Open access • Journal Article • DOI:10.1109/LGRS.2014.2355871

Upscaling Sensible Heat Fluxes With Area-to-Area Regression Kriging

Yong Ge, Yongzhong Liang, Jianghao Wang, Qianyi Zhao ...+1 more authors

Institutions: Beijing Normal University

Published on: 01 Mar 2015 - IEEE Geoscience and Remote Sensing Letters (IEEE)

Topics: Variogram, Sensible heat, Kriging and Scintillometer

Upscaling Sensible Heat Fluxes With Area-to-Area Regression Kriging

Yong Ge, *Member, IEEE*, Yongzhong Liang, Jianghao Wang, Qianyi Zhao, and Shaomin Liu

***Abstract*—Surface sensible heat flux (SHF) is a critical indicator for understanding heat exchange at the land–atmosphere interface. A common method for estimating regional SHF is to use ground observations with approaches such as eddy correlation (EC) or the use of a large aperture scintillometer (LAS). However, data observed by these different methods might have an issue with different spatial supports for cross-validation and comparison. This letter utilizes a geostatistical method called area-to-area regression kriging (ATARK) to solve this problem. The approach is illustrated by upscaling SHF from EC to LAS supports in the Heihe River basin, China. To construct a point support variogram, a likelihood function of four parameters (nugget, sill, range, and shape parameters) conditioned by EC observations is used. The results testify to the applicability of ATARK as a solution for upscaling SHF from EC support to LAS support.**

***Index Terms*—Area-to-area, kriging, sensible heat flux (SHF), upscaling.**

I. Introduction

SENSIBLE heat flux (SHF) is the process by which heat energy is transferred from the Earth's surface to the atmosphere by conduction and convection; it is a critical indicator for understanding heat exchange at the land–atmosphere interface. Many applications in meteorology, climatology, hydrology, and agriculture require estimates of SHF [1]. Currently, common methods for estimating SHF at the local, regional, and global scales can be categorized by one of four types. The first type uses remote sensing data [2], and the second makes use of ground observations [3]. The third and fourth types use land surface modeling and land data assimilation, respectively [4]. Such SHFs observed or estimated by these different methods might have different spatial supports. The issue of different spatial supports causes scaling conversion, which can occur in the same or different groups, for the purpose of validation, cross-validation, comparison, or subsequent analysis [5]. For example, for ground observations, the measurements of SHF commonly include the Bowen ratio, eddy correlation (EC), and lysimeter [6], [7]. These approaches can observe fluxes at ten to tens of thousands of square meters around a station.

This letter focuses on scaling conversions of SHF ground observations, particularly upscaling from EC to large aperture scintillometer (LAS) supports. In particular, LAS measurements are integrated over a long transect of approximately 500–5000 m from the same or different underlying surfaces, whereas EC measurements are normally a few to hundreds of meters [1], [8]–[14]. Compared with the support provided by LAS, EC support can be considered over a small area. Both supports change with time, and their shapes are irregular. More often, area-weighted [1], [15] and footprint-weighted [16] methods are used to investigate the scaling issue, with reasonable results. However, implementing these methods commonly involves physical process models or prior knowledge, which requires that the user is familiar with these. These methods do not consider spatial heterogeneity over complex terrain and only provide approximations as irregular footprints [4]. In the geostatistics field, upscaling EC observations to LAS support can be handled by a method called area-to-area kriging (ATAK) [17]–[21], which considers the spatial autocorrelation and supports of areal data to obtain the best linear unbiased estimation. However, different underlying surfaces produce different SHFs, which might lead to the nonstationarity of a random function. So it is not reasonable to upscale directly with ATAK. Because covariates contain information about the dependent variable, regression kriging [22] can produce more accurate predictions than ordinary kriging. Moreover, an assumption of stationarity for the regression kriging residual is usually more realistic than an assumption of stationarity for the target variable itself. Therefore, we propose to initially use a regression approach with environmental variables related to SHF and then apply ATAK to the residuals. This method is termed area-to-area regression kriging (ATARK). It is applied to the Heihe Watershed Allied Telemetry Experimental Research (HiWATER) experiment. HiWATER was launched by the National Natural Science Foundation of China in 2010 [8]–[10], [23] to validate remote sensing models, algorithms, and products. Furthermore, the Multi-scale Observation Experiment on Evapotranspiration over the heterogeneous land surfaces of HiWATER was undertaken from May to September 2012 [8], [10].

