



OPEN

Seasonal weather and climate prediction over area burned in grasslands of northeast China

Ali Hassan Shabbir^{1,2,3}, Jiquan Zhang^{1,2,3✉}, John W. Groninger⁴, Eddie J. B. van Etten⁵, Samuel Asumadu Sarkodie⁶, James A. Lutz⁷ & Carlos F. Valencia⁸

Grassland fire dynamics are subject to myriad climatic, biological, and anthropogenic drivers, thresholds, and feedbacks and therefore do not conform to assumptions of statistical stationarity. The presence of non-stationarity in time series data leads to ambiguous results that can misinform regional-level fire management strategies. This study employs non-stationarity in time series data among multiple variables and multiple intensities using dynamic simulations of autoregressive distributed lag models to elucidate key drivers of climate and ecological change on burned grasslands in Xilingol, China. We used unit root methods to select appropriate estimation methods for further analysis. Using the model estimations, we developed scenarios emulating the effects of instantaneous changes (i.e., shocks) of some significant variables on climate and ecological change. Changes in mean monthly wind speed and maximum temperature produce complex responses on area burned, directly, and through feedback relationships. Our framework addresses interactions among multiple drivers to explain fire and ecosystem responses in grasslands, and how these may be understood and prioritized in different empirical contexts needed to formulate effective fire management policies.

Human-populated grassland ecosystems are especially responsive to interrelated and dynamic physical, social, and biological forces. Climate, human-driven ignition and grazing dynamics, vegetation, and fuel characteristics, all interact to shape fire activity in grassland ecosystems^{1–9}. Incidence and seasonality of grassland fire are both associated with natural climate variations and anthropogenic ignition patterns^{3,9–12}. The potentially complex interactions among these drivers, along with anthropogenic climate change and the importance of grasslands ecosystems to sustaining regional livelihoods, add urgency to better understanding these relationships over multi-year periods.

Burned grassland extent is influenced by regional climate change through increasing temperature which reduces fuel moisture and, thereby, increases flammability^{13–16}, but climate change also alters productivity, biomass abundance, and carbon emissions^{17,18}, all factors which drive fuel accumulation dynamics^{13,19,20}. Further, effects of changing climates on fire activity may only manifest once critical thresholds are crossed^{21,22}. Such feedbacks and thresholds can mediate subsequent ecosystem responses to ecosystem drivers in the face of significant change in the annual extent of burned grassland incidence. Ongoing and coherent short-term change in climate across the Xilingol region provide a responsive example of potential direct anthropogenic and climatic drivers shifting critical ecosystem responses through measurable ecological traits via threshold or non-linear relationships²².

To ensure the provision of critical ecosystem services and the viability of grassland-dependent enterprises, ecologists and policymakers need reliable models to better understand the short- and long-term impacts of dynamic climatic and anthropogenic ecosystem drivers^{22,23}. Because ecological responses can be relatively insensitive to climate, at least until thresholds are crossed²², ecological modeling using ordinary least square regression (OLS) has proven to be of little value to wildfire policymakers and managers²⁴. Applicability of these models in grassland ecosystems have been limited by poor connectivity between climatic processes and ecological response

¹Institute of Natural Disaster Research, School of Environment, Northeast Normal University, Changchun 130024, China. ²State Environmental Protection Key Laboratory of Wetland Ecology and Vegetation Restoration, Northeast Normal University, Changchun 130024, China. ³Key Laboratory for Vegetation Ecology, Ministry of Education, Changchun 130024, China. ⁴Department of Forestry, Southern Illinois University, Mail Code 4411, Carbondale, IL 62901, USA. ⁵Centre for Ecosystem Management, Edith Cowan University, Joondalup, Perth 6027, Australia. ⁶Nord University Business School (HHN), Post Box 1490, 8049 Bodø, Norway. ⁷Wildland Resources Department, Utah State University, 5230 Old Main Hill, Logan, UT 84322-5230, USA. ⁸Industrial Engineering, University of Los Andes, Cra1 Este 19-40, Bogota, Colombia. ✉email: zhangjq022@nenu.edu.cn

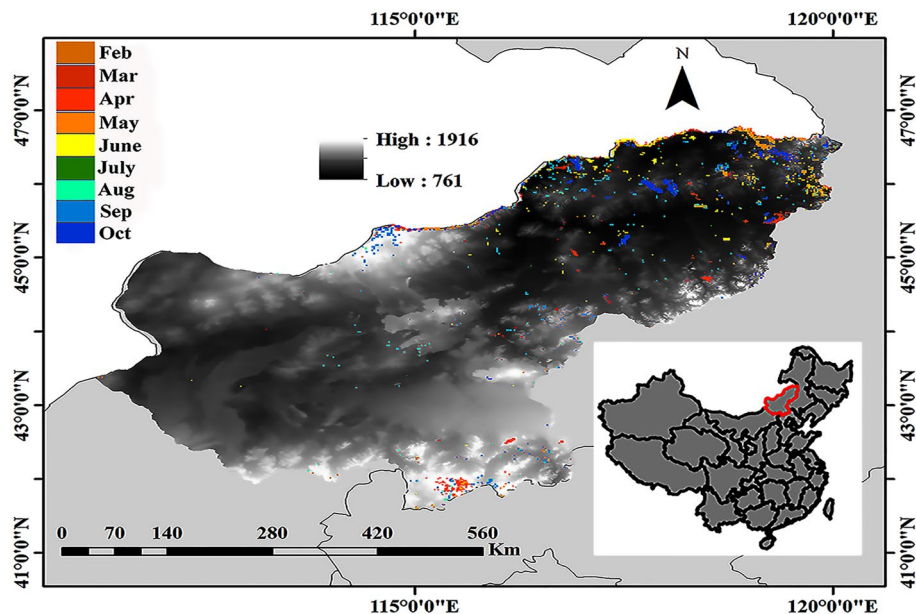


Figure 1. The location of the study area showing burned grasslands in the period 2000–2018 with each color representing a different month. The grey shading represents elevation from 761 to 1916 MAMSL. Map generated with ArcGIS 10.3, (<https://desktop.arcgis.com/>), with data sourced from https://lpaac.usgs.gov/data_access/data_pool.

for two reasons: First, is the need to understand the error structure which, if not properly estimated, especially in the case of highly skewed and heterogeneous data, can bias parameter estimations that produce false precision²⁵. Second, fire management requires appropriate information regarding human interactions with burned grasslands as these may influence seasonality and extent of the fire, and their impacts over time. The issue of stationarity in the dominant fire-climate relationships in time-series data in response to human-caused fire activity has not been well represented in the previous studies^{1,4,5,13,23}. Further, this information is often overlooked in summarized fire-climate relationships.

