

Dear Dr. Roger Moussa,

We would like to express our sincere thanks to you for the organization of the review on our revised manuscript. Now, the manuscript has been improved according to the comments from the two reviewers.

Below we provide the detailed replies to the comments. The original comments are in bold face and our replies are in regular font.

Thank you for your consideration of this paper. If you have any further requires, please let us know as soon as possible!

Best regards,

Xu-Sheng Wang

Yangxiao Zhou

Reply to Anonymous Referee #3:

1. It seems like there was a clear misconception between the authors and the reviewers, as the reviewers argued the authors present a (not novel) way to incorporate storage effects into the Budyko framework. The authors, however, wanted to point out in particular the effect of groundwater-dependent evapotranspiration. I still don't think that this kind of misconception is resolved within the present version of the manuscript. Please add a better and more concise motivation to your introduction (and abstract), that clearly points out the novelty of your approach and how this goes beyond previous studies (Istanbulluoglu, etc.). This could be done for example somewhere on page 4.

Reply: Good suggestion. We revised the introduction part on the page 4 accordingly.

2. P. 3, L. 11ff: For the discussion on interannual variability please check also: "Koster, R.D., and Suarez, M.J. (1999). A Simple Framework for Examining the Interannual Variability of Land Surface Moisture Fluxes. *Journal of Climate* 12, 1911 – 1917." and "Gentine, P., D' Odorico, P., Lintner, B.R., Sivandran, G., and Salvucci, G. (2012). Interdependence of climate, soil, and vegetation as constrained by the Budyko curve. *Geophysical Research Letters* 39".

Reply: Accept!

3. P. 3, L. 29: What do you mean by “extremely arid situations”? I think the term “arid” is more used to describe climatic conditions and not events such as droughts. Better use simply “dry” .

Reply: Accept! We replaced the words ‘extremely arid’ and ‘extremely dry’ with ‘dry’ in the text.

4. P. 4, L. 22ff: In the overview of what is done in the study it will help a lot if you refer to the respective sections.

Reply: Accept!

5. P.6, L. 4: From Figure 1 it seems like all six catchments are located in the part of the

Erdos Plateau with highest mean annual rainfall. Does this play a role for the overall interpretation of the results.

Reply: In the west part of the Erdos Plateau there were also some river basins but they are lack of hydrological data. Another risk in investigating the catchments in the west part is the complexity in the groundwater system. This was discussed in the Section 5.4 in the revised manuscript.

6. P.9, L. 10ff: You might point out again (in one or two sentences) what you will do different than Istanbuluoglu.

Reply: Accept! In a different way from Istanbuluoglu et al. (2012), we estimated ΔG from the monthly baseflow (groundwater discharge) data with a calibrated hydrological model, in which the groundwater dependent evapotranspiration is also incorporated.

7. P. 15, L30: You can leave out “not very high” , I guess. Sounds a bit sloppy.

Reply: Accept! The words were deleted.

8. P. 21, L27-30: Please also check "Greve, P., Gudmundsson, L., Orlowsky, B., and Seneviratne, S.I. (2016). A two-parameter Budyko function to represent conditions under which evapotranspiration exceeds precipitation. Hydrol. Earth Syst. Sci. 20, 2195 – 2205." in this context.

Reply: Accept! We referenced Greve et al. (2016) in the first paragraph of the Section 5.2.

9. P. 23, L. 26ff: I pretty much like the limitation remark. However, I think some (more general) shortcomings of your approach should be also mentioned here. These are especially related to the data you use. I especially miss a discussion on potential biases stemming from E0-estimates (since there are so many different ways to handle and estimate E0). Also biases in P could alter your results.

Reply: Good suggestion. Limitations in the data processing and the complexities in the hydrogeological conditions are discussed in the Section 5.4.

10. Figure 3: Please note in the caption that the x-axes are different between (a) and the other subfigures. And what does the grey form indicate in subplot (a)? Is it necessary?.

Reply: Accept! The ticks of the x-axes in Figure 3(a) have been extended to the same range as in the other sub-figures. The grey polygon in Figure 3(a) was removed.

11. Figure 5: The labeling of the y-axes actually does not correspond to the explanation in the caption.

Reply: Accept! The caption of the figure and title of the axes in Figure 5 have been revised to avoid this misleading.

12. Please revise the English usage and grammar throughout the manuscript. If possible by a native speaker.

Reply: We apologize for the language problems in the previous manuscript. The revised text has been carefully checked and significantly improved in English.

Reply to Anonymous Referee #5:

1. P 7. and fig 2b: Labels on the y-axis indicating time frame (monthly?) of these measurements – particularly runoff should be included.

Reply: Accept! “Monthly” is included in the titles of the y-axes in Figure 2b.

2. 8.4: Fig 2b seems to indicate that streamflow is pretty constant from year to year!

Reply: This is a special phenomenon after 1969. The streamflow was strongly influenced by dams in the catchment C1 after 1969, cutting off the floods in summer. Remain was the seasonal patterns dominated by groundwater discharge, which show smaller variation from year to year. This is an important reason for the discussions in the Section 5.3 on the role of landscape-driven and human-controlled streamflow shift.

3. 9.10: Personal opinion: “abnormal trend” is probably not the correct phrase for this. It

is only “abnormal” because annual water balance in the catchments doesn’t suit Budyko’s original hypotheses (which are that the points are based on long-term average with no storage changes, also Budyko’s original figure was variability between catchments rather than year-to-year variability within a catchment). Also “trend” implies (to me) a progression through time, whereas the figures have dryness index on the x-axis

Reply: The word ‘trend’ is revised to ‘relation’.

4. 11.26-28: It would be good to have an example catchment (C1?) with the groundwater contour lines plotted to help in understanding and visualising the zone-1/zone-2 boundary in your catchments!

Reply: Accept! In Figure 1(c) the groundwater level contours were shown.

5. 12.9: So E_{1m} depends on W only and not V ?

Reply: Yes. In the ABCD-GE model, the evapotranspiration in Zone-1 is estimated by W as the same in the ABCD model. V is the storage in the transition zone that could not be accessed by the evapotranspiration process.

6. Fig 4b: I may be wrong but the downward flow into the intermediate vadose zone in the diagram should be cR not R .

Reply: Thanks for reminding the mistake. It should be cR and the Figure 4b has been revised.

7. Fig 4b: Is there infiltration into zone-2 areas in the diagram? A labelled (I assume cP) flux arrow would help here.

Reply: Accept! The ‘ P ’ on the land surface of the zone-2 is modified to ‘ cP ’.

8. Fig 4b again: The way it is drawn looks like zone 2 represents evaporation from deep groundwater, whereas from your description, zone 2 is where the water table is very close to the surface. In this case, your diagram would be better drawn with a sloping ground surface with zone 2 at the valley bottom, i.e. something similar to

http://www.cgenarchive.org/uploads/2/5/2/6/25269392/2902862_orig.gif. A more accurate diagram would help with conceptual understanding of the abcd-ge model.

Reply: Accept! The land surface in Zone-2 was drawn slightly incline to the river.

9. 15.23-24: How can all the catchments have the same RMSE?

Reply: This RMSE refers to the aggregated result of the modeling monthly runoff (combine the data in to a single array for calculation). It was pointed out more clearly in the revised manuscript.

10. 16.23: It' s worth mentioning here that changes in climate characteristics (e.g. seasonal rainfall or extreme rainfall events) over the second time period could also lead to Qnat being different from Qobs (i.e. it' s not necessarily all due to irrigation or groundwater).

Reply: Accept!

11. Figure 6: It would be better to have the same x-axis tick mark spacing on the four panels. For panel d, the legend should read “Actual ET” not “Actrual ET” . Also, (and this applies to figure 3 as well) it would be better when presenting data for a particular catchment to have the catchment number (C1-C6) on the panel rather than a), b) etc. – similar to figure 7 and 10.

Reply: Good suggestions. Figure 6 has been modified. But in Figure 6 the labels (a)-(d) are remained to number the sub-figures. All of these sub-figures are just associated to C1 as indicated in the caption.

12. Figure 7: It would be good to have “F” on the figure itself, as you refer to “F” when talking about the figure but it' s not there (of course E/P is though...)?

Reply: Accept! The ‘E/P’ has been replaced by ‘F’ in all of the figures and in the text to avoid using different symbols for the same variables.

13. 17.16: Should be C4 and C6 here – also there still related to aridity index, but less so (“less sensitive” ?)

Reply: Accept! There was a mistake here. The words “C3 and C6” now are replaced

by “C4 and C6”. Also “trend” is replaced by “regression line”.

14. I am confused by Figure 8. Firstly (and this occurs throughout the paper), the HRC is defined on p5.14 as the Hailutu River catchment, which is also denoted C1. Is this correct – i.e. is C1 identical to the HRC? If so, it is confusing to interchange these two throughout the manuscript. Secondly, none of the catchments C1-C6 have E/P greater than about 2, whereas Figure 8a (HRC?) has E2/P close to 4 for very dry years. How can I reconcile the two figures?

Reply: Good suggestions. The term ‘HRC’ has been deleted from the text and figures, just use ‘C1’ because they denote the same catchment. For the second problem, it should be mentioned that E/P at the catchment scale is an weighted average value of E1/P and E2/P so that it is always less than E2/P. In the Zone-2, the evapotranspiration keeps in high levels, leading to very high values of E2/P in dry years. This has been explained in Section 4.3 in detail.

15. Figure 9: spelling mistake in the legend: “Cachments data”

Reply: We are sorry for the mistake. It has been revised in Figure 9.

16. 20.17-18: What about groundwater from originating in other catchments (lateral flow) or inter-basin transfers?

Reply: In the revised manuscript, we discussed this issue in the second paragraph of the Section 5.4. In this study, it is assumed that the boundary of a catchment determined on land surface is also the boundary of groundwater flow, *i. e.*, no inter-basin transfer of groundwater exists. The assumption is plausible acceptable in the east part of the Erdos Plateau because spatial variation in groundwater level is highly correlated with the fluctuation of ground surface in this area (Lv et al. 2013; Zhou et al., 2015). However, the groundwater flow system in the west part of the Erdos Plateau is more complex where the climate is more arid and water table divides could be different from topographic divides. How to handle the role of the inter-basin groundwater flow for the annual water balance in a catchment is not the objective of this study but of worth for further investigations.

17. 21.17-19: And also that your modeled change in storage may not be representative of actual change in storage for many different reasons (misspecified model, data uncertainty, problems with parameter calibration and so on...)

Reply: We discussed the limitations of the model and of the case study in Section 5.4.

