Modelling Biological Soil Crust based on Multiple Datasets

Doctoral thesis

for

the award of the doctoral degree

of the Faculty of Mathematics and Natural Sciences

of the University of Cologne

submitted by

Yifan Li

accepted in the year 2025

Abstract

Further understanding of biological soil crust (BSC) response to climate change requires BSC-climate models, which represent the relevant processes taking place in the atmosphere and land surface. In this study, a modelling system for biological soil crust and climate factors based on multi-datasets is developed in two approaches. The effects of climate variability on the long temporal and large spatial distribution of BSC are revealed by an improved BSC detection method and multiple linear regression. The models can be used to explain the dominant climatic factors associated with BSC changes. the short-term or long-term forecasts of regional-scale distribution of BSC, the assessment for the potential effects of climate change on the availability of BSC and the sustainable development of ecosystem, as well as the short-term or long-term forecasts of regional-scale distribution of BSC.

The long-time and large-scale distribution of biological soil crust is obtained in the study area. To this end, this study is divided into the following four steps: 1) Fusion of MODIS and Landsat7 satellite data using the Sspatial and Temporal Adaptive Reflectance Fusion Model (STARFM) to obtain multispectral data with high spatial and temporal resolution; 2) Calculation of the BSC Index (BSCI) and the NDVI from the fused satellite spectral data; 3) Extraction of the BSC for the study area based on the BSCI thresholds obtained from previous studies, as well as considered with NDVI. 4) Analyzing the extracted BSC data from multiple perspectives. The analysis for 19a shows that on the time scale, the BSC variations have an interannual periodicity, peaking in March and October of each year, and almost zero in winter. On the spatial scale, the BSC is mainly distributed in the desert-oasis transition zone, while the distribution become gradually sparse toward to the desert hinterland.

Lag-correlation and partial correlation between BSC and climate variables is analyses. In this study, five climatic variables (specific humidity, 10-meter wind speed, 2-meter temperature, surface solar radiation and total precipitation) and their time lags were

used as independent variables. The results show that in some areas the BSC is more strongly correlated with time-lagged climate factors when the time lag is taken into account, and this is most evident for specific humidity. The response of the BSC to this is usually delayed by 1 to 2 months. In principle, the time lag between the BSC and the climatic variables does not exceed three months. The BSC responds quickly to temperature, with a correlation coefficient of 0.7. The BSC also responds quickly to precipitation, while the correlation coefficient is relatively low at 0.46 and the significantly correlated areas are mainly in the east and south. These correlation analyses provide a good reference for the selection of variables for subsequent modelling.

The models of biological soil crust and climate factors is constructed using two approaches, in which the influence of time lag is considered. One approach is based on fixed climate factors, and the other slides over the time series to select more appropriate climate factors and coefficients for different time points. Multiple regression analysis is applied to both models. Statistical parameters are used to estimation. The results shows that the two approaches can explain about 40% and 75% of the BSC, respectively. Then applied models to paleoclimate (Last Glacial Maximum and Mid-Holocene) in the Gurbantunggut Desert and to historical climate in the Atacama Desert. Changes in biological soil crust during different time periods are also compared and analyzed.

In summary, the long-temporal and large-spatial distribution of BSC is obtained. Benefiting from it, the correlation between BSC and climatic factors is analyzed. And the model system developed captures well the climatological processes in the study area. The BSC-climate model can appropriately predict the BSC in paleoclimate and indicate the its response to the climate variables.

Zusammenfassung

Um ein besseres Verständnis der Reaktion biologischer Bodenkrusten (BSC) auf den Klimawandel zu gewinnen, sind BSC-Klima Modelle erforderlich, die die relevanten Prozesse in der Atmosphäre und an der Landoberfläche abbilden. In dieser Studie wird ein Modellierungssystem für biologische Bodenkrusten und Klimafaktoren auf der Grundlage von Multi-Datensätzen mit zwei Ansätzen entwickelt. Durch eine verbesserte BSC-Erkennungsmethode und eine multiple lineare Regression werden die Auswirkungen der Klimavariabilität auf die langfristige und großräumige Verteilung der BSC aufgezeigt. Diese Modelle können verwendet werden, um die dominierenden Klimafaktoren, die Veränderungen der BSC beeinflussen, zu identifizieren, kurz- und langfristige Prognosen zur regionalen Verteilung der BSC zu erstellen, die potenziellen Auswirkungen des Klimawandels auf die Verfügbarkeit der BSC und die nachhaltige Entwicklung des Ökosystems zu bewerten sowie regionale Prognosen der BSC-Verteilung für unterschiedliche Zeiträume zu ermöglichen.

In der Studie wird die langfristige und großräumige Verteilung der biologischen Bodenkrusten im Untersuchungsgebiet bestimmt. Dazu wird die Studie in die folgenden vier Schritte unterteilt: 1) Fusion von MODIS- und Landsat7-Satellitendaten unter Verwendung des Sspatial and Temporal Adaptive Reflectance Fusion Model (STARFM), um multispektrale Daten mit hoher räumlicher und zeitlicher Auflösung zu erhalten; 2) Berechnung des BSC Indikator(BSCI) und des NDVI aus den fusionierten Satellitenspektraldaten; 3) Extraktion der BSC im Untersuchungsgebiet auf Basis der aus früheren Studien ermittelten BSCI-Schwellenwerte sowie unter Berücksichtigung des NDVI; 4) Analyse der extrahierten BSC-Daten aus verschiedenen Perspektiven. Die Analyse der 19-jährigen Daten zeigt, dass die BSC-Veränderungen auf der Zeitskala eine jährliche Periodizität aufweisen, mit Spitzenwerten im März und Oktober eines jeden Jahres und nahezu null im Winter. Auf

der räumlichen Skala ist die BSC hauptsächlich in der Übergangszone zwischen Wüste und Oase verteilt, wobei die Verteilung in Richtung Wüsteninneres allmählich spärlicher wird.

Die Verzögerungskorrelation und Partialkorrelation zwischen BSC und Klimavariablen wird analysiert. In dieser Studie werden fünf Klimavariablen (spezifische Feuchtigkeit, Windgeschwindigkeit in 10 Metern Höhe, Temperatur in 2 Metern Höhe, Oberflächensonnenstrahlung und Gesamtniederschlag) sowie deren Zeitverzögerungen als unabhängige Variablen verwendet. Die Ergebnisse zeigen, dass die BSC in einigen Gebieten stärker mit zeitverzögerten Klimafaktoren korreliert ist, wenn die Zeitverzögerung berücksichtigt wird, was bei der spezifischen Feuchtigkeit besonders deutlich ist. Die Reaktion der BSC auf die spezifische Feuchtigkeit erfolgt in der Regel mit einer Verzögerung von 1 bis 2 Monaten. Grundsätzlich überschreitet die Zeitverzögerung zwischen der BSC und den Klimavariablen nicht drei Monate. Die BSC reagiert schnell auf Temperatur mit einem Korrelationskoeffizienten von 0,7. Auch auf Niederschlag reagiert die BSC schnell, wobei der Korrelationskoeffizient jedoch relativ niedrig ist 0,46 und die signifikant korrelierten Gebiete hauptsächlich im Osten und Süden liegen. Diese Korrelationsanalysen liefern eine gute Grundlage für die Auswahl der Variablen für die nachfolgende Modellierung.

Die Modelle für biologische Bodenkruste und Klimafaktoren werden unter Berücksichtigung der Zeitverzögerung mit zwei Ansätzen entwickelt. Ein Ansatz basiert auf festen Klimafaktoren, der andere gleitet über die Zeitreihe, um besser geeignete Klimafaktoren und Koeffizienten für verschiedene Zeitpunkte auszuwählen. Beide Modelle verwenden multiple Regressionsanalysen, die durch statistische Parameter bewertet werden. Die Ergebnisse zeigen, dass die beiden Ansätze etwa 40 % bzw. 75 % der BSC-Veränderungen erklären können. Anschließend werden die Modelle auf das Paläoklima (Letzteiszeitliches Maximum und Mittelholozän) in der Gurbantunggut-Wüste und auf das historische Klima in der Atacama-Wüste angewendet. Veränderungen der biologischen Bodenkrusten in verschiedenen

Zeiträumen werden ebenfalls verglichen und analysiert.

Zusammenfassend wird die langfristige und großräumige Verteilung der BSC bestimmt. Auf dieser Grundlage wird die Korrelation zwischen BSC und Klimafaktoren analysiert. Das entwickelte Modellsystem erfasst die klimatischen Prozesse im Untersuchungsgebiet gut. Das BSC-Klima Modell kann die BSC im Paläoklima angemessen vorhersagen und ihre Reaktion auf Klimavariablen aufzeigen.

Contents

1 Introduction

Biological soil crust (BSC) is formed by soil microorganisms, algae, lichens, mosses plants and soil forming an organic complex (Belnap et al., 1994; West, 1990). It is a common phenomenon in the desert and semi-desert areas of the world, and its formation makes the soil surface obviously different from loose sandy soil in physical, chemical and biological properties, with strong wind erosion resistance and important ecological and geological effects, which has become an important basis for the succession of vegetation in desert areas (Zhang et al., 2005). Biological soil crusts are important structure in the topological succession of soil in desert and semi-desert areas, which has a significant effect on the improvement of soil erosion resistance, and is also the first sign of sand fixation (Hu et al., 2000). Biological crusts can grow and reproduce in poor conditions and influence and change the environment through their own activities, which is of great importance and irreplaceable ecological significance (Chen et al., 2003).

Remote sensing optical images (visible (VIS) to shortwave infrared (SWIR)) from space-borne sensors have been widely used to monitor terrestrial ecosystem functions due to their synoptic coverage of the land surface at fixed intervals. However, monitoring drylands from satellite platforms has been quite challenging because arid and semi-arid regions are usually only sparsely vegetated, and the observed spectral signal is a mixture of soil, biocrusts and vascular plants (annuals and perennials) (Rozenstein & Adamowski, 2017; Weber & Hill, 2016). Remote sensing images at high spectral, temporal and spatial resolution are required to accurately map the spatial distribution of the different dryland components from space (Karnieli et al., 1996, 1999, 2002; Zaady et al., 2007). High spectral resolution is helpful to separate the biocrust unique spectral features from that of bare soils (Karnieli & Tsoar, 1995; Rozenstein & Karnieli, 2015). A high temporal resolution can help in separating different land covers because biocrusts and vascular plants have different phenological cycles. A high spatial

resolution is expected to reduce spectral mixing effects in the VIS and infrared regions (Qin et al., 2006), thus improving the characterization of the biocrust spatial distribution when their fractional cover is low. Ground-based spectral measurements and Normalized Difference Vegetation Index (NDVI)-derived values have also been used to link semi-arid ecosystem phenology to biocrust CO2 fluxes in order to assess the capability to detect biocrust activity from satellite (Burgheimer et al., 2006a, 2006b). Therefore, using suitable spectral indices and satellite images to detect BSC from background (rocks, bare soils or sand dunes) and map their distribution is critical to quantify biological crusts interactions to ecosystem.

The spatial and temporal distribution of biological crusts contributes to the understanding of the evolution trend of the ecological environment in desert areas and its response to global changes. The distribution of biological crusts has selective characteristics at different scales. At the landscape scale, well-developed biological soil crusts are found in arid desert areas such as Ordos, Shapotou and Junggar in China. On the other hand, in the Taklamakan Desert, which is also a temperate desert area, there is no distribution of biological soil crusts. At the regional scale, as in the Junggar Basin, the cover and distribution of biological soil crusts is much higher in the south than in the central and northern regions (Zhang, 2005; Zhang et al., 2007). At the dune scale, the cover and variety of biological soil crusts varied along the top, mid-slope, bottom and base of the interdune line (Zhang et al., 2004; Chen et al., 2005, 2007). At the microscale, for instance, at the meter and centimeter scales of homogeneity, lichen and moss crusts are also distributed in discontinuous patches (Boeker et al., 2006). However, current research lacks the study of the distribution of BSC over long temporal and spatial scales. Hence, it is important to study the characteristics and patterns of the spatial and temporal distribution of biological crusts.

Biological soil crust is a major component of desert ecosystems. The anatomical structure of biological soil crusts is very simple, which is extremely sensitive to external

disturbances and changes in environmental conditions. It is able to manifest the impact of climate change far ahead of vascular plants in desert ecosystems, and the most sensitive indicator organisms in desert ecosystems of environmental and climate change (Conti et al., 2001; Wu et al., 2002, 2003). As the earliest and most susceptible surface system in desert ecosystems, biological soil crust microcosms are ideal information carriers for studying the impacts of global change on desert ecosystems. At global and regional scales, the distribution of biological soil crusts shows a positive correlation with annual precipitation (Belnap et al., 2007), condensation and soil moisture content (Schieferstein & Loris, 1992; Jacobs et al., 1999, 2000); the effect of temperature on the distribution of biological soil crusts varies according to the species that make up the biological soil crusts (Spier & Van Herk, 1997; Van Herk et al., 2002). On a small scale, the distribution of biological soil crusts is limited by soil type, texture and nutrients (Eldridge, 1996; Bowker et al., 2005). Therefore, the selection of climatic factors, such as moisture and temperature, to carry out the changes in the distribution pattern of biological soil crusts under different climatic factors and multi-factor coupling conditions is an important issue to be addressed.

This study aims at developing a modelling system for biological soil crusts and climate factors based on multi-datasets, with improved BSC detection method and machine learning, to reveal the long-term and large spatial distribution of BSC and its interaction with the climate variability. This model can be used to explain the effects of climatic factors on biological crusts, the short-term or long-term forecasts of regional-scale distribution of BSC, the assessment for the potential effects of climate change on the availability of BSC and the sustainable development of ecosystem.

The objectives of this study are:

1. To obtain the long-time and large-scale distribution of biological soil crusts based on satellite data and improved crust detection method, and to analyze their changing patterns and characteristics.

2. To develop BSC-climate models using machine learning techniques and to evaluate the performance of this model;
3. To understand how the climate factors in the model affect the variability of biological soil crusts.
4. To apply the models to paleoclimate in Gurbantunggut Desert and modern climate in Atacama Desert.

This thesis is divided into 6 chapters. In Chapter 2, a relevant research review and motivation are presented. In Chapter 3, the spatiotemporal distribution of biological soil crusts described along with the introduction of the study area and data processing. In Chapter 4, a model of biological soil crusts and climate factors is constructed, in which the influence of time lag is considered. Statistical parameters are used to evaluate and select the model. Predictions are compared with observations in Gurbantunggut Desert to determine the validity of the model. Further, the effects of climatic variables on BSC are interpreted according to the model parameters. In Chapter 5, the BSC-climate model is applied to paleoclimate (Last Glacial Maximum and Mid-Holocene) and the Atacama Desert. Changes in biological soil crusts in different areas during different periods are also compared and analyzed. In the last chapter, the discussion, conclusion, and outlook are given.

2 Research review and motivation

2.1 Biological Soil Crust

2.1.1 Characteristics

The concept of biological crust is first mentioned by Fritsche (1907) in a study related to tropical desert areas. Vascular plants in desert areas are usually low and sparse, with large open spaces between individuals, and the amount of litter is relatively small, which has a limited role in resisting soil wind erosion in arid and semi-arid areas, and the role of BSC on soil stability is particularly important in this situation (Belnap, 2003). Numerous studies have confirmed that BSC can enhance soil stability and improve soil resistance to wind and water erosion (Eldridge, 2003; Eldridge & Leys, 2003; Chaudhary et al., 2009; Bu et al., 2015). In arid and semi-arid regions, water is one of the main factors limiting plant growth (Xiong et al., 2011), and the presence of BSC can lead to soil water redistribution (Yair, 2003), and the structure of BSC and physiological activities of microorganisms can change the aeration and permeability of the topsoil, affecting the hydrological processes of precipitation infiltration, flow production, and evapotranspiration (Warren, 2001; Belnap, 2006; Wu et al., 2002). Algae, lichens, and mosses are primary producers with carbon sequestration functions and are important contributors of organic carbon in arid ecosystems (Bowker et al. 2010; Castillo-Monroy et al., 2011). Cyanobacteria in the BSC have nitrogen fixation function, which can provide abundant nitrogen source for plant growth and contribute to soil nitrogen input in desert ecosystems (Belnap, 2002; Billings et al., 2003; Su et al., 2011). In addition, the presence of BSC also interacts with the surrounding vegetation and affects the stability of sandy ecosystems (Bowker et al., 2014; Chen, 2007). China researchers have carried out a lot of research work in the Gurbantunggut Desert, Mu Us sandy land, Ningxia Shapotou and Loess Plateau, and most of the current research is still in the stage of understanding the functions and roles of BSC (Li et al., 2009).

2.1.2 Distribution

Biological soil crusts are formed by soil particles tightly bound to algae, fungi, lichens, and mosses in varying proportions, and are a common ground cover in arid and semi-arid regions (Belnap et al., 2003; Belnap, 2006). In recent years, many studies in Chinese deserts have shown that BSC have a certain distribution in the region and plays an important role in preventing soil erosion and regulating hydrological processes (Xiong et al., 2011; Zhang et al., 2013; Wang et al., 2017).

Biological soil crusts are widely distributed across terrestrial surfaces, extending beyond arid and desert regions to semi-arid landscapes (Pickard, 1986; Blank & Camero, 1966). Their presence is largely attributed to the low vegetation cover in these ecosystems, where natural and semi-natural plant communities typically exhibit less than 40% coverage, leaving substantial open ground available for BSC colonization (Belnap et al., 1994). Additionally, the spatial heterogeneity and environmental variability across these landscapes contribute to the ecological significance of BSCs, making them a fundamental component of fragile ecosystems. The role of BSCs in ecosystem stability is particularly significant, as they enhance soil carbon and nitrogen fixation, improve fertility, and increase water infiltration capacity (Belnap & Lange, 2003). Given their ecological importance, a systematic analysis of BSC formation, development, and spatial distribution in natural environments is essential, alongside investigations into their successional dynamics and functional contributions to ecosystem processes.

Biological soil crusts have been widely studied across various habitats, with research focusing on their distribution, composition, diversity, and environmental interactions, providing key insights into their spatial patterns globally. In Australia, BSCs collected from 30 quadrats across three geomorphic types in a forest were analysed for their species composition, with a particular emphasis on lichens and bryophytes (Eldridge, 1999). In North America, research has primarily focused on semi-arid and arid

landscapes. In the Colorado Plateau, USA, BSCs have been characterized at millimetre resolution, revealing their bacterial biomass, diversity, and vertical stratification (Garcia-Pichel et al., 2003). In the Mojave Desert, USA, BSCs were mapped and analysed using GIS techniques, surface feature extraction, and spatial statistics to examine their relationships with topographic and soil properties, contributing to a conceptual model of BSC spatial distribution (Williams et al., 2013). In Latin America, studies have emphasized BSC composition and ecological roles in semi-arid environments. In Mexico's Tehuacán Valley, BSCs from 87 sample plots were examined to investigate the distribution and composition of algae, lichens, and bryophytes, highlighting their role in stabilizing desert soils (Rivera-Aguilar et al., 2006). Across southwestern Africa, research has focused on regional-scale climatic influences on BSC diversity and distribution. A large-scale study assessed BSC diversity and distribution patterns at 29 observation stations along an 1800 km climatic transect, capturing biogeographical variations across multiple climate zones (Zedda et al., 2011). Studies in the Middle East have explored BSC formation and classification based on morphological characteristics. In the Negev Desert, Israel, BSCs were categorized into different types, and their spatial distribution patterns and formation mechanisms were systematically evaluated (Kidron et al., 2010). European research has focused on the role of environmental factors in BSC successional dynamics. In central Spain, BSCs within a nature reserve were analysed to determine the effects of environmental variables on post-successional distribution patterns, highlighting their ecological importance in Mediterranean landscapes (Ochoa-Hueso et al., 2011). These studies collectively illustrate the global variation in BSC distribution and composition across diverse ecosystems.

The primary focus of biological soil crust research in China has been on the Gurbantunggut Desert, the Tengger Desert, the Kubuqi Desert, and the Mu Us Sandy Land, etc., which are situated in low-altitude desertification-prone regions in northern China (Weber et al., 2016; Li et al., 2020; X. Zhou et al., 2020; H. Zhou et al., 2020).

In the Gurbantunggut Desert, BSCs are predominantly distributed in the southern region (Zhang et al., 2007). In the Tengger Desert, BSCs beneath Hedysarum scoparium and Calligonum mongolicum were analyzed through field and laboratory investigations, revealing that crust thickness decreased with increasing distance from plant roots, further highlighting the role of vegetation in BSC formation (R. Zhang et al., 2024). In the Kubuqi Desert, induced BSCs were found to vary in distribution with slope aspect, gradient, and plant canopy, with better crust development observed on north-facing slopes and in vegetated areas (X. Zhou et al., 2020). In the Mu Us Sandy Land, BSCs within *Artemisia ordosica* communities were reported to be prevalent, with significantly lower coverage in semi-fixed sand than in fixed sand, indicating the influence of surface stability on BSC development (Zhang et al., 2010). The spatial distribution of BSCs beneath *Artemisia ordosica* was further analyzed, showing strong associations with wind direction, wind speed, and plant-root proximity (Liu et al., 2014). Beyond these desert systems, slope-associated BSCs in Liudaogou, a transitional zone between wind and water erosion in the northern Loess Plateau, were investigated, revealing that crusts develop more extensively in sandy soils with gentle slopes, higher moisture availability, lower solar radiation, and reduced erosion intensity (Bu et al., 2016).

With the advancement of research, the study of BSCs has become increasingly multidimensional, with scholars identifying key environmental factors that influence their spatial distribution at different scales. At the micro-scale, microtopography plays a critical role in BSC formation and the maintenance of community diversity (Li et al., 2010). At small to medium scales, factors such as atmospheric dust accumulation, light availability, soil moisture, and soil nutrients drive variations in BSC cover and diversity (Chen et al., 2007; Lan et al., 2015; Zhang et al., 2000; Zhang et al., 2015). At the landscape scale, precipitation serves as a dominant factor in determining the spatial distribution of BSCs, whereas at the regional scale, soil properties become the primary determinant of BSC status (Li et al., 2017). Additionally, at the sample site scale,

disturbances and vegetation cover influence the spatial patterns of BSCs (Li et al., 2017).

2.1.3 Detection and Methodology

Remote sensing plays a crucial role in mapping, classifying, and characterizing BSCs, as their patchy distribution across vast regions makes it challenging to accurately assess spatial patterns using only ground-based mapping techniques. Spectroscopy has been widely employed in BSC research, with insights from proximal sensing informing regional-scale remote sensing studies. Reflectance spectroscopy offers a superior alternative to conventional laboratory methods, as it is minimally invasive and enables the in situ detection of temporal changes in BSCs.

The application of remote sensing in BSC studies has been explored since 1986, when optical imagery from the Landsat TM sensor was analyzed for this purpose (Green, 1986; Wessels & Van Vuuren, 1986). Two primary spectral domains have been identified for analyzing BSC properties: the reflective domain and the thermal infrared (TIR) domain, both extensively used in soil, vegetation, and land cover analysis (Rozenstein & Adamowski, 2017). Spectroscopy in the reflective domain, covering the visible (VIS), near-infrared (NIR), and shortwave infrared (SWIR) regions, relies on reflected solar illumination to provide detailed surface composition data. In contrast, TIR spectroscopy, operating primarily in the long-wave infrared (LIR) region, exhibits greater sensitivity to soil properties. Although its field application remains limited due to the high cost and restricted availability of instrumentation, TIR spectroscopy offers an advantage through inherent self-emission, enabling remote sensing under shaded conditions, cloud cover, and even at night (Eisele et al., 2012; Eisele et al., 2015). Over time, various techniques have been developed to improve BSC identification and mapping. Methods for BSC identification and extraction primarily include spectral index-based techniques that leverage BSC spectral properties and object-oriented approaches (Crucil & Van Oost, 2021). The distribution of BSCs has been mapped using spectral mixture analysis, integrating aerial photos with limited spectral

information and hyperspectral imagery (Hill et al., 1998; Hill et al., 2008). Spectral analysis of BSCs, bare soil, and vegetation provides critical insights into how BSC cover influences spectral response in heterogeneous landscapes, significantly affecting vegetation indices and surface albedo (Rodríguez-Caballero et al., 2015). In China's Mu Us Sandy Land, remote sensing techniques have enabled the quantification of moss crusts, revealing a total coverage of approximately 6.43% (0.72×10^4 km^2) (Feng et al., 2015). In addition, multi-source UAV imagery enables high-precision mapping (>80% accuracy) of vegetation, bare ground, and BSC components in dryland ecosystems, as demonstrated in central Spain (Blanco-Sacristán et al., 2021). These advancements underscore the increasing role of remote sensing technologies in improving the detection, classification, and large-scale monitoring of BSCs.