To explore the ability of dynamic simulations to understand and predict fire activity in grasslands, we need to target existing knowledge gaps by comparing trends in climatically driven changes in grassland fire activity and confirm that the outcomes are indeed consistent with our expectations of the key disturbance drivers of climate and ecological change on burned grasslands. To do this, we must first apply unit root tests to resolve non-stationarity issues in time series fire data from anthropogenic grasslands in order to select whether Johansen Cointegration²⁶, OLS, or dynamic simulations such as Autoregressive Distributed Lags (ARDL)²⁷ models are most appropriate. The reason for selecting the dynamic simulated ARDL model over the others is because it not only addresses complicated interactions among fire-relevant climate variables at different levels but also accounts for dynamic processes and feedbacks in a system. This research also attempts to extract additional measures from the models that can improve system predictability and reliably inform stakeholders. In this study, we aim to clarify the key drivers of climate and ecological change on burned grasslands in Xilingol, China using the novel approach of dynamic simulations of autoregressive distributed lag models.

Study area and data description. The Xilingol League lies about 600 km north of Beijing in central Inner Mongolia, the most representative steppe grassland in northern China (Fig. 1). Grassland covers 70% of the total area of 2.03×10^7 ha. Grazing has been the extensive land use in the Xilingol League where climate variations and other factors, including fire-generating human-activities, are the principal contributing factors to increased incidence of grassland burning²⁸. Heterogeneous climate and human activities contribute to a complex burn matrix throughout Xilingol. Fire occurrence and severity is greatest from spring through autumn, promoted by drying effects of wind and other climatic drivers (Fig. 2).

We used eight climatic variables to evaluate the effects of key drivers on burned grasslands and ecological change: monthly average maximum temperature (in degrees C); monthly average sunlight (in hours per month); monthly average wind speed (in meters per second); the monthly average carbon emission (in grams per square centimeter²); monthly average vapor pressure deficit; monthly average precipitation (in millimeters); and monthly average relative humidity (percentage). Heteroscedasticity²⁹, autocorrelation³⁰, and multicollinearity³¹ are evident in their effect on the climate records of burned grasslands. We then reanalyzed, assessed the fire-relevant climate variables, and confirmed the trends that are consistent with our expectations.

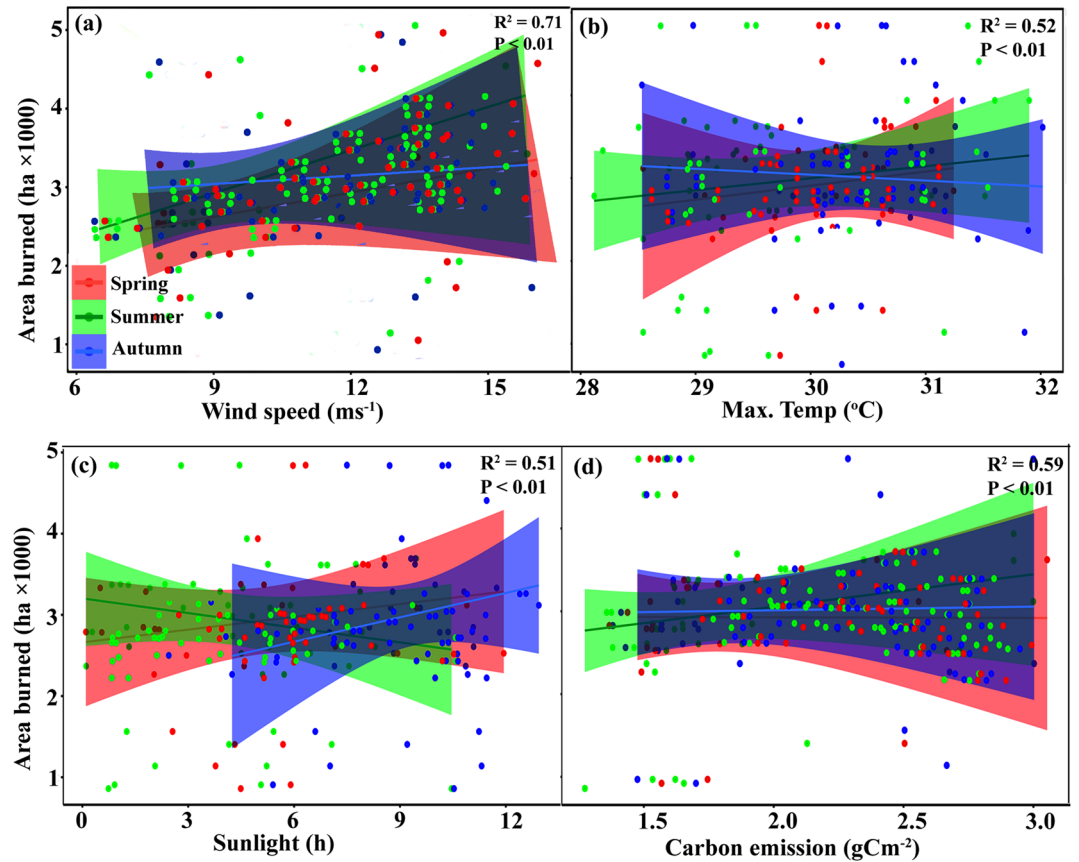


Figure 2. Influence of climate variables on burned grassland in Xilingol. Spring, Summer, and Autumn-burned grasslands are represented in red, green, and blue, respectively.

Variables	ADF		PP		KPSS		DF-GLS	
	Level	1st Diff	Level	1st Diff	Level	1st Diff	Level	1st Diff
ln T _{min}	-7.05	-19.05**	-9.91	-16.08**	0.255	0.052**	-1.41*	-16.08**
ln Rel.humidity	-8.63	-13.32**	-8.75	-24.93**	0.183	0.092**	-3.05	-23.63**
ln T _{max}	-9.14*	-14.41**	-8.55*	-12.43**	0.234*	0.011**	-4.55*	-23.63**
ln Precip	-8.96	-13.21**	-9.17	-28.49**	0.196	0.061**	-7.97	-32.39**
ln Sunlight	-5.91	-16.39**	-8.92	-26.92**	0.241	0.059**	-2.02	-34.62**
ln wind speed	-7.44*	-7.06**	-8.86*	-30.45**	0.344*	0.016**	-8.76*	-23.65**
ln Cem	-8.16	-11.01**	-9.47	-21.09**	0.126	0.044**	-8.17	-30.09**
ln Vpd	-6.01	-15.09**	-9.02	-25.02**	0.291	0.041**	-2.92	-31.69**

Table 1. Unit root test results of climatic variables for Augmented Dickey–Fuller (ADF), Phillips–Perron (PP), Kwiatkowski, Phillips, Schmidt, Shin (KPSS), and modified Dickey–Fuller generalized least squares (DF-GLS). All tests, except KPSS, use the null hypothesis of the unit root (non-stationarity), while KPSS tests the hypothesis against the alternative of stationarity. *1% level of significance. **5% level of significance.

