Estimating global solar radiation using wavelet and data driven techniques

Sungwon Kim*· Youngmin Seo**

ABSTRACT

The objective of this study is to apply a hybrid model for estimating solar radiation and investigate their accuracy. A hybrid model is wavelet-based support vector machines (WSVMs). Wavelet decomposition is employed to decompose the solar radiation time series into approximation and detail components. These decomposed time series are then used as inputs of support vector machines (SVMs) modules in the WSVMs model. Results obtained indicate that WSVMs can successfully be used for the estimation of daily global solar radiation at Champaign and Springfield stations in Illinois.

Keywords: support vector machines, wavelet decomposition, solar radiation, Illinois

INTRODUCTION

Solar radiation is the principal energy source for physical, biological and chemical processes, such as snow melt, plant photosynthesis, evaporation, crop growth and is also a variable needed for biophysical models to evaluate risk of forest fires, hydrological simulation models and mathematical models of natural processes. Solar radiation plays an important role in the design and analysis of energy efficient buildings in different climates. In cold and severe cold regions, passive solar designs and active solar systems help lower the reliance on conventional heating means using fossil fuels. In tropical and subtropical climates, solar heat gain is a major cooling load component, especially in cooling dominated buildings. The effects of prevailing climate and local topography would determine the actual amount of solar radiation reaching a particular location. The objective of the present study is to develop SVMs and WSVMs models that can be used to estimate daily solar radiation at two locations (Champaign and Springfield stations) in Illinois.

SUPPORT VECTOR MACHINES

The SVMs model has found wide applications in several areas, including pattern recognition, regression, multimedia, bio-informatics, and artificial intelligence. The SVMs model is a new kind of classifier that is motivated by two concepts. First, transformation of data into a high-dimensional space can transform complex problems into simpler problems that can use linear discriminant functions. Second, the SVMs model is motivated by the concept of training and uses only those inputs that are near the decision surface (Principe *et al*. 2000; Tripathi *et al*. 2006; Vapnik 2010). The solution of traditional neural networks models may tend to fall into a local optimal solution, whereas global

* Corresponding author, Ph.D./P.E., Associate Professor, Department of Railroad and Civil Engineering, Dongyang University, Yeongju, Republic of Korea, 750-711 (E-mail : swkim1968@dyu.ac.kr)

** Adjunct Professor, Department of Constructional Disaster Prevention Engineering, Kyungpook National University, Sangju, Republic of Korea, 742-711 (E-mail: ymseo@knu.ac.kr)

optimum solution is guaranteed for the SVMs model (Haykin 2009). The current study uses an ε -support vector regression (ε-SVR) model. It has been successfully applied for modeling hydrological processes (Tripathi *et al.* 2006; Kim *et al.* 2012; 2013a, b). During the ε-SVR model training performance, the purpose is to find a nonlinear function that minimizes a regularized risk function. This is achieved for the least value of the desired error criterion (e.g., RMSE) for various constant parameters C_C, and ε and various kernel functions with various constant σ values. Detailed information on the SVMs model can be found in Vapnik (2010), Principe *et al.* (2000), Tripathi *et al.* (2006), and Kim *et al.* (2012, 2013a, b).

WAVELET DECOMPOSITION

Wavelet analysis is a multi-resolution analysis in time and frequency domains. The wavelet transform decomposes a time series signal into different resolutions by controlling scaling and shifting. It provides a good localization properties in both time and frequency domains (Nejad and Nourani 2012). It also has an advantage in that it has flexibility in choosing the mother wavelet, which is the transform function, according to the characteristics of time series.

A fast DWT algorithm, developed by Mallat (1989), is based on four filters, including decomposition low-pass and high-pass, reconstruction low-pass and high-pass filters. For practical implementation of Mallat's algorithm, low-pass and high-pass filters are used instead of father and mother wavelets, which are also called scaling and wavelet functions, respectively. The low-pass filter, associated with the scaling function, allows the analysis of low frequency components, while the high-pass filter, associated with the wavelet function, allows the analysis of high frequency components. These filters, used in Mallat's algorithm are determined according to the selection of mother wavelets (González-Audícana *et al.* 2005). Multiresolution analysis by Mallat's algorithm is a procedure to obtain 'approximations' and 'details' for a given time series signal. An approximation holds the general trend of the original signal, while a detail depicts high-frequency components of it. A multilevel decomposition process (Figure 1) can be achieved, where the original signal is broken down into lower resolution components (Catalão *et al.* 2011). Detailed information for Mallat's algorithm can be found in Nason (2010).

