Proceedings of the 8th International Symposium on
Spatial Accuracy Assessment in Natural Resources and Environmental Sciences
Shanghai, P. R. China, June 25-27, 2008, pp. 206-212

Information-Theoretical Comparison between Actual and Potential Natural Vegetation

Yan Chen [1] [+], Zongjian Lin [2] and Jiong You [1]

[1] School of Remote Sensing Information Engineering, Wuhan University, 129 Luoyu Road, Wuhan 430079
[2] Chinese Academy of Surveying and Mapping 16 Beitaiping Road, Beijing 100039, China

Abstract. An information theory measure of shared information, average mutual information (AMI), was applied to compare the similarity between potential natural vegetation (PNV) and actual vegetation (AV), where terrain factors were incorporated for mapping PNV, and TM imagery and MNDVI for AV, respectively. Object-oriented approaches were used to produce categorical maps for distributions of PNV and AV. In order to find the maximum relation between PNV and AV, AMI and contingency tables were used. Results show that AMI is low between PNV and AV, and varies with different scales. Therefore, simple combination of PNV and AV should be implemented with care.

Keywords: potential natural vegetation (PNV); actual vegetation (AV); average mutual information (AMI); contingency table

1. Introduction

There is a long tradition in vegetation science that vegetation distribution is primarily determined by the physical environment, assuming that stable plant communities and species are in equilibrium with environmental parameters. Therefore the vegetation-environment relationship plays a crucial role in the research of global change and terrestrial ecosystems. PNV of a site can be defined as the terminal plant community that would develop if all human influences on the site and its immediate surrounds would stop at once. However the composition types of AV are the result of co-adaptation between vegetation and nature. In other words, due to the millennia impact of human activities the AV covers up the relationship between vegetation and environment. As an expression of environmental factors such as topography, soils and climate across an area, the conception and theory of the potential natural vegetation have been widely applied into naturalness assessment, ecosystem restoration and reconstruction, resource utilization and planning, vegetation zonation and global change. As a result, there are now hundreds of literatures to quantify various aspects of landscape structure, diversity, and the potential effects on ecological processes (Turner 1990; Riitters *et al.*1995; Ricotta. Carlo *et al*. 2000).

However, little attention has been paid to the description of the similarity between PNV and AV. Many approaches for map comparison have been used in remote sensing (*e.g.* Pontius 2002; Foody 2004). The concept of mutual information from information theory provides a way to quantify this type of similarity. Finn (1993) has applied the AMI to compare thematic maps. Davis and Dozier (1990) used mutual information to help identify land-cover classes using a number of GIS layers including Landsat.

The aim of this study is to evaluate the similarity degree between PNV and AV in the area of central Montana, USA, and find the maximum relation between them. In this paper the image of the AV and the PNV are classified by object-oriented method. There are steps in the course of finding the relation between the PNV and the AV. In the first place, the contingency table is established. Secondly, AMI of them is

[+] Corresponding author. Tel.: +86-27-63767722;
E-mail address: beanchen@163.com.

ISBN: 1-84626-170-8, **ISBN13:** 978-1-84626-170-1

computed. Thirdly entropy and change in entropy of one map given the class label information in the other map are computed. Then contingency table and the table of entropy change are compared.

2. Data Description

Our study region is an area of central Montana, USA, located at $46°29' \sim 48°23'N$ and $107°52' \sim 111°08'W$. A spur of Little Belt Mountains located at South-west in this region, while Big Snowy Mountains south and Bearpaw Mountains north. Missouri River meanders across this location from west to east. Diversiform land cover types exist in this area, with dominant types of grassland, shrubland, forest, pine, fir and rock, form a very complex landscape in the zone (Xiong Rao, 2007)

The AV data we used is derived from the 07/12/96 P38/R27 Landsat 5 Thematic Mapper (TM) image, with 30 meters resolution. They include variables: TM1, TM2, TM3, TM4, TM5, TM7 and the Modified normalized difference vegetation index (MNDVI).The formula is as follows :

$$MNDVI = (TM4 - TM3)/(TM4 + TM3 + 1) \times (256/(TM5 + 1)) \times 100$$

And the PNV include digital elevation model (DEM), slope and aspect of the same area as the AV. Vegetation type code whose definition is listed in Table 1.

