Determining the sensitivity of grassland area burned to climate variation in Xilingol, China, with an autoregressive distributed lag approach

Ali Hassan Shabbir\textsuperscript{A,B,C}, Jiquan Zhang\textsuperscript{A,B,C,G}, Xingpeng Liu\textsuperscript{A,B,C}, James A. Lutz\textsuperscript{D}, Carlos Valencia\textsuperscript{E} and James D. Johnston\textsuperscript{F}

\textsuperscript{A}Institute of Natural Disaster Research, School of Environment, Northeast Normal University, Changchun, 130024, China.
\textsuperscript{B}State Environmental Protection Key Laboratory of Wetland Ecology and Vegetation Restoration, Northeast Normal University, Changchun, 130024, China.
\textsuperscript{C}Key Laboratory for Vegetation Ecology, Ministry of Education, Changchun, 130024, China.
\textsuperscript{D}Wildland Resources Department, Utah State University, 5230 Old Main Hill, Logan, UT 84322-5230, USA.
\textsuperscript{E}Industrial Engineering Department, University of los Andes, Carrera 1 Este No. 19 A 40, Bogota, Colombia.
\textsuperscript{F}College of Forestry, Oregon State University, 140 Peavy Hall, 3100 SW Jefferson Way, Corvallis, OR 97333, USA.
\textsuperscript{G}Corresponding author. Email: zhangjq022@nenu.edu.cn

Abstract. We examined the relationship between climate variables and grassland area burned in Xilingol, China, from 2001 to 2014 using an autoregressive distributed lag (ARDL) model, and describe the application of this econometric method to studies of climate influences on wildland fire. We show that there is a stationary linear combination of non-stationary climate time series (cointegration) that can be used to reliably estimate the influence of different climate signals on area burned. Our model shows a strong relationship between maximum temperature and grassland area burned. Mean monthly wind speed and monthly hours of sunlight were also strongly associated with area burned, whereas minimum temperature and precipitation were not. Some climate variables like wind speed had significant immediate effects on area burned, the strength of which varied over the 2001–14 observation period (in econometrics terms, a ‘short-run’ effect). The relationship between temperature and area burned exhibited a steady-state or ‘long-run’ relationship. We analysed three different periods (2001–05, 2006–10 and 2011–14) to illustrate how the effects of climate on area burned vary over time. These results should be helpful in estimating the potential impact of changing climate on the eastern Eurasian Steppe.

Additional keywords: ARDL model, climate change, climate sensitivity, grassland fire.

Received 2 May 2018, accepted 8 May 2019, published online 1 July 2019

Introduction

Changing climate and land-use practices have had a significant impact on the productivity, community composition, structure and function of many terrestrial ecosystems (Mooney et al. 2009; Kipling et al. 2016; Zhang et al. 2016). Changes to ecosystem structure and function as a result of disturbance are more difficult to predict than directional climate change. An understanding of the sensitivity of grasslands – lands dominated by grasses as opposed to trees or shrubs – to climate-mediated changes such as fire is especially important because grasslands are one of the most spatially extensive vegetation types in the world, accounting for ~25% of the global land surface. Changes to grasslands could have global implications for resource cycling and carbon sequestration (Yang et al. 2016).

Anthropogenically derived activities are important drivers of grassland degradation, especially in developing countries (Yang et al. 2005; Verón et al. 2006; Bliege Bird et al. 2016; Gowlett 2016). Some studies have attributed grassland degradation to temperature increases and changes in precipitation seasonality (Ravi et al. 2010; Yeganeh et al. 2014), whereas other studies report that it may be difficult to differentiate between causal factors (Wessels et al. 2007; Gang et al. 2014). Less attention has been focused on fire-mediated change, and few studies have quantified the relationship between climate variability and fire extent in grasslands at large spatial scales (Zheng et al. 2006; Wessels et al. 2008).

Fire regimes of grasslands and forests have been analysed at multiple temporal and spatial scales (Abatzoglou and
Williams 2016; Francos et al. 2016; Garbolino et al. 2016; Gaudreau et al. 2016; Harris et al. 2016; Kukavskaya et al. 2016; Liu and Wimberly 2016; Tian et al. 2016; Liu et al. 2017). Previous research suggests that extreme weather is an important determinant of fire severity and extent. Increased temperatures and wind speeds and lower relative humidity are correlated with increases in fire size, intensity and severity (Argañaraz et al. 2015; Holsinger et al. 2016). In many regions, wildfire activity has fluctuated dramatically during the last century, and fire regimes have been altered in most geographic regions (Ikemori et al. 2015; Semeraro et al. 2016). Analysis of the eastern Eurasian Steppe has only recently become possible with the availability of national fire datasets (Eugenio et al. 2016).