Results and discussion

Consistent with best practices for homoscedastic stationary series using small data sets, we used the ADF, PP, KPSS, and DF-GLS tests to confirm model validation (Table 1). Maximum temperature and wind speed were stationary I(0) at a 1% level of significance while other variables were integrated to order one I(1) indicating stationarity at first difference. According to the DF-GLS test, the mean minimum monthly temperature was also stationary I(0). The absence of integration to order two I(2) for any of the variables indicated the suitability of the ARDL approach.

Published studies on fire–climate relationship^{23,32} generally do not resolve the non-stationarity issue that often arises in time series data that measures climate variables during recent years. Particularly, for explaining the variability of grassland area burned, ignoring the non-stationarity of the series leads to bad specified models that invalidate some of the results and generate high prediction error. In addition, modeling the series as if they were

stationary may produce spurious relations that, although perceived as statistically significant, are not real and do not have a correct interpretation^{33,34}. These potential problems generated by non-stationarity provide possible explanations for the large errors we observed in predictions and estimated relationships between area burned and climate. Cointegration is therefore a solution to model the complex dynamics of multivariate ecological time series when some of them are non-stationary. However, even under cointegration, all model parameters, including the lags of independent variables and the error correction term must be correctly specified. Moreover, the error term should not be serially dependent, in order to prevent spurious relationships arising from temporal correlations^{33,35}. Implementation of statistical tests for stationarity of time series, known as unit root tests, is then a fundamental tool for the ecologist to analyze longitudinal data.

The ADF is among the most used and powerful unit root tests under constant variance and if the lag p is correctly selected. Alternatively, PP unit root tests differ mainly in how they address serial correlation and heteroskedasticity in the errors. ADF tests use a parametric autoregression to approximate the autoregressive moving averages structure of the errors in the test regression, whereas the PP tests ignore any serial correlation in the test regression. Unlike ADF tests, PP tests are robust to general forms of heteroskedasticity and do not require a specified lag length for the test regression²³. The ADF and PP tests are asymptotically equivalent but may differ substantially in finite samples due to differences in correcting serial correlation in the test regression. In contrast, DF-GLS is a sensitive test that may increase the power of the previous, especially near stationarity and for small samples. Another recommendation, especially for short time series, is the simultaneous use of the KPSS test.

The lag selection criteria of log-likelihood³⁶ (LR), the final prediction error³⁷ (FPE), Akaike information criterion³⁸ (AIC), Schwarz Bayesian criterion³⁹ (SBC), and Hannan-Quinn information criterion⁴⁰ (HQ) were used to select the optimal lag on all variables prior to testing for cointegration (see Supplementary Information Table S1). All these criteria are based on the minus maximized value of the respective likelihood function including a penalization term that increases with the number of estimated parameters to favor a more parsimonious model. Balancing out these two terms creates a metric that can be used as a model selection criterion. Both SBC and HQ are harsher in the selection than AIC, tending to choose a more parsimonious model. Therefore, we preferentially select the lag(1) as the optimal lag for subsequent model estimation in order to keep the model more compact and interpretable. The final prediction error, Schwarz Bayesian criterion, and Hannan-Quinn information criterion were used to confirm lag(1) as the optimal lag for subsequent model estimation.

The ARDL bounds testing approach is advantageous over traditional cointegration methods because it can examine a mixed order of integrations. We used the F-statistic to test for the null hypothesis of no-level relationship⁴¹ (SI Table S2). The bounds test for cointegration that the F-statistic was above both lower [1(0)] and upper [1(1)] bounds critical values, thus, all the models presented reject the null hypothesis of no cointegration. To validate the cointegration test, we further examined the null hypothesis following^{42,43} in the light of response surface regressions that report critical p -value approximations⁴⁴. The p -values strongly rejected the null hypothesis at a 1% significance level, confirming a level relationship in the proposed models.

We implemented a graphical representation of the ARDL model simulation to provide an easily understandable representation of an instantaneous change (i.e., a shock) in one independent variable on the response in posterior observations (Fig. 3). This contrasts with the more elaborate interpretation of estimated coefficients, given their separation in long and short-run effects. For example, the dynamic plot shows that a positive shock (one standard deviation increase) at time $t = 10$ in wind speed produces a rapid positive short-run increase in burned grassland area.

At $t = 15$ this effect vanishes and the response value stabilizes thereafter (Fig. 3a). The confidence intervals (blue lines) indicate that the long-run change in burned grassland area is statistically significant using 75% confidence intervals, rather than significant with 90% confidence intervals. In contrast, the resulting plot of negative shock wind speed at time $t = 10$ decreased the mean response over several periods, resulting in a statistically significant equilibrium shift over four periods (Fig. 3b).

To validate the model, we used diagnostic tests to examine the residual independence (Table 2). The coefficient of determination (R-squared) resulted in a goodness of fit of $\sim 78.3\%$. The diagnostic tests showed that functional misspecification error, first-order serial correlation, heteroskedasticity, and non-normality assumptions are adequately controlled in the estimated model. To corroborate the independence of the residuals, we further employed the recursive cumulative sum (CUSUM) plots to examine the stability of the data series over the sampled period (Fig. 4).

Simulation and analysis of ARDL dynamics. The grassland ecosystems of Xilingol are characterized by highly dynamic and complex interactions among climate, human settlement patterns, ignition, and vegetation cover. Increased grazing constrains grassland burning leading to an “area burned deficit”. Otherwise, the system is highly responsive to high-speed seasonal winds with low relative humidity, quickly drying fuel, and increasing the potential for extensive grassland fires across the region^{28,42}.

The long-run and short-run dynamics illustrated in Table 2 show that wind speed represents the dominant variable that extensively increases grassland area burned over both the long and short run. Arithmetically, a 1% increase in mean monthly wind speed increases the burned grassland area by 6.9% (different from zero in a 99% confidence interval (99% CI)) in the long run, and 2.2% (99% CI) in the short run, while holding other variables constant. Besides the direct impact of wind speed on the grassland area burned, it is likely to affect other environmental conditions which also influence fire behavior throughout the Xilingol region. That is, the dynamics of the whole system show complex causal relations that may not be completely explicit on single equation coefficients. For example, our data suggest that wind speed increased following the burning of grasslands (see Supplementary Information Fig. S1). This is consistent with the observation of more extensive burned grasslands over the study period also associated with a dramatic increase in wind speed, an indicator of potential positive feedback not

