

Dear editors and reviewers,

Thank you very much for taking your time to review this manuscript. We are grateful for the detailed comments and suggestions, and we believe that these comments and suggestions will considerably improve our paper. Please find our point-by-point responses below.

Comment 1:

The uncertainties are discussed in a rather detailed and informative way, and the authors also state that they only apply poems when the poets are contemporary – this is a very important and valuable information, what should be in my opinion emphasized also earlier (maybe already in the abstract?). Does this mean that the (contemporary) poets are known in all cases? And what cases are we talking about? Are these the poems where the 86 phenological data are coming from, or do the authors have a broader-scale overview, so that they are able to provide a general picture for a larger region than the study area (and if yes, based on how many poems/data)?

Response 1:

Thank you for the comment. First, we would like to give brief introductions to the Quan-Tang-Shi (the poetry of the Tang Dynasty) and Quan-Song-Shi (the Poetry of the Song Dynasty), which are common sources for the poems of the Tang and Song Dynasties. Both Quan-Tang-Shi and Quan-Song-Shi are poetry collections of the two dynasties. The former was compiled in the Qing Dynasty (around 1705 AD), and the latter was compiled in modern times (after 1986). The titles, poets, and the verses were often recorded (in some cases, the titles and poets may also be unrecorded), but other information such as the writing time and places were not recorded in the Quan-Tang-Shi and Quan-Song-Shi. Just as we mentioned in Line 155-169 of the revised manuscript, the accessibility of phenological records of poems is relatively lower than that of other documents. For a specific poem, as we mentioned in Line 420-431, we cannot make sure whether it contains phenological information and is used in reconstructions before we read through the lines and related background information. Therefore, a standard procedure for extracting phenological records from poems is necessary to minimize the uncertainties of the records and filter out the useless records efficiently.

Our overall goals of this study are to demonstrate the validity and reliability of phenological records from poems as a proxy of past climate changes and to provide a reference in both theory and method for the extraction and application of phenological records from poems in China. Although we have only talked about the poems of the Tang and Song dynasties and involved a case study of 86 phenological records for climate reconstruction, we believe that the methods of data extraction and processing are applicable for other areas and periods.

We have rewritten the abstract to make clear the objective of this study as well as explaining the source and function of the 86 phenological records. Please find the details in Line 24-33. To present a broader-scale overview, we have changed Section 2.2 into “The numbers, spatial distributions and accessibility of phenological records from poems”. In this section, we have introduced the numbers of poems of the Tang and Song Dynasties (Line 144-147), the spatial distributions of phenological records of the Tang and Song Dynasties (Line 150-154) and the accessibility of phenological records from poems (Line 155-169).

Comment 2:

Just a side remark on the uncertainties subchapters: some of these uncertainties could be explained shorter and more accurately, once the authors involve a (Chinese) medieval social, economic or environmental historian as co-author of their study.

The authors discuss an over 600-year period covering the early and high-medieval period. Providing basic socio-economic background on how and why these poems were written (with reference), and the basic environmental characteristics (differences compared to recent times) of the environment the poets lived in and described should be an essential part of the presentation and analysis. As the topic is particularly sensitive on source dating, reliability and contemporary social/environmental background, the active participation of a trained (Chinese) medievalist, who can give a short concise historical overview, would be in my opinion essential.

Response 2:

Thanks for the suggestion. Two authors of this manuscript, Xiuqi Fang and Junhu Dai, have expertise both in Chinese history and past environmental changes and are qualified to give a concise historical review. The Tang and Song Dynasties were two powerful and prosperous dynasties of imperial China. During these periods, society was relatively open, with Confucianism, Taoism and Buddhism coexisting. The developed economy made people more educated and could express their thoughts through literature. The status of literature, especially poetry, had also been elevated. The Imperial Examination System, a civil service examination system in imperial China for selecting candidates for the state bureaucracy, had gradually improved, and poetry was incorporated into the examination subjects during this period (Zhang 2015). People had the opportunities to gain attention and change their lives if they could write beautiful poems. In these contexts, as a literary genre, poetry reached its highest level during the Tang and Song Dynasties in ancient China. People in the Tang and Song Dynasties preferred to record their thoughts and daily lives in poems. The change in the Tang and Song Dynasties environment was controversial, spurring us to improve data resolution by extracting phenological records from poems for environmental reconstruction. We have also discussed this in Line 407-419.