Table 1 Vegetation type code definition

Category code	Land cover type
1	Low Cover Grasslands
2	Moderate/ High Cover Grasslands
3	Mixed Broadleaf / Cottonwood Forest
4	Mixed Conifer Forest
5	Limber Pine/ Ponderosa Pine
6	Douglas-fir/ Lodgepole Pine
7	Xeric Shrublands/rock

3. Methods and results

3.1. Classification methods

In this paper, the object-oriented classification was used. In contrast to traditional image processing methods, the basic processing units of the object-oriented classification are image objects or segments, and not single pixels. The common, pixel-based approaches is done according to the spectral information, and it result in the mixture of different types and the well-known salt and pepper effect in the classification map. The object-oriented classification can eliminate the influence of salt and pepper effect. Because it considers a lot of information, beyond spectral information, there is shape information, texture information and many different relational or context features. For each feature this information is computed per-object considering its actual shape and size. Thus, the typical failures of filter operations, especially on transitions between different types of areas, are avoided.

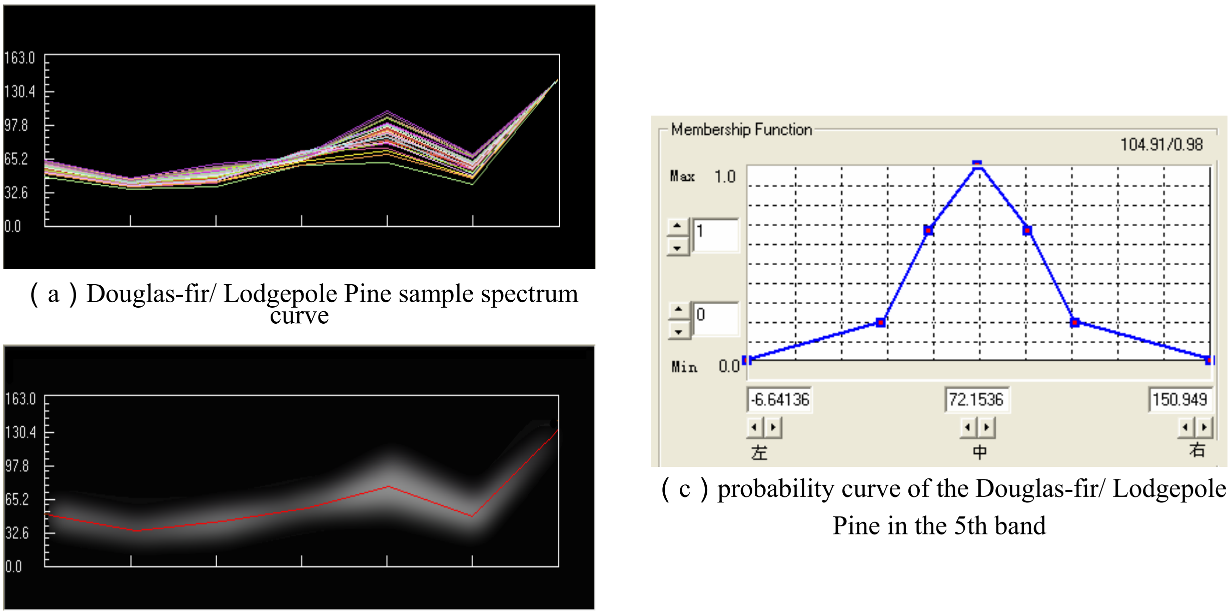
In object-oriented classification, image objects are obtained adopted the multiscale image segmentation. More details about the multiscale image segmentation can be found in [Kaimin Sun, 2006]. The feature is a complex structure which includes a set of parameters used to describe the essential of ground object in real world and it is the most important part in class definition.