II. Materials and Methods

A. Study Area and Data Description

The Heihe River basin lies in an arid region of northwestern China. It was selected as an experimental watershed for conducting HiWATER because it is a typical inland river basin that has long served as a test region for integrated watershed studies and land surface or hydrologic experiments [23]. The

Manuscript received January 30, 2014; revised May 30, 2014; accepted June 25, 2014. This work was supported in part by the 863 High Technology of China under Grant 2012AA12A305 and in part by the National Natural Science Foundation of China under Grants 40971222 and 91125002.

Y. Ge, Y. Liang, and J. Wang are with the State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China (e-mail: gey@lreis.ac.cn).

Q. Zhao and S. Liu are with the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing 100875, China.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2014.2355871

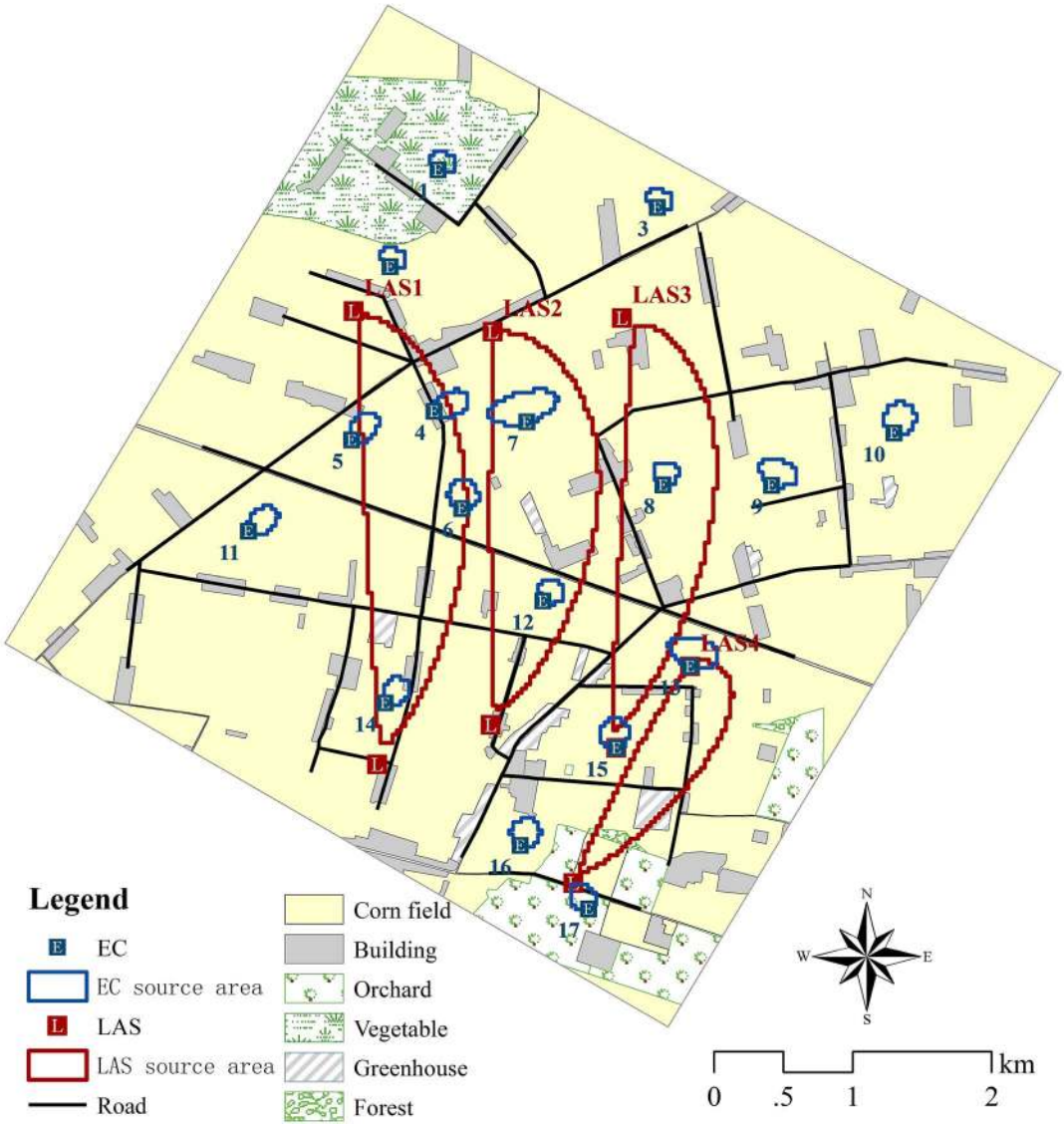

Fig. 1. Yingke-Daman irrigation district.

study area here is the Yingke-Daman irrigation district (see Fig. 1) in the middle reaches of the Heihe River. This district is between 38°50′–38°54′ N and 100°19′–100°24′ E. The terrain slopes gently downward from west to east, with elevations ranging from 1560 to 1447 m. Cornfields cover most of the area, but land cover also includes buildings, roads, orchards, greenhouses, forests, and other vegetation. The core experimental area covers a roughly 5.5 by 5.5 km region, centered near 100°22′ E, 38°52′ N. There were four LAS system groups (eight sets, with two sets in each group) installed within 3×3 and 2×1 Moderate Resolution Imaging Spectroradiometer pixels, and 17 EC systems are installed within the core experimental area. Details concerning specific performances and intercomparisons are found in [9]. Spatial distributions of EC and LAS systems are shown in Fig. 1: EC1 is in a vegetable field, EC4 is on building ground, EC17 is