18. I find section 5.3 quite difficult to follow – and the aim of this section is not really clear to me. It is not obvious what “Landscape-driven change” is, and why it’s important. If this section is needed, a better description of what you’re trying to achieve here would be good.

Reply: As shown in the data of the catchment C1 and in the previous studies of this catchment (Zhou et al., 2013; Yang et al., 2012), the streamflow in the catchment was significantly influenced by change in land-use and hydraulic engineering. One of the reviewers of this manuscript, Dr. Jaramillo, proposed to identify the different roles of the landscape-driven and human-controlled shifts of annual water balance. Section 5.3 was presented for this purpose. In reply on the second comments, it was noticed that after 1969 the streamflow show significantly smaller variation from year to year, which is caused by the two impacts. This is an important reason for the discussions in Section 5.3.

19. 23.1: CVq looks like coefficient of variability rather than variability (variance). Is it?

Reply: Accept! The term is revised to "The coefficient of intra-annual variation". These words in the caption of Figure 11 were also revised.

20. General comments about the figures: There are a few spelling mistakes in the figures as noted above. Spending some time to get the figures more consistent would make the paper much better, e.g. you have “aridity index” , “phi” and “E0/P” in different figures and these are all the same thing. Similarly the “F” and “E/P” issue discussed above. (See also comments above about the panel labelling)

Reply: Good suggestion. The figures are revised to just use “aridity index” and ‘F’.

Shift of annual water balance in the Budyko space for catchments with groundwater dependent evapotranspiration

X.-S. Wang¹ and Y. Zhou²

[1]{Ministry of Education Key Laboratory of Groundwater Circulation and Evolution, China University of Geosciences, Beijing 100083, China }

[2]{UNESCO-IHE Institute for Water Education, Delft, The Netherlands }

Correspondence to: X.-S. Wang (wxsh@cugb.edu.cn)

Abstract

The Budyko framework represents the general relationship between the evapotranspiration ratio (F) and the aridity index (ϕ) for the mean annual steady state water balance at the catchment scale. It is interesting to investigate if this standard F - ϕ space can be also applied to capture the shift of annual water balance in catchments with the varying dryness. Previous studies have made significant progress in incorporating the storage effect in the Budyko framework for the non-steady conditions whereas the role of groundwater dependent evapotranspiration was not investigated. This study investigates how groundwater dependent evapotranspiration causes the shift of the annual water balance in the standard Budyko space. A widely used monthly hydrological model, the ABCD model, is modified to incorporate groundwater dependent evapotranspiration in the zone with shallow water table and delayed groundwater recharge in the zone with deep water table. This model is applied in six catchments in the Erdos Plateau, China, to estimate the actual annual evapotranspiration. Results show that the variations in the annual F value with the aridity index do not satisfy the standard Budyko formulas. The shift of the annual water balance in the standard Budyko space is a combination of the Budyko-type response in the deep groundwater zone and the quasi-energy limited condition in the shallow groundwater zone. Excess evapotranspiration ($F > 1$) could occur in dry years, which is contributed by the significant supply of groundwater

for evapotranspiration. Use of groundwater for irrigation can increase the frequency of the $F > 1$ cases.

1 Introduction

Estimating catchment water balance is one of the fundamental tasks in hydrology. Efforts have long been devoted to construct the physical, empirical, and statistical models to explain the general relationship among precipitation (P), runoff (Q), potential evapotranspiration (E_0) and actual evapotranspiration (E) in terms of mean annual fluxes at the catchment scale (Budyko, 1948, 1958, 1974; Mezentsev, 1955; Fu, 1981; Porporato et al., 2004; Gerrits et al., 2009). A simple and highly intuitive approach widely used for estimating E at the mean annual steady state is the Budyko framework, in which the mean annual evapotranspiration ratio (E/P) was presumed as a function of the climatic dryness:

$$\frac{E}{P} = F\left(\frac{E_0}{P}\right) = F(\phi) \quad (1)$$

where ϕ is the aridity index defined as E_0/P , and $F(\phi)$ is an empirical function that relates E/P to ϕ based on general water-energy balance. The proposed formula by Budyko (1958; 1974) was:

$$F(\phi) = \sqrt{\phi[1 - \exp(-\phi)] \tanh(1/\phi)} \quad (2)$$

which indicates a nonlinear relation between F and ϕ . This F - ϕ curve has been called the Budyko curve (Zhang et al., 2004; Roderick and Farquhar, 2011) and the F - ϕ space was called Budyko space (Renner et al., 2012).

Instead of using a single curve determined by Eq. (2) in the Budyko space, researchers have introduced a specific catchment parameter in $F(\phi)$ to consider the impacts of catchment properties such as soils and vegetation (Mezentsev, 1955; Fu, 1981; Zhang et al., 2001). For example, Fu's equation (Fu, 1981) was derived following the idea of Mezentsev (1955), which can be expressed as follows:

$$F(\phi, w) = 1 + \phi - (1 + \phi^w)^{1/w} \quad (3)$$

where w is a parameter represents the catchment conditions. F increases with w , leading to reduced Q/P as w grows (Fu, 1981). Fu's equation has been widely used in the last decade (Zhang et al., 2004; Yang et al., 2006; Yang et al., 2007; Zhang et al., 2008; Greve et al., 2015). Donohue (2007) highlighted the role of vegetation dynamics in the application of the Budyko framework. Wang and Tang (2014) also developed a one-parameter Budyko model based on the proportionality hypothesis and revealed a complex relationship between the catchment specific parameter and the remote sensing vegetation index. These modified formulas suggested a group of Budyko curves instead of the single original Budyko curve, in which a curve represents a specific type of the catchments with similar features controlling the mean annual water balance. Nevertheless, Gentine et al. (2012) argued that the original Budyko curve reflects the interdependence among vegetation, soil and climate and could be applied as a strong constraint on land-surface parameterizations.

Budyko hypothesis has been directly used to analyze the interannual change in water balance in catchments (Koster and Suarez, 1999; Arora, 2002; Zhang et al., 2008; Potter and Zhang, 2009) ignoring the change in storage (ΔS) under the assumption of steady state water balance. One can plot annually the estimated F data in the standard Budyko space to check whether the standard Budyko curves are sufficient or not to represent the interannual variability of evapotranspiration with the varying dryness. In this way, Potter and Zhang (2009) found that the Budyko framework is generally applicable for the catchments in Australia and the optimal Budyko curve of the annual F - ϕ data is highly dependent on the seasonal variations in rainfall. However, this approach should be carefully used when the annual F value is approximated by the annual $(P-Q)/P$ value. Wang et al. (2009) and Istanbuluoglu et al. (2012) reported that the annual data of $(P-Q)/P$ in some basins are negatively related to the aridity index, exhibiting an inverse relation in comparison with the standard Budyko curves. According to long-term groundwater observations in the North Loup River basin, Nebraska, USA, Istanbuluoglu et al. (2012) demonstrated that the annual F data estimated by $(P-Q-\Delta G)/P$ basically follow the Budyko hypothesis, where ΔG is the change in groundwater storage. However, in some other studies, an unexpected high evapotranspiration ratio ($F>1$) was observed (Cheng et al., 2011; Wang, 2012; Chen et al., 2013). Among the 12 watersheds investigated by Wang (2012), half of them had such high F values in two or more drought years. The physical base of the phenomena is the significant contribution of storage in dry periods by which the high level of evapotranspiration is maintained. Although some of the

cases were triggered by extracting groundwater for irrigation in farmlands (Cheng et al., 2011; Wang, 2012), it could occur in natural conditions as a result of the temporal redistribution of water from seasonal patterns (Chen et al., 2013). Wang (2012) and Chen et al. (2013) proposed an approach to extend the Budyko framework for annual or even intra-annual water balance by considering the soil water storage as a potential source of water supply for evapotranspiration. They defined $P-\Delta S$ for the selected time scales as the effective rainfall in building the modified Budyko space with $E/(P-\Delta S)$ and $E_0/(P-\Delta S)$, instead of E/P and ϕ , respectively. In summary, the previous studies made significant progress in incorporating storage effects into the Budyko framework, but the role of groundwater dependent evapotranspiration was not yet investigated.

The excess annual evapotranspiration may be originated from both soil water and groundwater. As reported by Wang (2012), during the drought year in 1988, two watersheds in Illinois, USA, showed $F=1.1$ with ~100 mm depletion in soil water and ~200 mm decrease in groundwater storage, respectively. It seemed that the contribution of groundwater was more significant (partially enhanced by groundwater pumping). Small depth to water table is an advantage to keep a high level of soil water content near ground surface for evapotranspiration (Chen and Hu, 2004). Therefore, it could be argued that the existence of shallow groundwater in a catchment would enhance the occurrence of the $F>1$ cases in drought years. Groundwater dependent evapotranspiration at the regional scale has been noticed in the previous studies (York et al, 2002; Chen and Hu, 2004; Cohen et al., 2006; Yeh and Famiglietti, 2009). Nevertheless, little has been known on the role of groundwater in the interannual variability of the evapotranspiration ratio with the varying dryness. Chen et al. (2013) did not identify the change in groundwater storage to explain the controls of the $F>1$ cases. Wang (2012) mentioned the potential role of groundwater in occurrence of the $F>1$ cases, but the individual contribution of groundwater dependent evapotranspiration was not soundly analyzed.

This study develops a method to analyze the effect of groundwater dependent evapotranspiration on the annual water balance of catchments in the standard Budyko space, in addition to the storage effect that proposed in previous studies. In Section 2, the location, characteristics and data of the typical six studied catchments in the Erdos Plateau, China, are presented. Preliminary analysis showed the abnormal F - ϕ relations in some of the catchments, which probably indicates the effect of groundwater dependent evapotranspiration. In Section

3, a monthly hydrological model is developed from the widely used ABCD model (Thomas, 1981) to incorporate the groundwater dependent evapotranspiration as well as the deep infiltration in the vadose zone. The monthly E is partitioned into two components in accounting for the individual roles of the normal soil water dependent and the specific groundwater dependent evapotranspiration. Then, the modified model is calibrated for the study catchments in Section 4 and used to produce the annual time series of the evapotranspiration components linking with the variable soil water and groundwater storages. With varying climatic dryness, the shifts of the interannual water balance in the standard and modified Budyko spaces are analyzed and discussed in Section 5. The impacts of human activities and the limitations of the approach are also discussed in Section 5.

2 Study Area, Data and Preliminary Analysis

2.1 River Basins

The study area is located in the Erdos Plateau in the north-central China (Fig. 1a), belongs to the middle part of the Yellow River Basin. The climate of the Erdos Plateau is typically inland semiarid to arid with a significant gradient of the mean annual precipitation, from 150 mm in the west to 450 mm in the east (Fig. 1b). More than half of the annual precipitation is received in the warm season (from June to September). Six catchments with available data, numbered as C1-C6 (Fig. 1b), are selected for this study. The areas of the catchments range between 1,272 km² and 3,253 km².

