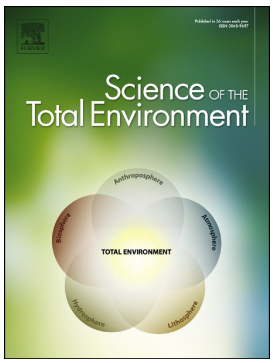


Causal relationship in the interaction between land cover change and underlying surface climate in the grassland ecosystems in China

Zhouyuan Li, Zezhong Wang, Xuehua Liu, Brian D. Fath, Xiaofei Liu, Yanjie Xu, Ronald Hutjes, Carolien Kroeze

PII: S0048-9697(18)32896-1
DOI: doi:10.1016/j.scitotenv.2018.07.401
Reference: STOTEN 27995

To appear in: *Science of the Total Environment*

Received date: 25 April 2018
Revised date: 6 July 2018
Accepted date: 29 July 2018

Please cite this article as: Zhouyuan Li, Zezhong Wang, Xuehua Liu, Brian D. Fath, Xiaofei Liu, Yanjie Xu, Ronald Hutjes, Carolien Kroeze , Causal relationship in the interaction between land cover change and underlying surface climate in the grassland ecosystems in China. Stoten (2018), doi:10.1016/j.scitotenv.2018.07.401

Causal relationship in the interaction between land cover change and underlying surface climate in the grassland ecosystems in China

Zhouyuan Li[1 2 3], Zezhong Wang[4], Xuehua Liu[1]*, Brian D. Fath[2 5]*, Xiaofei Liu[1], Yanjie Xu[6], Ronald Hutjes[3], Carolien Kroeze[3]

1. State Key Joint Laboratory of Environment Simulation and Pollution Control, and School of Environment, Tsinghua University, Beijing, 100084, People's Republic of China.
2. Department of Biological Sciences, Towson University, Towson, MD 21252 USA.
3. Water and Global Change Group, Wageningen University & Research, 6700 AA Wageningen, The Netherlands
4. Institute of Remote Sensing and Geographic Information System, Peking University, Beijing 100871, China.
5. Advanced Systems Analysis Program, International Institute for Applied Systems Analysis, Laxenburg, Austria.
6. Resource Ecology Group, Wageningen University & Research, 6708PB, Wageningen, The Netherlands

* Corresponding authors: xuehua-hjx@mail.tsinghua.edu.cn; bfath@towson.edu (They co-supervised the research and will handle the correspondence at the stage of post-publication. Dr. Xuehua Liu, as the host institution faculty, will handle the correspondence all stages of refereeing and publication

Abstract

Land-climate interactions are driven by causal relations that are difficult to ascertain given the complexity and high dimensionality of the systems. Many methods of statistical and mechanistic models exist to identify and quantify the causality in such highly-interacting systems. Recent advances in remote sensing development allowed people to investigate the land-climate interaction with spatially and temporally continuous data. In this study, we present a new approach to measure how climatic factors interact with each other under land cover change. The quantification method is based on the correlation analysis of the different order derivatives, with the canonical mathematical definitions developed from the theories of system dynamics and practices of the macroscopic observations. We examined the causal relationship between

the interacting variables on both spatial and temporal dimensions based on macroscopic observations of land cover change and surface climatic factors through a comparative study in the different grassland ecosystems of China. The results suggested that the interaction of land-climate could be used to explain the temporal lag effect in the comparison of the three grassland ecosystems. Significant spatial correlations between the vegetation and the climatic factors confirmed feedback mechanisms described in the theories of eco-climatology, while the uncertain temporal synchronicity reflects the causality among the key indicators. This has been rarely addressed before. Our research show that spatial correlations and the temporal synchronicity among key indicators of the land surface and climatic factors can be explained by a novel method of causality quantification using derivative analysis.

Keywords: remote sensing; grassland; land-climate; cause-effect; correlation analysis; eco-climatology

1 Introduction

In the field of climate change, elucidating the mechanism of regional climate, caused by biophysical processes at the land surface, is essential but still inadequate (Anderson et al. 2011). Observations on the coupled land–atmosphere system have generated considerable data, especially with remote sensing techniques for hydrothermal environmental mapping. However, in the process of data analysis and mechanism exploration, the analysis of causal relationships has been mostly focused on correlations, but no breakthroughs have emerged yet (Teuling et al. 2010; Zhao et al. 2014).

At the microclimate scale, individual plants transport water through transpiration and utilize gaseous water that is converted from liquid water. Surface evaporation is also used to transport gaseous water from the land surface to the atmosphere. Isotope labeling methods have been

used to corroborate the scale and rate of water circulation in different temporal and spatial dimensions (Hoffmann, Werner, and Heimann 1998; Xu et al. 2015). Environmental factors, such as precipitation, humidity, the underlying surface, and temperature, have direct impacts on the physiological processes of vegetation communities, including vegetation succession, and phenology. The micro-mechanism of interaction between these communities is well understood (Xu et al. 2015). However, at the ecosystem level, the cumulative function should be integrated as the effects at the community level are not equal to the sum of the effects at the individual level. Therefore, it is essential to assess, quantitatively, the interactive cause-effect relationship among these effects in a more holistic way.

In statistical models, causality analysis was driven mostly from correlation analysis. There are also many well-developed methodologies based on a combination of the relevant principles (Anderson et al. 2011)[1]. In economics, the Granger causality test was developed to judge the causation with a hypothesis on prediction capabilities (Granger 1969). These conventional statistical methods for causality assessment have been applied widely in economics and social sciences indicating patterns of relationships with the sampled data (Shipley 2000). Recently, as spatial analysis techniques are developing, multi-variate regression has been widely applied to geographic studies to test the association and causation (Mooij et al. 2016; Thomas and Beierkuhnlein 2013). For example, using mechanistic models, theoretical ecologists introduced a method with nonlinear state space reconstruction to distinguish causation from correlation. This shed light on a feasible way to detect the causality for non-separable, weakly-connected dynamic systems, i.e., a class of systems not previously covered by the Granger causality approach (Sugihara et al. 2012). Cause-effect demonstration and quantification is one of the key aims of system dynamics. The simulation of causality with feedback provides an analysis tool for the structure and behavior of the real systems (Li et al. 2015, 2017; Ford 2000). For ecological

applications, based on an input-output food web model, network analysis reveals direct and indirect causation of the interacting components in the ecosystems (Fath and Patten 1999). In the latest ecological studies, researchers tried to calculate the different order derivatives correlation to measure the relationship between bird migration and phenological dynamic (Wijk et al. 2012; Si et al. 2015). The methods applied in such studies are quite instructive as concordant with the thoughts in system dynamics constructing the model with the dependent relationship between the stock and flow variables, and similar to the higher-order pathway analysis in network analysis as well (Jørgensen and Fath 2011). Inspired by the derivative correlation analysis, in this study, we developed a method to measure the causation using the different order derivatives among the key factors of the land-climate interaction.

