Intercomparison of surface energy flux measurement systems used during the HiWATER-MUSOEXE

Ziwei Xu,[1] Shaomin Liu,[1] Xin Li,[2] Shengjin Shi,[3] Jiemin Wang,[2] Zhongli Zhu,[1] Tongren Xu,[1] Weizhen Wang,[2] and Mingguo Ma[2]

Received 22 May 2013; revised 18 November 2013; accepted 20 November 2013; published 13 December 2013.

[1] Agreement among instruments is very important for the Multi-Scale Observation Experiment on Evapotranspiration over heterogeneous land surfaces of The Heihe Watershed Allied Telemetry Experimental Research (HiWATER-MUSOEXE), particularly in regard to radiation and turbulent flux measurements. Before HiWATER-MUSOEXE was conducted, 20 eddy covariance (EC) system sets, 18 radiometer sets, and seven large aperture scintillometers (LASs) sets were intercompared over the Gobi desert between 14 and 24 May 2012. For radiometers, the four-component radiation measurements exhibited good agreement —the average root-mean-square error (RMSE) and mean relative error (MRE) for the net radiation were 10.38 W m^{-2} and 1.24%, respectively. With regard to the EC systems, the best consistency for sensible heat fluxes was found among CSAT3 sonic anemometers and Li7500A/Li7500/EC150 combinations (average RMSE, 12.30 W m^{-2} and MRE, −1.36%), followed by Gill sonic anemometers and Li7500A/Li7500 combinations when a proper angle of attack correction method was applied (average RMSE, 16.75 W m^{-2} and MRE, −5.52%). The sensible heat flux measured using different LASs agreed well with high correlation coefficients —the average RMSE and MRE values were 10.26 W m^{-2} and 5.48% for boundary layer scintillometer (BLS) 900, 16.32 W m^{-2} and 10.47% for BLS450, and 14.38 W m^{-2} and −3.72% for ZZLAS, respectively. The EC and LAS measurements were compared and agreed well over homogeneous underlying surfaces, which also indicated that the EC and LAS measurements would be comparable in the follow-up experiment. The intercomparison results can be used to determine instrument placement and are very helpful for subsequent data analysis.

Citation: Xu, Z., S. Liu, X. Li, S. Shi, J. Wang, Z. Zhu, T. Xu, W. Wang, and M. Ma (2013), Intercomparison of surface energy flux measurement systems used during the HiWATER-MUSOEXE, *J. Geophys. Res. Atmos.*, *118*, 13,140–13,157, doi:10.1002/2013JD020260.

1. Introduction

[2] The Heihe Watershed Allied Telemetry Experimental Research (HiWATER) program was designed as a comprehensive ecohydrological experiment within the framework of the Heihe Plan, based on the diverse needs of interdisciplinary studies of the research plan and the existing observation infrastructures in the Heihe River Basin [*Li et al.*, 2013]. The first thematic experiment launched in HiWATER was the Multi-Scale Observation Experiment on Evapotranspiration over heterogeneous land surfaces 2012 (HiWATER-MUSOEXE), which involved a flux observation matrix in the middle reach of the Heihe River Basin between May and September in 2012. HiWATER-MUSOEXE was composed of two nested matrices: one large experimental area (30 km × 30 km) and one kernel experimental area (5.5 km × 5.5 km). The large experimental area contained one superstation (within the oasis, cropland) and four ordinary stations (around the oasis), with underlying desert, desert steppe, Gobi, and wetland surfaces. The 5.5 km × 5.5 km kernel experimental area was located in the oasis. Moreover, 17 elementary sampling plots were divided according to the distribution of, for example, crop structure, windbreak, resident area, soil moisture, and irrigation status. There were also four large aperture scintillometer (LAS) groups (eight sets with two sets in each group) installed in 3 × 3 and 2 × 1 Moderate Resolution Imaging Spectroradiometer (MODIS) pixels within the kernel experimental area. Overall, 22 eddy covariance (EC) system sets, eight LAS sets, and 21 automatic weather station sets were involved in the experiment. HiWATER-MUSOEXE was focused on studying the spatial-temporal variations of evapotranspiration (ET), the effects of advection in the oasis-desert ecosystem (30 km × 30 km), the heterogeneity of ET in the irrigated oasis (5.5 km × 5.5 km), and the ET

[1]State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing, China.
[2]Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, China.
[3]College of Resources and Environmental Sciences, China Agricultural University, Beijing, China.

Corresponding author: S. Liu, State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, No.19 Xinjiekouwai St., Beijing 100875, China. (smliu@bnu.edu.cn)

2169-897X/13/10.1002/2013JD020260

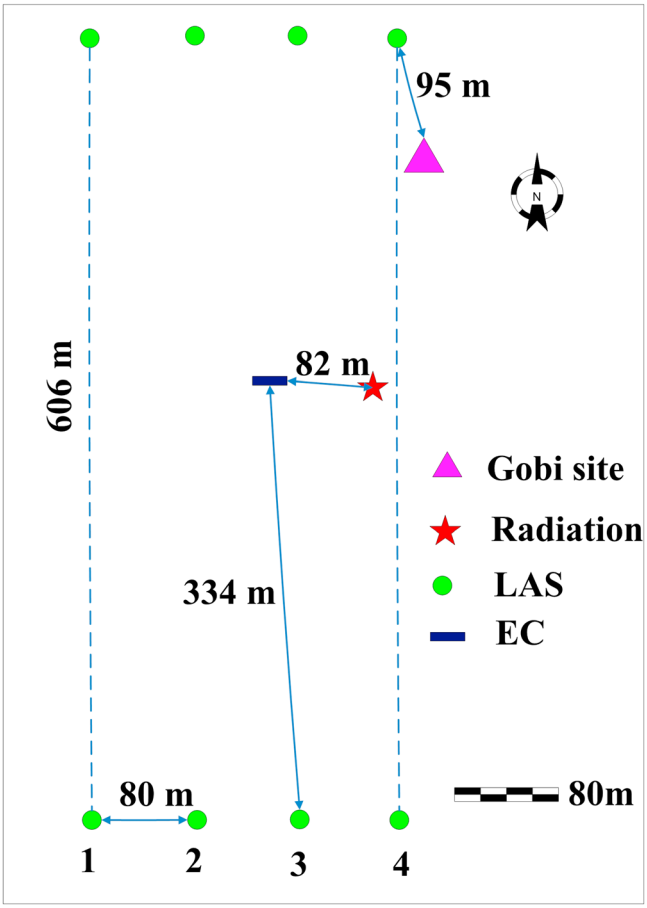

Figure 1. The layout of instruments in the intercomparison field.

acquisition at the pixel scale. A detailed description of HiWATER-MUSOEXE can be found in Liu, S.M., et al. (manuscript in preparation, 2013).

[3] To acquire reliable data, it is important to know how one instrument measurement differs from another (i.e., the measurement uncertainty) before an experiment is conducted, particularly regarding instruments that directly measure energy balance components and turbulent fluxes. Indeed, instrument intercomparisons are typically performed prior to or during large experiments. The First International Satellite Land Surface Climatology Project Field Experiment (FIFE) employed seven different designs of net radiometer from five different manufacturers (Precision Spectral Pyranometer (PSP) and Precision Infrared Radiometer (PIR), Eppley; Q3, Q4, and Q5, Radiation and Energy Balance Systems (REBS); Didcot; Swissteco; and Thornthwaite). *Field et al.* [1992] examined these radiometers at three sites via side by side comparisons for durations of one to three days and found that the net radiation regression slopes (deviation from the 1:1 regression line, the same as below) were as large as 5% to 7% in daytime for instruments from the same manufacturers and 10% to 15% between different manufacturers. After careful calibration, the root-mean-square error (RMSE) ranged from 15 to 35 W m^{-2} during the day. During FIFE, a radiometer intercomparison was also performed by *Nie et al.* [1992] using a mobile set of instruments at selected sites, finding a 10% mean relative error (MRE) in the daily averaged net radiation. Through several years of development, considerable effort has been focused on increasing radiation measurement accuracy [*Ohmura et al.*, 1998]. The radiometers in the Energy Balance Experiment (EBEX-2000) were intercompared by *Kohsiek et al.* [2007], who investigated shortwave radiometers (PSP, Eppley; CM11, CM14, and CM21, Kipp & Zonen; CNR1 and CM3, Kipp & Zonen), longwave radiometers (PIR, Eppley; CNR1 and CG3, Kipp & Zonen), and net radiometers (CNR1, Kipp & Zonen; Q7, REBS; Däke, Schulze) in a flat, irrigated cotton field. By plotting the daily course of differences between the measurement and reference (the involved instrument measurement minus the reference) for the four components, the difference of incoming/outgoing shortwave radiation was estimated at max (5 W m^{-2}, 1% of the reference value)/(5 W m^{-2}, 6%), respectively. For the incoming/outgoing longwave radiation, the difference was estimated to be 10 W m^{-2} (daytime) and 5 W m^{-2} (nighttime). The difference between the net radiation and the reference was up to 50 W m^{-2} or more during irrigation periods. Moreover, the difference of net radiation was estimated at max (25 W m^{-2}, 5%) during the day and 10 W m^{-2} at night if the correlated errors were not considered.

[4] *Nie et al.* [1992] investigated the surface energy flux measurement system intercomparison used during FIFE, showing that the average MRE of the latent heat flux was 20% for the 14 Bowen ratio (BR) system sets, four EC system sets, and the rover system (one BR was used as the reference). Currently, the EC system is the main instrument for directly measuring turbulent fluxes and is primarily composed of a fast response three-dimensional sonic anemometer and a fast response scalar sensor. *Mauder et al.* [2007] found that the impact of different instrumentation combinations in an EC system (e.g., USA-1 and KH20/LI-7500, CSAT3 and KH20, Solent-HS and KH20, and ATI-K and KH20) on the resulting heat flux estimates was very large. The regression slopes of sensible heat flux were between 0%

Table 1. The Information for the Radiometers Used[a]

No.	Type	Serial Number	Duration
		Line 1	
1	NR01	1242	16 May 15:50 to 23 May 19:50
2	CNR1	040890	16 May 15:50 to 23 May 19:50
3	CNR1	040772	16 May 15:50 to 23 May 19:50
4	CNR1	011489	16 May 15:50 to 23 May 19:50
5	CNR1	011490	16 May 15:50 to 23 May 19:50
6	CNR1	090255	16 May 15:50 to 23 May 19:50
7	CNR1	071488	16 May 15:50 to 23 May 19:50
8	CNR1	030623	16 May 15:50 to 23 May 19:50
9	CNR1	071494	16 May 15:50 to 23 May 19:50
		Line 2	
1	PSP and PIR	36995F3,36996F3 36165F3,36166F3	16 May 15:50 to 23 May 19:50
		Line 3	
1	Q7	10111	16 May 15:50 to 23 May 19:50
2	CNR4	100414	16 May 15:50 to 23 May 19:50
3	CNR4	110635	16 May 15:50 to 23 May 19:50
4	CNR4	110636	16 May 15:50 to 23 May 19:50
5	CNR4	110637	16 May 15:50 to 23 May 19:50
6	CNR4	110638	16 May 15:50 to 23 May 19:50
7	CNR4	100424	16 May 15:50 to 23 May 19:50
8	CNR4	110481	16 May 15:50 to 23 May 19:50

[a]The first line of nos.1–9 was from west to east; the third line of nos.1–8 was from west to east.

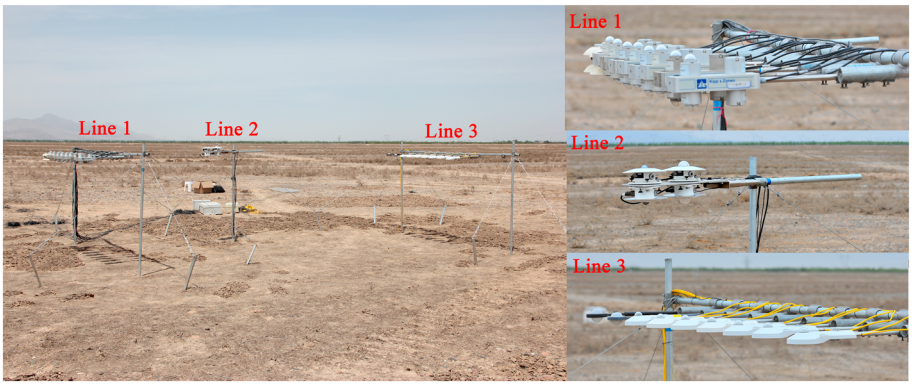

Figure 2. Pictures of the radiometers used in the intercomparison field.

and 22%, with an average of 7%, and between 1% and 45% for latent heat flux, with an average of 17%. The RMSE of the sensible heat flux ranged from 6.1 to 12.8 W m^{-2}, with an average of 8.65 W m^{-2}; the range was 35.7 to 115.0 W m^{-2} for the latent heat flux, with an average of 59.02 W m^{-2} [see *Mauder et al*., 2007, Tables 9 and 10]. For some sensor combinations, the maximum regression slope was as high as 50%. However, after proper calibration and maintenance, a regression slope of less than 5% for sensible and latent heat flux measurements could be achieved for some sensor combinations. *Mauder et al*. [2007] also demonstrated that data postprocessing methods were very important for the final fluxes, and the different methodologies could result in deviations of up to 10% and 15% for 30 min average sensible and latent heat fluxes, respectively.