In particular, C1 is the Hailiutu River Catchment, with an area of 2,645 km², which lies on the southeast edge of the Mu Us Desert and is a sub-catchment of the Wuding River basin (Fig. 1b). The main channel in C1 has a length of approximately 85 km and flows southwards to the Hanjiamao hydrological station, as shown in Fig. 1c. Due to the arid climate and desert landscape, the land cover within the catchment is characterized by desert sand dunes with patches of mostly shrublands. Depression areas and terraces with shallow groundwater are covered by meadows and some farmlands. Wind-breaking trees (*Salix matsudana* and *Populus tomentosa*) can be found along the roads and crop areas. Farmlands are mainly located in the southern area and especially in the river valley. Crops cover only ~3% of the total catchment area. Maize is the dominant crop and is irrigated with streamflow and/or

groundwater. Several diversion dams have been constructed along the Hailiutu River for irrigation since the early 1970s.

The other catchments have different characteristics. C2 is also located in the Wuding River basin but mainly covered by loess soils and shows a hilly landscape than that in C1. The landscape of C3 is a mixture of the desert sand dunes in the west and the loess hills in the east. C4-C6 are mainly located in the loess hills region with steep slopes and deep valleys. Among them, C4 is the upstream area of a larger river basin with more flat landscape than the downstream topography. A lot of ravines exist in the loess hills region by which floods can be released to the rivers. Soil and water conservation projects have been conducted to control the floods and sediment loss. A comparison of the hydrological behaviors between C1 and C6 has been presented by Zhou et al. (2015). In the west part of the Erdos Plateau there are also some river basins but they are lack of hydrological data for a proper analysis.

In the study area, groundwater is stored in complex aquifer systems. In general, the Erdos Plateau is characterized by shallow groundwater in the sandy sediments and deep groundwater in the underlying sandstones. In C1-C2, the Cretaceous sandstones form a thick aquifer enabling active groundwater circulation. In C3~C6, the Cretaceous sandstones is limited or replaced by the Jurassic sandstone-mudstone formations with lower permeability so that the movement of deep groundwater is restricted. Meanwhile, shallow groundwater exists in the valleys or near-lake areas covered by sandy or loess sediments. Regional groundwater level distribution in the catchment C1 has been investigated in Lv et al. (2013) based on a hydrogeological survey carried out in 2010 and was shown in Figure 1(c). According to this investigation, depth to water table (DWT) in C1 varies in a large range from zero to 110 m. In more than half of the area, DWT is less than 10 m. The shallow groundwater zone, where DWT is no more than 2 m, occupies 16.0% of the whole catchment area. As investigated in Yin et al. (2015) at a research site in this catchment, when DWT is less than 2 m, the actual evapotranspiration is generally 80% higher than the potential evapotranspiration. This investigation confirmed that groundwater dependent evapotranspiration is an essential process in the Erdos Plateau.

2.2 Data

Daily streamflow data since 1957 for the hydrological stations at the outlets of the 6 catchments were collected from the Yellow River Conservancy Commission. A rainfall gauge

was also installed at the Hanjiamao hydrological station (Fig. 1c) in 1961, providing daily precipitation.

To better account for the variability of rainfall in space and time, we develop gridded monthly precipitation data with 1-km resolution between 1957 and 2010 from the data of 15 national meteorological stations in the Erdos Plateau (Fig. 1b). Monthly rainfall data at these stations were downloaded from the China Meteorological Data Sharing Service System (CMDSSS, <http://cdc.nmic.cn>). We construct the gridded data using the inverse distance square weighting (IDSW) method due to the moderate topography of the Erdos Plateau in the form of low-relief rolling hills. Fig. 1b shows the mean annual precipitation contours of the Erdos Plateau obtained from the gridded data. In this study, the area-averaged monthly data of the precipitation in the 6 catchments for the period 1957-2010 are estimated by imposing the basin boundaries on the gridded monthly precipitation data and taking the arithmetic average of the cell values within the catchment.

The method applied in constructing the gridded precipitation data is further applied in constructing a 1-km resolution gridded dataset for the monthly pan evaporation between 1957 and 2010 covering the Erdos Plateau. The pan evaporation data were based on observations from 200-mm diameter pans that were installed in most stations in the Erdos Plateau and can be also downloaded from CMDSSS (<http://cdc.nmic.cn>). The average monthly data of the potential evapotranspiration (E_0) in the 6 catchments are estimated from the spatially averaged data of the pan evaporation using a local pan coefficient (0.58) for the 200-mm diameter pan. This coefficient was suggested by various investigations of pan coefficients for Chinese meteorological stations (Shi et al., 1986; Fan et al., 2006).

In summary, the mean annual values of P , E_0 and Q during the period in 1957-1978 for the 6 catchments are listed in Table 1. In this period, the streamflow was not significantly influenced by the hydraulic engineering, irrigation water use and coal mine industry so that the hydrological behavior was close to the natural state. It can be estimated from the data that the mean annual Q/P values in this 'natural' state ranged between 0.1 and 0.3 for the catchments. Accordingly, the mean annual F values are higher than 0.7 with respect to the mean values of the aridity index (E_0/P) varying between 2.5 and 3.4.

In Fig. 2a, the variation patterns of the monthly rainfall and potential evapotranspiration during 1957-2010 are shown for the catchment C1. Both rainfall and evapotranspiration are

high in the summer and low in the winter. However, there is a difference in the patterns by which the seasonal variation in runoff may be influenced: the rainfall peak normally arrives in the August but the highest evaporation is exhibited in the June. With respect to these meteorological patterns, the total runoff drops in the Spring and in the early Summer until the heavy rainfall comes in the August, as shown in Fig. 2b. In comparison with the rainfall and the potential evapotranspiration, the mean monthly runoff (2.6 mm) and its fluctuation amplitude (0.8-11.9 mm) are quite small. This indicates that most of the precipitation in C1 returns to the atmosphere by evapotranspiration. During 1957 to 2010, the annual aridity index in the catchment showed a large variation range (between 1 and 10), covering the semi-humid, semi-arid and arid climatic conditions as classified in the scheme recommended by the United Nations Environment Programme (UNEP) (Middleton and Thomas, 1992).

In the study area, there are significant interannual fluctuations in streamflow. For the catchment C1, Yang et al. (2012) investigated the annual regime shifts in streamflow and found that the shifts were caused largely by land use policy changes and river water diversions for irrigation. Tab. 2 shows the mean annual fluxes in four typical periods with different numbers of water diversions in the Hailutu River and major branches during 1957-2010. These diversions influenced the hydrological behavior in C1 and will be discussed in the following sections. However, before 1967, the Hailutu River was free of hydraulic engineering, and the studied area was mostly close to the natural conditions. In the other catchments, the changes in the streamflow regime were also mainly contributed by human activities but in more complex ways. In C6, the river discharge was also influenced by a large number of check dams that constructed to reduce water and sediment loss (Zhou et al., 2015). In C3-C5, the impacts of coal mining industry were significant. To analyze the natural hydrological behaviors, the study period should not be later than 1978.

2.3 Preliminary analysis using $(P-Q)/P$

In many cases, it is possible to estimate the annual E in a catchment from the annually observed P and Q by $P-Q$ when the change in storage is significantly small. Then it could be treated as the "real" data of the annual E and the shift of annual water balance in the Budyko space could be investigated with the plot of $(P-Q)/P$ versus ϕ . In this section, we check the validity of this approach in the study area.

The plots of annual $(P-Q)/P$ versus ϕ during 1957-1878 (close to the natural state) for the 6 catchments are shown in Fig. 3. In particular, Fig. 3a shows the locations of the mean annual $(P-Q)/P$ data of the 6 catchments in the Budyko space. The data points fall below the original Budyko curve given by Eq. (2) but can be bounded between two modified Budyko curves that determined using Eq. (3). They approximately exhibit a positive relation between $(P-Q)/P$ and ϕ . It indicates that the behaviors of the catchments in long-term average (22-years) satisfied the steady state water balance assumption in the original Budyko framework.

If the original Budyko formula (Budyko, 1958) is available for the annual water balance, the annual $(P-Q)/P$ value should be equal to $F(\phi)$ so that the shift path of $(P-Q)/P$ in each catchment should be an increasing curve with a positive slope in the Budyko space. However, as shown in Fig. 3(b-d), the annual $(P-Q)/P$ data in C1-C4 follow a negative relation. The annual $(P-Q)/P$ value of C3 significantly decreased from ~0.8 to ~0.5 when the aridity index increased from 1.2 to 5.8 (Fig. 3c). It seems that C5 (Fig. 3e) and C6 (Fig. 3f) showed a positive relation but the data points did not fall closely along the original Budyko curve. The negative relation in C1-C4 are contrary to the positive relation in the original Budyko framework, indicating the false of taking the annual $(P-Q)/P$ as the replica of the annual F value in the study area. In a previous study, Istanbulluoglu et al. (2012) also highlighted this abnormal relation in the North Loup River basin, Nebraska, USA, and they demonstrated that it was caused by ignoring the change in storage. They used long-term monitoring data of groundwater level to estimate the inter-annual change in groundwater storage (ΔG) and replaced the $(P-Q)/P$ data with the $(P-Q-\Delta G)/P$ data to reproduce a normal Budyko curve for the basin. However, groundwater dependent evapotranspiration was not explicitly considered in Istanbulluoglu et al. (2012).

It is a good idea to estimate the change in groundwater storage using groundwater monitoring data. However, long-term groundwater level monitoring data are not available for the catchments in this study. In addition, the approach of using $(P-Q-\Delta G)/P$ data ignores the inter-annual change in the soil moisture storage. In a different way from Istanbulluoglu et al. (2012), we estimate ΔG from the monthly baseflow (groundwater discharge) data with a calibrated hydrological model, in which the groundwater dependent evapotranspiration is also incorporated. Using the model, the storage components and the contribution of groundwater for the annual E can be simultaneously obtained at the catchment scale.

3 Hydroclimatologic models

3.1 The ABCD model

The ABCD model is a conceptual hydrological model with 4 parameters (a , b , c , and d) developed by Thomas (1981) to account for the actual evapotranspiration, surface/sub-surface runoff and storage changes. The ABCD model was originally applied at an annual time step but has been recommended as a monthly hydrological model (Alley, 1984). It was widely applied as a hydroclimatologic model to investigate the response of catchments to climate change (Vandewiele et al., 1992; Fernandez et al., 2000; Sankarasubramanian and Vogel, 2002; Li and Sankarasubramanian, 2012).