And all detailed parameters of image objects are recorded. In image classification, it is these statistics of every band that determine what class an image object can be classified into. remote sensing image is uncertain and all the image objects corresponding to the same class may have different values for the same band. That is to say the feature value of nth band defined in the class can not set a fixed value because the actual values in corresponding image objects are changeful in a range. In order to describe the value of nth band in class, we replace the fixed value with a probability curve $p(x)$ based fuzzy theory. The x-coordinate of the probability curve $p(x)$ is the gray value of specified band and the y-coordinate is the corresponding probability p . So if the nth band value of an image object is V, it is more rational to say that the image object probably belongs to a class with a probability P than to say it does or not. It's certain that the probability P is a weighted average, as shown in following formula.

$$\begin{cases} P=\sum_{j=1}^{m}(W_j\sum_{i=1}^{n}p_{ji}w_i); \\ \sum_{j=1}^{m}W_j=1; \\ \sum_{i=1}^{n}w_i=1; \end{cases} \tag{1}$$

Where m is the amount of kind of statistics in one class definition and n is the amount of band, the p_{ji} is the probability of the value which is the j th kind of statistics in the i th band, the w_i is the weight of the i th band, generally speaking, it is $1/n$, the W_j is the weight of the j th kind of statistics. So the essential of the classification is to find the maximal probability by matching an image object with all classes according to above formula (1).

Fig 1 shows the Douglas-fir/ Lodgepole Pine sample spectrum curve, the curve of average feature and probability distribution of Fig.1 (a), and probability curve of the Douglas-fir/ Lodgepole Pine in 5th band. The probability curve in 5th band is regarded as the Douglas-fir/ Lodgepole Pine feature knowledge. Using these features knowledge, the images are classified based on fuzzy rules. The result of classification shows in Fig2. Fig.2(a) and (b) represses AV and PNV.

