Validation of evapotranspiration and its long-term trends in the Yellow River source region

Rong Liu, Jun Wen, Xin Wang and Zuoliang Wang

ABSTRACT

In this paper, the ground observations were compared to the ERA-Interim, NCEP-DOE AMIP-II Reanalysis, MODIS ET product, and emerging offline SEBS ET data sets in the Yellow River source region of the Tibet Plateau. In general, the slopes of linear least squares exhibit differences, with ERA-Interim, NCEP–DOE, MOD16, and SEBS slopes of 0.88 ± 0.05, 0.64 ± 0.07, 0.66 ± 0.17, and 1.24 ± 0.97 respectively. ERA-Interim was found superior with ground observations to others; therefore, it provided a good representation of the study area. Based on the ERA-Interim ET product, the Sen's slope estimator and the Mann–Kendall (MK) test were applied to quantify the significance of the shifts in trends, while the moving t-test and MK test characterized abrupt changes. The results show that the Yellow River source region experienced a statistical increase in evapotranspiration (ET) in the northern part and a decrease in the southern part of the region from 1979 to 2014 at rates of approximately 1.65 and −0.50 mm/yr, respectively. The shift in the annual ET trend was more pronounced, and abrupt changes were detected in the 1980s. Precipitation was the most dominant factor affecting ET variation, whereas surface temperature was the least influential.

Key words | evapotranspiration, long-term trend, reanalysis data sets, Yellow River source region

Rong Liu
Jun Wen (corresponding author)
Xin Wang
Zuoliang Wang
Key Laboratory of Land Surface Process and Climate Change in Cold and Arid Regions, Cold and Arid Regions Environmental and Engineering Research Institute,
Chinese Academy of Sciences,
Lanzhou 730000,
China
E-mail: *jwenen@lzb.ac.cn*

Jun Wen
College of Atmospheric Sciences,
Plateau Atmosphere and Environment Key Laboratory of Sichuan Province,
Chengdu University of Information Technology,
Chengdu, Sichuan 610225,
China

INTRODUCTION

Evapotranspiration (ET) is one of the main hydrological processes. One of the distinguishing factors of ET is its role as a linchpin between the energy and water cycles. The physics and dynamics of ET are crucial for understanding the role of the hydrological cycle in the climate system (Mojid *et al.* 2015). Although it is possible to observe daily ET dynamics on a local scale, it is challenging to translate these point measurements to natural landscapes because of the lack of ET variability over large spatial scales and long temporal series. Over the past decade, reanalysis data, models, and satellite remote sensing products have been developed as potential tools to produce spatially and temporally consistent ET information, and these tools have been widely applied (Mendoza *et al.* 2008; Mo *et al.* 2014; Su *et al.* 2014). However, the estimates these tools produce require validation with intensive ET measurements to quantify the feedback between the water cycle and climate change.

In recent years, operational centers at the European Centre for Medium-Range Weather Forecasts (ECMWF) and the US National Centers for Environmental Prediction (NCEP) have produced uniform, long-term reanalysis data sets (Mooney *et al.* 2011). Jiménez *et al.* (2011) found that the ERA-Interim and NCEP-Department of Energy (NCEP-DOE) reanalysis data sets of ET contain fundamental observational differences. Large ET differences were observed in rainforest regions, and some of the lowest correlations were associated with the NCEP-DOE reanalysis. The Moderate Resolution Imaging Spectroradiometer (MODIS) global terrestrial ET (MOD16) algorithm is based on the Penman–Monteith (P-M) equation. MOD16 products were validated at 17 flux tower locations in Asia and showed

doi: 10.2166/wcc.2017.134

overestimations due to non-closure of the energy balance (Kim *et al.* 2012). The Surface Energy Balance System (SEBS) developed by Su (2002) was designed to estimate energy partitioning using satellite and meteorological data. It has been used to estimate ET on a regional scale and contains a physically based turbulent flux parameterization scheme that can overcome the shortcomings of statistical methods and produce spatially continuous distributions of land–surface energy fluxes using meteorological forcing data. Cleugh *et al.* (2007) demonstrated that the P-M model is superior to SEBS models that use MODIS remote sensing data to provide regionally distributed ET information in an evergreen forest and a tropical savanna in Australia.

The measurement scale of the Eddy covariance system (EC) is generally a hundred meters, and it is widely used around the world. It can directly measure the water vapor, heat and carbon dioxide flux, and it is considered as the standard method for measuring water vapor and heat fluxes (Liu *et al.* 2011). In recent studies, it was found that the non-closure of the energy balance was explained by the energy fluxes from secondary circulations and larger eddies that cannot be captured by EC measurement at a single station (Liu *et al.* 2013). Many reports have been published about the use of the EC system to measure the energy and water vapor fluxes in a variety of ecosystems, including grasslands (Wever *et al.* 2002). Nevertheless, the EC method has limitations, such as data processing and quality control methods under complex conditions and surface energy imbalance problems in EC observations (Foken *et al.* 2010). Few studies have compared these four data products, especially in the Yellow River source region catchment, where an EC system has provided continuous temporal observations since 2009.

The objective of this paper is to exhibit the temporal-spatial variation pattern of ET over the past 30 years in the Yellow River source region catchment, then to detect the contributions of changes in meteorological variables. The paper is structured as follows. The next section presents the data sources and different methodologies. The spatial structure of the trends given by Sen's slope estimator and the Mann–Kendall (MK) test are presented, and abrupt changes in the ET series detected by the moving t-test and the Sequential Mann–Kendall (SQMK) test are shown. The following section introduces the study area, EC instrumentation. The first subsection in the Results and discussion section present the correlations between ET products and observations. According to the results, ET product from ERA-Interim is superior to the others. This is followed by a subsection giving the spatial–temporal variation of ET from ERA-Interim in the Yellow River source region catchment. The causes of these characteristics are determined, and the contribution of climatic variation to ET is quantified in the next subsection. The final section gives the conclusions. Our work represents the first time that such a comprehensive evaluation has been performed to assess the qualities of the four most widely used data products over the Yellow River source region catchment, which is crucial to the climate and weather of Asia.

DATA AND METHODOLOGY

The global terrestrial ET data from ERA-Interim, NCEP–DOE Reanalysis 2, MODIS, and SEBS are validated in the Yellow River source region catchment. Sen's slope estimator and the nonparametric MK test were applied to detect significant trends. The moving t-test and the SQMK test were applied to detect abrupt temporal trend shifts. The analysis was conducted over the entire Yellow River source region.

Reanalysis data sets and satellite-based ET products

ERA-Interim is the latest global atmospheric reanalysis performed by the ECMWF, and the project is being conducted in part to prepare for a new atmospheric reanalysis that will replace ERA-40. The ERA-Interim system is based on a recent release of the Integrated Forecast System cycle 31r2 (IFS Cy31r2), which contains many improvements to both the forecasting model and analysis methodology. The surface fluxes in ERA-Interim are based on the Tiled ECMWF Scheme for Surface Exchanges over Land model (Jones & Lister 2014) and forced by atmospheric analysis and short-range forecasts. We collected the fluxes as monthly mean values in W/m^2 at a resolution of $0.125 \times 0.125^{\circ}$, which is very close to that of the native ERA-Interim T255 Gaussian reduced grid.

NCEP-DOE Reanalysis 2 is an improved version of the NCEP-National Center for Atmospheric Research (NCEP-NCAR) Reanalysis I model that fixed errors and updated parameterizations of physical processes (You *et al.* 2013). Unlike the NCEP-NCAR reanalysis, the NCEP-DOE reanalysis utilizes the pentad mean observed precipitation to correct model precipitation in driving the soil model, leading to a more realistic evolution of soil moisture. NCEP-DOE Reanalysis 2 fluxes were provided as daily averages in W/m^2 at a resolution of $2.0 \times 2.0^{\circ}$ and are available from 1979 to the present.