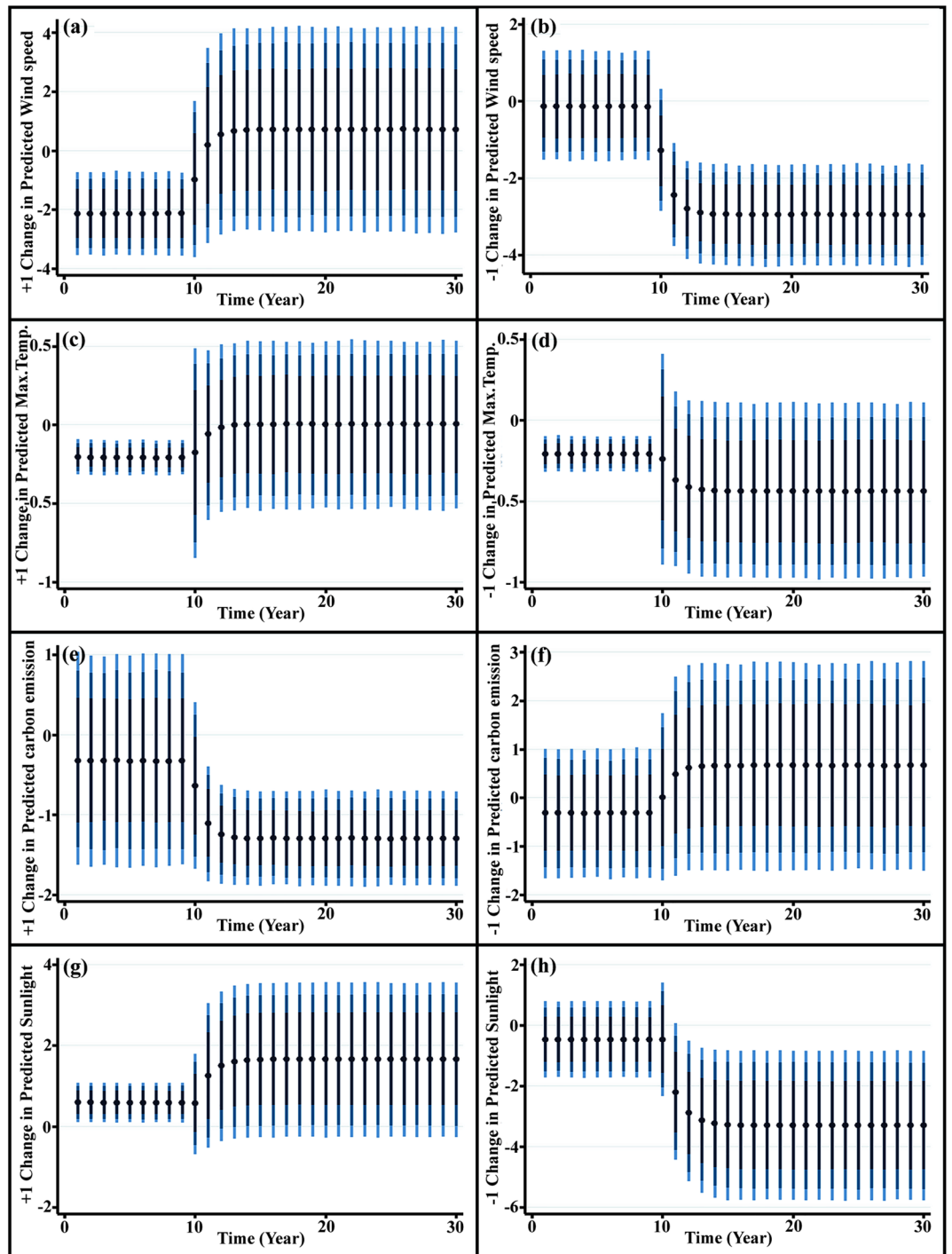


Figure 3. Plots of the dynamic simulated ARDL model showing the effect of: (a,b) positive and negative one standard deviation change in predicted wind speed on burned grasslands (c,d) positive and negative one standard deviation change in predicted maximum temperature on burned grasslands (e,f) positive and negative one standard deviation change in predicted carbon emission on burned grasslands (g,h) positive and negative one standard deviation change in predicted sunlight on burned grasslands. Dots represent average predicted value while dark blue to light blue lines denote 75, 90 and 95% confidence intervals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

considered in previous studies²³. Importantly, high wind speeds were measured in May–August accompanied by the dry land surface conditions and atmospheric aridity, contributing to the influence of burning grasslands feeding back to further promote pyrogenic conditions. Above the threshold, each incremental increase in wind

Variable	Coefficient	Std. error	t-Statistic	Prob
$\ln Wind_t$	6.929*	1.591	4.352	0.01
$\Delta \ln Wind_{t-1}$	2.237*	0.106	21.06	0.01
$\ln Tmax_t$	2.361**	0.095	24.521	0.01
$\Delta \ln Tmax_{t-1}$	2.281*	0.519	4.393	0.01
$\ln Rel.humanity_t$	0.097	0.064	1.514	0.134
$\Delta \ln Rel.humanity_{t-1}$	0.068	0.067	1.034	0.305
$\ln Precip_t$	0.106	0.099	1.070	0.288
$\Delta \ln Precip_{t-1}$	0.091	0.103	0.883	0.380
$\ln Sunlight_t$	1.986*	0.412	4.814	0.01
$\Delta \ln Sunlight_{t-1}$	0.239*	0.112	2.128	0.01
$\ln Tmin_t$	0.043	0.083	0.519	0.605
$\Delta \ln Tmin_{t-1}$	0.073	0.126	0.578	0.564
$\ln Bpd_t$	0.180	0.155	1.161	0.249
$\Delta \ln Vpd_{t-1}$	0.017	0.064	0.271	0.786
$\ln Cem_t$	0.913**	0.155	5.872	0.01
$\Delta \ln Cem_{t-1}$	0.864*	0.189	4.562	0.01
ECT_{t-1}	-0.601*	0.124	-4.81	0.000
R^2	0.783			
Adjusted R^2	0.714			
Simulations	5000			
Durbin-Watson-stat	2.082			
Diagnostics				
$\chi LM - ARCH^2$	0.38			
$\chi LM - B - G^2$	0.41			
Functional form	2.131 (0.121)			
Ramsey RESET test	3.421 (0.154)			
Normality	0.761 (0.681)			

Table 2. Short-run and long-run coefficients of dynamic ARDL. ARDL (1,1,0,1,1,1) was selected based on the Schwarz Bayesian Criterion. **1% level of significance. *5% level of significance.

speed on a monthly scale leads to larger burned grasslands across the Xilingol region, overriding other ecological properties, and grassland functions.

Threshold maximum temperature causes uncertainty in ecological change. The maximum temperature is also an important predictor of the increasing extent of burned grasslands. The estimated elasticity of maximum temperature in the long-run indicates that a 1% increase in maximum temperature results in a 2.4% increase in burned grassland extent (99% CI) while holding other variables constant (Table 2). The short-run estimated elasticity of maximum temperature is similar to the long-run response, as a 1% increase in maximum temperature leads to a 2.3% (99% CI) increase in burned grasslands. The results corroborate similar studies^{13–15,22}.