CASE STUDY

The daily weather data obtained from two weather stations, Champaign (latitude, 40.0840° N; longitude, 88.2404° W; altitude, 219 m) and Springfield (latitude, 39.7273° N; longitude, 89.6106°W; altitude, 177 m) operated by the Illinois State Water Survey (ISWS), were used in this study (http://www.isws.illinois.edu/warm/).

The ISWS is a division of the Prairie Research Institute of the University of Illinois at Urbana-Champaign and has flourished for more than a century by anticipating and responding to new challenges and opportunities to serve the citizens of Illinois. The weather data consisted of six years (January 2007 to December 2012, N=2,192 days) of daily records of air temperature (TEM), solar radiation (RAD), relative humidity (HUM), dew point temperature (DEW), wind speed (WIN), and potential evapotranspiration (ETO). Air temperature and relative humidity have been measured at 2 m above the ground, whereas wind speed has been measured at 10 m above the ground (prior to winter 2011/2012 measurement made at 9.1 m). Potential evapotranspiration has been calculated using the Food and Agricultural Organization (FAO) of the United Nations Penman–Monteith equation as outlined in FAO Irrigation and Drainage Paper No. 56 "Crop Evapotranspiration" (Allen et al. 1998) since 1 December 2012 (Water and Atmospheric Resources Monitoring Program 2011). Prior to that time, the

Van Bavel method was used for calculating potential evapotranspiration (Van Bavel 1956). Figure 2 shows the comparison of the observed and estimated daily solar radiation values using SVMs and WSVMs models (three input).

CONCLUSIONS

This study develops and evaluates data-driven models for estimating daily solar radiation at Champaign and Springfield stations in Illinois. The SVMs and WSVMs models are developed for the three input combination. Results obtained indicate that WSVMs model can successfully be used for the estimation of daily global solar radiation at Champaign and Springfield stations in Illinois. In this study, it can be found that the data-driven models can estimate daily solar radiation.

<References>

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration guidelines for computing crop water requirements. FAO Irrigation and Drainage. Paper No. 56. Food and Agriculture Organization of the United Nations, Rome.

Catalão, J.P.S., Pousinho, H.M.I., Mendes, V.M.F., 2011. Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting. IEEE Trans. Power Syst. 26 (1), 137-144.

González-Audícana, M., Otazu, X., Fors, O., Seco, A., 2005. Comparison between Mallat's and the 'à trous' discrete wavelet transform based algorithms for the fusion of multispectral and panchromatic images. Int. J. Remote Sens. 26 (3), 595-614.

Haykin, S., 2009. Neural networks and learning machines, 3rd ed.PrenticeHall,NJ,USA.

Kim, S., Shiri, J., Kisi, O., 2012. Pan evaporation modeling using neural computing approach for different climatic zones. Water Resour. Manag. 26 (11), 3231-3249.

Kim, S., Seo, Y., Singh, V.P., 2013a. Assessment of pan evaporation modeling using bootstrap resampling and soft computing methods. J. Comput. Civ. Eng. doi:10.1061/(ASCE)CP. 1943–5487.0000367.

Kim, S., Shiri, J., Kisi, O., Singh, V.P., 2013b. Estimating daily pan evaporation using different data-driven methods and lag-time patterns. Water Resour. Manag. 27 (7), 2267-2286.

Mallat, S.G., 1989. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. Pattern. Anal. Mach. Intell. 11 (7), 674-693.

Nason, G., 2010. Wavelet methods in statistics with R. Springer, NY, USA.

Nejad, F.H., Nourani, V., 2012. Elevation of wavelet denoising performance via an ANN-based streamflow forecasting model. Int. J. Comput. Sci. Manag. Res. 1 (4), 764-770.

Principe, J.C., Euliano, N.R., Lefebvre, W.C., 2000. Neural and adaptive systems: fundamentals through simulation. Wiley, John & Sons, Inc., NY, USA.

Tripathi, S., Srinivas, V.V., Nanjundish, R.S., 2006. Downscaling of precipitation for climate change scenarios: a support vector machine approach. J. Hydrol. 330 (3-4), 621–640.

Vapnik, V.N., 2010. The nature of statistical learning theory, 2nd Edition.Springer,NY,USA.

Van Bavel, C.H.M., 1956. Estimating soil moisture conditions and time for irrigation with the evapotranspiration method. USDA, ARS 41-11, U.S. Dept. of Agric., Raleigh, NC, 1-16.