The MOD16 ET product is a land surface ET product that provides critical information regarding global terrestrial water and energy cycles and environmental changes. It expresses all transpiration based on vegetation and evaporation from canopy and soil surfaces in 1-dimensional vertical units (mm/day). Mu *et al.* (2007) adapted Cleugh's PM-based MODIS ET approach and produced the first regular 1-km^2 land surface ET data set for 109.03 million km^2 of global vegetated land area at 8-day intervals. Mu *et al.* (2011) further improved the MODIS ET approach and evaluated its ability to generate global ET data products.

The SEBS model (Su 2002) uses a variety of satellite-based sensor data and meteorological forcing data to estimate monthly energy and water fluxes across China. The forcing data can come from satellite-based or reanalysis data sets, while the other data can be obtained from meteorological forcing data sets (e.g. Global Land Data Assimilation Systems (GLDAS), ERA, and NCEP reanalysis data). Various temporal and spatial gap-filling algorithms have been developed to produce continuous monthly data from 2001 to 2010 at a $0.1 \times 0.1^{\circ}$ spatial resolution using multi-satellite and meteorological forcing data, and detailed information about each input variable is given in the report by Chen *et al.* (2014).

Trend

Sen's slope estimator

Sen's slope estimator has been widely applied to hydrometeorological time series. Sen (1968) developed the nonparametric procedure as a method of estimating the slope of a trend in a sample of N pairs of data:

$$Q_i = \frac{X_j - X_k}{j - k} \quad for\ i = 1, \ldots, N \tag{1}$$

where X_j and X_k are the data values at times j and k ($j > k$), respectively.

If there is only one data point in each time period, then $N = n(n-1)/2$, where n is the number of time periods. If there are multiple observations in one or more time periods, then $N\langle n(n-1)/2$, where n is the total number of observations.

The N values of Q_i are ranked from smallest to largest, and the median of the slope estimated by Sen's slope estimator is computed as follows (Gocic & Trajkovic 2013):

$$Q_{med} = \begin{cases} Q[(N+1)/2], & if\ N\ \ is\ \ odd \\ \dfrac{Q[N/2] + Q[(N+2)/2]}{2}, & if\ N\ \ is\ \ even \end{cases} \tag{2}$$

The sign of Q_{med} reflects the trend in the data, and its magnitude indicates the steepness of the trend. To determine whether the median slope is significantly different from zero, one should obtain a confidence interval of Q_{med} at a specific probability.

For more detailed information, please refer to Gocic & Trajkovic (2013).

Mann–Kendall test

The nonparametric MK test is recommended by the World Meteorological Organization to assess trends in environmental time series data (Yu *et al.* 2002). The method has been widely used for trend detection in meteorological and hydrological data sets (Samba & Nganga 2012). This test has the advantage of not assuming any special form of the distribution function of the data, while having nearly as much power as its parametric competitors.

The computation of the MK S-statistic value from the raw data can yield a large positive or negative value of S, indicating a positive or negative trend, respectively. The null hypothesis (H_0) of the MK test assumes that the ranked data (X_c, $c = 1, 2, 3, \ldots, n-1$) and (X_d, $d = c+1, \ldots, n$) belong to a sample of n independent and identically

distributed random variables (Nalley *et al.* 2012). The alternative hypothesis (H_1) of the two-sided test assumes that the distributions of X_c and X_d are not identical for all c, and $d \leq n$ with $c \neq d$. The S-statistic of the MK test can be expressed as follows:

$$S_t = \sum_{c=1}^{n-1} \sum_{d=c+1}^{n} sign(X_d - X_c) \tag{3}$$

$$Sign_t(X_d - X_c) = \begin{cases} +1, & X_d > X_c \\ 0, & X_d = X_c \\ -1, & X_d < X_c \end{cases} \tag{4}$$

where X_c and X_d denote the ranked values of the data and n is the length of the data record. For data that are distributed independently with a zero mean, the variance of the S_t statistic is as follows:

$$\text{Variance}\,(S_t) = \frac{\{n(n-1)(2n+5) - \sum_{c=1}^{n} t_c(c)(c-1)(2c+5)\}}{18} \tag{5}$$

where t_c represents the summation of t, which is the number of values associated with the extent of c. The Z statistic of the MK test is computed as follows:

$$Z = \begin{cases} \dfrac{S_t - 1}{\sqrt{Var(S_t)}} & S_t > 0 \\ 0 & S_t = 0 \\ \dfrac{S_t + 1}{\sqrt{Var(S_t)}} & S_t < 0 \end{cases} \tag{6}$$

In Equation (6), the Z statistic of the MK test can be used when the number of records n is larger than 10. The trend's significance is assessed by comparing the Z value with the standard normal variation at the pre-specified level of statistical significance (Hamed 2008). In a two-sided trend test, with α representing the significance level, the null hypothesis should not be accepted if $|Z| > Z_{\alpha/2}$, suggesting that the trend is significant. A positive Z value at the significance level implies a positive trend, whereas a negative value indicates a negative trend. The significance of a trend can be verified by the probability value (p-value) obtained from the MK Z value. If the p-value is less than the predetermined significance level (e.g. $\alpha = 5\%$) or greater than the confidence level (if $\alpha = 5\%$, confidence level = 95%), the null hypothesis of the trend cannot be accepted.

Abrupt change

Moving t-test

The moving t-test detects abrupt changes in a series by assessing the significant difference between the averages of two groups of samples. For an n-sample series, selecting a data point, we obtain subsequence X_1 of n_1 samples before the data point, with an average of $\overline{X_1}$ and a variance of S_1^2, and subsequence X_2 of n_2 samples after the data point, with an average of $\overline{X_2}$ and a variance of S_2^2. The statistic t is as follows (Fraedrich *et al.* 1997):

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} / (1/n_1 + 1/n_2)} \tag{7}$$

Given a significance level α, the null hypothesis of no difference will be rejected if $|t| > t_{\alpha/2}$. Because different choices of the lengths of the subsequences can affect the location of abrupt change, we try two conditions: $n_1 = n_2 = 5$ and $n_1 = n_2 = 10$. In this study, the typical significance level of 0.05 was used. All points satisfying $|t| > t_{\alpha/2}$ constitute the time scope of an abrupt change, and in a year with maximum $|t|$, an abrupt change may appear (Nalley *et al.* 2012). The difference between the 10-year averages before and after the abrupt change is the magnitude of change. The abrupt changes determined by the MK method cannot represent the 10-year periods before and after the change point; thus, all magnitudes in this study are calculated using the moving t-test.

Sequential Mann–Kendall (SQMK) test

The application of the SQMK test has the following four steps in sequence (Zang & Liu 2013).

During each comparison, the number of cases $X_i > X_j$ is counted and indicated by n_i, where X_i ($i = 1, 2, \ldots, n$) and X_j ($j = 1, \ldots, i-1$) are the sequential values in a series.

The test statistic t_i of the SQMK test can be expressed as follows:

$$t_i = \sum_i n_i \tag{8}$$

The mean and variance of the test statistic are given in Equations (9) and (10), respectively:

$$E(t) = \frac{n(n-1)}{4} \tag{9}$$

$$Var(t_i) = \frac{i(i-1)(2i+5)}{72} \tag{10}$$

Sequential progressive values can be expressed as follows:

$$u(t) = \frac{t_i - E(t)}{\sqrt{Var(t_i)}} \tag{11}$$

Similarly, sequential backward ($u'(t)$) analysis of the MK test is calculated starting from the end of the time series data.

STUDY AREA AND GROUND OBSERVATION

The Yellow River source region is the most important catchment in the Yellow River, accounting for 15% of the area but providing over 40% of the entire watershed runoff (Liu *et al.* 2015). Figure 1 shows a map of the Yellow River source region catchment. The six stations represent different catchments in the Yellow River source region. MaduoX and Maduo represent the headwaters of the Yellow River source region; Dari, Jiuzhi, and Maqu are considered the primary runoff stations; and Tangnaihai is a typical station located at the outlet of the Yellow River source region. Areas of mixed bushes and grasses surround the source region. During the experimental period, the Yellow River source region was covered primarily by grass; however, it has complex geomorphic features consisting of diverse landscapes, including mixed forest, shrubland, grassland, sparsely wooded pastures, and bodies of water. Gyaring Lake and Ngoring Lake are the largest lakes in the region.