The objective of the present study was to examine changes in area burned in northern China given intra- and interannual climate trends. To accomplish this objective, we used an econometrics method to study the relationship between area burned and climate time series that are non-stationary, that is, the mean and the variance of the series do not remain constant during the observation period. In econometrics, it is common to estimate the effect on a response of interest, e.g. stock performance, given non-stationary trends in time series predictor variables, e.g. inflation and market capitalisation. In this financial market example, investigators are interested in relating past and present observations of predictor and response variables to future observations of the response. These effects can be characterised either as ‘short-run’, or immediate fluctuations in the response given a change in a predictor variable, or ‘long-run’ relationships from which variables deviate but always return to over the entire period of observation. In the case of grassland area burned in Xilingol over time, we used an autoregressive distributed lag (ARDL) approach to determine the sensitivity of grassland area burned to climate trends. We present a detailed account of the application of this methodology and its potential usefulness to the fire ecology community.

Data description and methodology

Study area and climate data

The Xilingol League (a league is an administrative unit of the autonomous region of Inner Mongolia within China) is located on the eastern Eurasian steppe in northern China (Fig. 1). The total area of Xilingol is 2.03 × 10^5 ha, of which 70% is grasslands. Xilingol grasslands are among the most productive prairies in China (Kobayashi et al. 1994; Chen et al. 2008), and have been shown to be highly sensitive both to climate change and human activity, including human-caused fire (Hardy 2005; Kaloudis et al. 2005; Zhijun et al. 2009).

Between 2001 and 2014, 564 fires within Xilingol burned 29,092 ha. Most of the fires (94.6%) occurred between March and October. Fire activity was somewhat bimodal, with large areas burned in April–May and also in September–October (Fig. 2). Area burned in Xilingol increased from 2001 to 2014, although the interannual and spatial pattern of fire varied during this time period (Fig. 3). This variability suggests potentially complex interactions between different climate drivers of fire.

We used six climate variables to predict area burned: monthly average maximum temperature ($T_{\text{max}}$), monthly average minimum temperature ($T_{\text{min}}$), monthly precipitation (Precip), monthly average wind speed, monthly average sunlight, and monthly average relative humidity (percentage of saturation humidity). Two of these variables were weakly correlated (sunlight hours and relative humidity) and two variables were strongly correlated ($T_{\text{max}}$ and $T_{\text{min}}$). Following the practice of most econometric analyses, we ignored multicollinearity of variables in implementing the ARDL model because the degree of differencing of data using this method tends to decompose model residuals and eliminate multicollinearity. Also, in theory, multicollinearity only inflates the errors of regression coefficients but coefficients remain unbiased (Farrar and Glauber 1967; Goldberger 1991). Climate data from 2001 to 2014 were taken from the China Meteorological Data Sharing Service Centre (https://data.cma.cn/en, last accessed 20 May 2019). Data for grassland area burned were acquired from the Monitoring Centre of the Ministry of Agriculture (http://www.moa.gov.cn/, last accessed 20 May 2019). The basic meteorological and burned area data used in the present paper are included in the online supplemental material.

ARDL model

The available climate data that are potentially explanatory of grassland area burned in Xilingol League are time series. A critical assumption of ordinary least-squares analysis methods that could be used to evaluate the relationship between area burned and climate is stationarity, i.e. the means and variances of the time series are constants that are independent of time. Failure to account for non-stationarity in time series risks spurious regression results in which t-statistics and measures of model fit are misleadingly significant (Granger and Newbold 1974; Phillips 1986).

This problem can be avoided by determining if there is a stationary linear combination of non-stationary climate variables, or cointegration of time series. Analysing multiple time series that are cointegrated can provide important information about system behaviour over long time periods. We used the ARDL cointegration model, which allowed us to test for the existence of long-term climate–fire relationships, irrespective of whether the underlying regressors require no differencing to obtain a stationary series (I(0)), require one order of differencing (I(1)), or are mutually cointegrated (Pesaran et al. 2001; Oteng-Abayie and Frimpong 2006). Model fitting was performed using Microfit 5.5 (Pesaran and Pesaran 2010; Pesaran 2015).