We have added a paragraph to introduce the socio-economic background from the perspective of data sources. Please find the details in Line 71-79. We have also added another paragraph to introduce the environmental background from the perspective of climate reconstruction. Please find the details in Line 80-87.

Comment 3:

Even if it is clear that the authors would like to present the potentials of Chinese poems, and these potentials are valid for entire China, based on the information presented in line 305 and on their previous paper(s) in the subject, they have tested source potentials only in one area of one province. There is no problem with that but, please, do indicate this information at the beginning of the paper (i.e. you should have a “Study area” chapter at the beginning, which is a usual part of papers in CP), too. Because it is a rather important information that the authors do discuss this topic based on a database regarding entire China, but only one area within a province, and in fact you suggest that this might have relevance for the entire China. China is huge, and even in your study period there were long periods when China was not one empire, but an area divided to separate states. So, it would be also useful to discuss shortly

why you think that in this rather eventful period of China's history this source was written in the same way and out of the same reasons when historical background (and also the level of literacy) in faraway regions could be rather different. Again, a (Chinese) medievalist would be able to answer this latter question easily and adequately.

Response 3:

We are grateful for the suggestion and appreciate your profound understanding for China's history and related characteristics. As explained in Response 1, the *Quan-Tang-Shi* and *Quan-Song-Shi* were compiled according to the dynasties of the poets, and the spatial information was usually unknown in poems. In addition, the national border lines of the Tang and Song Dynasties had changed many times in history, but it did not affect the introduction of extracting and processing phenological records from poems. However, there were some features of the spatial distribution of phenological records from poems. The spatial distributions of phenological records were highly consistent with the ruling regions of the dynasties, and more developed areas had more records .

Furthermore, we do believe the methodologies put forward in this study are not only applicable to the Guanzhong Area, as you put forward kindly one area in one province, and the Tang Dynasty but also applicable to other periods and areas. The reasons are as follows. First, modern phenological studies have confirmed that the primary driving factor of phenophases in whole China is the climatic factors, especially the temperatures in most areas of temperate China (Chmielewski et al., 2001; Schwartz et al., 2006; Ge et al., 2015). Thus, the phenological records from poems can be used as indicators of climate changes for the places they obtained. Our work was to make the phenological records from poems meet the needs of quantitative reconstruction rather than change the expression of phenological phenomena. Second, historians all agreed that the feudal society in Chinese history had not fundamentally changed during different dynasties (Liu, 1981; Tian, 1982; Feng, 1994). The relatively stable feudal system is also the reason why the feudal society has lasted for more than 2000 years. Although historical China varied its border lines from dynasty to dynasty, its core social-economic closely aligned with the major agricultural area throughout history. This geographic and temporal overlap allows for continuous comparison across the Chinese core areas (Fang et al., 2019). Correspondingly, the essence of literature, especially poetry, has not changed, although different dynasties may have various characteristics of poetry such as the limitations on poetic forms, the number of words, the need for rhymes and sounds, etc.

We have made it clear in the Abstract (Line 29-33) that the methodologies proposed in this study were applicable for other periods and areas. The reasons were discussed in the Discussion section (Line 438-453). We have added the introduction of the spatial distributions of phenological records from poems in Section 2.2. Please find details in Line 150-154. In addition, in the case study, we have added a "Study area" Section to introduce the area of reconstruction. Please find details in Line 353-358.

Comment 4:

The authors present both biological and physical phenological information. The biological information consists of plant and animal related phenological data. At first, I really needed to search a lot to figure out how many and what (wild) animal-related phenological data the authors

actually used in the (case) study, and then I realised this was one bird type. It would be useful to state such information, because based on the main text (about source potentials of entire China and the entire study period) one expects several different types of animals. As for the plant-related phenological information, the authors mention different types: ornamental and cultivated plants.

Response 4:

Thank you for the comment. There is abundant phenological evidence of different types of animals in the poems of the Tang and Song Dynasties. However, when it comes to climate reconstruction, it is another story. On the one hand, as mentioned in Response 1, only when we read through all the poems of the Tang and Song Dynasties can we know how many types of animals there are. On the other hand, modern phenological observation in China focuses mainly on plants. Only a small number of early observational records refer to animal phenology. According to our historical and modern phenological data at hand, only the following animal phenology has the potential to participate in climate reconstruction: *Cuculus micropterus* (Indian Cuckoo), *Oriolus chinensis* (Black-naped oriole), *Hirundo rustica* (barn swallow) and *Mogannia conica* (a type of cicada).

