P1.2 SELF-ORGANIZING NONLINER OUTPUT MAP (SONO): AN ARTIFICIAL NEURAL NETWORK SUITABLE FOR CLOUD-PATCH BASED RAINFALL ESTIMATION

Kuo-Lin HSU*, Yang Hong, and S. Sorooshian

University of California, Irvine, California

1. INTRODUCTION

The spatial and temporal distribution of precipitation is crucial for water management for agriculture, electrical power, and for drought and flood control. For many regions where ground observations are lacking, the development of satellite remote sensing techniques provides a unique opportunity to extend precipitation measurement.

Remote sensing of precipitation usually depends on visible/infrared (IR) and microwave (MW) radiometers to interpret rainfall over the surface. From previous intercomparison project, it has been demonstrated that MW-based algorithms give better instantaneous estimates from their direct sensing capability, while IR/VIS-based algorithms give better long-term average estimates from the advantage of large amount of temporal samples from geostationary orbits (Ebert, 1996). Many recent developments have been in the merging both IR and MW sensor's information to generate better precipitation fields.

One particular example of merging IR and MW information to generate improved rainfall estimates is the PERSIANN (Precipitation Estimation from Remotely Sensed Information using a Artificial Neural Network) algorithm. This algorithm calculates rainfall rate (RR) based on the local texture of cloud top IR brightness temperatures (T_b) near the calculation pixels, while limited MW rainfall estimates from low orbital satellites are regularly used in the adjustment of algorithm parameters (Hsu et al., 1997; Sorooshian et al., 2000). In this study, Self-Organizing Nonlinear Output (SONO) is presented in the retrieval of surface rainfall. Instead of extracting local texture information and facilitating one adjustable fitting function in PERSIANN, SONO extracts information from the whole cloud patch and provides multiple T_b-RR functions for different cloud types. Based on a step-by-step data processing through satellite IR cloud image segmentation, classification, and nonlinear mapping, surface rainfall rates are generated.

* Corresponding author address: Kuo-Lin Hsu, University of California at Irvine, Department of Civil and Environmental Engineering, E/4130 Engineering Gateway, Irvine, CA 92697-2175.

In this study, the development of SONO and its application to the rainfall retrieval using geostationary satellite IR imagery is introduced. Evaluation of SONO rainfall retrieval based on WSR-88 radar data and a comparison of model performance with UAGPI method is discussed.

2. METHODOLOGY

The SONO model is shown in Figure 1. This model has an input preprocessor which extracts IR cloud features as input variables. SONO performs two basic functions as "switchboard" and "approximator". First, the SOFM (self-organizing feature map) is used as a classification tool to select or assign the candidate unit in the Nonlinear Output layer, while this output layer consists a number of exponential fitting functions, which served as an approximator to the T_b-RR relationship.

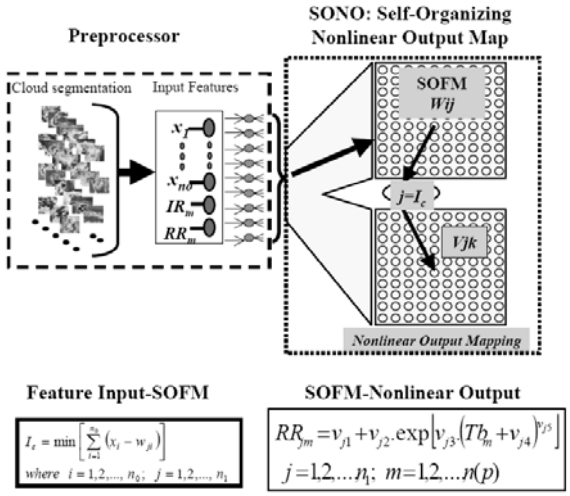

Figure 1. Preprocessor and Self-Organizing Nonlinear Output (SONO) map

2.1 Image Preprocessing

Segmentation of IR imagery is a preprocess step for cloud analysis. Incremental Temperature Threshold (ITT; Hong et al. 2003) method is proposed for this purpose. The ITT is a hybrid patch segmentation approach, which includes the advantages of both

hierarchical thresholding and Seeded Region Growing (SRG; Adams 1994). Given a snapshot of satellite IR image, ITT first locates the minimum temperature (Tb_{min}) as seeds (illustrated by cross marker), and then starts to iteratively expand each seed's area one neighborhood size at a time until touching neighboring clouds or temperature threshold that delineates cloud from the clear sky. Figure 2 shows an example of cloud segmentation using ITT method.