NDVI has been widely used to estimate and map BSC coverage, as it effectively represents different vegetation types and their physiological conditions. However, for BSCs, both dry and wet states must be considered, as spectral reflectance varies with biological activity and surface moisture content. Escribano et al. (2010) demonstrated that chlorophyll significantly influences spectral reflectance in BSCs. NDVI is influenced by chlorophyll absorption, with spectral features around 680 nm correlating with chlorophyll content in cyanobacteria-, lichen-, and moss-dominated crusts (Weber, 2008; Chamizo, 2012). NDVI values for wet BSCs are notably higher than those for dry BSCs. After six years of disturbance through soil redistribution, NDVI values of wet crusts reached 0.15 (Zaady et al., 2007). NDVI has been shown to vary significantly across different BSC types, with wet cyanobacteria-dominated crusts on sand reaching values of up to 0.22, while moss-dominated crusts exhibit even higher values of 0.3, in contrast to bare soil at 0.08 (Weber & Hill, 2016). These NDVI differences suggest that dry BSCs share spectral characteristics with bare soil, whereas wet BSCs display enhanced chlorophyll absorption and increased NIR reflectance, making their spectral response more comparable to that of vascular plants. The spectral similarity between BSCs and other land cover types, such as bare soil and sparse woody vegetation, often

leads to classification ambiguities when using NDVI. To address these limitations, alternative spectral indices have been developed over the past few decades to enhance the accuracy of BSC identification and differentiation.

Numerous studies have explored the calculation methods for biocrust-related indices. Karnieli (1997) developed the spectral Crust Index (CI) as a remote sensing-based approach for detecting and mapping biological soil crusts in arid dune environments. This index was specifically designed to distinguish BSCs from bare sand and other land surface features by leveraging their unique spectral reflectance characteristics. Utilizing data from multispectral sensors, CI enhances the spectral contrast between crust-covered and non-crust areas, providing a more effective method for large-scale BSC mapping. The study demonstrated the feasibility of satellite-based BSC detection, laying a foundation for subsequent advancements in remote sensing applications for dryland ecosystem monitoring. The Biological Soil Crust Index (BSCI) was developed based on Landsat ETM+ imagery, enabling the effective identification and extraction of lichen biocrusts in the Gurbantunggut Desert (Chen et al., 2005). The interpretation achieved a Kappa coefficient of 0.82, with an overall detection accuracy of 94.7% for the presence or absence of biocrusts. However, lichen crusts could only be effectively distinguished when biocrust coverage exceeded 33%. The CI and BSCI were evaluated using field spectra and Compact Airborne Spectrographic Imager (CASI) hyperspectral images, but both indices exhibited limitations in differentiating biocrusts from bare soil. When applying CI, roads devoid of biocrusts were erroneously classified as biocrust-covered areas, whereas with BSCI, certain vegetation zones without biocrusts were misclassified as biocrusts. To address these issues, the Continuum Removal Crust Identification Algorithm (CRCIA) was developed, which, when applied in South Africa, achieved a Kappa index of 0.831 and demonstrated strong classification performance (Weber et al., 2008). Using CASI hyperspectral imagery, the support vector machine (SVM) classification method was shown to effectively differentiate bare soil, algal crust, lichen crust, green vegetation, and dry vegetation, while spectral mixture analysis

accurately quantified the proportion of each feature within a pixel (Rodríguez-Caballero, 2014). The Thermal Crust Index (TCI) was developed based on the principle that variations in reflectivity and absorptivity induced by BSCs alter surface temperature. Designed to differentiate BSCs from bare sand in desert environments, the TCI demonstrates enhanced effectiveness when combined with the NDVI and the CI (Rozenstein & Karnieli, 2015). Applied to Sentinel-2 imagery, the Biocrust Greenness Index (BGI) and Maximum Vegetation Development Index (MVDI) enhance dryland ecosystem monitoring by capturing the greening dynamics of BSCs and vascular plants. These indices provide valuable insights into BSC responses to climatic variability, particularly precipitation, advancing the understanding of climate change impacts on BSCs and vegetation (Panigada et al., 2019). The Sandy Land Ratio Crust Index (SRCI) and the Desert Ratio Crust Index (DRCI) were employed to map moss-dominated crusts in the Mu Us Sandy Land, China. Compared to mapping without these indices, this approach improved the overall classification accuracy for sandy land and desert areas by 6% (Wang et al., 2022). The Fractional Biocrust Cover Index (FBCI) is derived from Sentinel-2 imagery with a 10 m spatial resolution, based on radiative transfer theory. The estimated fractional biocrust coverage exhibits a high level of agreement with field measurements, with a RMSE of 0.0774 and a systematic deviation of −4.05% (Sun et al., 2024). Spectral characterization methods provide a practical and efficient approach for monitoring the long-term spatiotemporal dynamics of biocrust distribution. However, it is important to note that their applicability is primarily limited to environments where BSC coverage exceeds 30% and vegetation cover remains below 10% (Beaugendre et al., 2017).

While remote sensing has advanced BSC classification and mapping, long-term monitoring and large-scale distribution analysis remain limited. Few studies have developed robust methodologies to assess the long-temporal dynamics of BSCs across large-scale spatial distributions using satellite observations, highlighting the need for further research. Additionally, BSCs are often analysed independently through either

qualitative descriptions or quantitative comparisons, hindered by spatial-temporal variability and environmental complexity. Bridging these gaps requires integrating diverse analytical approaches with advanced remote sensing techniques to improve the understanding of BSC dynamics.

2.2 Satellite data processing

In recent years, the research of spatial-temporal fusion models has gained extensive attention from scholars at home and abroad. According to the different principles of their algorithms, spatial-temporal fusion models can be mainly classified into three categories: spatial-temporal fusion models based on weight function, spatial-temporal fusion models based on hybrid image element decomposition and spatial-temporal fusion models based on dictionary pair learning (Dong & Meng, 2018).

Among spatiotemporal fusion models based on weighting functions, the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) was introduced to improve reflectance prediction. This model accounts for temporal and spatial distance as well as spectral similarity between similar image elements and target image elements, constructing a weighting function to determine their relative contributions. By integrating the high temporal resolution of MODIS remote sensing images with the high spatial resolution of Landsat remote sensing images, STARFM generates reflectance data that simultaneously retain both MODIS's temporal resolution and Landsat's spatial resolution (Gao et al., 2006). To enhance transient surface change detection, the Spatial Temporal Adaptive Algorithm for Mapping Reflectance Change (STAARCH) was introduced, allowing the extraction of relatively short-term surface change information (Hilker et al., 2009). The Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM) further improved the methodology by incorporating the temporal trend of surface reflectance and introducing the concept of correlation coefficients, thereby enhancing adaptability to complex surfaces (Zhu et al., 2010). Additionally, the Spatial-Temporal Vegetation Index Image Fusion Model

(STVIFM) was developed to integrate the rate of change of vegetation indices across different growth periods, facilitating the construction of vegetation index time series (Liao et al., 2017).

Among spatiotemporal fusion models based on hybrid image element decomposition, several approaches have been developed to improve reflectance prediction and surface change detection. The Spatial and Temporal Data Fusion Approach (STDFA) utilizes a hybrid image element decomposition method to extract temporal change information from time-series low-resolution imagery. An unsupervised classification method is then applied to classify image elements and construct spectral unmixing equations, thereby obtaining the average reflectance change for each category to enhance image prediction (Wu et al., 2012). The Spatial and Temporal Reflectance Unmixing Model (STRUM) was later introduced to directly decompose temporal changes in low-resolution image elements to estimate endmember variations. This approach incorporates Bayesian theory to constrain the estimation process, enhancing prediction accuracy (Gevaert et al., 2015). The Flexible Spatiotemporal Data Fusion (FSDAF) model employs a thin-plate spline function to interpolate low-resolution data at the prediction moment, deriving the spatial component of surface information. The temporal component is then extracted using a hybrid image decomposition method, and the two components are combined to generate the final prediction. This method effectively captures surface type changes over time (Zhu et al., 2015). To further enhance STDFA, the Improved Spatial and Temporal Data Fusion Approach (ISTDFA) was developed. This model introduces a weighted hybrid image decomposition equation to address reflectance change volatility and applies a linear model to account for sensor discrepancies, thereby improving prediction accuracy (Wu et al., 2016). More recently, the Improved Flexible Spatiotemporal Data Fusion Model (IFSDAF) was proposed, integrating a constraint-based least-squares method to combine temporal and spatial variations of surface information, enabling the generation of high-resolution spatiotemporal data with improved accuracy (Liu et al., 2019).

Compared to the previously discussed methods, spatiotemporal fusion models based on dictionary pair learning were developed relatively late. These models establish correspondences between high-resolution and low-resolution images by leveraging structural similarities, allowing them to capture key predictive features, including land cover changes. The Sparse-Representation-Based Spatiotemporal Reflectance Fusion Model (SPSTFM) was introduced to construct relationships between two pairs of high-resolution and low-resolution images through dictionary pair learning, using the trained dictionary to predict high-resolution imagery (Song et al., 2012). Building upon SPSTFM, a modified spatiotemporal fusion model based on dictionary pair learning was proposed, requiring only a single pair of high-resolution and low-resolution images for dictionary training. This approach utilizes sparse coding technology to downscale the predicted low-resolution data before synthesizing the final high-resolution image (Song et al., 2013). In recent years, with advancements in deep learning algorithms, researchers have integrated these techniques into spatiotemporal fusion. Due to the computational complexity of such algorithms, their applicability remains largely confined to small-scale study areas. The Spatiotemporal Fusion Using Deep Convolutional Neural Network (STFDCNN) model applies convolutional neural networks (CNNs) to extract spatial and temporal information from large-scale trained remote sensing datasets, thereby improving prediction accuracy (Song et al., 2018). Additionally, the Deep Convolutional Spatiotemporal Fusion Network (DCSTFN) fully leverages CNNs to establish a complex yet direct nonlinear mapping between input and output images. This method not only enhances fusion accuracy but also exhibits greater robustness compared to traditional spatiotemporal fusion algorithms (Tan et al., 2018).

2.3 Climate change and Biological Soil Crust response

Biological soil crusts are widely distributed across global desert regions due to their remarkable resilience to desiccation, extreme temperatures (up to 70 °C), high pH, and salinity (Friedmann & Galun, 1974; Wessels & Van Vuuren, 1986; West, 1990; Zedda et al., 2011; Yu et al., 2022). Their distribution at global and regional scales is positively

correlated with precipitation (Belnap et al., 2007), condensate water availability (Jacobs et al., 1999, 2000), and soil moisture content (Schieferstein & Loris, 1992). The influence of temperature on BSC distribution varies depending on the species composition of the crusts (Spier & Van Herk, 1997; Van Herk et al., 2002). At local and landscape scales, as well as other finer spatial levels, BSC distribution is shaped by environmental factors, such as soil type (Eldridge, 1996; Bowker et al., 2005), texture (Duan et al., 2003), stability (Stovall et al., 2022), fertility (Bowker et al., 2006), vegetation cover (Seitz et al., 2017), topography (Su et al., 2020), and solar radiation exposure (Durham et al., 2018). The ecological functionality of BSCs is closely linked to their developmental stage, which follows a hierarchical succession from cyanobacteria-dominated crusts to lichen- and moss-dominated crusts. This successional trajectory is primarily driven by precipitation rates and, in some cases, by prevailing temperatures (Rozenstein & Adamowski, 2017; Weber & Hill, 2016; Maestre et al., 2011; Maestre et al., 2012). However, the role of vascular plants in BSC distribution remains inconclusive, with conflicting evidence on their impact (Maestre & Cortina, 2002; Bowker, 2007). Additionally, disturbance intensity significantly influences BSC dynamics, as moderate disturbance has minimal effects, whereas high-intensity disturbances lead to structural degradation and functional decline, ultimately reducing BSC coverage (Wang et al., 2004, 2007).

Understanding these factors is essential for predicting their ecological roles and responses to environmental change. Precipitation has long been recognized as a key determinant of BSC distribution, with increasing precipitation generally promoting lichen and moss crust cover, while algal crusts initially expand before declining as moisture availability increases (Eldridge & Tozer, 1997; Marsh et al., 2006; Büdel et al., 2009; Zhao et al., 2014). However, precipitation also stimulates vascular plant growth, and their canopy cover and litter accumulation can suppress BSC development (Bowker et al., 2005). Beyond total precipitation, seasonal patterns and event frequency play a crucial role, with winter precipitation and moderate rainfall events (5 mm) being

most conducive to BSC formation (Read et al., 2014), whereas frequent light rains (1.2 mm) can hinder bryophyte crust survival (Chamizo et al., 2016; Jia et al., 2019). In hyper-arid environments like the Namib Desert, where precipitation is extremely scarce, lichen and bryophyte cover can still reach ~70%, primarily due to non-precipitation water sources such as fog and dew condensation, which facilitate atmospheric humidity retention (Eldridge et al., 2020; Kidron, 2019a; Li et al., 2021). Temperature further influences BSC distribution and species richness by modifying vegetation cover and soil pH (Eldridge and Delgado-Baquerizo, 2019), while elevated soil temperatures increase evaporation, creating conditions that limit biological colonization (Garcia-Pichel, 2013). The influence of temperature on the structural composition of biological soil crust communities becomes increasingly pronounced when interacting with time and precipitation (Ferrenberg et al., 2015). Additionally, legacy effects of past climatic conditions, which shape contemporary ecological patterns, have been shown to influence both the distribution and community composition of biological soil crusts (Eldridge & Delgado-Baquerizo, 2019). Under projected future climate scenarios of reduced precipitation, rising temperatures, and increased drought frequency, BSC cover is expected to decline by approximately 25% - 40% by the end of the century, with communities shifting towards early-stage cyanobacterial crusts (Rodriguez-Caballero et al., 2018). In addition to precipitation and temperature, potential evapotranspiration plays a crucial role in shaping BSC composition and distribution. In the hot deserts of the southern United States, such as the Mojave and Chihuahuan Deserts, high evapotranspiration favours the dominance and widespread distribution of algal crusts. As evapotranspiration decreases further north, in regions such as the Colorado Plateau, Great Basin Desert, and Columbia Plateau, lichen and moss crusts gradually replace algal crusts, increasing in both distribution and coverage (Rosentreter & Belnap, 2003). Beyond climatic factors, spatial heterogeneity and dispersal limitations of propagules, such as fungal spores and cyanobacteria, play a crucial role in the establishment and composition of biological soil crusts (Garcia-Pichel et al., 2013). Soil texture has traditionally been considered a key determinant, with finer particles thought to facilitate

BSC development (Williams et al., 2013; Belnap et al., 2014). However, this assumption has been challenged, as some studies suggest that dust and fine-grained soils are not essential for BSC formation (Kidron, 2019b). Additionally, soil parent material influences surface weathering and water retention capacity, thereby shaping BSC distribution patterns (Bowker & Belnap, 2008). A comprehensive review of experimental studies further indicates that bryophyte- and lichen-dominated crusts are more prevalent on gypsum and calcareous soils (Elbert et al., 2012), whereas cyanobacterial crusts tend to dominate sandy substrates (Root & McCune, 2012).

In recent years, extensive research has examined the environmental factors influencing the growth and development of biological soil crusts in northern China. Higher plants have been found to have minimal impact on BSC degradation in interplant spaces, whereas the thickness and coverage of plant litter play a crucial role in crust degradation both beneath litter and in general (Zhang et al., 2008). Changes in the physical and chemical properties of BSCs and the underlying topsoil are influenced by vegetation type and crust development, with vegetation promoting crust formation and soil improvement in sandy areas (Zhao et al., 2009). In the Mu Us Sandy Land, BSCs in early developmental stages are highly influenced by vegetation presence. Light grazing disturbance has no significant effect on BSC cover or thickness, whereas moderate grazing disturbance leads to a substantial reduction in BSC cover. Additionally, BSCs in semi-fixed sandy lands are more sensitive to grazing disturbances (Zhang, 2014). On the Loess Plateau, BSC development varies significantly across different vegetation types in terms of crust thickness, shear strength, and bulk density. Moreover, slope aspect has been identified as a key factor affecting BSC development (Meng et al., 2011). In the agricultural-pastoral transition zone of northern China, plant litter plays a significant role in promoting BSC growth. In habitats with poor soil texture, vegetation facilitates crust formation; however, as environmental conditions improve, this relationship becomes significantly negative. Different plant communities exert varying effects on BSC cover, with the highest coverage observed in the *Xinjiang Salsola*

community, followed by the *Artemisia capillaris* + *Xinjiang Salsola* community, and the lowest in the *Artemisia capillaris* community. Furthermore, disturbance negatively impacts BSC cover, with the highest coverage found in core areas (complete enclosure), followed by edge areas (seasonal enclosure), and the lowest in peripheral areas (control area) (Lu et al., 2007). A survey conducted on the slopes of the Liudaogou small watershed in the Loess Plateau identified solar radiation, erosion, soil texture, and secondary soil properties as the four key factors influencing BSC development (Zhang et al., 2015). Among environmental factors, slope aspect and disturbance have the most significant direct impact on BSC cover, whereas slope position, vegetation height, plant species diversity, and litter cover exert indirect effects through other mediating factors (Wang, 2011).

Collectively, these studies demonstrate the diverse environmental drivers shaping BSC distribution across arid and semi-arid landscapes. However, variations in BSC distribution patterns across different study areas, each with distinct ecological and geomorphological conditions, suggest that findings may not always be consistent (Guo et al., 2012). Therefore, a comprehensive approach integrating ecological conditions, distribution patterns, and developmental processes across multiple scales is essential for a systematic and realistic assessment of BSC dynamics. Such an approach can provide deeper insights into the current status and functional roles of biological crusts in desert ecosystems.

2.4 Motivation

BSC play a significant role to assure the regular functioning of desert ecosystem, such as involvement in the process of formation, stability and fertility of soil, prevention of soil erosion caused by water or wind, augment of vascular plant colonization, and stabilization of sand dunes (Belnap, 2003; Belnap et al., 2001; Eldridge & Greene, 1994). BSC have been recommended as the top management priority in desert regions

(Belnap, 2003), especially under recent conditions in which desert regions are experiencing global warming and increasing human activity.

Many typical problems in the interrelationship between human activities and the natural environment in semi-arid and arid sandy areas (Zhou et al., 2022). The multiple ecological functions of BSC are closely related to the ecosystem's ability to simultaneously maintain a variety of functions and services such as nutrient cycling, hydrological cycling, and material export, especially in arid ecosystems where vascular plant growth is restricted, and the role played by BSC should not be underestimated. Up to now, most of the studies on BSC are at the sample site scale, which could only reflect the distribution of BSC in local areas, and there is a lack of understanding of the distribution of BSC at a large scale. A regional-scale study on the distribution of BSC would help us better understand the distribution of BSC in the desert, which is of great significance in the assessment of the ecological function of BSC and the scientific management of BSC resources. Therefore, taking into account the actual situation of arid ecosystems in China, conducting regional-scale studies to understand the distribution and growth of BSC in the Gurbantunggut Desert as well as the response of BSC to climatic factors at the bioclimatic level is not only a useful supplement to the existing research work, but also provides a reference to comprehensively assess the relationship between BSC and the environmental system. In addition, in the context of global change, understanding the relationship between BSC and climatic variables is not only conducive to a comprehensive understanding of the global environment and ecosystem impacts on biomes, but also conducive to scientifically combining BSC with other biological resources for ecosystem conservation, so that they can give full play to their ecological benefits and better serve human beings.

Hence, a modelling system for biological soil crusts and climate factors based on multi-datasets is required to bridge the gap between the long-term and large spatial distribution of BSC and its interaction with the climate variability. In this study, two

models with improved BSC detection method and machine learning are constructed. These models can be used to explain the effects of climatic factors on biological crusts, the short-term or long-term forecasts of regional-scale distribution of BSC, the assessment for the potential effects of climate change on the availability of BSC and the sustainable development of ecosystem.

3 Study area description and data processing

3.1 Introduction to the studied desert

3.1.1 Gurbantunggut Desert

The Gurbantunggut Desert, the study area of this research, is analyzed for the distribution of biological soil crusts (BSCs) and serves as the basis for constructing BSC-climate models. Geographically, it is located between 44º11′–46º50′N and 84º31′–90º00′E, positioned at the center of the Junggar Basin in the Xinjiang Uygur Autonomous Region, China. Covering an area of 48,800 km², it is the second-largest desert in China and the largest fixed and semi-fixed desert in the country (Fig. 3-1). The southern boundary of the Gurbantunggut Desert transitions into alluvial and floodplain fan margins formed by sediments from the Tianshan Mountains, shaping the natural vertical zonation from the northern slopes of the Tianshan Mountains to the basin floor.

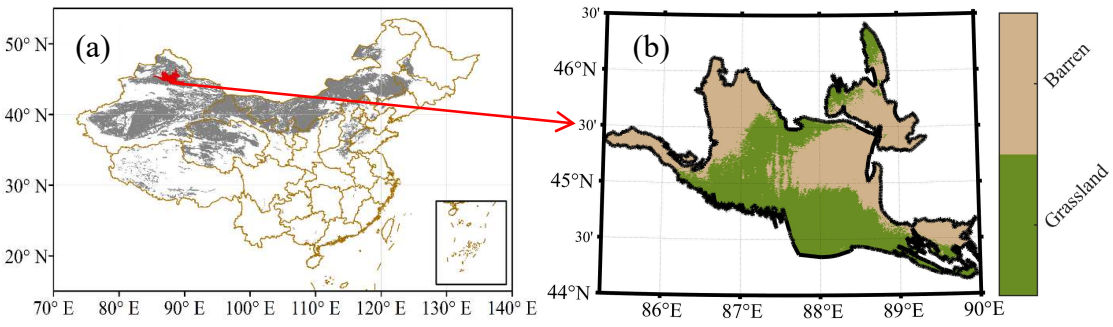


Fig. 3-1 Study area (a) and its land type (b).

The Gurbantunggut Desert is characterized by a temperate continental desert climate, shaped by the "rain shadow effect" of the Himalayas, which prevents humid air currents from the Indian Ocean from reaching the region, resulting in an extensive arid landscape. Annual precipitation ranges from 150 to 250 mm, with only 70–100 mm occurring in the desert hinterland. Precipitation is slightly higher in spring and summer compared to autumn and winter, while annual evaporation far exceeds precipitation, ranging from 2,000 to 2,800 mm. The annual mean temperature is 6–10°C, with extreme values exceeding 40°C in summer and dropping below −40°C in winter. The

region exhibits a climatic pattern of low rainfall with high evapotranspiration (Li et al., 2001; Zhang et al., 2005; Liu et al., 2023). During winter, snow depth generally ranges between 10 and 30 cm, with snow accumulation beginning in November and persisting until mid-March of the following year. Snowmelt contributes 30%–50% of the annual precipitation. Surface runoff is nearly absent, and the water table is relatively deep, exceeding 5 m at the edges and 16 m in the hinterland of the desert (Qian et al., 2007; Jian et al., 2019). Fig. 3-2 illustrates the monthly variations in temperature and precipitation in the Gurbantunggut Desert. In general, biological soil crusts (BSCs) remain dormant or fail to grow when temperatures fall below 0°C (Piao et al., 2006). Therefore, this study primarily focuses on analyzing the spatiotemporal distribution patterns and dynamics of BSCs during the growing season (March to November), encompassing spring, summer, and autumn.

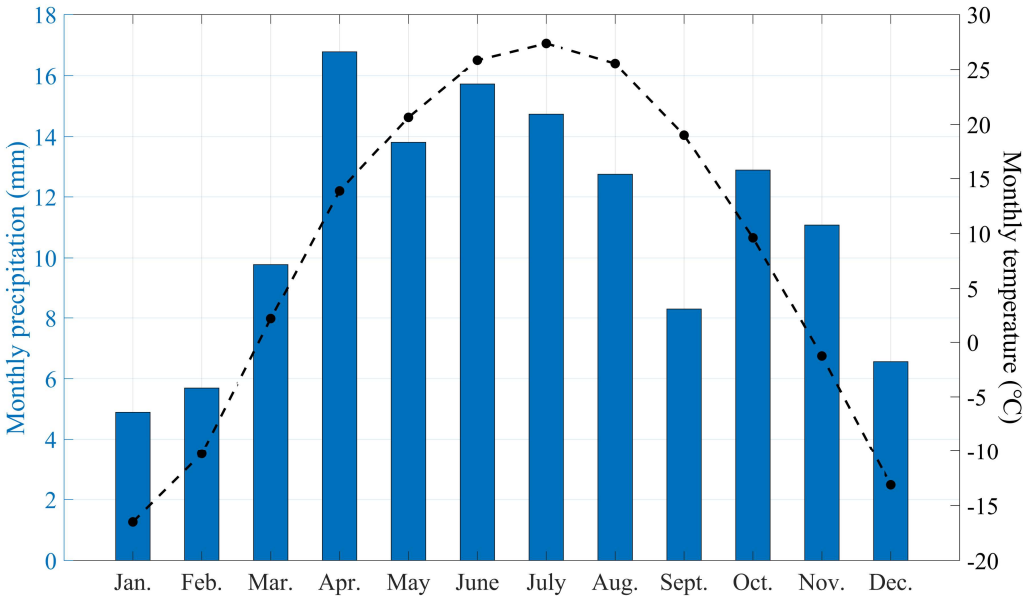


Fig. 3-2 The monthly average temperature and monthly precipitation changes in the Gurbantunggut Desert.