The ARDL model was originally proposed by Pesaran et al. (1999, 2001). The ARDL methodology is graphically illustrated in Fig. 4. According to several authors (e.g. Pesaran et al. 2001; Haug 2002; Oteng-Abayie and Frimpong 2006), the ARDL methodology is preferable to test cointegration when there is a single long-run relationship among the response and covariates and there are no long-run relations in the conditional error correction form of the covariates in terms of the response. In other words, there is a reduced form on which all underlying series are weakly exogenous and taken as explanatory variables. Among these explanatory variables, there can be a mix of I(0) and I(1) series, and one or more cointegrated relations. It is important to note that in a case where one of the time series in question is integrated of order ≥2, the ARDL method cannot be applied.
In the case of our study, some variables were demonstrated to be stationary, whereas some of them required one order of differencing. In addition, we were interested in the impact of the climate variables on grassland burned area through a single equation, and therefore, the ARDL cointegration approach was the most pertinent econometric model to use in order to estimate long-run and short-run dynamics from the error correction model (ECM) specification. We also selected this methodology.
Fig. 3. Spatial pattern of grassland area burned between 2001 and 2014 in Xilingol League, China. The distribution of area burned varied from the annual averages for the entire study region. Area burned is expressed in deviations from the 14-year mean. Red areas indicate area burned greater than two standard deviations from the mean, orange between one and two standard deviations greater than the mean, yellow within one standard deviation of the mean, green between one and two standard deviations less than the mean, and purple more than two standard deviations less than the mean.
because ARDL modelling performs better with a smaller sample size than other cointegration techniques (Haug 2002; Abdullah Yusof and Duasa 2010). For instance, the ARDL methodology was shown to achieve good performance in measurement of environmental quality related indicators (such as sandy desertification and deforestation) (Ahmed et al. 2015; Ge et al. 2016).

As a first step (see Fig. 4), we used augmented Dickey–Fuller (ADF) (Dickey and Fuller 1979) and Phillips–Perron (PP) (Phillips 1987) unit root tests to test the order of integration for each variable. ARDL bounds testing requires that all variables should be integrated at purely order I(0), purely order I(1) or mutually cointegrated (Oteng-Abayie and Frimpong 2006). Therefore, it is necessary to test the integrating order of all variables before applying ARDL bounds testing; otherwise, the calculation of the $F$ statistic of ARDL becomes invalid (Baum 2004).

The general formulation of an ARDL $(n,m)$ model with a lag $m$ for exogenous variables, $\mathbf{X} = (x_1, x_2, \ldots, x_p)$, and lag $n$ for response variable $Y$ is:

$$Y_t = z_0 + \sum_{i=1}^{n} a_i Y_{t-i} + \sum_{i=0}^{m} c_i X_{t-i} + U_t,$$  

(1)

where $c_t = (c_{11}, c_{12}, \ldots, c_{1p})$ are the effects of the level and lagged exogenous variables on the response $Y$ at time $t$. Coefficients $a_1, \ldots, a_n$ represent the autoregressive effect of past realisations of $Y$, $z_0$ is the intercept term and $U_t$ is the white noise error. This model can be reformulated as an error correction model (ECM) considering the marginal changes with the lag $-1$ difference operator (with $\Delta Y_t = (Y_t - Y_{t-1})$, and $\Delta \mathbf{X}_t = (\mathbf{X}_t - \mathbf{X}_{t-1})$), as:

$$\Delta Y_t = z_0 + \sum_{i=1}^{n-1} a_i \Delta Y_{t-i} + \sum_{i=0}^{m-1} \theta_i \Delta \mathbf{X}_{t-i} + \delta \Delta \mathbf{ECM}_{t-1} + \epsilon_t,$$  

(2)

where, in the case of cointegration, $\delta$ represents the speed of adjustment to long-run equilibrium deviations. That is, the error correction term ($\Delta \mathbf{ECM}$) specifies the magnitude of the observed divergence from the equilibrium equation between $Y_t$ and $\mathbf{X}_t$ that should be corrected in each following period, and $\delta$ determines how much time it will take to come again to the long-run equilibrium position. In order to estimate the parameters $z$, $\theta$, and $\delta$, that correspond to the short-run dynamics, the unobservable variables $\Delta \mathbf{ECM}_t$ are estimated from the long-run parameters, and for a significant ECM, the parameter $\delta$ should be negative.