Both the soil water and groundwater are considered in the ABCD model, as shown in Fig. 4a. At the monthly time step, the change in the soil water storage is determined by:

$$W_m - W_{m-1} = P_m - E_m - R_m \quad (4)$$

where W_{m-1} and W_m are the effective soil water storages at the beginning and at the end of the m -th month, respectively; P_m and E_m are the monthly precipitation and evapotranspiration, respectively; and R_m is the monthly loss of soil water via direct runoff and groundwater recharge. The change in groundwater storage is determined by:

$$G_m - G_{m-1} = cR_m - dG_m \quad (5)$$

where G_{m-1} and G_m represent the groundwater storage at the beginning and the end of the m -th month, respectively; c and d are two parameters that account for groundwater recharge and discharge from R_m and G_m , respectively. The monthly streamflow is the summation of the monthly direct runoff and groundwater discharge, as follows:

$$Q_m = (1 - c)R_m + dG_m \quad (6)$$

The change in storage in the ABCD model is the summation of the changes in the soil water storage and groundwater storage, which can be expressed as $(W_m - W_{m-1}) + (G_m - G_{m-1})$.

Thomas (1981) proposed a nonlinear function to estimate $(E_m + W_m)$ from $(P_m + W_{m-1})$ as follows:

$$E_m + W_m = \frac{P_m + W_{m-1} + b}{2a} - \sqrt{\left(\frac{P_m + W_{m-1} + b}{2a}\right)^2 - \frac{(P_m + W_{m-1})b}{a}} \quad (7)$$

where a is a dimensionless parameter, and b is the upper limit of $(E_m + W_m)$. In addition, Thomas (1981) assumed:

$$W_m = (E_m + W_m) \exp(-E_{0m}/b) \quad (8)$$

where E_{0m} is the monthly potential evaporation for the m -th month. Substituting Eq. (8) into Eq. (7), the monthly evapotranspiration can be estimated as:

$$E_m = \left[\frac{P_m + W_{m-1} + b}{2a} - \sqrt{\left(\frac{P_m + W_{m-1} + b}{2a}\right)^2 - \frac{(P_m + W_{m-1})b}{a}} \right] \left[1 - \exp\left(-\frac{E_{0m}}{b}\right) \right] \quad (9)$$

Wang and Tang (2014) demonstrated that Eq. (9) can be derived from the generalized proportionality principle and yield an equivalent Budyko-type model.

3.2 The ABCD-GE model

To investigate the effect of groundwater dependent evapotranspiration in basins with both shallow and deep groundwater, the original ABCD model is extended in this study as the ABCD-GE model, where ‘GE’ denotes groundwater dependent evapotranspiration. As shown in Fig. 4b, a catchment is conceptually divided into two zones where the Zone-1 and Zone-2 represent different areas with deep and shallow groundwater, respectively. Direct runoff is originated from both zones. Surface water (water in river, canals, lakes, etc.) is also included in the Zone-2. The soil water reservoir in the Zone-1 is the same as that in the ABCD model. In addition, a transition vadose zone is specified between the soil layer and water table to represent the delayed groundwater recharge. The transition zone is included to handle the existence of the thick unsaturated zone (>10 m) in a basin. Soil water in this zone is dominated by the vertical downward flow. In the Zone-2, the rainfall and evapotranspiration are the components directly involved in the water balance of groundwater. Thus, three storage components are considered as a chain in the hydrological processes. It is assumed that the

potential change in groundwater storage by the lateral flow coming in and out of a catchment is negligible. For a large river basin ($>1000 \text{ km}^2$) this assumption is generally acceptable.

Dividing the Zone-1 and Zone-2 in a catchment depends on how groundwater can be accessed by the evapotranspiration process. It is controlled by the depth of plant roots and the rise of capillary water above water table. In the case study of the catchment C1, it was observed that some trees have long roots penetrated 2-3 m or more into the soil (Lv et al., 2013), but in general the dominant root zone is less than 2 m below ground surface for shrubs and grasses. When the DWT is larger than 2 m, the contribution of groundwater for evapotranspiration will dramatically decrease to a negligible level (Yin et al., 2015). Thus, it is reasonable to use the contours of 2-m-depth of groundwater as the approximate boundary between the Zone-1 and Zone-2 in the study area. In the Zone-1, the transition vadose zone is roughly defined as the zone between 2-m-depth below ground surface and 2-m-height above water table. In the assumptions of the ABCD-GE model, this zone could not be influenced by both the evapotranspiration and groundwater flow processes. Thus, the thickness of the soil layer would be less than 2 m in the model. However, one should be aware of that it is not necessary to find the distinct and exact boundaries for the zones, since the ABCD-GE model is a conceptual hydrological model.

Similar to that in the ABCD model, the change in the soil water storage in the Zone-1 is determined by:

$$W_m - W_{m-1} = P_m - E_{1m} - R_m \quad (10)$$

where E_{1m} is the monthly evapotranspiration in the Zone-1 determined with Eq. (9); R_m becomes the summation of the leaking soil water to the transition vadose zone, cR_m , and the direct runoff, $(1-c)R_m$, in the Zone-1. The change in storage of the vadose zone is described with:

$$V_m - V_{m-1} = cR_m - kV_m \quad (11)$$

where V_m and V_{m-1} represent the values of the storage in the transition vadose zone at the end and beginning of the m -th month, respectively; and k is the parameter that accounts for the groundwater recharge rate as kV_m .

In the ABCD-GE model, the source of runoff is partitioned by the fractions of the area of the two zones. The total runoff in the catchment is the summation of the direct runoff and groundwater discharge as follows:

$$Q_m = (1-\alpha)(1-c)R_m + \alpha(1-c)P_m + dG_m \quad (12)$$

where α is the ratio of the Zone-2 area to the whole catchment area. In Eq. (12), the term $(1-\alpha)(1-c)R_m$ denotes the direct runoff contributed by the Zone-1 whereas the term $\alpha(1-c)P_m$ denotes the direct runoff contributed by the Zone-2. In considering of the gain-loss processes of groundwater, the change in the effective groundwater storage is yielded by:

$$G_m - G_{m-1} = (1-\alpha)kV_m + \alpha(cP_m - E_{2m}) - dG_m \quad (13)$$

where E_{2m} is the monthly evapotranspiration in the Zone-2, which depends on the effective groundwater storage as follows:

$$E_{2m} = gG_mE_{0m} \quad (14)$$

where g is a parameter controlling the intensity of groundwater dependent evapotranspiration. Eq. (14) assumes that the evapotranspiration rate in the Zone-2 is simply proportional to both the groundwater storage (which is positively related to groundwater level) and the potential evapotranspiration rate. Thus, the evapotranspiration rate as a whole in the catchment is summarized as:

$$E_m = (1-\alpha)E_{1m} + \alpha E_{2m} \quad (15)$$

Eqs. (10)-(13) are solved one by one and finally the value of G_m is substituting into Eq. (12) to obtain the runoff. The results of the ABCD-GE model are controlled by 7 parameters as: a , b , c , d , g , k and α . The parameter values can be identified with the model calibration process.

4 Model Calibration and Results

4.1 Model calibration

We apply the ABCD-GE model to estimate the monthly evapotranspiration and the change in the storage components in the 6 catchments after the model parameters were calibrated. The

monthly evapotranspiration data are then summed up to estimate the annual evapotranspiration for further analysis. The model calibration is based on the observed monthly streamflow data at the hydrological stations and the separated baseflow data.

Because groundwater discharge has been included in the model, a baseflow analysis was performed to obtain the expected groundwater discharge for the model calibration. Using the automated hydrograph separation method HYSEP (Sloto and Crouse, 1996) on the daily streamflow data, such ‘observed’ groundwater discharge data were obtained. For C1, these data were partly reported in Zhou et al. (2013). The mean values of the baseflow index for C1-C6 range between 0.13 and 0.88 (Table 1). The catchment C1 has the highest baseflow index (0.88), indicating that groundwater discharge is the dominant hydrological process in this catchment. Variation patterns of the monthly groundwater discharge in C1 are shown in Fig. 2b. In C5 and C6, the baseflow index values are smaller than 0.2 because most of the streamflow in the two catchments are contributed by direct runoff.

The ordinary least squares (OLS) criterion is applied for parameter estimation. The errors of both log-streamflow and log-baseflow were included in the OLS objective function, as follows:

$$\min U = \sum_{m=1}^N (e_m^2 + q_m^2) \quad (16)$$

where:

$$e_m = \ln(\hat{Q} / Q)_m, \quad q_m = \ln(\hat{Q}_b / Q_b)_m \quad (17)$$

and U is the value of the objective function; N is the number of the months; \hat{Q} and Q are the simulated and observed monthly streamflow, respectively; \hat{Q}_b is the simulated monthly groundwater discharge through dG_m in Eq. (12); and Q_b is the ‘observed’ monthly groundwater discharge obtained from the baseflow analysis. The log form errors given in Eq. (17) are used to obtain the homoscedastic residuals (rather than the residual errors) of the normal absolute differences between the observed data and the model outputs (Alley, 1984). The nonlinear optimization algorithm Generalized Reduced Gradient (GRG) (Lasdon et al., 1978) is applied to determine the optimum values of the parameters. The Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970) is also used to evaluate the performance of the

model in simulating the monthly runoff. It ranges in $(-\infty, 1)$ whereas a higher than zero value is required for a well-perform model.

For the catchment C1, the model parameters are firstly identified using the 1957-1966 data, and this calibrated model is considered to be a 'natural' model due to the minimum impact of human activities during this 10-year period. For C2-C6, the 1957-1978 period is applied to roughly identify the 'natural' model. The initial storage values are treated as the unknown parameters to be determined in the calibration process. Changes in the initial conditions generally influenced the simulated results in the first and second years. Therefore, the residual errors in the later years are applied to estimate the parameter values with the less influence from the initial conditions. A sensitivity analysis is carried out to schematically capture the ranges of the parameter values.