The geomorphology of the Gurbantunggut Desert exhibits distinct east-west and north-south variations, characterized by a low-elevation trend in the east-west direction and higher elevations along the north-south axis (Fig. 3-3a). Sand ridges oriented east-west are distributed in a disordered manner, exhibiting significant variability in

geomorphological types. The relative heights of these ridges range from 10 to 50 m, with ridge lengths varying from several hundred meters to over a dozen kilometers. In contrast, north-south-oriented landforms are predominantly composed of chain, crescent, and beehive-shaped ridges. The geomorphological diversity is relatively low, with fixed and semi-fixed dunes dominating the landscape (Qian et al., 2007; Wang et al., 2015). Fixed and semi-fixed sand dunes account for approximately 87% of the total desert area, with dunes generally aligned along a northwest-southeast orientation. The majority of dunes range between 15 and 20 m in height, with vegetation cover on dune surfaces varying between 15% and 55%. Among these, fixed sand dunes support vegetation cover of 40%–55%, while semi-fixed dunes exhibit a lower vegetation cover of 15%–25%. The summits of dunes in the Gurbantunggut Desert are predominantly composed of windswept sandy soils, whereas saline soils and intermountain gravelly soils are primarily found in interdune areas. The sandy substrate mainly consists of medium to fine sands, contributing to the region's overall sedimentary characteristics.

The Gurbantunggut Desert exhibits a vegetation cover of less than 30%, yet it harbors a diverse assemblage of plant species. Its floristic composition includes elements from Central Asia, the Mediterranean, and the Caspian-Kazakhstan-Mongolian region. Recognized as one of the most species-rich temperate deserts globally, the Gurbantunggut Desert is a typical region characterized by high plant species diversity and genetic resources. The vegetation is predominantly composed of sandy and drought-tolerant species with a diverse composition (Fig. 3-3b), including *Stipa glareosa, Agriophyllum squarrosum, Ephedra distachya, Calligonum mongolicum,* and *Reaumuria soongorica*. Additionally, sandy ephemeral species such as *Rudbeckia soongorica, Eremurus inderiensis,* and *Ceratocarpus arenarius* contribute to the region's seasonal plant dynamics. Among the most dominant and ecologically significant species are *Haloxylon persicum* and *Haloxylon ammodendron*, which characterize vast stretches of the desert landscape (Zhang & Chen, 2001; Song & Hu, 2011). Compared to other deserts, the Gurbantunggut Desert supports relatively high

biological diversity and hosts well-preserved desert plant communities, making it China's only designated desert vegetation nature reserve. The prevalence of extensive, densely packed semi-fixed sand dunes contributes to stable soil moisture conditions. Precipitation is relatively evenly distributed throughout the year, with winter and spring rainfall supporting the development of short-lived and short-lived-like plant species. The surface soils of the Gurbantunggut Desert exhibit distinct coloration patterns, ranging from black and dark brown to white and yellowish-green, reflecting variations in moisture availability, biological composition, and soil development stages. Additionally, the desert surface is rich in biological soil crusts, with lichen-dominated BSCs being the most prevalent, alongside moss-dominated BSCs. These BSCs exhibit peak growth during cool, humid periods in autumn and early spring, when dew, fog, and temporary rainfall serve as critical moisture sources that sustain species involved in BSC formation (Kidron et al., 2002). Overall, the Gurbantunggut Desert is a species-rich temperate desert ecosystem, characterized by diverse drought-tolerant vegetation, extensive semi-fixed sand dunes that help stabilize soil moisture, and biological soil crusts that thrive during cool, humid periods.

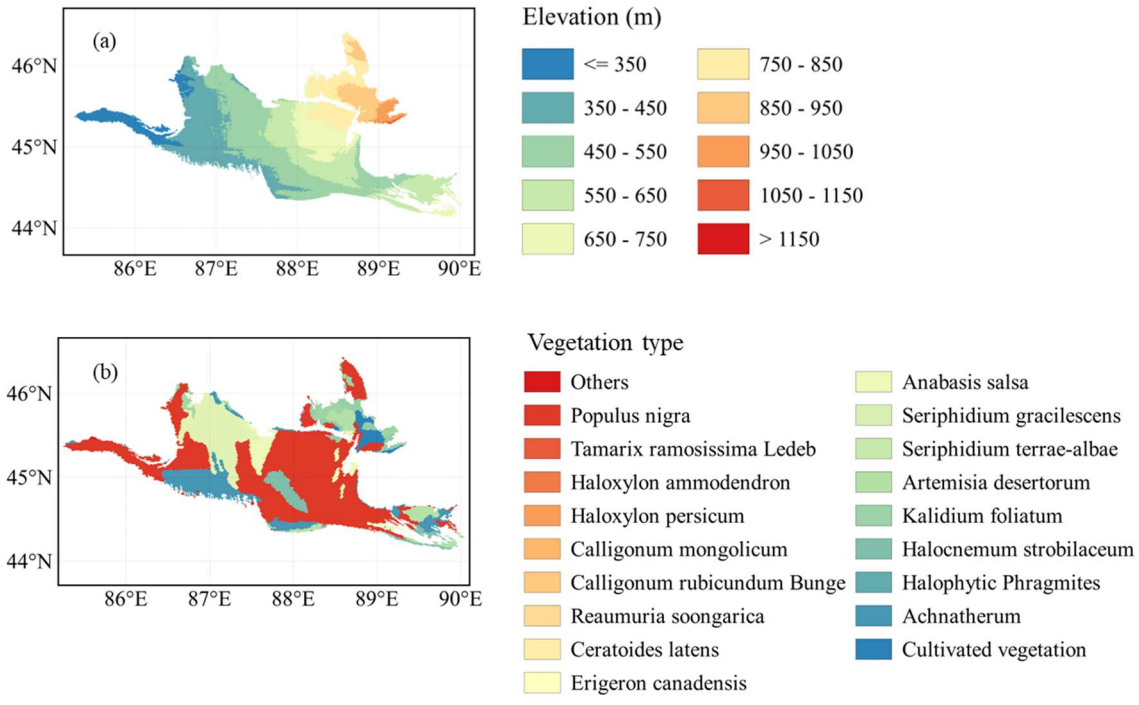


Fig. 3-3 Elevation (a) and vegetation communities (b) in the Gurbantunggut Desert.

3.1.2 Atacama Desert

The Atacama Desert is another area where BSC-climate model applied in this study. The Atacama Desert located along the western border of South America, which runs nearly 1000 km along the Pacific coast of South America from 30°S to 20°S, with a temperate desert climate. This region has an annual mean temperature of 14 - 16 °C (McKay et al., 2003). It is known for its arid to hyper-arid climate (Dunai et al., 2005; Clarke, 2006). Average annual rainfall is less than 200mm. Rainfall is concentrated in summer (January to February), with almost no rainfall recorded during the rest of the year (Eshel et al., 2021). The distinctive climate of this region is the result of a complex interplay of factors. These include the presence of subtropical high-pressure zones, the influence of the cold coastal Humboldt Current, the offshore winds (Clarke, 2006), the Andean rain shadow effect and the latitudinal position of the region (Houston & Hartley, 2003). The constant temperature inversion due to the cool north- flowing Humboldt Current and the presence of the strong Pacific anticyclone (Rundel et al.,1991; Miller, 1976). The position of the Pacific anticyclone is generally stable with a small shift of a few degrees south in the summer (Trewartha, 1961). Geological and soil mineralogical evidence suggests that extreme arid conditions have persisted in the Southern Atacama for 10–15 million years (Ericksen, 1983; Houston & Hartley, 2003; Clarke, 2006) making it one of the oldest deserts on Earth.

The diversity, occurrence, and distribution of microbial life in the Atacama Desert are shaped by multiple environmental factors, including salinity, ultraviolet (UV) radiation, water availability, and temperature. Among the microbial communities in this extreme environment, lichenized fungi are particularly dominant, representing a significant source of fungal diversity. Early research on fungi in the Atacama Desert provided comparative insights into lichenized fungal species from the coastal Atacama region and Baja California, Mexico (Rundel, 1978). Subsequent studies expanded the understanding of microbial diversity in this desert ecosystem. For example, two new species of lichenized fungi—*Lecanographa azurea* and *Roccellina ochracea*—were

described (Follmann, 2008). More recently, a survey along two altitudinal transects in Alto Patache identified 77 lichenized fungal species, further highlighting the region's microbial richness (Vargas Castillo et al., 2017).

The flora and vegetation distribution of the Atacama Desert is shaped by three distinct topographic units. The first is the Coastal Range, a faulted escarpment that rises abruptly along the coast. Atmospheric conditions under stable subtropical high-pressure systems create a mild yet arid climate. The northern Chilean coastline is frequently covered by stratocumulus cloud layers, which are blocked by the Coastal Range, leading to the formation of fog-dependent plant communities known as Lomas. These Lomas ecosystems are distributed across approximately 50 sites in the Atacama Desert, primarily along the north-central Coastal Range at elevations of 600–1100 m. The second unit is the Central Valley, which represents the core of the Atacama Desert. In El Niño years, these localized oases can expand to cover up to 5,000 km². Extreme drought conditions dominate this region, particularly north of 26°S, where vegetation is almost entirely absent, except for oases and riparian plant communities along river channels. The third unit is the western slope of the Andes, which is stratified into four distinct vegetation zones based on altitude. The unique geographical and climatic conditions of the region have led to the development of a specialized biological structure and the occurrence of highly localized Lomas plant communities. Therefore, among the 550 documented desert plant species in the Atacama Desert, more than 60% are endemic, underscoring the region's exceptional biodiversity (Dillon & Hoffmann, 1997).

3.2 Satellite data

3.2.1 Datasets

This study uses multi-satellite datasets from 2000 to 2018: 1) Landsat 7 (Enhanced Thematic Mapper, ETM) Level 2 surface reflectance data at 30 m resolution acquired from United States Geological Survey (USGS). Landsat scenes the Earth Explorer

(http://earthexplorer.usgs.gov/) archive are processed by the USGS with Standard Terrain Correction. These images contain 4 visible and near-infrared (VNIR) bands and 2 short-wave infrared (SWIR) bands processed to orthorectified surface reflectance, and one thermal infrared (TIR) band processed to orthorectified surface temperature. The data provides systematic radiometric and geometric accuracy by incorporating ground control points while employing a digital elevation map (DEM) for topographic accuracy. In order to reduce the impact of clouds, only selected clear sky (cloud coverage is less than 10%) data. 2) MOD09A1 (https://ladsweb.nascom.nasa.gov/search) provides MODIS (Moderate Resolution Imaging Spectroradiometer) band 1-7 surface reflectance at 500 m resolution. It is a level-3 composite of 500 m resolution MOD09GA. Each product pixel contains the best possible L2G observation during an 8-day period as selected on the basis of high observation coverage, low view angle, absence of clouds or cloud shadow, and aerosol loading. 3) In the case study of extreme dust events, the FY-3 meteorological satellite (http://satellite.nsmc.org.cn/PortalSite/Default.aspx), as the second generation of polar-orbiting meteorological satellites in China, carries a Visible and Infra-Red Radiometer (VIRR) that can provide dust monitoring daily data with a resolution of 1km.

3.2.2 Satellite data fusion

BSC monitoring require high resolution remote sensing imagery in both time and space - a requirement that cannot currently be satisfied by any single Earth observing sensor in isolation. The MOD09A1 provides daily global observations at 500m spatial resolution. While imagery from coarse resolution sensors such as MODIS are typically superior to finer resolution data in terms of their revisit frequency, they lack spatial detail to capture surface features for many applications. The Landsat satellite series provides medium spatial resolution (30m) imagery which is well suited to capturing surface details, but a long revisit cycle (16-day) has limited its use in describing daily surface changes. Therefore, this study used data fusion to utilize observations from multiple sensors. Fig. 3-4 shows the flow chart of satellite data fusion. At first, selecting

the Landsat 7 images containing the study area, a total of four scenes (Path: 142-143, Row: 28-29). Since the Scan Line Corrector (SLC) failed the Landsat 7 images collected after May 31, 2003 have data gaps (Landsat 7 ETM+ SLC-off), but are still useful and maintain the same radiometric and geometric corrections as data collected prior to the SLC failure. The data gaps are filled based on Geospatial Data Abstraction Library (GDAL) in this study. Then, the MOD09A1 data needs to be reprojected, clipped and resampled to have the same characteristics as the Landsat data. At last, from one Landsat image and two MOD09A1 images, a fusion image with high spatial-temporal resolution can be obtained.

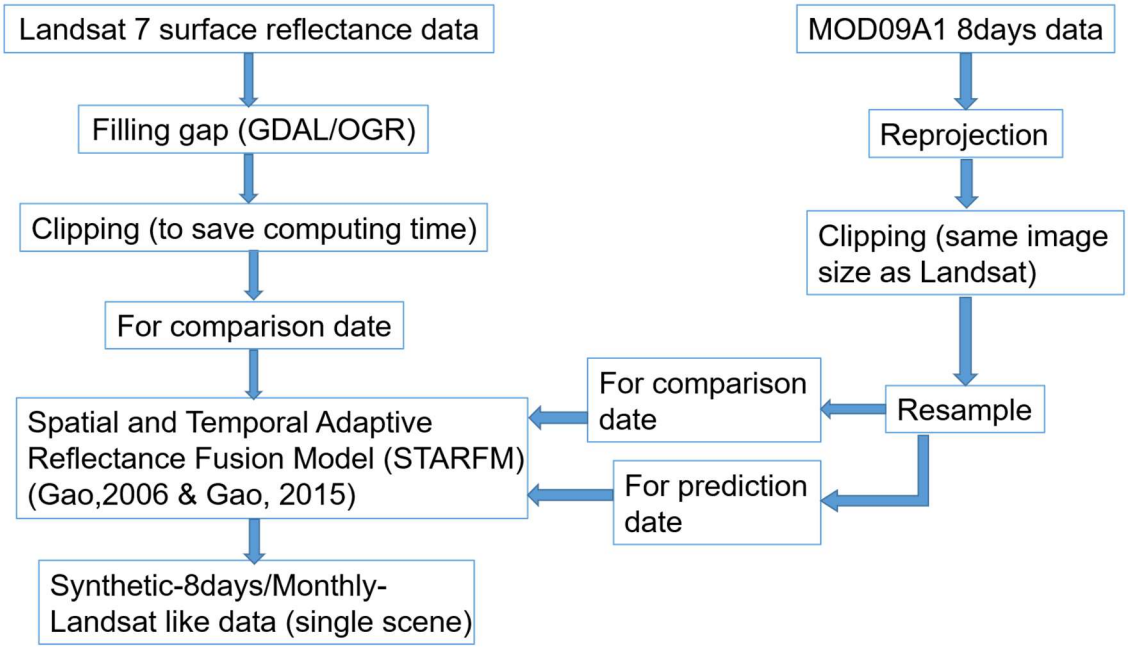


Fig. 3-4 Flow chart of satellite data fusion.

3.2.3 BSC detection method

Chen et al. (2005) developed Biological Soil Crusts Index (BSCI) to detect BSC for each pixel when the BSC coverage rate is 33% or more and determined lower and upper detection thresholds of BSCI of Landsat 7 sensor experimentally, which is 3.69-6.59 in Gurbantunggut desert. The Biological Soil Crust Index (BSCI) is calculated using the

reflectance values of different spectral bands, incorporating an empirical parameter to enhance its sensitivity to biocrust detection. The formula is defined as:

$$BSCI = \frac{1 - L \times \left| R_{red} - R_{green} \right|}{R_{GRNIR}^{mean}} \quad (3\text{-}1)$$

where:

- R_{red} and R_{green} represent the reflectance values in the red and green spectral bands, respectively.
- R_{GRNIR}^{mean} is the mean reflectance of the green and near-infrared (NIR) bands.
- L is an empirical parameter, which adjusts the sensitivity of the index based on the spectral characteristics of biological soil crusts. In this study, the empirical parameter L is set to 2.
-

This index is designed to distinguish biological soil crusts from bare soil by leveraging the differences in spectral reflectance among vegetation and non-vegetated surfaces, particularly focusing on the reflectance contrast between the red and green bands while normalizing using the green and NIR bands. The Landsat BSCI shows higher values for the presence BSC, relative to the background of bare sand and dry plant material. The higher the BSC percent coverage, the higher the BSCI value would be expected.

Biological soil crusts (BSCs) exhibit peak growth during wet and cool periods, relying on dew, fog, or temporary rainfall as primary water sources (Zhang et al., 2009). At the onset of these favorable conditions, BSC signals are most prominent, as annual plants have not yet germinated, and perennials remain dry. Under these circumstances, BSCI values for BSCs—particularly moss-dominated BSCs—can be similar to those of dry plants, whereas the Normalized Difference Vegetation Index (NDVI) values show greater differentiation. Furthermore, land cover data (Fig. 3-1b) indicate that nearly 50% of the Gurbantunggut Desert consists of grassland, where vascular plants such as mosses exhibit higher NDVI values, potentially leading to misclassification with BSCs.

Neglecting these spectral characteristics may result in erroneous interpretations of vegetation phenology. To address this issue, NDVI is introduced as an additional indicator, alongside land cover classification and the Kappa index (Cohen, 1960), to refine Chen's method and enhance the discrimination between BSCs and vascular plants. The formula for the NDVI is expressed as:

$$NDVI = \frac{(R_{NIR} - R_{red})}{R_{NIR} + R_{red})} \tag{3-2}$$

where:

- R_{NIR} represents the reflectance in the near-infrared (NIR) band.
- R_{red} represents the reflectance in the red band.

The Kappa index ($\widehat{K}$), also known as Cohen's Kappa coefficient, is a statistical measure widely used in remote sensing, land cover classification, and ecological studies to assess the reliability and accuracy of classification results. It is computed using the following formula:

$$\widehat{K} = \frac{p_o - p_c}{1 - p_c} \tag{3-3}$$

where:

- p_o (Observed agreement) represents the proportion of actual agreement between the classification and the reference data.
- p_c (Chance agreement) represents the expected agreement due to random chance.
- For further details, see Appendix 7.1.

The Kappa index ranges from -1 to 1:

$\widehat{K} = 1$ indicates perfect agreement between classification and reference data.

$\widehat{K} > 0$ indicates better-than-random agreement, with higher values indicating stronger reliability.

$\widehat{K} = 0$ indicates that the agreement is no better than random chance.

$\widehat{K} < 0$ indicates worse-than-random classification performance, suggesting systematic disagreement.

$\widehat{K} = -1$ indicates the classification is entirely incorrect compared to the reference data.

Higher NDVI values indicate dense, healthy vegetation, while lower values correspond to sparse or non-vegetated surfaces, such as bare soil or water. Similar to BSCI, NDVI also requires an appropriately defined threshold to ensure accurate application. To determine the optimal NDVI thresholds, values were selected based on those yielding the highest Kappa index, using validation points for accuracy assessment. NDVI values were evaluated through a confusion matrix, which served as a prerequisite for Kappa index calculation. The NDVI range of 0.15–0.22 was identified as achieving the highest Kappa index (0.93), indicating a strong agreement between classification results and ground truth data.

3.2.4 Gridding satellite data

Applying the aforementioned improved detection method to the satellite fusion imagery enables the long-term, large-scale assessment of BSC coverage across the study area. To facilitate comparative analysis and modelling, the satellite-derived BSC data are aggregated into the same grid system as the climate data (Fig. 3-5). The BSC coverage fraction for each grid cell is computed as the ratio of BSC-classified pixels to the total number of pixels within the grid cell. After applying a masking procedure, the study area contains 55 valid grid points, which are sequentially numbered from left to right and top to bottom. These gridded datasets serve as the basis for subsequent BSC-climate modelling, ensuring data format consistency between satellite-derived BSC coverage and climatic variables.

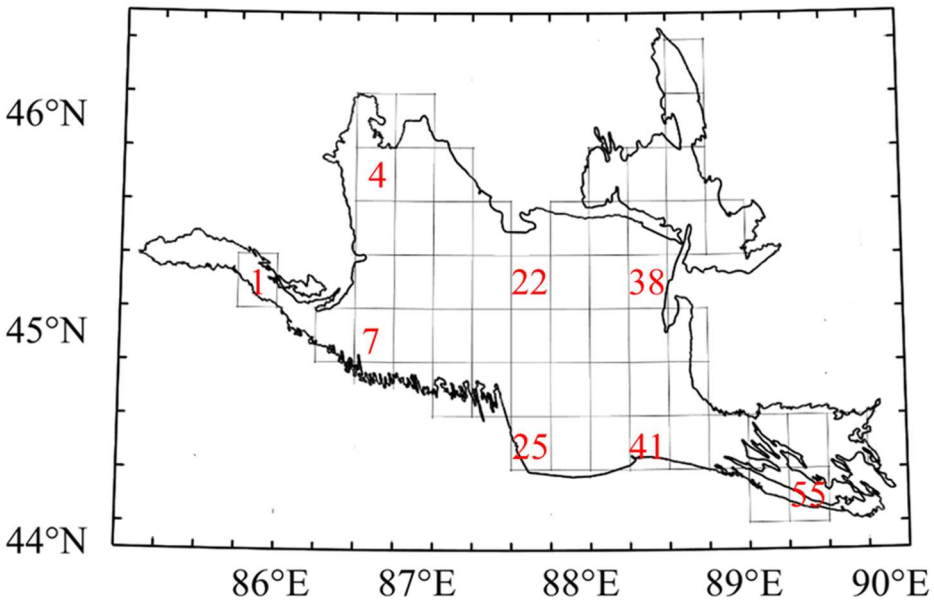


Fig. 3-5 Data grid schematic. The grey grid divides the study area into 55 grid points. The red numbers are examples of grid point numbers, showing that the grid point numbering sequence is from left to right and top to bottom.

3.3 Climate data

The formation of biological soil crusts is driven by a complex interplay of interrelated causes and effects, broadly categorized into natural and anthropogenic factors (Belnap, 2006; Belnap et al., 2016). Numerous studies have demonstrated that natural drivers play a dominant role, particularly at large and medium temporal scales (Belnap & Lange, 2003). In the context of global climate change, climatic factors have become a central focus of research, as BSC development is highly dependent on specific climatic conditions. While climate serves as a background factor, the formation and persistence of BSCs occur only under certain combinations of climatic variables. To investigate BSC-climate relationships, this study selects five key long-term climatic variables that are closely associated with BSC dynamics: specific humidity (SH) (kg/kg), 10-meter wind speed (WS) (m/s), 2-meter temperature (TMP) (K), surface solar radiation downward under clear-sky conditions (SR) (W/m^2), and total precipitation (PRE) (mm/day). Additionally, time-lag effects of these climatic factors are considered. To ensure comparability and consistency, the spatial resolution and units of all climate datasets are standardized. Temporally, climate data begin in 1999, one year earlier than

BSC data (2000), to account for lagged responses of BSC formation to climatic variations. Furthermore, to enhance the analysis of the water balance in the Gurbantunggut Desert, evapotranspiration data are incorporated into the Section 3.5 analysis, although they are not included in the BSC-climate modelling framework.

3.3.1 ERA5 data

ERA5 (https://cds.climate.copernicus.eu/) reanalysis data is the fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis for the global climate and weather for the past 4 to 7 decades. Radiative inversion using the RRTM rapid radiative transfer model. Data is available from 1940 onwards. In this study, the ERA5 data covers a total of 20 years from 1999-2018 with a spatial resolution of 0.25° x 0.25°. The specific humidity is obtained from ERA5 hourly data on pressure levels from 1940 to present, selected at 1000hPa near the ground surface; other data is from ERA5 hourly data on single levels from 1940 to present, where the 10-meter wind speed is calculated from the components in the u (eastward) and v (northward) directions. In terms of temporal resolution, all climate data are processed as monthly data for constructing the BSC-climate model.

3.3.2 CMIP6 data

Coupled Model Intercomparison Project Phase 6 (CMIP6) is a project coordinated by the Working Group on Coupled Modelling (WGCM) as part of the World Climate Research Programme (WCRP). In order to apply BSC-climate model to paleoclimate, Monthly average data from three periods under three scenarios from CMIP6 are selected (Table 3-1). The global attribute of AWI and MPI is “r1i1p1f1” and of MIROC is “r1i1p1f2”, where “r” is realisation, “i” is initialisation method, “p” is physics and “f” is forcing. "1" represents Initial conditions taken from a prior simulation. “2” As "1", but with an additional random perturbation applied to the initial state of the atmosphere.

Table 3-1 Scenario of CMIP6 data

Scenario	AWI-ESM-1-1-LR	MPI-ESM1-2-LR	MIROC-ES2L
Spatial resolution	250km	250km	500km
Time scale	Historical (1999-2014), LGM, Mid-Holocene		

Climate models exhibit systematic error (biases) due to the limited spatial resolution, simplified physics and thermodynamic processes, numerical schemes or incomplete knowledge of climate system processes. Hence, it is important to bias-correct the raw climate model outputs in order to produce climate projections that are better fit for modelling. The CMIP6 data are bias-corrected using the ERA5 data as reference data. The method of bias correction will be described in detail in Section 4.2.2.