Fig. 4. Flowchart for an implementation of the autoregressive distributed lag (ARDL) model showing the preliminary tests required before applying the model to the dataset. The order of integration is represented by $d$, and SIC is the Schwarz information criterion.
We natural-log-transformed response and predictor variables, which allowed us to report model coefficients as ‘elasticities’, or measures of how a variable changes in response to another variable. Unlike partial slope coefficients, which represent the unit change in the response variable for a unit change in the predictor variable, coefficients in the log–log model represent the estimated percentage change in the dependent variable for a percentage change in the independent variable. If cointegration exists, the long-run relation of the response with the covariates, in the particular case of our study, can be written as:

\[
\ln \text{Area burned}_t = \beta_0 + \beta_1 \ln \text{Humidity}_t + \beta_2 \ln T_{\min,j} + \beta_3 \\
\ln T_{\max,j} + \beta_4 \ln \text{Precip}_t + \beta_5 \ln \text{Sunlight}_t + \beta_6 \ln \text{Wind}_t + \epsilon_t
\]

where \( \epsilon_t \) is a zero mean stationary process and grassland area burned is in units of hectares, \( T_{\min} \) and \( T_{\max} \) in degrees Celsius, Precip in millimetres, Humidity in percentage, Wind in metres per second, and Sunlight in hours.

For testing cointegration, the conditional error correction model in Eqn 2 is expressed as:

\[
\Delta \ln \text{Area burned}_t = \delta_0 + \sum_{i=1}^{m-1} \delta_i \Delta \ln \text{Area burned}_{t-i} + \sum_{i=0}^{m-1} \theta_i \Delta \ln \text{Humidity}_{t-i} + \sum_{i=0}^{m-1} \theta_i \Delta \ln T_{\min,j-i} + \sum_{i=0}^{m-1} \theta_i \Delta \ln \text{Precip}_{t-i} + \sum_{i=0}^{m-1} \theta_i \Delta \ln \text{Sunlight}_{t-i} + \sum_{i=0}^{m-1} \theta_i \Delta \ln \text{Wind}_{t-i} + \delta_1 \Delta \ln \text{Area burned}_{t-1} + \delta_1 \ln \text{Humidity}_{t-1} + \delta_1 \ln T_{\min,j-1} + \delta_3 \ln \text{Precip}_{t-1} + \delta_3 \ln \text{Sunlight}_{t-1} + \delta_6 \ln \text{Wind}_{t-1} + \epsilon_t
\]

The parameters \( \delta \) and \( \theta \) represent the short-run dynamics of the model. For each of the covariates, the long-run parameter \( \beta_j \) is estimated from Eqn 4 as \( \beta_j = \delta_j \beta_0 \) for \( j = 1, 2, \ldots, 6 \). Parameterisation of the model in Eqn 4 is preferred in order to implement the bound test for cointegration. Therefore, the null hypothesis of no long-run relationship between the variables is \( H_0 : \delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = \delta_6 = 0 \) and is checked against the alternative hypothesis \( H_1 \): at least one \( \delta_j \neq 0, j = 0, 1, \ldots, 6 \). Note that this test allows for degenerate long-run relations in the model, as stated in the alternative hypothesis.

Parameters in Eqn 4 can be estimated by ordinary least-squares (OLS) and the test statistics correspond to the partial \( F \) statistics for simultaneous significance of the lagged level of the variables. Narayan (2005) provides critical boundaries for the asymptotic distribution of the \( F \) statistic appropriate for small sample sizes. If the computed \( F \) statistic is greater than Narayan’s upper critical boundary, then we conclude there is cointegration between the variables. If the computed \( F \) statistic is less than the lower critical boundary, then there is no cointegration among the variables. We used the Akaika information criterion, Schwarz Bayesian criterion and Hannan–Quinn information criterion to determine the optimal lag selection (Schwarz 1971; Hannan and Quinn 1978; Akaika 1981).

After confirming long-run relationship among the variables, we applied the cumulative sum (CUSUM) and cumulative sum of squares (CUSUMSQ) tests developed by Brown et al. (1975). These tests are used to check the goodness of fit for ARDL as suggested by Pesaran et al. (2001). These tests are incorporated on the residuals of the error correction model and display results in graphical form. Plots of CUSUM and CUSUMSQ that fall within the 5% critical band demonstrate model stability.

Area burned increased over time and the spatial pattern of fire within Xilingol varied over time (Fig. 3). The ARDL model ultimately demonstrated both long-run and short-run effects of climate on area burned. To provide an easy to interpret demonstration of the effect of climate signals on area burned at different temporal scales, we calculated correlation coefficients of climate parameters and area burned for three different time periods: 2001 to 2005, 2006 to 2010, and 2010 to 2014.