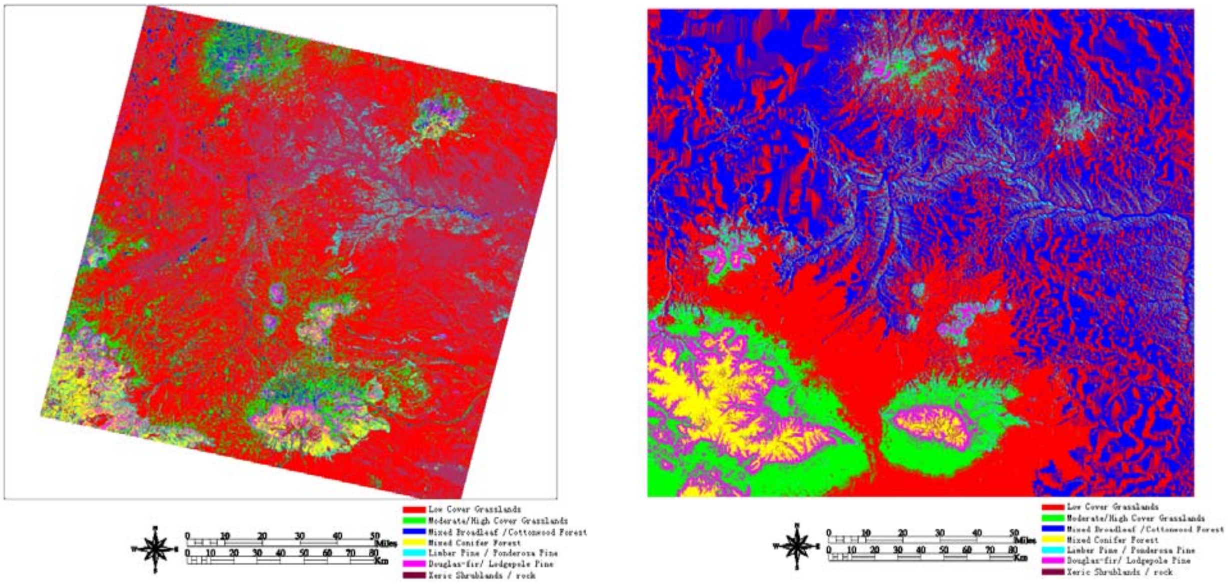


（a）Douglas-fir/ Lodgepole Pine sample spectrum curve

（b）average feature curve and probability distribution of Douglas-fir/ Lodgepole Pine sample

（c）probability curve of the Douglas-fir/ Lodgepole Pine in the 5th band

Fig. 1: Schematic Diagram of Douglas-fir/Lodgepole Pine feature curve


0 10 20 30 40 50 Miles
0 10 20 30 40 50 60 70 80 Km
Low Cover Grasslands
Moderate/High Cover Grasslands
Mixed Broadleaf /Cottonwood Forest
Mixed Conifer Forest
Limber Pine / Ponderosa Pine
Douglas-fir/ Lodgepole Pine
Xeric Shrublands / rock
0 10 20 30 40 50 Miles
0 10 20 30 40 50 60 70 80 Km
Low Cover Grasslands
Moderate/High Cover Grasslands
Mixed Broadleaf /Cottonwood Forest
Mixed Conifer Forest
Limber Pine / Ponderosa Pine
Douglas-fir/ Lodgepole Pine
Xeric Shrublands / rock

(a) actual vegetation

(b) potential natural vegetation

Fig.2: Classification Image

3.2. The calculation of the average mutual information

The methodology of mutual information analysis employs the concept of common entropy, which has been used in many fields. One means of comparing similarity of maps is to computer the average mutual information of maps. In other words, the average mutual information can be expressed the amount of shared information in the maps. In this regard, the information content of a map is its uncertainty. And uncertainty may be expressed in terms of entropy for a probabilistic system. Given image A, the entropy $H(A)$ can be calculated as:

$$H(A) = -K\sum_{i=1}^{n} p_i \log_2(p_i) \tag{2}$$

Where p_i is the proportion of area in the i th class and K is a constant that is often, and here, equal to 1 (Finn，1993).

For a pair of maps of the same location, A and B, let there be i class in map A and j class in map B. Supposed $p(a_i)$ is the proportion of class i in map A and $p(b_j)$ is the proportion of class j in map B. Overlapping Map A on to Map B, we can computer the average mutual information. The average mutual information involves the conditional probability $p(a_i \mid b_j)$ and the joint probability $p(a_i, b_j)$. The conditional probability $p(a_i \mid b_j)$ expresses that a pixel in map A belongs to class i given it is in class j on map B. And the joint probability $p(a_i, b_j)$ represents a pixel being class i in map A and class j in map B. Then the amount of information shared between the two maps is calculated by these formulas:

$$\begin{aligned} H(A;B) &= H(A) + H(B) - H(AB) \\ &= H(A) - H(A \mid B) \\ &= H(B) - H(B \mid A) \end{aligned} \tag{3}$$

where $H(A)$ and $H(B)$ are the entropy of map A and map B; $H(AB)$ is the joint entropy, $H(A \mid B)$ is the conditional entropy.

Sometimes it may be useful to focus on specific class(es) in the map. On an individual classes basis, a posteriori entropies for one map given the class label information from the other may be used to evaluate the amount of information shared by the maps (Foody, 2006). The a posteriori entropy of A once b_j is known may be calculated from

$$H(A|b_j) = -K\sum_{i=1}^{n} p(a_i|b_j)\log[p(a_i|b_j)] \tag{4}$$

The difference between $H(A)$ and $H(A|b_j)$ shows that the information can be provided through the knowledge of b_j. If the difference is less than 0, it shows the uncertainty increase given the knowledge of b_j.