The effect of maximum temperature could have more uncertainty because significant relationships between it and burned grasslands occur only when the regional summer maximum temperature exceeded a threshold level, producing a nonlinear relationship²². That is, slight changes in maximum temperature are unlikely to change the ecosystem except when a threshold is crossed, above which effects can be dramatic.

A more meaningful understanding of the effect of maximum temperature on burned grassland areas can be achieved by the use of the simulations of the ARDL model, given that all the relationships among variables are considered. At time $t = 10$, a small increase in maximum temperature in the short run is not statistically significant, but becomes statistically significant in the long run. Following a negative shock on the temperature at $t = 10$, the maximum temperature decreases, stabilizing three periods later. The equilibrium temperature shift over the long-run scenario time is statistically significant (95% CI). Discrepancies between burned grassland area and maximum temperature trends in Xilingol may be rooted in urbanization, vegetation type transitions, and/or natural variability to climate response. Additionally, the frequency of human ignitions, along with other anthropogenic disturbances, strongly supported the abrupt impact of maximum temperature overpowering the influence of other environmental variables (SI Fig. S2).

Influence of sunlight on burned grasslands and ecological change. The amount of sunlight exacerbates changes in burned grassland area both directly via the drying effect of radiant heat on fuel and indirectly by influencing other climatic parameters such as temperature and humidity. A 1% increase in sunlight resulted in a 1.98% and 0.24% (99% CI) increase in burned grasslands in the long and short-runs, respectively (Table 2).

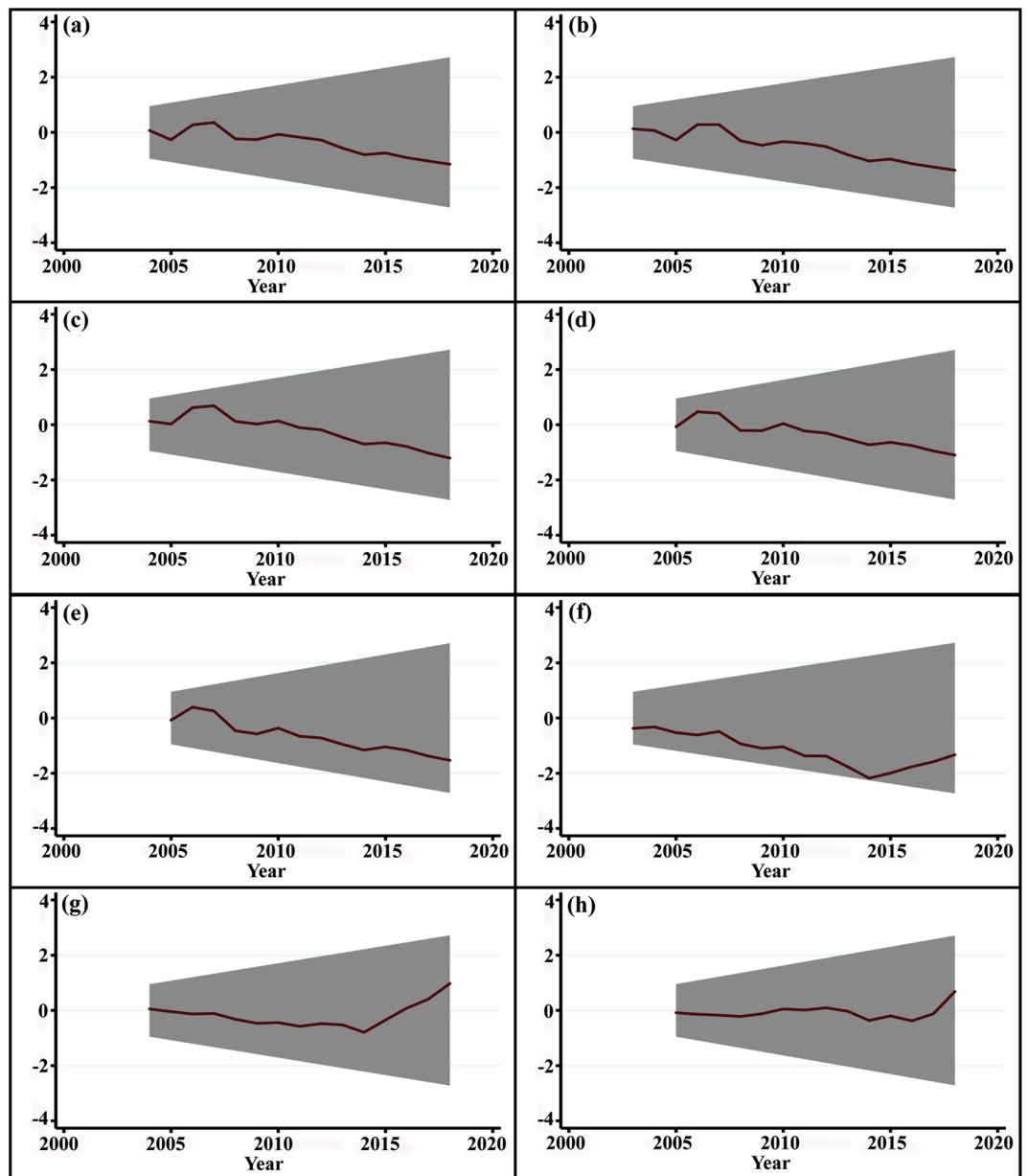


Figure 4. Recursive CUSUM plot of (a) Maximum temperature, (b) minimum temperature, (c) sunlight, (d) wind speed, (e) vapor pressure deficit, (f) relative humidity, (g) precipitation, (h) carbon emission with 95% confidence bands around the null. For all variables, residuals fall between the limits, validating the ARDL bounds testing model and showing stability of the variances of coefficients over time.

A negative shock in $\ln(\text{sunlight})$ at time ($t = 10$), insignificantly impacted burned grassland area over the next several periods (Fig. 3). This rapid response is explained by the inclusion of sunlight with a lagged-difference as a dependent variable. Conversely, a positive shock to $\ln(\text{sunlight})$ at time ($t = 10$) increases area burned, which, in turn, shifts the equilibrium and predicts the ecological change just beyond $t = 10$.

Impacts of burned grasslands to carbon emissions and ecological change. The consequence of carbon emissions from burned grasslands was found to be an important relationship in the long run. Carbon emissions, whether from these fires or other sources in Xilingol may influence the extent of burning in this region by contributing to the forcing of atmospheric temperature. More directly, grassland fires resets succession, potentially increasing ecosystem flammability (SI Fig. S3). The estimated elasticity of carbon emission, in the long run, is 0.91, hence, a 1% increase in carbon emissions is related to a 0.86% increase in burned grasslands (95% CI) in the long run and 0.91% (99% CI) in the short run (Table 2).