[5] There are only a few worldwide large aperture scintillometer (LAS) comparison experiments. *Kleissl et al*. [2008] discussed the discrepancies in the sensible heat flux measured using LASs, reporting regression slopes between 6% and 21% for LASs produced by Kipp & Zonen. *Kleissl et al*. [2009] also compared the boundary layer scintillometer (BLS) series LASs from Scintec, reporting regression slopes of less than 3%.

[6] The EC technique is generally considered the most accurate method for measuring fluxes and is used to assess the accuracy of sensible heat fluxes measured using LASs, by investigating the discrepancy between flux values, particularly on flat and homogeneous terrain. *Zeweldi et al*. [2010] compared the sensible heat flux measurements between one LAS and two EC in dry conditions and found that the measurements were very similar (the LAS measurements were overestimated by 6% and 2% compare with the two EC measurements). *Liu et al*. [2011] investigated the causes of the discrepancy between the LAS and EC measurements in an alpine meadow in the Heihe River Basin and reported that the LAS measurements were overestimated by 6% compared to the EC measurements when the EC energy balance closure was greater than 0.75 and the LAS and EC measurement values greater than 50 W m^{-2}. The heterogeneity of the underlying surfaces, the different source areas of LAS and EC measurements and the energy imbalance of the EC system, were the primary reasons for the discrepancy.

[7] The above intercomparisons revealed that discrepancies exist with measurements from the same and different sensor types. Therefore, it is necessary to analyze and characterize the discrepancies between the instrument systems prior to conducting HiWATER-MUSOEXE.

[8] An instrument's intercomparison experiment (20 EC sets, 18 radiometer sets, and 7 LAS sets) was performed at the Gobi site near Zhangye city, Gansu Province, China, on flat and homogeneous terrain. The instruments used in the intercomparison field were the primary instruments used to measure global surface energy fluxes. Some instruments, e.g., CSAT3 and Li7500A, CSAT3 and EC150, and BLS900, have only recently begun commercial manufacturing, making the intercomparison more encouraging. The purpose was to assess the consistency of the instruments (i.e., radiometers, ECs, and LASs) used in the flux observation matrix during HiWATER-MUSOEXE to enable the comparison of flux measurements from different instruments at different sites and estimate the probable site discrepancy considering these comparisons. A description of the intercomparison is introduced in section 2; the data processing method is discussed in section 3. Section 4 presents the results, including the (i) meteorological conditions and a comparison of (ii) radiation sensors, (iii) eddy covariance sensors, (iv) large aperture scintillometers, and (v) the sensible heat fluxes derived from the EC and LAS measurements. A discussion and conclusions are provided in section 5.

2. Experimental Description

[9] The field selected for the intercomparison campaign was located in the Baji Gobi desert, west of Zhangye city (100°18′15.17″E; 38°54′53.87″N). The Baji Gobi desert is nearly flat, open land covered by coarse grain sand and small pebbles with withered sparse scrub vegetation. The intercomparison experiment was performed from 14 to 24 May in 2012. An area of 300 m × 600 m was selected as the field site—four LAS groups were located along the north and south sides (green dots in Figure 1), 20 EC sets were located in the center (blue rectangles area in Figure 1), and 18 radiometer sets were located in the east-central region (red five-pointed star in Figure 1). A 10 m meteorological tower was constructed in the northeast region of the field site, including two heights for wind speed (010C-1, 5 m and 10 m) and air temperature and humidity (HMP45AC, 5 m and 10 m) measurements. Moreover, wind direction (020C-1, 10 m), the rain gage (TE525MM), and air pressure (PTB110) were measured at one height. Soil temperature profile (AV-10T, 0, 2, 4, 10, 20, 40, 60, 100 cm), soil moisture profile (ECH_2O-5, 2, 4,

Table 2. The Information for the Eddy Covariance Systems Used[a]

No.	Type	Serial Number	Data Logger	Duration
1	CSAT3 and Li7500	0859-2/75H-0118	CR5000	14 May 18:00 to 24 May 13:00
2	CSAT3 and Li7500	0854-2/75H-0781	CR5000	14 May 18:00 to 24 May 13:00
3	CSAT3 and Li7500A	2309/75H-2099	CR1000	14 May 17:30 to 24 May 13:00
4	CSAT3 and Li7500	7871/75H-544	CR3000	14 May 17:00 to 24 May 13:30
5	CSAT3 and Li7500A	2311/75H-2224	CR1000	18 May 19:30 to 24 May 13:00
6	CSAT3 and Li7500A	6907/75H-2210	CR3000	17 May 18:00 to 24 May 13:00
7	CSAT3 and Li7500	1525/75H-0347	CR3000	16 May 11:00 to 24 May 13:00
8	CSAT3 and Li7500	2169/75H-1769	CR3000	14 May 17:30 to 24 May 13:00
9	CSAT3 and Li7500	0817-2/75H-0559	CR3000	16 May 11:00 to 24 May 13:00
10	CSAT3 and Li7500	0818-2/75H-0575	CR3000	16 May 11:00 to 24 May 13:00
11	CSAT3 and Li7500	1101-2/75H-0901	CR5000	14 May 18:30 to 24 May 13:30
12	CSAT3 and Li7500	1585/75H-1465	CR5000	14 May 18:30 to 24 May 13:00
13	CSAT3 and Li7500A	2257/75H-2132	CR3000	14 May 19:00 to 24 May 13:00
14	CSAT3 and Li7500	0929/75H-0698	CR5000	14 May 18:00 to 24 May 13:00
15	CSAT3 and Li7500A	1982/75H-1931	CR3000	Only CSAT3 21 May 15:30 to 23 May 09:30 CSAT3 + Li7500A 23 May 10:30 to 24 May 13:00
16	Gill-WindMaster and Li7500A	105101/75H-2199	CR1000	17 May 20:00 to 24 May 13:30
17	CSAT3 and EC150	1093/1028	CR3000	14 May 19:30 to 24 May 13:00
18	CSAT3 and Li7500A	2254/75H-1954	CR3000	14 May 18:30 to 24 May 13:00
19	Gill-WindMaster and Li7500A	1205/75H-2104	CR1000	16 May 18:00 to 24 May 13:00
20	Gill-WindMaster and Li7500	0823/75H-1201	CR1000	16 May 18:30 to 24 May 13:00

[a]The instruments were named 20 serial numbers with the order from west to east.

10, 20, 40, 60, 100 cm), and soil heat flux (HFT3, 6 cm, three replicates) measurements were also performed. The Gobi site is located in the northeast side of the intercomparison field site (pink triangle in Figure 1).

[10] HiWATER-MUSOEXE contained 21 radiometer sets, 22 EC sets, and 8 LAS sets; 3 radiometer sets, 2 EC sets, and a LAS set were not included in the comparison due to their late arrival. The late-arriving radiometers, ECs, and LASs were compared with the corresponding instruments installed in the superstation.

2.1. Radiometers

[11] A total of 18 radiometer sets were involved in the comparison experiment, including 17 four-component radiometer sets and one net radiometer: eight CNR1 sets and seven CNR4 sets from Kipp & Zonen, one PSP and PIR manufactured by the Eppley, one NR01 set produced by the Hukseflux, and one Q7 set manufactured by Radiation and Energy Balance Systems (REBS). Most of the radiometers were new, and all of the radiometers were precalibrated by the manufacturers. The sampling frequency was 0.5 Hz; the output data were stored at 10 min intervals. The radiometers were installed in three lines. The first line included CNR1 and NR01 (nos.1–9 from west to east). The second line included PSP and PIR. Lastly, the third line included CNR4 and Q7 (nos.1–8 from west to east) (Table 1 and Figure 2), orientated in a southern direction and installed at heights of 1.44 m (the first and third lines) and 1.5 m (the second line). Moreover, a separate protected area was constructed. Hay and small pebbles were removed from this area. The area was subsequently covered with the same type of soil to ensure the homogeneity of the underlying surface.

2.2. Eddy Covariance Systems

[12] Twenty eddy covariance system sets were included in the intercomparison: CSAT3 and Li7500 (10 sets, Campbell and Licor), CSAT3 and Li7500A (six sets, Campbell and Licor), WindMaster (Gill) and Li7500 (one set, Gill and Licor), WindMaster (Gill) and Li7500 A (two sets, Gill and Licor), and CSAT3 and EC150 (one set, Campbell). During the intercomparison, the EC installation was divided into four groups with five sets in each group. The first three groups contained CSAT3 and Li7500 /Li7500A systems. The last group was a combination of Gill and Li7500/ Li7500A, CSAT3 and EC150, and CSAT3 and Li7500A. The EC for each group was arranged at approximately 0.7 m intervals; each

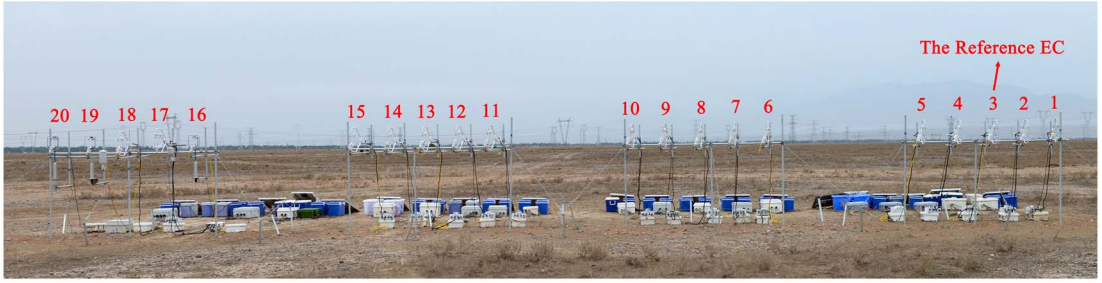

Figure 3. Pictures of the ECs used in the intercomparison field.

Table 3. The Information for the Large Aperture Scintillometers Used

No.	Type	Start Time	End Time
LAS1	BLS900, Scintec (BNU1); LAS, ZZLAS1	18-05-2012 12:36 16-05-2012 19:55	24-05-2012 11:00 24-05-2012 10:30
LAS2	BLS900, Scintec (CAU); LAS, ZZLAS2	15-05-2012 19:30 16-05-2012 19:48	24-05-2012 11:00 24-05-2012 10:30
LAS3	BLS450, Scintec (AR)	17-05-2012 00:00	24-05-2012 10:30
	BLS450, Scintec (NQ)	17-05-2012 00:00	24-05-2012 10:30
LAS4	BLS900, Scintec (BNU2)	18-05-2012 11:29	24-05-2012 10:28

group covered an interval of approximately 2 m. Therefore, all ECs were placed within 20 m. Moreover, the EC height was 1.7 m, and the sonic anemometer sensors faced north. The instruments were named with 20 serial numbers, ordered from west to east. The EC sample frequency was 10 Hz. The EC data were stored with an average time of 30 min and 10 Hz in the compact flash card. All EC infrared CO_2/H_2O gas analyzers used for the comparison were calibrated indoors prior to the field test. Pictures of the instruments and detailed information are shown in Table 2 and Figure 3.

2.3. Large Aperture Scintillometers

[13] Seven large aperture scintillometer (LAS) sets were involved in the intercomparison and divided into four groups: three BLS900 sets and two BLS450 sets manufactured by Scintec and two scintillometer sets constructed by our group (ZZLAS) [*Shi et al*., 2010]. The two BLS450 sets were erected together (the transmitter and receiver were reversed, LAS3), and all but one BLS900 set was erected together with a ZZLAS (Table 3 and Figure 4). The BLS900 (BNU1) and ZZLAS1, BLS900 (CAU) and ZZLAS2, BLS450 (AR) and BLS450 (NQ), and BLS900 (BNU2) combinations were installed from west to east at the field site. Each LAS group was arranged at 80 m intervals, a path length of 606 m, and an installation height of 2.02 m. The transmitter frequency was 5 Hz and 1 Hz for the BLS series and ZZLAS, respectively, and the output data were stored at 1 min intervals.