in an orchard, and the others are in cornfields. Different underlying surfaces produce different SHFs, particularly in daylight. SHF observed on building ground is much greater than SHFs observed in fields (see Fig. 2). EC footprints were calculated with a Eulerian analytic flux footprint model [24], whereas LAS footprints were calculated by combining the path-weighting function of LAS [12] with a Eulerian analytic flux footprint model [10], [25]. Footprints will change over time because they are affected by such factors as instrument height, atmospheric stability, wind speed, and wind direction. The small circles in Fig. 1 represent the source areas (the size and extent of footprints) of EC, and the large circles covering different underlying surfaces represent the source areas of LAS. The total flux contribution of the chosen total source area was set at 95%. Our objective is to upscale SHF from EC support to LAS support and then make a comparison with the observations of LAS.

Data used include the mean values of SHF observations and footprints of 17 EC and 4 LAS between 12:00 and 12:30 Beijing standard time (BST) and the mean values of wind speed observations of 17 EC between 12:00 and 12:30 BST. These EC and LAS data were preprocessed. Corresponding methods and quality control procedures are given in [25]. Land Surface Temperature (LST), Normalized Difference Vegetation Index (NDVI), and Fractional Vegetation Cover (FVC) of the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) product [26] are from about 12:15 BST. We

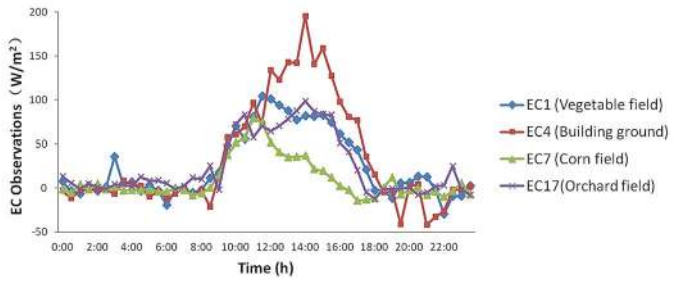

Fig. 2. EC SHF observations of different underlying surfaces.

chose observations from nine periods, i.e., June 15, June 24, July 10, August 2, August 11, August 18, August 27, September 3, and September 12, 2012, when the weather was clear and ASTER satellite passes were available for the study area. The resulting remote sensing images captured on these days had high quality.

