Figure 4.9a-b), channel 4 (Figure 4.9c-d) and channel 5 (Figure 4.9e-f) for all 98 PCs (left panels) and the first 10 PCs (right panels).

Figure 4.9: Distributions of power spectral density of each PC exceeding 95% confidence level for different recognized periods for MWSH (a)-(b) channel 3, (c)-(d) channel 4 and (e)-(f) channel 5 for all 98 PCs (left panels) or only the first 10 PCs (right panels) calculated for the same swath as in Figure 4.7.

Results in Figure 4.9 are calculated for a single swath at 0100 UTC April 2, 2011. The following three features are noticed: (i) The larger the PC mode number, the smaller the period of large power spectral density; (ii) The high-frequency oscillatory signal noise with a periodicity of 2.6 FOVs seen in Figure 4.5a and 4.6 is contained in the 1st PC mode; and (iii) It requires more PC modes to capture most of the variations in channel 5 data than upper level modes.

A Fourier analysis is carried out for all the PCs. Results for the first five PCs are presented in Figure 4.10. For the first PC at all 3 channels, there are sharp increases of power spectrum near 2.5 FOVs. It further confirms that the cohesive (? coherent??) noise in the MWHs data is largely in the first PC.

Figure 4.10: Spectra of the first five PCs for the MWHs channels 3-5 extracted by Fourier analysis using the same swath as in Figure 4.7.

4.3 Elimination of Line-shape Noise in MWHS Data

Since the high-frequency noise is contained primarily in the first PC mode (see Figure 4.9 and 4.10), a five-point smooth is applied to the first PC. Then equation (6) is then used to reconstruct the MWHS data. The effectiveness of noise filtering is shown in Figure 4.11 for a single swath. It is found that the high-frequency oscillations in the O-B variations with respect to scan position are effectively removed in the PCA-reconstructed data for all three MWHS sounding channels.

Figure 4.11: The O-B variations averaged over the same swath as in Figure 4.7 with respect to scan position of the MWHS data (black) and the PCA-reconstructed data (colored) for (a) channel 3, (b) channel 4 and (c) channel 5 or the PCA-reconstructed “observations” (black curve).

The convergence of the PCA-reconstructed MWHS data is shown in Figure 4.12. The standard deviations of the differences between the MWHS observations and the PCA reconstructed fields

$$A_m^{reconstructed} = \overline{e_1^{sm}} \overline{u_1} + \sum_{i=2}^m \overline{e_i} \overline{u_i} \quad (7)$$

decrease rapidly with the number of PC modes employed for the data reconstruction. It can thus be concluded that the reconstruction of the radiance spectra with smoothing the 1st PC provides an accurate and robust algorithm for MWHS data noise filtering. The scan angle dependent bias from reconstructed MWHS data becomes smooth and compares favorably with NOAA-18 MHS data.

Figure 4.12: Variations of the standard deviation of the differences between the MWHs observations and the PCA reconstructed fields with the number of PC modes employed for the data reconstruction for (a) channel 3, (b) channel 4, and (c) channel 5.

Differences between the MWHS observations and the PCA-reconstructed fields for channels 3-5 show a line-shape noise (Figure 4.13, 4.14), implying that the MWHS data noise oscillation is fixed in phase with respect to the FOV. This will lead to high correlations between different scanlines, which explains why the MWHS data noise is contained in the 1st PC mode.

Figure 4.13: Differences between MWHS observations and the PC-reconstructed fields for (a) channel 3, (b) channel 4 and (c) channel 5 on 0114-0255 UTC, April 2, 2011.

Figure 4.14: Same as Figure 4.13 but for an enlarged region.

Having demonstrated that the noise indeed does not change over land versus over ocean, as well as over global versus over regional scales, we examine if the above-mentioned line-shape noise change under cloudy conditions. The cloudy MWHS data points are identified using the cloud liquid water path (LWP) retrieved from surface-sensitive channels of AMSU-A on board NOAA-18. A collocation criterion of less than 30 km spatial separation is imposed. An example of how this collocation is done is provided in Figure 4.15.

Figure 4.15: (a) Cloud liquid water path (LWP) retrieved from surface-sensitive channels of AMSU-A on board NOAA-18 during 0000-1200 UTC April 2, 2011. (b) Cloudy data points of FY-3B MWHS (cyan) identified by their collocation with cloudy points in (a) under the criteria that $CLW > 0.01 \text{ kg/m}^2$ and spatial separation being less than 30 km.

The cloud LWP distribution from NOAA-18 during 0000-1200 UTC April 2, 2011 (Figure 4.15a) is compared with the cloudy MWHS data point distribution (Figure 4.15b). It is seen that locations of cloudy data points found for FY-3B MWHS data are consistent with places where $CLW > 0.01 \text{ kg/m}^2$.