3. Data Processing

[14] The Edire software developed at the University of Edinburgh was used to process the EC data (http://www.geos.ed.ac.uk/abs/research/micromet/EdiRe/). The raw 10 Hz data were processed, including spike detection, lag correction of H_2O/CO_2 relative to the vertical wind component, sonic virtual temperature correction, coordinate rotation (2-D rotation), corrections for density fluctuation (Webb-Pearman-Leuning correction), and frequency response correction. The angle of attack (AOA) error correction was applied to the EC systems with Gill-WindMaster sonic anemometer using the Eddypro software developed by the Li-Cor (http://www.licor.com/env/products/eddy_covariance/software.html), which provided the latest AOA correction method proposed by *Nakai and Shimoyama* [2012]. The EC data were finally averaged over 30 min periods. Moreover, the observation data quality was divided into three classes according to the quality assessment method of stationarity (Δst) and integral turbulent characteristics test (ITC) proposed by *Foken and Wichura* [1996]: class 1 (level 0: $\Delta st < 30$ and $ITC < 30$), class 2 (level 1: Δst 100 and $ITC < 100$), and class 3 (level 2: $\Delta st > 100$ and $ITC > 100$), representing high-, medium-, and low-quality data, respectively. Class 1 was selected for the analysis. In addition to the above processing steps, the half-hourly flux data were screened in a five-step procedure: (i) data from periods of sensor malfunction were rejected (e.g., when there was a faulty diagnostic signal, and the automatic gain control value was greater than 65); (ii) data within 1 h of precipitation were rejected; (iii) incomplete 30 min data were rejected when the missing data constituted more than 3% of the 30 min raw record; (iv) data were rejected at night when the friction velocity (u_*) was less than 0.1 m s^{-1} [*Blanken et al*., 1998]; and (v) the orientation of the sonic anemometer was northward (consistent with the prevailing wind direction). To avoid the effects of brackets or other nearby EC sensors, only wind directions of 315°–0° and 0°–45° were considered.

[15] The structure parameter of the refractive index of air C_n^2 was directly output by the BLS900/450. For the ZZLAS, C_n^2 was calculated using the equation $C_n^2 = 10^{\left(U_{CN2}-12+1.15\sigma_{U_{CN2}^2}\right)}$, where U_{CN2} is the logarithm of C_n^2 and $\sigma_{U_{CN2}}$ is the variance of U_{CN2}. The data were carefully screened to ensure good LAS data quality using the following procedure: (i) data for C_n^2 beyond the saturation criterion were rejected, which was determined according to *Ochs and Wilson* [1993], and the upper limit of C_n^2 saturation was 1.01×10^{-12} m$^{-2/3}$; (ii) data obtained during periods of precipitation were rejected; (iii) data were rejected when the average X intensity was less than 1000 for the BLS series (for one BL450 (NQ) purchased in 2006, the data were deleted when the minimum value of the demodulated signal was less than 50), and data were rejected when the demodulated signal was greater than −30 mV for ZZLAS; (iv) data were rejected at night during periods of weak turbulence (when u_* was less than 0.1 m s^{-1}); and (v) data were rejected if collected when the sensor was malfunctioning.

Figure 4. Pictures of the LASs used in the intercomparison field.

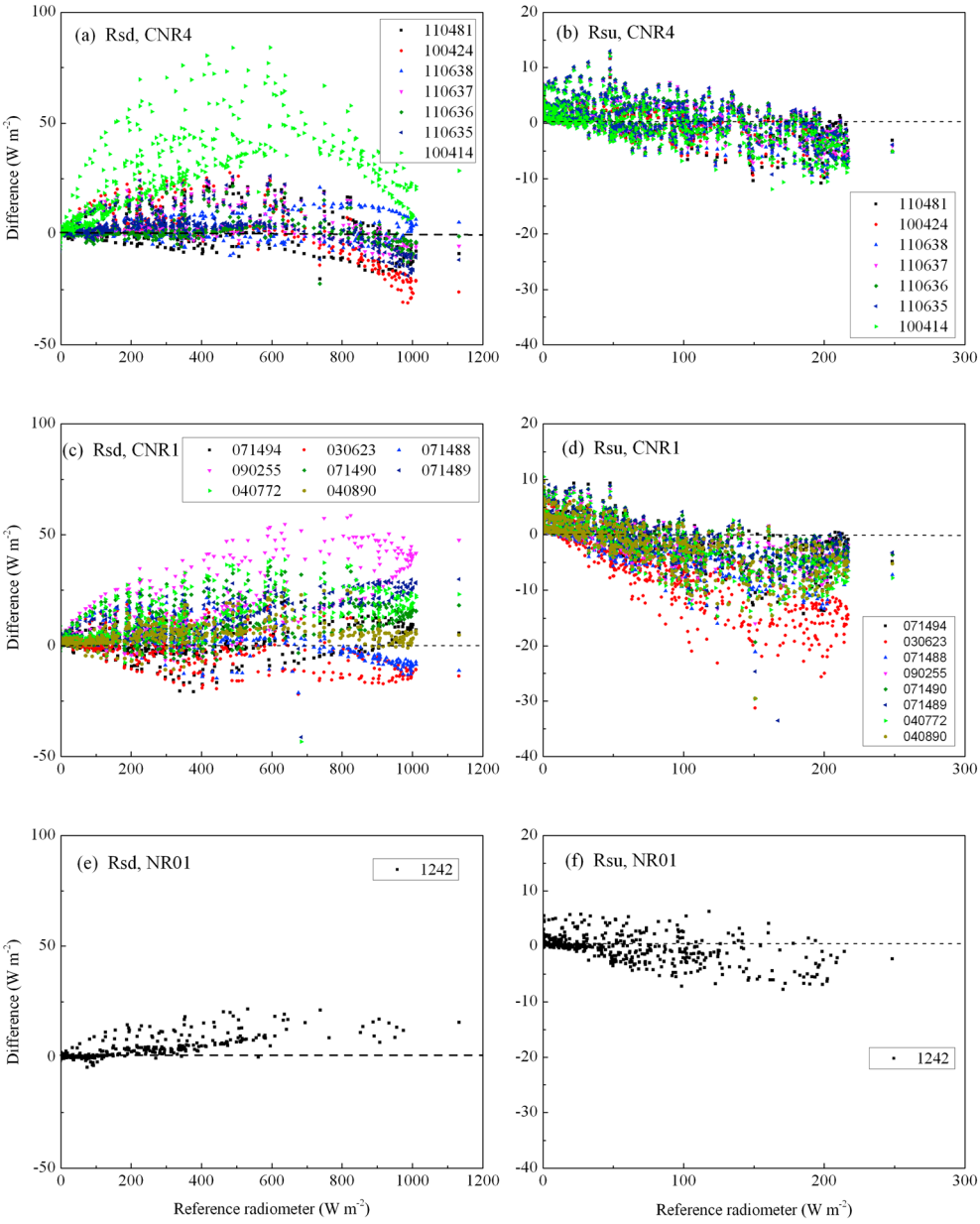

Figure 5. Scatterplots of the differences in 10 min incoming/outgoing shortwave radiation as a function of the 10 min reference measurements (16 to 22 May 2012): (a) incoming shortwave radiation for CNR4, (b) outgoing shortwave radiation for CNR4, (c) incoming shortwave radiation for CNR1, (d) outgoing shortwave radiation for CNR1, (e) incoming shortwave radiation for NR01, and (f) outgoing shortwave radiation for NR01.

[16] C_n^2 is related to the temperature structure parameter C_T^2 ($K^2\,m^{-2/3}$), the humidity structure parameter C_q^2 ($kg^2\,m^{-6}\,m^{-2/3}$), and a covariant term C_{Tq} ($K\,kg\,m^{-3}\,m^{-2/3}$). The optical scintillometer is more sensitive to variations in temperature than humidity. As a simplification, C_T^2 can be calculated together with the air temperature, air pressure, and Bowen ratio [*Wesely*, 1976]. According to the Monin-Obukhov similarity theory, the sensible heat flux can be obtained through an iterative method, together with meteorological elements (e.g., wind speed, air temperature, and air pressure) and additional parameters (e.g., zero-plane displacement height, the aerodynamic roughness length, and Bowen ratio). The LAS data were averaged over 30 min intervals. The reader is referred to *Liu et al.* [2011, 2013] for details.

[17] The output radiometer data were produced using the *.dat format and saved in an Excel spreadsheet as incoming/outgoing shortwave radiation, incoming/outgoing longwave radiation, and net radiation. The diurnal radiation variation was assessed to ensure the radiometer measurement quality. Moreover, data clearly beyond the range of what is physically possible were rejected.

[18] In this study, the root-mean-square error (RMSE), mean relative error (MRE), and correlation coefficient (R) were used to measure the average discrepancy between each

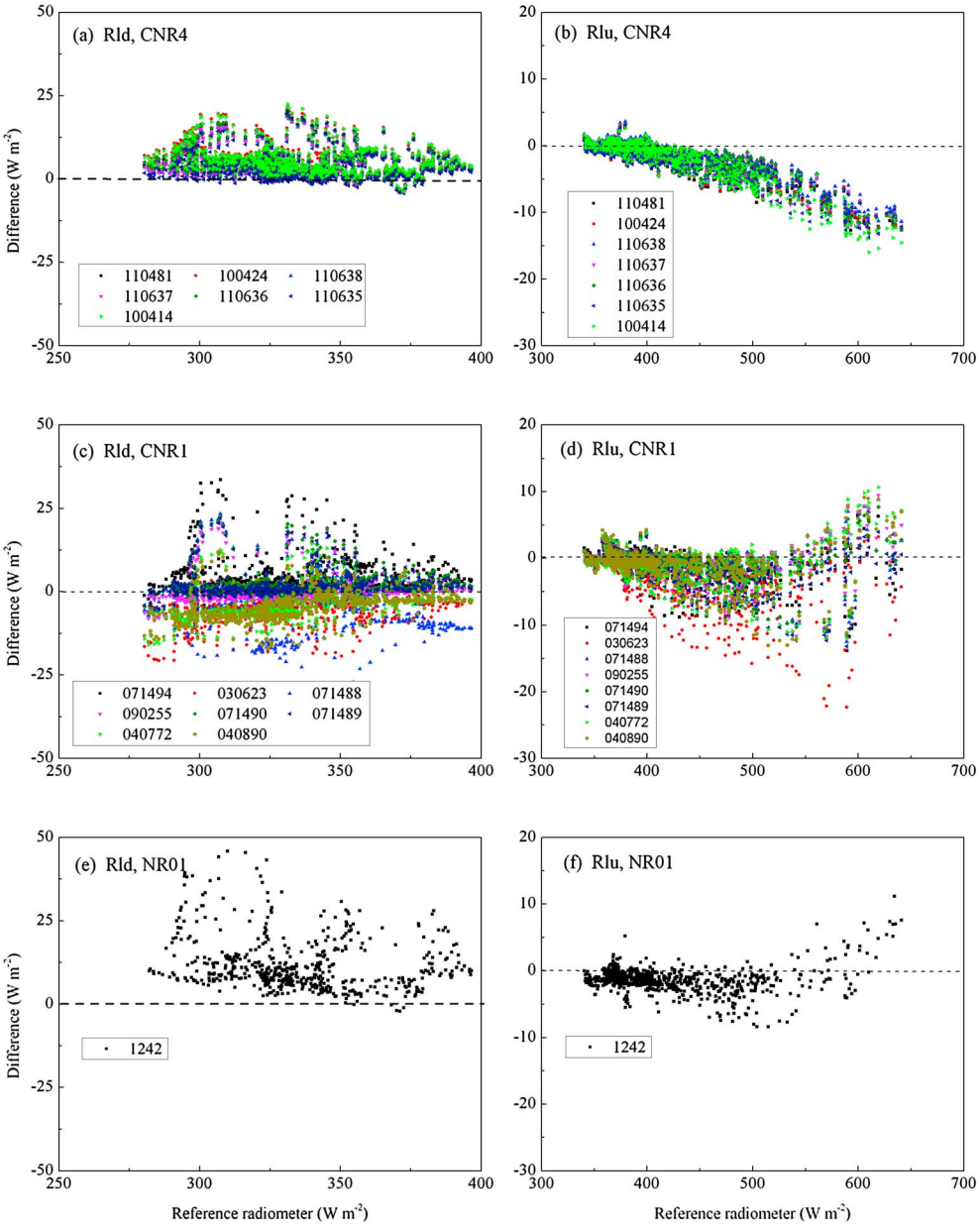

Figure 6. Scatterplots of the differences in 10 min incoming/outgoing longwave radiation as a function of the 10 min reference measurements (16 to 22 May 2012): (a) incoming longwave radiation for CNR4, (b) outgoing longwave radiation for CNR4, (c) incoming longwave radiation for CNR1, (d) outgoing longwave radiation for CNR1, (e) incoming longwave radiation for NR01, and (f) outgoing longwave radiation for NR01.

instrument and the reference. RMSE, MRE, and R were calculated according to *Jia et al*. [2012], as follows:

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{n} (P_i - O_i)^2 \Big/ n} \tag{1}$$

$$\mathrm{MRE} = \frac{100}{n} \sum_{i=1}^{n} \frac{(P_i - O_i)}{\overline{O}} \tag{2}$$

$$R = \sum_{i=1}^{n} (P_i - \overline{P})(O_i - \overline{O}) \Big/ \left[\sum_{i=1}^{n} (P_i - \overline{P})^2 \sum_{i=1}^{n} (O_i - \overline{O})^2 \right]^{1/2} \tag{3}$$

where P_i is the measured value of each instrument, O_i is the measured value of the reference, $\overline{P}$ is the mean measured value of each instrument, $\overline{O}$ is the mean value of the reference, and n is the number of samples. The RMSE is used to measure the discrepancy between the comparison instrument measurements and the reference measurement, MRE is used to measure the relative degree of their discrepancy, and R is used to measure the measurement consistency.