B. Area-to-Area Regression Kriging

The ATARK rescaling procedure was divided into four steps according to the situation in our study area.

Step 1. Trend Modeling: We first extract the auxiliary environmental variables using the footprints of 17 EC systems from nine periods. Then, a linear multiple regression model is established

$$H_{it}=\beta_{0t}+\beta_{1t}\cdot\overline{FVC_{it}}+\beta_{2t}\cdot\overline{LST_{it}}+\beta_{3t}\cdot\overline{NDVI_{it}}+\beta_{4t}\cdot\overline{WS_{it}} \quad (1)$$

where H_{it} is the SHF of the ith EC during the tth period and $\overline{FVC_{it}}$, $\overline{LST_{it}}$, $\overline{NDVI_{it}}$, and $\overline{WS_{it}}$ are auxiliary environmental variables extracted by the ith EC footprints during the tth period. We can achieve full coverage of wind speed over the study area using simple interpolation such as ordinary kriging. Stepwise regression would be adopted in case of collinearity.

Step 2. Derivation of Point Variogram: After the regression procedure, we obtained 17 residuals (EC support) and one SHF trend for each period. Based on the assumption of stationarity, residuals of LAS support can be estimated with ATAK. Before the application of ATAK, the most important—and difficult—procedure is to calculate the point support covariance, or equivalently the point support variogram, of the residuals. One popular method is the deconvolution procedure [27], which can derive a point support variogram based on the variogram calculated by the areal observations using their centroids. Because only 17 EC support residuals are available during a single period, the centroid-based variogram would be unstable, which could cause unreliable deconvolution results. In this letter, we use restricted maximum likelihood (REML) estimation to assess the point support variogram. The solution is to assume the second-order stationarity of the residual at the point support and use four parameters (nugget, sill, range, and shape parameters) to characterize a Matérn variogram [28]. After that, we can calculate the area-to-area covariance of the EC support using

$$\bar{C}(v_i,v_j)=\frac{1}{N(v_i)}\frac{1}{N(v_j)}\sum_{k=1}^{N(v_i)}\sum_{l=1}^{N(v_j)}C(p_k,p_l), \quad p_k\in v_i, p_l\in v_j \quad (2)$$

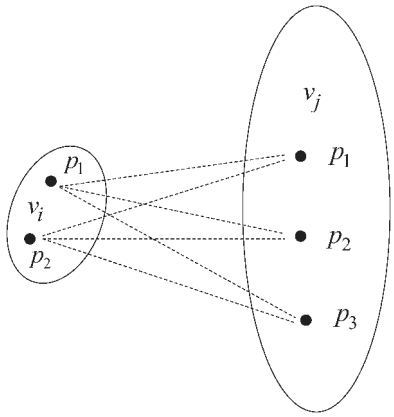

Fig. 3. Area-to-area covariance between any two areas.

where $\bar{C}(v_i,v_j)$ is the area-to-area covariance between any two areas v_i and v_j (see Fig. 3). $N(v_i)$ and $N(v_j)$ are the number of points used to discretize the two areas v_i and v_j, respectively. Fig. 3 is just a sketch map. In practice, there are many more points to be used to discretize the areas. Equation (2) means that $\bar{C}(v_i,v_j)$ is computed between any two points discretizing the areas v_i and v_j.

Assuming that the target variable follows a normal distribution and $\boldsymbol{\theta}$ is the vector of four parameters of the Matérn variogram, the likelihood function conditioning on the EC support residuals is given by

$$L(\boldsymbol{\theta}|\bar{Z}_v)=2\pi^{-\frac{N}{2}}|\mathbf{C_{vv}}|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\bar{Z}_v^T\mathbf{C_{vv}}^{-1}\bar{Z}_v\right) \quad (3)$$

where $\bar{Z}_v$ corresponds to 17 EC support residuals and $\mathbf{C_{vv}}$ is the area-to-area covariance matrix calculated using (2). With the aforementioned equation, we can derive the point support variogram model. Next, ATAK will be used to upscale the residuals from EC support to LAS support.