3.4 Distribution of BSC in Gurbantunggut Desert

3.4.1 Variability of BSC coverage

Fig. 3-6 illustrates the temporal distribution, annual cycle, and anomaly of monthly BSC coverage in the Gurbantunggut Desert from 2000 to 2018. Throughout the year, BSC coverage exhibits a bimodal pattern, with peaks occurring in March and October, indicating that BSCs experience their highest growth during wet, cool periods (early spring and autumn). However, some anomalies deviate from this overall trend. In 2012 and 2013, a weaker summer peak was observed, whereas in 2002 and 2016, the early spring peak was less pronounced, possibly due to snow cover at high latitudes in early March of those years (Hui et al., 2019; Zhou et al., 2009). BSCI has limitations in distinguishing biological soil crust from snow cover, which may have contributed to this pattern. Over the 19-year period, the overall trend of BSC coverage is increasing. Notably, in 2000, 2001, 2010, and 2015, BSC coverage was below the annual average, whereas in 2013, 2017, and 2018, it was above the annual average. In other years, fluctuations remained relatively minor. During winter (December, January, and February), BSCs enter dormancy, making their detection via remote sensing nearly

impossible. Consequently, this study excludes winter BSC analysis from further discussion.

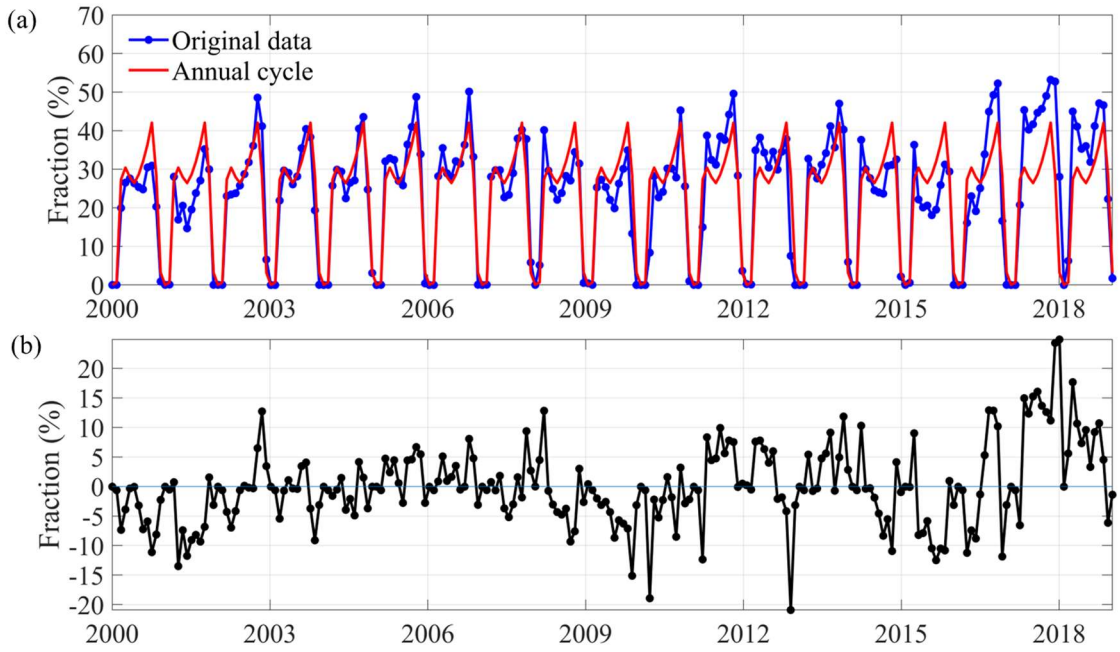
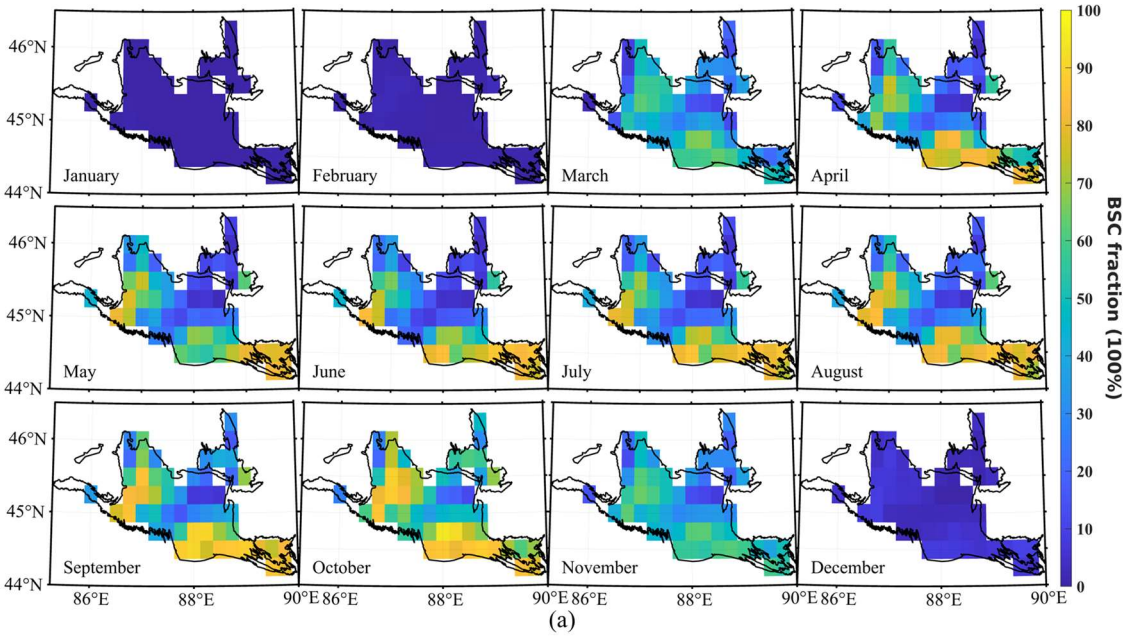


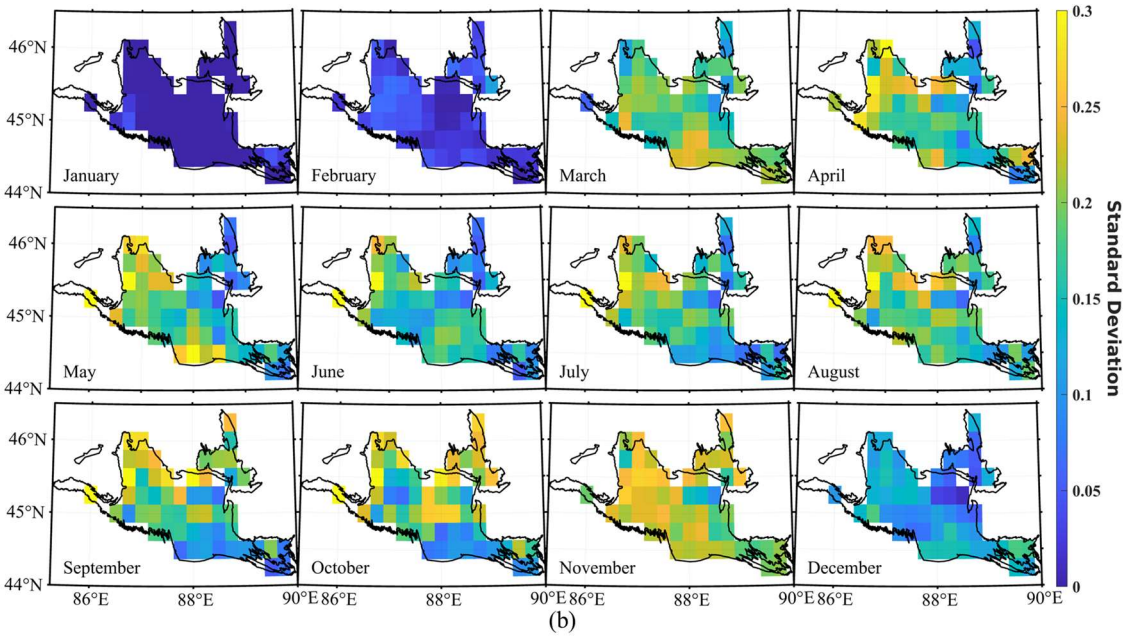
Fig. 3-6 Temporal distribution of the BSC fraction from 2000 to 2018: (a) The monthly temporal distribution (blue line) and the annual cycle (red line); (b) The anomaly.

Fig. 3-7 illustrates the spatial distribution of BSC fraction and its standard deviation in the Gurbantunggut Desert from 2000 to 2018. Overall, BSC coverage is relatively high in the northwestern and southern regions, with the southern area exhibiting particularly dense BSC coverage (70–80%). This pattern corresponds to the grassland distribution observed in the land cover classification (Fig. 3-1b), as BSCs often develop beneath vascular plants (Eldridge & Greene, 1994). Despite this broad similarity, notable differences exist between the two high-BSC regions. In the northwest, the standard deviation of BSC coverage is large, indicating greater temporal variability. In contrast, BSCs in the south exhibit more stability, with a lower standard deviation. This stability may be attributed to more favorable moisture and temperature conditions in the interdunal lowlands of the southern desert, which create a more suitable microenvironment for BSC persistence. Additionally, the southern region hosts a

greater diversity of plant species, potentially contributing to the enhanced stability of BSC communities (Zhang et al., 2002).



(a)



(b)

Fig. 3-7 Monthly spatial distribution (a) and standard deviation (b) of BSC fraction from 2000 to 2018 in the Gurbantunggut.

Fig. 3-8 shows the probability of BSC occurrence in the Gurbantunggut Desert during 2000-2018. It is calculated for each pixel using the following formula:

$$Probability_{month} = \frac{N_{occurrence,month}}{N_{total}} \quad (3\text{-}4)$$

where:

$N_{occurrence,mont}$ represents the number of years in which BSC was detected in a given $month$.

N_{total} is 19 in this study, denotes the total number of years considered in the analysis (2000–2018).

The colors in the figure represent the probability of BSC occurrence, with yellow indicating a probability of 1, signifying areas where BSCs are consistently present throughout the study period. This suggests that in these regions, BSCs enter a dormant state during the undetectable winter months rather than undergoing senescence or dying off. Spatially, BSC occurrence is highest in the southern and central regions, followed by the northern region, with the lowest occurrence in the western region. The higher probability of BSC occurrence near the oasis-desert transition zone is likely influenced by the favorable environmental conditions at the oasis edge, where sufficient moisture and moderate climatic conditions create an optimal microenvironment for BSC development. Temporally, BSC occurrence during winter (December–January) is low (0–20%), primarily due to sub-zero winter temperatures in the Gurbantunggut Desert. Consequently, the subsequent analyses focus on the distribution of BSCs during the growing season (March–November). Notably, the areas with a probability of 1 are larger in March and October than in other months. Additionally, the higher BSC occurrence in the northern part of the desert in March may be attributed to water availability from the nearby Ulungur River in early spring.

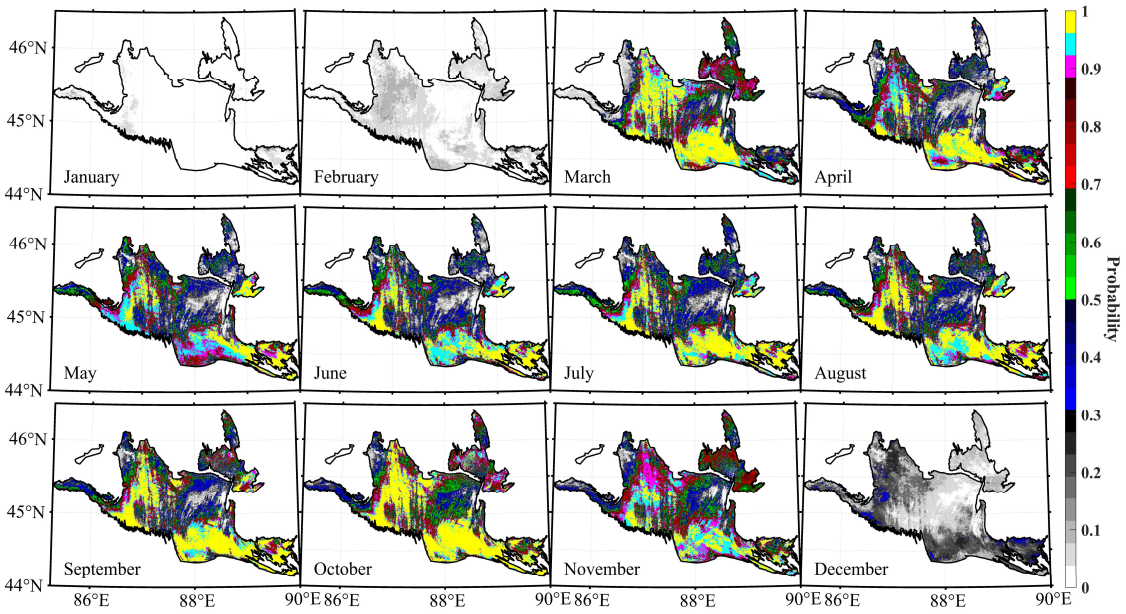


Fig. 3-8 Probability of BSC occurrence in the Gurbantunggut Desert over the period 2000–2018.

Fig. 3-9 shows the changes in BSC coverage fraction during the growing season from 2000 to 2018 relative to 2000. BSC coverage fraction in Fig. 3-9 is defined as the proportion of pixels with detected BSC presence relative to the total number of pixels within the study area. Based on the BSC coverage fraction data from 2000, changes in BSC presence or absence were analyzed at the pixel level. For each pixel, if the BSC detection value transitioned from 0 to 1 (from absence to presence) in a given year relative to its 2000 status, indicating newly detected BSC coverage, it was classified as an increase. Conversely, if the detection value changed from 1 to 0 (from presence to absence), indicating BSC loss, it was categorized as a decrease. Pixels that remained unchanged relative to their 2000 status—either consistently 0 (absence) or 1 (presence)—were classified as stable (no change). Finally, the annual proportions of increased, decreased, and stable (no change) pixels were computed to quantify long-term BSC dynamics over the study period. The blue section of the bar graph represents the area where BSC coverage remained unchanged since 2000, indicating relatively stable BSC presence. This suggests that these areas have maintained long-term BSC coverage in the corresponding months. The green section represents areas where BSCs

were newly detected in regions that lacked BSC coverage in 2000, while the yellow section denotes areas where BSC coverage has declined relative to 2000. For instance, in March 2018, there was a net gain of approximately 20% in BSC cover in certain areas, while a net loss of about 4% occurred in other regions of the Gurbantunggut Desert, resulting in an overall net increase of approximately 16% in BSC coverage over this period. The solid black line represents the average annual change in BSC coverage, enabling comparisons of BSC coverage variations for a given month across consecutive years. For example, the annual average BSC coverage in May 2018 was approximately 48%, while in May 2017, it was around 43%. Based on this, the decrease in BSC coverage in May 2018 relative to May 2017 can be determined to be approximately 5%. The dashed line illustrates the long-term trend in BSC coverage, indicating an overall increase from March to November between 2000 and 2018.

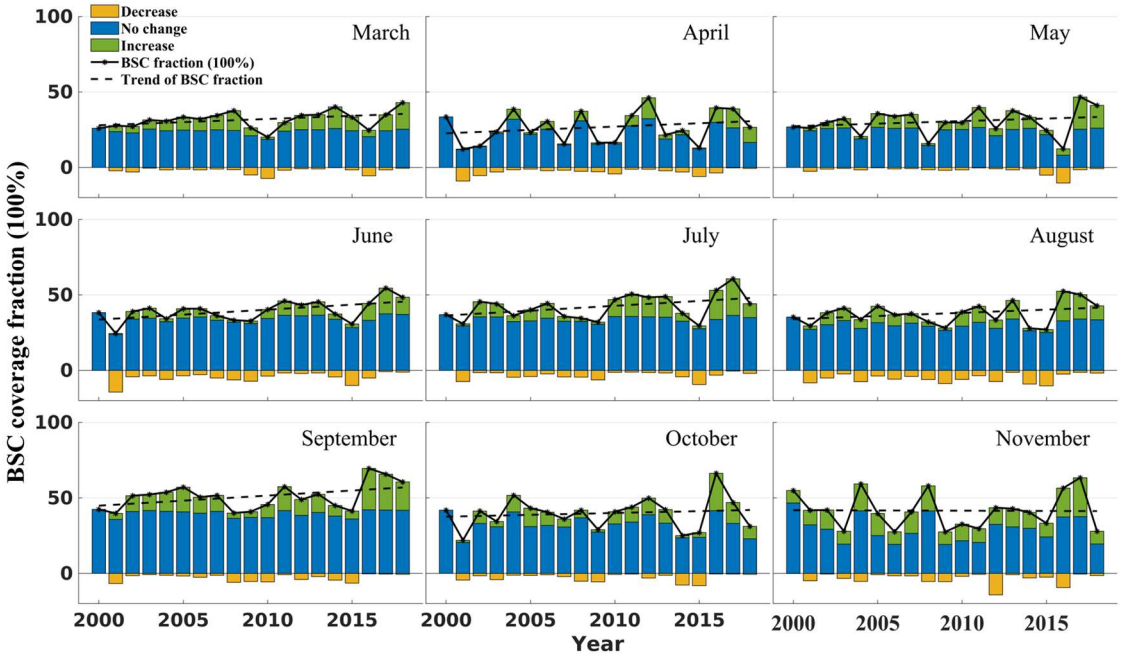


Fig. 3-9 Changes in BSC coverage fraction during the growing season from 2000 to 2018 (baseline: 2000). The yellow, blue, and green regions represent the proportion of pixels where BSC coverage fraction has increased, remained unchanged, and decreased, respectively. The black solid line indicates the annual average change in BSC coverage fraction, while the dashed line represents the overall trend over time.

Fig. 3-10 highlights the relatively stable BSC areas extracted from Fig. 3-9 (blue regions) in the Gurbantunggut Desert. The dark green regions represent pixels where relatively stable BSCs were detected, while the gray areas indicate either the absence of BSCs or high variability in BSC coverage. The proportion of relatively stable BSC coverage from March to November is 23.7%, 16.9%, 12.5%, 16.2%, 20.7%, 21.4%, 26.7%, 31.9%, and 11.3%, respectively. The stable BSC area is most extensive in March and October, predominantly located in the southern and northwestern regions of the desert, aligning with the land cover classification (Fig. 3-1b). The presence of grasslands in these regions likely facilitates BSC formation. Additionally, the southern desert is characterized by widespread sand dunes, where fine sand dominates the interdunal lowlands. In these areas, the silt and clay content significantly increases, affecting soil porosity, bulk density, water retention, permeability, and cohesion. The finer the soil particles and poorer the sediment sorting, the lower the permeability, which in turn enhances conditions for BSC establishment. The presence of fine-grained materials reduces soil porosity, forming a barrier that limits water infiltration, thereby promoting moisture retention—a crucial factor for BSC growth and development (Chen et al., 2005). Moreover, the low-lying interdunal terrain provides an ideal microenvironment for BSC colonization, as temporary water accumulation is common during spring snowmelt and summer precipitation, fostering biological reproduction and vegetative growth (Anderson, 1983). In the southeastern part of the desert, where grasslands and barren land coexist, some relatively stable BSCs are also observed, suggesting that land type heterogeneity plays a role in BSC distribution and persistence.

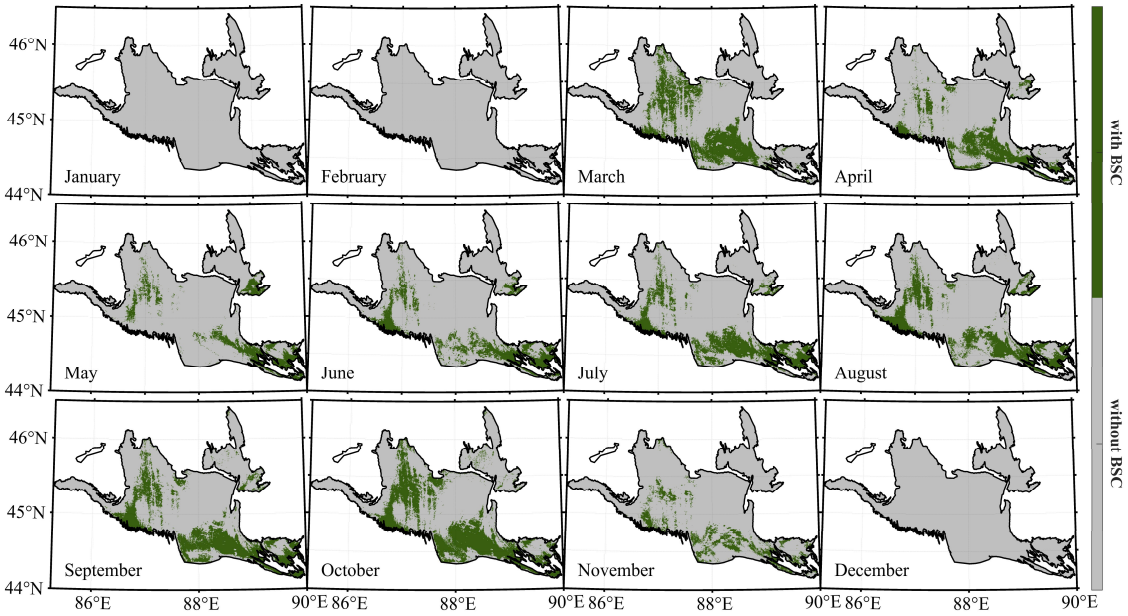


Fig .3-10 Relatively stable BSC area from 2000 to 2018 (baseline: 2000).

Fig. 3-11 presents slope of BSC fraction variability from 2000 to 2018, derived using Sen's method (see Appendix 7.2) and the Mann-Kendall significance test (see Appendix 7.3). In the context of BSC fraction variability, Sen's slope represents the rate of change in BSC coverage over time (2000–2018). A positive slope indicates an increase in BSC coverage, while a negative slope signifies a decline. The steeper the slope, the greater the rate of change. Combined with the Mann-Kendall significance test, it helps determine whether these trends are statistically significant, distinguishing between natural fluctuations and long-term directional changes. In Fig. 3-11, a comparative analysis with Fig. 3-10 shows that Sen's slope values indicate largely stable BSCs in the southern part of the desert, characterized by minimal slope values and little change in coverage over time. In contrast, while relatively stable BSCs are also present in the northwest, they exhibit notably higher variability, as reflected by greater slope values, indicating more pronounced fluctuations in BSC coverage.

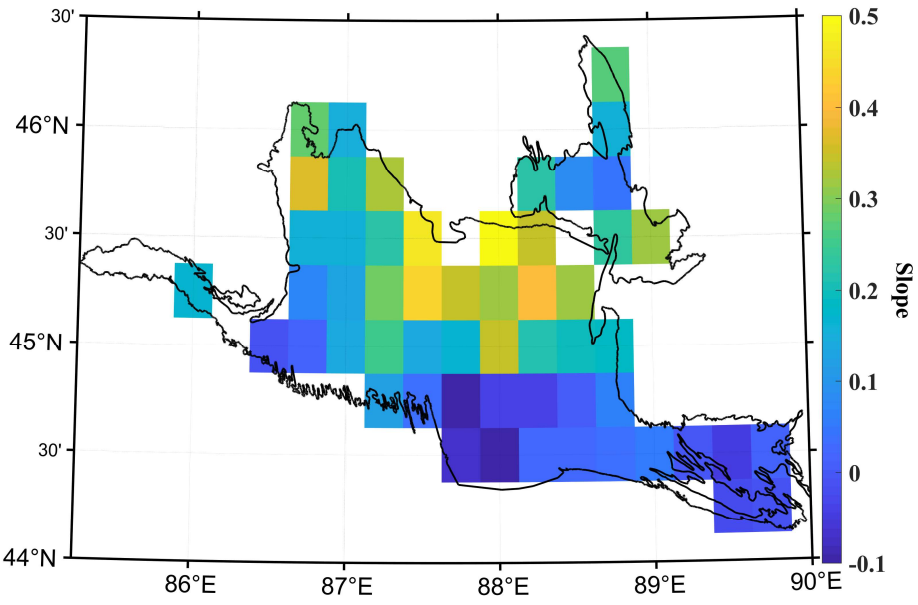


Fig. 3-11 Sen's slope of BSC fraction variability from 2000 to 2018.

3.4.2 Case study: Changes in BSC cover after an extreme dust event

Although BSC has a certain degree of mechanical resistance and sand fixation, the dust event will still affect the distribution of BSC. The FY-3 satellite detected a strong dust event on April 27, 2015 (Fig. 3-12). From the figure, there is a high concentration of dust in the eastern of the study area. Combined with related news reports (http://www.cma.gov.cn/), it can be confirmed that the source area of dust in this event is Gurbantungut Desert.