MaduoX, Maduo, and Maqu are the stations that contain EC instrumentation. The EC instrumentation is oriented to the north, which is the predominant wind direction during the summer. For quality assurance, the EC

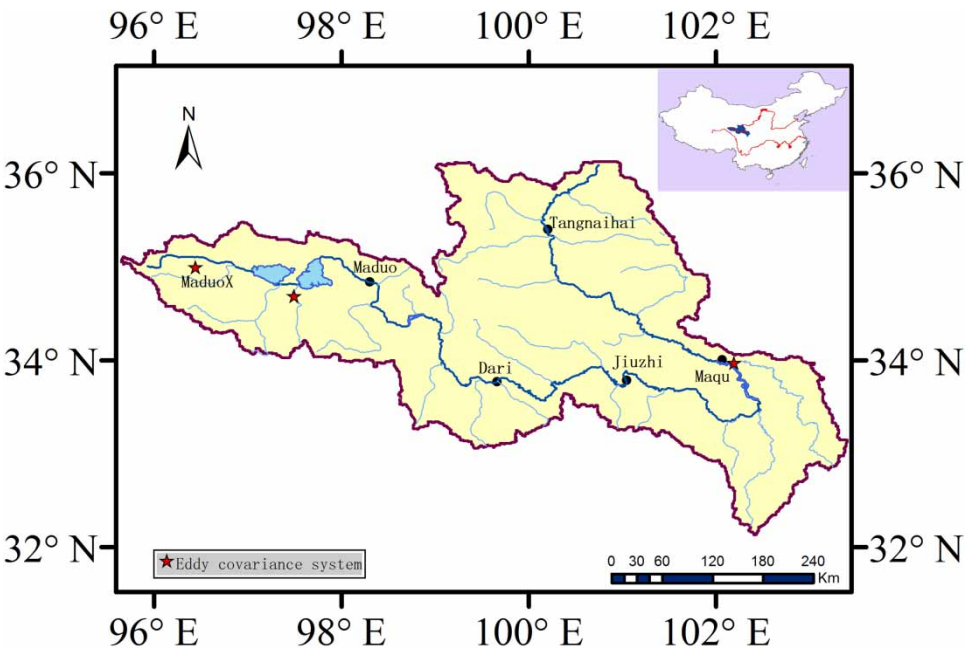

Figure 1 | Map of the observation site and the Yellow River source area (the circles represent the six stations over the different catchments of the Yellow River source region and the stars represent the stations that contain EC instrumentation).

system measurement data were checked routinely. Various filtering techniques were applied during data preprocessing. The flux data were then examined at 10-day intervals, and extreme values and values more than ±3 standard deviations from the mean were removed. For the missing values, the daytime gaps were filled via linear interpolation between the nearest temporal measurements. In addition, an automatic weather station was also installed during the experimental period. The measurements included wind velocity and direction at a height of 2.0 m, as well as temperature and humidity at a height of 1.5 m. All instruments were calibrated after 2 years of continuous observation to ensure the quality of the observational data. Other data, such as precipitation, daily sunshine hours, and daily cloud cover, were also collected. In addition, hourly cloud cover was recorded on days without precipitation. In this paper, the grid point values were compared with flux measurements extracted from ET products during 2010. For the time scale, the data logger recorded an ET value every 30 min, and the daily ET was obtained by accumulation of all 30 min values within 1 day.

RESULTS AND DISCUSSION

The accuracies of remote sensing-based measurements of land surface heat fluxes are questionable without validation using ground-based measurements. To make this comparison possible, different products have been aggregated to common spatial and temporal resolutions. First, the spatial resolution of the products was reduced to the $2.0 \times 2.0^{\circ}$ resolution of the coarser product by spatially averaging the original estimates. Only pixels with fluxes were retained from all products. Next, the products were matched based on units. Reanalysis data sets provides latent heat fluxes in units of W/m^2, these values converted to ET with units of mm. The latent heat flux (λE), which is measured in energy units (W/m^2). E is identical to ET, which is the amount of water transferred from land surface to atmosphere and is measured in water column units. λ is the latent heat of vaporization, $2.49 \times 10^6\ J/kg^{-1}$, these values are converted to ET with units of mm. Finally, the products are temporally matched using observations from the EC instrumentation. For reanalysis data (ERA-Interim and NCEP-DOE Reanalysis 2 data sets) which are daily scale, the data from EC instrument was examined as daily scale. For the ET product from the MOD16 and the SEBS model, the time scales are 8-day and monthly intervals, the data from EC instrument were matched as 8-day and monthly before being compared. This guarantees that differences in the statistics are not due to differences in spatial or temporal coverage. After these operations, the correlations between ET products and observations are compared and analyzed to show which product is more suitable for the study area.

Correlation between $ET_{products}$ and $ET_{observations}$

As is typical at stations operating under Tibetan conditions, flux measurements were compared with grid point values extracted from ET products during 2010. Figure 2(a) and 2(c) compare the time series of $ET_{products}$ and $ET_{observations}$ from 2010. ERA-Interim and NCEP-DOE provide daily steps, whereas the MOD16 and SEBS data sets give 8-day averages and monthly time steps, respectively. In general, the four data sets show good agreement with the observed data and with each other. ERA-Interim showed closer agreement with the observed data than did the other data sets. This indicates that ERA-Interim has a stronger linear relationship with the observations than do the NCEP-DOE, MOD16, or SEBS data sets. In Figure 2(a) and 2(c), the highest value is 5.73 mm on June 29, 2010, but none of the data sets reflected this peak value, and NCEP–DOE illustrated more scatter than the ERA-Interim data set from April to October. The ERA-Interim (Figure 2(a)), NCEP–DOE (Figure 2(a)), and MOD16 (Figure 2(c)) data sets all exhibited increasing trends through May. The values decreased slightly after the middle of May and increased substantially until July, illustrating the peak during the year. After this, the data decreased slowly; however, these data sets showed a gradual increase in September. The SEBS data set (Figure 2(c)) was relatively high throughout the whole year, which could be caused by the coarse temporal sampling of the station's satellite sensor.

Figure 2(b) and 2(d) show scatter plots of the ERA-Interim, NCEP–DOE, MOD16, and SEBS data sets versus the observed data at the Maqu Climate and Environment Comprehensive Observation Station during 2010. Linear