The best fitting parameter values of the 'natural' models for C1-C6 are shown in Table 3. The a value ranges between 0.91 and 0.97. In previous studies using the ABCD model, the a value was found generally to be higher than 0.9 (Alley, 1984; Sankarasubramanian and Vogel, 2002; Li and Sankarasubramanian, 2012). The b and c values in this study are generally higher than that obtained by Alley (1984) for ten catchments in the USA. The higher c value indicates the more significant role of groundwater in the hydrological behaviors. However, the d values of C1-C6 fall into the range suggested by Alley (1984). The optimized α values in the 'natural' models range between 0.09 and 0.27. In particular, the α value of C1 (0.21) in this 'natural' state was larger than the current value (DWT in 16.0% of the area is less than 2 m). Such a difference is reasonable because the groundwater level in the 1950s and 1960s should be higher than that at present as indicated by the higher baseflow (Fig. 2b). The k value controls the rate of groundwater recharge below the transition vadose zone. The transition vadose zone is a necessary component in C1-C3 as demonstrated by the sensitivity analysis. When an extremely high value of k is used ($k > 100$), the kV_m value would be almost equal to cR_m so that the transition vadose zone does not make sense. However, in this situation the model could not capture the seasonal variation patterns of groundwater discharge. Thus, the delayed groundwater recharge is an essential process for the catchments. The best fitting k values for C4-C6 are significantly higher than that for C1-C3, indicating a weaker delay effect. This confirms the hydrogeological conditions in C4-C6: active groundwater flow is limited in the near-surface zone.

For the calibration period, the root-mean-square error (RMSE) of the 'natural' models with respect to the aggregated results of the 6 catchments is 39% for the monthly runoff. For the aggregated annual runoff data, the RMSE is 21%. The errors include the streamflow observation error and the meteorological data treatment error. It is more reasonable to evaluate the model performance according to the observation-simulation correlation coefficient and the NSE value. A comparison between the observed and simulated monthly runoff (including groundwater discharge) for all of the 6 catchments can be seen in Fig. 5. The coefficient of determination ($R^2=0.89$) is high. The NSE values of the model range between 0.48 and 0.81 for the different catchments (Table 3), indicating that the model performs well in the study area. It is usually difficult to obtain a high NSE value for a catchment with weak seasonal variation in runoff (Mathevet et al., 2006), such as that in C1 and C2.

4.2 Modelling results

We use the 'natural' models to estimate the monthly hydrological components during the whole 1957-2010 period. As an example, typical results of the catchment C1 are shown in Fig. 6. The modeling results of the monthly runoff after 1970s are generally higher than the observed data (Fig. 6a) due to ignoring the impacts of land use changes and increased utilization of water for irrigation. However, the simulated patterns of groundwater discharge are similar to the observations (Fig. 6b): falling in the summer, rising in the winter. This agreement between the simulated and observed patterns demonstrates the ability of the ABCD-GE model in simulating the hydrological behaviors in the studied catchment: significant groundwater-dependent evapotranspiration occurs in the summer, and a strong recovery of storage in the shallow-groundwater zone occurs in the winter due to delayed recharge from the thick vadose zone.

For the catchment C1, there are significant differences between the observed and model calculated annual runoff after 1966, as shown in Fig. 6c. This deviation could be interpreted as the excess evapotranspiration induced by the increasing agricultural water use. Enhanced evapotranspiration also occurred in the shallow groundwater zone due to groundwater pumping for irrigation. To evaluate the actual water balance, the following equation:

$$E_{ACT} \approx E_{NAT} + (Q_{NAT} - Q_{OBS}) \quad (18)$$

is applied to approximately estimate the actual annual evapotranspiration (E_{ACT}) after 1966 from the ‘natural’ model result (E_{NAT}) plus the difference of the annual runoff between the ‘natural’ model (Q_{NAT}) and the observation (Q_{OBS}). This difference may be partly induced by the multiple time-scale variations in the climate conditions but would be mainly caused by the irrigation water use. Results are shown in Fig. 6d. It seems that the relative difference between E_{NAT} and E_{ACT} is not significant. The maximum $Q_{NAT}-Q_{OBS}$ value is less than 10% of the mean annual evapotranspiration (~315 mm). Accordingly, the irrigation water use in this catchment did not significantly influence the annual evapotranspiration. However, it dramatically influenced the streamflow. As shown in Fig. 6a, almost all of the direct runoff was removed from the total runoff after 1987 and the groundwater discharge was significantly decreased even though the seasonal patterns basically remained (Fig. 6b).

4.3 Annual water balance in the standard Budyko space

In Fig. 7, the annual F data for the annual water balance obtained from the ‘natural’ models over the 1957-2010 period are plotted in the standard Budyko space. It is obvious that with the increasing aridity index (ϕ), the evapotranspiration ratio (F) for all of the catchments increased in almost a linearly patterns with the different slopes. When $\phi < 4$, most of the data points fall below the original Budyko curve, indicating that $F < 1$ is generally satisfied in this situation. When $\phi > 4$, the original Budyko curve gives $F \approx 1$, indicating the limitation ($0 < F \leq 1$) for the mean annual evapotranspiration ratio. However, some of the data points fall above the line of $F=1$ while ϕ is larger than 3 but even less than 4, indicating that the $F > 1$ cases were not only occurred in dry years. The maximum F value (2.2) was obtained in C1 when the aridity index jumped from $\phi=1.5$ in 1964 to $\phi=9.8$ in 1965. It means that C1 lost a volume of water in 1965 by evapotranspiration, which is more than twice of the gained water from precipitation in the same year. In comparison, the F values in C4 and C6 are not sensitive to the change in the aridity index since the slopes of the F - ϕ regression lines are less than 0.1.

The effect of groundwater dependent evapotranspiration can be clearly observed when the evapotranspiration ratio is divided into two parts and plotted in the Budyko space separately with respect to the shallow and deep groundwater zones. The annual E values in the Zone-1 and Zone-2 are estimated, respectively, as:

$$E_1 = \sum_{m=1}^{12} E_{1m}(W_{m-1}, a, b) \quad \text{and} \quad E_2 = \sum_{m=1}^{12} E_{2m}(G_m, g) \quad (19)$$

for every year, where E_{1m} is calculated with Eq. (9), whereas E_{2m} is calculated with Eq. (14). Typical results are shown in Fig. 8 for C1 and C4. The data in Fig. 8 are estimated with the parameter values of a , b and g for the ‘natural’ models. It is obvious that the annual E_1/P values in the Zone-1 (deep groundwater) for the whole range of the aridity index ($0 < \phi < 8$) are smaller than 1.0 and fall below the original Budyko curve determined by Eq. (2). The low E_1/P value in the Zone-1 is mainly due to the water limited condition in the soils without the supply of groundwater, especially in C1 where the landscape is dominated by sand dunes. The variation of E_1/P can be approximately fitted by the Budyko curve determined with Eq. (3) using $w = 1.4$ and 1.6, respectively, for C1 and C4. The higher w value in C4 is linked with the higher b value of the model for C4, as listed in Table 3. However, the relationship between E_2/P and the annual aridity index in the shallow groundwater zone definitely could not be explained by any of the standard Budyko formulas, because for the Zone-2 all the annual F values are higher than 1.0 in C1 and most of the data points fall above the bound in C4.

In the catchment C1, the E_2/P data points follow a linear regression line with the slope of 0.52. This agrees with the relationship between E_2 and E_0 ($E_2 \propto E_0$) that described in Eq. (14). Since the groundwater storage, G , is relatively stable (small d and k values in the model), the annual E_2/P value would be proportional to the ϕ value and the slope is close to the annual mean value of gG . In C1, the annual mean value of gG is 0.65 according to the ‘natural’ model. Such a groundwater dependent evapotranspiration process is the reason for the occurrence of the $F > 1$ cases at the catchment scale. Note that in the original Budyko framework, the $F = \phi$ case denotes an energy-limited condition when water supply (only precipitation for mean annual water balance) is sufficient for the evapotranspiration process. The slope of the E_2/P line in C1 is less than 1.0 but is closer to the $F = \phi$ line than the water limited line represented by $F = 1$. It indicates that in the Zone-2 the evapotranspiration process is in a quasi-energy limited condition, rather than in a water limited condition, because shallow groundwater can effectively serves as an external source of water supply.

In C4, the E_2/P data points show a scatter distribution around the regression line. This is mainly caused by the significant variability in the groundwater storage, G , at the monthly and annual scales. The d and k values in the model for C4 are quite larger than that for C1 by

which the model could capture the significant fluctuation of **the** baseflow. As a result, the annual mean gG value of C4 **vary** in a large range between 0.18 and 0.84. Similar unstable E_2/P data also exist for C5 and C6 as indicated by the high d and k values (**Table 3**).

5 Discussions

5.1 Controls on the $F > 1$ cases

It has been demonstrated in **Fig. 7** that the annual evapotranspiration ratio, F , could be higher than 1.0 when the aridity index, ϕ , is larger than 4.0 in the studied catchments. In the literature, the $F > 1$ cases were also observed when ϕ is just higher than 1.0 (**Cheng et al., 2011; Wang, 2012; Chen et al, 2013**). Thus, it is interesting to discuss how the occurrence of the $F > 1$ cases is controlled by the catchment properties when shallow groundwater plays an important role.

The equation for the annual evapotranspiration ratio can be derived from Eqs. (15) and (19) as follows:

$$F = (1 - \alpha) \frac{E_1}{P} + \alpha g \frac{E_0}{P} \sum_{m=1}^{12} \left(\frac{E_{0m}}{E_0} G_m \right) \quad (20)$$

where the term E_{0m}/E_0 denotes the proportion of **the** monthly potential evaporation to the annual one with respect to the m -th month. It has been known that the relationship between E_1/P and ϕ determined by the ABCD model is similar to that predicted by the standard Budyko formulas, as shown in **Fig. 8**, where E_1/P is less than 1.0. For the groundwater dependent term, defining:

$$G_a = \sum_{m=1}^{12} \left(\frac{E_{0m}}{E_0} G_m \right) \quad (21)$$

as the weighted average of the monthly groundwater storage, Eq. (20) can be replaced by:

$$F(\phi) = (1 - \alpha)[1 + \phi - (1 + \phi^w)^{1/w}] + \alpha g G_a \phi \quad (22)$$

where E_1/P is represented by Eq. (3). According to Eq. (22), the function $F(\phi)$ is controlled by the parameters, g , w , α and the status of groundwater represented by G_a . As indicated in Eq.

(14), gG_a is a dimensionless parameter to describe the intensity of groundwater dependent evapotranspiration related to the potential evaporation. The recommended range of gG_a is 0.2-1.0 according to Table 3. In Eq. (22), the term with w indicates the normal energy-water limited process in the Zone-1, whereas the term with gG_a indicates the quasi-energy limited process in the Zone-2. The actual F value is a mixed result of the different processes.

Typical F - ϕ curves obtained with Eq. (22) are plotted in Fig. 9. It can be seen that the proportion of shallow water table area (α) has large effect on the occurrence of the $F>1$ case. When the shallow water table area is small ($\alpha=0.1$), the $F>1$ case occurs only during dry years. When the gG_a value increases, the $F>1$ case occurs at the smaller aridity index. The specific catchment parameter (w) for E_1/P also influences the occurrence of the $F>1$ case. A larger w value shifts the F - ϕ curves (comparing Fig. 9b with Fig. 9a) to the left side indicating that the $F>1$ case could occur at smaller aridity index.