Step 3. Prediction and Error Variance of Residual at LAS Support: Assuming the second-order stationary of the residual, the ATAK value over v_0 was estimated as a linear combination of K neighboring units

$$z^*(v_0)=\sum_{i=1}^{K}\lambda_i(v_0)z(v_i) \quad (4)$$

where v_0 is the unknown area support to be estimated and v_i is the area support of the regression residual. In this letter, v_0 and v_i represent the supports of LAS and EC, respectively. $\lambda_i(v_0)$ is the weight assigned to $z(v_i)$ for the prediction at v_0. The K weights are the solution of the following system of linear equations:

$$\begin{cases}\sum_{j=1}^{K}\lambda_j(v_0)\bar{C}(v_i,v_j)+\mu(v_0)=\bar{C}(v_i,v_0), & i=1,2,\ldots,K\\ \sum_{j=1}^{K}\lambda_j(v_0)=1.\end{cases} \quad (5)$$

The ATAK prediction error variance for v_0 is computed as

$$\delta(v_0)=\bar{C}(v_0,v_0)-\sum_{i=1}^{K}\lambda_i(v_0)C(v_i,v_0)-\mu(v_0). \quad (6)$$

TABLE I
SIGNIFICANT TEST FOR COEFFICIENTS AND EQUATION

	R	P					
	Square	Constant	FVC	LST	NDVI	WS	Model
15 Jun	0.404	0.000	0.006	-	-	-	0.006
24 Jun	0.596	0.000	-	-	0.000	-	0.000
10 Jul	0.878	0.000	-	-	0.000	-	0.000
2 Aug	0.952	0.044	0.033	0.024	-	-	0.000
11 Aug	0.411	0.000	-	-	0.006	-	0.006
18 Aug	0.391	0.009	-	0.007	-	-	0.007
27 Aug	0.925	0.020	0.003	0.012	0.030	-	0.000
3 Sep	0.374	0.109	-	-	-	0.096	0.096
12 Sep	0.386	0.001	0.096	-	0.043	-	0.092

Step 4. Obtaining the Estimation: Finally, the estimated value of LAS supports can be obtained by adding the regression estimations (step 2) to the ATAK results (step 3).

III. RESULTS AND DISCUSSION

Based on the data and methods described in Section II, environmental variables were extracted using EC footprints. A stepwise procedure was used for selecting environmental variables, and the regression models and regression coefficients were tested for significance. As shown in Table I, most of the regression coefficients and models are considered significant. The symbol "-" means that the variable has been excluded from the regression model by the stepwise method.

To derive a point support variogram in step 2, EC source areas must be discretized first. Then, using (2) and (3), the point support variograms of SHF residuals at nine periods were estimated using REML (see Fig. 4).

Although a point support variogram is estimated using REML here, Goovaerts [27], Truong and Heuvelink [28], and Nagle *et al.* [29] have pointed out that area support observations retain little information to infer the nugget component of a point support variogram. Fortunately, area-to-area covariance will not be influenced by the nugget variance because nugget variability cancels out for both areas. With the point variogram models available, SHF estimations of LAS supports were derived according to steps 3 and 4. A comparison of SHF observations and estimations of LAS supports is shown in Fig. 5.

Assuming no uncertainty in the regression trend of SHF, ATAK prediction error variances of LAS support were calculated and showed as a 95% confidence interval. The dashed lines are models forced to pass through the origin and fit to the points. The results of upscaling for LAS1, LAS2, and LAS4 are close to these observations, with coefficients of determination larger than 0.89 and the slope of the dashed line close to 1. One of the main reasons is that the underlying surfaces of LAS1, LAS2, and LAS4 are relatively homogeneous, which means that the SHF observations of EC and LAS are similar. However, LAS3 was obviously underestimated. Although the coefficient of determination was greater than 0.8, the slope of the dashed line was less than 0.79. All of the upscaling results for LAS3 were smaller than the observations. The main reason for this is that the underlying surface of LAS3 is more complex, particularly with more buildings within the source area, which produced higher SHF observations, but the EC systems could not capture completely. Meanwhile this higher SHF contributed a lot and could be observed by LAS3. These smaller measurements from the EC systems might cause underestimations in the regression step and then lead to smaller upscaling results than the LAS3 measurements.