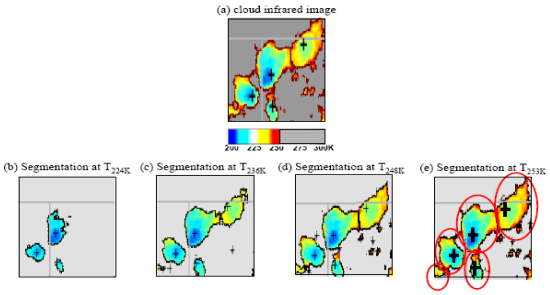

Figure 2. Cloud patches from (a) GOES infrared imagery are separated slowly by increasing IR temperature thresholds from low to high (b)-(e) in the ITT segmentation algorithm.

2.2 Input Feature Extraction

The characteristics of cloud patches, relevant to precipitation, are grouped into three categories: coldness, geometry, and texture. The first category is generally associated with the geophysical variables—cloud brightness temperature; the second one is derived from the geometric properties of cloud patch; and the third is the texture variation of cloud brightness temperature (see Table 1). Statistical analysis found that the first category is mostly relevant to the rainfall intensity in a manner of negative correlation and that the size in the second category is positively correlated to rainfall volume. Although the features in third category are not necessary directly related to rain rate or rainfall volume, they do improve the discrimination of cloud clusters.

Table 1. The input features extracted from cloud patches

Category	Features
Coldness	Minimum temperature of a cloud patch (IR_{min})
	Mean temperature of a cloud Patch (IR_{mean})
Geometric	Cloud patch size (Size)
	Cloud Patch Shape Index (SI)
Texture	Standard deviation of cloud patch (STD)
	Mean value of local standard deviation of cloud patch (MSTD(5x5))
	Standard deviation of local standard deviation (STD_{std}(5x5))
	Gradient of cloud top brightness temperature (TOPG)

2.3 Image Classification

Given a cloud patch p, the preprocessor provides three sets of input vectors for SONO:

$$p=([x_i], [Tb_m], [RR_m]),\ i=1,\ldots,n_0;\ m=1,\ldots,N(p) \quad (1)$$

Where $[x_i] = (x_1, x_2,..., x_{n0})$ is the feature vector listed in Table 1; n_0 is the number of input features; $[Tb_m]$ is the IR brightness temperature vector within the patch and $[RR_m]$, if available, is the corresponding observation of rainfall vector; and $N(p)$ is the number of pixels of the cloud patch p.

SONO consists of two major components: SOFM and nonlinear regression fittings (see Figure 1). Classification layer is comprised of n_1 units or groups: $\mathbf{y}=[y_1, y_2, ..., y_{n1}]$; the weight matrix w_{ij} is a set of parameters connecting from input node x_i to the classification units y_j. The mapping layer, $\mathbf{z}=[z_1, z_2, ..., z_{n1}]$ contains n_1 nonlinear fitting functions, where v_{jk} are the k^{th} parameter of a nonlinear function connecting from unit y_j to output unit z_j. The SOFM classification proceeds based on the similarity of the cloud patch features **x** and the connection weights **w.** The training of SOFM projects input patterns of many variables into an organized cloud classification map. Detail of the training procedure is described in Kohonen (2001) and Hsu et al. (1999). A brief summary of the training algorithm is listed below:

Step 1: Prepare the input **x** and initialize the weights w_{ij} as normalized random number.

Step 2: Determine the winner node that has minimum distance between **x** and weight **w**:

$$j^* = \arg_j(\min\|x_{ij} - w_{ij}\|) \quad (2)$$

Step 3: Update weight for the neighborhood (radius r) nodes of j^* with learn rate α :

$$w_{ij}=w_{ij}+\alpha\,(x_{ij}-w_{ij}) \quad (3)$$

Step 4: Terminate if the w_{ij} is converged, or reduce r and α and go to step 3.