[19] To compare with previous studies, the regression slope (deviation from the 1:1 regression line) was added in

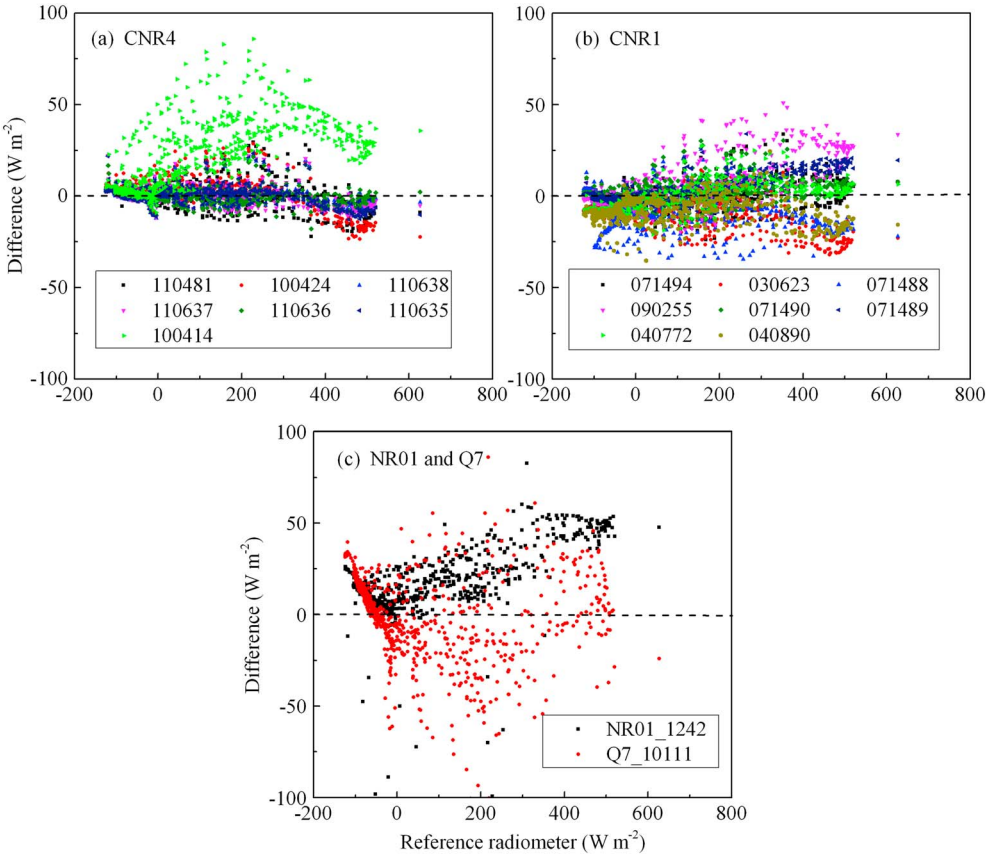

Figure 7. Scatterplots of the differences in 10 min net radiation as a function of the 10 min reference measurements (16 to 22 May 2012): (a) net radiation for CNR4, (b) net radiation for CNR1, and (c) net radiation for NR01 and Q7.

the ensuing analysis, representing the relative discrepancy between each instrument value and the reference measurement.

4. Results

4.1. Meteorological Conditions

[20] The weather during the intercomparison period was characterized by cloudy skies. It was rainy on the morning of 21 May. This rainy period was preceded by a gray sky with drizzle from the afternoon of 19 May to the evening of 20 May. The air temperature typically ranged from 5°C at night to 30°C during the day; the average temperature was 17.1°C. The air humidity was less than 10% during the day and approximately 40% at night. However, the air humidity reached 70% on the rainy day (21 May); the average air humidity was 27% during the comparison period. The wind speed was generally less than 6 m s^{-1}; the maximum wind speed was 13 m s^{-1} during the study period. The prevailing wind direction was north and northwest during the day (the proportion was greater than 35%). Moreover, the air pressure was approximately 840 hPa. The soil moisture was very low (the volumetric water content was less than 10%). However, the soil temperature was high, greater than 50°C during day.

4.2. Radiation Sensors

[21] Five types of radiometers were used for the intercomparison: PSP and PIR (Eppley, U.S.), CNR4 and CNR1 (Kipp & Zonen, Netherlands), NR01 (Hukseflux, Netherlands), and Q7 (REBS, U.S.). The PSP and PIR radiometer was the WMO first-class radiometer. Therefore, this radiometer was used for the reference. However, some problems occurred with the PSP and PIR shortwave radiation after 22 May. Therefore, the period from 16 May to 22 May was selected for the analysis. The peak incoming shortwave radiation (Rsd) was greater than 1000 W m^{-2}; the peak outgoing shortwave radiation (Rsu) was less than 200 W m^{-2}. The incoming longwave radiation (Rld) varied daily between 250 and 400 W m^{-2}, while outgoing longwave radiation (Rlu) was between 340 and 630 W m^{-2}. Moreover, the net radiation was between −100 and 600 W m^{-2}.

[22] For the radiometer intercomparison, the differences (the involved radiometers minus the reference, 10 min) were plotted as a function of the reference measurements during the study period. Figures 5, 6, and 7 showed the incoming/outgoing shortwave radiation, incoming/outgoing longwave radiation, and net radiation, respectively. The statistics for the different radiometers were listed in Table 4. Moreover, the RMSE, MRE, and the correlation coefficient (R) were selected to assess the radiometer performances. The regression slopes were also analyzed.

4.2.1. Incoming/Outgoing Shortwave Radiation

[23] Figure 5 depicts the differences as a function of the reference for incoming/outgoing shortwave radiation (Rsd/Rsu) from 16 to 22 May. Table 4 shows the radiometer statistics.

Table 4. Statistics of Radiometers for 16 to 22 May 2012 (Using the PSP and PIR Measurements as the Reference, With RMSE ($W\,m^{-2}$), MRE (%))

	Rsd				Rsu				Rld				Rlu				Rn			
Type	Slope	RMSE	MRE	R	Slope	RMSE	MRE	R	Slope	RMSE	MRE	R	Slope	RMSE	MRE	R	Slope	RMSE	MRE	R
CNR4_110481	1.0	6.36	0.32	0.9998	0.99	3.34	3.55	0.9994	1.01	5.68	1.22	0.9897	0.99	3.56	−0.49	0.9998	1.0	6.13	−0.41	0.9994
CNR4_100424	1.0	8.22	0.37	0.9997	0.99	3.44	3.05	0.9994	1.02	7.13	1.70	0.9881	0.99	3.46	−0.47	0.9998	1.0	7.31	2.51	0.9993
CNR4_110638	1.0	5.17	0.82	0.9999	0.99	3.52	3.54	0.9994	1.01	5.55	1.24	0.9910	1.00	3.01	−0.34	0.9998	1.0	3.78	0.33	0.9998
CNR4_110637	1.0	5.81	0.75	0.9998	1.00	3.42	3.91	0.9994	1.01	5.47	1.18	0.9905	0.99	3.41	−0.45	0.9998	1.0	4.44	0.24	0.9999
CNR4_110636	1.0	4.86	0.53	0.9999	0.99	3.53	4.09	0.9994	1.01	6.08	1.41	0.9903	1.00	3.28	−0.41	0.9998	1.0	3.73	0.56	0.9998
CNR4_110635	1.0	6.18	0.80	0.9998	0.99	3.52	4.05	0.9994	1.01	5.19	0.99	0.9897	0.99	3.61	−0.53	0.9998	1.0	4.38	0.29	0.9997
CNR4_100414	1.05	24.58	7.08	0.9989	0.98	3.57	2.64	0.9994	1.02	7.08	1.71	0.9886	0.99	3.96	−0.48	0.9998	0.99	24.94	24.88	0.9974
CNR1_071494	1.0	5.93	0.72	0.9998	0.98	3.98	1.89	0.9992	1.01	7.27	1.32	0.9787	1.00	2.78	−0.22	0.9993	1.0	6.02	0.20	0.9994
CNR1_030623	0.99	6.81	−0.12	0.9998	0.92	7.61	−4.02	0.9989	0.98	6.79	−1.55	0.9899	0.99	4.77	−0.56	0.9985	1.0	11.45	−12.69	0.9993
CNR1_071488	1.0	5.90	0.91	0.9998	0.96	4.50	0.33	0.9993	0.99	8.56	−1.09	0.9617	1.00	2.54	−0.15	0.9993	1.0	11.91	−10.99	0.9986
CNR1_090255	1.04	17.71	4.78	0.9997	0.98	3.61	1.95	0.9994	1.00	3.92	0.04	0.9908	1.00	2.57	−0.15	0.9993	1.0	11.96	5.33	0.9991
CNR1_071490	1.02	9.39	2.60	0.9998	0.98	3.68	2.76	0.9994	1.01	4.64	0.63	0.9893	1.00	2.44	−0.17	0.9994	1.0	5.52	1.74	0.9996
CNR1_071489	1.03	5.14	−0.72	0.9999	0.98	4.00	2.62	0.9993	1.00	4.23	0.46	0.9901	1.00	2.79	−0.30	0.9994	1.0	7.02	3.69	0.9998
CNR1_040772	1.03	3.60	0.18	0.9998	0.96	4.80	0.52	0.9993	0.99	5.89	−1.18	0.9904	1.00	3.03	−0.21	0.9991	1.0	7.54	−3.97	0.9995
CNR1_040890	1.01	8.04	−1.73	0.9999	0.97	4.43	0.88	0.9992	0.99	6.57	−1.38	0.9896	1.00	3.05	−0.22	0.9991	1.0	10.25	−12.15	0.9994
NR01_1242	1.02	4.00	−1.75	0.9999	0.99	2.30	1.43	0.9995	1.03	12.79	3.21	0.9682	1.00	2.35	−0.29	0.9996	0.99	26.72	23.47	0.9950
Q7_10111																	0.98	23.35	−1.96	0.9909
Average	1.01	7.98	0.97	0.9998	0.98	3.95	2.07	0.9993	1.01	6.43	0.62	0.9860	0.9963	3.16	−0.34	0.9995	0.9976	10.38	1.24	0.9986

[24] The differences between the Rsd measurements were within ±25 $W\,m^{-2}$ of the reference (except for the CNR4-ID100414 and CNR1-ID090255). The differences varied as a function of the reference according to a parabolic shape for the CNR4-type instruments; the maximum was approximately 25 $W\,m^{-2}$ when the reference radiation was approximately 500 $W\,m^{-2}$. The differences for the CNR1/NR01-type instruments were between −25 and 37 $W\,m^{-2}$; the differences generally increased with increasing reference radiation. The CNR4-ID100414 had a relatively large amplitude that was approximately 80 $W\,m^{-2}$ greater than the reference. Moreover, the CNR1-ID090255 exhibited relatively large difference of approximately 60 $W\,m^{-2}$. Table 4 shows that the average regression slope was 1%, typically less than 5%, and the average RMSE was 7.98 $W\,m^{-2}$, ranging between 3.6 and 24.58 $W\,m^{-2}$. The average MRE was 0.97%, ranging from −1.75% to 7.08%. Moreover, R was very high (greater than 0.99), indicating that the incoming shortwave radiometer measurements were very consistent.

[25] Figure 5 (b,d,f) shows that all radiometers had good consistency with each other except for the CNR1-ID030623; the differences between the Rsu measurements were between −10 and 10 $W\,m^{-2}$. The differences were between 0 and 10 $W\,m^{-2}$ when the reference measurement was less than 50 $W\,m^{-2}$, decreasing to −10 $W\,m^{-2}$ when the reference increased to 250 $W\,m^{-2}$. The statistics in Table 4 indicate that the regression slopes were primarily less than 4% (one CNR1-ID030623 was slightly greater than 8%), with an average of 2%. Moreover, the average RMSE was 3.95 $W\,m^{-2}$, ranging from 2.30 to 7.61 $W\,m^{-2}$, and the average MRE was 2.07%, ranging from −4.02% to 4.09%. The outgoing shortwave radiometers were also very consistent with each other; R was greater than 0.99.

4.2.2. Incoming/Outgoing Longwave Radiation

[26] Figure 6 shows the differences for incoming/outgoing longwave radiation (Rld/Rlu). The differences in Rld were between −25 and 50 $W\,m^{-2}$. Moreover, the difference had not clear variation as the reference increased. The differences for the CNR4-type were concentrated between −5 and 25 $W\,m^{-2}$; the differences varied consistently with each other for all the CNR4-type radiometers. The differences for most CNR1-type agreed within ±25 $W\,m^{-2}$. However, the differences were more discrete, especially for the CNR1-ID071488. The NR01 radiometer exhibited a maximum overestimation (relative to the reference) of approximately 45 $W\,m^{-2}$. The average regression slopes for incoming longwave radiation was 1%, typically less than 3%, and the average RMSE was 6.43 $W\,m^{-2}$, ranging from 3.92 to 12.79 $W\,m^{-2}$. The MRE ranged from −1.55% to 3.21%, with an average of 0.62%. Moreover, R was greater than 0.98, except for two radiometers (0.96 for the CNR1-ID071488 and NR01-ID1242) (Table 4).