**Results**

**Unit root tests**

The results of the unit root test are shown in Table 1. In the ADF test, \( \ln T_{\min} \) and \( \ln \text{Wind} \) were stationary at a 5% significance level, whereas in the PP test, \( \ln T_{\max} \) and \( \ln \text{Precip} \) were stationary at a 10% significance level. However, all variables became stationary after considering the first difference, which is confirmed by most of our unit root test approaches. Thus, they are indicated at order 1. The absence of a (1,2) variable corroborates the application of the ARDL bounds-testing technique.

**Bound testing for cointegration analysis**

Before testing for cointegration, it is important to select the correct number of lags in \( y(n) \) considered in the ARDL model in order to avoid misspecification problems (Pesaran et al. 2001). The use of one lag in the model was indicated by most selection criteria (Table 2).

The results of the ARDL bounds testing and critical values are presented in Table 3. The calculated value of the \( F \) statistic (5.11) was greater than the upper bound values (\( F_{1,5} \)) at 5 and 10% levels of significance respectively, which demonstrates the existence of long-run relationships among climate variables irrespective of whether the covariates are I(0), I(1) or mutually cointegrated. From this result, it is possible to find long-run coefficients and short-run dynamics from the conditional vector ARDL equation with error correction term.

**Long-run coefficients**

Once cointegration is proved, there is a linear combination of all variables, including predictors and response, that is stationary. This long-run relationship is scaled such that the coefficient of the grassland area burned is 1 (Table 4). Results are estimated from the selected ARDL model according to the Schwarz Bayesian criterion. Significant long-run effects calculated at a 5% significance level are wind, maximum average temperature and sunlight expressed in logarithmic scale. From these long-run elasticities, it is estimated that a 1% change in wind is associated with a 1.05% change in area burned. This long-run elasticity is not expressed instantaneously, but represents the long-term...
Table 1. Unit root test of grassland area burned as a function of climate variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Augmented Dickey-Fuller (ADF) test statistics</th>
<th>Phillips–Perron (PP) test statistics</th>
<th>Level of significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnArea burned</td>
<td>−6.359</td>
<td>−6.819</td>
<td>−14.001</td>
</tr>
<tr>
<td>lnRel. humidity</td>
<td>−7.133</td>
<td>−4.658</td>
<td>−24.824</td>
</tr>
<tr>
<td>lnTmin</td>
<td>3.880</td>
<td>−4.864</td>
<td>−11.821</td>
</tr>
<tr>
<td>lnTmax</td>
<td>−3.449</td>
<td>−6.053</td>
<td>−24.936</td>
</tr>
<tr>
<td>lnPrecip</td>
<td>−7.668</td>
<td>−9.874</td>
<td>−32.864</td>
</tr>
<tr>
<td>lnSunlight</td>
<td>−9.612</td>
<td>−7.826</td>
<td>−38.326</td>
</tr>
<tr>
<td>lnWind</td>
<td>−5.331</td>
<td>−8.667</td>
<td>−27.942</td>
</tr>
</tbody>
</table>

A5% level of significance.
B10% level of significance.

Table 2. Selection criteria of lag order of variables for the autoregressive distributed lag (ARDL) approach

<table>
<thead>
<tr>
<th>Lag</th>
<th>Log L</th>
<th>LR</th>
<th>FPE</th>
<th>AIC</th>
<th>SBC</th>
<th>HQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>138.86</td>
<td>NA</td>
<td>8.01 $\times 10^{13}$</td>
<td>−9.11</td>
<td>−8.91</td>
<td>−4.11</td>
</tr>
<tr>
<td>1</td>
<td>123.92</td>
<td>270.51</td>
<td>1.31 $\times 10^{13}$</td>
<td>−16.14</td>
<td>−14.64</td>
<td>−21.71</td>
</tr>
<tr>
<td>2</td>
<td>119.50</td>
<td>67.25</td>
<td>6.22 $\times 10^{12}$</td>
<td>−15.21</td>
<td>−14.11</td>
<td>−21.27</td>
</tr>
<tr>
<td>3</td>
<td>104.54</td>
<td>93.09</td>
<td>8.1 $\times 10^{11}$</td>
<td>−18.22</td>
<td>−17.41</td>
<td>−21.04</td>
</tr>
</tbody>
</table>

AIndicates lag order selected by the criterion.