Groundwater dependent evapotranspiration estimated in the ABCD-GE model does not violate the $F<1$ rule for the long-term steady state water balance, because the model will yield $E=P-Q$ for the average flux in a long-term period. As shown in Fig. 9, more than half of the data points fall below the line of $F=1$, indicating that the less-than-1 rule for the average F value is satisfied in the whole study area. For each catchment, in addition, the mean annual F - ϕ data for the 1957-2010 period have been shown in Fig. 7 (intersection points of the green cross lines). None of these mean annual F values is higher than 1.

5.2 Using effective precipitation and modified Budyko space

The standard Budyko space assumes that the potential water supply for evapotranspiration is only rainfall in a catchment. This is valid for the mean annual water balance, but exceptions might exist for the annual or intra-annual behaviors. Several previous studies attempted to modify the Budyko framework for the short time scale. Wang (2012) and Chen et al. (2013) argued that the reduction of storage in a period should be regarded as one of the water supply components. They suggested an approach to replace the evapotranspiration ratio and the aridity index by $E/(P-\Delta S)$ and $E_0/(P-\Delta S)$, respectively, where ΔS is the storage depletion in a studied period and $P-\Delta S$ is regarded as the effective precipitation. In this modified Budyko space, evapotranspiration is always less than the water supply so that the original Budyko hypothesis could be satisfied for the small time-scale problems. Alternatively, Greve et al.

(2016) proposed a two-parameter Budyko function to explain the cases of evapotranspiration exceeds precipitation in unsteady state conditions. However, groundwater flow is not included in their model so that the role of groundwater dependent evapotranspiration could not be assessed by such a two-parameter Budyko function.

In this section, we attempt to check the characteristics of the annual water balance data in the study area using the modified Budyko space suggested by Wang (2012) and Chen et al. (2013). With the results of the ABCD-GE model, the total change in storage for a year can be estimated as:

$$\Delta S = \sum_{m=1}^{12} [(1 - \alpha)(W_m + V_m - W_{m-1} - V_{m-1}) + (G_m - G_{m-1})] \quad (23)$$

where m is the number of the months in the year, W_0 , V_0 and G_0 for $m=0$ denoting the respective storage components at the end of the previous year.

Results are shown in Fig. 10 for the 6 studied catchments. It can be seen that in the modified Budyko space the annual water balance data fall into the zone below the limitation: $E/(P-\Delta S) < 1$. However, in any one of the catchments, the shift path of the data points could not be captured by a single Budyko curve in the modified Budyko space. For the cases in C1 and C2, the increase of $E/(P-\Delta S)$ with the increasing $E_0/(P-\Delta S)$ seems too small in comparison with any one of the standard Budyko curves determined by Eq. (3). Similar difference between the data and the standard Budyko curves also exists in the other catchments but not as significant as that in C1 and C2. Furthermore, in C1 and C2 the $E/(P-\Delta S)$ value approaches a stable value around 0.90 with the high $E_0/(P-\Delta S)$ values. It indicates that at least 10% of $P-\Delta S$ is contributed to the annual runoff, in terms of $Q/(P-\Delta S)$. This portion of the water supply seems to be inaccessible for the evapotranspiration process. Similar bounds of the $E/(P-\Delta S)$ value also exist in C3-C6. In particular, this limitation is lower than 0.8 in C3, implying significant contribution of change in storage to the streamflow in dry years.

The difficulties in using the effective precipitation defined by Wang (2012) and Chen et al. (2013) are the unknown ΔS for an investigated time step and the possible existence of the inaccessible part of ΔS for the evapotranspiration process. Consequently, the estimation of $E/(P-\Delta S)$ value is not straightforward, but requires a complex iteration process. In the original

Budyko framework for the steady state water balance, the water supply (only precipitation) does not depend on both evapotranspiration and runoff so that the aridity index is an independent variable in assessing the behaviors of the catchments. However, the water supply represented by the effective precipitation is influenced by the evapotranspiration-runoff processes due to the feedback mechanism. This interdependence between the water supply and evapotranspiration significantly reduces the efficiency of using the modified Budyko space in analyzing the shift of annual water balance in a catchment. In contrast, it would be an efficient and straightforward approach to extend formulas for annual water balance in the standard Budyko space, such as Eq. (22), keeping an independent index (ϕ) for the climatic conditions.

5.3 Landscape-driven and human-controlled shifts of annual water balance

As illustrated in Fig. 6d, the actual evapotranspiration in the catchment C1 has been enhanced by human activities. This impact might exist in both the shallow and deep groundwater zones. Crops in C1 are mainly planted in the depressions and terrace lands with shallow groundwater, especially in the river valley. Crops require much more water than the precipitation for growing. For example, maize could consume more than 3 times of rainfall water in growing seasons (Zhou et al., 2013). Thus, irrigation is necessary to maintain the agricultural production. In the croplands far away from the rivers, groundwater was abstracted from wells for irrigation. In the river valley, irrigation was realized with the diversions and channels. Therefore, an increase in the evapotranspiration in the shallow groundwater zone is dominated by irrigation. Along the river, the area of the surface water body was significantly enlarged in the reservoirs, leading to an increase in the surface water evapotranspiration loss. It is equivalent to the increase in groundwater dependent evapotranspiration in this study because surface water is also included in the shallow groundwater zone. As a result, the shift of the annual water balance in the Budyko space was partly caused by change in land use and controlled by regulation of river water for irrigation.

Recently, Jaramillo and Destouni (2014) developed a method to assess the landscape-driven change in the mean evapotranspiration ratio using the difference between the actual change in the F value and the climate-driven change in the F value following the Budyko framework. In this section, we extend their method to assess the landscape-driven change in annual water balance in the catchment C1. The period between 1957-1966 is selected from Table 2 as the

reference period. Changes are evaluated for the different average values of the annual F data in the different periods listed in Table 2. The climate-driven change is estimated with the annual E_{NAT} values obtained from the 'natural' model, using a formula similar to Jaramillo and Destouni (2014), as follows:

$$\Delta\left(\frac{E_{LD}}{P}\right) = \Delta\left(\frac{E_{ACT}}{P}\right) - \Delta\left(\frac{E_{NAT}}{P}\right) \quad (24)$$

where $\Delta(E_{LD}/P)$ denotes the landscape-driven change in comparison with the 1957-1966 period. However, this quantity index includes the landscape changes driven by both of the climatic force and human activities. To check how this index is correlated with the increasing impacts from the reservoirs and diversions in rivers, following Jaramillo and Destouni (2015), the coefficient of intra-annual variation of the monthly runoff (CV_Q) was applied. The CV_Q/CV_P value was estimated to reveal the separate influence of such a human-controlled flow regulation from the mixed human-climate controlling, where CV_P is the coefficient of intra-annual variation of the monthly rainfall.

Results of the $\Delta(E_{LD}/P)$ and $\Delta(CV_Q/CV_P)$ data between the three periods 1968-1987, 1988-1997, 1998-2010 and the reference period 1957-1967 are shown for the catchment C1 in Fig. 11. The $\Delta(E_{LD}/P)$ values are all positive but not big (less than 6%), indicating a slight increase in the evapotranspiration ratio after 1966 driven by the changes in the natural landscape conditions of water storage and/or human controlled land use. The $\Delta(CV_Q/CV_P)$ values show a significant fluctuation around zero but also limited in a small range ($\pm 5\%$). Both the $\Delta(E_{LD}/P)$ and $\Delta(CV_Q/CV_P)$ values are largest in 1988-1997. Fluctuations of these data could not be fully explained by the increasing number of diversions in the rivers. The negative $\Delta(CV_Q/CV_P)$ value in 1968-1987 may be caused by the construction of two reservoirs since reservoirs commonly smooth the variation of the streamflow. In 1988-1997, the $\Delta(CV_Q/CV_P)$ value turned to positive when 5 new diversions were built, indicating the opposite impacts of the reservoirs and diversions. It is possible that the streamflow was disturbed by the regulation of water for irrigation on these diversions with small overflow dams. The decrease in the $\Delta(CV_Q/CV_P)$ value from 4.72% in 1988-1977 to 0.72% in 1998-2010 may be caused by the control of the river water use under some government policies to prevent the desertification (Yang et al., 2012; Zhou et al., 2015). The following decrease in the $\Delta(E_{LD}/P)$ value from 5.05% in 1988-1997 to 3.73% in 1998-2010 is not significant, seems indicating the alternative

irrigation practice in the croplands (for example pumping groundwater) so that the real water consumption was reduced but still on a high level. As a result, utilization of surface water and groundwater for irrigation can increase the frequency of the $F>1$ cases.

5.4 Limitation Remarks

Attention should be paid to the simplifications in the conceptual model extended from the ABCD model, when the equations and formulas are applied in complicated catchments. The ABCD model assumes that the storage-evapotranspiration relationship is controlled by the parameters a and b whereas the physical interpretation of them is difficult (Alley, 1984). Eq. (8) in the ABCD model is also hypothesized from a simplified storage-loss model that controlled by the parameter b (Thomas, 1981). Sankarasubramanian and Vogel (2002) suggested that the b value for the annual water balance could be approximately represented by the maximum soil moisture field capacity plus the maximum E_0 for $\phi<1$ or the maximum P for $\phi\geq 1$. The a value is generally estimated in a small range between 0.95 and 1.0. In this study, the model output is not sensitive to the a value. The correlation between a and b may exist because both of them are positively related with E_m+W_m in Eq. (7). The ABCD model neglects the possibility of groundwater-dependent evapotranspiration which has been incorporated in the ABCD-GE model. The ABCD-GE model divides the area into shallow and deep groundwater zones, without considering a complicated spatial distribution of groundwater depth. For the shallow groundwater zone, the evapotranspiration is assumed to be proportional to the groundwater storage. Nonlinear behavior in groundwater dependent evapotranspiration could be further included if it can be successfully parameterized. Linear groundwater storage-discharge relationship is adopted in both of the ABCD and ABCD-GE models. These simplifications could cause systematic errors in modeling a catchment where the nonlinear behaviors in the hydrological processes are significant.

Limitations in the data processing and complexities in the hydrogeological conditions could also influence the accuracy of the modeling results. The average potential evapotranspiration data, E_0 , and the average precipitation data, P , in some degree, are dependent on the estimation methods and may introduce biases. For the hydrogeological conditions, this study assumed that the boundary of a catchment determined along the terrain divides is also the boundary of groundwater flow, *i. e.*, no inter-basin transfer of groundwater exists. The assumption is plausible acceptable in the east part of the Erdos Plateau because the spatial

variation in the groundwater level is highly correlated with the land surface elevation in this area (Lv et al. 2013; Zhou et al., 2015). However, the groundwater flow system in the west part of the Erdos Plateau is more complex where the climate is more arid and water table divides could be significantly different from the terrain divides.