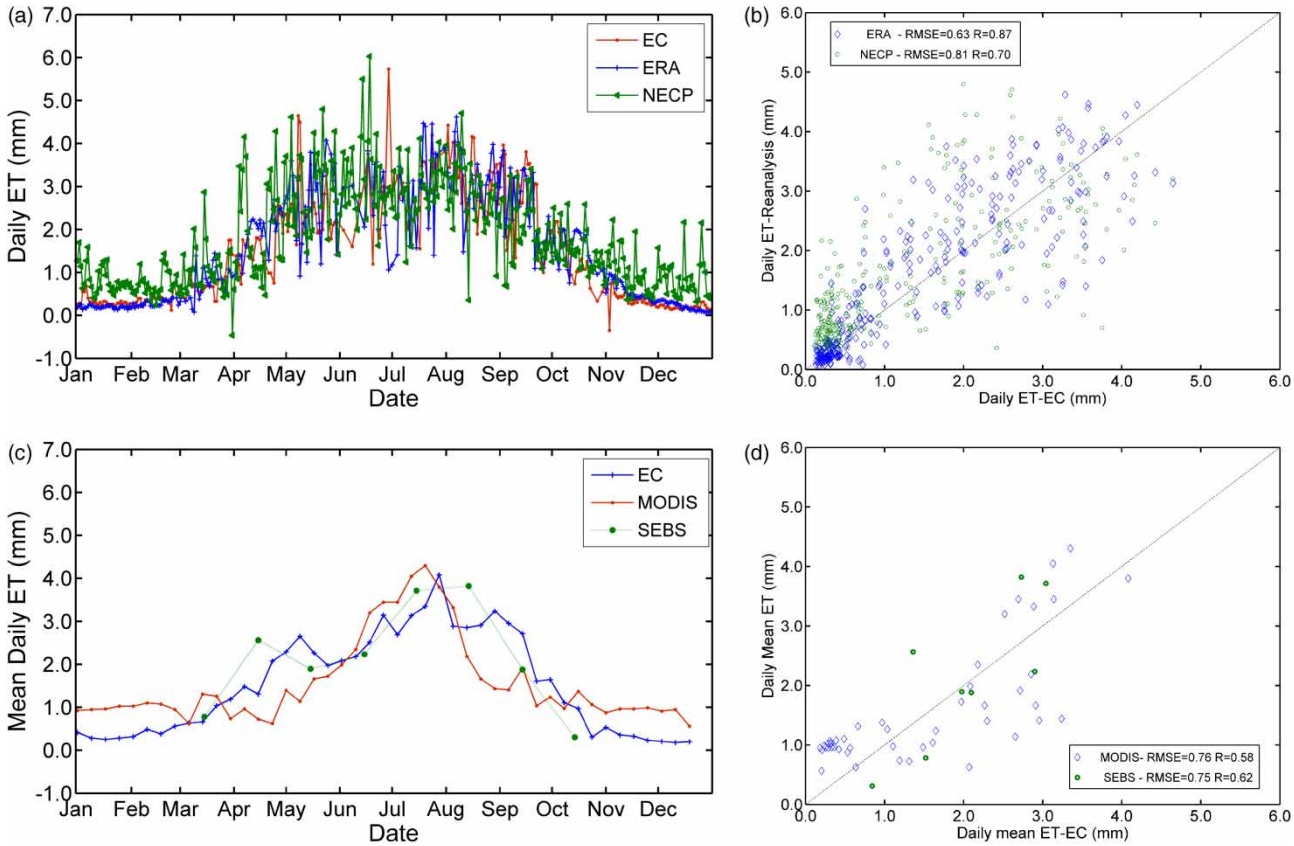

Figure 2 | Comparison of time series of $ET_{products}$ and $ET_{observations}$ from 2010.

least squares fitting is also shown for each of the four data sets. Slopes of the fitting lines based on scatter plots of these data versus observational data show differences, with slopes of the ERA-Interim, NCEP–DOE, MOD16, and SEBS ET data sets of 0.88 ± 0.05, 0.64 ± 0.07, 0.66 ± 0.17, and 1.24 ± 0.97, respectively. The lines fitted to the data sets could have slopes equal to, less than, or greater than 1.0. Slopes equal to 1.0, in conjunction with an intercept value of 0.0, imply that the reanalysis is in perfect agreement with the observations. If the slope is less than 1.0, the reanalysis models indicate warmer winters and colder summers than those in the observations. Slopes greater than 1.0 indicate that compared with the observations, the reanalysis values reflect warmer summers and colder winters.

We also analyzed other statistics (Table 1). The sum of squared errors (SSE) indicates the deviation of the responses from the fitted values of the responses. A value closer to 0 suggests a better fit. The SSEs of ERA-Interim, NCEP–DOE, MOD16, and SEBS are 147.8, 239.7, 1190, and 3701, respectively. The R-square value (R) measures how successful the fit is in explaining the variation in the data. A value closer to 1 suggests a better fit. The R value of ERA-Interim is 0.87, which is significantly higher than the root mean square error (RMSE) of this product. The RMSE represents the standard deviation of the differences between predicted values and observed values. A value closer to 0 suggests a better fit. The results show that ERA-Interim has an RMSE value of slightly less than

Table 1 | Error analysis of four products vs. observations

Product	Regression equation	SSE (mm)	R	RMSE
ERA-Interim	$y = (0.88 \pm 0.05)X + 0.18$	147.8	0.87	0.63
NCEP–DOE	$y = (0.64 \pm 0.07)X + 0.82$	239.7	0.70	0.81
MOD16	$y = (0.66 \pm 0.17)X + 4.494$	1190	0.58	0.76
SEBS	$y = (1.24 \pm 0.97)X - 11.9$	3701	0.62	0.75

SSE: sum of squares due to error, R: coefficient of multiple determination, RMSE: root-mean-square error.

0.63 mm, which is lower than the RMSE values of the other products.

Spatial–temporal variation of ET_{ERA} in the Yellow River source region catchment

The ability to capture the inter- and intra-annual variations of ET is of interest to climate change researchers. Understanding variations in ET is essential for studies of water cycle and water resource issues. Comparisons between the measured and modeled fluxes show that the ERA-Interim ET product is more accurate than other data sets in the Yellow River source region catchment. We calculated the trends and abrupt ET changes in ERA-Interim products from 1979 to 2014 using Sen's slope estimator, the nonparametric MK test, the moving t-test, and the SQMK test.

Trend variation in annual ET series

To assess the spatial variation in ET trends in the Yellow River region, Sen's slope estimator was used to calculate the trends in the annual ET at $0.125 \times 0.125^{\circ}$ in latitude and longitude (Figure 3(a)). The ET trends in the periods of May–June (spring; Figure 3(b)), July–August (summer; Figure 3(c)), and September–October (autumn; Figure 3(d)) are also illustrated.

In general, the results of the ET trends using Sen's slope estimator can be divided into three parts. Significant positive trends are found mostly in the northern part of the study area in the upper reaches and in the southeast part of the study area, whereas significant negative trends are found in the western part of the study area. For the annual ET, the trend at MaduoX, which is in the headwaters of the Yellow River source region catchment located in the western part of the study area, has the lowest value (−0.10 mm/yr) because of its high elevation (over 4,450 m). The trend at Tangnaihai, which is located in the downstream area of the Yellow River source region catchment, has the highest value at 3.50 mm/yr, which is five times higher than the regional average. The fact that most tendencies are positive indicates that an increasing trend is universal for the annual ET in the Yellow River source region catchment, which is probably an indication of the extension of the regional hydrological