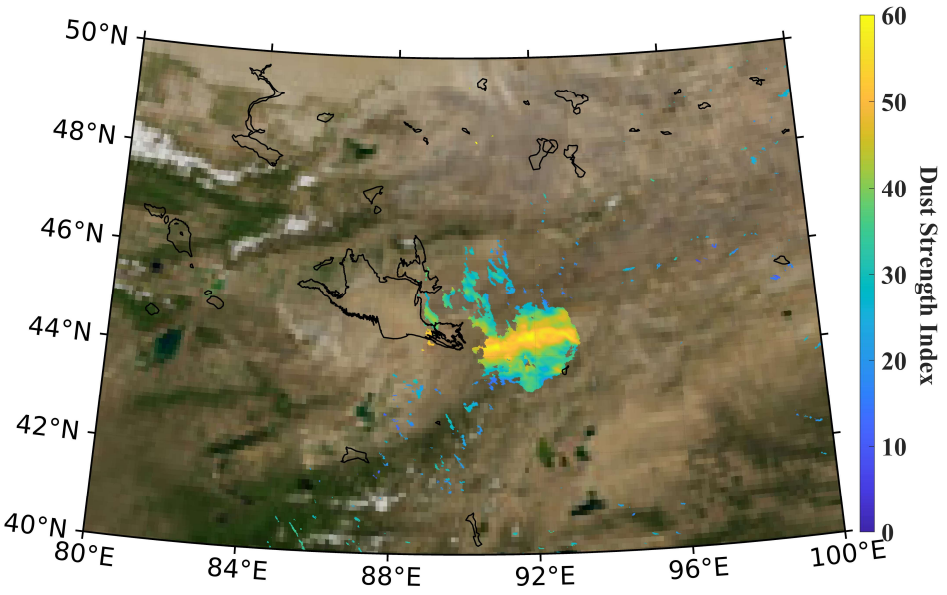


Fig. 3-12 Dust event on April 27, 2015.

Fig. 3-13 presents a comparison of BSC spatial distribution before (a, 2015-04-23) and after (b, 2015-05-01) the dust event on April 27, 2015. In Fig. 3-13a, a distinct stitching line is visible in the central part of the desert, resulting from the edge alignment of the satellite image during the stitching process. The figure illustrates the changes in BSC coverage before and after the dust event, showing a significant reduction in BSC coverage in the central part of the desert, whereas BSCs in the southeastern region remain largely unchanged. Considering Fig. 3-10, the BSCs significantly affected by the dust event in the central Gurbantunggut Desert tend to be less stable. In contrast, the BSCs in the southern desert, which exhibited little change, align with previously identified stable BSC areas, suggesting that dust events have a greater impact on unstable BSCs while having minimal effects on more stable ones. Another potential explanation for the observed differences between the central and southern regions is variation in BSC composition. While previous studies suggest that the southern desert is dominated by lichen-dominated BSCs (Chen et al., 2005), limited research has been conducted on BSC composition in other regions of the desert. Further studies are needed to better understand the distribution and resilience of different BSC types across the Gurbantunggut Desert.

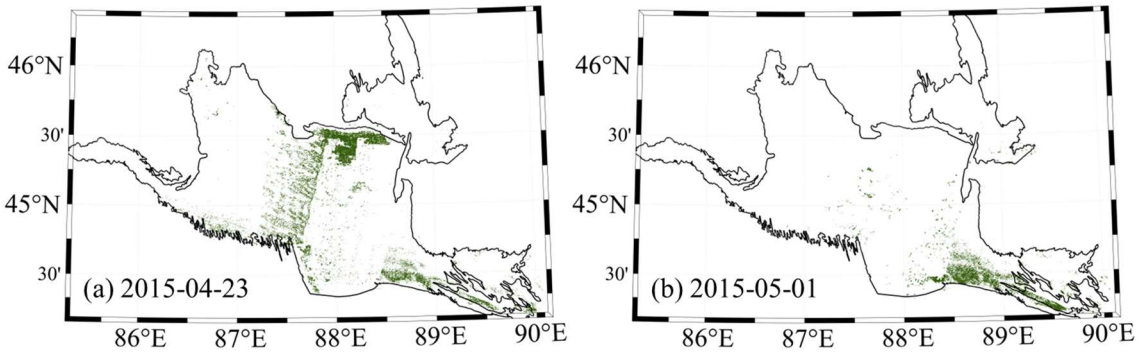


Fig. 3-13 The Changes of BSC spatial distribution before (a) and after (b) the dust event. Green indicates BSC coverage.

3.5 Spatiotemporal variability of climatological processes

In the previous section, the distribution and variability of BSC in the Gurbantunggut Desert were described. Numerous studies have demonstrated that BSCs are highly sensitive to changes in abiotic factors, such as temperature and humidity (Maestre et al., 2013; Shen et al., 2015; Wang et al., 2016). This section analyses changes in various climatic factors in the Gurbantunggut Desert from 2000 to 2018. The analysis is conducted using the same temporal and spatial resolution as the BSC grid data to investigate the spatiotemporal variability of climatic processes in the desert and their correlations.

3.5.1 Hydrologic features

(1) Precipitation

It is important to examine the spatial and temporal patterns of precipitation over the desert, as BSCs are highly sensitive to this critical environmental input. Fig. 3-14 illustrates the seasonal mean precipitation from 2000 to 2018 within the Gurbantunggut Desert. In this figure, precipitation is measured in millimeters per day to account for the varying number of days across the four seasons, ensuring consistency in comparison.

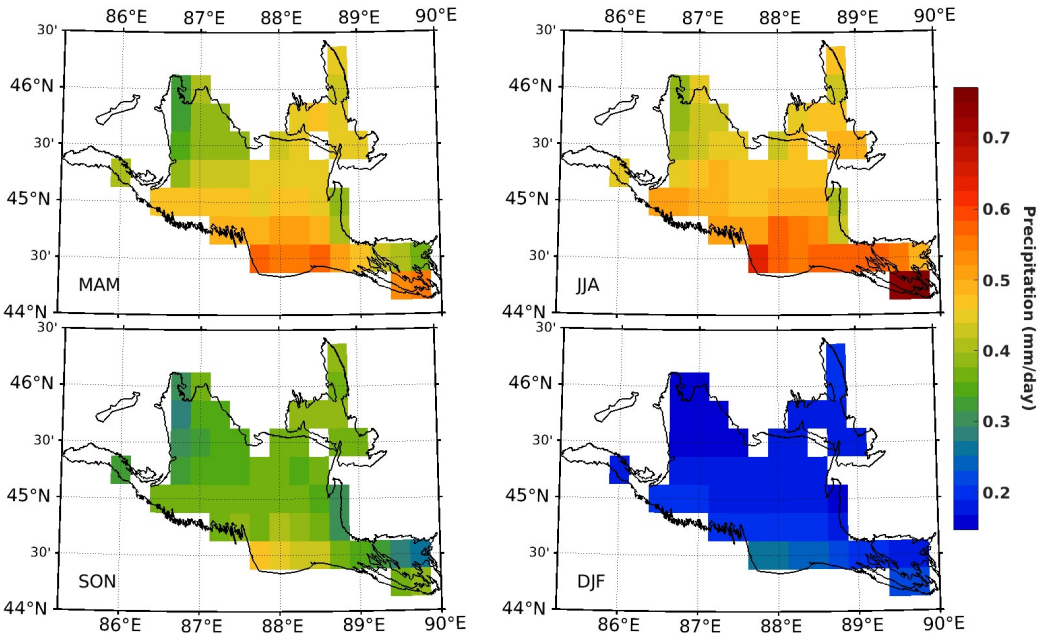


Fig. 3-14 Seasonal spatial distribution of precipitation in the Gurbantunggut Desert from 2000 to 2018.

As shown in Fig. 3-14, precipitation within the desert generally decreases from southeast to northwest. Across all four seasons, regions with higher precipitation are concentrated in the southern and southeastern parts of the desert, particularly in the northern region of the Hutubi River and the desert-oasis ecotone. The annual precipitation pattern is characterized by the highest rainfall occurring in summer, though this difference is not statistically significant compared to spring. In contrast, the lowest precipitation is recorded in winter.

Similarly, the spatially averaged monthly mean precipitation from 1999 to 2018 is presented in Fig. 3-15, along with its relative contributions to annual precipitation. The 20-year average annual precipitation is 132.97 mm. Within the desert, precipitation during the growing season (March to November, generally covering spring, summer, and autumn) accounts for approximately 85% of the total annual precipitation.

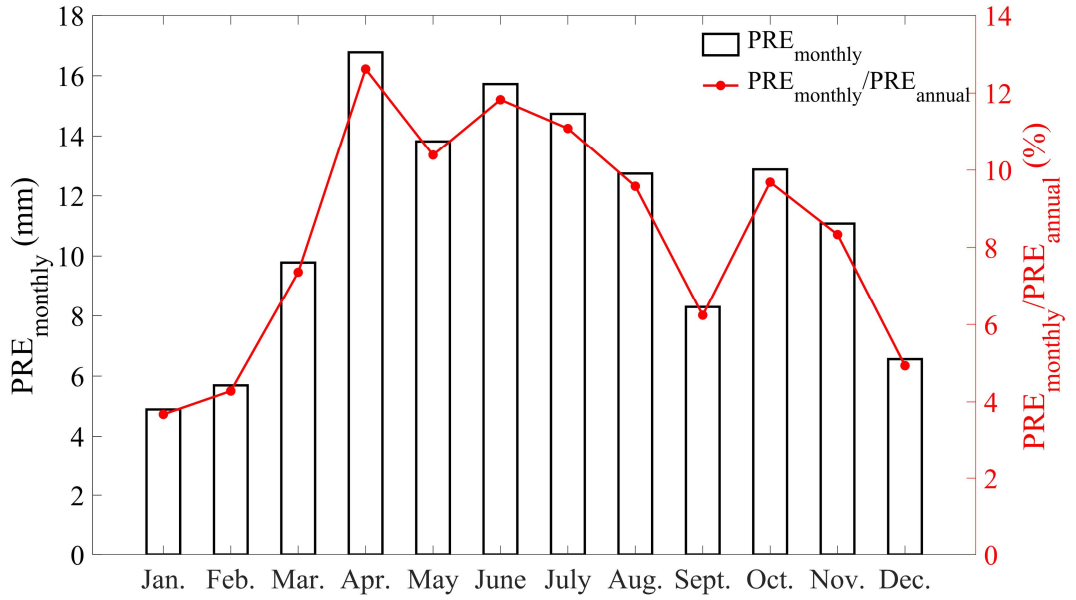


Fig. 3-15 Mean monthly precipitation amount ($P_{monthly}$) in the Gurbantunggut Desert, and the corresponding percentage to annual precipitation ($P_{monthly}/P_{annual}$).

(2) Evapotranspiration

Evapotranspiration (ET) is defined as the total flux of water transferred from the land and ocean surface to the atmosphere through evaporation and plant transpiration. In the following analysis, land evapotranspiration is calculated as the sum of contributions from three evaporation components: direct evaporation from bare soil, evaporation of precipitation intercepted by the vegetation canopy, and vegetation transpiration.

The seasonal mean evapotranspiration from 2000 to 2018 and its spatial distribution within the desert are presented in Fig. 3-16. Similar to precipitation, evapotranspiration is highest in summer and lowest in winter in the Gurbantunggut Desert. Spatially, evapotranspiration increases southeastward across the desert. The lowest evapotranspiration occurs in the northwestern region, whereas the highest values are observed in the interdune lowlands at the southern desert margin. Although the potential annual evaporation in the Gurbantunggut Desert approaches 2000 mm, the actual annual evapotranspiration is limited to 100–200 mm due to soil moisture constraints

and low precipitation. Fig. 3-16 presents evapotranspiration in millimetres per day.

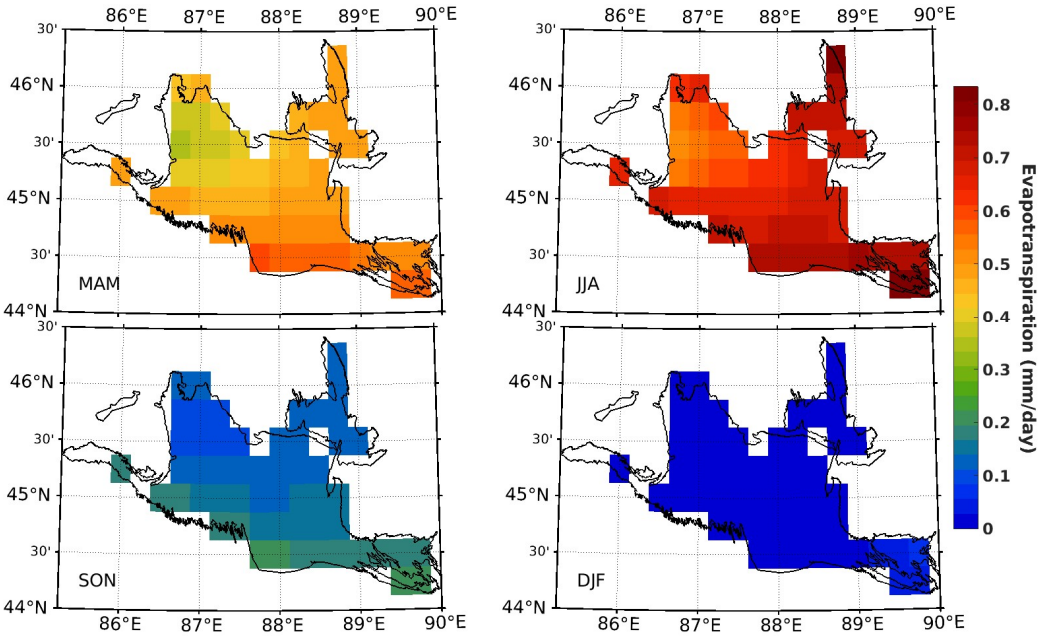


Fig. 3-16 Seasonal spatial distribution of evapotranspiration in the Gurbantunggut Desert from 2000 to 2018.

The mean monthly evapotranspiration from 2000 to 2018 is spatially averaged over the desert and presented in Fig. 3-17, which decomposes evapotranspiration into contributions from bare soil evaporation, plant transpiration, and canopy interception evaporation. Based on the bar heights corresponding to these components, evapotranspiration from canopy interception accounts for the largest fraction, significantly exceeding that from bare soil, while vegetation transpiration remains negligible. This finding suggests that monthly variations in canopy-intercepted evapotranspiration play a crucial role in arid ecosystems, which are typically characterized by sparse vegetation cover. Zhang et al. (2018) demonstrated that the monthly variability of canopy evapotranspiration is closely linked to the vegetation growth cycle in the Gurbantunggut Desert.

The red line in Fig. 3-17 indicates that evapotranspiration peaks in summer (June–August), contributing 52% of the total annual evapotranspiration. Evapotranspiration

in spring exceeds that in autumn, accounting for 36% and 11% of annual evapotranspiration, respectively. The significant decline in evapotranspiration during autumn is attributed to reduced precipitation, which limits soil moisture infiltration and causes water to remain near the surface. Although temperatures remain high and potential evapotranspiration is elevated, soil moisture is not adequately replenished, leading to water stress and subsequently lower evapotranspiration rates.

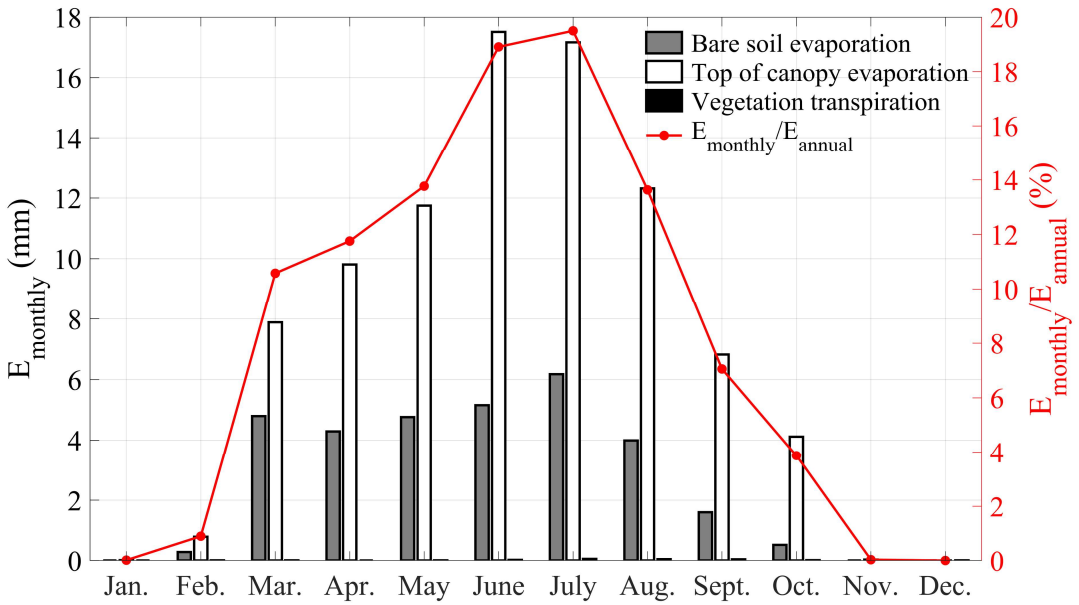


Fig. 3-17 Mean monthly evapotranspiration amount ($E_{monthly}$) in the Gurbantunggut Desert with the amounts for various components, and the percentage of monthly to annual evapotranspiration ($E_{monthly}/E_{annual}$).

(3) Relationship between precipitation and evapotranspiration

Precipitation and evapotranspiration are the primary components of water input and loss in natural desert ecosystems and play a crucial role in regulating their water balance.

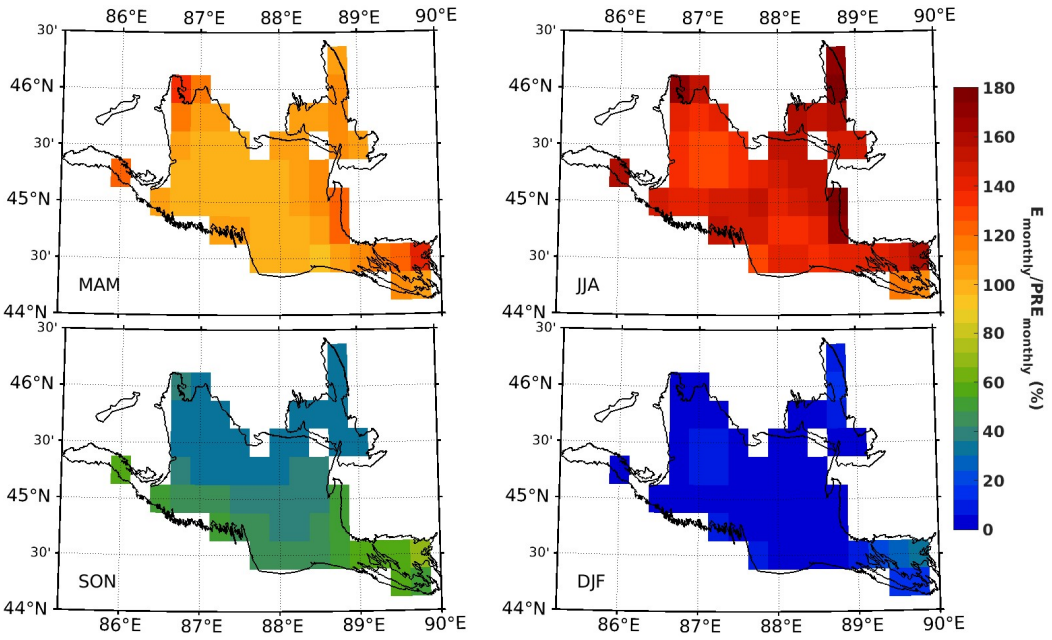


Fig. 3-18 Seasonal distribution of evapotranspiration as a percentage of precipitation in the Gurbantunggut Desert from 2000 to 2018.

As shown in Fig. 3-18, the seasonal distribution of evapotranspiration as a percentage of precipitation in the Gurbantunggut Desert is approximately balanced in spring. In summer, evapotranspiration is nearly twice the amount of precipitation, whereas in autumn, evapotranspiration is lower than precipitation and is proportionally higher in the south than in the north. Winter, as a non-growing season, is characterized by minimal evapotranspiration and precipitation, both approaching zero.

Fig. 3-19 illustrates that in January, February, November, and December, despite some precipitation, evapotranspiration remains negligible due to the dormancy or wilting of biocrusts and vegetation, as well as low temperatures. During the growing season, evapotranspiration falls below precipitation in April and October. Precipitation during these months not only fully meets evapotranspiration demand but also partially replenishes soil moisture storage, providing a water source for subsequent BSC growth phases. Since soil retains moisture, evapotranspiration increases significantly in March. Additionally, evapotranspiration reaches peak values in June and July, coinciding with higher rainfall. However, throughout the growing season, $E_{monthly}/PRE_{monthly}$ typically

remains below 100%, indicating that precipitation alone cannot fully meet evapotranspiration demands. All precipitation is consumed through evapotranspiration, and soil water storage is required to compensate for the evapotranspiration deficit.

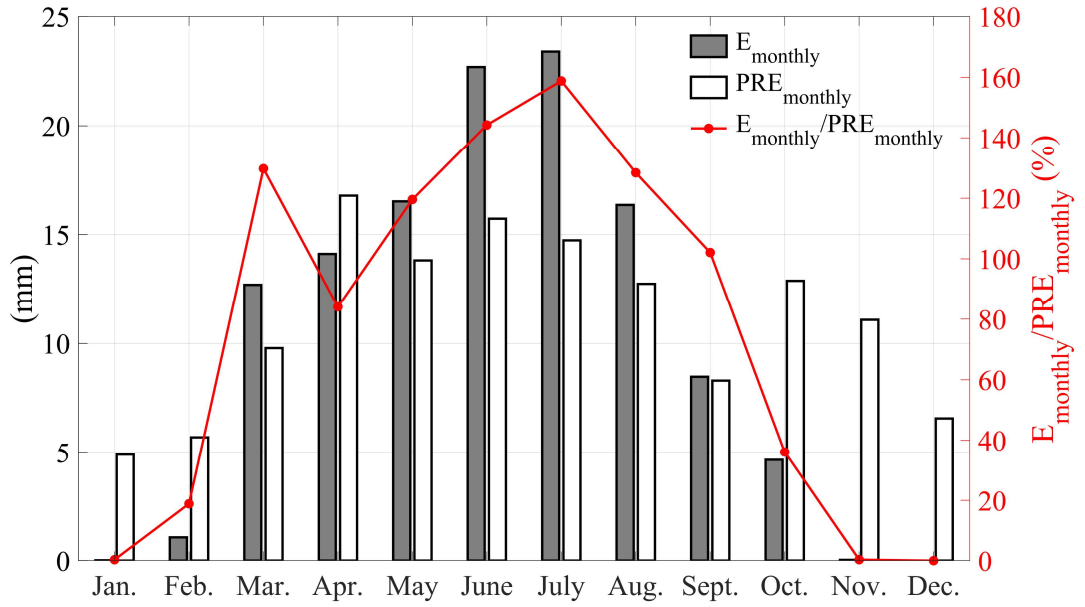


Fig. 3-19 Mean monthly evapotranspiration ($E_{monthly}$) and Precipitation ($PRE_{monthly}$) in the Gurbantunggut Desert, and evapotranspiration as a percentage of precipitation ($E_{monthly}/PRE_{monthly}$).

3.5.2 Humidity

Specific humidity, also referred to as moisture content, is defined as the ratio of the mass of water vapor to the total mass of an air parcel (Byers, 1965). The Gurbantunggut Desert is an extremely arid region characterized by low precipitation, sparse vegetation cover, and low specific humidity. Fig. 3-20 illustrates a pronounced seasonal trend in specific humidity, with the highest values occurring in summer and the lowest in winter. As autumn and spring serve as transitional seasons, the average specific humidity during these periods falls between winter and summer, with autumn exhibiting slightly higher values than spring. Spatially, specific humidity increases from the northeast to the southwest, reaching its maximum in the desert-oasis transition zone along the southern desert margin. Overall, the annual specific humidity in the Gurbantunggut Desert remains low, ranging from 1 to 7 × 10^{-3} kg/kg.

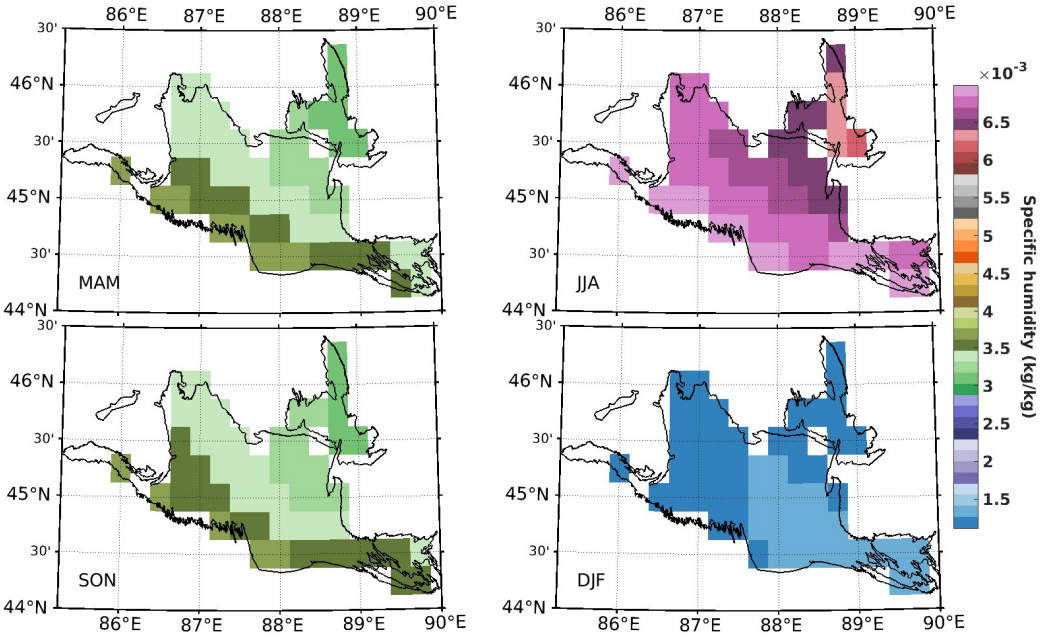


Fig. 3-20 Seasonal spatial distribution of specific humidity in the Gurbantunggut Desert from 2000 to 2018.