[27] The Rlu differences were approximately between −20 and 10 $W\,m^{-2}$; the differences decreased as the reference measurements increased, especially for the CNR4-type instruments. Differences between −15 and 5 $W\,m^{-2}$ were found between the CNR4 and the reference; the differences for all CNR4-type radiometers exhibited the same variations. The differences between most CNR1/NR01-type radiometers and the reference were between −10 and 10 $W\,m^{-2}$. However, a relatively large difference (up to 24 $W\,m^{-2}$) was

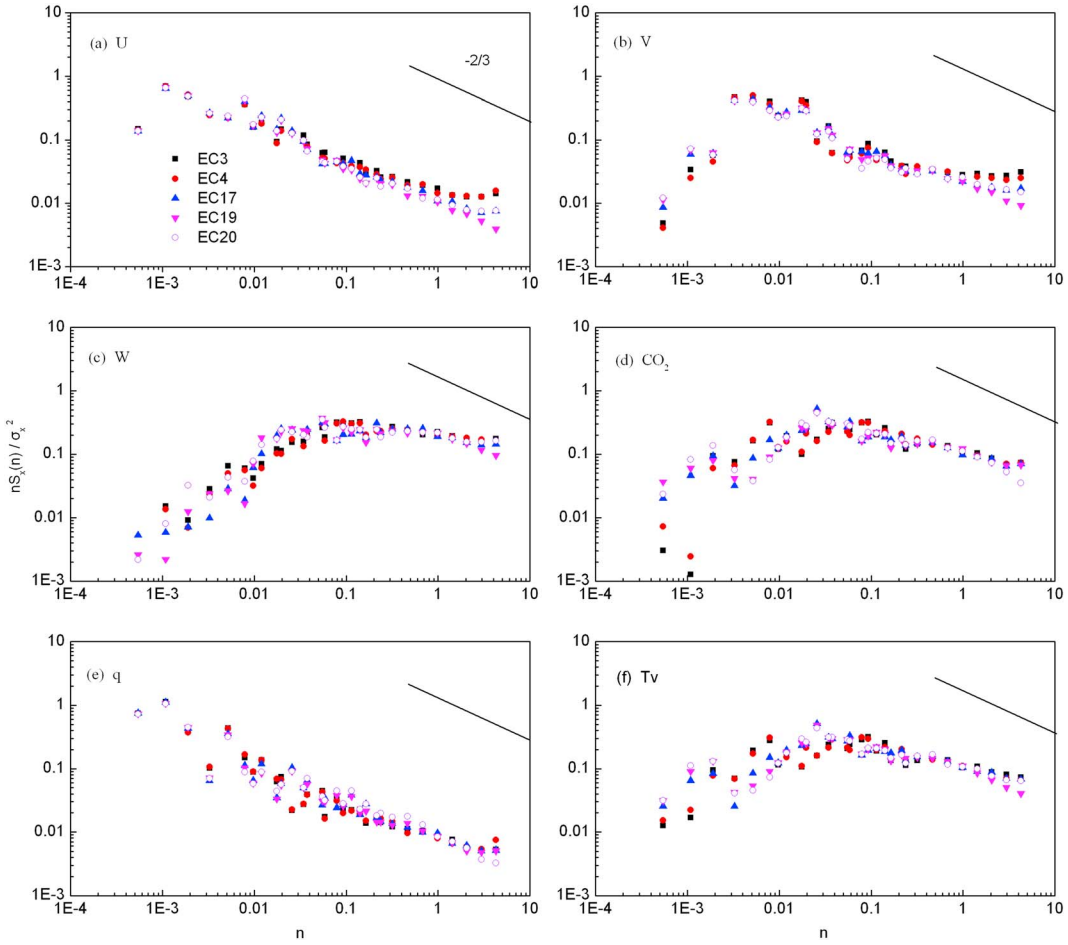

Figure 8. Normalized power spectra of (a–c) the wind velocity components (u, v, w), (d) carbon dioxide (CO_2) density, (e) water vapor (q), and (f) virtual temperature (T_v) (the data with unstable stability (approximately −0.2) on 18 May 2012 were selected) (EC3: CSAT3 and Li7500A, EC4: CSAT3 and Li7500, EC17: CSAT3 and EC150, EC19: Gill and Li7500A, and EC20: Gill and Li7500).

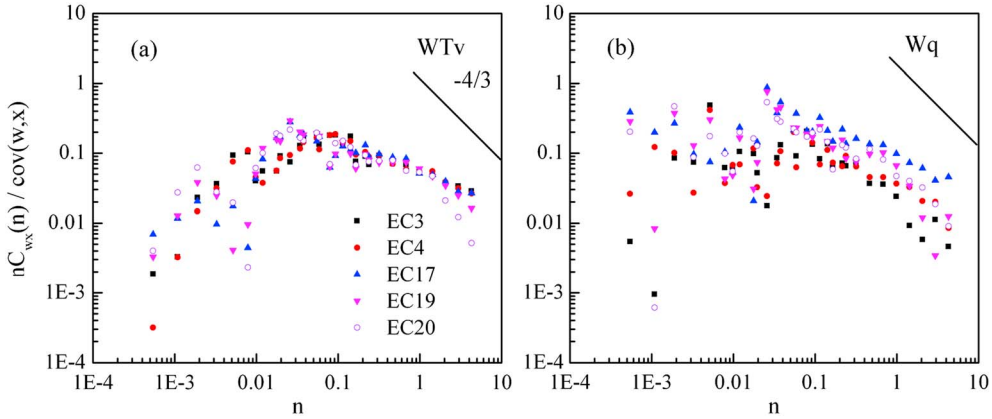

Figure 9. Normalized cospectra of (a) vertical velocity and virtual temperature (WT_v) and (b) water vapor (Wq) (the data with unstable stability (approximately −0.2) on 18 May 2012 were selected) (EC3: CSAT3 and Li7500A, EC4: CSAT3 and Li7500, EC17: CSAT3 and EC150, EC19: Gill and Li7500A, and EC20: Gill and Li7500).

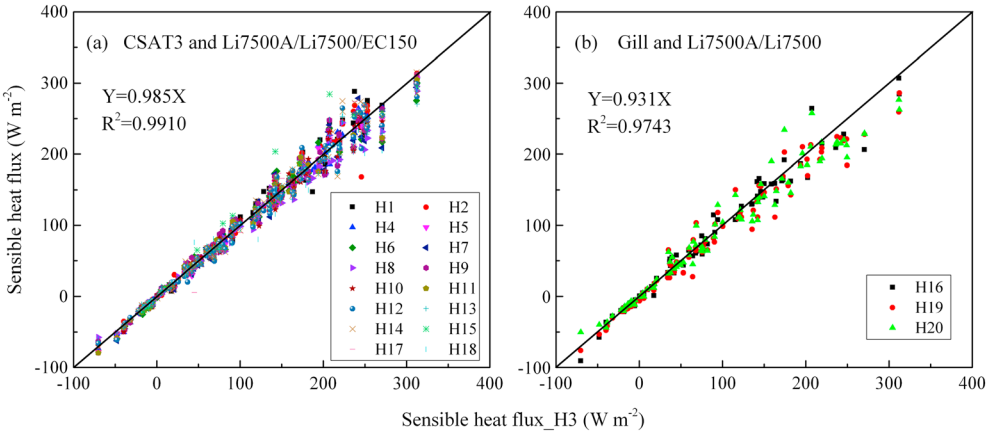

Figure 10. The intercomparison of sensible heat fluxes measured by the EC systems (using the EC3 measurements as the reference, 14 to 24 May 2012): (a) CSAT3 and Li7500A/Li7500/EC150,and (b) Gill and Li7500A/Li7500.

found for the CNR1-ID030623 radiometer. As shown in Table 4, the outgoing longwave radiation regression slopes were less than 1%, with an average of 0.4%. Moreover, the average RMSE was 3.16 W m^{-2}, ranging between 2.35 and 4.77 W m^{-2}; the average MRE was −0.34%, ranging from −0.53% to −0.15%. Good consistency was observed for outgoing longwave radiation; *R* was greater than 0.99.

[28] During the instrument comparison, the underlying surface in the radiation sensor field of view was recovered with the same soil. However, there was also a small difference in the underlying surface in the three radiometer lines (Figure 2) that may have caused the differences in outgoing longwave/shortwave radiation.

4.2.3. Net Radiation

[29] The net radiation (Rn) differences are shown in Figure 7; the corresponding statistics are presented in Table 4. Generally, Figure 7 shows that the Rn differences ranged from −50 to 50 W m^{-2} (except for the CNR4-ID100414); no pronounced increase was detected with the reference for CNR1- and CNR4-type radiometers. Most of the CNR4- and CNR1-type instruments exhibited the differences of −25 to 25 W m^{-2}. However, the CNR1-type radiometer differences were more discrete. The CNR4-ID100414 radiometer exhibited differences of up to 80 W m^{-2} compared with the reference. The NR01/Q7-type differences were also very large and more discrete. The average net radiation difference was estimated of 25 W m^{-2} (4% of the maximum value) except for CNR4-ID100414, NR01, and Q7 radiometers. Table 4 shows that the net radiation regression slopes were less than 2%, with an average of 0.2%. Moreover, the average RMSE was 10.38 W m^{-2}, ranging from 3.73 to 26.72 W m^{-2}; the average MRE was 1.24%, ranging from −12.69% to 24.88%. All net radiometers exhibited very good consistency; *R* was greater than 0.99.

[30] Better consistency was observed with most CNR4-type sensors compared to the CNR1, NR01, and Q7 (except the CNR4-ID100414). The intercomparison results were better than those of the FIFE comparison [*Field et al.*, 1992; *Nie et al.*, 1992] and were consistent with the EBEX-2000 comparison [*Kohsiek et al.*, 2007]. Overall, most radiometers were consistent with the reference and showed reliable accuracy. Therefore, these radiometers can be used in HiWATER-MUSOEXE.

4.3. Eddy Covariance Sensors

[31] Five EC system sensor combinations (20 sets) were compared (CSAT3 and Li7500, CSAT3 and Li7500A, WindMaster (Gill) and Li7500, WindMaster (Gill) and Li7500A, and CSAT3 and EC150); these classifications were used in the latent heat flux intercomparison. The sensible heat flux was calculated from the sonic anemometer measurements. Two classes were defined in the sensible heat flux intercomparison, i.e., CSAT3 and WindMaster (Gill) sonic anemometers. Because the EC3 system was a new combination (CSAT3 and Li7500A) and contained the longest data series, with a proper and stable working status, this EC set was chosen as the reference.

[32] After carefully screening the data (described in section 3), the percentage of valid data was approximately 27% (taking EC3 as an example). The data were mainly rejected by the

Table 5. Statistics of Sensible Heat Fluxes Measured by the EC Systems for 14 to 24 May 2012 (Using the EC3 Measurements as the Reference)

No.	Type	RMSE (W m^{-2})	MRE (%)	*R*
H1	CSAT3 and Li7500	11.13	1.65	0.9939
H2	CSAT3 and Li7500	10.96	0.50	0.9937
H4	CSAT3 and Li7500	6.22	−0.05	0.9979
H5	CSAT3 and Li7500A	7.62	−3.83	0.9968
H6	CSAT3 and Li7500A	12.69	−2.57	0.9897
H7	CSAT3 and Li7500	13.57	−4.47	0.9902
H8	CSAT3 and Li7500	14.28	−5.83	0.9914
H9	CSAT3 and Li7500	11.80	−0.83	0.9920
H10	CSAT3 and Li7500	10.94	−2.29	0.9932
H11	CSAT3 and Li7500	9.55	−1.46	0.9932
H12	CSAT3 and Li7500	11.97	−3.05	0.9923
H13	CSAT3 and Li7500A	12.22	−2.46	0.9921
H14	CSAT3 and Li7500	13.40	−1.21	0.9902
H15	CSAT3 and Li7500A	22.66	10.67	0.9829
H17	CSAT3 and EC150	12.55	−2.49	0.9909
H18	CSAT3 and Li7500A	15.16	−4.03	0.9878
Average		12.30	−1.36	0.9918
H16	Gill and Li7500A	13.87	−3.36	0.9877
H19	Gill and Li7500A	17.43	−8.11	0.9854
H20	Gill and Li7500	18.96	−5.10	0.9823
Average		16.75	−5.52	0.9851