There are two types of data in this study. One is used as examples to illustrate the characteristics of phenological records in poems and the handling methodology in the studies of past climate changes. The locations involved in this kind of poem (No.1-29 in Appendix D) are across China. Five of this kind of poem (No.1, 7, 13, 23, 24) described the phenology of 4 types of animals (cicada, geese, cuckoo and swallow). The other kind of phenological records are used as evidence to reconstruct the temperature anomalies in Guanzhong Area during 618-900 AD. The locations involved in this kind of poem (No. 12, 14, 21, 29, 30-93 in Appendix D) are distributed in the Guanzhong area. One of this kind of poem (No. 86) was related to the phenology of cicada. To eliminate potential misunderstandings, we have listed all 86 pieces of original records of the temperature reconstruction in Appendix A with Gregorian dates, sites, phenophases, and the translations of the original verses.

Comment 5:

What do you mean under “ornamental plant”? The only case where I saw any explanation was Table 1, where an example was added: “Plum blossoms begin to bloom in early winter”. But plum is a fruit tree and as such, it is part of the cultivated vegetation, and fruit production is usually part of the agriculture. Why is it considered separately? Similarly, “ornamental animal” comes at one point in the picture, but it is not clear what it means and why it is mentioned.

Response 5:

Thanks for the comment. The phrase “ornamental plant” was used to express the concept as opposed to agricultural plants. We have changed the phrase to “natural plant” in the revised manuscript.

The English word “plum” can refer to two different plants when translated into Chinese. One of them is “mei”(pronunciation in Chinese pinyin), which usually refers to *Chimonanthus praecox* or *Armeniaca mume*. They are the species we would like to express here. The other is “li”, which usually refers to *Prunus salicina*, the species you understood here. To eliminate potential misunderstandings, we have changed the word to “ume” for the meaning of “mei”. The word

“plum” has remained for the translation of “li”. Both “mei” and “li” in the Tang and Song Dynasties were natural plants because their phenophases were rarely affected by human activities at that time.

Comment 6:

I have some problems with the presentation of phenological information related to cultivated plants, as it seems the authors treat them as if they were similar modern cultivated plants. There is no any indication in the paper that early and high-medieval agriculture used rather different grain and other cultivated plant types/varieties (even plum or almond trees) than modern agriculture, not talking about the fact that medieval agriculture was on a totally different level than its modern equivalent. Although these differences usually have an effect on a temperature reconstruction, there is no any indication in the paper that the authors would have taken these differences into consideration. Again, the related knowledge of a Chinese historian expert would have basic importance. To some extent, the same is true for some of the physical indicators, particularly for the development of river ice (e.g. differences in streamflow due to river regulations, dams can strongly affect temperature-river ice relationship).

Response 6:

We are grateful for the comment and benefit by your expertise at this. We acknowledge the uncertainties caused by the difference in cultivated plant types and crop management. We would like to make some explanations. First, one of our data processing steps is “identifying the animals and plants to species level” (Section 3.2.2 in Line 297 to 315), which requires that the plants compared from modern observation and poems should be the same species. Second, modern phenological studies discuss a lot on the phylogenetic conservatism of phenology in response to climate changes, which has proved that phenological responses to climate changes are often shared among closely related species (Davies et al., 2013; Du et al., 2017; Davis et al., 2018). These studies indicate that even considering that evolution will lead to differences between historical and modern plants, the plants recorded in the poems and their corresponding modern observation plants still have similar responses to climate changes. Not to mention that a thousand years of time is too short for the evolutionary cycle of plants (Calderon, 1995; Donoghue, 2008). As for cultivated plants, many modern studies have proved that the phenophases of crops are mainly affected by climatic factors, especially temperature, compared with other factors such as agricultural management (Lobell et al., 2012; Tao et al., 2014; Liu et al., 2018). Third, though the traditional calibration procedure may make a contribution to this problem, it seems not suitable for our study. The common calibration procedure in climate reconstruction relies on the statistical calibration of climate proxy data against representative instrumental data based on data in a long period of overlap between the two datasets. However, the phenological data for producing phenological data series from poems are neither continuous nor from the same species. In addition, there is no overlapping period between phenological records from poems and observational data. For physical indicators, there is no record of ice phenology in our reconstruction of the Guanzhong Area (Appendix A).