2 Methodology

2.1 Study areas and land cover change

In this study, the three typical grassland ecosystems of China were selected including a) the alpine grassland in Madoi in the Tibetan Plataea, b) the arid grassland in the middle of Inner Mongolia, and c) the temperate grassland in Hulubuir in North-eastern China (Figure 1). The study areas were comparable in the middle latitude zone, having a similar ecological function. The altitudes and the distance to the sea of the three areas were very different as with the various climate types. The heat-hydro patterns, the vegetation communities, and the distribution of river basins have obvious spatial heterogeneity, which provides exemplary cases to discover common regulators of land-atmosphere interactions under varied environmental gradients.

The alpine grassland is located at the headwaters of the Yellow River on the first-level Tibetan Plateau of the Chinese mainland terrain in Qinghai Province. From the 1980s to the 2010s, the alpine grasslands experienced ecological degradation and restoration. The regional climate has

experienced significant changes including rainfall, runoff, and temperatures (Li et al. 2013, 2015, 2017). The mid-section of the arid grassland, Inner Mongolia, which partially covers the Siziwang Banner, is located on the second-level of the Chinese mainland terrain, adjacent to the Xilingol grassland and Hohhot city. The arid grassland area is located on the agro-pastoral ecotone, where the land cover types is diverse and changeable. That area has rapidly urbanized during the past 20 years. The regional climate is drought, amplified by strong winds and adjacent to sand deserts making the grassland quality comparably lower than the other two grasslands. The temperate grasslands were selected in the Middle Eastern Region of the Hulunbuir grassland on the northeastern third-level step of the Chinese terrain. It mainly covers the most of Hulunbuir city, Old Barag Banner, and Evenk Autonomous Banner. The presence of a local ground water system provides abundant water and the grass is good; the regional climate is relatively warmer and wetter. Urbanization has been developing rapidly. The area is also on the agro-pastoral ecotone and is adjacent to the Greater Khingan Forest area. The Hulunbuir grassland is a world-class area of high-quality grassland.

The land cover change, representing the basic structural ecosystem dynamics, was mapped with satellite imagery (Beck et al. 2015). The remote sensing data consisted of five scenes of Landsat-5 TM/ 7 ETM images (30 m × 30 m resolution) of each study area acquired annually from 1988 to 2011 (Table S-1 of supplementary file). The land covers of the images were classified into five types by applying the method of maximum likelihood in ERDAS IMAGINE 2011. There were five categories of the land cover types in general, including 1) natural vegetation, 2) agricultural land, 3) bare land, 4) developed land, and 5) waters. The accuracy assessment of the land classification results was in Table S-2 of the supplementary file. The specific land cover types were varied in the different grassland ecosystems.

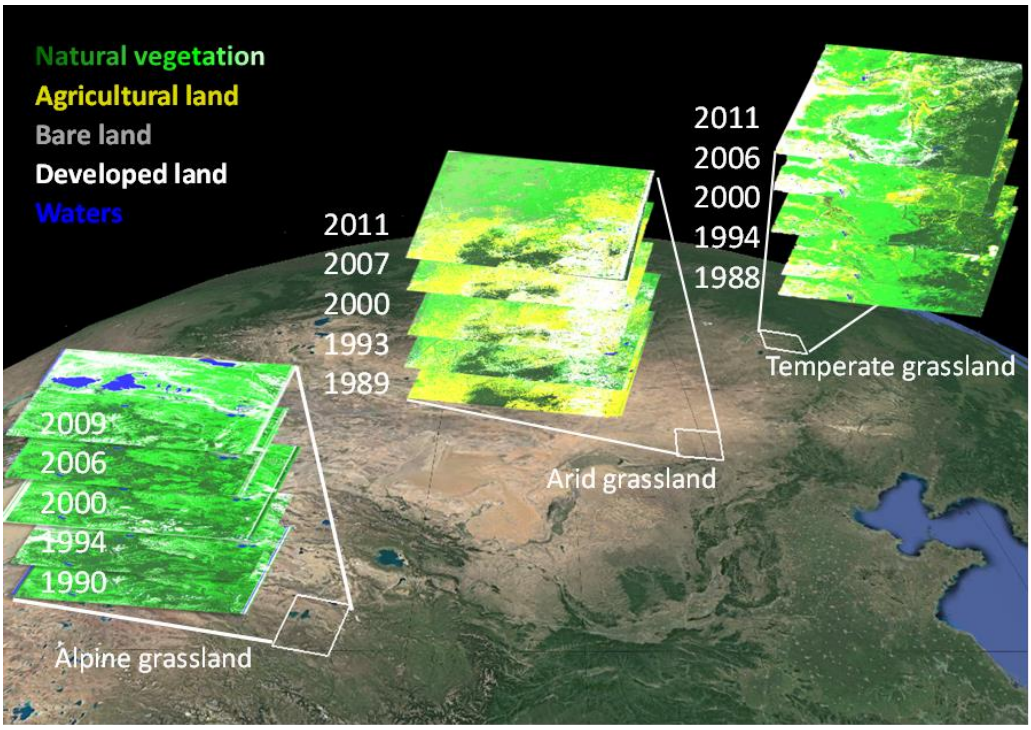


Fig. 1. The study area and the land cover change of the study areas in the alpine grassland, the arid grassland, and the temperate grassland.