The graphical representation of Fig. 3 shows that a change in predicted carbon emissions on burned grasslands and ecological change. The positive shock of carbon emission on burned grasslands and ecological change was associated with a small increase in carbon emission in the short run that is statistically significant at time $t = 10$.

Consistent with other modeled changes, carbon emissions also increase significantly over the predicted value in the long run.

Ecological properties under threshold-governed ecosystems. The mechanism underpinning ecological properties can vary widely with significant uncertainty where threshold relationships among different biophysical processes occur²². The model suggests significant changes in the frequency of climatic variations prior to and in response to, the increased areal extent of burned grasslands (SI Figs. S1–S3). Our results highlight the importance of varying patterns in dominant burned grassland climatic factors (e.g., SI Figs. S1–S3) may also hint at broader implications, suggesting that sudden ecosystem state changes, fire-related or otherwise, may naturally exhibit significant variability across space and time²². The results strongly suggest that area burned increases mainly due to increases in wind speeds during the summer season when characteristically warm temperatures and low humidity facilitate fuel drying and fire spread (see SI Fig. S2) and not due to concurrent changes in anthropogenic activities such as fire suppression practices, land management, and human ignitions. This is not to say that anthropogenic activities are negligible in dictating the modern extent of annually burned grasslands. Rather, increased human ignition is associated with increased burned grassland extent, further confounding a modeling effort based solely on physical ecosystem drivers.

Our long-run analyses suggested that predicted burned grassland extent is insensitive to change in maximum temperature, vapor pressure deficit, precipitation, or relative humidity. These findings contrast with studies that predict increased fire impacts in north-eastern China in response to higher maximum temperature, decreased precipitation, greater vapor pressure deficit, and lower humidity^{28,42,45–48}. We posit four potential explanations for this divergence: First, although inter-annual precipitation outside the fire season varies significantly in Xilingol grasslands, little precipitation arrives during the fire season itself. Therefore, inter-annual variation in fire season precipitation is insufficient to impact burned grassland extent. Second, grassland soil field capacity in this region is quite high. Therefore, fire season fuel moisture levels are accordingly unresponsive to precipitation variation. Third, the vast majority of fires in Xilingol originate from human activities^{23,28,42}. These ignitions, and wind speeds and temperatures coincident with them, rather than fuel conditioning, may explain burned grassland extent. Finally, it is also possible that our statistical approach is insufficiently sensitive to detect a precipitation response under the conditions experienced across the 18-year time series used in this study.

Model estimation. Our empirical strategy for estimating the proposed models follows a series of tests required before the application of the novel dynamic simulations of ARDL²⁷. First, the dynamic ARDL simulations require that the series are integrated below order two, $I(2)$, meaning that the series can be integrated of order one, $I(1)$; integrated of order zero, $I(0)$, or a mixture of both $I(1)$ and $I(0)$. To test the level of integration in the series, we utilized the Augmented Dickey–Fuller (ADF)⁴⁹, Phillips–Perron (PP)⁵⁰, Kwiatkowski, Phillips, Schmidt, Shin (KPSS)⁵¹, and modified Dickey–Fuller generalized least squares (DF-GLS)⁵² unit root techniques. While all the unit root techniques are tested based on the null hypothesis of unit root, KPSS is tested based on the null hypothesis of stationarity. Testing for unit root is an important component in the model estimation to prevent a spurious relationship between the series. A simplified version of testing for unit root can be expressed in AR (1) as:

$$y_t = a + \rho y_{t-1} + \varepsilon_t, \quad t = 1 \dots T \quad (1)$$

where y_t is the series to be tested for unit root, ε_t is the error term. ε_t is independent and identically distributed, independent of the initial detected value y_0 , and has a zero-mean if the previous observation of y is: $E(\varepsilon_t | y_{t-1} \dots y_0) = 0$. The null hypothesis of the unit root can be tested by $H_0: \rho = 1$ against the alternative $H_1: \rho < 1$. If the null hypothesis is true, the model is non-stationary and would have an order of integration larger than 0. The ADF test includes lagged values of the response allowing for testing stationarity in higher-order autoregressive processes. The PP test makes a similar correction as ADF but uses a non-parametric modification of the t statistic that makes it more robust for unspecified autocorrelation and non-constant variance. The KPSS test may be used as a complementary test for stationarity due to its null hypothesis, which can be used with data on which other unit root tests are not conclusive. The DF-GLS test presents an improvement over ADF by using a generalized least squares rationale. On several empirical studies^{53,54} it has been shown to have better power performance for small-sample sizes and especially when the coefficient ρ is close to 1, that is, the series is near to non-stationarity.

After confirming the order of integration, the dynamic ARDL simulations require the series in the models to be cointegrated. The traditional concept of cointegration exists when a set of series are $I(1)$ but there are some linear combinations of them that are $I(0)$. This will, in effect, control for spurious regression that affects statistical inferences. Validation of the existence of cointegration confirms a possible long-run relationship between the series. Contrary to the traditional forms of cointegration that require all series to exhibit both $I(0)$ and $I(1)$, Pesaran proposed a vibrant cointegration method that deals with a sophisticated order of integration⁴¹. This means that the autoregressive distributed lag (ARDL) bounds testing cointegration do not require strict pre-conditions, except that the order of integration should not be $I(2)$. Before cointegration was tested, we utilized lag selection criteria to determine the optimal lag. We used the ARDL bounds testing approach to examine the level relationship between the variables. The null hypothesis of no cointegration is tested using Fisher's test⁵⁵ statistic matched against the lower and upper bound critical values. The null hypothesis of no level relationship is rejection if the F-statistic is extreme compared to the upper bound critical values. To validate the bounds testing procedure for cointegration, we utilized the novel critical value bounds test and approximate p-values using response surface regression⁴⁴.

Third, after validating the existence of a level relationship, the dynamic ARDL simulations require that the series are fitted with the error correction based-ARDL, expressed as:

$$\Delta y_t = \alpha_0 + \delta_0 y_{t-1} + \delta_1 x_{1,t-1} + \dots + \delta_k x_{k,t-1} + \sum_{i=1}^p \alpha_i \Delta y_{t-1} + \sum_{j=0}^{q_1} \beta_{1,j} \Delta x_{1,t-j} + \dots + \sum_{j=0}^{q_k} \beta_{k,j} \Delta x_{k,t-j} + \varepsilon_t \quad (2)$$

where y_t denotes a change in the target variable, α_0 is the constant, y_{t-1} represents the lagged target variable at levels, x represents the independent series in time (\cdot) _{$t-1$} at levels, Δ is the first difference, p and q are the optimal lags, and ε_t is the error term. Model 2 corresponds to a particular formulation called Error Correction Model (ECM), where long and short-run effects are combined. Parameters δ_j for $j=0, 1, \dots, k$; are related to the long-run relations that were found by cointegration. Expressing and for $j=1, \dots, k$, then all variables related to δ coefficients may be rearranged as: $(\alpha_0 y_{t-1} + \alpha_1 x_{1,t-1} + \dots + \alpha_k x_{k,t-1}) = \alpha_0 EC_{t-1}$ where EC_{t-1} is known as the error correction term, that allows to estimating parameters in Model 2 given cointegration.