The correlation between monthly hydrological variables—precipitation, evapotranspiration, and specific humidity—in the Gurbantunggut Desert from 2000 to 2018 is examined. Fig. 3-21(a) illustrates that there is no significant linear correlation between the monthly average precipitation and specific humidity. The coefficient of determination (R^2) is 0.22, with numerous data points falling outside the 95% confidence interval, indicating that a linear model is unsuitable for describing the relationship between precipitation and specific humidity. In other words, these two variables exhibit minimal collinearity. Similarly, Fig. 3-21(b) shows that although the linear fit between evapotranspiration and specific humidity yields an R^2 value of 0.65, many data points still lie outside the confidence interval, suggesting that the relationship between the two remains weak and lacks strong collinearity.

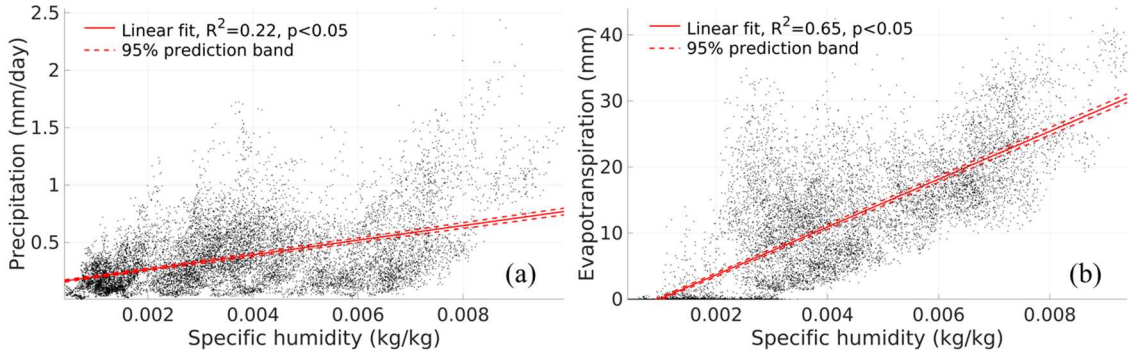


Fig. 3-21 The correlation between monthly precipitation (a) and evapotranspiration (b) and specific humidity from 2000 to 2018 in the Gurbantunggut Desert.

3.5.3 Temperature

The seasonal spatial distribution of temperature in the Gurbantunggut Desert from 2000 to 2018 is presented in Fig. 3-22. The seasonal temperature pattern follows the order: summer > spring > autumn > winter. The annual temperature variation is substantial, with a difference exceeding 40°C between the highest and lowest recorded temperatures. Spatially, temperatures in spring, summer, and autumn exhibit a strong correlation with elevation, as depicted in Fig. 3-3. From north to south and east to west, temperature increases as elevation decreases. However, in winter, the spatial pattern differs, with higher temperatures in the east and lower temperatures in the west.

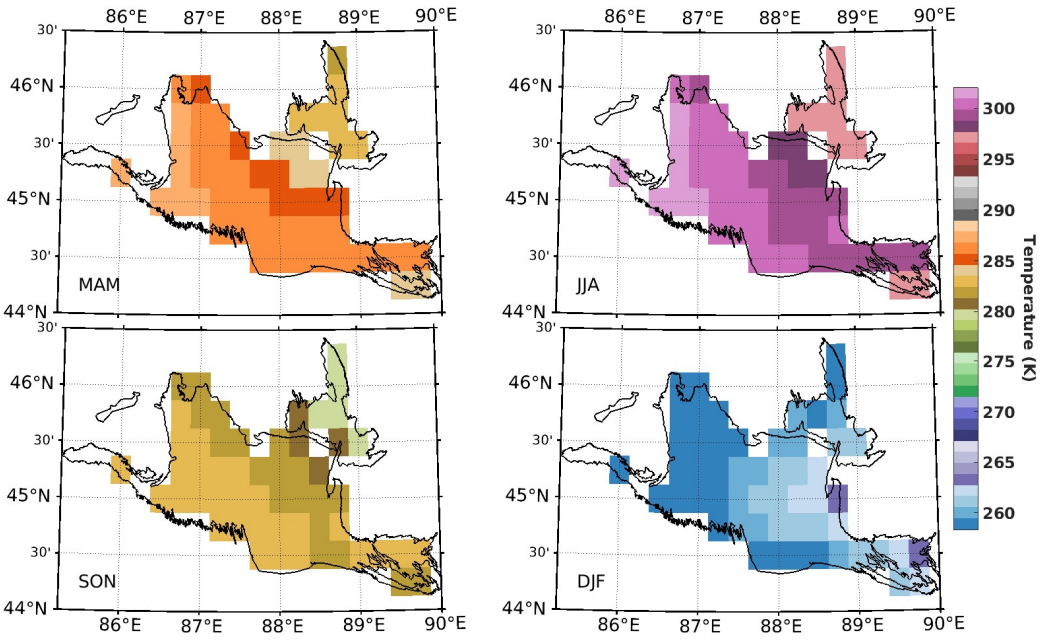


Fig. 3-22 Seasonal spatial distribution of temperature in the Gurbantunggut Desert from 2000 to 2018.

As shown in Fig. 3-23(a), the correlation between monthly specific humidity and temperature in the Gurbantunggut Desert follows a non-linear relationship, with a coefficient of determination (R^2) of 0.89 and a significance level of $p < 0.05$. In Fig. 3-23(b), the linear correlation between wind speed and temperature is less pronounced, with an R^2 value of 0.63 and a greater degree of scatter.

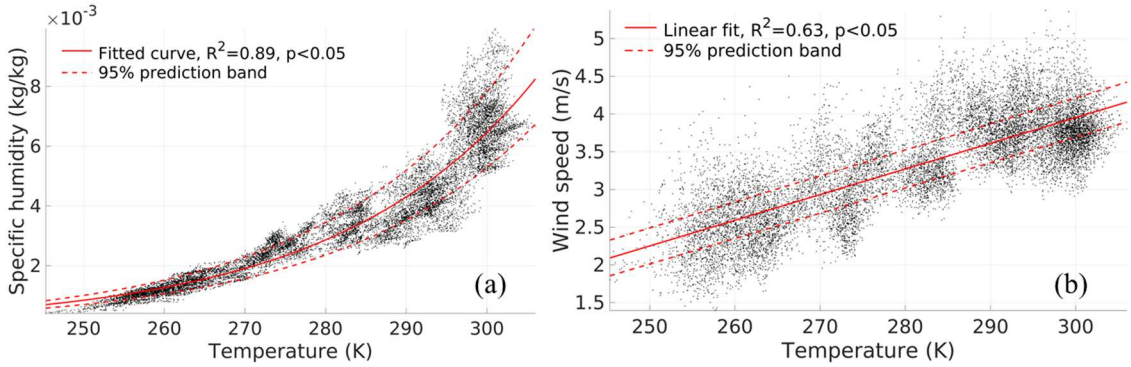


Fig. 3-23 The correlation between monthly specific humidity (a) and wind speed (b) and temperature from 2000 to 2018 in the Gurbantunggut Desert.

3.5.4 Wind speed

The Gurbantunggut Desert is influenced by westerly winds year-round. In winter, the Siberian-Mongolian high-pressure system dominates, resulting in northwesterly and westerly winds. In summer, subtropical high-pressure systems prevail, generating predominantly westerly winds. As shown in Fig. 3-24, average wind speed is higher in summer and autumn, ranging from 3.5 to 4.5 m/s, and lower in winter, particularly in the southwestern part of the desert, where it averages around 2 m/s. Overall, wind speeds in the interior of the desert are higher than those in the surrounding areas. Additionally, high wind speeds are observed along the northeastern and southeastern desert margins.

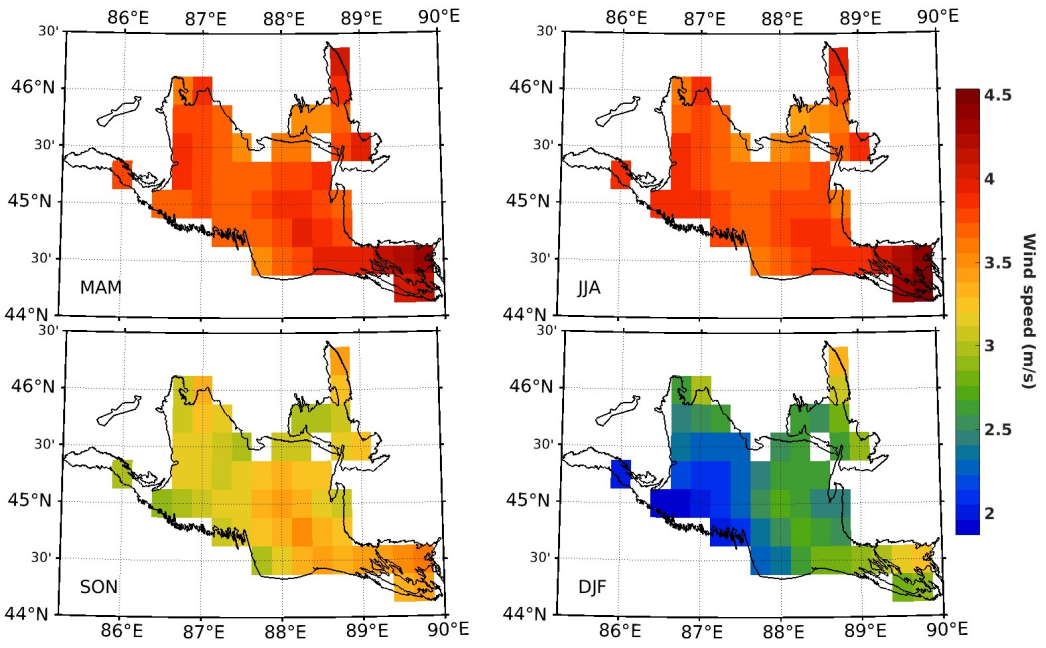


Fig. 3-24 Seasonal spatial distribution of 10-meter wind speed in the Gurbantunggut Desert from 2000 to 2018.

The wind speed distribution in the Gurbantunggut Desert, as depicted in Fig. 3-25, indicates that 97% of recorded wind speeds are below 8 m/s. Wind tunnel experiments (Wang et al., 2004) have demonstrated that the threshold friction velocity for bare sand in this desert is 8.42 m/s. When wind speed exceeds 18 m/s, surface wind erosion occurs under most disturbance conditions, except in areas with undisturbed algae, algal-lichen

crusts, or lichen crusts with less than 20% damage. This suggests that, in most cases, wind speed is not the primary factor influencing biological crust dynamics in the Gurbantunggut Desert.

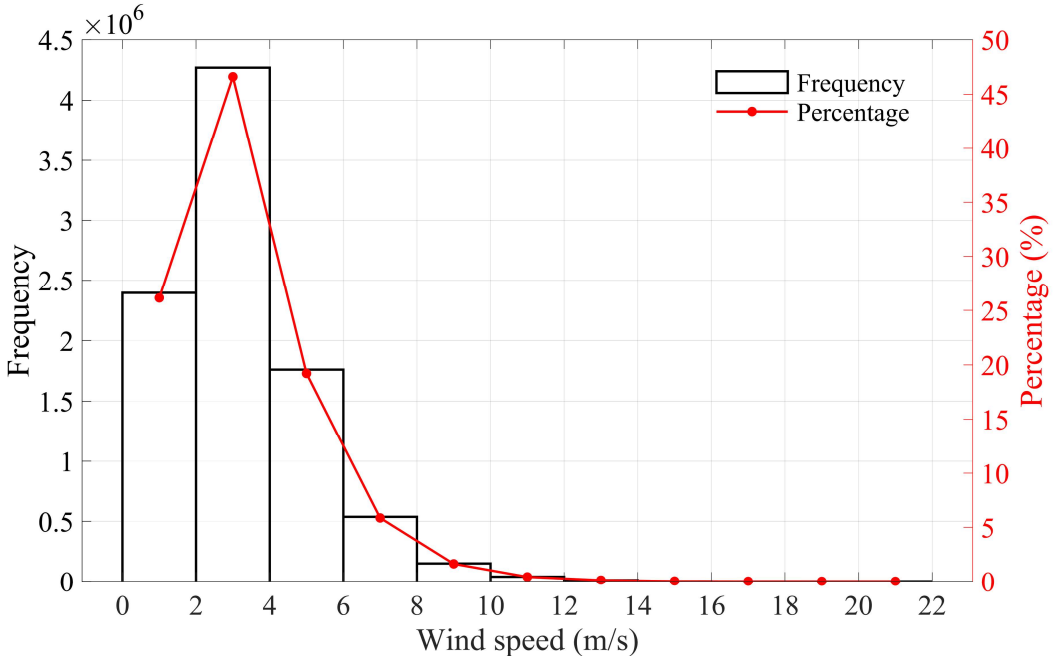


Fig. 3-25 The frequency histogram of wind speed in the Gurbantunggut Desert from 2000 to 2018.

This section analysed the temporal and spatial distribution of climate variables in the Gurbantunggut Desert and their interrelationships. The climatic processes in the Atacama Desert are discussed in the model application chapter (see Section 5.2.1 for details).

4 Construction of the BSC-Climate model

4.1 Lag-correlation between BSC and climate variables

4.1.1 Temporal effects of each climatic factor

Dust events in Asia predominantly occur in the Southern Xinjiang Basin, Mongolian Plateau, and Hexi Corridor, while strong sand and dust events are relatively rare in the Gurbantunggut Desert. Therefore, the impact of local environmental factors on BSC is primarily considered in this study. To evaluate the temporal effects of climatic factors on BSC, we applied linear regression to each climatic variable separately (Eq. 4-1):

$$BSC_{t,grid} = \beta_{t,grid} \times Climate_{t-i} + \alpha_{t,grid} \quad (4\text{-}1)$$

where α and β are regression coefficients, t represents time, i ranges from 0 to 3 ($i = 0$ represents no time lag, while $i = 1$ to 3 represents a one- to three-month lag), and *grid* ranges from 1 to 55 (corresponding to the climate data grid points introduced in Section 3.2.4). *Climate* represents the climatic variables, including specific humidity (SH), 10-meter wind speed (WS), 2-meter temperature (TMP), surface downward solar radiation under clear-sky conditions (SR), and total precipitation (PRE). By testing different values of i, we comprehensively assess the temporal effects of climate variables. The correlation coefficient is used to determine the optimal lag time (i) for the best model fit.

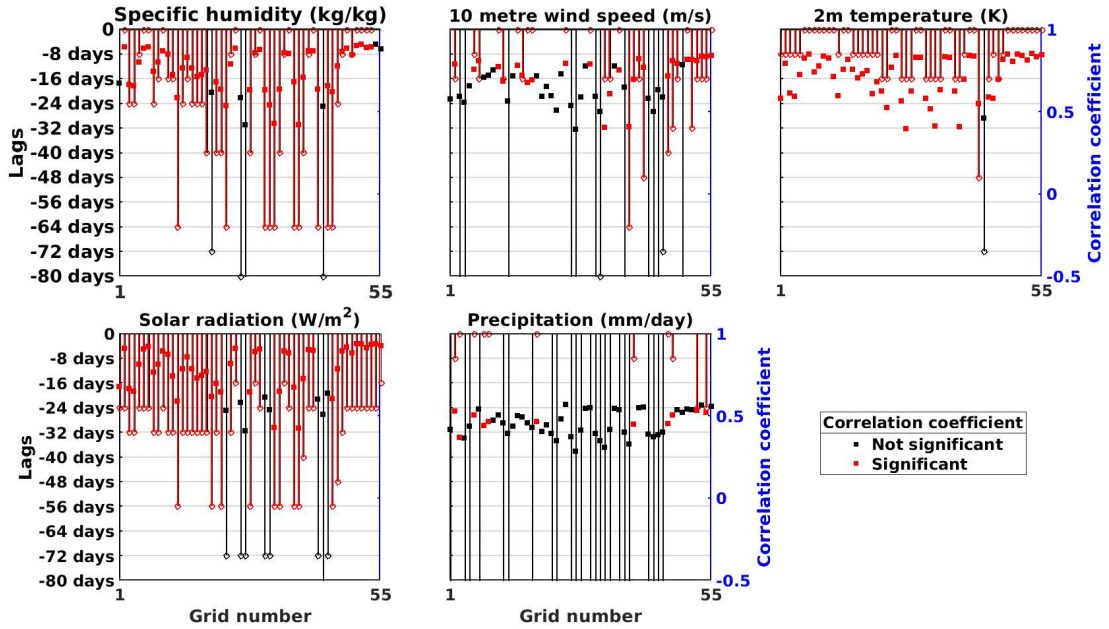


Fig. 4-1 Lag-correlation between BSC and climate variables during 2000-2018 with the time window of 8 days.

The lag correlation between BSC and climatic variables from 2000 to 2018, using an 8-day time window, is presented in Fig. 4-1. By calculating the sliding correlation coefficient, we identified the climatic variables most strongly correlated with BSC and their corresponding time lags. In Fig. 4-1, the horizontal axis represents the grid point number, the left vertical axis denotes the time lag (days), and the right vertical axis indicates the correlation coefficient. The lengths of the red and black lines illustrate the dominant time lag for each climatic variable at each grid point. The red and black solid points represent the correlation coefficient corresponding to that dominant time lag, where red indicates statistical significance and black represents non-significance. Among the climatic variables, BSC exhibits the fastest response to temperature, typically within 0–8 days. Although precipitation correlations are relatively low (around 0.5), the time lag between significant grid points and BSC dynamics is minimal or absent. Overall, the time lags of all significant climatic variables do not exceed 64 days, and the corresponding correlation coefficients are generally above 0.5.

Fig. 4-2 illustrates the spatial distribution of time lags across the study area. Along the

southern desert-oasis boundary, the response of BSC to all climatic variables is relatively rapid, with correlation coefficients exceeding 0.7. The longest time lag for specific humidity occurs in the northern part of the desert, exceeding 56 days. A similar pattern is observed for solar radiation, where the lag generally ranges from 24 to 32 days, but extends beyond 40 days in the northern desert region. The spatial distribution of time lags for wind speed and precipitation exhibits a similar band-like pattern, with a zero-time lag zone extending from northwest to southeast. In contrast, the time lag for temperature does not exceed 16 days across the desert, with the longest lag occurring in the central hinterland and the northern bare soil region.

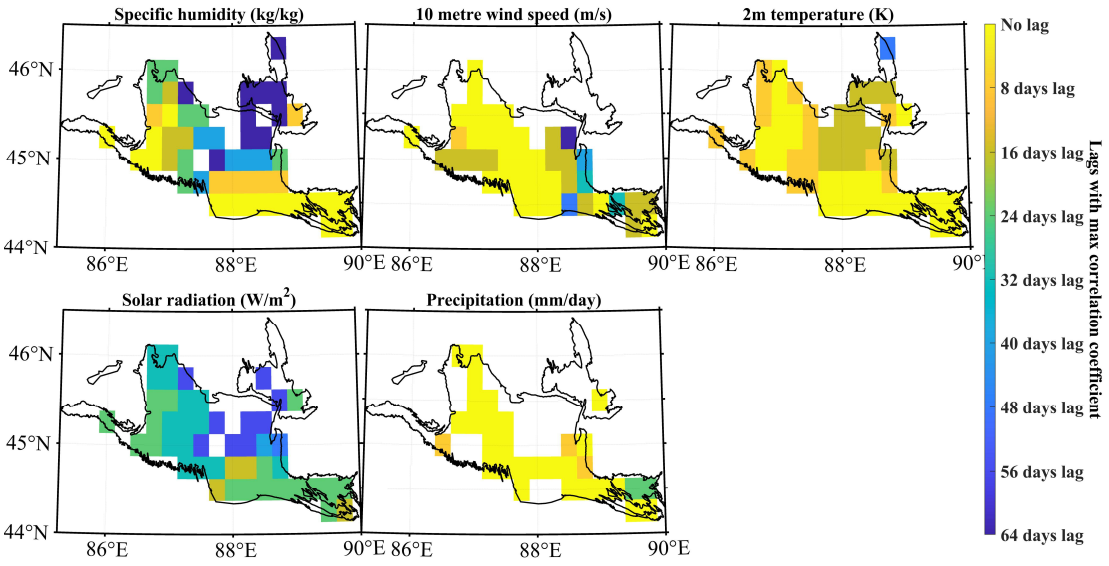


Fig. 4-2 Distribution of lags between BSC and climate variables during 2000-2018 with the time window of 8 days.

Fig. 4-3 presents the lagged correlation between BSC and climatic variables from 2000 to 2018, using a monthly time window. This figure follows a similar pattern to Fig. 4-1, but with an extended time window designed for modeling based on monthly data. The time lag for specific humidity and solar radiation is primarily 1–2 months, while for wind speed and temperature, it is 0–1 month. For precipitation, the lag is effectively zero. Most of the significant correlation coefficients exceed 0.5, aligning with the results obtained from the 8-day dataset.

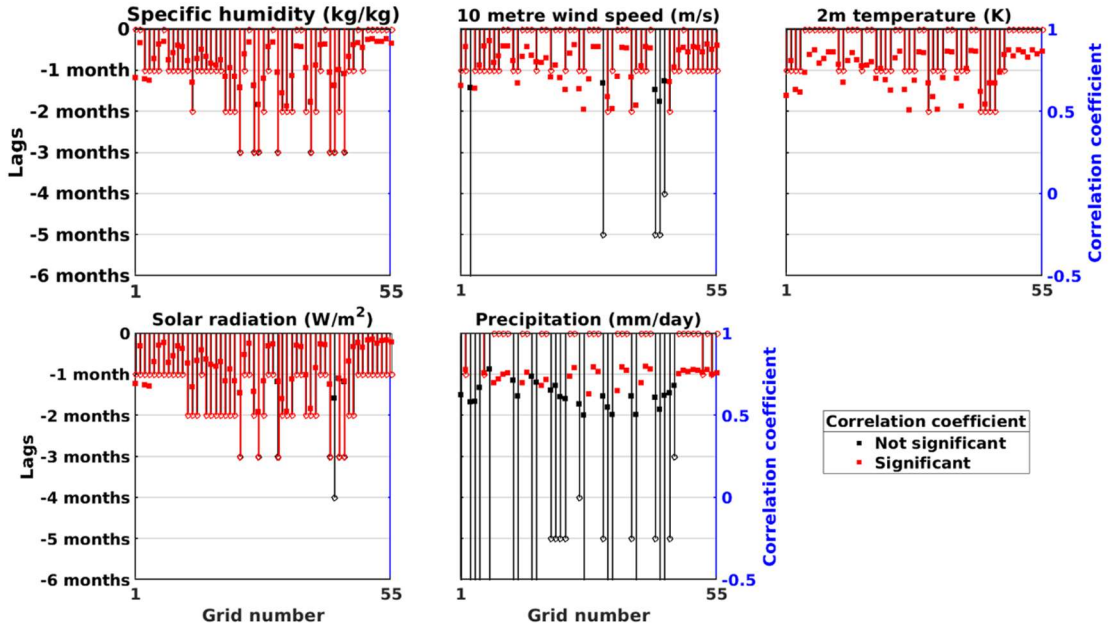


Fig. 4-3 Lag-correlation between BSC and climate variables during 2000-2018 with the time window of month

The spatial distribution of time lags for climatic variables in Fig. 4-4 is consistent with that in Fig. 4-2. Time lags are shorter in the southern and northwestern regions of the desert compared to the central and northeastern areas.