2.4 Mapping Pixel Rainfall

As a result of SOFM classification, a dataset of Tb-RR pairs is classified to the unit j and the data pairs are sorted based on the Probability Matching Method (PMM; Atlas et al. 1990). It is assumed that higher rain rates are associated to lower IR brightness (Arkin and Meisner 1987). Thus, the PMM matches the cumulative

distribution functions (CDF) of the data pairs of Tb-RR as follows:

$$\int_0^{RR_m} P(RR)\,dRR = \int_{IR_{min}}^{IR_{max}} P(Tb)\,dTb - \int_{IR_{min}}^{IR_m} P(Tb)\,dTb = 1 - \int_{IR_{min}}^{IR_m} P(Tb)\,dTb \quad (4)$$

Where the *P(.)* is the probability distribution function and the estimated value of Tb and RR is in the range [0 RR_{max}] and [Tb_{min} Tb_{max}], respectively.

The use of power-law mapping between Tb and RR was suggested by the study of Gagin et al. (1985) and implemented by several others (Martin et al., 1990; Goodman et al., 1994; and Vicente et al., 1998). Here the nonlinear IR-RR fitting function is:

$$RR_{jm} = f(v_{jk}, Tb_m) = v_{j1} + v_{j2} \cdot \exp[v_{j3}(Tb_m + v_{j4})^{v_{j5}}] \quad (5)$$

The above function consists of five parameters [v_{jk}], k=1,2, …, 5. In this case, SONO processes a large number (n_1) of IR-R mapping relationships corresponding to different classified cloud group in the SOFM layer.

3. CASE STUDY

The SOFM classification layer is assigned as a 20 x 20 matrix, which classifies cloud patches into 400 groups (see Figure 3a). Note that both the SOFM classification layer and nonlinear mapping layer consist of the same arrangement of units in a 2-D coordinate. Therefore, a matrix (20 x 20) of T_b-RR relationships are calibrated according to Equation (5) and all the 400 curves are plotted on an T_b-RR plane (Figure 3b). Flat curves indicate that the cloud top temperatures are usually cold but produce little or no rain; steep curves are relevant to the convective clouds that are capable of producing significant rainfall.

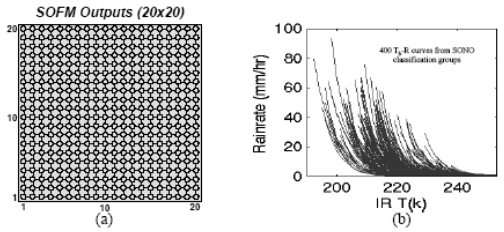

Figure 3. (a) 400 classification groups in SOFM and (b) their IR-RR fitting curves in each classified data group.

Validation and comparison of SONO was conducted at sub-daily, daily, and monthly temporal resolutions for spatial scales such as 0.04^o, 0.12^o, 0.5^o, and 1.0^o. Another cloud patch-based algorithm, Universal Adjusted GPI (UAGPI) (Xu et al., 1999) was used in the comparison.

3.1 Sub-Daily Rainfall

The rainfall observations used to evaluate SONO estimates are collected from Weather Surveillance Radar-1988 Doppler (WSR-88D), which is accessible from the National Center for Environmental Prediction (NCEP). Evaluation of rainfall estimates from SONO is conducted of various spatial and temporal scales. Here we have listed part of the performance evaluation at four 1^o x 1^o grids (28^{o}N-30^{o}N and 102^{o}W-100^{o}W) located at Texas from July 1st through 10th, 2002. Figure 4 presents the scatter plots of hourly, 3-hour, 6-hour, and daily rainfall of those four grids. It shows that at 1^{o}x1^o scale, SONO estimates are with high correlation coefficients (CC) to the radar rainfall around 0.75 at hourly scale and 0.85 at 24-hour (daily) scale. Comparison of model performance with the UAGPI (see Table 2), SONO is consistent with lower RMSE and Bias errors. Both UAGPI and SONO consist of high CCs in all scales. CCs of both models are high (around 0.9), especially for rainfall at daily scale,