We have added a short discussion about these uncertainties. Please find the details in Line 454-460.

Comment 7:

Moreover, it is not clear exactly what phenological phenomena the authors relate to what temperatures (i.e. what periods of the year), because the authors simply refer to Chinese Meteorological Administration, and do not give any further information. It would be useful to conclude shortly the information taken from these official records. I also have problem with using only 30 years (1961-1990) to identify the exact relationship between temperature (of what period?) and phenophase information. Phenology based temperature reconstruction studies usually consider 50-60 years, at least, to identify this relationship. I understand that it is not possible to have longer overlap in some cases, but at least in those cases when it is possible to extend this control period, it would be useful to do it, and try out whether a longer control period gives the same relationship as 30 years.

Response 7:

Thank you for the comment. Many studies have concluded that the starting dates of the phenological phases are highly correlated with the temperature of the previous 2-3 months (Ahas et al., 2000; Piao et al., 2006; Dai et al., 2013). The 86 records in our study belonged to 34 different phenological phases corresponding to temperature in different periods of the year (Appendix A). However, in order to obtain a relatively uniform and comparable series of reconstructed temperatures, the mean annual temperature anomaly was selected as the reconstruction index. The correlation coefficients between the phenological phases and annual mean temperature were shown in Appendix C. We agree with you totally that the longer the time frame, the more accurate the relationship between phenology and a certain climatic variable will be. The period of 1961-1990 was selected as the reference period to calculate the mean annual temperature. By changing the time series of mean annual temperatures and phenological phases to anomalies with respect to 1961-1990, more data than 30 years were used to identify the exact relationship between anomalies of temperatures and phenophases. Taking the beginning date of spring cultivation as an example (Appendix C), as mentioned in Appendix B, the beginning date of spring cultivation was defined based on meteorological data, the data of 62 years (1951-2013, the data of 2006 was missing) were used to develop the transfer functions between anomalies of the beginning dates of spring cultivation and mean annual temperatures. However, as we also mentioned in Appendix B (Line 771-772), the China Phenological Observation Network (CPON) began in 1963 and was stopped during 1997-2002. In addition, some phenophases may lack observations in specific years. Thus, some time series we used were less than 30 year but we have tried our best to use all available data.

We have rewritten the Data and Methods Section. The reason for the reconstructed period and the reference period we used were explained. Please find the details in Line 359-367.

Comment 8:

In the abstract, the authors refer to the abundance of the source (poems) and phenology information, but this abundance does not reflect on the applied database and the correlation statistics, where only 86 phenological data are available, covering only 38 years out of 300 years with any temperature-related information. Moreover, according to Appendix C, correlation statistics is based on a database where more than 2/3 of the phenological data types are calculated

with the number of observations under 30, and 1/3 is under 20 – thus, in most cases the number of observations in fact does not reach the value to have any statistical significance. Moreover, sometimes even with the low observation number, correlations are rather low. In these cases, it would be useful to provide more information on why the authors think these data have further potentials. While in line 303 the authors suggest that they have selected 86 phenological records for validation, in line 382 the number of records is 85. So, is it 85 or 86? Either 85 or 86, this sounds like a rather low number for a reconstruction. Especially if we consider the fact that the authors used a number of different phenological data. I find the temperature reconstruction methodology a bit problematic. Based on Appendix B, in the reconstruction the authors applied the simple method of linear regression. However, in case of non-continuous datasets, as it is clearly the case with poem-based phenological information, the method of linear regression is not really a good method to apply. Could you explain why you think linear regression is the most suitable method to apply in this particular case? In fact (as I mentioned before), I also do not particularly like the fact that the authors treat this rather mixed set of early medieval phenological data automatically similar to those of the late 20th century.