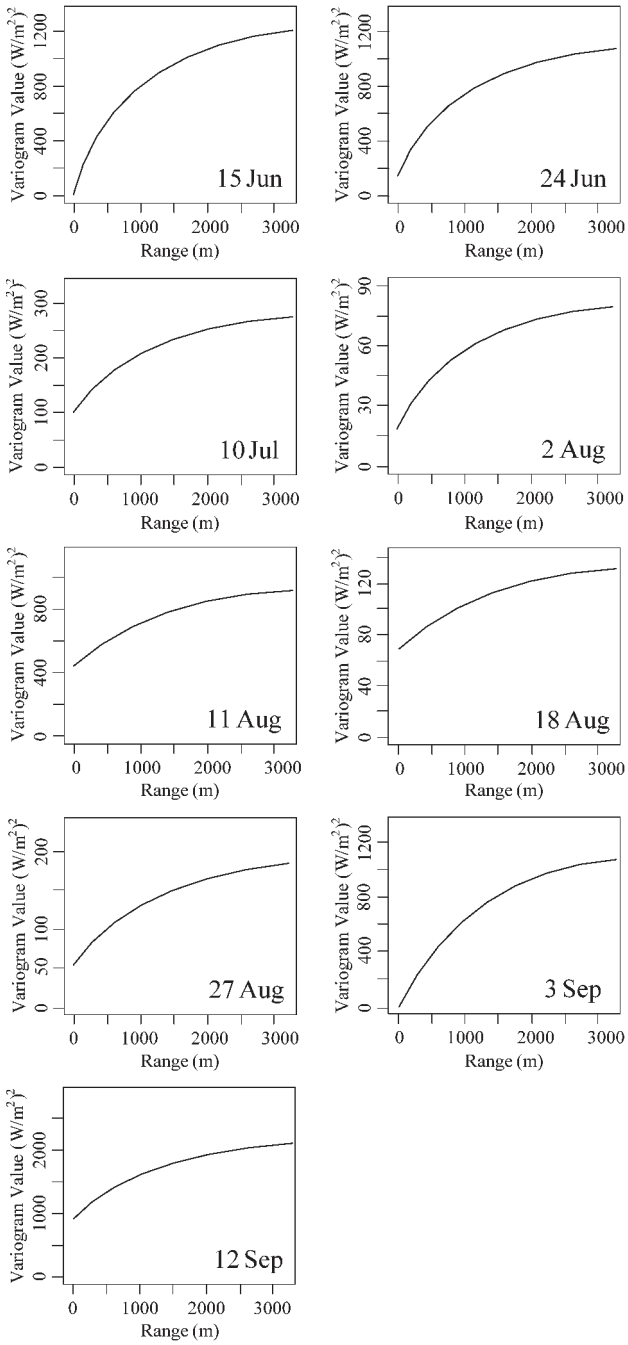

Fig. 4. Point support variograms for nine time periods in 2012.

IV. CONCLUSION

Because the accuracy of different SHF measuring methods can be assessed by scaling issues, upscaling ground-based SHF measurements is of great importance for the validation, cross-validation, comparison, or subsequent analysis of these methods. In this letter, the ATARK method has been presented as a way to upscale SHF from EC support to LAS support. The method has been illustrated using data from the HiWATER