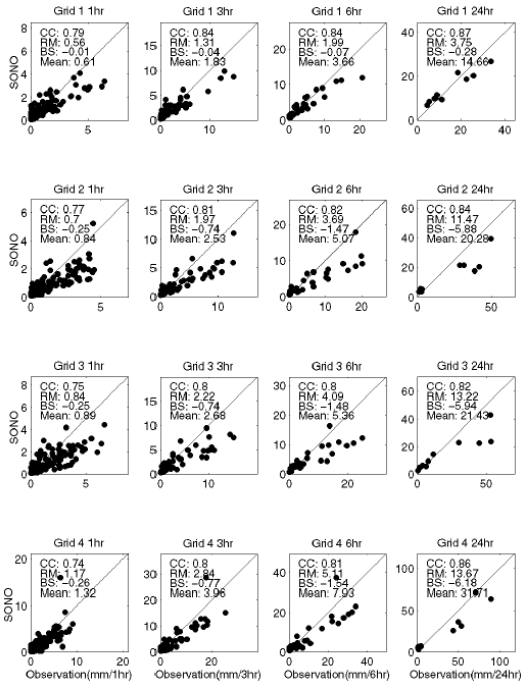

Figure 4. Scatterplots of time series of hourly, 3-hour, 6-hour, and 24-hour rainfall at four 1o x 1o grids for July 1-10 2002 located in Texas, 28oN-30oN and 102oW-100oW region

Table 2 Statistical comparison of time series averaged 1°x1° SONO estimates vs. WSR-88D radar located Texas region 29°-30°N and 101°-100°W at 1-hour, 3-hour, 6-hour, and daily intervals from 1st July 2002 through 10th July 2002.

Time Period	Radar mean (mm)	RMSE (mm)		CORR		Bias (mm)	
		UA GPI	SO NO	UA GPI	SO NO	UA GPI	SO NO
Hour	1.32	1.51	1.39	0.69	0.79	-0.6	-0.1
3-hrs	3.96	4.23	3.12	0.77	0.85	-1.8	-0.5
6-hrs	7.93	7.56	5.42	0.80	0.87	-3.7	-1.1
Daily	31.71	23.2	12.9	0.89	0.93	-14.	-4.4

3.2 Monthly Rainfall

Evaluation of SONO and UAGPI monthly rainfall covers a 10° x 10° region (30°-40°N and 105°-115°W) at central Texas plain using WSR-88D network data. One month of data in July 2002 was used in the case study. Figure 5 shows the scatter plots of monthly rainfall total at 0.04° ~ 1.0° grid scales generated from SONO and UAGPI rainfalls. Overall, UAGPI fails to catch high rainfall but its performance improved when spatial resolution is accumulated to a low-resolution scale; at spatial scales 0.5°~1.0°, CCs are around 0.6~0.7.

SONO, on the other hand, is able to catch the rainfall at higher spatial resolution (e.g. 0.04°). The model performance is also improved at lower spatial resolution. Correlation coefficients of SONO are as high as 0.75 at all scales. In comparison with UAGPI, SONO estimates consistently show better performance in terms of the listed evaluation statistics (CCs, RMSEs, and BIASs).

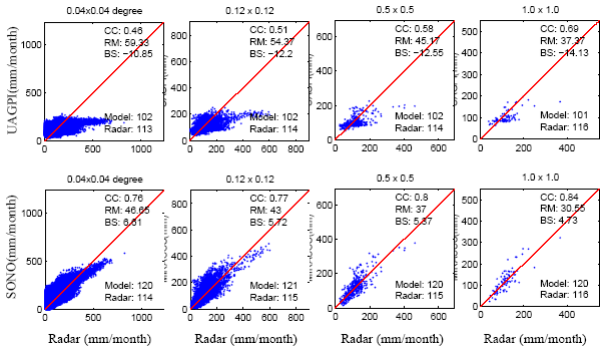

Figure 5. Scatterplots of July 2002 monthly rainfall total derived from Exponential function, UAGPI, and SONO vs. radar at 30N0-400N and 105W0-1150W region.

3.3 Adaptability of SONO model

The re-calibration of parameter v_{jk} enables a better fit to the data provided by surface radar or TMI and SSM/I rainfall estimates. In this study, SONO model was trained using one-month (June 1999) WSR-88D radar rainfall, which might not represent the generalized behavior of those classified T_b-RR curves. For future operation of SONO algorithm beyond the test regions and seasons, it will be useful to include additional information from MW TMI and SSM/I rainfalls.