Response 8:

Thanks for your wise guidance at this point, challenging, but very constructive for the further modification of our manuscript. As mentioned in Section 2.2 “The numbers, spatial distributions and accessibility of phenological records from poems” (Line 141-169) and Response 1, the abundance of phenological evidence from poems is accompanied by relatively low data resolution for quantitative reconstruction. We have applied all modern and historical data at hand to the reconstruction. Besides, the reconstruction of the Guanzhong Area for the Tang Dynasty is just a case study to prove the validity of the quantitative reconstruction of past climate changes. The number of original records for quantitative reconstruction should be 86. Regarding the method of the linear regression, on the one hand, it is one of the common methods for reconstructions based on phenological data (Ge et al, 2003; Možný et al, 2012; Wetter et al., 2013). On the other hand, considering factors such as discontinuity of data, no overlapping period and limitation of modern data mentioned in Response 6 and 7, other methods such as process-based phenological model are either not applicable or would bring in more uncertainties. Although linear regression is not the most suitable reconstruction method but it is the only method we could use here. Our future work will focus on extracting more records from poems, and developing integration methods for different phenophases at different sites to explore the overall phenological change and climate change over larger regions. Regarding the comment about treating early medieval phenological data similar to those of the late 20th century, please find the explanations in Response 6.

Comment 9:

I have read several times the validation subchapter and the related Appendix parts, but I still do not fully understand how the authors were able to reconstruct annual temperature anomalies. Do I understand well that – based on Fig. 3a, the Validation subchapter and the Appendices – the authors reconstructed annual temperature anomalies of over 300 years in a study area, based on 85 or 86 phenological data (if I understood well, covering only 38 years)? How? This sounds far too little evidence for any temperature anomaly reconstruction. Such a temperature reconstruction would require that the database (near-)systematically cover the study period or at least a

significantly higher number of observations. So, here a bit more explanation would be needed why the authors think 38 years of data can adequately describe the weather anomalies of 300 years.

Response 9:

We are grateful for the comment pointed out doubts about the core outcome of our reconstruction. Per your guidance and suggestions, we have rewritten the part of the reconstruction and changed the title of chapter 4 into “Validation of the phenological records from poems for reconstructing the past climate changes: a case study of temperature reconstruction in the Guanzhong Area for specific years during 600-900 AD”. Furthermore, the chapter has been further divided into three subchapters named 4.1 study area, 4.2 data and methods and 4.3 results and the comparisons with other reconstructions, in order to make the fact clearer. The reconstruction was introduced, analyzed and compared from the perspective of specific years instead of the whole period of 600-900 AD. Please find the details in Line 346-405.

Comment 10:

In the Validation subchapter and in Fig. 3(b) the authors referred to another paper (Liu et al. 2016): this paper contains an annual temperature anomaly reconstruction for the period 600-902, in the Guanzhong Area – practically the same study area and period the current paper discusses. In Liu et al. 2016, the temperature reconstruction was based on 271 (phenological, weather and climate, and human response) data, from which 87 was phenological data. As we received little information on the exact 86 (or 85) phenological data the current study utilizes, the question arises whether or not there is an overlap of phenological data between the database of the current study and the phenology data part of the Liu et al. 2016 database. Especially, because the only phenological source quotation Liu et al. (2016) provides as an example is quoted from a poem. It is also not clear for me how and why this temperature reconstruction – or even the comparison with the Liu et al. 2016 paper – provides any validation for the utilisation potentials of poem-based phenological data. The authors used modern phenology-measured temperature relationship, applying it on early-medieval poembased phenological data, to reconstruct early medieval annual temperature anomalies. As for the validation, as described above, it is not clear whether or not the Liu et al. (2016) reconstruction is independent from the current reconstruction. If not, the Liu et al. 2016 reconstruction should be applied with caution. Second: while comparing the two reconstructions in Fig. 3, the authors suggest that “There were approximately simultaneous temperature fluctuations between the two reconstructions,...” –well, looking at the Figure, this “simultaneous fluctuations” are not so easily and obviously recognisable. A statistically significant correlation would be a stronger proof for simultaneous fluctuation, but the authors do not provide any information on that. Dear authors, please, give correlation data.

Response 10:

Thanks for the very kind and innovative suggestions. Although the study of Liu et al., 2016 was one of our previous works, it is independent of this study. In Liu et al., 2016, we obtained 87 phenological records (other records of weather, climate and human response were used to verify the results of temperature reconstruction) from diverse historical documents such as the Xin-Tang-Shu (New Book of Tang, the official history of the Tang Dynasty) to quantitatively reconstruct the winter half-year (from October to next April) temperatures in the Guanzhong Area

from 600 to 902 AD. Except for one piece of data from a poem (No. 14 in Appendix A), there is no overlap between the two databases. We believed that the reconstruction by Liu et al., 2016 is a good case for comparison with ours because of the same study area, the similar reconstruction period, the same data type from different sources, the similar data amounts (87 and 86), the same reconstruction index (We have obtained the original data from Liu et al., 2016 to reconstruct the mean annual temperature anomalies) and same transfer functions (For the same phenological evidence involved in both studies such as the first date of frost, they share the same transfer function). Thus, it proves the validation of phenological records from poems if the two studies have similar features in temperature variations. Meanwhile, the differences between the two reconstructions caused by the above factors could be eliminated. Regarding the modern phenology-measured temperature relationship on early-medieval poem-based phenological data, we have explained in Response 6. The statistically significant correlation between the two reconstructions was not applicable, because only a few reconstructed years of the two studies were overlapped.

We have reintroduced the reconstruction of Liu et al, 2016 and explained the reason why it was used as a comparison in Line 370-378. Moreover, we have also rewritten the comparison between our study and relevant reconstructions in Line 379-405.

Comment 11:

Accounting with so low data density and so many uncertainties, to me it seems somewhat surprising to state that annual temperatures were “0.43_C and 0.29_C higher during the study period (600-902 AD) than at present (1961-1990).” I doubt one can give such exact statements (without an estimation of uncertainties), when temperature related information is available only for 76 and 38 years out of 300 years. Based on these statements, I assume that the years for which information is not available were regarded as “average”. However, if there is no poem referring to any phenophasis dates for 2-3 (or more) years in a row, this does not mean there could be no negative or positive temperature anomalies or even extremes in these years. It means only that no poem dealt with this question. In this respect, it would be useful to know how many different authors these 86 phenological data come from.

Response 11:

Thank you for the comment and we are inspired by your questions. We have rewritten this part. The reconstructed temperature anomalies by phenological records from poems were treated as the temperature variations of specific years during the period of 600-900 AD. We also compared the occurrences of relatively cold and warm periods and the amplitudes of reconstructed temperatures with other relevant reconstructions. The uncertainties from transfer functions were shown in Figure 3 and Appendix C. Please find details in Line 391- 402. All the 86 pieces of original records of the temperature reconstruction have been listed in Appendix A and they belong to 69 poems from 39 poets.

Comment 12:

The authors do not compare their reconstruction to any other reconstructions from China. Is it because there are no other annually-resolved temperature reconstructions available in (Central-)China that cover the period 600-900? Because if there is at least one other, independent

reconstruction (documentary based or natural scientific), then it would be useful to compare (and correlate) the current reconstruction results to that reconstruction (or reconstructions, if more than one exists).

Response 12:

Thanks for this very constructive suggestion and we made it clearer in this modification. As discussed in Line 407-419, the reconstructions based on natural evidence either cannot cover the whole period, or they have relatively low temporal resolutions. It is also the reason why we try to improve the spatiotemporal resolution of proxy by extracting phenological records from poems. We have added two relevant reconstructions for comparison. One of them was winter half-year temperature anomalies at a 30-year resolution reconstructed from documentary evidence in the middle and lower reaches of the Yellow and Yangtze Rivers of China (Ge et al., 2003). The other was annual temperature anomalies reconstructed from tree rings in Asia (Ahmed et al., 2013). All the four reconstructions have been converted to temperature anomalies with respect to the mean climatology between 1961 and 1990 for comparison. Please find the details in Figure 3 and Line 393-405.

Comment 13:

And finally an addition: poems and songs are also applied in historical climatology in Europe, but it is not used independently for reconstruction, and poems very rarely contain phenological information (but it is not without an example).

Response 13:

Thank you for your confirmation of our main proxy data in this manuscript and the related comment. We really appreciate this. Although phenology and poems have been applied in historical climatology since Chu (1973), few studies have relied solely on phenological records or poetic content to reconstruct historical climate changes in China quantitatively. When we were finishing the work of Liu et al, 2016, we found that most of the phenological evidence in the traditional documents, such as the history books was non-organic. The idea of using poetry, which is the most popular literary form at that time, as a data source came to our minds. As mentioned in Line 120-140, poems in China contain abundant phenological evidence. However, as mentioned in Line 170-224 and Response 1, most of the essential information required for climate reconstruction such as the species, time and sites were hidden. That is also the reason why we try to provide a reference in both principle and methodology for the extraction and application of phenological records from poems.

