Preliminary Results of Atmospheric Temperature Retrievals with Least Squares Support Vector Regression

Yao Zhigang, Chen Hongbin, Gu Hongfang

Institute of Atmospheric Physics, CAS, Beijing, China

Background

- **One technique for retrieving temperature profiles from microwave radiances is a linear statistical inversion.**
- **Back propagation neural networks were employed in some studies (Churnside et al. 1994; Motteler et al. 1995; Butler et al. 1996; Shi 2001).**
- **Recently, Support Vector Regression emerges as an alternative regression tool. SVR is a derivation of Support Vector Machines (SVM), introduced by Vapnik (1995).**

Characteristics of LS-SVR

- **Least Squares Support Vector Regression (LS-SVR) is a reformation of SVR (Sukykens, 2002)**
- **It gives a global and perhaps unique solution**
- **It finds a general solution and thus avoiding overtraining.**
- **It gives a solution which is sparse**
- **It is able to model non-linear relations**
- **It solves the regression problem by a set of linear equations.**

Objective

- **Investigate the performance of Least Squares Support Vector Regression for retrieving atmospheric temperatures from AMSU-A measurements.**

Data

- **NOAA-16 ATOVS/AMSU data received at Beijing, China**
- **The absolute distance between the position (latitude and longitude) of the radiosonde and the ATOVS retrieval FOV is less than 0.5º.**
- **The time difference between radiosonde and ATOVS measurements is less than 1.5 h.**
- **The radiosonde observations from reporting stations with terrain heights less than 500 m are selected.**
- **Brightness temperature records of AMSU-A are complete.**
- **The radiosonde observation has no missing layers from 1000- to 100-hPa pressure levels.**
- **4649 pairs from Jan 2002 to Dec 2003 and 2944 pairs from Jan 2004 to Sep 2004**

Tuning-grid search algorithm

In case of LS-SVR, there are two parameters to be tuned

- **For each set of values of the parameters, leave-n-out-cross validation on the training set is performed to predict the prediction error.**
- **Select the set of values of the parameters that produced the model that gave the smallest prediction error (optimal parameter settings).**
- **Train the model with the optimal parameter settings with the whole training set and test it with a test set (test is not used for training).**

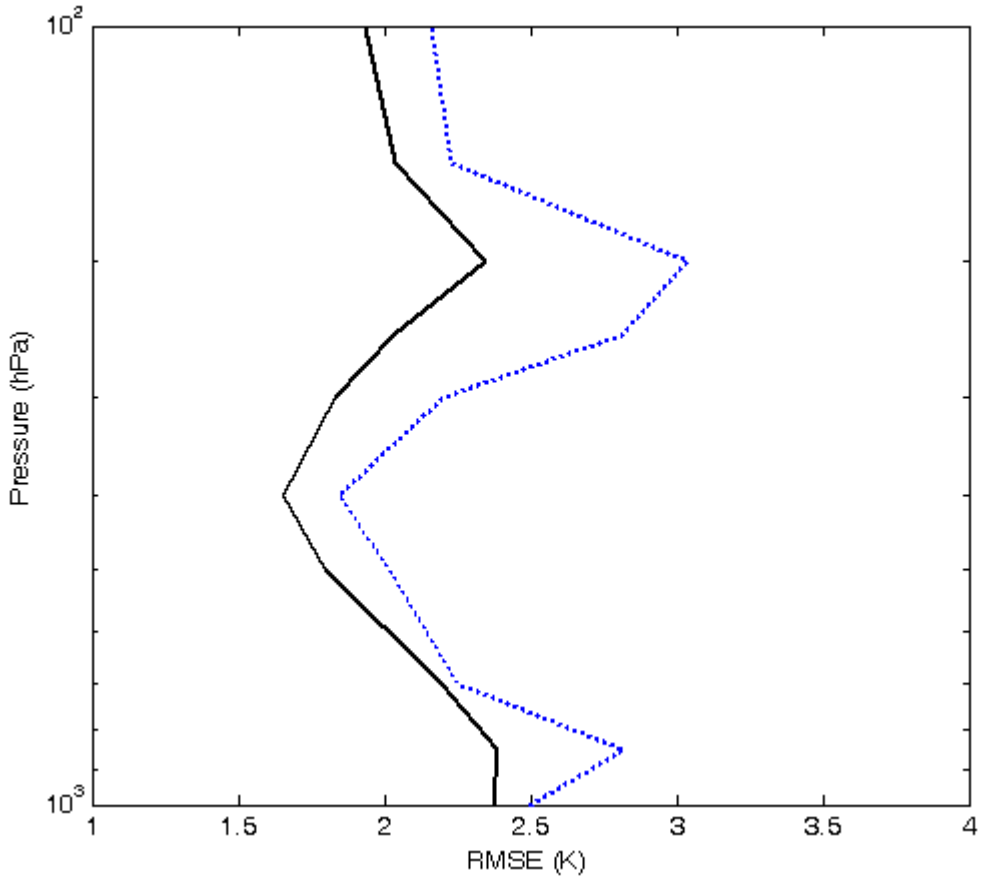
Software

LS-SVR is calculated using the LS-SVMlab1.5 Toolbox developed by Suykens.