2.2 Retrieval of the underlying surface climatic indicators

Satellite-derived vegetation index as the normalized difference vegetation index (NDVI) have been closely associated with primary vegetation production. NDVI is defined as follows:

$$NDVI = (r_{nir} - r_{red})/(r_{nir} + r_{red})$$

where r_{nir} and r_{red} represent surface reflectance levels averaged over wavelength ranges of infrared and visible infrared regions of the spectrum, respectively (Li et al. 2016).

Regional underlying surface climatic indicators regarding surface energy and water budget levels were examined using the same set of Landsat imagery. Surface temperature (T_s) was calculated as an indicator of surface energy conditions based on the following formula:

$$T_s = K_2/\ln(K_1\varepsilon/L_6+1)$$

where K_1=607.76×10^6 W/cm^2/sr/μm; K_2=1260.56×10^6 W/cm^2/sr/μm, as K_1 and K_2 are radiation constants for Landsat-5 images; L_6 is the spectral radiance of band 6 in Landsat-5 images; and ε is the atmospheric emissivity level determined based on the NDVI (Li et al. 2016, 2015).

We calculated the heat fluxes of the land surface to retrieve the ET (cm/day), the indicator for regional hydro-recycle, in the atmospheric correction (ATCOR2) module platform of the ERDAS 2011 remote sensing processing software. The ET retrieval was based on the following surface energy balance equations (Li et al. 2015):

$$R_n = H+G+\lambda ET$$

where the terms denote composites of net radiation (R_n, W/m^2), sensible heat flux (H, W/m^2), ground heat flux (G, W/m^2), and latent heat flux (λET, W/m^2). The modeling method involved two main tasks: 1) calculating the radiation balance based on remote sensing pixel reflectance levels and 2) calculating the heat balance using field knowledge that includes surface vegetation and meteorological conditions. Albedo, α, was calculated as the intermedia key variable in the ATCOR2 module based on the reflectance of the bands of the Landsat imagery.

The four indicators, NDVI, α, T_s, and ET, were calculated by the method above. The details of the calculation can be found in the previous study (Li et al. 2015) and its supplementary materials. To perform the statistical analysis, we extracted 800 random sampling points of the indicator values along with the information of the land types using ArcGIS. The causality measurement was based on the mean value curve of the different land types.

2.3 Definition and quantification on causalities

According to the principles and methods in systems dynamics, as well as the higher-order pathway analysis in network analysis, we conceptualized a coupled model representing the structure-function of land-climate interactions (Figure 2). In the horizontal planes, the structural dynamics showed the shifting of the mutually exclusive components inside the land or the

climate subsystem. On the spatial dimension, the variables of the land and the climate subsystems have the correlation when we overlapped the layers. In the vertical direction, the two subsystems are coupled and geared as a dynamic whole, which could be mathematically described using a cause-effect chain. Here, we defined the direct causality, indirect causality, and the semblance causality, with a canonical mathematical description, using the correlation analysis between the different order derivatives. The constructed causality index ***C***(X,Y) was used to measure the direct causation between the key indicators in the dynamics of the coupled systems.

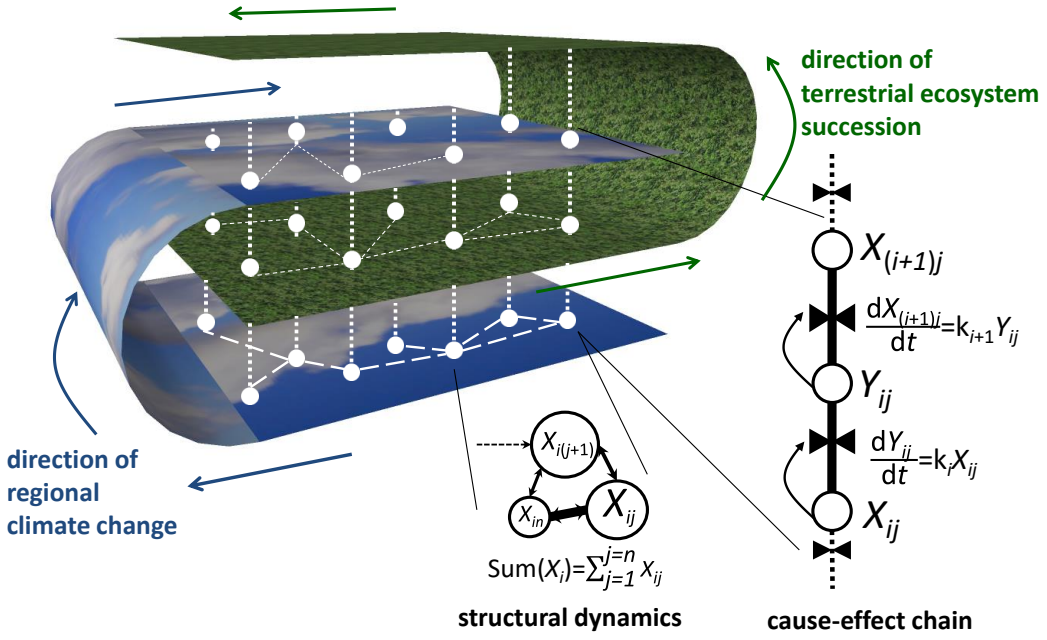


Fig. 2. Conceptual diagram structure-function interactive mechanism between land cover succession and underlying surface climate.

Within the framework of system dynamics (Jørgensen and Fath 2011) and based on the previous empirical practices applying the different order derivatives to detect the causal relationship, we made the canonical mathematical definition for this method (Wijk et al. 2012; Si et al. 2015).

Firstly, we defined a kind of direct causality. If the changing rate of *Y* has linear correlation with *X,* that is,

$$dY/dt = kX+c$$

where *k* and *c* are constants, then it is defined that *X* is the causation for *Y*. In a more general statement, if $Y^{(n)}$ has linear correlation with $X^{(n-1)}$, where n means the order of derivative for the variables, then *X* is defined as the causality to *Y,* noted as $\mathbf{D}_{X\rightarrow Y}$.