The O-B bias as a function of scan angle averaged for the MWHS data during the entire month of April 2011 in cloudy conditions is provided in Figure 4.16a. Similar to Figure 4.5a for all-weather conditions, a high-frequency oscillatory signal is found in the O-B bias in cloudy conditions. This implies that variations associated with cloud systems cancel out when a large amount of data over a one-month period of time is averaged. A power spectral density analysis of the results in Figure 4.16a is shown in Figure 4.16b.

Figure 4.16: (a) Scan-angle dependence of the O-B bias in cloudy conditions during the entire month of April 2011 for channel 3 (red), channel 4 (blue) and channel 5 (green). (b) Power spectral density (solid) and 95% confidence level (dashed) for the O-B bias as in (a).

Similar to Figure 4.16, Figure 4.17 indicates under clear conditions, the monthly mean O-B bias with respect to scan angle has the 2.5 FOVs high-frequency noise too, which further confirms that the line-shape noise found in the MWHs data is independent

of weather systems.

Figure 4.17: Same as Figure 4.16 but for clear-sky conditions.

As shown in Figure 4.11, the cohesive, periodic data noise is removed by applying a five-point smoothing to the 1st PCA, which explains more than 99.99% total variance. A natural question arises as whether a direct five-point smoothing to the MWHs data

without a PCA-decomposition would work. Figure 4.18 presents the differences between MWHS observations and the five point smoothed fields at the same time as in Figure 4.13.

Compared with Figure 4.13, a direct five-point smoothing to MWHS data modified the weather-related features at a magnitude that is an order of magnitude greater than the data noise seen in Figure 4.13.

Figure 4.18: Differences between MWSH observations and the five point smoothed fields for (a) channel 3, (b) channel 4 and (c) channel 5 on 0114-0255 UTC, April 2, 2011. (d) Same as (c) except for the entire swath.

A power spectral density analysis (Figure 4.19) indicates that the PCA method eliminates the signal with their wavelength centered around 2.5 FOVs, but a direct smoothing on data removes those features whose wavelength varies from 3 to 10 FOVs. The main features eliminated by the PCA-assisted smoothing have a fixed scale for all the scanlines, while those eliminated by a direct smoothing on data are different for different scanlines. This further demonstrates the effectiveness of the proposed method for the elimination of the line-shape noise in MWHS data while the true signal does not project onto the first PC and is left untouched by the PC smoothing.

Figure 4.19: Power spectral density (color shaded) obtained for data along all the complete scanlines in (a) Figure 4.14c and (b) Figure 4.18c. The contour in (a)-(b) indicates the 95% confidence level.

The scan-angle dependence of MWHs sounding channels 3-5 O-B biases calculated from PCA-reconstructed MWHs observations and model simulations are shown in Figure 4.20. Compared with Figure 4.5, the O-B bias calculated from reconstructed MWHs data varies smoothly with scan angle without changing the large-scale feature.

Figure 4.20: Same as Figure 4.5 except for PCA-reconstructed fields.

CHAPTER 5

CONCLUSIONS

5.1 Summary and Conclusions

Detecting, characterizing and removing observation noise in FY-3B MWHS data are extremely important before they could be effectively applied in NWP and climate studies. PCA is employed for characterizing the noise in the MWHS data, and a five-point filter is used for filtering the noise contained primarily within the first PC mode. The reconstruction of the brightness temperature spectra with the first term in the PCA reconstructed field smoothed provides an effective data filter. This PCA-based filter is not only accurate and robust, but also flexible in its implementation. It can be applied on a swath-by-swath basis or granule-by-granule basis.

5.2 Future Work

The root-cause of the high-frequency noise is not clear and still in need of further investigations. The work will also be extended to surface channels, i.e., channel 1 and channel 2 of FY-3B MWHS, for which surface emissivity must be considered. FY-3A MWHS data will be checked to see if similar noise is contained in observation and could be removed by PCA reconstruction.