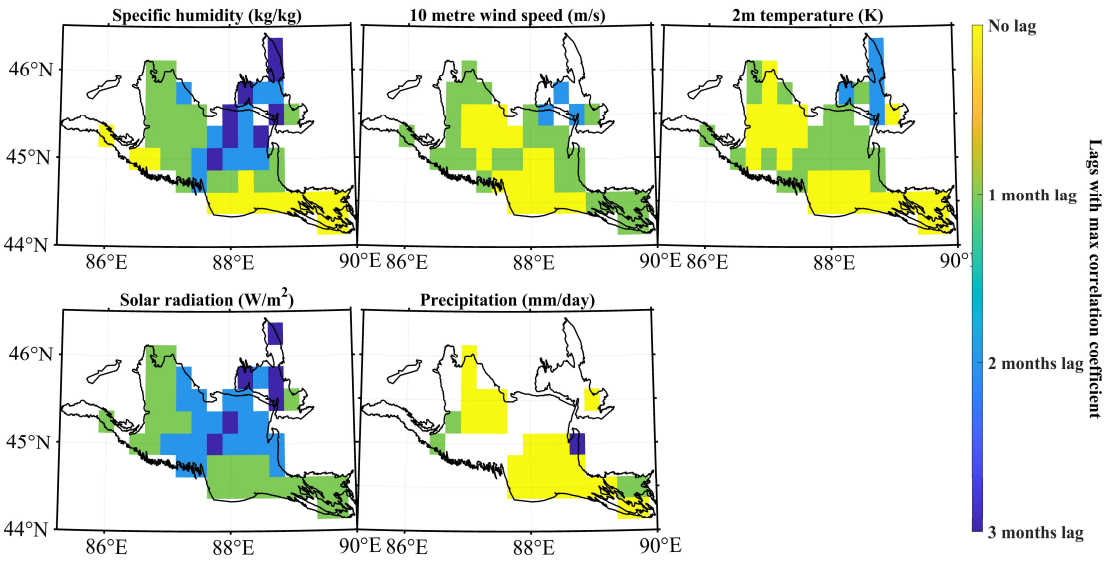


Fig. 4-4 Distribution of lags between BSC and climate variables during 2000-2018 with the time window of month

At this stage, we can infer that climatic factors exert time-lagged effects on BSC growth. In the Gurbantunggut Desert, where BSC survival is constrained by climatic variability, fluctuations in climate conditions have a significant impact on BSC coverage. The results indicate that the time-lag effect at a monthly scale is generally shorter than one quarter, aligning with previous research (Anderson et al., 2010; Chen et al., 2014; Ivits et al., 2016; Rundquist and Harrington, 2000; Vicente-Serrano et al., 2013; Wu et al., 2015). Consequently, this study considers time-lag effects up to a maximum of 3 months.

4.1.2 Partial correlation analysis of climate variables and BSC

To identify the dominant climatic factors driving BSC variation during the growing season, while accounting for time-lag effects, we perform partial correlation analysis between BSC and individual climatic variables. This analysis controls for the influence of other climatic factors, ensuring that the observed relationships reflect independent effects.

Table 4-1 Time lag (months) across different land types.

	Mean					Standard deviation				
	SH	WS	TMP	SR	PRE	SH	WS	TMP	SR	PRE
Whole aera	1.70	2.04	0.60	0.82	1.09	0.78	0.75	0.63	0.71	0.81
Barren	1.65	2.20	0.59	0.75	0.81	0.86	0.64	0.60	0.64	0.80
Grassland	1.74	1.93	0.60	0.91	1.40	0.91	0.93	0.68	0.76	0.83

Specific humidity exhibited a 1.70 ± 0.78 (mean ± standard deviation) month lag in the whole area (Table 4-1). The time lag is consistently two months throughout the desert, except at the southwestern and northeastern edges (Fig. 4-5). In terms of correlation, values exceed 0.5 across the study area, except in the central hinterland (Fig. 4-6), where the time lag is two months and specific humidity correlation remains low. Notably, at the southern edge, the correlation exceeds 0.8, which may be related to

agricultural land, roads, livestock production, and other human activities in the vicinity, all of which are known to contribute to biocrust disturbance (Zaady et al., 2016).

Although BSC enhances surface roughness, thereby increasing frictional resistance to wind-driven sand movement, it becomes vulnerable when wind speed exceeds its threshold. Across the region, wind speed has a mean lag of 2.04 ± 0.75 months (Table 4-1). In terms of land use, most grid points with significant wind speed effects are located in grassland areas. Overall, BSC responds slowly to wind speed, with a three-month lag in the northwestern desert (Fig. 4-5). Furthermore, correlations are very low, approaching zero, with some areas even displaying a negative correlation (Fig. 4-6). This suggests that wind speed exerts a weak and negative influence on BSC dynamics, primarily because wind speeds in the Gurbantunggut Desert generally remain below 6 m/s, insufficient to exceed the frictional threshold required to degrade BSC.

Temperature exhibits an average time lag of 0.60 ± 0.71 months (Table 4-1), with minimal variation across different land types. The temperature lag is consistently zero across most of the desert, except at the northeastern and northwestern fringes (Fig. 4-5). In grassland areas, the correlation is significantly higher than in other regions, reaching 0.5 (Fig. 4-6). These findings suggest that temperature plays a crucial role in shaping BSC distribution in the Gurbantunggut Desert.

The time lag for solar radiation is 0.82 ± 0.63 months (Table 4-1). Its spatial distribution closely resembles that of 2-meter temperature, with a near-zero lag throughout most of the desert. However, unlike temperature, the solar radiation time lag abruptly increases to three months in the southwest. Similarly, the spatial correlation distribution of solar radiation mirrors that of temperature, with higher correlations (~0.4) in grassland areas (Fig. 4-6). In contrast, north of the central desert hinterland, the correlation is close to zero with a one-month lag.

Precipitation exhibits an average lag of 1.09 ± 0.81 months (Table 4-1), with its most significant effects occurring at the desert edges. The western edge shows no lag, the southern edge exhibits a one-month lag, and the northeastern edge experiences a 2–3 month lag. In terms of correlation, most areas show low values, except for grassland at the southern edge, where the correlation coefficient is approximately 0.3. Two possible explanations exist for this pattern: (1) The arid climate and low precipitation levels in the Gurbantunggut Desert result in rainfall amounts that are insufficient to surpass evapotranspiration, thereby limiting BSC growth. (2) Reanalysis data may fail to accurately capture the variability of low-magnitude rainfall events, leading to discrepancies when compared to observational datasets (Guo et al., 2024).

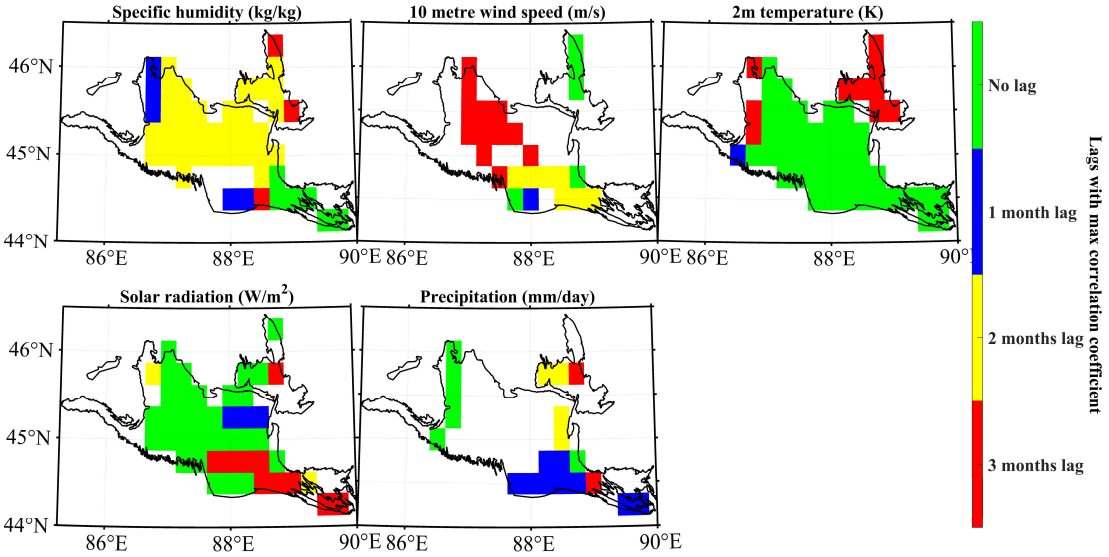


Fig. 4-5 Time lag across the entire study area for different climate variables.

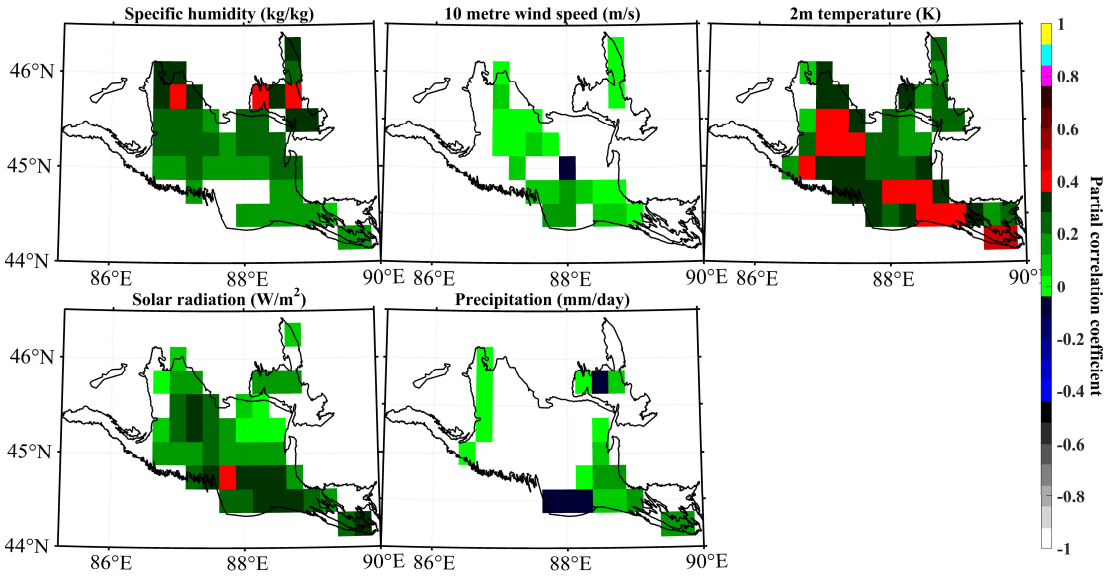


Fig. 4-6 Partial coefficient coefficients across the entire study area for different climate variables.

By employing partial correlation for analysis, the lag times and correlation coefficients obtained differ from those derived using Pearson's correlation (as discussed in Section 4.1.1). This difference arises because partial correlation removes the linear effects of other variables. In certain cases, Pearson's correlation may not accurately capture the true relationship between independent and dependent variables due to multicollinearity.

4.2 Identification of model variables

The climate data used for both training and testing the model are sourced from the ERA5 dataset. Given that the model will be applied to the Last Glacial Maximum (LGM) and the Mid-Holocene in future research, paleoclimate data from the CMIP6 dataset will be required. Therefore, all variables must first be normalized, followed by bias correction for the CMIP6 data.

4.2.1 Normalization

To eliminate the influence of magnitude on the model, all data must be normalized. For CMIP6 data, the normalization of AWI, MPI, and MIROC model data is performed using the maximum and minimum values from the ERA5 dataset. This ensures that the

training and prediction sets remain independent and identically distributed. The normalization process is conducted separately for each grid point, as the regression model is independently established for each grid point.

4.2.2 Bias correction

The comparison of the annual cycle and standard deviation between the ERA5 and CMIP6 datasets is presented in Fig. 4-7 and Fig. 4-8. Solar radiation and temperature exhibit similar patterns, with only minor differences. However, specific humidity and wind speed show noticeable discrepancies in their mean values. Precipitation, in particular, exhibits significant differences, with AWI and MPI values in summer being much lower than those of ERA5. Additionally, the standard deviation distribution of precipitation is even opposite in trend. Therefore, bias correction based on ERA5 data is essential when using ERA5 as the sole dataset for statistical model training.

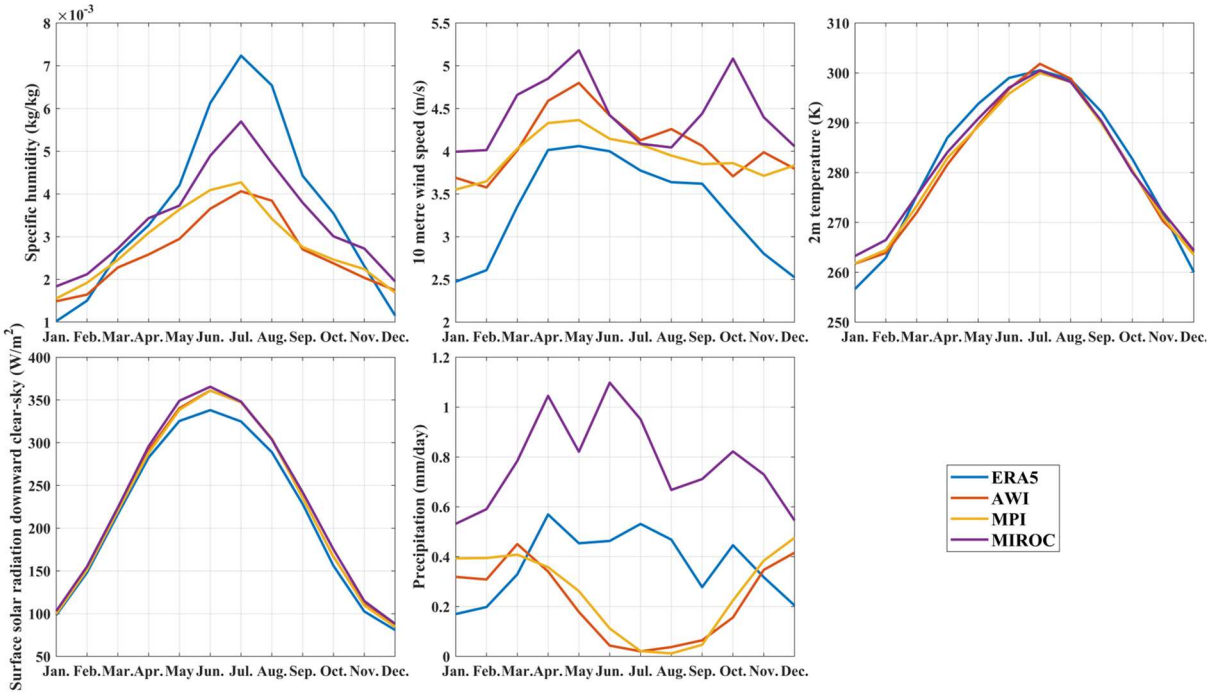


Fig. 4-7 The comparison of annual cycle among ERA5 and CMIP6 datasets for 1999-2014.

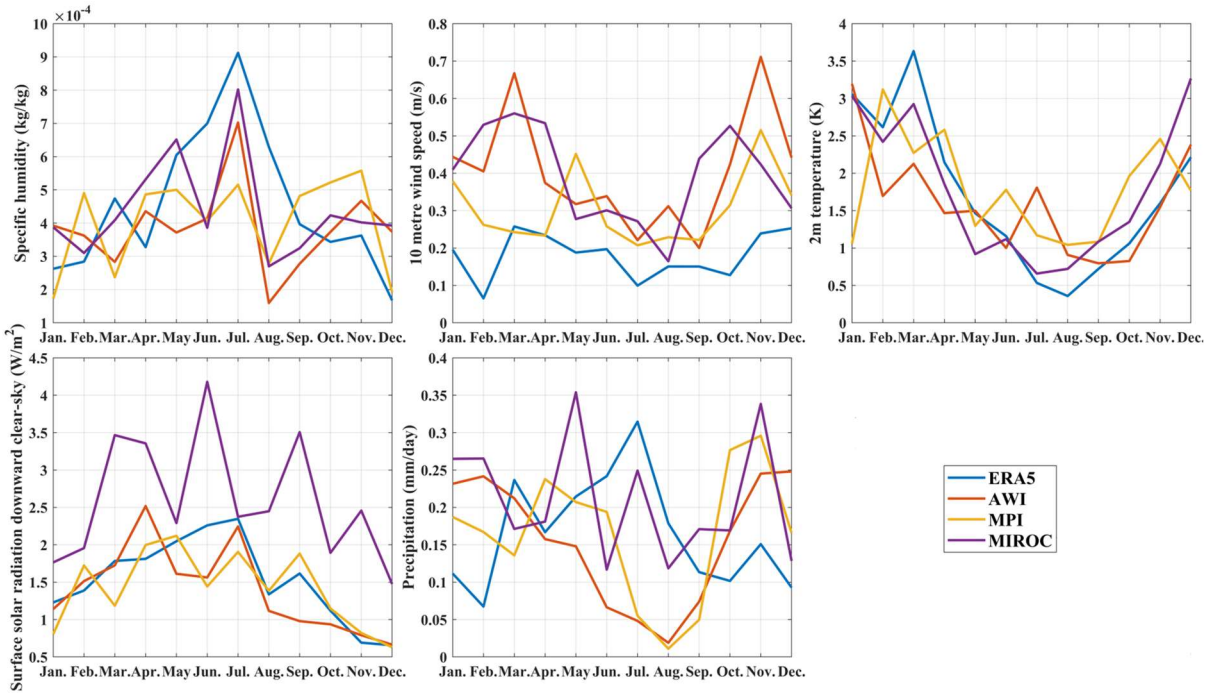


Fig. 4-8 The comparison of standard deviation among ERA5 and CMIP6 datasets for 1999-2014.

The general form of the bias correction method (Eq. 4-2) is applied to climate variables other than precipitation. This method adjusts not only the mean values but also the temporal variability of the model output to align with observations (Hawkins et al., 2013; Ho et al., 2012):

$$M_{BC}(t) = \overline{O_{REF}} + \left(\frac{\sigma_{O_{REF}}}{\sigma_{M_{REF}}}\right) * \left(M_{RAW}(t) - \overline{M_{REF}}\right) \quad (4\text{-}2)$$

In this equation, O_{REF} represents the ERA5 data for the historical reference period, M_{REF} denotes the CMIP6 output from the historical reference period, and M_{RAW} corresponds to the raw CMIP6 output for the historical, future, or paleo period. M_{BC} is the bias-corrected CMIP6 output. Additionally, $\sigma_{O_{REF}}$ and $\sigma_{M_{REF}}$ represent the standard deviations of the observations and daily CMIP6 output in the reference period, respectively.

Precipitation is a stochastic variable in the study area and requires a more sophisticated approach for bias correction. To accurately bias-correct CMIP6 output for monthly

totals and wet-day frequency while maintaining realistic daily and interannual variability, the Equiratio Cumulative Distribution Function (EQCDF) matching method (Eq. 4-3) (Wang & Chen, 2014) is implemented in this study. The EQCDF technique removes systematic biases in CMIP6 simulations and accounts for biases in all statistical moments. However, as with all statistical downscaling approaches, it is assumed that biases relative to historical observations remain constant in the projection period. Precipitation values below 0.01 mm/day are set to 0 (dry-day) during the correction process.

$$M_{BC} = M_{RAW} * \frac{F_O^{-1}[F_{RAW}(M_{RAW})]}{F_{REF}^{-1}[F_{RAW}(M_{RAW})]},\ M_{BC\_adj} = M_{BC} * \frac{\overline{M_O}}{\overline{M_{BC}}} \quad (4\text{-}3)$$

In this equation, F_{raw} represents the Cumulative Distribution Function (CDF) of the model for a future projection period, while F_O^{-1} and F_{REF}^{-1} denote the quantile functions for observations and the model in the reference period, respectively.

For both bias correction methods, ERA5 data is used as the observational dataset. The reference period is 1999–2006, and the output period is 2007–2014. The results of applying Eq. 4-2 and Eq. 4-3 are presented in Table 4-2, Table 4-3, Fig. 4-9, and Fig. 4-10.

Table 4-2 Correlation coefficient between biased CMIP6 data and ERA5 data.

	Temperature	Solar radiation	Specific humidity	Wind speed	Precipitation
AWI	0.980	0.999	0.936	0.869	0.32
MPI	0.982	0.999	0.934	0.868	0.31
MIROC	0.978	0.999	0.917	0.861	0.19

Table 4-3 RMSE between biased CMIP6 data and ERA5 data.

	Temperature	Solar radiation	Specific humidity	Wind speed	Precipitation
AWI	3.05	2.20	0.00075	0.34	0.31
MPI	2.96	2.16	0.00076	0.34	0.33
MIROC	3.21	2.42	0.00086	0.35	0.34

For temperature, the correlation coefficients between the bias-corrected outputs of each climate model and the observed data exceed 0.97, with lower dispersion in summer (standard deviation between 0 and 2) and a root mean square error (RMSE) of approximately 3 K. For solar radiation, the correlation coefficient is close to 1, with a standard deviation ranging between 0.5 and 3, and an RMSE of about 2 W/m². For specific humidity, the correlation coefficient is approximately 0.93, and the bias-corrected model data are slightly lower than the ERA5 data. Dispersion is higher than that of ERA5 in October and November, with an RMSE of 0.0008 kg/kg. For wind speed, the correlation coefficient is 0.86, and the RMSE is 0.34 m/s. The dispersion of AWI and MPI data is highest in November and May, respectively. For precipitation, the correlation coefficients between the bias-corrected outputs of each climate model and the observed data are relatively low, and the standard deviations also vary due to the stochastic nature of precipitation events, with an RMSE of 0.31 mm/day. From an annual cycle perspective, the error between the bias-corrected data and ERA5 data is significantly reduced.

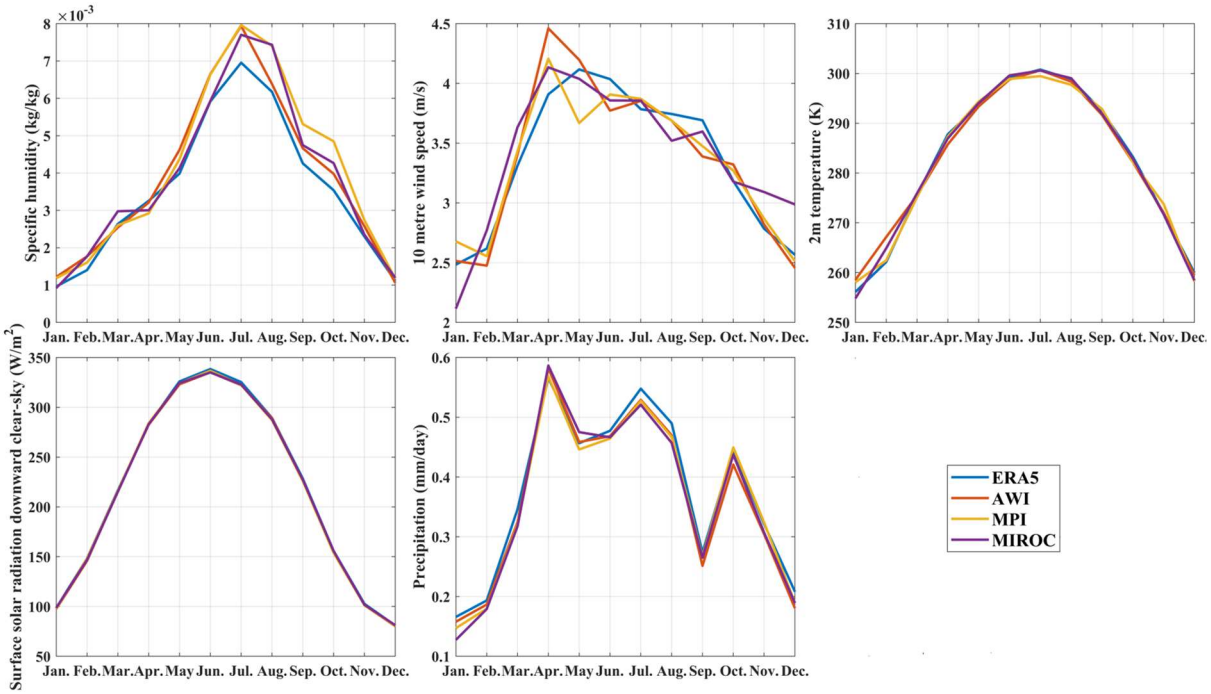


Fig. 4-9 Annual cycle of climate variables after bias correction.

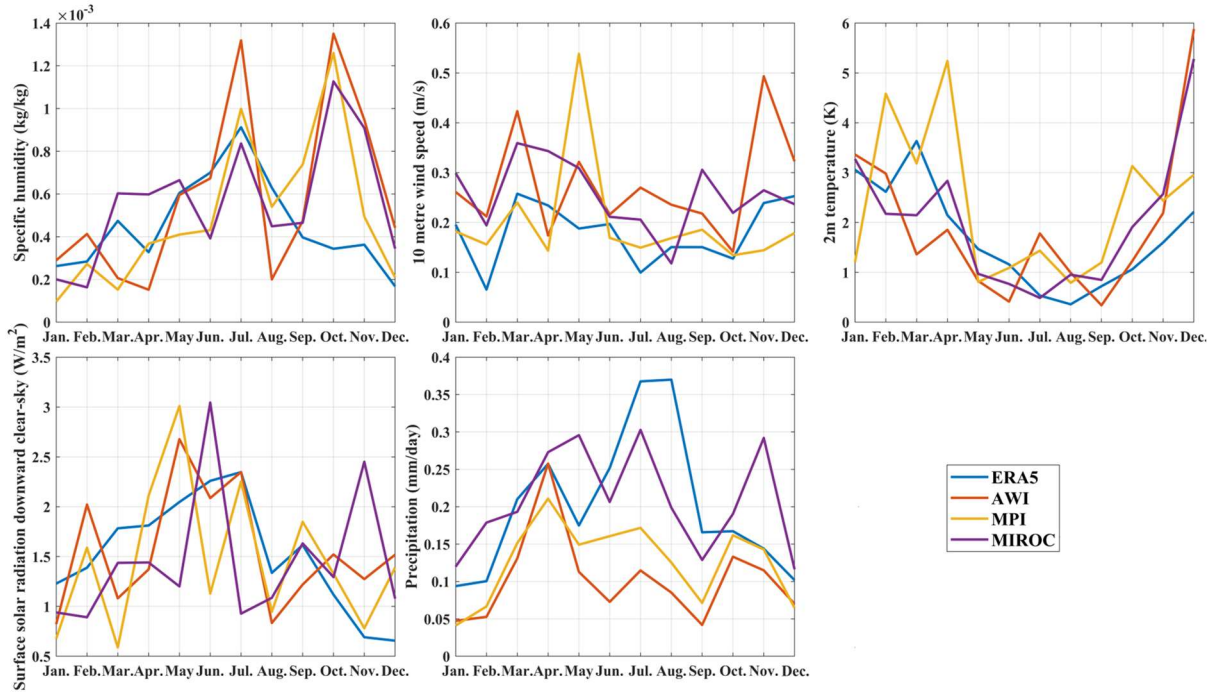


Fig. 4-10 Standard deviation of climate variables after bias correction.