Secondly, similarly to above, we define a kind of indirect causality (I).If the acceleration or higher order derivatives of *Y* has linear correlation with *X*, such as,

$$d^2Y/dt^2 = kX+c$$

more generally, if $Y^{(n)}$ has linear correlation with $X^{(n-2)}$, then *X* is defined as the indirect causality to *Y*, noted as $\mathbf{I}_{X\rightarrow Y}$.

Particularly, if there is a linear correlation between two sets of data of *X* and *Y* of the same order derivatives, then it will be defined as an association between *X* and *Y* with no direction of dependence, noted as $\mathbf{A}_{XY}$.

Based on the definition above, we used the Pearson regression coefficient to measure the linear relation between the variables, *X* and *Y*, and their different order derivatives, $X^{(n)}$ and $Y^{(n)}$. The absolute values of the Pearson regression coefficient of the two variables which had the difference of a one-order derivative, e.g., $Y^{(n)}$ and $X^{(n-1)}$, were used to quantify the direct causation from −2 to 2; the absolute values of the Pearson regression coefficient of the two variables which had the difference of second-order derivatives, like $Y^{(n)}$ and $X^{(n-2)}$, were used to quantify the indirect causation from −1 to 1. The absolute values of the Pearson regression coefficient of the original functions and the functions of the same-order derivative, *X* and *Y*, or $X^{(n)}$ and $Y^{(n)}$, were used to quantify the association between X and Y, as a measurement space from −3 to 3. We formulated a comprehensive indicator ***C***(X, Y) using the residue of the direct

causation minus indirect causation and divided by the association to assess the overall direct causality between the two variables. When $\boldsymbol{C}$(X, Y) is below zero, it means the indirect causality is over the direct causality. According to the statistical distribution of $\boldsymbol{C}$(X, Y), the value over 3.00 was regarded as the strong direct causal contribution.

$$\boldsymbol{C}(X,Y) = \frac{\sum_{i=0,j=i+1}^{1}\left|P_{ij}\left(X_i,Y_j\right)\right| - \sum_{i=0,j=i+2}^{0}\left|P_{ij}\left(X_i,Y_j\right)\right|}{\sum_{i=0,j=i}^{2}\left|P_{ij}\left(X_i,Y_j\right)\right|}$$

Where, $P_{ij}(X_i,Y_j)$— the Pearson regression coefficient between X_i and Y_j; X_i,Y_j —$X_i=v^{(i)}(t)$ or $Y_j=v^{(i)}(t)$, i=0, 1, 2; $X_i=\alpha^{(i)}(t)$ or $Y_j=\alpha^{(i)}(t)$, i=0, 1, 2; $X_i=\tau^{(i)}(t)$ or $Y_j=\tau^{(i)}(t)$, i=0, 1, 2; $X_i=\varepsilon^{(i)}(t)$ or $Y_j=\varepsilon^{(i)}(t)$, i=0, 1, 2; where, $v^{(i)}$(t), $\alpha^{(i)}(t)$, $\tau^{(i)}(t)$ and $\varepsilon^{(i)}(t)$, represent the fitted original functions (i=0) or the derivative functions of i order by time (t) of the four key indicators, NDVI, α, T_s , and ET.

Results

Land cover in the three grasslands shows structural changes over the past two decades (Fig. 3). It is evident that the alpine grassland was restored in 1994 and the recovery accelerated after 2000. For the arid grassland, the high coverage grassland and low coverage grassland shifted up and down alternatively under the impact of annual precipitation, which reflected adaptivity of the low coverage vegetation succession in the drought years. The original quality of the natural vegetation in the temperate grassland was better due to its advantageous location and hydrological endowment; however, since 2000, the ecological degradation in that area became more severe. The proportion of grassland and open water decreased, while the bare and sandy land increased rapidly. The land cover reflected the different phases in the three grassland ecosystems.