REFERENCES

- Antonelli, P., H. E. Revercomb, L. A. Sromovsky, W. L. Smith, R. O. Knuteson, D. C. Tobin, R. K. Garcia, H. B. Howell, K.-L. Huang, and F. A. Best, "A principal component noise filter for high spectral resolution infrared measurements." *J. Geophys. Res.* (2004), vol. 109, D23102, doi: 10.1029/2004JD004862.
- Atkinson, N. C, "Calibration, monitoring and validation of AMSU-B", *Advances in Space Research* (October 2001), vol.28, no.1, pp.117-126.
- Bonsignori R. (2007), "The Microwave Humidity Sounder (MHS): in-orbit performance", *Proc. SPIE*, 6744, 67440A
- Bormann, N. and P. Bauer, "Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to ATOVS data", *Q. J. R. Meteorol. Soc.* (April 2010), vol.136, pp.1036-1050.
- Boukabara, S. A., K. Garrett and W. Chen, "Global coverage of total precipitable water using a microwave variational algorithm", *Geoscience and Remote Sensing, IEEE Transactions on* (October 2010), vol.48, no.10, pp.3608-3621.
- Buehler, S. A., M. Kuvatov, and V. O. John, "Scan asymmetries in AMSU-B data." *Geophys. Res. Lett.* (2005), vol. 32, L24810, doi:10.1029/2005GL024747.
- Dong, C., J. Yang, W. Zhang, Z. Yang, N. Lu, J. Shi, P. Zhang, Y. Liu, and B. Cai , "An overview of a new Chinese weather satellite FY-3A." *Bulletin of the American Meteorological Society* (2009), vol. 90, no. 10, pp. 1531–1544.
- Ferraro, R.R., F. Weng, N.C. Grody., L. Zhao, H. Meng; C. Kongoli, P. Pellegrino, S. Qiu and C. Dean (May 2005), "NOAA operational hydrological products derived from the advanced microwave sounding unit," *Geoscience and Remote Sensing, IEEE Transactions on* , vol.43, no.5, pp. 1036- 1049.
- Guan, L., X. Zou, F. Weng, and G. Li (2011), "Assessments of FY-3A Microwave Humidity Sounder measurements using NOAA-18 Microwave Humidity Sounder", *J. Geophys. Res.*, 116, D10106, doi:10.1029/2010JD015412.
- Han, Y., F. Weng, Q. Liu, and P. V. Delst, "A fast radiative transfer model for SSMIS upper atmosphere sounding channel", *J. Geophys. Res.* (2007), vol. 112, D11121, doi: 10.1029/2006JD008208.

- Lu, Q., W. Bell, P. Bauer, N. Bormann and C. Peubey, “Characterizing the FY-3A Microwave Temperature Sounder using ECMWF Model”, *J. Atmospheric and Oceanic Technology* (November 2011), vol. 28, pp: 1373-1389.
- Mo, T., and K. Jarva, “Microwave humidity sounder calibration algorithm.”, *NOAA Tech. Report NESDIS 116* (October 2004): 22.
- Moradi, I., S. A. Buehler, V. O. John, and S. Eliasson (2010), “Comparing upper tropospheric humidity data from microwave satellite instruments and tropical radiosondes”, *J. Geophys. Res.*, 115, D24310, doi:10.1029/2010JD013962.
- Tobin, D. C., P. Antonelli, H. E. Revercomb, S. Dutcher, D. D. Turner, J. K. Taylor, R. O. Knuteson, and K. Vinson, “Hyperspectral data noise characterization using principal component analysis: application to the atmospheric infrared sounder.” *J. Appl. Remote Sensing* (June 2007), vol. 1, 013515.
- Tobin, D.C., H. E. Revercomb, and P. Antonelli, “Principal component analysis of IASI spectra with a focus on non-uniform scene effects on the ILS.” *AIP Conference Proceedings* (2009), 1100, 16, doi: 10.1063/1.3116941.
- Weng, F., “Advances in radiative transfer modeling in support of satellite data assimilation.” *J. Atmos. Sci.* (2007), vol. 64, no. 11, pp. 3799-3807.
- Weng, F., R. R. Ferraro, and N. C. Grody, “Effects of AMSU-A cross track asymmetry of brightness temperatures on retrieval of atmospheric and surface parameters.” in *Microwave Radiometry and Remote Sensing of the Earth’s Surface and Atmosphere* (1999), edited by P. Pampaloni and S. Paloscia, VSP Intl. Sci. Publ., ISBN 90- 6764-318-1, pp. 255–262.
- Weng, F., L. Zhao, R. Ferraro, G. Poe, X. Li, and N. Grody, “Advanced microwave sounding unit cloud and precipitation algorithms”, *Radio Sci.* (2003), vol. 38, no.4, MAR33–1 – MAR33-13.
- Zhang, P., J. Yang, C. Dong, N. Lu, Z. Yang, and J. Shi, “General introduction on payloads, ground segment and data application of Fengyun 3A.” *Frontiers of Earth Science in China* (2009), vol. 3, pp. 367–373.

BIOGRAPHICAL SKETCH

Yuan Ma grew up in Yangzhou, Jiangsu Province, China. Fascinated with weather and excelling in math and physics, Yuan attended School of Atmospheric Sciences in Nanjing University as a freshman and completed her Bachelor's degree in June 2010.

Yuan started graduate study at Florida State University in August 2010, under supervision of Dr. Xiaolei Zou. Her research interests include satellite data analysis and data assimilation. Yuan is going to further her study as a Ph.D. student.