4. CONCLUSIONS

In this study, we introduced the SONO model and its application to the retrieval of surface rainfall using geostationary imagery. Ground radar rainfall data was used for the calibration and validation of SONO. The SONO includes image processing, feature extraction, classification at rough cloud patch scale and provides multiple functions of mapping pixel image to rainfall at small scale. The different T_b-RR mapping functions for the classified cloud patch groups generate variable rain intensities at a given IR brightness temperature. Comparison of SONO and UAGPI shows that SONO estimates outperform UAGPI in terms of three evaluation statistics (RMSE, BIAS, CC).

Future work is to develop SONO as an operational satellite-based estimation system to produce rainfall information at small scales. Given the more accurate rainfall estimates from low-orbital MW imagers, an adaptive procedure will be used to adjust SONO model parameters.

ACKNOWLEDGEMENTS

Financial support for this study was provided by the NASA NRA02OES05 grant, NSF for Sustainability of semi-Arid Hydrology and Riparian Areas (SAHRA) (grant EAR-9876800), and NASA-EOS (grant NA56GPO185).

REFERENCES

Admas, R. and L. Bischof (1994), Seeded region growing, *IEEE Trans. Pattern Anal. Machine Intell.*, *16(6)*, 641-647.

Arkin, P.A. and B. N. Meisner (1987), The relationship between large-scale convective rainfall and cold cloud over the Western Hemisphere during 1982-84, *Mon. Wea. Rev., 115*, 51-74.

Atlas, D., D. Rosenfeld, and D.B. Wolff (1990), Climatologically tuned reflectivity-rainrate relations and links to area-time integrals, *J. Appl. Meteor.*, *29*, 1120-1135.

Ebert, E.E., M.J. Manton, P.A. Arkin, R.J. Allam, C.E. Holpin, and A. Gruber (1996) Results from the

GPCP Algorithm Intercomparison Programme**.** *Bulletin of the American Meteorological Society*, 77(12),. 2875–2888.

Gagin, A., D. Rosenfeld, and R.E. Lopez (1985), The relationship between height and precipitation characteristics of summertime convective cells in South Florida, *J. Atmos. Sci.*, *42*, 84-94.

Goodman, B., D.W. Martin, W.P., Menzel, and E.C. Cutrim (1994), A non-linear algorithm for estimating 3-hourly rain rates over Amazon from GOES/VISSR observations, *Remote Sens. Rev.*, *10*, 169-177.

Hsu, K., X. Gao, S. Sorooshian, and H. V. Gupta (1997), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks, *J. Appl. Meteorol*., *36*, 1176-1190.

Hsu, K., H. V. Gupta, X. Gao, and S. Sorooshian (1999), Estimation of physical variables from multichannel remotely sensed imagery using a neural network: Application to rainfall estimation, *Water Resour. Res., 35(5)*, 1605-1618.

Hong, Y., K. Hsu, S. Sorooshian (2003), An Automatic Segmentation Algorithm for Cloud Infrared Satellite Images: Incremental Temperature Threshold Technique, Report, *03-020*, 55 pp, Dept. of Hydrology and Water Resources, The University of Arizona, Tucson, AZ.

Kohonen, T (2001), Self-Organizing Maps, Springer Series in Information Sciences, Vol. 30, 3rd ed., 501pp, Springer, New York, ISBN 3-540-67921-9.

Martin, D.W., B. Goodman, T.J. Schmit, and E.C., Cutrim (1990), Estimates of daily rainfall over the Amazon Basin, *J. Geophys. Res., 95(D10)*, 17043-17050.

Sorooshian, S., K. Hsu, X. Gao, H.V. Gupta, B. Imam, and D. Braithwaite (2000), Evaluation of PERSIANN system satellite-based estimates of tropical rainfall, *Bul. Am. Meteorol. Soc.*, *81(9)*, 2035-2046.

Vicente, G.A., R. A. Scofield, and W.P. Menzel (1998), The Operational GOES Infrared Rainfall Estimation Technique, *Bul. Am. Meteorol. Soc.,* 79(9), 1883–1898.

Xu, L., S. Sorooshian, X. Gao, H. Gupta (1999), A Cloud-Patch Technique for Identification and Removal of No-Rain Clouds from Satellite Infrared Imagery, *J. Appl. Meteor., 38*, 1170–1181.