The estimated models were validated using the cumulative sum and cumulative sum of square⁵⁶, to examine the stability of the estimated models. After meeting all the pre-conditions, the novel dynamic simulations of ARDL can then be applied using Eq. (2) but modified to include counterfactual change using a single shocked regressor, amount of shock to be applied (\pm), length of the scenario for the simulation, shock scenario time, number of simulations, burn-in rate, and the resulting dynamic plot.

Conclusion

This study demonstrates the importance of addressing non-stationarity of time series data, using the unit root test to elucidate short- and long-term ecosystem responses. Models for fire-prone grassland, and other landscapes existing near-threshold levels for key ecosystem drivers are susceptible to compounding data inaccuracies. The dynamic simulated ARDL approach identified wind speed as the most important climatic variable in determining the model outputs in the complex fire-climate-human relationship, followed by maximum temperature, with sunlight and carbon emission having only minor to modest influence. Potential positive and negative feedbacks were identified and these are needed to inform best management practices. Adaptation of this approach would help generate the more precise information needed by ecologists and stakeholders to plan and respond appropriately to management challenges in fire-prone landscapes under future environmental scenarios.

Data availability

The data of climatic variables from 2001 to 2018 were taken from the China Meteorological Data Sharing Service Centre (<https://data.cma.cn/en>; accessed 28 June 2020). Xilingol burned grasslands records were collated from the Monitoring Centre of the Ministry of Agriculture (<https://www.moa.gov.cn/>; accessed 28 June 2020). The field data of Xilingol burned grasslands are available from the corresponding author upon request.

Received: 12 July 2020; Accepted: 26 October 2020

Published online: 17 November 2020

References

- Elzein, T., Larocque, G. R., Sirois, L. & Arseneault, D. Comparing the predictions of gap model with vegetation and disturbance data in south-eastern Canadian mixed forests. *For. Ecol. Manage.* **455**, 117649 (2020).
- Slingsby, J. A., Moncrieff, G. R., Rogers, A. J. & February, E. C. Altered ignition catchments threaten a hyperdiverse fire-dependent ecosystem. *Glob. Change Biol.* **26**, 616–628 (2020).
- Fernández-García, V. *et al.* Fire regimes shape diversity and traits of vegetation under different climatic conditions. *Sci. Total Environ.* **716**, 137137 (2020).
- Corrêa Scalon, M. *et al.* Diversity of functional trade-offs enhances survival after fire in Neotropical savanna species. *J. Veg. Sci.* **31**, 139–150 (2020).
- Moradzadeh, H., Heydari, M., Omidipour, R., Mezban, A. & Prévosto, B. Ecological effects of fire severity and time since fire on the diversity partitioning, composition and niche apportionment models of post-fire understory vegetation in semi-arid oak forests of Western Iran. *Ecol. Eng.* **143**, 105694 (2020).
- Kowaljow, E. *et al.* A 55-year-old natural experiment gives evidence of the effects of changes in fire frequency on ecosystem properties in a seasonal subtropical dry forest. *L. Degrad. Dev.* **30**, 266–277 (2019).
- Keeley, J. E., van Mantgem, P. & Falk, D. A. Fire, climate and changing forests. *Nat. Plants* **5**, 774–775 (2019).
- Hayashi, N., Kawano, T. & Inoue, J. Long-term response of respective grass types to variations in fire frequency in central Japan, inferred from phytolith and macrocharcoal records in cumulative soils deposited during the Holocene. *Quat. Int.* **527**, 94–102 (2019).
- Turco, M. *et al.* Climate drivers of the 2017 devastating fires in Portugal. *Sci. Rep.* **9**, 1–8 (2019).
- Wahl, E. R., Zorita, E., Trouet, V. & Taylor, A. H. Jet stream dynamics, hydroclimate, and fire in California from 1600 CE to present. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 5393–5398 (2019).
- Rogers, B. M., Balch, J. K., Goetz, S. J., Lehmann, C. E. R. & Turetsky, M. Focus on changing fire regimes: interactions with climate, ecosystems, and society. *Environ. Res. Lett.* <https://doi.org/10.1088/1748-9326/ab6d3a> (2020).
- Rogers, B. M., Soja, A. J., Goulden, M. L. & Randerson, J. T. Influence of tree species on continental differences in boreal fires and climate feedbacks. *Nat. Geosci.* **8**, 228–234 (2015).
- Williams, A. P. *et al.* Observed impacts of anthropogenic climate change on wildfire in California. *Earth's Fut.* <https://doi.org/10.1029/2019EF001210> (2019).
- Hurteau, M. D., Liang, S., Westerling, A. L. R. & Wiedinmyer, C. Vegetation–fire feedback reduces projected area burned under climate change. *Sci. Rep.* <https://doi.org/10.1038/s41598-019-39284-1> (2019).
- Syphard, A. D. *et al.* The relative influence of climate and housing development on current and projected future fire patterns and structure loss across three California landscapes. *Glob. Environ. Chang.* **56**, 41–55 (2019).
- Seager, R. *et al.* Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases. *Nat. Clim. Change* **9**, 517–522 (2019).