References:

- Ahas R, Jaagus J, Aasa A. The phenological calendar of Estonia and its correlation with mean air temperature. *International Journal of Biometeorology*, 2000, 44(4): 159-166.
- Ahmed M, Anchukaitis K, Buckley B M et al. Continental-Scale Temperature Variability during the Past Two Millennia. *Nature Geoscience*, 2013, 6(5): 339-346.
- Calderon W O . Phylogenetic Patterns among Tropical Flowering Phenologies. *Journal of Ecology*,

- 1995, 83(6):937-948.
- Chmielewski F, Rotzer T. Response of tree phenology to climate change across Europe. *Agricultural and Forest Meteorology*, 2001, 108(2): 101-112.
- Chu, K. A preliminary study on the climatic fluctuations during the last 5,000 years in China, *Science China*, 1973,16: 226-256,.
- Dai J H, Wang H J, Ge Q S. Multiple phenological responses to climate change among 42 plant species in Xi'an, China. *International journal of biometeorology*, 2013, 57(5): 749-758.
- Davies T J, Wolkovich E M, Kraft N J B, et al. Phylogenetic conservatism in plant phenology. *Journal of Ecology*, 2013, 101(6): 1520-1530.
- Davis C C , Willis C G , Primack R B, et al. The importance of phylogeny to the study of phenological response to global climate change. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2010, 365: 3201-3213.
- Donoghue M J. A phylogenetic perspective on the distribution of plant diversity, *Proceedings of the National Academy of Sciences*, 2008, 105, 11549-11555.
- Du Y, Chen J, Willis C G, et al. Phylogenetic conservatism and trait correlates of spring phenological responses to climate change in northeast China. *Ecology & Evolution*, 2017, 7: 6747-6757.
- Fang X, Su Y, Wei Z et al. Social impacts of climate change in historical China, in: *Socio-Environmental Dynamics along the Historical Silk Road*, edited by: Yang L. E., Bork H.-R., Fang X., and Mischke S., Springer, Cham, Switzerland, 2019, 231-245.
- Feng E. *The evolution of Chinese social structure*, Henan People's Publishing House, Zhengzhou, China, 1994(in Chinese).
- Ge Q S, Zheng J Y, Fang X Q et al. Winter half-year temperature reconstruction for the middle and lower reaches of the Yellow River and Yangtze River, China, during the past 2000 years. *The Holocene*, 2003, 13(6): 933-940.
- Ge Q S, Wang H J, Dai J H. Phenological response to climate change in China: a meta-analysis. *Global change biology*, 2015, 21(1): 265-274.
- Liu C: *The reasons for the long-term continuation of Chinese feudal society*, *Historical Research*, 1981, 2, 15-28 (in Chinese).
- Liu Y, Chen Q, Ge Q et al. Spatiotemporal differentiation of changes in wheat phenology in China under climate change from 1981 to 2010. *Science China Earth Sciences*, 2018, 61: 1088-1097.
- Lobell D B, Sibley A, Ivan Ortiz-Monasterio J. Extreme heat effects on wheat senescence in India. *Nature Clim Change*, 2012, 2: 186-189.
- Možný M, Brázdil R, Dobrovolný P et al. Cereal harvest dates in the Czech Republic between 1501 and 2008 as a proxy for March-June temperature reconstruction. *Climatic Change*, 2012, 110(3-4):801-821.
- Piao S L, Fang J Y, Zhou L M et al. Variations in satellite-derived phenology in China's temperate vegetation. *Global Change Biology*, 2006, 12(4): 672-685.
- Schwartz M D, Ahas R, Aasa A. Onset of spring starting earlier across the Northern Hemisphere. *Global Change Biology*, 2006, 12(2): 343-351.
- Tao F, Zhang Z, Xiao D et al. Responses of wheat growth and yield to climate change in different climate zones of China, 1981-2009. *Agricultural and Forest Meteorology*, 2014, 189: 91-104.
- Tian J. A summary of the discussion on the reasons for the long-term continuation of Chinese

- feudal society, *Historical Research*, 1, 1982, 103-110 (in Chinese).
- Wetter O, Pfister C. An underestimated record breaking event – why summer 1540 was likely warmer than 2003. *Climate of the past*, 2013, 9, 41-56.
- Zhang, G. *History of Chinese Poetry*. Hebei Education Press, Shijiazhuang, China, 2015(in Chinese).