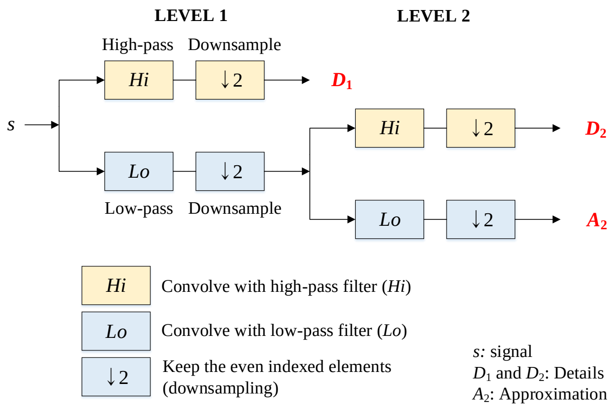
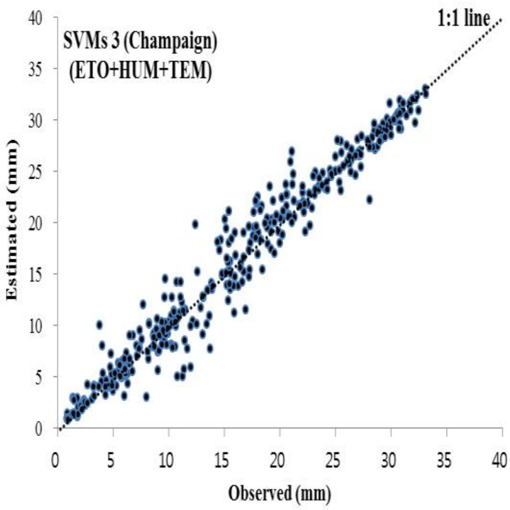
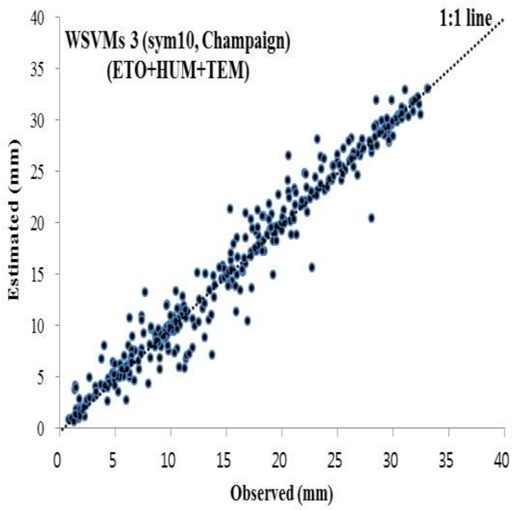


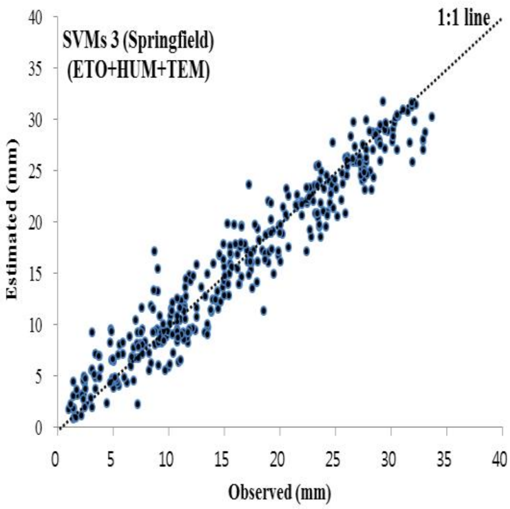
Figure 1 Mallat's algorithm for two-level decomposition of a signal



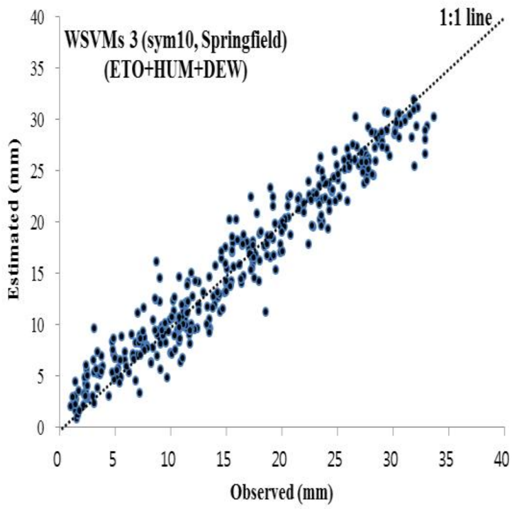
(a) SVMs 3 (Champaign)



(b) WSVMs 3 (Champaign)



(c) SVMs 3 (Springfield)



(d) WSVMs 3 (Springfield)

Figure 2 Comparison of the observed and estimated solar radiation values using SVMs and WSVMs models