17. Holdo, R. M., Onderdonk, D. A., Barr, A. G., Mwita, M. & Anderson, T. M. Spatial transitions in tree cover are associated with soil hydrology, but not with grass biomass, fire frequency, or herbivore biomass in *Serengeti savannahs*. *J. Ecol.* <https://doi.org/10.1111/1365-2745.13303> (2020).
18. Rowell, E. *et al.* Coupling terrestrial laser scanning with 3D fuel biomass sampling for advancing wildland fuels characterization. *For. Ecol. Manage.* **462**, 117945 (2020).
19. Golovynskiy, S., *et al.* ced Muspt. *Nanotechnology* 0–22 (2019).
20. Murphy, B. P., Prior, L. D., Cochrane, M. A., Williamson, G. J. & Bowman, D. M. J. S. Biomass consumption by surface fires across Earth's most fire prone continent. *Glob. Change Biol.* **25**, 254–268 (2019).
21. Hantson, S. *et al.* Rare, intense, big fires dominate the global tropics under drier conditions. *Sci. Rep.* **7**, 7–11 (2017).
22. Young, A. M., Higuera, P. E., Abatzoglou, J. T., Duffy, P. A. & Hu, F. S. Consequences of climatic thresholds for projecting fire activity and ecological change. *Glob. Ecol. Biogeogr.* **28**, 521–532 (2019).
23. Liu, X. P., Zhang, G. Q., Lu, J. & Zhang, J. Q. Risk assessment using transfer learning for grassland fires. *Agric. For. Meteorol.* **269–270**, 102–111 (2019).
24. MacKay, D. J. C. Comparison of approximate methods for handling hyperparameters. *Neural Comput.* <https://doi.org/10.1162/089976699300016331> (1999).
25. Koenker, R. & Bassett, G. Robust tests for heteroscedasticity based on regression quantiles. *Econometrica* <https://doi.org/10.2307/1912528> (1982).
26. Johansen, S. Statistical analysis of cointegration vectors. *J. Econ. Dyn. Control* **12**, 231–254 (1988).
27. Jordan, S. & Philips, A. Q. Cointegration testing and dynamic simulations of autoregressive distributed lag models. *Stata J.* **18**, 902–923 (2018).
28. Shabbir, A. H. *et al.* Determining the sensitivity of grassland area burned to climate variation in Xilingol, China, with an autoregressive distributed lag approach. *Int. J. Wildl. Fire* **28**, 628–639 (2019).
29. Harvey, A. C. Estimating regression models with multiplicative heteroscedasticity. *Econometrica* <https://doi.org/10.2307/1913974> (1976).
30. Andrews, D. W. K. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* <https://doi.org/10.2307/2938229> (1991).
31. Farrar, D. E. & Glauber, R. R. Multicollinearity in regression analysis: The problem revisited. *Rev. Econ. Stat.* **49**, 92 (1967).
32. Abatzoglou, J. T., Dobrowski, S. Z. & Parks, S. A. Multivariate climate departures have outpaced univariate changes across global lands. *Sci. Rep.* **10**, 1–9 (2020).
33. Ferson, W. E., Sarkissian, S. & Simin, T. T. Spurious regressions in financial economics?. *J. Finan.* <https://doi.org/10.1111/1540-6261.00571> (2003).
34. Deng, A. Understanding spurious regression in financial economics. *J. Finan. Econom.* <https://doi.org/10.1093/jjfinec/nbs025> (2013).
35. Lourenço, V. M., Pires, A. M. & Kirst, M. Robust linear regression methods in association studies. *Bioinformatics* <https://doi.org/10.1093/bioinformatics/btr006> (2011).
36. Grossman, Y. L. *et al.* Critique of stepwise multiple linear regression for the extraction of leaf biochemistry information from leaf reflectance data. *Remote Sens. Environ.* [https://doi.org/10.1016/0034-4257\(95\)00235-9](https://doi.org/10.1016/0034-4257(95)00235-9) (1996).
37. Tschernig, R. & Yang, L. Nonparametric lag selection for time series. *J. Time Ser. Anal.* <https://doi.org/10.1111/1467-9892.00193> (2000).
38. Pan, W. Akaike's information criterion in generalized estimating equations. *Biometrics* <https://doi.org/10.1111/j.0006-341X.2001.00120.x> (2001).
39. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* <https://doi.org/10.1214/aos/1176344136> (1978).
40. Hannan, E. J. & Quinn, B. G. The determination of the order of an autoregression. *J. R. Stat. Soc. Ser. B* **41**, 190–195 (1979).
41. Pesaran, M. H., Shin, Y. & Smith, R. J. Bounds testing approaches to the analysis of level relationships. *J. Appl. Econom.* **16**, 289–326 (2001).
42. Shabbir, A. H. *et al.* Predicting the influence of climate on grassland area burned in Xilingol, China with dynamic simulations of autoregressive distributed lag models. *PLoS ONE* **15**, 1–19 (2020).
43. Sarkodie, S. A. & Adams, S. Renewable energy, nuclear energy, and environmental pollution: Accounting for political institutional quality in South Africa. *Sci. Total Environ.* **643**, 1590–1601 (2018).
44. Kripfganz, S. & Schneider, D. C. ARDL: Stata module to estimate autoregressive distributed lag and equilibrium correction models. in *Proceedings of the 2018 London Stata Conference* (2018).
45. Zhang, F., Nilsson, C., Xu, Z. & Zhou, G. Evaluation of restoration approaches on the Inner Mongolian Steppe based on criteria of the Society for Ecological Restoration. *L. Degrad. Dev.* **31**, 285–296 (2020).
46. Mckelvey, K. S. *et al.* Climate change and rocky mountain ecosystems. *Adv. Glob. Change Res.* **63**, 143–167 (2018).
47. Meng, C. *et al.* Analyses of observed features and future trend of extreme temperature events in Inner Mongolia of China. *Theor. Appl. Climatol.* **139**, 577–597 (2020).
48. Wang, J., Brown, D. G., Riolo, R. L., Page, S. E. & Agrawal, A. Exploratory analyses of local institutions for climate change adaptation in the Mongolian grasslands: An agent-based modeling approach. *Glob. Environ. Change* **23**, 1266–1276 (2013).
49. Dickey, D. A. & Fuller, W. A. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* **74**, 427 (1979).
50. Phillips, P. C. B. Understanding spurious regressions in econometrics. *J. Econom.* **33**, 311–340 (1986).
51. Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. & Shin, Y. Testing the null hypothesis of stationarity against the alternative of a unit root. How sure are we that economic time series have a unit root?. *J. Econom.* **54**, 159–178 (1992).
52. Wu, S. Lag length selection in DF-GLS unit root tests. *Commun. Stat. Simul. Comput.* **39**, 1590–1604 (2010).
53. Oteng-Abayie, E. F. & Frimpong, M. J. Bounds testing approach to cointegration: An examination of foreign direct investment trade and growth relationships. *Am. J. Appl. Sci.* <https://doi.org/10.3844/ajassp.2006.2079.2085> (2006).
54. Elliott, G., Rothenberg, T. J. & Stock, J. H. Efficient tests for an autoregressive unit root. *Econometrica* <https://doi.org/10.2307/2171846> (1996).
55. Mangiafico, S. *An R Companion for the Handbook of Biological Statistics* (Rutgers Cooperative Extension, New Brunswick, 2015).
56. Brown, R. L., Durbin, J. & Evans, J. M. Techniques for testing the constancy of regression relationships over time. *J. R. Stat. Soc. Ser. B* **37**, 149–163 (1975).

Acknowledgements

This research was funded by The International (Regional) Cooperation and Exchange Programs of National Natural Science Foundation of China under grant no. 41961144019.

Author contributions

A.H.S. drafted the manuscript, tables, and figures and performed the statistical analysis. J.Z. provided grassland area burned and climatic data. J.W.G., E.J.B.V.E. and J.A.L. writing review, editing manuscript in scientific and

contributed equally. The methodology was performed by S.A.S. Model interpretation was performed by C.F.V. and critically reviewed by J.A.L. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-76191-2>.

Correspondence and requests for materials should be addressed to J.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020