Using these formulas, the assessment of similarity may be undertaken on an overall map or individual class basis. To evaluate the level of correspondence between classes, a contingency table of the cross-tabulated labels is used. The classes in map A were labeled A-G and those in map B labeled I－VII, so they are all corresponding code definition of vegetation type in Table 1.

Table 2. Contingency table of the cross-tabulated labels in A and B.

Map A	Map B								
	1	2	3	4	5	6	7	total	Agreement (%)
1	9845955	1735009	399265	92929	526700	78891	9361718	22040467	44.67
2	1229408	624278	129928	216878	194382	178509	2102664	4676047	13.35
3	6950592	678208	338074	44285	590346	83658	12155903	20841066	1.62
4	107365	38490	9118	375008	157174	99691	921290	1708136	21.95
5	694432	275199	101817	67269	221708	88637	947269	2396331	9.25
6	190300	167061	57852	405113	242075	341370	1208483	2612254	13.07
7	236654	63070	39827	55301	133430	45819	945934	1520035	62.23
total	19254706	3581315	1075881	1256783	2065815	916575	27643261	55794336	
Agreement (%)	51.14	17.43	31.42	29.84	10.73	37.24	3.42		

Table 2 helped in the assessment of the degree of correspondence between the class labels in the two maps and the degree of difference between them. It shows, for example, that class G in map A corresponded closely to class VII in map B, with 3.42% (945934 out of 27643261) of pixels of class VII lying in class G and 62.23%(945934 out of 1520035) of cases of class G lying in class VII. The entropy of the observed map, $H(A)$,is 2.1489 while the reference map has an entropy of $H(B)$= 2.0559. The class A corresponded closely to class I, with 51.14% of pixels of class I lying in class A， and 37.24% of pixels of class VI lying in class E. An overall assessment of the degree of shared information in the maps may be expressed in terms of the amount of shared information in the maps, represented by the AMI. The AMI of map A and map B is 0.0452, only 2.2% of the information in the reference map.

Sometimes it may be useful to focus on specific class in the map. On an individual class basis, a posteriori entropies of one map given the class label from the other may be used to evaluate the amount of information shared by the maps. Using Eq. 3, the a posteriori entropy, the change uncertainty and the percentage change were calculated (Table 3 and Table 4).

Table 3 shows the a posteriori entropies for AV given each class in PNV, the change in entropy for AV given PNV class, and the percentage change in entropy for AV given each class in PNV. Class 1 has the best reduction in uncertainty, 12.57%. Knowing that a point on the map is class II,III,IV,Vor VI would increase the uncertainty about map A class on that spot. In other words, the class I, VII on the map B can play a role identifying the vegetation type of map A.

Table 4 shows the a posteriori entropies for map B given each class in map A, the change in entropy for map B given map A class, and the percentage change in entropy for map B given each class in map A.

Class C is the highest at 30.17 per cent overlapping with 7 class of map B, because 12155903 of 20841066 of C fell in map B's VII class. Class A is next highest at 21.06 per cent, which means that the identity of the class map B will become more easier given class A of map A. Class F has the lowest change in entropy (-11.09 per cent), actually increasing uncertainty about the identity of the class of map B.

Table3 Entropy and change in entropy of map A given the class label information in map B

b_j	$H(A\|b_j)$	$H(A)-H(A\|b_j)$	$(H(A)-H(A\|b_j))/H(A)$
1	1.8787	0.2702	12.57%
2	2.3397	-0.1908	-8.88%
3	2.2068	-0.0579	-2.69%
4	2.3568	-0.2079	-9.67%
5	2.5860	-0.4371	-20.34%
6	2.5003	-0.3514	-16.35%
7	2.0273	0.1216	5.66%