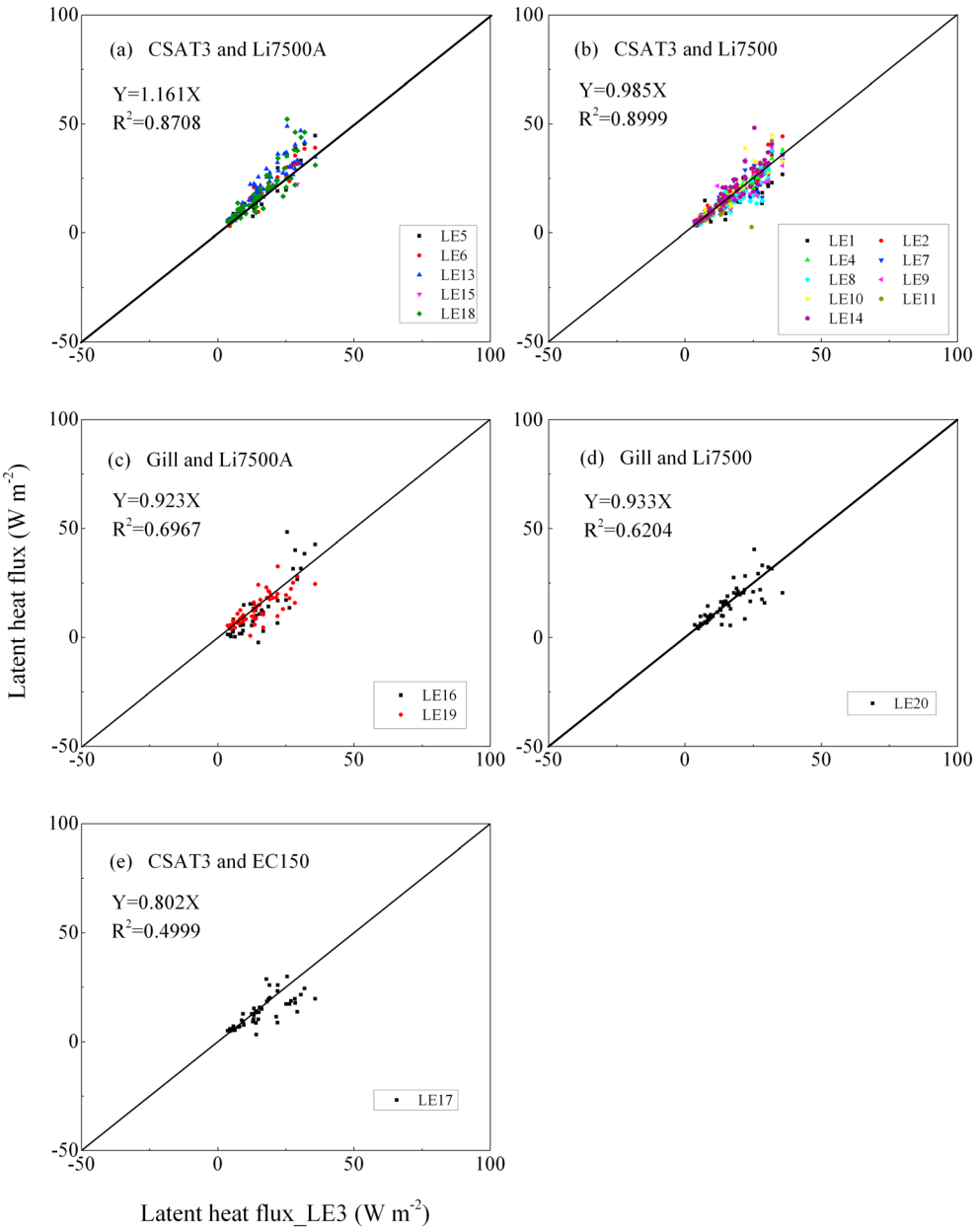

Figure 11. The intercomparison of latent heat fluxes measured by the EC systems (using the EC3 measurements as the reference, 14 to 24 May 2012): (a) CSAT3 and Li7500A, (b) CSAT3 and Li7500, (c) Gill and Li7500A, (d) Gill and Li7500, and (e) CSAT3 and EC150.

wind sector selection (approximately 67%). The rejected percentages of bad data quality and the u_* filter were 5% and 1%, respectively.

4.3.1. Power Spectra and Cospectra

[33] The power spectra and cospectra of each EC class were calculated using the fast Fourier transform and Hamming filter methods. EC3 (CSAT3 and Li7500A), EC4 (CSAT3 and Li7500), EC17 (CSAT3 and EC150), EC19 (Gill and Li7500A), and EC20 (Gill and Li7500) were chosen from different EC types. The power spectra, showing the frequency contribution of the wind velocity components (u, v, w), virtual temperature (T_v), water vapor (q) and carbon dioxide (CO_2) density, and the cospectra (WT_v, Wq) are shown in Figures 8 and 9, respectively, where the x axis is the natural frequency (n, Hz) and the y axis is the normalized power spectra ($S_x(n)/\sigma_x^2$) or cospectra ($C_{wx}(n)/\mathrm{Cov}(w, x)$) multiplied by n. A typical day (18 May 2012) was selected for this analysis. As expected, the spectra fell together in the inertial subrange and followed the $-2/3$ slope quite well. The cospectra WT_v collapsed in the inertial subrange; both cospectra (WT_v, Wq) followed a slope close to $-4/3$. The upward trending values of power spectra at high-frequency end may indicate the presence of noise. However, this noise did not affect the covariance. The frequencies of the spectral peaks can be related to a length scale of the dominant energy-containing eddies by the local average wind speed. The w spectral peak frequency was 0.06 Hz (Figure 8), yielding a dominant

Table 6. Statistics of Latent Heat Fluxes Measured by the EC Systems for 14 to 24 May 2012 (Using the EC3 Measurements as the Reference)

No.	Type	RMSE ($W m^{-2}$)	MRE (%)	*R*
LE5	CSAT3 and Li7500A	4.36	12.71	0.9491
LE6	CSAT3 and Li7500A	4.21	15.44	0.9513
LE13	CSAT3 and Li7500A	6.38	27.19	0.9242
LE15	CSAT3 and Li7500A	3.61	−11.62	0.9533
LE18	CSAT3 and Li7500A	6.06	15.10	0.8810
Average		4.92	11.76	0.9318
LE1	CSAT3 and Li7500	4.28	−13.81	0.8899
LE2	CSAT3 and Li7500	2.20	4.82	0.9741
LE4	CSAT3 and Li7500	1.44	0.99	0.9838
LE7	CSAT3 and Li7500	3.46	−0.95	0.9170
LE8	CSAT3 and Li7500	3.98	−8.89	0.8906
LE9	CSAT3 and Li7500	3.55	−1.54	0.8989
LE10	CSAT3 and Li7500	3.99	8.02	0.9182
LE11	CSAT3 and Li7500	5.01	−7.59	0.8100
LE14	CSAT3 and Li7500	4.22	9.33	0.9169
Average		3.57	−1.07	0.9110
LE16	Gill and Li7500A	7.25	−17.05	0.8436
LE19	Gill and Li7500A	5.02	−11.95	0.7838
Average		6.13	−14.50	0.8137
LE20	Gill and Li7500	5.29	−3.62	0.7973
LE17	CSAT3 and EC150	6.07	−14.58	0.7520

eddy length scale of approximately 37 m (the wind speed was 2.22 $m s^{-1}$).

4.3.2. Intercomparison of Sensible Heat Fluxes

[34] Figure 10 shows the comparison of sensible heat fluxes measured by the 20 EC sets; the corresponding statistics are provided in Table 5. A very good correlation was found in the measured sensible heat fluxes from the different EC types (Figure 10).

[35] Table 5 shows that a high correlation between the EC systems was found, which was significant at the 0.01 level. Moreover, *R* was approximately 0.99 for all classes, indicating that the sensible heat fluxes measured by different ECs were in agreement with each other. For the CSAT3 sonic anemometer, the average RMSE, MRE, and *R* were 12.30 $W m^{-2}$, −1.36%, and 0.9918, respectively. The regression slopes were less than 2% for the CSAT3 sonic anemometers (Figure 10). The RMSE and MRE values for Gill sonic anemometer were relatively large, with an average of 16.75 $W m^{-2}$ and −5.52%, respectively. The regression slope was approximately 7%, and the average *R* was 0.9851. The average regression slopes, RMSE, and MRE of the sensible heat flux measured by all ECs were 3.21%, 13.00 $W m^{-2}$, and −2.02%, respectively, with a high correlation coefficient compared with the reference of 0.99.

[36] As reported above, good consistency was observed for all ECs, especially for the CSAT3 sonic anemometers.

4.3.3. Intercomparison of Latent Heat Fluxes

[37] Most latent heat fluxes were less than 40 $W m^{-2}$ during the comparison campaign due to the low soil moisture in the Gobi underlying surface (see section 4.1). As shown in Figure 11, a relatively large discrepancy among the EC systems was observed in the latent heat fluxes compared to the sensible heat fluxes. However, the data were approximately distributed near the 1:1 line. Due to calibration problems, the EC12 measured H_2O density was very low; these data points are not shown in Figure 11 (this sensor was recalibrated before the formal observation).

[38] The latent heat flux statistics for the ECs are summarized in Table 6, showing good correlation in the latent heat fluxes measured by each EC (significant at the 0.01 level). The average RMSE, MRE, and *R* of the latent heat fluxes

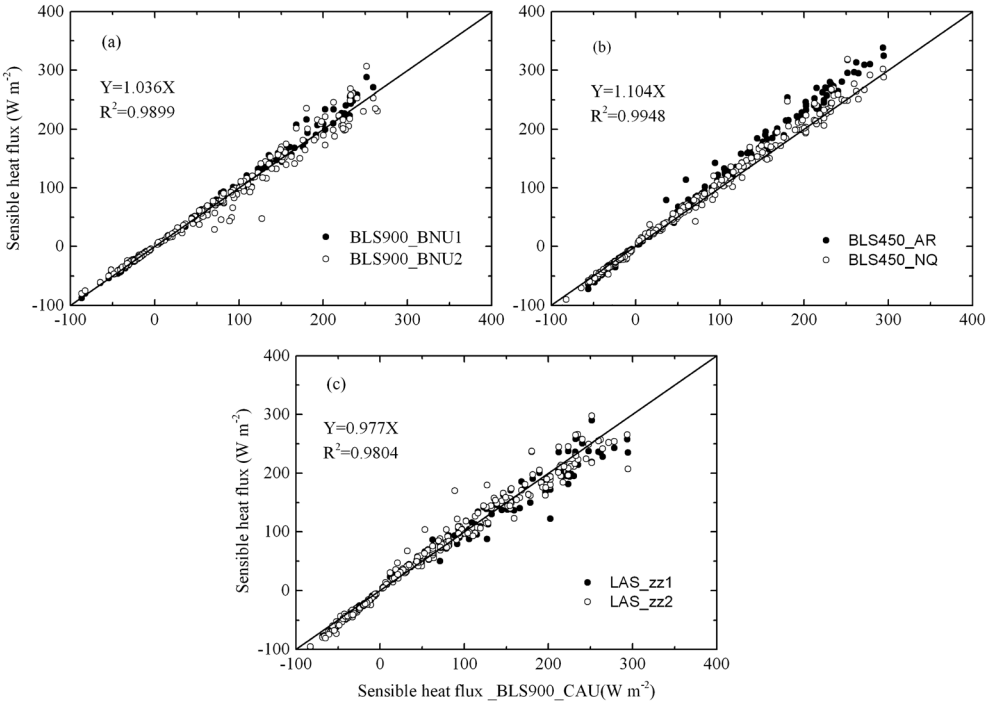

Figure 12. The intercomparison of sensible heat fluxes measured by the LASs (using the BLS900_CAU measurements as the reference, 14 to 24 May 2012): (a) BLS900, (b) BLS450, and (c) ZZLAS.

Table 7. Statistics of Sensible Heat Fluxes Measured by LASs for 16 to 24 May 2012 (Using the BLS900_CAU Measurements as the Reference)

	BLS900			BLS450			ZZLAS		
	BNU1	BNU2	Average	AR	NQ	Average	ZZLAS1	ZZLAS2	Average
RMSE($W m^{-2}$)	8.52	12.00	10.26	21.48	11.16	16.32	14.79	13.97	14.38
MRE (%)	6.71	4.25	5.48	15.33	5.60	10.47	−3.62	−3.81	−3.72
R	0.9978	0.9917	0.9948	0.9955	0.9960	0.9958	0.9881	0.9899	0.9890

for CSAT3 and Li7500A was 4.92 $W m^{-2}$, 11.76%, and 0.9318, respectively. The average RMSE, MRE, and *R* for CSAT3 and Li7500 were 3.57 $W m^{-2}$, −1.07%, and 0.9110, respectively. For Gill and Li7500A/Li7500, the average RMSE, MRE, and *R* were 6.13/5.29 $W m^{-2}$, −14.50%/−3.62%, and 0.8137/0.7973, respectively. Moreover, the RMSE, MRE, and *R* for CSAT3 and EC150 were 6.07 $W m^{-2}$, −14.58%, and 0.7520, respectively. The measured latent heat fluxes were larger in the CSAT3 and Li7500A measurements than the reference. Furthermore, CSAT3 and Li7500 observed lower measurements; the regression slopes were less than 16% and 2%, respectively. The latent heat fluxes measured by the Gill-type sonic anemometers were less than the reference; the regression slope was less than 7%. The CSAT3 and EC150 measurements were substantially less than the reference, with a regression slope of approximately 20% (Figure 11). The average regression slope, RMSE, and MRE of the latent heat fluxes measured by all ECs were 10.94%, 4.47 $W m^{-2}$, and 0.11%, respectively; the correlation coefficient was 0.8908. The latent heat fluxes were very small and had relatively large fluctuations and possibly large uncertainties in the drying Gobi desert. Therefore, we primarily focused on analyzing the sensible heat flux in this study. To fully understand the surface energy fluxes, the latent heat fluxes were also introduced.