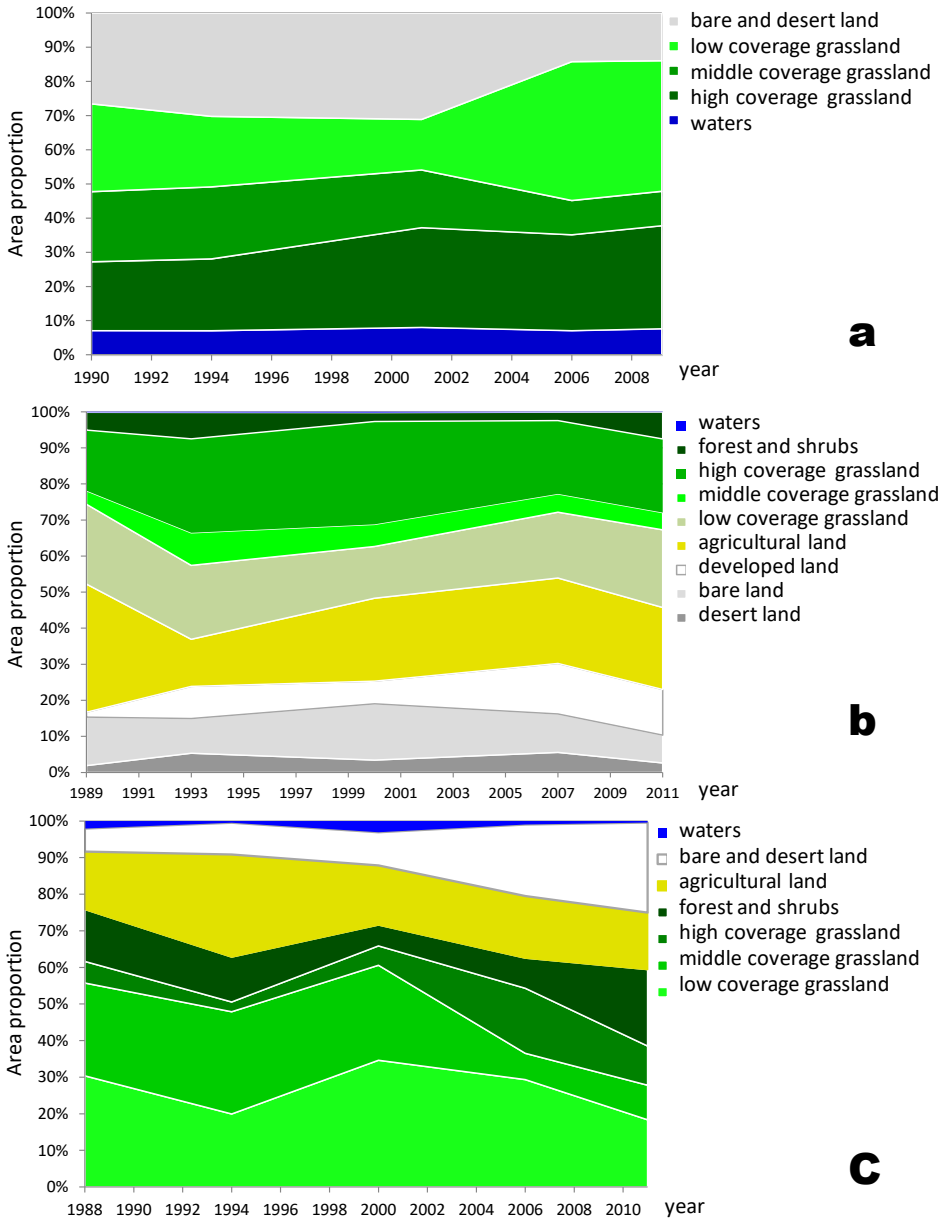


Fig. 3. The statistic results of the land cover change of the three grasslands, alpine grassland (a), arid grassland (b), and temperate grassland (c).

We selected and calculated the key indicators of the regional climate, including NDVI, α, T_s, and ET. They represented the vegetation, the underlying surface physical features of the energy balance, the heat environment, and the intermedia link of the hydrological cycle, respectively. The correlation analysis based on the spatially extracted data shows a concordance in the four

pairs of the key indicators, with negative correlation between NDVI-α, positive correlation between α-T_s, negative correlation between T_s-ET, and positive correlation between ET-NDVI (Fig. 4). Almost all correlations are statistically significant (p<0.05) (Table S-3 in the supplementary shows the detailed numbers). From Fig. 4, in the planes of the pair variables, the dots with different color represented the different land cover types. They distributed differently by the effects of the land covers and composed together to form the correlation as a whole. The general pattern is comet-like. The shrubs and trees, and the high-coverage grassland concentrated in the head, while the bare land scattered in the tail part, the low-coverage grassland and the agricultural land being in the middle. These reflected how closely the land and climate coupled, as the vegetation matters most. This result among these key steps in the three areas over the past decades proved the universality of the feedback loop.

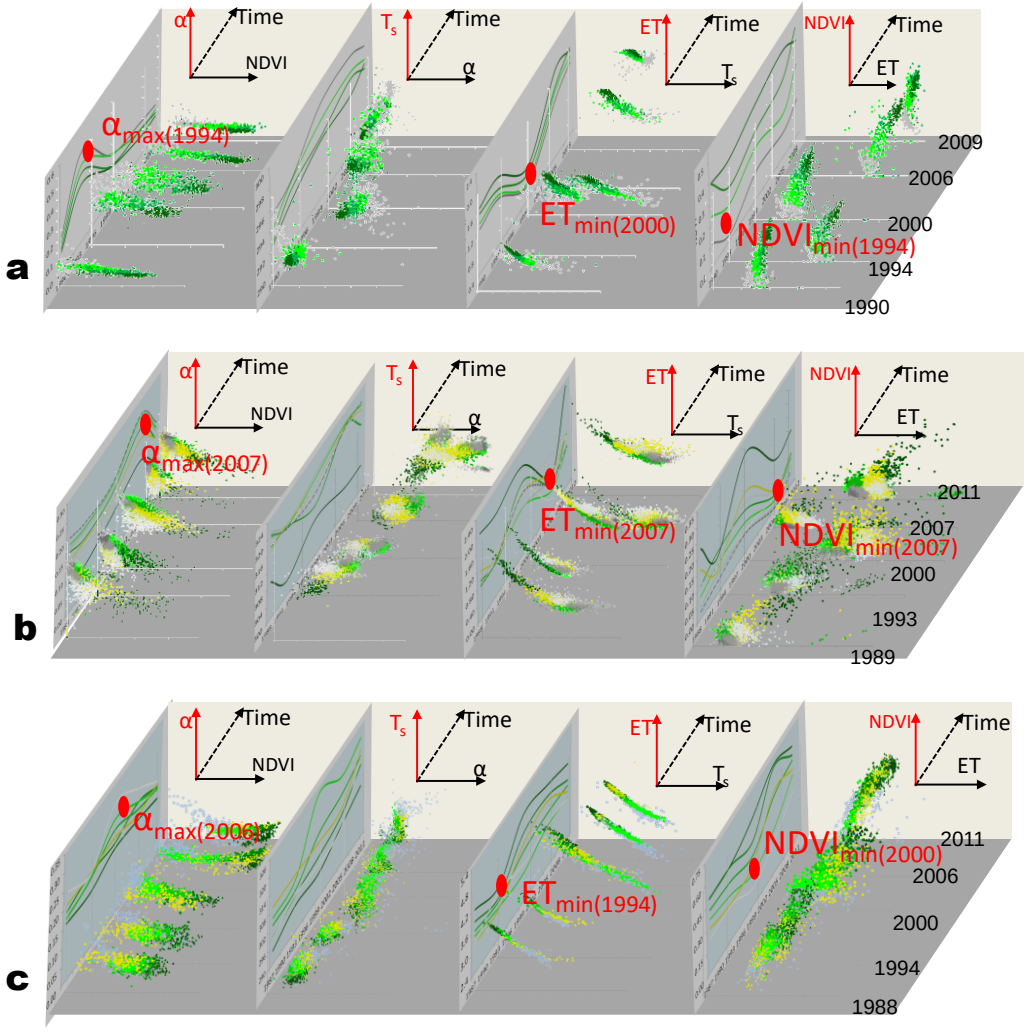


Fig. 4. Spatial correlation and temporal synchronicity among the four key indicators, NDVI, albedo (α), surface temperature (T_s), and evapotranspiration (ET), underlying surface climate subsystems by the different land cover types in the ecosystems of alpine grassland (a), arid grassland (b) and temperate grassland (c). (The color scheme of the dots is the same as the Fig. 3, meaning the different land cover types)