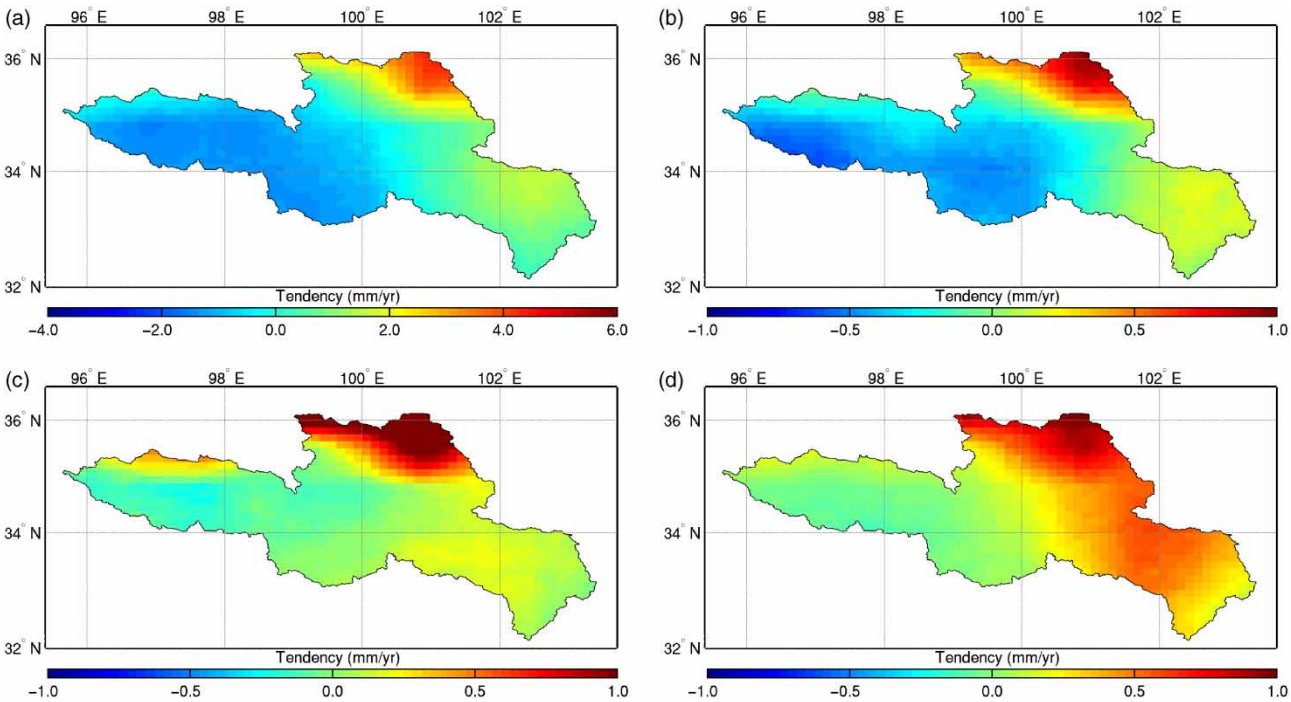

Figure 3 | Distribution of yearly ET trends by the Sen's slope analysisfrom 1979 to 2014 in the Yellow River region catchment: (a) annual, (b) spring, (c) summer, and (d) autumn.

cycle. The upper reach contributes most to the increasing trend in the annual regional ET.

During the growing season, the ET trend is more obvious than the annual value. The highest increase was found in the downstream area of the Yellow River source region catchment, with a value of 1.00 mm/yr. The trend at Maqu station, which is located at the bend in the Yellow River source region, was 0.20 mm/yr in the spring and summer series. Notably, the ET during the spring in the southwestern part of the region decreased significantly, with a trend of approximately −0.50 mm/yr, whereas ET in the northern part of the study area during the summer experienced a large increase of more than 1.00 mm/yr. Additionally, Maqu station showed the greatest increase during the autumn, with a value approximately twice that of the annual value.

The MK method was applied to the series of annual average fluxes to assess variations from 1979 to 2014. Areas with high MK values have increasing ET concentrations, and areas with low MK values generally exhibit decreasing trends. Therefore, the resulting slope indicates that ET generally increased, both annually and during the spring, in the studied decades over the majority of the northern part of the Yellow River source region catchment, whereas the strongest decreasing trend occurred in the southwestern area and expanded slowly toward the southeast (Figure 4(a) and 4(b)). The annual ET values exhibited varying temporal trends, which can be characterized as spatially heterogeneous among the study areas (Figure 4(a)). The results show the number of ET gauges that demonstrate trends (positive and negative) at three different significance levels. The majority of the headwaters of the Yellow River source region catchment (western part of Figure 4) experienced a decreasing MK trend, although these findings are not statistically significant at a significance level of 0.05. More than 90% of the pixels in the headwaters exhibited negative annual ET trends, and approximately 45% showed significant changes at a significance level of 0.1. The annual ET exhibited a statistically significant decreasing trend at a significance level of 0.05 ($|Z| > 1.96$). Specifically, the upper mainstream section was dominated by a significant decreasing trend at a significance level of 0.01 ($|Z| > 2.57$).

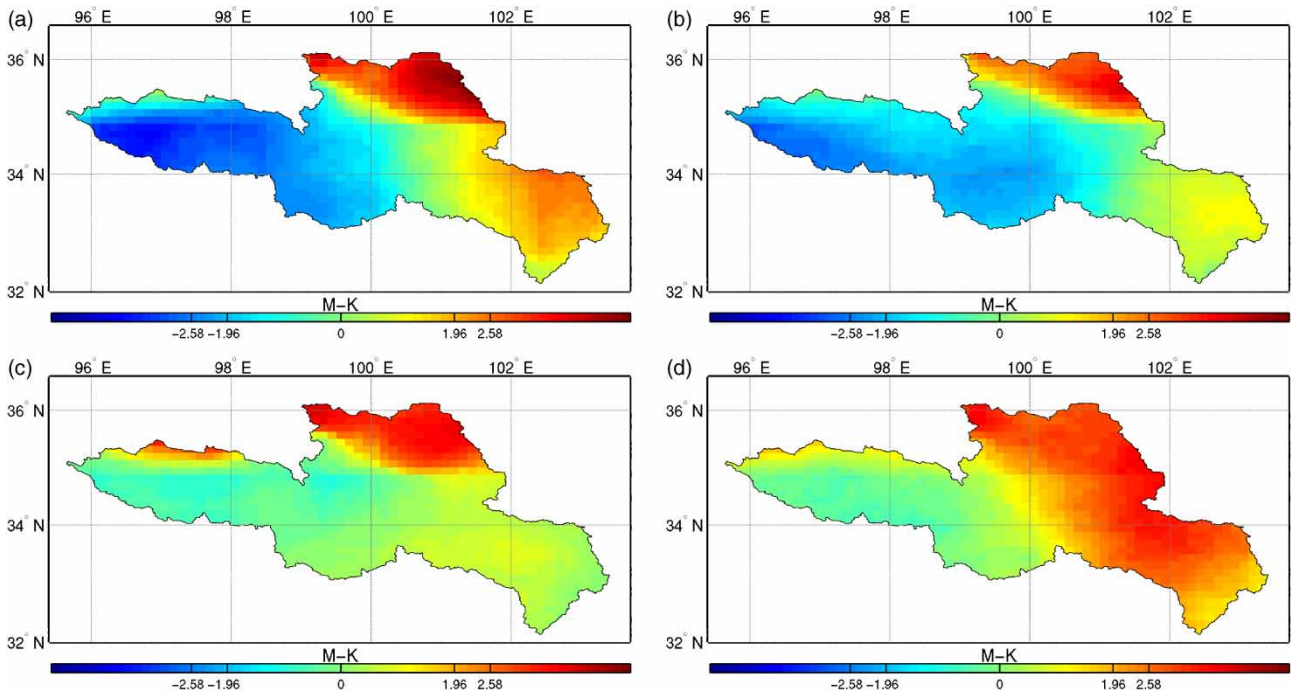

Figure 4 | MK analysis of ET from 1979 to 2014 in the Yellow River region catchment: (a) annual, (b) spring, (c) summer, and (d) autumn.

The upper region of the Yellow River is located at higher elevations close to the Tibetan Plateau. Influenced by monsoons and snowmelt, the distribution of MK values was insensitive to climate change and showed a negative trend in most areas of the upper reaches of the river. Figure 4(a) also shows that there are more gauges with positive ET trends than those with negative trends in the northern part of the study area. The distribution of annual ET trends indicates that a significant increasing trend mainly occurred in the upper Yellow River source region catchment, which is dominated by high ET values. These findings indicate that this area is at risk for extreme ET events. Note that the southeastern region displayed a larger increasing trend in the autumn (Figure 4(d)) compared to those in other seasons (Figure 4(b) and 4(c)) based on the MK test. Thus, the ET trend was more concentrated in the autumn than those during the spring and summer in this area.

As Matthew (2010) demonstrated, soil moisture on the Tibetan Plateau showed an increasing trend from 1987 to 2008. Wetter soil can cause the ground surface to absorb more net radiation, increasing the latent heat flux. The increase in net radiation and soil moisture may also explain the increasing trend in the latent heat on the Tibetan Plateau. The Yellow River source region catchment is located to the northeast of the Tibetan Plateau, and our research has shown that the upper reaches of the Yellow River source region are experiencing accelerated decreases in ET, which is the opposite of Matthew's conclusion.

Abrupt change in annual and growing season ET

To examine how the trends have progressed over time, the moving t-test and the SQMK test were applied to the original data and to the time series of the different periodic components obtained from the discrete wavelet decomposition. It is important to examine the moving t-test and SQMK test values because a mix of positive and negative trends may be present in the same time series. The two tests can help determine how the trend of a specific component may explain the trends found in the original data. In this study, the behaviors of the trend lines of the components are important. Therefore, this method of determining the most influential periodic component(s) considers not only the MK values of these components but also how similar the fluctuations of their trend lines are compared to those of the original data.