Table 3. Results of the Wald test of cointegration following Pesaran et al. (2001)

<table>
<thead>
<tr>
<th>F statistic</th>
<th>Degrees of freedom</th>
<th>Critical value</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.11</td>
<td>(7,37)</td>
<td>5%</td>
<td>3.14</td>
<td>4.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10%</td>
<td>2.76</td>
<td>4.65</td>
</tr>
</tbody>
</table>

Table 4. Estimated long-run coefficients for the effect of climate variables on grassland area burned from 2001 to 2014 using an autoregressive distributed lag (ARDL) model

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Coefficient</th>
<th>s.e.</th>
<th>T ratio [Probability]</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnWind</td>
<td>1.045$^A$</td>
<td>0.267</td>
<td>3.902[0.001]</td>
</tr>
<tr>
<td>lnTmax</td>
<td>1.977$^A$</td>
<td>0.532</td>
<td>3.695[0.001]</td>
</tr>
<tr>
<td>lnTmin</td>
<td>0.988</td>
<td>0.983</td>
<td>0.985[0.327]</td>
</tr>
<tr>
<td>lnRel. humidity</td>
<td>0.387</td>
<td>0.544</td>
<td>0.711[0.479]</td>
</tr>
<tr>
<td>lnPrecip</td>
<td>−8.669</td>
<td>8.758</td>
<td>0.989[0.325]</td>
</tr>
<tr>
<td>lnSunlight</td>
<td>1.977$^A$</td>
<td>0.533</td>
<td>3.695[0.001]</td>
</tr>
</tbody>
</table>

Diagnostic tests

<table>
<thead>
<tr>
<th>Test</th>
<th>Probability Chi-square value (P value)</th>
<th>Probability Chi-square value (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial correlation</td>
<td>3.421 (0.154)</td>
<td>1.004 (0.371)</td>
</tr>
<tr>
<td>Functional form Ramsey RESET test</td>
<td>2.131 (0.121)</td>
<td>1.418 (0.152)</td>
</tr>
<tr>
<td>Normality</td>
<td>0.761 (0.681)</td>
<td>Not applicable</td>
</tr>
<tr>
<td>Heteroscedasticity</td>
<td>5.649 (0.987)</td>
<td>0.281 (0.971)</td>
</tr>
</tbody>
</table>

$^A$5% significance level.
recovery to the stationary equilibrium. Similarly, a 1% change in average maximum temperature and sunlight are expected to produce a 1.98% change in area burned.

The reliability of the long-run results depends on an accurate specification of the ARDL model that meets all assumptions of this method. We performed several tests to validate the model. First, we checked the functional form of the model by applying Ramsey’s RESET test (Ramsey 1969). The results demonstrate the appropriateness of the functional form of the model both in the linear model (LM) version and $F$-version of diagnostic checks. We also performed Durbin–Watson tests, which indicated no first-order serial correlation ($P = 0.154$). Diagnostics also revealed no heteroscedasticity ($P = 0.987$). In addition, model residuals were shown to follow a normal distribution. Tests results are presented in the lower part of Table 4.

**ARDL error correction model**

Short-run coefficients from the ECM are displayed in Table 5. A 1% increase in sunlight (which can be interpreted as fires occurring closer to 21 June, the period of maximum sunlight) was associated with a 0.27% increase in area burned. A 1% increase in wind speed was associated with a 0.13% increase in area burned. A 1% change in relative humidity was associated with a 0.18% change in area burned. The lagged ECM term is negative as expected and statistically significant, implying that 60% of changes in grassland area burned were explained by short-run deviations from the long-run equilibrium. The relationship between area burned and different climate variables is shown in Fig. 5.

To check the stability of the ECM, we incorporated CUSUM and CUSUMSQ tests. The results of the tests are presented in graphical form in Fig. 6. These plots demonstrate that the CUSUM and CUSUMSQ lines are within the critical band of 5% significance level over time, which confirms the stability of the model.

Correlation analysis (shown in Table 6) demonstrates that simple models of grassland burned area produced more significant results when dividing the study period (2010–14) into three parts, 2001 to 2005, 2006 to 2010 and 2011 to 2014. Interestingly, we found that significance of some climate variables changed over these three periods. For example, during 2011 to 2014 sunlight was highly correlated with the area burned, but over the remaining periods, it was not.

**Table 5. Error correction representation for the autoregressive distributed lag (ARDL) model**

ARDL (1,1,1,1,0,1) was selected based on the Schwarz Bayesian criterion. The dependent variable was ln (grassland area burned) with 564 observations over the period 2001 to 2014.