In fact, when the Budyko framework is applied for small time-scale water balance in a catchment, the other additional sources of water supply should be considered, apart from groundwater. Significant changes in soil moisture, snow cover or frozen water in cold regions could also cause an 'abnormal' shift of the annual water balance for a catchment in the standard Budyko space (Jaramillo and Destouni, 2014). The effects of these storage components are negligible in this study but may be essential in other study areas. In particular, the special processes in the cold regions are not included in the ABCD-GE model. However, one can refer to Martinez and Gupta (2010) where the snow-augmented ABCD model was proposed and can be incorporated into an extension of the ABCD-GE model.

6 Conclusions

The Budyko framework was developed for the long-term steady state water balance in catchments, which estimates the evapotranspiration ratio (F) as a function of the aridity index (ϕ). It can be represented by curves for the F - ϕ relationship in the standard Budyko space that were determined by the original Budyko formula without any parameter or the formulas with a catchment specific parameter. It is interesting to investigate whether the Budyko space can be also applied to capture the annual water balance in a catchment with varying dryness. However, the shift of the annual water balance in the standard Budyko space could be significantly different from that presumed from the standard Budyko curves, in particular, when the cases of $F > 1$ occur as that have been observed in a number of catchments.

In this study, we highlight the effect of groundwater dependent evapotranspiration in triggering the abnormal shift of the annual water balance in the standard Budyko space. A conceptual monthly hydrological model, the ABCD-GE model, is developed from the widely used ABCD model to incorporate groundwater-dependent evapotranspiration in the zone with shallow water table and delayed groundwater recharge in the zone with deep water table. The

model is successfully applied to analyze the behaviors of 6 catchments in the Erdos Plateau, China.

The results show that the standard Budyko formulas are not applicable for the interannual variability of catchment water balance when groundwater dependent evapotranspiration is significant. The shift of the annual water balance in the F - ϕ space is a combination of the Budyko-type response in the deep groundwater zone and the quasi-energy-limited condition in the shallow groundwater zone. Shallow groundwater supplies excess evapotranspiration during dry years, leading to the $F > 1$ cases. The occurrence of the $F > 1$ cases depends on the proportion area of the shallow groundwater zone, the intensity of groundwater dependent evapotranspiration and the catchment properties determining the Budyko-type F - ϕ relationship in the deep groundwater zone. Water utilization for irrigation may enhance this excess evapotranspiration phenomenon. The modified Budyko space with the effective precipitation incorporating the change in storage can force F values below 1.0. However, the computation is complicated in dealing with the gain-loss feedback and uncertain with the inaccessible storage for the evapotranspiration process. The empirical formula proposed in this study for the standard Budyko space provides a straightforward method to predict the changes in the annual water balance with the varying dryness.

Acknowledgements

This study is supported by the Program for New Century Excellent Talents in University (NCET) that granted by the Ministry of Education, China, and partly supported by the Honor Power Foundation, UNESCO-IHE. The authors are grateful to the constructive comments from Dr. Donohue, Dr. Jaramillo and the other anonymous reviewers.

1 **References**

- 2 Arora, V. K.: The use of the aridity index to assess climate change effect on annual runoff, J.
3 Hydrol., 265, 164-177, 2002.
- 4 Alley, W. M.: On the treatment of evapotranspiration, soil moisture accounting, and aquifer
5 recharge in monthly water balance models, Water Resour. Res., 20, 1137–1149, 1984.
- 6 Budyko, M. I.: Evaporation under natural conditions, Isr. Program for Sci. Transl., Jerusalem,
7 1948.
- 8 Budyko, M. I.: The heat balance of the earth's surface, U.S. Department of Commerce,
9 Washington, D. C., U.S.A., 1958.
- 10 Budyko, M. I.: Climate and life, Academic, New York, U.S.A., 1974.
- 11 Chen, X., and Hu, Q.: Groundwater influences on soil moisture and surface evaporation, J.
12 Hydrol., 297, 285–300, 2004.
- 13 Cheng, L., Xu, Z., Wang, D., and Cai, X.: Assessing interannual variability of
14 evapotranspiration at the catchment scale using satellite-based evapotranspiration data sets,
15 Water Resour. Res., 47, W09509, doi:10.1029/2011WR010636, 2011
- 16 Chen, X., Alimohammadi, N., and Wang D.: Modeling interannual variability of seasonal
17 evaporation and storage change based on the extended Budyko framework, Water Resour.
18 Res., 49, 6067–6078, doi:10.1002/wrcr.20493, 2013
- 19 Cohen, D., Person, M., Daannen, R., Sharon, L., Dahlstrom, D., Zabielski, V., Winter, T.C.,
20 Rosenberry, D.O., Wright, H., Ito, E., Nieber, J., and Gutowski Jr, W.J: Groundwater-
21 supported evapotranspiration within glaciated watersheds under conditions of climate change.
22 J. Hydrol., 320, 484–500, 2006.
- 23 Fan, J., Wang, Q., and Hao, M.: Estimation of reference crop evapotranspiration by Chinese
24 pan, Transactions of the CSAE, 22(7), 14-17, 2006. (in Chinese)
- 25 Fernandez, W., Vogel, R. M. and Sankarasubramanian A.: Regional calibration of a
26 watershed model, Hydrol. Sci. J., 45, 689–707, 2000.

1 Fu, B. P.: On the calculation of the evaporation from land surface, *Sci. Atmos. Sin.*, 5, 23-31,
2 1981. (in Chinese)

3 Gentine, P., D'Odorico, P., Lintner, B. R., Sivandran, G. and Salvucci, G.: Interdependence of
4 climate, soil, and vegetation as constrained by the Budyko curve, *Geophysical Research*
5 *Letters*, 39, L19404, 2012. doi:10.1029/2012GL053492.

6 Gerrits, A. M. J., Savenije, H. H. G., Veling, E. J. M., and Pfister, L.: Analytical derivation of
7 the Budyko curve based on rainfall characteristics and a simple evaporation model, *Water*
8 *Resour. Res.*, 45, W04403, 2009. doi: 10.1029/2008wr007308.

9 Greve, P., Gudmundsson, L., Orlowsky, B., and Seneviratne, S. I.: Introducing a probabilistic
10 Budyko framework, *Geophysical Research Letters*, 42: 2261-2269, 2015.

11 Greve, P., Gudmundsson, L., Orlowsky, B., and Seneviratne, S. I.: A two-parameter Budyko
12 function to represent conditions under which evapotranspiration exceeds precipitation, *Hydrol.*
13 *Earth Syst. Sci.* 20, 2195–2205, 2016.

14 Istanbuloglu, E., Wang, T. Wright, O. M. and Lenters, J. D.: Interpretation of hydrologic
15 trends from a water balance perspective: The role of groundwater storage in the Budyko
16 hypothesis, *Water Resour. Res.*, 48, W00H16, 2012. doi:10.1029/2010WR010100.

17 Jaramillo, F. and Destouni, G.: Developing water change spectra and distinguishing change
18 drivers worldwide. *Geophys. Res. Lett.*, 41, 8377–8386, 2014.

19 Jaramillo, F., and Destouni, G.: Local flow regulation and irrigation raise global human water
20 consumption and footprint. *Science* 350, 1248–1251, 2015.

21 Koster, R. D. and Suarez, M. J.: A Simple framework for examining the interannual
22 variability of land surface moisture fluxes. *Journal of Climate*, 12, 1911–1917, 1999.

23 Li, W. and Sankarasubramanian, A.: Reducing hydrologic model uncertainty in monthly
24 streamflow predictions using multimodel combination, *Water Resour. Res.*, 48, W12516,
25 2012. doi:10.1029/2011WR011380.

26 Lv, J., Wang, X.-S., Zhou, Y., Qian, K., Wan, L., Derek, E., and Tao Z.: Groundwater-
27 dependent distribution of vegetation in Hailu River catchment, a semi-arid region in China,
28 *Ecohydrology*, 2013, 6: 142-149, 2013.

- 1 Mathevet, T., Michel, C., Andreassian, V. and Perrin C.: A bounded version of the Nash-
2 Sutcliffe criterion for better model assessment on large sets of basins, In: Large Sample Basin
3 Experiments for Hydrological Model Parameterization: Results of the Model Parameter
4 Experiment–MOPEX, IAHS Publ. 307, p.211-218, 2006.
- 5 Martinez, G. F., and Gupta, H. V.: Toward improved identification of hydrological models: A
6 diagnostic evaluation of the “abcd” monthly water balance model for the conterminous
7 United States, Water Resour. Res., 46, W08507, 2010. doi:10.1029/2009WR008294
- 8 Mezentsev, V. S.: More on the calculation of average total evaporation, Meteorol. Gidrol., 5,
9 24–26, 1955.
- 10 Middleton, N. and Thomas, D. S. G. (Eds.): World atlas of desertification, United Nations
11 Environment Programme, Edward Arnold, 1992.
- 12 Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A
13 discussion of principles, J. Hydrol., 10 (3), 282–290, 1970.
- 14 Porporato, A., Daly, E. and Rodriguez-Iturbe, I.: Soil water balance and ecosystem response
15 to climate change, American Naturalist, 164, 625-632, 2004.
- 16 Potter, N. J., and Zhang, L.: Interannual variability of catchment water balance in Australia, J.
17 Hydrol., 369, 120–129, 2009.
- 18 Renner, M., Seppelt, R. and Bernhofer, C.: Evaluation of water-energy balance frameworks to
19 predict the sensitivity of streamflow to climate change, Hydrol. Earth Syst. Sc., 16, 1419-
20 1433, 2012.
- 21 Roderick, M. L., and Farquhar, G. D.: A simple framework for relating variations in runoff to
22 variations in climatic conditions and catchment properties, Water Resour. Res., 47, W00G07,
23 2011. doi: 10.1029/2010WR009826.
- 24 Sankarasubramanian, A., and Vogel, R. M.: Annual hydroclimatology of the United States,
25 Water Resour. Res., 38, 1083, 2002. doi:10.1029/2001WR000619.
- 26 Shi, C., Niu, K. Chen, T. and Zhou, X.: The study of pan coefficients of evaporation pans,
27 Scientia Geographica Sinica, 6(4), 305-313, 1986. (in Chinese)

1 Sloto, R. A. and Crouse, M. Y.: HYSEP: a computer program for streamflow hydrograph
2 separation and analysis, US Geological Survey Water-Resources Investigations Report 96-
3 4040, 1996.