[39] Based on the above analysis, most the eddy covariance systems exhibited good agreement and accuracy for sensible and latent heat flux measurements. Moreover, the results of our intercomparison were better than the EBEX-2000 comparison [*Mauder et al*., 2007]. These eddy covariance systems can be used in HiWATER-MUSOEXE.

4.4. Large Aperture Scintillometers

[40] Seven LAS sets were involved, including three BLS900 sets, two BLS450 sets produced by Scintec (Germany), and two scintillometers sets built by our group. BLS900 in line 2 was selected as the reference (BLS900_CAU); the other scintillometers were compared with the reference (Figure 12). The valid BLS900_CAU (the reference) data was approximately 67%. The rejected data were mainly caused by the u_* filter (32%) and by instrument malfunction and precipitation (1%). Figure 12 shows that most points were distributed around the 1:1 line, demonstrating very good agreement between instruments.

[41] Table 7 shows that all LASs exhibited good consistency with high correlation coefficients (*R* values greater than 0.98). Among the BLS900 measurements, the regression slope was less than 4% (Figure 12); the average RMSE and MRE were 10.26 $W m^{-2}$ and 5.48%, respectively. For the BLS450 series, the BLS450_NQ agreed well with the reference, i.e., the regression slope, RMSE, MRE, and *R* were 5%, 11.16 $W m^{-2}$, 5.60%, and 0.9960, respectively. In contrast, relatively large differences were observed for BLS450_AR, with an RMSE and MRE of 21.48 $W m^{-2}$ and 15.33%, respectively (the regression slope was 15%). The ZZLAS performance was good. The regression slope was less than 4%. Moreover, the average RMSE, MRE, and *R* were 14.38 $W m^{-2}$, −3.72%, and 0.9890, respectively.

[42] Overall, the seven scintillometer sets exhibited good agreement with each other and were consistent with the results reported by *Kleissl et al*. [2008, 2009] except for the BLS450_AR scintillometer. Therefore, the examined scintillometers can be used in HiWATER-MUSOEXE.

4.5. Comparison of Sensible Heat Fluxes Derived From ECs and LASs

[43] In this section, the mean sensible heat fluxes measured by all ECs and LASs were compared; the EC and LAS standard deviation were also computed. Near neutral stratification, the sensible heat fluxes may have some uncertainties, especially for the LAS measurements. Therefore, the regression line was shown for sensible heat fluxes measured by ECs and LASs (H_{EC} and H_{LAS}) larger than 50 $W m^{-2}$, and the values less than 50 $W m^{-2}$ were only plotted. The results are shown in Figure 13. The EC and LAS measurements were consistent with each other; the data points were distributed around the 1:1 line. The LAS measurements were slightly larger than the EC measurements. Moreover, the regression slope was less than 8%, and the coefficient of determination R^2 was 0.8925. The results indicate that both the EC and

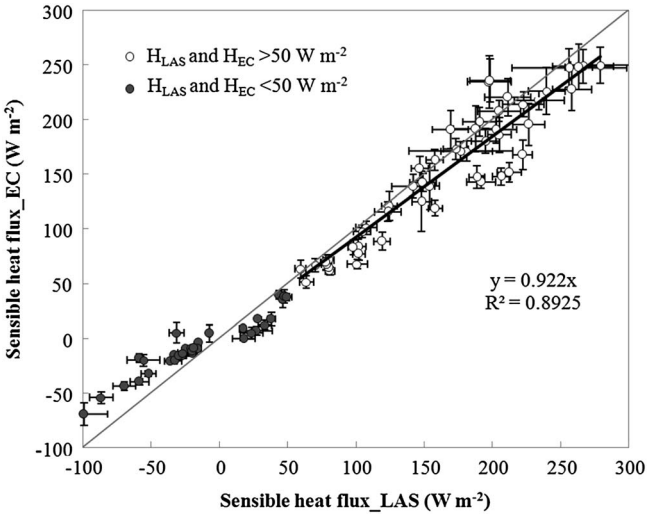

Figure 13. Comparison of sensible heat fluxes between the LAS mean and the EC mean (a wind sector of 315–0 and 0–45). The error bar was calculated as the standard deviation.

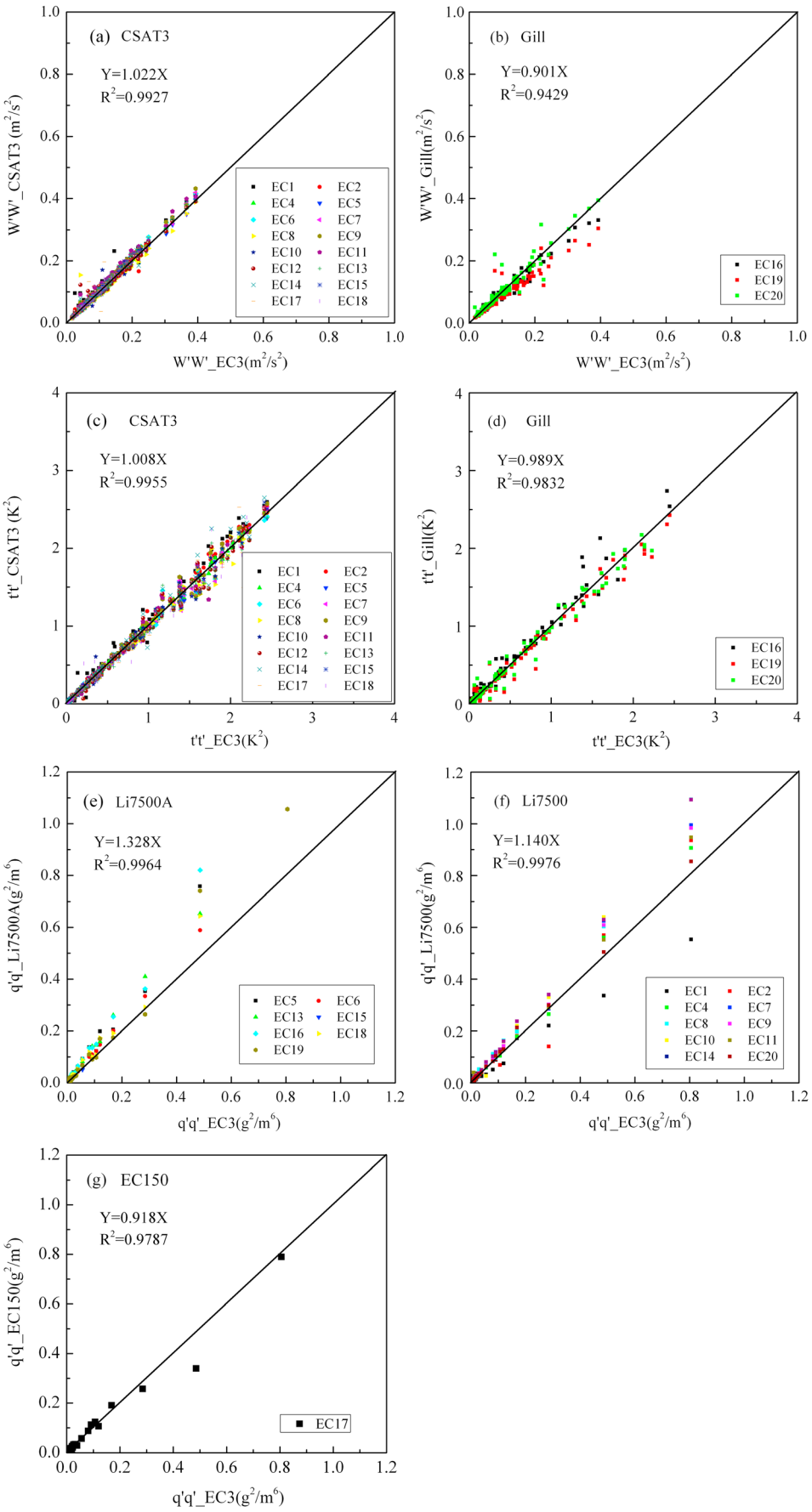
(a) CSAT3
Y=1.022X
R2=0.9927
W'W'_CSAT3 (m2/s2)
W'W'_EC3(m2/s2)
EC1 EC2 EC4 EC5 EC6 EC7 EC8 EC9 EC10 EC11 EC12 EC13 EC14 EC15 EC17 EC18
(b) Gill
Y=0.901X
R2=0.9429
W'W'_Gill(m2/s2)
W'W'_EC3(m2/s2)
EC16 EC19 EC20
(c) CSAT3
Y=1.008X
R2=0.9955
t't'_CSAT3 (K2)
t't'_EC3(K2)
EC1 EC2 EC4 EC5 EC6 EC7 EC8 EC9 EC10 EC11 EC12 EC13 EC14 EC15 EC17 EC18
(d) Gill
Y=0.989X
R2=0.9832
t't'_Gill(K2)
t't'_EC3(K2)
EC16 EC19 EC20
(e) Li7500A
Y=1.328X
R2=0.9964
q'q'_Li7500A(g2/m6)
q'q'_EC3(g2/m6)
EC5 EC6 EC13 EC15 EC16 EC18 EC19
(f) Li7500
Y=1.140X
R2=0.9976
q'q'_Li7500(g2/m6)
q'q'_EC3(g2/m6)
EC1 EC2 EC4 EC7 EC8 EC9 EC10 EC11 EC14 EC20
(g) EC150
Y=0.918X
R2=0.9787
q'q'_EC150(g2/m6)
q'q'_EC3(g2/m6)
EC17

Figure 14

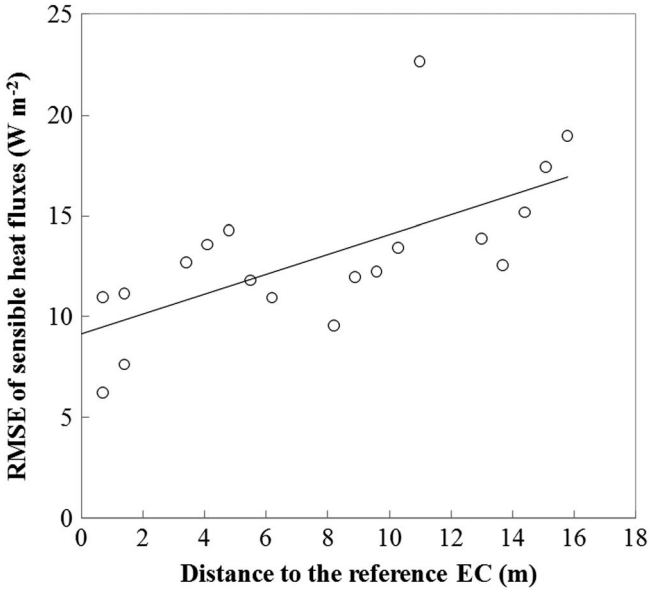

Figure 15. The RMSE variation in sensible heat fluxes as a function of the distance from the reference EC.

LAS system measurements were reliable and are comparable during HiWATER-MUSOEXE.

5. Discussion and Conclusions

[44] The radiometer intercomparison results showed that the five radiometer types (PSP and PIR, CNR1, CNR4, NR01, and Q7) used in HiWATER-MUSOEXE agree well with each other. The four-component radiation discrepancy was separately investigated, showing that the average RMSE values of incoming/outgoing shortwave radiation and the incoming/outgoing longwave radiation were 7.98/3.95 W m^{-2} and 6.43/3.16 W m^{-2}, respectively; the average MRE values were 0.97%/2.07% and 0.62%/−0.34%, respectively. The underlying surface differences in the three radiometer lines caused the differences in outgoing shortwave/longwave radiation between the radiometers. The accuracy of the net radiation measurements is very important to researchers. The difference was estimated of 25 W m^{-2} (4% of the maximum value); the average RMSE and MRE were 10.38 W m^{-2} and 1.24%, respectively. Another important issue is the temporal stability of the radiation sensors. In future field campaigns, radiometers should be regularly checked and periodic comparisons should be conducted for long-term measurements.

[45] The radiometers with relative large discrepancies (e.g., CNR4-ID 100414, NR01, and Q7) were calibrated using a linear regression method in a follow-up study; these radiometers were installed in the periphery of the kernel experimental area (5.5 km × 5.5 km), particularly outside the 3 × 3 MODIS pixel area during HiWATER-MUSOEXE. The three radiometers not involved in the intercomparison campaign were compared to the radiometer installed on the superstation (PSP and PIR) before implemented for formal observations.