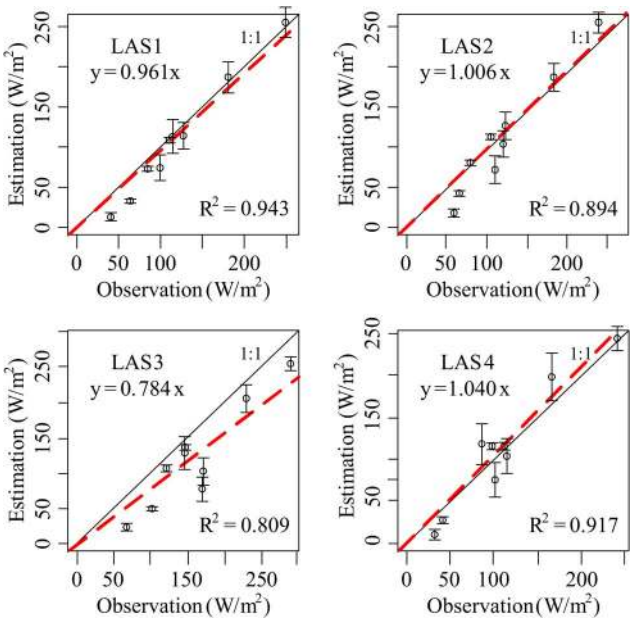

Fig. 5. SHF comparison of observations and estimations of LAS supports.

experiment to upscale 17 EC SHFs to LAS supports. By incorporating environmental covariates in the regression procedure, ATARK can produce more accurate predictions, and the assumption of stationarity of regression residuals is more realistic. The comparison between SHF estimations and observations of LAS supports testifies to the suitability of ATARK as a method for upscaling, except for the underestimations caused by a highly heterogeneous underlying surface.

Acknowledgment

The authors would like to thank three anonymous reviewers and the editor for their precise and valuable comments that greatly improved this letter.

References

[1] F. Beyrich, H. De Bruin, W. Meijninger, J. Schipper, and H. Lohse, "Results from one-year continuous operation of a large aperture scintillometer over a heterogeneous land surface," *Boundary-Layer Meteorol.*, vol. 105, no. 1, pp. 85–97, Oct. 2002.

[2] Y. Song *et al.*, "A revised surface resistance parameterisation for estimating latent heat flux from remotely sensed data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 17, pp. 76–84, Jul. 2012.

[3] S. Liu *et al.*, "Measurement, analysis and application of surface energy and water vapor fluxes at large scale," *Adv. Earch Sci.*, vol. 125, no. 11, pp. 1113–1127, Nov. 2010.

[4] T. Xu, S. Liang, and S. Liu, "Estimating turbulent fluxes through assimilation of geostationary operational environmental satellites data using ensemble Kalman filter," *J. Geophys. Res., Atmos.*, vol. 116, no. D9, pp. D09109-1–D09109-16, May 2011.

[5] R. P. Haining, *Spatial Data Analysis: Theory and Practice*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[6] D. Baldocchi *et al.*, "FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities," *Bull. Amer. Meteorol. Soc.*, vol. 82, no. 11, pp. 2415–2434, Nov. 2001.

[7] W. Massman and X. Lee, "Eddy covariance flux corrections and uncertainties in long-term studies of carbon and energy exchanges," *Agric. Forest Meteorol.*, vol. 113, no. 1, pp. 121–144, Dec. 2002.

[8] Y. Ma *et al.*, "Estimations of regional surface energy fluxes over heterogeneous oasis-desert surfaces in the middle reaches of the Heihe River during HiWATER-MUSOEXE," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 671–675, Mar. 2015.

[9] Z. Xu *et al.*, "Intercomparison of surface energy flux measurement systems used during the HiWATER-MUSOEXE," *J. Geophys. Res., Atmos.*, vol. 118, no. 23, pp. 13 140–13 157, Dec. 2013.

[10] S. Liu *et al.*, "A comparison of eddy-covariance and large aperture scintillometer measurements with respect to the energy balance closure problem," *Hydrol. Earth Syst. Sci.*, vol. 15, no. 4, pp. 1291–1306, Apr. 2011.

[11] C. Watts *et al.*, "Comparison of sensible heat flux estimates using AVHRR with scintillometer measurements over semi-arid grassland in northwest Mexico," *Agric. Forest Meteorol.*, vol. 105, no. 1, pp. 81–89, Nov. 2000.

[12] W. Meijninger *et al.*, "Determination of area-averaged sensible heat fluxes with a large aperture scintillometer over a heterogeneous surface–Flevoland field experiment," *Boundary-Layer Meteorol.*, vol. 105, no. 1, pp. 37–62, Oct. 2002.