4 Thomas, H. A.: Improved methods for national water assessment, report, Contract WR
5 15249270, U.S. Water Resour. Council, Washington, D. C.,1981.

6 Vandewiele, G. L., Xu, C.-Y. and Ni-Lar-Win: Methodology and comparative study of
7 monthly water balance models in Belgium, China and Burma, *J. Hydrol.*, 134, 315–347, 1992.

8 Wang, D.: Evaluating interannual water storage changes at watersheds in Illinois based on
9 long-term soil moisture and groundwater level data, *Water Resour. Res.*, 48, W03502, 2012.
10 doi:10.1029/2011WR010759.

11 Wang, D., and Tang, Y.: A one-parameter Budyko model for water balance captures emergent
12 behavior in darwinian hydrologic models, *Geophys. Res. Lett.*, 41, 4569–4577, 2014.
13 doi:10.1002/2014GL060509.

14 Wang, T., Istanbuloglu, E., Lenters, J. and Scott D.: On the role of groundwater and soil
15 texture in the regional water balance: An investigation of the Nebraska Sand Hills, USA,
16 *Water Resour. Res.*, 45, W10413, 2009. doi:10.1029/2009WR007733.

17 Yang, D., Shao, W., Yeh, P. J. F., Yang, H., Kanae, S., and Oki, T.: Impact of vegetation
18 coverage on regional water balance in the nonhumid regions of China, *Water Resour. Res.*, 45,
19 W00a14, 2009. doi: 10.1029/2008wr006948.

20 Yang, D., Sun, F., Liu, Z., Cong, Z. and Lei Z.: Interpreting the complementary relationship
21 in non-humid environments based on the Budyko and Penman hypotheses, *Geophys. Res.*
22 *Lett.*, 33(18), L18402, 2006. doi: 10.1029/2006gl027657.

23 Yang, D., Sun, F., Liu, Z., Cong, Z., Ni, G., and Lei Z.: Analyzing spatial and temporal
24 variability of annual water-energy balance in nonhumid regions of China using the Budyko
25 hypothesis, *Water Resour. Res.*, 43(4), W04426, 2007. doi: 10.1029/2006wr005224.

26 Yang, H., Yang, D., Lei, Z., and Sun, F.: New analytical derivation of the mean annual water-
27 energy balance equation, *Water Resour. Res.*, 44, W03410, 2008. doi:
28 10.1029/2007WR006135.

- 1 Yang, Z., Zhou, Y., Wenninger, J. and Uhlenbrook, S.: The causes of flow regime shifts in the
2 semi-arid Hailu River, Northwest China, *Hydrol. Earth Syst. Sc.*, 16, 87-103, 2012.
- 3 Yeh, P. J.-F., and Famiglietti J. S.: Regional Groundwater Evapotranspiration in Illinois.
4 *Journal of Hydrometeorology*, 10, 464-478, 2009.
- 5 Yin, L., Zhou, Y., Huang, J., Wenninger, J., Zhang, E., Hou, G., and Dong, J.: Interaction
6 between groundwater and trees in an arid site: Potential impacts of climate variation and
7 groundwater abstraction on trees, *J. Hydrol.*, 528, 435–448, 2015.
- 8 York, J.P., Person, M., Gutowski, W.J. and Winter, T.C.: Putting aquifers into atmospheric
9 simulation models: An example from the Mill Creek Watershed, Northeastern Kansas,
10 *Advances in Water Resources*, 25, 221–238, 2002.
- 11 Zhang, L., Dawes, W. R. and Walker G. R.: Response of mean annual evapotranspiration to
12 vegetation changes at catchment scale, *Water Resources Research*, 37(3), 701-708, 2001. doi:
13 10.1029/2000wr900325.
- 14 Zhang, L., Hickel, K., Dawes, W. R., Chiew, F. H. S., Western, A. W. and Briggs P. R.: A
15 rational function approach for estimating mean annual evapotranspiration, *Water Resources*
16 *Research*, 40, W02502, 2004. doi: 10.1029/2003WR002710.
- 17 Zhang, L., Potter, N., Hickel, K., Zhang, Y. and Shao, Q.: Water balance modeling over
18 variable time scales based on the Budyko framework - Model development and testing,
19 *Journal of Hydrology*, 360(1-4), 117-131, 2008. doi: 10.1016/j.jhydrol.2008.07.021.
- 20 Zhou, Y., Wenninger, J., Yang, Z., Yin, L., Huang, J., Hou, L., Wang, X., Zhang, D., and
21 Uhlenbrook, S.: Groundwater–surface water interactions, vegetation dependencies and
22 implications for water resources management in the semi-arid Hailu River catchment,
23 China – a synthesis, *Hydrology and Earth System Sciences*, 17, 2435–2447, 2013.
- 24 Zhou, Y., Yang, Z., Zhang, D., Jin, X., and Zhang, J.: Inter-catchment comparison of flow
25 regime between the Hailu and Huangfuchuan rivers in the semi-arid Erdos Plateau,
26 Northwest China, *Hydrological Sciences Journal*, 60(4), 688-705, 2015.

1 Table 1. The characteristics of the study catchments shown in Fig. 1.

Catchments	Area (km ²)	Mean annual Flux [†]			
		<i>P</i> (mm)	<i>E</i> ₀ (mm)	<i>Q</i> (mm)	Baseflow Index
C1	2645	367	1245	37.7	0.88
C2	2415	386	1218	38.8	0.64
C3	3253	447	1162	127.2	0.72
C4	3065	381	1227	75.8	0.31
C5	1272	466	1146	81.8	0.13
C6	3246	412	1186	53.0	0.16

2 [†] According to the data in the 1957-1978 period.

3

4

1 Table 2. Mean annual fluxes in the Hailutu River **Catchment (C1)** in different periods

Periods	$P(\text{mm})$	$E_0(\text{mm})$	$Q(\text{mm})$	Number Of Diversion(reservoirs) [†]
1957-1966	387.0	1230.2	42.3	0(0)
1967-1987	337.0	1269.6	32.6	4(2)
1988-1997	329.9	1240.2	23.4	9(2)
1998-2010	352.8	1234.0	28.0	10(2)

2 [†] According to [Yang et al. \(2012\)](#).

3

4

1 Table 3. Best fitting parameters of the ‘natural’ models for the study catchments.

Catchments	a	b (mm)	c	d	g	k	α	$\alpha g G_{\text{mean}}^{\dagger}$	NSE
C1	0.97	33	0.97	0.05	0.010	0.017	0.21	0.14	0.51
C2	0.91	75	0.92	0.11	0.058	0.025	0.09	0.12	0.43
C3	0.91	41	0.90	0.18	0.017	0.017	0.18	0.15	0.71
C4	0.94	83	0.68	0.19	0.049	0.076	0.21	0.09	0.67
C5	0.97	155	0.67	0.10	0.070	0.214	0.27	0.14	0.81
C6	0.93	179	0.67	0.11	0.077	0.162	0.18	0.11	0.71

2 $^{\dagger}G_{\text{mean}}$ is the mean value of the effective groundwater storage in the calibration period

3

Figure Captions:

Figure 1. Geographic information of the study site: (a) location of the study area in north central China; (b) Distribution of meteorological stations in the Erdos Plateau (green points) and the study catchments numbered as C1-C6. (c) Characteristics of landscape in the catchment C1 (the Hailiutu River Catchment) according to Lv et al. (2013).

Figure 2. The monthly meteorological data (a) and streamflow-baseflow data (b) from 1957 to 2010 in the catchment C1.

Figure 3. The plots of the annual $(P-Q)/P$ data versus the aridity index in the study catchments for the 1957-1978 period: (a) the mean annual data points for the 6 catchments bounded by the two Budyko curves (dashed lines) according to Eq. (3) with $w=1.6$ and $w=2.5$; and (b)-(f) are the annual data points of the different catchments. C1-C6 are the numbers of the catchments shown in Fig. 1. The solid line is the original Budyko curve determined with Eq. (2). The dashed lines are the regression curves of the scatter data points with the slope values shown nearby.

Figure 4. Schematic representations of the ABCD model (a) and ABCD-GE model (b). W and V are the effective soil water storage and the effective storage in the transition vadose zone, respectively. G is the effective groundwater storage.

Figure 5. Scatter plot for the observation and simulation results of the monthly runoff and groundwater discharge in the study catchments in the calibration period.

Figure 6. Simulated results of the ‘natural’ ABCD-GE model in comparison with the observation data in the catchment C1 from 1957 to 2010, including: Monthly runoff (a),

groundwater discharge (b), annual runoff (c) and annual evapotranspiration (d). The actual evapotranspiration in (d) was estimated with Eq. (18).

Figure 7. Plots of the annual $F-\phi$ data in the study catchments in the standard Budyko space for the 1957-2010 period. The actual evapotranspiration is estimated using the 'natural' models. The solid blue line is the original Budyko curve determined with Eq. (2). The dashed red lines are the linear regression curves of the data points with the slope data shown nearby. The intersection point of the green lines denotes the mean annual data.

Figure 8. Plots of the $F-\phi$ data in the standard Budyko space using the E_1 data for the Zone-1 and the E_2 data for the Zone-2 that estimated with Eq. (19) in the catchments C1 (left) and C4 (right). The red curves are the Budyko curves determined with Eq. (3), which could approximately represent the variations of the E_1/P data. The dashed blue lines are the linear regression lines of the E_2/P data.

Figure 9. The typical $F-\phi$ curves for annual water balance in the standard Budyko space determined with Eq. (22) when $w=1.5$ (a) and $w=2.0$ (b). The solid and dashed curves are estimated using $gG_a=0.2$ and $gG_a=1.0$, respectively. Dots are the data points of the study catchments.

Figure 10. The annual water balance data in the modified Budyko space with the effective precipitation defined by Wang (2012). Dots are the data obtained for the catchments using the 'natural' models. The solid curves represent the standard Budyko curves determined with Eq.(3) using $E_0/(P-\Delta S)$ and $E/(P-\Delta S)$, respectively, instead of F and ϕ . Budyko curves of $w=1.9$ and $w=2.5$ are selected to bound the data points of the catchment C1 and applied for the comparison with the other catchments. The dashed lines approximately represent the limitations of the $E/(P-\Delta S)$ data.

1 Figure 11. Histogram of the $\Delta(E_{LD}/P)$ data determined with Eq. (24) and the $\Delta(CV_Q/CV_P)$ data
2 determined with the coefficients of intra-annual variation of the monthly runoff (CV_Q) and
3 rainfall (CV_P) for the different periods in the catchment C1. The numbers of diversions
4 (reservoirs) are shown on the top of the blocks according to Table 2.