Table4 Entropy and change in entropy of map B given the class label information in map A

a_i	$H(B\|a_i)$	$H(B)-H(B\|a_i)$	$(H(B)-H(B\|a_i))/H(B)$
1	1.6286	0.43304	21.06%
2	2.1328	-0.0769	-3.74%
3	1.4357	0.6202	30.17%
4	1.9311	0.1248	6.07%
5	2.2377	-0.1818	-8.84%
6	2.2839	-0.2280	-11.09%
7	1.8061	0.2498	12.15%

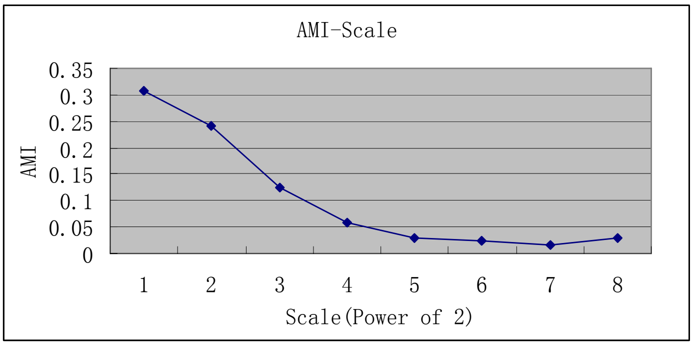


Fig. 3: The scale and the corresponding average mutual information (AMI)

At the same time, there is an phenomenon in computing average mutual information. From fig.3 it is clear that the amount of average mutual information is changing with the scale. There the scale represses the different meaning compared with the scale of multiscale image segmentation. Here, the scale represents the compared window. The scales are $2^n \times 2^n$ $(n=1,2,...8)$.

4. Conclusion

In this paper, the similarity between PNV and AV has been discussed. AMI and contingency table were applied to assess the similarity. The main findings of the paper are that the simple combined classifier may not be suitable for mapping PNV and AV all together, as they may be extremely dissimilar and should be treated separately. In the paper, PNV is defined by terrain factors, such as elevation, slope, and aspect. The results of the experiments showed that the similarity between PNV and AV is very low. The main reason for this is that spectral properties of vegetation are the classification criteria, but the criteria of PNV are terrain factors. Consequently, there is a low similarity in the two maps. Therefore, simple combination of PNV and AV should be implemented with care.

Proper mapping of AV and PNV involves semantics. Further room for improvement may lie in bridging the gaps between class definitions inherent to AV and PNV. Because the same class can be expressed the different class due to the diversity source and define of data. On the other hand, the degrees of AMI are varied with scales. At a certain scale, the maximum similarity between PNV and AV may be obtained, which should used as guidance for constructing classifiers.

5. Acknowledgements

The research is partially supported by a “973 Program” grant (2007CB714402－5), and a grant received from the Key Lab in Geospatial Information Engineering, Chinese Academy of Surveying and Mapping, State Bureau of Surveying and Mapping (2007020).

6. References

[1] Turner, M.G. Spatial and temporal analysis of landscape patterns. *Landscape Ecol*. 1990, **4**: 21-30.

[2] Riitters, K.H. *et al.* A factor analysis of landscape pattern and structure metrics. *Landscape Eecol*. 1995, **10**: 23-39.

[3] Ricotta, Carlo. e*t al.* Quantitative comparison of the diversity of landscapes with actual vs. potential natural vegetation. *Applied Vegetation Science*, 2000, **3**: 157-162.

[4] Foody GM. Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. *Photogrammetric Engineering and Remote Sensing,* 2004, **70**: 627-633.

[5] Foody GM. What is the difference between two maps? A remote senser’s view. *Journal of Geographical Systems*. 2006, **8**: 119-130.

[6] Pontius RG. Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogrammetric Engineering and Remote Sensing,* 2002, **68**: 1041-1049

[7] Finn JT. Use of the average mutual information index in evaluating classification error and consistency. *Int. J. Geographical information systems*. 1993, **7**: 349-366.