[13] L. Jia *et al.*, "Estimation of sensible heat flux using the Surface Energy Balance System (SEBS) and ATSR measurements," *Phys. Chem. Earth, Parts A/B/C*, vol. 28, no. 1, pp. 75–88, Jan. 2003.

[14] Z. Jia, S. Liu, Z. Xu, Y. Chen, and M. Zhu, "Validation of remotely sensed evapotranspiration over the Hai River Basin, China," *J. Geophys. Res., Atmos.*, vol. 117, no. D13, pp. D13113-1–D13113-21, May 2012.

[15] L. Mahrt, D. Vickers, J. Sun, and J. H. McCaughey, "Calculation of area-averaged fluxes: Application to BOREAS," *J. Appl. Meteorol.*, vol. 40, no. 5, pp. 915–920, Sep. 2001.

[16] Ü. Rannik *et al.*, "Footprint analysis for measurements over a heterogeneous forest," *Boundary-Layer Meteorol.*, vol. 97, no. 1, pp. 137–166, Oct. 2000.

[17] P. C. Kyriakidis, "A geostatistical framework for area-to-point spatial interpolation," *Geograph. Anal.*, vol. 36, no. 3, pp. 259–289, Jul. 2004.

[18] X. Liu, P. C. Kyriakidis, and M. F. Goodchild, "Population-density estimation using regression and area-to-point residual kriging," *Int. J. Geograph. Inf. Sci.*, vol. 22, no. 4, pp. 431–447, Jul. 2008.

[19] R. Kerry, P. Goovaerts, B. G. Rawlins, and B. P. Marchant, "Disaggregation of legacy soil data using area to point kriging for mapping soil organic carbon at the regional scale," *Geoderma*, vol. 170, pp. 347–358, Jan. 2012.

[20] P. M. Atkinson, "Downscaling in remote sensing," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 22, pp. 106–114, Jun. 2012.

[21] C. A. Gotway and L. J. Young, "A geostatistical approach to linking geographically aggregated data from different sources," *J. Comput. Graph. Stat.*, vol. 16, no. 1, pp. 115–135, Jul. 2007.

[22] T. Hengl, G. Heuvelink, and D. G. Rossiter, "About regression-kriging: From equations to case studies," *Comput. Geosci.*, vol. 33, no. 10, pp. 1301–1315, Oct. 2007.

[23] X. Li *et al.*, "Heihe Watershed Allied Telemetry Experimental Research (HiWATER): Scientific objectives and experimental design," *Bull. Amer. Meteorol. Soc.*, vol. 94, no. 8, pp. 1145–1160, Aug. 2013.

[24] R. Kormann and F. X. Meixner, "An analytical footprint model for non-neutral stratification," *Boundary-Layer Meteorol.*, vol. 99, no. 2, pp. 207–224, May 2001.

[25] S. Liu, Z. Xu, Z. Zhu, Z. Jia, and M. Zhu, "Measurements of evapotranspiration from eddy-covariance systems and large aperture scintillometers in the Hai River Basin, China," *J. Hydrol.*, vol. 487, pp. 24–38, Apr. 2013.

[26] J. Zhou, J. Li, L. Zhang, D. Hu, and W. Zhan, "Intercomparison of methods for estimating land surface temperature from a Landsat-5 TM image in an arid region with low water vapour in the atmosphere," *Int. J. Remote Sens.*, vol. 33, no. 8, pp. 2582–2602, Oct. 2012.

[27] P. Goovaerts, "Kriging and semivariogram deconvolution in the presence of irregular geographical units," *Math. Geosci.*, vol. 40, no. 1, pp. 101–128, Jan. 2008.

[28] P. Truong and G. B. M. Heuvelink, "Bayesian area-to-point kriging using expert knowledge as informative priors," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 30, pp. 128–138, Aug. 2014.

[29] N. N. Nagle, S. H. Sweeney, and P. C. Kyriakidis, "A geostatistical linear regression model for small area data," *Geograph. Anal.*, vol. 43, no. 1, pp. 38–60, Jan. 2011.