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Coefficient</th>
<th>s.e.</th>
<th>T ratio</th>
<th>[Probability]</th>
</tr>
</thead>
<tbody>
<tr>
<td>D ln Sunlight</td>
<td>0.268 $^{A}$</td>
<td>0.051</td>
<td>4.103</td>
<td>[0.001]</td>
</tr>
<tr>
<td>D ln Tmax</td>
<td>0.008</td>
<td>0.013</td>
<td>0.121</td>
<td>[0.542]</td>
</tr>
<tr>
<td>D ln Tmin</td>
<td>0.051</td>
<td>0.031</td>
<td>1.681</td>
<td>[0.104]</td>
</tr>
<tr>
<td>D ln Rel.humidity</td>
<td>0.189 $^{A}$</td>
<td>0.087</td>
<td>2.157</td>
<td>[0.033]</td>
</tr>
<tr>
<td>D ln Precip</td>
<td>0.001</td>
<td>0.018</td>
<td>0.421</td>
<td>[0.946]</td>
</tr>
<tr>
<td>D ln Wind</td>
<td>0.127 $^{A}$</td>
<td>0.046</td>
<td>2.744</td>
<td>[0.001]</td>
</tr>
<tr>
<td>ECM (-1) / C0</td>
<td>$-0.601^{A}$</td>
<td>0.124</td>
<td>$-4.8[0.000]$</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.783</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.714</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s.e. of regression</td>
<td>0.065</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum squared residuals</td>
<td>0.131</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durbin–Watson-statistic</td>
<td>2.082</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$A$5% level of significance.

**Fig. 5.** Relationship between area burned and climate variables between 2001 and 2014 in Xilingol League, China. Area burned is in red and corresponds to the left $y$ axis. Climate variables correspond to the right $y$ axis. Vertical grey lines divide 2001–14 into three time periods (see text).
The significance of precipitation also varied during different time periods.

**Discussion**

Time series climate data with different temporal properties present several challenges for researchers. In the case of the present study, some climate time series data used to explain grassland area burned were stationary, whereas other data were stationary after differencing. ADRL bound testing methodology is well suited for these mixed time series orders. In addition, given that no variable had integration order larger than 1, and that cointegration in a single equation including the response grassland burned area was proved, it was possible to reparametrise the model as an ECM. With this type of model, relevant parameters can be estimated in a unique methodological framework. In particular, the elasticities (long-run coefficients), the short-run dynamics and the tendency of the series to recover the long-run stationary level explain the relevant information about the relations between the climate covariables and the response.

Absent cointegration among the grassland area burned and climate variables, regression analysis would have resulted in spurious relations with no reliable interpretation. Regression analysis of climate–fire relationships is only appropriate with integrated time series when there exists a linear combination of the variables at a level (same time) that behaves as a stationary process. Cointegration matches the order of non-stationarity of different variables in a form that combined with a specific linear equation produces a new predictive series. In addition, the same coefficients of this combination represent the long-run equilibrium relations that give important information about the effects on the response variable.

ARDL analysis revealed that maximum temperature is the most important variable affecting grassland fires in Xilingol (Table 4), a finding consistent with other studies in northern China and worldwide that show a relationship between higher temperatures and increased fire activity (Kobayashi et al. 1994; Su and Liu 2004; Hardy 2005; Kaloudis et al. 2005; Zhijun et al. 2009; Chen et al. 2008). Guo et al. (2016) found that as the temperature increased, the amount of boreal forest burned in China increased, even when increased temperatures were accompanied by higher than normal precipitation. Wind and sunlight also explain significant long-term changes in area burned.

Climate influences on grassland area burned in the eastern Eurasian steppe may vary given the temporal scale examined. Liu et al. (2017) analysed climate influences on daily fire progression in Hulun Buir, a region of the Inner Mongolia Autonomous Area north-east of Xilingol. They found positive correlations between grassland area burned and higher temperature, wind speed and sunlight, and negative correlations with precipitation and relative humidity. At the monthly and annual