The toolbox can be obtained from http://www.kernel-machines.org.

The Toolbox works under Matlab (The Mathworks, Inc.).

Preliminary results



The collocated data from Jan 2002 to Dec 2003 are used to train the regression model, and the data from Jan 2004 to Sep 2004 are used to test the model.

Fig. 1 Comparison of the RMS retrieval errors of LS-SVM with MLR (Dotted: MLR; Solid: LS-SVR)

Training 400 Testing 2944

Effects of the volume of the training dataset on the testing results

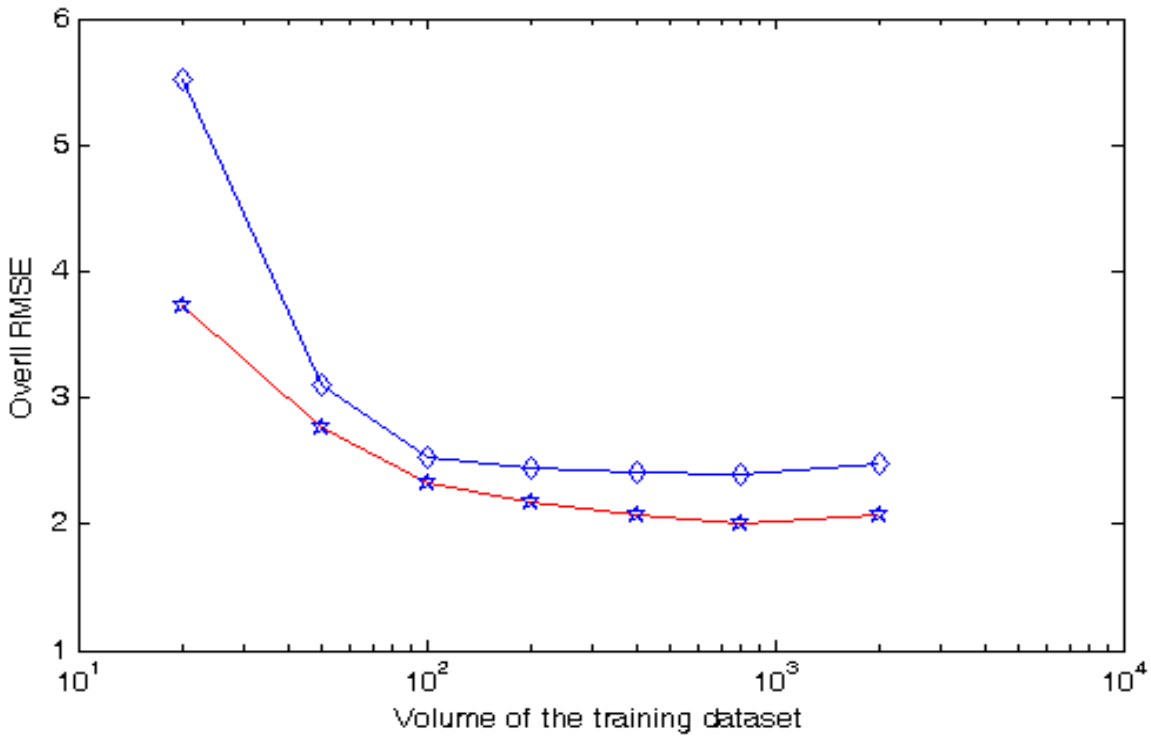


Fig.2 Comparison of the overall RMS errors of LS-SVM with MLR for the different volume of the training data set (Blue: MLR; Red: LS-SVR)

Effects of noise and calibration offset on the results

Table 1. Variation of the overall error with the offset (Training: 400; Testing: 2944)

	0.0	0.2	0.5	1.0	1.5	2.0
MLR	2.41	+0.02	+0.07	+0.21	+0.41	+0.66
LS-SVR	2.07	+0.02	+0.08	+0.29	+0.59	+0.95

Table 2. Variation of the overall error with the additional noise (Training: 400; Testing: 2944)

	0.0	±0.1	±0.2	±0.3	±0.5	±0.8	±1.0
MLR	2.41	+0.01	+0.05	+0.10	+0.25	+0.60	+0.88
LS-SVR	2.07	+0.01	+0.05	+0.11	+0.25	+0.61	+0.89

Summary and Future Work

- **The overall root mean square (RMS) error in the retrieved profiles of a testing dataset is remarkably smaller than the overall error using a multi-linear regression.**
- **The experiments of the variation of the training data show that for the small training dataset LS-SVR could obtain significantly more information from the sounding data than the method of the linear regression**
- **While the offset of 0.5 K or the noise of ±0.2 K is added to all channels simultaneously, the increase in the overall RMS error is less than 0.1 K**
- **Plan to use the infrared and microwave data to retrieve temperature and other geophysical parameters simultaneously.**

International TOVS Study Conference, 14th, ITSC-14, Beijing, China, 25-31 May 2005. Madison, WI, University of Wisconsin-Madison, Space Science and Engineering Center, Cooperative Institute for Meteorological Satellite Studies, 2005.