Figure 5 shows the results of the moving t-test for ET from 1979 to 2014 at Maqu station, which is where the instruments were located. This station mainly represents the runoff catchment. When a 5-year subsequence was applied, three abrupt changes in annual ET appeared in 1985, 1988, and 1989 (Figure 5(a)); one appeared in spring ET in 1985 (Figure 5(b)); and one appeared in autumn ET in 1990 (Figure 5(d)). For the spring values, abrupt changes in 1988 and 1989 were not indicated by the moving t-test. Similarly, for the autumn values, abrupt changes were not shown in 1985, 1988, or 1989, which were the 3 years indicated in the annual values, although there was an abrupt change in 1990. For the summer values, no abrupt changes were shown from 1979 to 2014. According to the location of the minimum absolute t statistic, we found that abrupt changes occurred in 1985 for both the annual (Figure 5(a)) and spring (Figure 5(b)) ET values. In terms of ET, the hydrological regimes are quite different before and after the detected abrupt changes. The entire Maqu area shows a general increase in ET after 1979. Then, ET starts to decline until 1990. The magnitude of the abrupt change in annual ET is -126.15 mm.

The SQMK test was used to detect the trend and abrupt changes in the annual, growing season, spring, summer, and autumn ET data series from 1979 to 2014 in the Yellow River source region. In Figure 6, horizontal dashed lines correspond to the confidence limits at a significance level of 95% (± 1.96). Note that there is a statistically significant trend (increasing or decreasing trend) when the function passes over the dashed lines. Notably, the SQMK values of the spring (Figure 6(b)), autumn (Figure 6(d)), and annual (Figure 6(a)) data were in very good agreement with each other, which might suggest that annual ET is mainly impacted by spring and autumn ET in the Maqu area. In addition, a strong increasing trend was observed from 1988 to 1990 in both the annual and autumn ET data series. In the long term, the trend was significant because the SQMK values in both the annual and autumn series exceeded the significance level threshold ($U0.005 = 1.96$), and they even surpassed the significance level of $U0.001 = 2.56$ in the short term. Before 1984, a decreasing trend in ET was observed at the Maqu station in the annual,

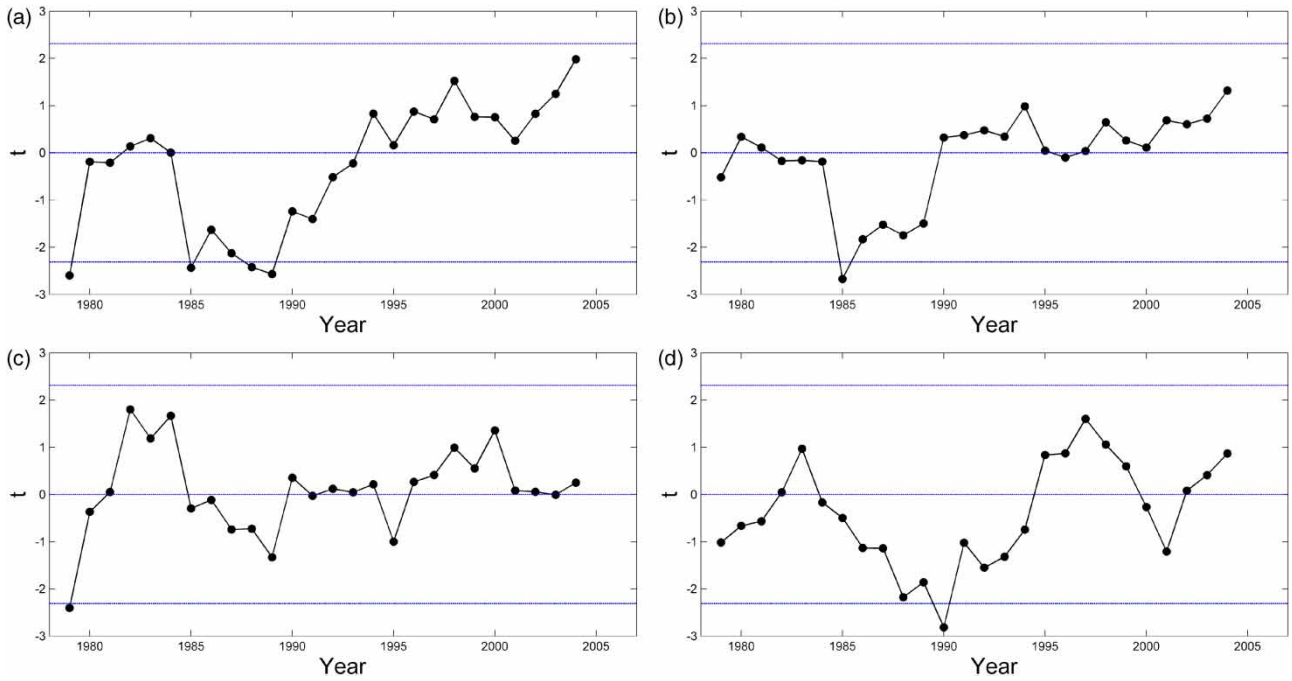

Figure 5 | Abrupt ET changes using the moving t-test from 1979 to 2014 in the Yellow River region catchment: (a) annual, (b) spring, (c) summer, (d) autumn. The horizontal blue lines indicate a significance level of 0.05 when $n = 5$. Please refer to the online version of this paper to see this figure in colour: http://dx.doi.org/10.2166/wcc.2017.134.

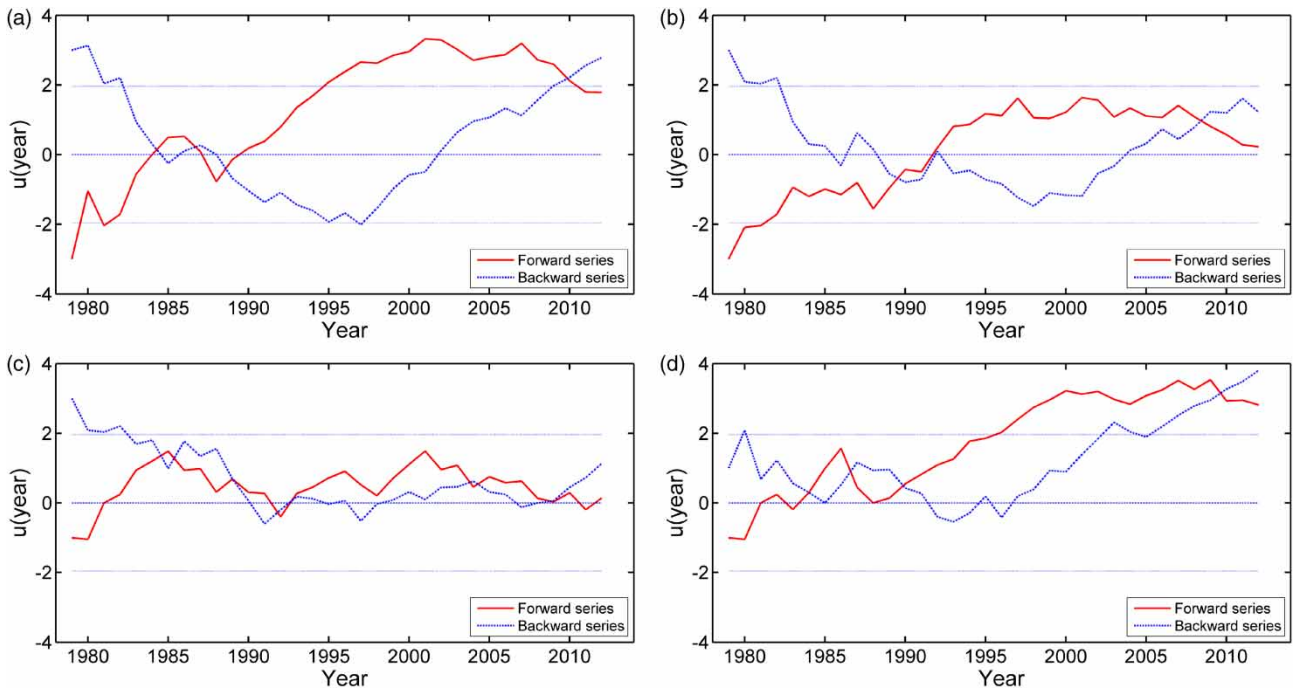

Figure 6 | SQMK analysis of ET from 1979 to 2014 in the Yellow River region catchment: (a) annual, (b) spring, (c) summer, (d) autumn. Solid line: forward series, dashed line: backward series.

spring, and autumn SQMK values, which were also beyond the significance threshold over short periods.