We measured the trends of changes in the key indicators with their annual means (Fig. 4). For the alpine grassland, NDVI started to rise from the bottom as vegetation restored since 1994, and α declined from its peak in 1994 as the roughness increased with expanding vegetated cover. The condition of the hydrological cycle improved after 2000, indicated by increasing ET,

which demonstrates a lag effect between the land surface structure and its eco-hydrological function. For the arid grassland, the turning points of the indicators for the most land covers were in 2007, shown as the synchronicity. It was recorded by meteorological observation that the arid grassland experienced a severe drought around 2007. For the temperate grassland, the minimum ET appeared in 1994, then NDVI reached its lowest point and turned upwards in 2000. Later in 2006, the maximum α appeared. Through an overall comparison, we found that the key indicators of the underlying surface climate system have a strong correlation on the spatial dimension while it has an uncertain synchronicity on the temporal dimension. This prompted further investigation to measure the causality between these factors to explore the interaction processes.

Applying the above method, the causality of the 4 key indicators were quantitatively assessed. In the three grassland ecosystems, all of them have strong direct causation with the value of ***C***(X, Y) larger than 3.00 (Fig. 5), mostly in the feedback of NDVI-ET in the vegetated underlying surface. The negative value of ***C***(X,Y) means the indirect causation was stronger than the direct one (shown with the dash-line linkages in Fig. 5). For all the three grassland ecosystems, the causality quantification results demonstrated that the vegetated underlying surface engaged more in land-atmosphere coupling than the un-vegetated surface, especially for the hydrological cycle indicated by ET. Because vegetation is the dominant land cover in the grassland ecosystems, and the biophysical processes like photosynthesis and transpiration of vegetation are directly involved in the surface water circulation and energy, the direct contribution in the causality of the vegetated land to the regional heat-hydro environment and the coupling relationship is higher than the un-vegetated land.

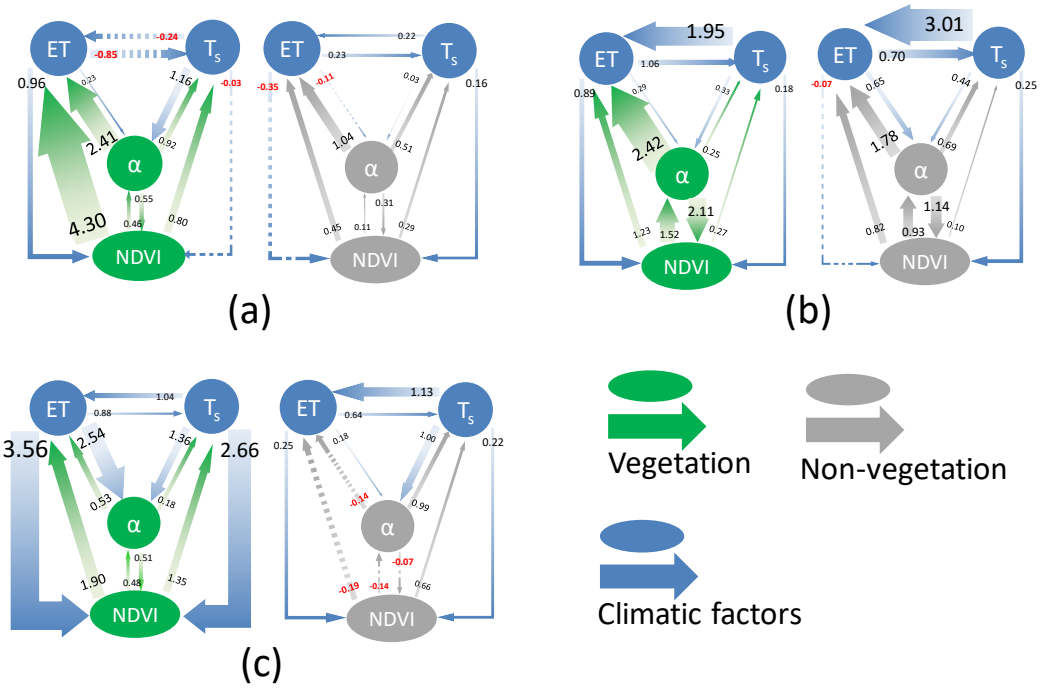


Fig. 5. Causality quantification results in the interaction between the land cover and the four key indicators, NDVI, α, T_s, and ET of the underlying surface climate subsystems in the ecosystems of alpine grassland (a), arid grassland (b) and temperate grassland (c).

It was found that the total direct causality of the arid grassland was higher than the other two grassland ecosystems. Through further comparison, we knew that the feedback loop within the internal sub-systems of land and underlying surface climate, represented by NDVI–α and T_s–ET, were significantly higher in the arid grassland than those in the other two grasslands. When we compared the coupling feedbacks between those sub-systems, NDVI–α-T_s and NDVI–α-ET, to the internal feedbacks within the sub-systems, NDVI–α and T_s–ET, it was found that the difference of these calculated causalities in the arid grassland is much smaller than the other two types of grassland.

Discussion and Conclusions

Novel methods were used to demonstrate patterns of the different coupling features in the various grassland ecosystems. Based on the coupling features identified, we categorized the arid grassland as one type while we took the alpine and the temperate ones as the other type. For the grassland with lower vegetation coverage and drought climate, like the arid grassland in our study, both the direct causality within the internal sub-systems and between the coupled subsystems are strong. Meanwhile, the alpine and temperate grassland ecosystems, with higher vegetation coverage and cooler and wetter climate, showed relatively loose causal connection within their internal sub-systems. In this type of grassland ecosystem, the indirect cause-effect was detected as comparatively higher. The difference in the quantified direct causality could explain why the climatic functional indicators of alpine and temperate grasslands have lag effect on the temporal dimension while those indicators have the synchronicity in arid grasslands. Since the causality measurement was based on the different-order derivative calculation, it could be formulated with a system dynamics approach, which could link to the mechanism of climatology (Mavrommati, Bithas, and Panayiotidis 2013; Byakatonda, Parida, and Kenabatho 2018). However, it should be clarified that such a linkage would be general and limited. The main model used in our study is a statistical approach aiming to measure and quantify the numerical relationship in the dynamic of vegetation-climate as in a holistic manner, ours is different from the biophysical models which are bottom-up calculations based on the physiological and physical mechanism.