[8] Kaimin Sun, Yan Chen, and Deren Li. Multiscale image segmentation and its application in image information extraction. *Proceedings of SPIE,* Volume 6419*,Geoinformatics 2006: Remotely Sensed Data and Information.*

[9] Xiong, Rao, Jinping, Zhang, Brian M. Steele, Roland L. Redmond. Generalized linear models for mapping land cover using satellite measurement and digital terrain data. *Proceedings of SPIE,* Volume 6751*, Geoinformatics 2007: Cartographic Theory and Models.*

Proceedings of the 8th International Symposium on
Spatial Accuracy Assessment in Natural Resources and Environmental Sciences
Shanghai, P. R. China, June 25-27, 2008, pp. 213-220

Accuracy Issues Associated with Satellite Remote Sensing Soil Moisture Data and Their Assimilation

Xiwu Zhan

NOAA NESDIS Center for Satellite Applications and Research, Camp Springs, MD 20746, USA

Abstract. Satellite remote sensing is widely used for monitoring the changing planet Earth. Many remote sensing data products are being generated and used every day. Among these data products are the microwave remote sensing data of land surface soil moisture. Soil moisture often limits the exchanges of water and energy between atmosphere and land surface, controls the partitioning of rainfall among evaporation, infiltration and runoff, and impacts vegetation photosynthetic rate and soil microbiologic respiratory activities. Their accuracy plays essential role for the success of their applications. Accurate measurement of this variable across the global land surface is thus required for global water, energy and carbon cycle sciences and many civil and military applications. Currently available satellite soil moisture data products have been generated from the low frequency channel observations of the currently flying microwave sensors (the TRMM Microwave Imager-TMI; Aqua Advanced Microwave Scanning Radiometer-AMSR-E, and Navel Research Lab's WindSat). However, because of several accuracy issues all of these soil moisture data have not yet been used in operational applications. The most apparent accuracy issue is that the soil moisture data retrievals from the three different sensors are significantly different from each other even when they are retrieved with the same algorithm. This might have been caused by the calibration errors in their brightness temperatures. A Simultaneous Conical-scanning Overpass (SCO) method is tested to address this issue. Secondly, satellite sensor footprints are usually several orders larger than the local points where in situ soil moisture measurements for validation are obtained. How to appropriately compare the satellite soil moisture retrievals of large spatial areas with the in situ measurements becomes an important issue. A point-to-pixel mapping approach is examined for a solution of this issue. The third issue is how to handle biases of the soil moisture retrievals from land surface model (LSM) simulations when they are assimilated into the LSM. Existing solutions for this issue are summarized and whether these error-handling strategies are effective or reliable are discussed. Finally general conclusions of this study are presented for users who are interested in satellite soil moisture data assimilation.

Keywords: data accuracy, soil moisture, satellite remote sensing, data assimilation

1. Introduction

Satellite remote sensing has become the most effective tool for monitoring the changing planet Earth. Since the launch of Earth Observing System Terra and Aqua satellites in 1999 and 2002 respectively, extensive amounts of remote sensing data of land surface properties have become available [1 Parkinson, 2002]. For example, land surface temperature data from the moderate resolution imaging spectroradiometers (MODIS) on both Aqua and Terra satellites, the operational advanced very high resolution radiometers (AVHRR) on NOAA polar satellites, and the imagers on the geostationary operational environmental satellites (GOES) have been available for many years. Satellite precipitation data from the various microwave sensors onboard the tropical rainfall monitoring mission (TRMM) and other satellites are freely available at increasing space and time resolutions. Data for snow cover, snow depth, vegetation dynamics and surface soil moisture have become available from a number of satellites. However, the potential maximal returns of these costly investments has not yet achieved because the advances of technology for the analysis and interpretation of these data have not progressed at proportional paces [2 Walker & Houser, 2003].

Among these data products are the microwave remote sensing data of land surface soil moisture. Soil moisture often limits the exchanges of water and energy between atmosphere and land surface, controls the

ISBN: 1-84626-170-8, **ISBN13:** 978-1-84626-170-1