Abrupt changes found using the SQMK method reflect significant changes in the trends of the analyzed series, while those detected by the moving t-test indicate a significant difference between the averages of subsequence series. Consequently, the results of the two methods are not necessarily consistent with each other, as reported by Liang *et al.* (2010). Nonetheless, the abrupt change in 1984 in the increasing ET series was detected by both methods, suggesting the accuracy of this detected abrupt change. Two abrupt changes in the annual ET series in 1989 and 1988, which were detected only by the SQMK method, may also exist. It can be concluded that there were two abrupt changes in annual ET series in 1985 and 1989, showing decreasing and increasing changes, respectively. To some extent, the SQMK method supplements and verifies the moving t-test results.

Variation patterns of climate variables

Little research has been conducted on the contributions of changes in meteorological variables to changes in ET in different global regions. ET variation is influenced by both radiometric and aerodynamic variables (McVicar *et al.* 2007), which reflect the combined effects of topography, soil, vegetation, and climate. Explanations for causes of temporal trends in ET lie mainly in declining global solar irradiance, the complementary relationship between actual evaporation and potential evaporation, and decreasing wind speed (Roderick & Farquhar 2002; Hobbins *et al.* 2004; McVicar *et al.* 2008). In our study, a sensitive coefficient analysis was performed based on the impacts of climate change on ET. To detect the causes of ET characteristics and quantify the contribution of climatic variation, the influences of precipitation (mm), runoff (mm), soil moisture (mm), albedo, downward surface solar radiation (W/m^2), sunshine duration (s), land surface temperature (K), air temperature (K), and wind speed (m/s) are discussed. All variables were measured by ERA-Interim at the Maqu station.

Simple correlation does not prove to be an all-encompassing technique, especially under the above circumstances. To obtain an accurate picture of the relationships between ET and certain variables, we should first eliminate the influences of other variables. Partial correlation analysis involves studying the linear relationship between two variables after excluding the effect of one or more independent factors. Thus, the absolute values of partial correlation coefficients between ET and nine climate variables are also examined.

Table 2 shows the results of partial correlation analysis between ET and other variables. The variation in ET was dominated by precipitation, which accounted for 0.78, followed by soil moisture and runoff, which accounted for 0.38 and 0.14, respectively. Precipitation is the most important input component of both hydrological processes and energy cycles in the study area. Because soil moisture is a physical characteristic at the interface of the atmosphere and the earth, soil moisture influences many land surface and hydrological processes, affecting ET, partitioning of rainfall into infiltration and runoff, and precipitation, although the existence and strength of this influence has been debated. In the Yellow River source region, the soil moisture did not readily impact the ET, with a partial correlation of −0.38. Given the scarcity of long-term observations of runoff and ET in China, the spatiotemporal variation in the hydrological flux and its underlying mechanisms require further investigation. Several studies have discussed the effects of climate change on the water cycle in China; however, they have rarely considered ET in these estimations. According to our research, runoff and ET do not have a strong relationship, with a partial correlation of 0.14.

A sensitivity analysis estimating the impact of albedo uncertainties on climate modeling showed that an absolute

Table 2 | Partial correlation analysis between ET and other variables

Variables	Partial correlation
Precipitation	0.78
Runoff	0.14
Soil moisture	−0.38
Albedo	−0.77
Surface solar radiation downwards	0.54
Sunshine duration	−0.35
Land surface temperature	−0.04
Air temperature	0.44
Wind speed	0.16

albedo accuracy of ±0.02 to ±0.03, which is equivalent to a net radiation uncertainty of ±10 W/m^2, results in significant changes in regional climate simulations (Sellers *et al.* 1996). In fact, the surface albedo accuracy directly affects the precision of the net radiation and indirectly affects daily ET and other land surface fluxes. Nevertheless, no previous research has analyzed the impact of albedo accuracy on surface energy balance parameters. In the case of ET models, when no *in situ* data are available, the surface albedo can be estimated in different ways. From our study (Table 2), the annual anomalies of albedo display maximum and minimum levels in 2011 and 1998, with values of 0.52 and −0.30, respectively. The partial correlation between ET and albedo is −0.77, and the partial correlation between ET and downward surface solar radiation is 0.54. The lowest value of the anomalies in both downward surface solar radiation and ET occurred in 1989. For the monthly data, the downward surface solar radiation changed from 11,421,470 W/m^2 in December 1995 to 26,240,590 W/m^2 in April 1982, and the ET reflected this change, with values of 21.37 and 10.10 mm in the same respective months. The partial correlation between ET and sunshine duration is −0.35 from 1979 to 2014. The maximum and minimum monthly mean values of sunshine duration were 38,745 s in June 1990 and 28,440 s in September 2011, and the corresponding ET values were 97.10 and 50.35 mm, respectively. The anomalies in both ET and sunshine duration reached minimum values in 1989 and maximum values in 2002.

The land surface temperature is considered more influential in determining the energy fluxes than are the other input parameters, such as vegetation parameters or surface albedo, that are derived from remote sensing data and used to estimate ET because the land surface temperature provides an indication of the net effect of land–atmosphere interactions, which result in ET. However, according to our research (Table 2), the land surface temperature does not closely correspond to changes in ET, with a partial correlation of −0.04. The maximum and minimum values of the anomalies in the land surface temperature occurred in 1998 and 1980, with values of 14.23 and −14.64 K, respectively. During these 2 years, the values of the anomalies in ET were 34.5 and −54.2 mm. The air temperature exhibited larger variations as ET changed, with a partial correlation of 0.44 from 1979 to 2014. Additionally, wind speed is an important input variable that has been used to calculate ET worldwide. Tabari *et al.* (2011) analyzed the spatiotemporal variations of ET in arid and semi-arid regions of Iran and revealed that the main factor associated with decreasing ET was decreasing wind speed. However, based on our research, there is no agreement between these parameters in the Yellow River source catchment. The partial correlation between ET and wind speed is only 0.16, suggesting that the wind speed did not have a significant impact on ET.

CONCLUSIONS

With the goal of generating optimal land surface ET, reanalysis data, the satellite-based MODIS global ET product, and the SEBS model have generated quality-controlled data. Using correlation analysis, the ERA-Interim data were shown to be superior to the observations from the EC system in the Yellow River source region. Temporally and spatially consistent terrestrial data, including ET and other variables that affect evaporation and transpiration, were also analyzed from the ERA-Interim data. The annual mean ET from 1979 to 2014 in the Yellow River source region was 551.3 mm, and the maximum and minimum values were 595.9 mm in 2002 and 485.9 mm in 1989. The results show that the Yellow River source region catchment experienced a statistical decrease in ET at a rate of −0.50 mm/yr in the upstream area and an increase of 1.65 mm/yr in the runoff and outlet areas from 1979 to 2014.

At the Maqu station where our instruments were located, both the moving t-test and the SQMK test were able to detect an abrupt change in ET in 1985. The absolute values of the partial correlation coefficients between ET and nine climate variables indicate that precipitation played the most important role in the changes in ET in the Yellow River source region, followed by albedo, solar radiation, and downward surface solar radiation, while land surface temperature was the least sensitive factor among the climate variables.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant Nos 41405079, 41405017,

41530529, 41375022 and 41675015), the Key Research Program of the Chinese Academy of Sciences (Grant KZZD-EW-13). We thank the anonymous reviewers for their constructive comments.