[46] Good correlations in the measured sensible and latent heat fluxes were observed for different ECs (CSAT3/Gill and Li7500/Li7500A/EC150); the average R values were 0.99 and 0.89, respectively. Good consistency was found for sensible heat fluxes observed for the EC with CSAT3 sonic anemometers (the average RMSE and MRE were 12.30 W m^{-2} and −1.36%, respectively); the EC with a Gill sonic anemometer also exhibited good consistency (the average RMSE and MRE were 16.75 W m^{-2} and −5.52%, respectively). Moreover, the average RMSE and MRE for all EC measurements were 13.00 W m^{-2} and −2.02%, respectively. The latent heat fluxes were very small and had relatively large fluctuations and possibly large uncertainties in the drying Gobi desert. Therefore, we primarily focused on analyzing sensible heat fluxes in this study. The average RMSE and MRE for latent heat flux were 4.47 W m^{-2} and 0.11% for all of the EC measurements. The variance of the measured turbulent quantities, i.e., vertical wind speed, sonic temperature, and water vapor density, are compared in Figure 14. Compared with the reference EC, the discrepancy in the sonic temperature variance ($t't'$) measured by the Gill and CSAT3 sonic anemometers was similar (the regression slope was less than 2%). The vertical wind speed variance ($w'w'$) measured by the Gill sonic anemometer was less than the CSAT3 measurements (Figure 14b). Moreover, the discrepancy in sensible heat fluxes between the Gill and CSAT3 sonic anemometer measurements may be related to the vertical wind speed discrepancy (Figure 14). Although the correlation in the H_2O variance ($q'q'$) was very good (R was larger than 0.99), the discrepancy was relatively large; the regression slopes for Li7500A, Li7500, and EC150 were approximately 32%, 14%, and 9%, respectively. Furthermore, the water vapor density was low, and the measurements from different ECs were not in good agreement. All these reasons can cause the large discrepancy in the latent heat fluxes. However, the discrepancy between EC systems is actually a mix of discrepancies in the systematic differences and random components due to the fact they do not see the same eddies, which means the sensors closer to each other could see more often the same eddies. Figure 15 shows the RMSE variation in sensible heat fluxes as a function of the distance from the reference EC, demonstrating that the RMSE increased as the distance increased from the reference. Moreover, this finding also indicated that the close ECs observed the same eddies more often and have small differences. However, the ECs were all located within 20 m during the intercomparison experiment, and the correlation relationships for all ECs compared with the reference were greater than 0.98 (Table 5); the maximum RMSE was approximately 20 W m^{-2} (Figure 15). Therefore, the eddies observed by the 20 EC systems were nearly identical, and

Figure 14. The difference of scalar variance measured by different eddy covariance systems for 14 to 24 May 2012: (a) vertical wind speed variance ($w'w'$) measured by the CSAT3 sonic anemometers, (b) $w'w'$ measured by the Gill sonic anemometers, (c) sonic temperature variance ($t't'$) measured by the CSAT3 sonic anemometers, (d) $t't'$ measured by the Gill sonic anemometers, (e) H_2O variance ($q'q'$) measured by the Li7500A infrared gas analyzer, (f) $q'q'$ measured by the Li7500 infrared gas analyzer, and (g) $q'q'$ measured by the EC150 infrared gas analyzer.

the discrepancy primarily came from systematic differences between each EC and the reference in this study.

[47] The measurement accuracy of the CO_2/H_2O infrared gas analyzer also depends on the calibration procedure. To ensure the consistency of the EC systems in the formal observations, the abnormal CO_2/H_2O infrared gas analyzer sensor was recalibrated for zero drift (e.g., EC12). Moreover, the EC systems with Gill sonic anemometers (EC16, EC19, and EC20) were placed along the periphery of the kernel experimental area (5.5 km × 5.5 km), particularly outside the 3 × 3 MODIS pixel area during HiWATER-MUSOEXE. The two EC sets not involved in the intercomparison campaign were compared to the EC installed in the superstation (no.5 in the intercomparison campaign) prior to installation.

[48] There was good agreement among the seven scintillometer set, with high correlation coefficients (all R values for the LASs were larger than 0.98). The best performance was found for the BLS900 measurements (the average RMSE and MRE were 10.26 W m^{-2} and 5.48%, respectively). The present study also obtained good consistency for ZZLAS measurements (the average RMSE and MRE were 14.38 W m^{-2} and −3.72%, respectively). For the BLS450 measurements, the BLS_NQ measurements were consistent with the reference (the RMSE and MRE were 11.16 W m^{-2} and 5.60%, respectively), whereas the BLS450_AR measurements showed a relatively large discrepancy (the RMSE and MRE were 21.48 W m^{-2} and 15.33%, respectively). To assure the consistency of the scintillometers, BLS450_AR was recalibrated before installation. The scintillometer (LAS, Kipp & Zonen) not involved in the comparison campaign was compared to the BLS900 (BNU2) measurement prior to installation. To ensure data continuity, the eight scintillometer sets were divided into four groups during HiWATER-MUSOEXE. Each group contained two scintillometer sets with the transmitter and receiver reversed. Moreover, the three BLS900 sets were installed within the 3 × 3 MODIS pixel area.

[49] The comparison of average sensible heat fluxes measured by all ECs and LASs indicated that the EC and LAS measurements agreed well for homogeneous surfaces (the regression slope was less than 8% and R^2 was 0.8932); the results were consistent with those reported by *Zeweldi et al*. [2010] and *Liu et al*. [2011]. The discrepancy may be primarily caused by the heterogeneity of the underlying surfaces, the larger source area of LAS, and possibly more contribution of larger turbulent structures to the sensible heat flux (the energy balance closure was approximately 0.93 at the Gobi site) [*Liu et al*., 2011]. The results indicated that the EC and LAS measurements were comparable during HiWATER-MUSOEXE.

[50] High correlation coefficients between different instrument intercomparisons (radiometers, ECs, and LASs) suggest that these instruments performed well. However, careful data processing and a detailed analysis should be considered to ensure good performance. Most CNR4-type radiometers exhibited good agreement with PSP and PIR radiometer; the CNR1-type also performed fairly well. Relative large discrete were found in NR01- and Q7-type radiometers for net radiation. The PSP and PIR-, CNR4-, and CNR1-type radiometers are recommended for fundamental radiation flux research if correctly operated. If proper AOA correction was applied on the Gill anemometer, the EC systems with both the CSAT3 anemometer and Gill anemometer performed well for sensible heat flux measurements; the CSAT3 measurements exhibited better agreement. The EC with relatively new combinations of CSAT3 and Li7500A/EC150 operated well. Moreover, a careful intercomparison in relatively wet conditions should be conducted in the future. The sensible heat fluxes measured by the LAS of BLS series were reliable, especially for BLS900.

[51] Although most radiometers, ECs, and LASs were consistent with each other, there were some deviations. Therefore, an instrument intercomparison study should be conducted as an integral part of experimental design and post data processing and analyzing to provide valuable baseline information for the observed data comparison at different sites. Due to the surface characteristic and environmental variability, the instrument uncertainty found in the intercomparison campaign cannot completely characterize the HiWATER-MUSOEXE experimental period. Therefore, a continuous uncertainty evaluation for flux observations is necessary in the further analysis to characterize the measurement uncertainty accurately over time and changing environmental conditions [*Alfieri et al*., 2011].

[52] **Acknowledgments.** We thank the two anonymous reviewers for their valuable comments, which substantially improved our paper. We also thank all the scientists, engineers, and students who participated in HiWATER field campaigns, especially for the instrument providers of Zhenghui Xie, Li Jia, Zhongjing Wang, Bingfang Wu, Xiaoping Xin, Xiaomin Sun, Zhongqin Li, Guoyu Qiu, Zhiqiu Gao, Jianping Huang, and Qiang Zhang. This work was supported by the National Natural Science Foundation of China (91125002), the Chinese Academy of Sciences (KZCX2-XB3-15), and the National Natural Science Foundation of China (41301355).

References

Alfieri, J. G., W. P. Kustus, J. H. Prueger, L. E. Hipps, J. L. Chávez, A. N. French, and S. R. Evett (2011), Intercomparison of nine micrometeorological stations during the BEAREX08 field campaign, *J. Atmos. Oceanic Technol.*, *28*, 1391–1406.

Blanken, P. D., T. A. Black, H. H. Neumann, C. D. Hartog, P. C. Yang, Z. Nesic, R. Staebler, W. Chen, and M. D. Novak (1998), Turbulence flux measurements above and below the overstory of a boreal aspen forest, *Boundary-Layer Meteorol*, *89*, 109–140.

Field, R. T., L. J. Fritschen, E. T. Kanemasu, E. A. Smith, J. B. Stewart, S. B. Verma, and W. P. Kustas (1992), Calibration, comparison, and correction of net radiation instruments used during FIFE, *J. Geophys. Res.*, *97*, 18,681–18,695.

Foken, T., and B. Wichura (1996), Tools for quality assessment of surface-based flux measurements, *Agric. For. Meteorol*, *78*, 83–105.

Jia, Z. Z., S. M. Liu, Z. W. Xu, Y. J. Chen, and M. J. Zhu (2012), Validation of remotely sensed evapotranspiration over the Hai River Basin, China, *J. Geophys. Res.*, *117*, D13113, doi:10.1029/2011JD017037.

Kleissl, J., J. Gomez, S. H. Hong, J. M. H. Hendrickx, T. Rahn, and W. L. Defoor (2008), Large aperture scintillometer intercomparison study, *Boundary-Layer Meteorol.*, *128*, 133–150.

Kleissl, J., C. J. Watts, J. C. Rodriguez, S. Naif, and E. R. Vivoni (2009), Scintillometer intercomparison study-continued, *Boundary-Layer Meteorol.*, *130*, 437–443.

Kohsiek, W., C. Liebethal, T. Foken, R. Vogt, S. P. Oncley, C. Bernhofer, and H. A. R. Debruin (2007), The energy balance experiment EBEX-2000. Part III: Behaviour and quality of the radiation measurements, *Boundary-Layer Meteorol.*, *123*, 55–75.

Li, X., et al. (2013), Heihe Watershed Allied Telemetry Experimental Research (HiWATER): Scientific objectives and experimental design, *Bull. Am. Meteorol. Soc.*, *94*(8), 1145–1160.

Liu, S. M., Z. W. Xu, W. Z. Wang, Z. Z. Jia, M. J. Zhu, J. Bai, and J. M. Wang (2011), A comparison of eddy-covariance and large aperture scintillometer measurements with respect to the energy balance problem, *Hydrol. Earth Syst. Sci.*, *15*, 1291–1306.

Liu, S. M., Z. W. Xu, Z. L. Zhu, Z. Z. Jia, and M. J. Zhu (2013), Measurements of evapotranspiration from eddy-covariance systems and large aperture scintillometers in the Hai River Basin, China, *J. Hydrol.*, *487*, 24–38.

Mauder, M., S. P. Oncley, R. Vogt, T. Weidinger, L. Ribeiro, C. Bernhofer, T. Foken, W. Kohsiek, H. A. R. De Bruin, and H. P. Liu (2007), The energy balance experiment EBEX-2000. Part II: Intercomparison of eddy-covariance sensors and post-field data processing methods, *Boundary-Layer Meteorol.*, *123*, 29–54.

Nakai, T., and K. Shimoyama (2012), Ultrasonic anemometer angle of attack errors under turbulent conditions, *Agric. For. Meteorol.*, *162–163*, 14–26.

Nie, D., E. T. Kanemasu, L. J. Fritschen, H. L. Weaver, S. E. A. Smith, S. B. Verma, R. T. Field, W. P. Kustas, and J. B. Stewart (1992), An intercomparison of surface energy flux measurement systems used during FIFE 1987, *J. Geophys. Res.*, *97*, 18,715–18,724.

Ochs, G. R., and J. J. Wilson (1993), A second-generation large aperture scintillometer, NOAA Tech. Memor. ERL ETL-232, NOAA Environmental Research Laboratories, Boulder, CO USA, 24 pp.

Ohmura, A., et al. (1998), Baseline Surface Radiation Network (BSRN/WCRP): New precision radiometry for climate research, *Bull. Am. Meteorol. Soc.*, *79*, 2115–2136.

Shi, S. J., B. X. Huang, S. M. Liu, Y. Yang, Y. B. Huang, and Z. W. Xu (2010), Development of a measuring system for surface energy and water vapor fluxes at large scale [in Chinese with English abstract], *Adv. Earth Sci.*, *25*, 1128–1138.

Wesely, M. L. (1976), The combined effect of temperature and humidity fluctuations on refractive index, *J. Appl. Meteorol.*, *15*, 43–49.

Zeweldi, D. A., M. Gebremichael, J. M. Wang, T. Sammis, J. Kleissl, and D. Miller (2010), Intercomparison of sensible heat flux from large aperture scintillometer and eddy covariance methods: Field experiment over a homogeneous semi-arid region, *Boundary-Layer Meteorol*, *135*, 151–159.