We may develop these summarized findings with the principles of restoration ecology (Dobson, Bradshaw, and Baker 1997; Downing et al. 2012; Lamb, Erskine, and Parrotta 2005). In the arid grassland, the drought environment is not very suitable for plants habitat, as there are both relatively lower biomass per unit area and thinner vegetation indicated by the lowest NDVI

(0.091±0.066) among the three cases. The artificial land use and cover (agricultural, developed, and bare lands) takes a larger proportion (47.6%) than the natural ones, so the biodiversity and complexity in the underlying ecological structure was assumed as lower. Such ecosystems have a lower resistance and higher resilience. They would be more vulnerable to get an impact from the external inference of the climate or humans, but turn back to the original level sooner, such as what happened around 2007 in the arid grassland. In the alpine and temperate grasslands, the plant communities flourished under better hydrological condition. The vegetation is dense and has a higher biomass per unit area indicated by NDVI (0.279±0.135 for the alpine, and 0.401±0.148 for the temperate regions, respectively). The natural vegetation area proportion is larger (69.4% for the arid, and 66.5% for the temperate), so the biodiversity and complexity there was assumed as higher. These ecosystems have higher resistance and lower resilience, as having more indirect causation in the systems. The effects of external interferences on functional factors are delayed after a longer interval of accumulation, such as the lag effect that we found in the alpine and the temperate grasslands.

The generalized inference above is based on the logic of ecological cause an effect, leading to a new hypothesis about the relationship between the structural features and the stability of ecosystems. Additional scientific questions regarding the measurement approach need be specified in future research to test the ideas and to reveal more intrinsic principles regarding the coupled land-atmosphere system.

Acknowledgements This research was supported by State Key Joint Laboratory of Environment Simulation and Pollution Control, China (No.11Y02ESPCT). Z. L. acknowledges the support from '985' Future Scholarship of Tsinghua University, and the WIMEK Fellowship of Wageningen University and Research.

References

Anderson, Ray G, Josep G Canadell, James T Randerson, Robert B Jackson, Bruce A Hungate, Dennis D Baldocchi, George A Ban-Weiss, Gordon B Bonan, Ken Caldeira, and Long Cao. 2011. "Biophysical Considerations in Forestry for Climate Protection." *Frontiers in Ecology & the Environment* 9 (3): 174–82.

Beck, Inga, Ralf Ludwig, Monique Bernier, Esther Lévesque, and Julia Boike. 2015. "Assessing Permafrost Degradation and Land Cover Changes (1986–2009) Using Remote Sensing Data over Umiujaq, Sub-Arctic Québec." *Permafrost & Periglacial Processes* 26 (2): 129–41.

Byakatonda, Jimmy, B P Parida, and Piet K Kenabatho. 2018. "Relating the Dynamics of Climatological and Hydrological Droughts in Semiarid Botswana." *Physics & Chemistry of the Earth Parts A/b/c*.

Dobson, Andy P, A D Bradshaw, and A J M Baker. 1997. "Hopes for the Future: Restoration Ecology and Conservation Biology." *Science* 277 (5325): 515–22.

Downing, A S, E H van Nes, W M Mooij, and M Scheffer. 2012. "The Resilience and Resistance of an Ecosystem to a Collapse of Diversity." *Plos One* 7 (9): e46135.

Fath, Brian D, and Bernard C Patten. 1999. "Review of the Foundations of Network Environ Analysis." *Ecosystems* 2 (2): 167–79.

Ford, Andrew. 2000. *Modeling the Environment. An Introduction to System Dynamic Modeling of Environmental Systems*. University of Chicago Press,.

Granger, C W J. 1969. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." *Econometrica* 37 (3). [Wiley, Econometric Society]: 424–38. https://doi.org/10.2307/1912791.

Hoffmann, G, M Werner, and M Heimann. 1998. "Water Isotope Module of the ECHAM Atmospheric General Circulation Model: A Study on Timescales from Days to Several Years." *Journal of Geophysical Research Atmospheres* 103 (D14): 16871–96.

Jørgensen, Sven Erik, and Brian David Fath. 2011. *Fundamentals of Ecological Modelling: Applications in Environmental Management and Research*. 4th ed. Elsevier.

Lamb, D, P D Erskine, and J A Parrotta. 2005. "Restoration of Degraded Tropical Forest Landscapes." *Science* 310 (5754): 1628–32.

Li, Z., X. Liu, T. Ma, D. Kejia, Q. Zhou, B. Yao, and T. Niu. 2013. "Retrieval of the Surface Evapotranspiration Patterns in the Alpine Grassland-Wetland Ecosystem Applying SEBAL Model in the Source Region of the Yellow River, China." *Ecological Modelling* 270. https://doi.org/10.1016/j.ecolmodel.2013.09.004.

Li, Z., W. Wu, X. Liu, B.D. Fath, H. Sun, X. Liu, X. Xiao, and J. Cao. 2017. "Land Use/cover Change and Regional Climate Change in an Arid Grassland Ecosystem of Inner Mongolia, China." *Ecological Modelling* 353. https://doi.org/10.1016/j.ecolmodel.2016.07.019.

Li, Zhouyuan, X. Liu, T. Niu, D. Kejia, Q. Zhou, T. Ma, and Y. Gao. 2015. "Ecological Restoration and Its Effects on a Regional Climate: The Source Region of the Yellow River, China." *Environmental Science and Technology* 49 (10). https://doi.org/10.1021/es505985q.