REFERENCES

Chen, X., Su, Z., Ma, Y., Liu, S., Yu, Q. & Xu, Z. 2014 Development of a 10 year (2001–2010) 0.1 dataset of land-surface energy balance for mainland China. *Atmos. Chem. Phys.* **14** (23), 14471–14518.

Cleugh, H. A., Leuning, R., Mu, Q. & Running, S. W. 2007 Regional evaporation estimates from flux tower and MODIS satellite data. *Remote Sens. Environ.* **106** (3), 285–304.

Foken, T., Mauder, M., Liebethal, C., Wimmer, F., Beyrich, F., Leps, J. P., Raasch, S., De Bruin, H. A. R., Meijninger, W. M. L. & Bange, J. 2010 Energy balance closure for the LITFASS-2003 experiment. *Theor. Appl. Climatol.* **101**, 149–160.

Fraedrich, K., Jiang, J., Gerstengarbe, W. & Werner, P. C. 1997 Multiscale detection of abrupt climate changes: application to Nile River flood levels. *Int. J. Climatol.* **17**, 1301–1315.

Gocic, M. & Trajkovic, S. 2013 Analysis of changes in meteorological variables using Mann–Kendall and Sen's slope estimator statistical tests in Serbia. *Glob. Planet. Change* **100**, 172–182.

Hamed, K. H. 2008 Trend detection in hydrologic data: the Mann–Kendall trend test under the scaling hypothesis. *J. Hydrol.* **349**, 350–363.

Hobbins, M. T., Ramírez, J. A. & Brown, T. C. 2004 Trends in pan evaporation and actual evapotranspiration across the conterminous U.S: paradoxical or complementary? *Geophys. Res. Lett.* **31**, L13503.

Jiménez, C., Prigent, C., Mueller, B., Seneviratne, S. I., McCabe, M. F., Wood, E. F., Rossow, W. B., Balsamo, G., Betts, A. K., Dirmeyer, P. A., Fisher, J. B., Jung, M., Kanamitsu, M., Reichle, R. H., Reichstein, M., Rodell, M., Sheffield, J., Tu, K. & Wang, K. 2011 Global intercomparison of 12 land surface heat flux estimates. *J. Geophys. Res.* **116**, D02102.

Jones, P. D. & Lister, D. H. 2014 Antarctic near-surface air temperatures compared with ERA-Interim values since 1979. *Int. J. Climatol.* **35**, 1354–1366.

Kim, H. W., Hwang, K., Mu, Q. Z., Lee, S. O. & Choi, M. H. 2012 Validation of MODIS 16 global terrestrial evapotranspiration products in various climates and land cover types in Asia. *J. Civil Eng.* **16** (2), 229–238.

Liang, L. Q., Li, L. J. & Liu, Q. 2010 Temporal variation of reference evapotranspiration during 1961–2005 in the Taoer River basin of Northeast China. *Agric. Forest Meteorol.* **150**, 298–306.

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

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

Liu, R., Wen, J., Wang, X. & Zhang, Y. 2015 Evapotranspiration estimated by using datasets from the Chinese FengYun-2D geostationary meteorological satellite over the Yellow River source area. *Adv. Space Res.* **55**, 60–71.

Matthew, O. 2010 Characterization of the Effects of Climate Variation on Land Surface Temperature and Soil Moisture Through Stochastic Analysis of Long Term SSM/I Observations Over the Tibetan Plateau. Master's thesis, International Institute for Geo-information Science and Earth Observation, University of Twente, Enschede, The Netherlands, pp. 1–67.

McVicar, T. R., Niel, T. G. V., Li, L. T., Hutchinson, M. F., Mu, X. M. & Liu, Z. H. 2007 Spatially distributing monthly reference evapotranspiration and pan evaporation considering topographic influences. *J. Hydrol.* **338**, 196–220.

McVicar, T. R., van Niel, T. G., Li, L. T., Roderick, M. L., Rayner, D. P., Ricciardulli, L. & Donohue, R. J. 2008 Wind speed climatology and trends for Australia, 1975–2006: capturing the stilling phenomenon and comparison with near-surface reanalysis output. *Geophys. Res. Lett.* **35**, L20403.

Mendoza, V. M., Villanueva, E. E., Garduño, R., Nava, Y., Santisteban, G., Mendoza, A. S., Oda, B. & Adem, J. 2008 Thermo-hydrological modelling of the climate change effect on water availability in two hydrologic regions of Mexico. *Int. J. Climatol.* **29** (8), 1131–1153.

Mo, X. G., Liu, S. X., Meng, D. J. & Lin, Z. H. 2014 Exploring the interannual and spatial variations of ET and GPP with climate by a physical model and remote sensing data in a large basin of Northeast China. *Int. J. Climatol.* **34** (6), 1945–1963.

Mojid, M. A., Rannu, R. P. & Karim, N. N. 2015 Climate change impacts on reference crop evapotranspiration in North-West hydrological region of Bangladesh. *Int. J. Climatol.* **35**, 4041–4046.

Mooney, P. A., Mulligan, F. J. & Fealy, R. 2011 Comparison of ERA-40, ERA-Interim and NCEP/NCAR reanalysis data with observed surface air temperatures over Ireland. *Int. J. Climatol.* **31** (4), 545–557.

Mu, Q., Heinsch, F., Zhao, M. & Running, S. 2007 Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sens. Environ.* **111** (4), 519–536.

Mu, Q., Zhao, M. & Running, S. W. 2011 Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sens. Environ.* **115** (8), 1781–1800.

Nalley, D., Adamowski, J. & Khalil, B. 2012 Using discrete wavelet transforms to analyze trends in streamflow and precipitation in Quebec and Ontario (1954–2008). *J. Hydrol.* **475**, 204–228.

Roderick, M. L. & Farquhar, G. D. 2002 The cause of decreased pan evaporation over the past 50 years. *Science* **298** (15), 1410–1411.

Samba, G. & Nganga, D. 2012 Rainfall variability in Congo-Brazzaville: 1932–2007. *Int. J. Climatol.* **32** (6), 854–873.

Sellers, P. J., Los, S. O., Tucker, C. J., Justice, C. O., Dazlich, D. A. & Collatz, G. J. 1996 A revised land surface parameterization (sib2) for atmospheric GCMs. 2. The generation of global fields of terrestrial biophysical parameters from satellite data. *J. Climate* **9**, 706–737.

Sen, P. K. 1968 Estimates of the regression coefficient based on Kendall's tau. *J. Am. Stat. Assoc.* **63**, 1379–1389.

Su, Z. 2002 The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. *Hydrol. Earth Syst. Sci.* **6** (1), 85–99.

Su, T., Feng, G. L., Zhou, J. & Ye, M. 2014 The response of actual evaporation to global warming in China based on six reanalysis datasets. *Int. J. Climatol.* **35**, 3238–3248.

Tabari, H., Marofi, S., Aeini, A., Hosseinzadeh Talaee, P. & Mohammadi, K. 2011 Trend analysis of reference evapotranspiration in the western half of Iran. *Agric. Forest Meteorol.* **151** (2), 128–136.

Wever, L. A., Flanagan, L. B. & Carlson, P. J. 2002 Seasonal and interannual variation in evapotranspiration, energy balance and surface conductance in a northern temperate grassland. *Agric. Forest Meteorol.* **112**, 31–49.

You, Q. L., Fraedrich, K., Ren, G. Y., Pepin, N. & Kang, S. C. 2013 Variability of temperature in the Tibetan Plateau based on homogenized surface stations and reanalysis data. *Int. J. Climatol.* **33** (6), 1337–1347.

Yu, P. S., Yang, T. C. & Wu, C. K. 2002 Impact of climate change on water resources in southern Taiwan. *J. Hydrol.* **260**, 161–175.

Zang, C. F. & Liu, J. G. 2013 Trend analysis for the flows of green and blue water in the Heihe River basin, northwestern China. *J. Hydrol.* **502**, 27–36.

First received 27 September 2015; accepted in revised form 3 March 2017. Available online 19 April 2017