---

**Table 6. Correlation between grassland area burned and climatic variables from 2001 to 2014**

<table>
<thead>
<tr>
<th>Time period</th>
<th>$T_{\text{max}}$</th>
<th>$T_{\text{min}}$</th>
<th>Precip</th>
<th>Rel. humidity</th>
<th>Wind</th>
<th>Sunlight</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001–05</td>
<td>0.718</td>
<td>0.011</td>
<td>0.036</td>
<td>0.031</td>
<td>0.461</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.311</td>
<td>0.189</td>
<td>0.303</td>
<td>0.001</td>
<td>0.301</td>
</tr>
<tr>
<td>2006–10</td>
<td>0.699</td>
<td>−0.108</td>
<td>0.041</td>
<td>0.004</td>
<td>0.791</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.186</td>
<td>0.911</td>
<td>0.541</td>
<td>0.001</td>
<td>0.219</td>
</tr>
<tr>
<td>2011–14</td>
<td>0.598</td>
<td>0.019</td>
<td>−0.003</td>
<td>0.008</td>
<td>0.684</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.416</td>
<td>0.727</td>
<td>0.941</td>
<td>0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>

---

**Fig. 6.** Plots of the cumulative sum of recursive residuals for the autoregressive distributed lag (ARDL) model for short run (a), and long run (b) showing model stability. Straight lines represent the 95th percentile confidence intervals.
time steps analysed in the present study, we found precipitation to be less influential on fire extent. But our study and the findings of Liu et al. suggest that temperature exerts a powerful influence on fire behaviour at every spatial and temporal scale examined to date.

Directional temperature increases are likely to significantly increase winter temperatures relative to summer temperatures, and warmer winter and spring temperatures are likely to extend fire seasons (Westerling et al. 2006; Deser et al. 2012). Increases in area burned in Xilingol in May through October were associated with higher temperatures, but there were weak correlations between climate signals and area burned in February, March and April, suggesting that a threshold for increased winter and early spring fire activity in the eastern Eurasian steppe has not yet been reached.

Previous research in China has not incorporated wind into fire models (Yang et al. 2008; Zhang et al. 2010; Li et al. 2017). Wind played a key role in increasing grassland area burned in Xilingol, similarly to results from other regions (Moritz et al. 2010; Rolinski et al. 2016; Keeley and Syphard 2017). Higher wind speeds in the early season (e.g. February, $R^2 = 0.018, P > 0.012$) can outweigh the effect of low temperatures. From 2001 to 2014, there were four monthly peaks in wind speed (May, August, September and October), with the highest peak in May. We found that the number of extreme events that occurred in 2003, 2004 and 2014 was associated with higher area burned in the northern regions ($R^2 = 0.810, P < 0.002$).

Climate change and human factors are likely to continue to transform grasslands in northern China. Climate projections for north-eastern China predict higher temperatures along with lower precipitation and relative humidity Niu and Zhai (2012). Liu et al. (2012) also indicated that the expected change in population density is correlated with temperature as well as the change in precipitation. However, the effects of altered human ignitions cannot be easily predicted (Guo et al. 2016). Given the natural range of variation in this region, neither precipitation, relative humidity nor minimum temperature affected grassland area burned. But wind speed was a key factor in grassland area burned in northern areas, and the ARDL methods identified both long-term and short-term sensitivities.

Conclusion

Both climate change and human factors are likely to continue to transform grasslands in northern China. Climate projections for north-eastern China include higher temperatures along with lower precipitation and relative humidity (Niu and Zhai 2012). Projected changes in the wind regime are as yet unknown. Increases in population density are likely to occur in conjunction with increasing temperature as well as changes in precipitation patterns (Liu et al. 2012). However, the effects of altered human ignitions cannot be easily predicted (Guo et al. 2016).

The ARDL method identified both long-term and short-term sensitivities, showing that it can be a valuable tool in investigating trends in area burned, even in the presence of high variability and short periods of climatological data. In particular, the present study found that prevailing winds were an important determinant of area burned and also suggests that the wind regime may not be stationary.

Conflicts of interest

The authors declare they have no conflicts of interest.

Acknowledgements

This research was funded by the National Natural Science Foundation of China under grant no. 41301584, the National Key Technology R&D Program of China under grant nos 2013BAK05B01 and 2013BAK05B02, and the Youth Foundation of Jilin province (no. 20140520156H), and the Program for Introducing Talents to Universities (B16011). We are grateful to the associate editor and two anonymous reviewers for extremely helpful comments and suggestions that significantly improved the manuscript.

References

Francesco M, Ubeda X, Tort J, Panareda JM, Cerdá A (2016) The role of forest fire severity on vegetation recovery after 18 years. Implications for forest management of Quercus suber L. in Iberian Peninsula.
ARDL model for sensitivity to climate change