Li, Zhouyuan, Wenzhao Wu, Xuehua Liu, Brian D Fath, Hailian Sun, Xinchao Liu, Xinru Xiao, and Jun Cao. 2016. "Land

Use/cover Change and Regional Climate Change in an Arid Grassland Ecosystem of Inner Mongolia, China." *Ecological Modelling*.

Mavrommati, G, K Bithas, and P Panayiotidis. 2013. "Operationalizing Sustainability in Urban Coastal Systems: A System Dynamics Analysis." *Water Research* 47 (20): 7235–50.

Mooij, Joris M, Jonas Peters, Dominik Janzing, and Jakob Zscheischler. 2016. "Distinguishing Cause from Effect Using Observational Data: Methods and Benchmarks." *Journal of Machine Learning Research* 17 (1): 1103–1204.

Shipley, B. 2000. "Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations, and Causal Inference." *Quarterly Review of Biology* 82 (4): 646–49.

Si, Y, Q Xin, W F de Boer, P Gong, R C Ydenberg, and H H Prins. 2015. "Do Arctic Breeding Geese Track or Overtake a Green Wave during Spring Migration?" *Sci Rep* 5 (1): 8749.

Sugihara, G, R May, H Ye, C H Hsieh, E Deyle, M Fogarty, and S Munch. 2012. "Detecting Causality in Complex Ecosystems." *Science* 338 (6106): 496–500.

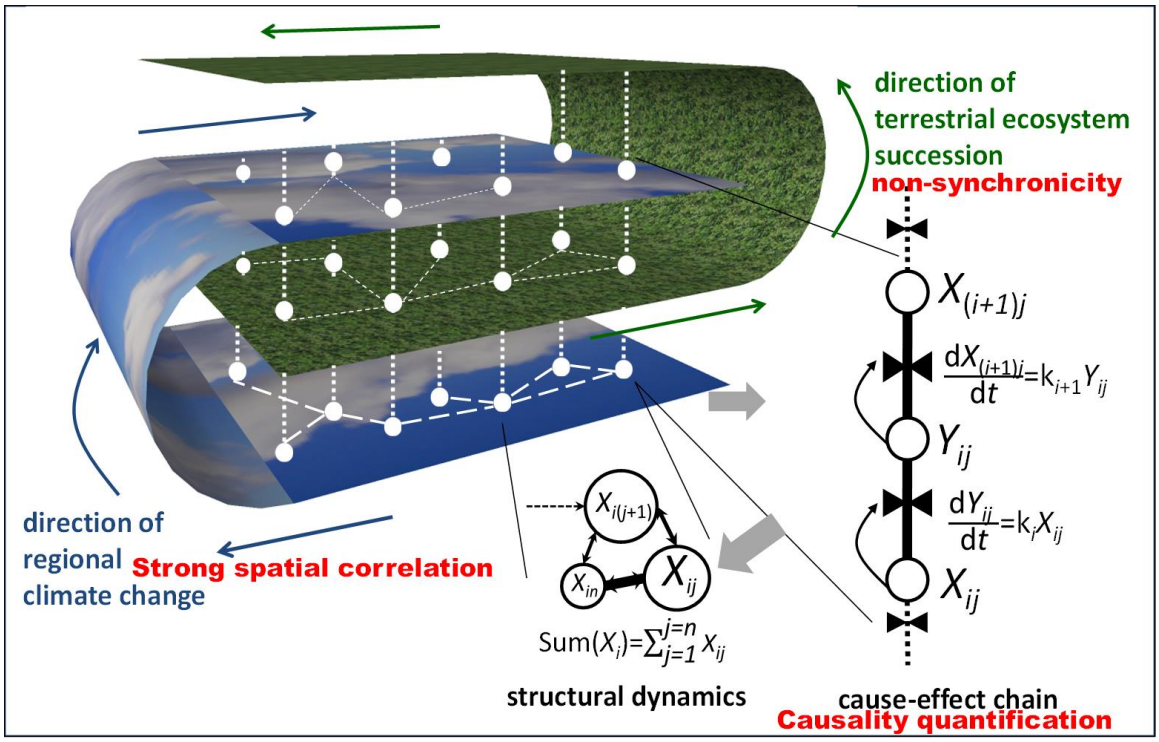
Teuling, Adriaan J, Sonia I Seneviratne, Reto Stockli, Markus Reichstein, Eddy Moors, Philippe Ciais, Sebastiaan Luyssaert, Bart Van Den Hurk, Christof Ammann, and Christian Bernhofer. 2010. *Contrasting Response of European Forest and Grassland Energy Exchange to Heatwaves*. Nature Publishing Group.

Thomas, Stephanie Margarete, and Carl Beierkuhnlein. 2013. "Predicting Ectotherm Disease Vector Spread—benefits from Multidisciplinary Approaches and Directions Forward." *Naturwissenschaften* 100 (5): 395–405.

Wijk, Rien E Van, Andrea Kölzsch, Helmut Kruckenberg, Barwolt S Ebbinge, Gerhard J D M Müskens, and Bart A Nolet. 2012. "Individually Tracked Geese Follow Peaks of Temperature Acceleration during Spring Migration." *Oikos* 121 (5): 655–664.

Xu, Chenxi, Nathsuda Pumijumnong, Takeshi Nakatsuka, Masaki Sano, and Zhen Li. 2015. "A Tree-Ring Cellulose δ 18 O-Based July–October Precipitation Reconstruction since AD 1828, Northwest Thailand." *Journal of Hydrology* 529: 433–41.

Zhao, L, X Lee, R B Smith, and K Oleson. 2014. "Strong Contributions of Local Background Climate to Urban Heat Islands." *Nature* 511 (7508): 216.



Graphical abstract

Causal relationship in the interaction between land cover change and underlying surface climate in the grassland ecosystems in China

Zhouyuan L, Zezhong Wang, Xuehua Liu, Brian D. Fath, Xiaofei Liu, Yanjie Xu, Ronald Hutjes, Carolien Kroeze

Highlights

1. We discovered that spatially highly correlated factors changed not simultaneously.
2. The lag effects was quantified to assess the causality contribution.
3. This study compared the land-climate dynamics among three different types of grassland.