The results above indicate that the bias-corrected climate model exhibits strong simulation capabilities for temperature, solar radiation, relative humidity, wind speed, and precipitation in the Gurbantunggut Desert. Furthermore, the corresponding model

performance evaluation confirms that, for each climate variable, the bias-corrected CMIP6 data outperform the uncorrected data (Xu et al., 2012). Therefore, the corrected dataset enhances the reliability of future climate change projections in the Gurbantunggut Desert to a certain extent.

4.2.3 Climate variables selection

Two approaches are employed to select the model variables influencing biological soil crust (BSC) variation in the study area. Climatic variables and their time lags are considered in both approaches.

The first approach is static selection, in which climatic factors are incorporated using adjusted time series with the optimal time-lag coefficient, as identified by Eq. 4-1. For each grid point, the climate variables used in the models remain fixed.

The second approach is best subset selection. Starting with the null model, different combinations of features are fitted, and one of the best models is selected from each subset—ranging from a single-feature model to an n-feature model. The optimal model is then chosen from a total of n+1 models based on prediction accuracy. The features configured in this optimal model are the final selected variables.

Three methods are used in this study to evaluate prediction accuracy. They are:

1) Adjusted r-squared (Ezekiel, 1930):

$$\mathrm{R}^2_{\mathrm{adj}} = 1 - (1 - \mathrm{R}^2)\frac{(\mathrm{n} - 1)}{(\mathrm{n} - \mathrm{k} - 1)} \tag{4-4}$$

Where R^2 represents the coefficient of determination (regular R-squared) and was first introduced by Pearson (1896). n denotes the number of observations (sample size), and k refers to the number of independent variables (predictors)

in the model. This improves upon R^2 preventing it from increasing with each additional predictor. Using this metric, the best model is the one with the highest R^2_{adj}. Maximizing R^2_{adj} is equivalent to minimizing the standard error. R^2_{adj} is an effective method for selecting predictors, particularly when comparing models with different numbers of predictors.

2) Schwarz's Bayesian Information Criterion (BIC) (Schwarz, 1978):

$$BIC = Tlog\left(\frac{SSE}{T}\right) + (k+2)\log(T) \tag{4-5}$$

Minimizing the Bayesian Information Criterion (BIC) is intended to identify the optimal model. The model selected by BIC is either the same as that chosen by the Akaike Information Criterion (AIC) (Akaike, 1973, 1974) or a more parsimonious one. This occurs because BIC imposes a stronger penalty on the number of parameters compared to AIC. For large values of T, minimizing BIC closely approximates leave-v-out cross-validation, where $v = T[1 - \frac{1}{\log(T)-1}]$.

3) Mallows's C_p (Mallows, 1973):

$$C_p = \frac{SSE_p}{S^2} - (N - 2p) \tag{4-6}$$

Where $SSE_p = \sum_{i=1}^{N} (Y_i - \hat{Y}_{pi})^2$ is the error sum of squares for the model with p predictors (including the intercept), Y_{pi} is the predicted value of the ith observation of Y from the p predictors, S^2 is the estimation of residuals variance after regression on the complete set of K predictors and can be estimated by $\frac{1}{N-K} = \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2$, and N is the sample size.

After the initial reduction of multicollinearity among climate variables based on correlation analysis, the appropriate models are selected using the three measures mentioned above. For different months at each grid point, the model incorporates varying climate factors with corresponding time lags.

4.3 BSC-climate model based on fixed variables

To quantify the contribution of climate change to BSC variation, a multiple linear regression model is constructed to analyze the relationship between BSC and five climatic factors (Eq. 4-7):

$$BSC_{var} = \sum_{var=}^{5} \beta_{var} \times Climate_{var} + \varepsilon \quad (4\text{-}7)$$

where β represents regression coefficients and ε denotes the error term. The variable *var* ranges from 0 to 5, representing different climatic factors. These climatic factors incorporate the time-lag effects identified by Eq. 4-1.

Based on the identified time-lag effects that best predict BSC responses to climate, a multiple linear regression model is established to examine the relationship between climatic factors and BSC variation. Figure 4-11 presents the determination coefficients of the multiple linear regression models for the entire season. In terms of spatial distribution (Fig. 4-11a), the model provides a better explanation for the southern and western regions of the study area, where the determination coefficient exceeds 80%. In contrast, in the central and northern regions, the model's explanatory power is weaker, with determination coefficients around 20%. On average (Fig. 4-11b), climatic factors account for 40% of the variation in BSC when considering both time-lag and accumulation effects.

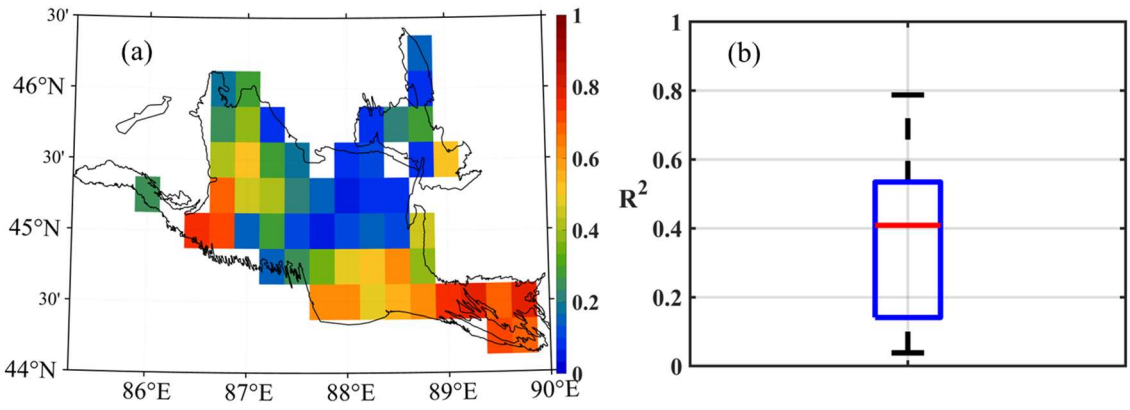


Fig. 4-11 Distribution (a) and box plot (b) of r-squared. The red line in the box plot indicates the average value of the data.

4.4 BSC-climate model based on machine learning

4.4.1 Statistical evaluation and model selection

Considering the quality and size of the dataset, three machine learning methods are selected for model development: multiple linear regression (MLR) (Draper & Smith, 1998), support vector regression (SVR) with a linear kernel (Drucker et al., 1997), and support vector regression with a radial basis function (RBF) kernel (Smola & Schölkopf, 2004). During the machine learning process, 80% of the data is allocated as the training dataset, which also serves as the validation dataset due to the use of k-fold cross-validation. The remaining 20% of the data is set aside as the testing dataset to evaluate the model's performance. To identify the best-performing model among these three approaches, three predictive accuracy metrics and two interpretability measures are applied for model selection (Table 4-4).

Table 4-4 Measures for model selection

This table presents the accuracy and interpretability measures used for model evaluation. The accuracy metrics include Root Mean Squared Error ($RMSE$), Standard Error (SE), and Correlation Coefficient (R), while the interpretability measures include Determination Coefficient (R^2) and Adjusted R-squared (R^2_{adj}).

Accuracy	Root Mean Squared Error	$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Predictied_i - Actual_i)^2}{N}}$
	Standard Error	$SE = \frac{\text{RSS}}{sqrt(n)},\ RSS = \sum(Actual_i - Predict_i)^2$
	Correlation Coefficient	$R = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$
Interpretability	Determination Coefficient	$R^2 = 1 - \frac{\sum(Actual_i - Predict_i)}{\sum(Actual_i - Actural\_mean)}$
	Adjusted R-squared	$R^2_{adj} = 1 - (1 - R^2)\frac{(n-1)}{(n-k-1)}$

Note. *N = total number of observations. $Actual_i$, $Predict_i$ = predicted and actual values for the i-th observation. RSS (Residual Sum of Squares) = $\sum(Actual_i - Predict_i)^2$. n = total number of observations. X_i, Y_i = predicted and actual values for the i-th observation. $\bar{X}$, $\bar{Y}$ = mean of the predicted and actual values. $Actural\_mean$ = mean of actual observed values. k = number of predictors, including the intercept.*

By calculating the values of the above five measures on both monthly and yearly scales, it is found that Multiple linear regression is the best performing model in relative terms. The performance and selection of the three modelling approaches for each indicator is shown in the Table 4-5 (See the Section 7.4 for more information.). Sometimes more than one approach performs well under a measure, for example, multiple linear regression and SVR with RBF kernel are close in performance on the standard error and adjusted r^2. Therefore, the multiple regression approach is chosen to model the relationship between the BSC and climate variables. it's worth noting that the multiple

regression model trained based on machine learning has different climate variables at each grid point, which is different from the fixed climate variable multiple regression model in Section 4.3.

Table 4-5 Performance of three machine learning models in the study area

	Multiple linear regression	SVR with linear kernel	SVR with RBF kernel
RMSE	0.058	0.078	0.085
SE	0.020	0.020	0.024
Correlation Coefficient	0.334	0.312	0.162
Determination Coefficient	0.704	0.597	0.666
Adjusted R-squared	0.500	0.543	0.566

***Units**: RMSE and SE have the same unit as the BSC fraction (%). Correlation Coefficient, Determination Coefficient (R^2), and Adjusted R-squared ($\boldsymbol{R^2_{adj}}$) are dimensionless.*

4.4.2 Model validation

Fig. 4-12 illustrates the determination coefficient between observed and predicted values. The model exhibits better interpretability from spring to autumn compared to winter, explaining over 75% of the variation in BSC. Thus, in terms of interpretability, this model outperforms the one based on the selection of fixed climate variables presented in Section 4.3.

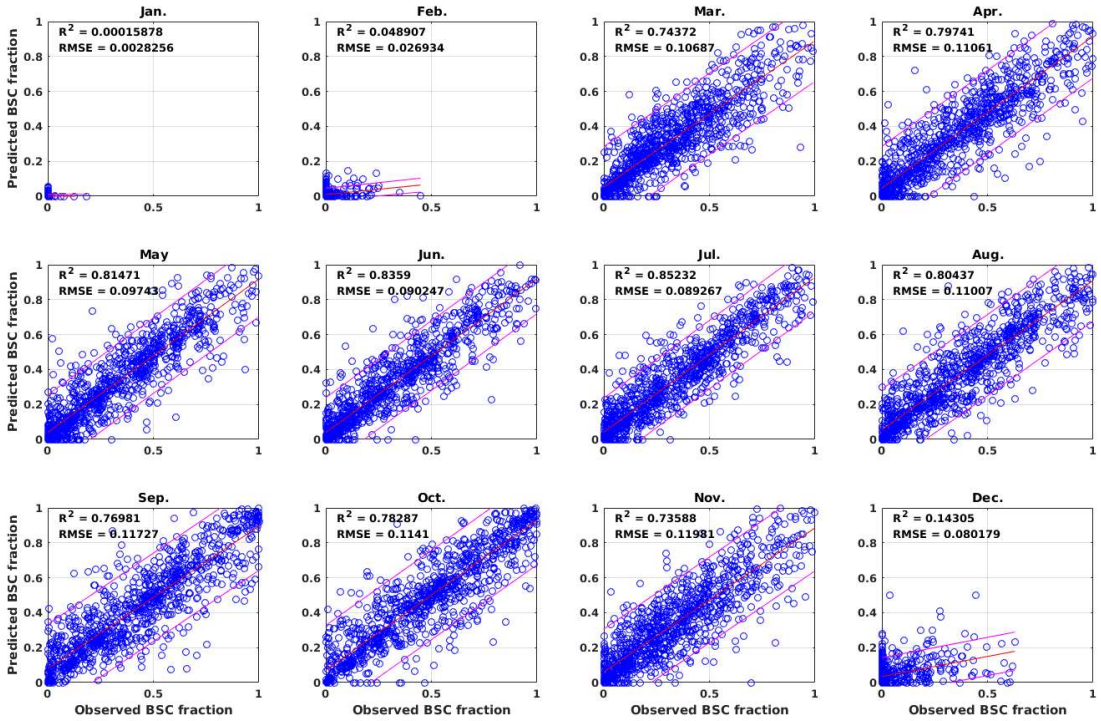


Fig. 4-12 Determination coefficient between observation and prediction.

4.5 Model reliability on the CMIP6 dataset

As the models are intended for application to the LGM and Mid-Holocene, ensuring the reliability of the CMIP6 paleoclimate dataset is crucial for subsequent model applications. This section evaluates the reliability of the corrected CMIP6 data in both the fixed-variable selection model and the machine learning model.

4.5.1 In the fixed-variable model

The prediction results of the CMIP6 data using the fixed-variable model are presented in Fig. 4-13 and Fig. 4-14. The predictions closely align with satellite observations on an annual cycle. While the standard deviation is lower than that of the satellite data, the linear variation trends remain consistent, with peaks occurring in March and November. However, the MIROC data exhibit an additional peak in September. Regarding the seasonal spatial distribution, the predicted values from the model generally match those from satellite observations. Despite the overall similarity, the CMIP6 data differ in that they predict lower values in the south-central part of the desert during spring and

summer. In winter, BSC is essentially unobservable due to vegetation dormancy.

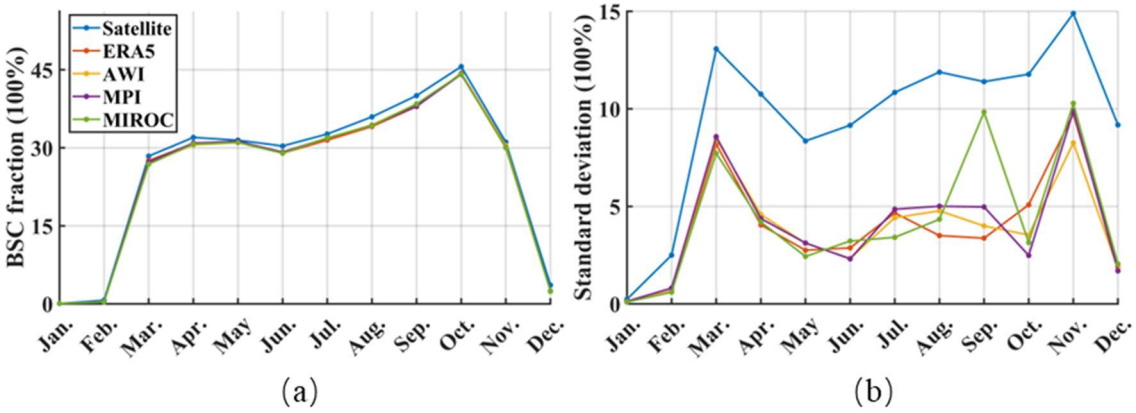


Fig. 4-13 Annual cycle (a) and standard deviation (b) of the prediction results of the regression model and compared with satellite data.

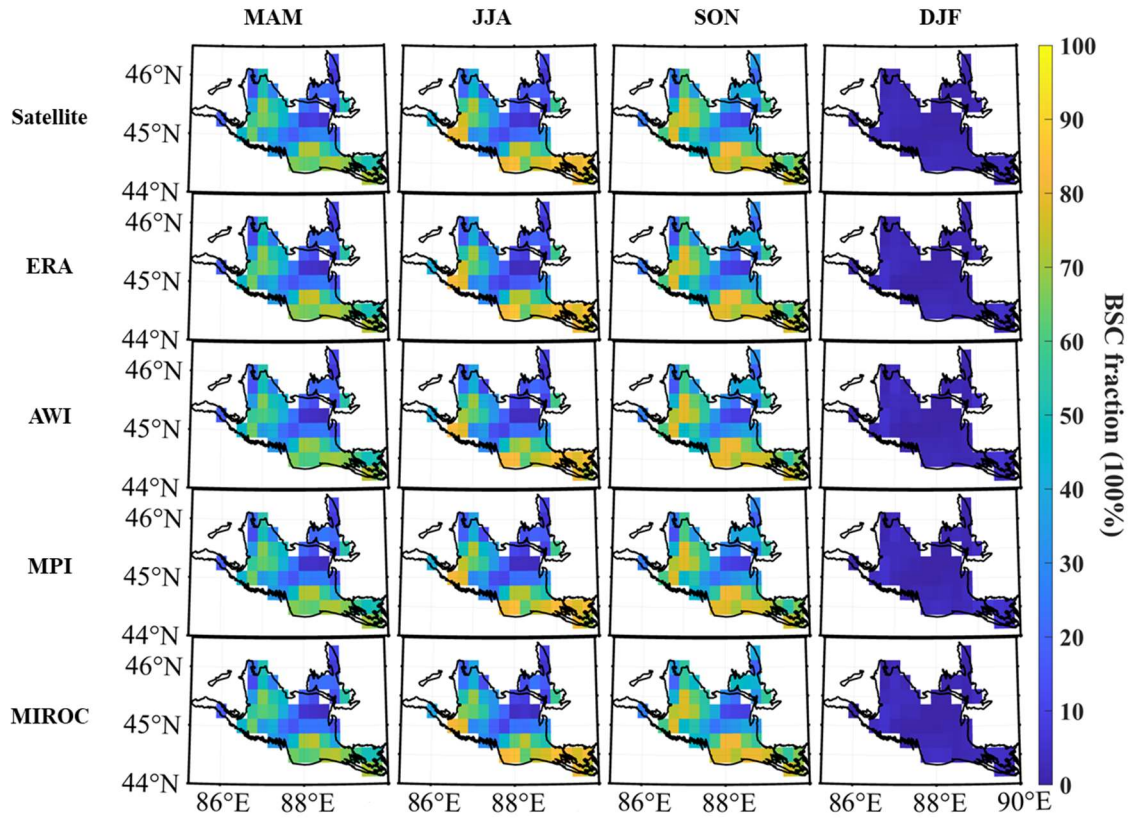


Fig. 4-14 Seasonal distribution of BSC fraction in the Gurbantunggut Desert in historical.

4.5.2 In the machine learning model

Fig. 4-15 presents the annual cycle and standard deviation of the machine learning

model's prediction results, compared with satellite data. Consistent with the satellite observations, the BSC fraction begins to increase significantly in March and peaks in October during the annual cycle. The standard deviation of the model predictions is generally lower than that of the observations, with the largest difference occurring in September.

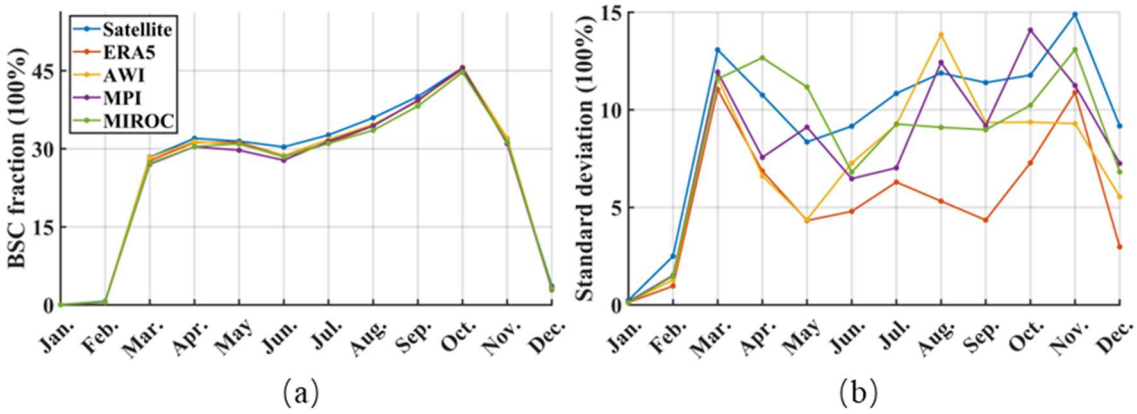


Fig. 4-15 Annual cycle (a) and standard deviation (b) of the prediction results of the regression model and compared with satellite data.

The seasonal distribution in Fig. 4-16 aligns with the annual cycle in Fig. 4-15. The BSC fraction values are generally lower than those observed in the satellite data, particularly in the southwestern region of the desert and along its southern edge.

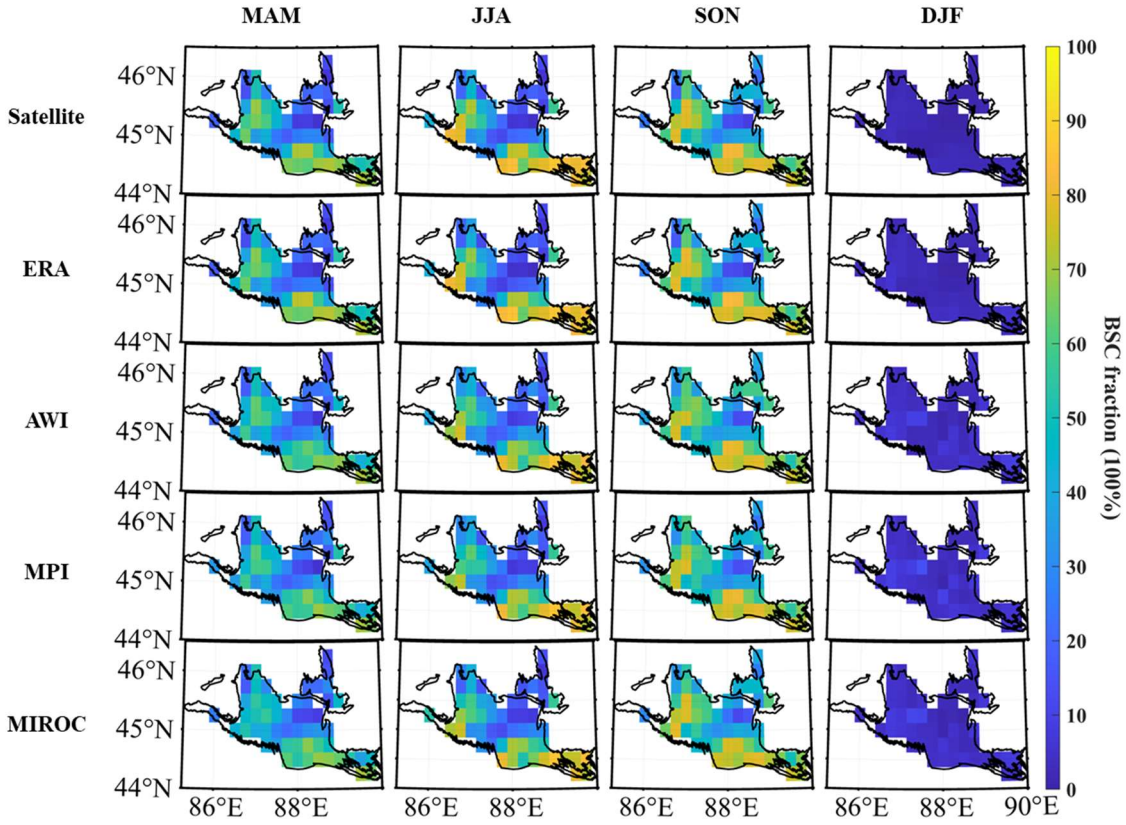


Fig. 4-16 Seasonal distribution of BSC fraction in the Gurbantunggut Desert in historical.

In summary, the outputs of both the fixed-variable model and the machine learning model, trained using ERA5 data as input, closely match the satellite-observed BSC data when the corrected and normalized CMIP6 data are used as input for prediction. This suggests that the regression model trained with ERA5 data can be reliably applied to the corrected AWI, MPI, and MIROC model data.

4.6 Interpretation the effects of climatic variables on BSC

For the linear trend analysis, changes in BSC, specific humidity, wind speed, temperature, solar radiation, and precipitation are first estimated through linear fitting, incorporating the identified optimal time lags ($p < 0.05$). Subsequently, regions exhibiting significant BSC changes are compared with those experiencing significant climate changes, allowing for the identification of dominant climatic factors in regions

where BSC variations are most pronounced.

Based on the spatial patterns of linear trends, regions with significant changes in both BSC and climate variables accounted for 70% of the areas with significant BSC changes from 2000 to 2018 (Fig. 4-17). Panel (a) of Fig. 4-17 illustrates the climatic factors most strongly correlated with significant BSC changes, with temperature emerging as the dominant driver. In contrast, only a few areas along the northwestern edge are primarily influenced by precipitation. Panel (b) of Fig. 4-17 depicts the second most influential climatic factor when BSC changes significantly. A comparison with panel (a) reveals that in the western region, specific humidity serves as the primary driver, while precipitation acts as the secondary driver. Additionally, in the northern half of the desert, the roles of temperature and precipitation as driving factors have been reversed. These findings indicate that temperature, specific humidity, and precipitation are key drivers of major BSC changes.

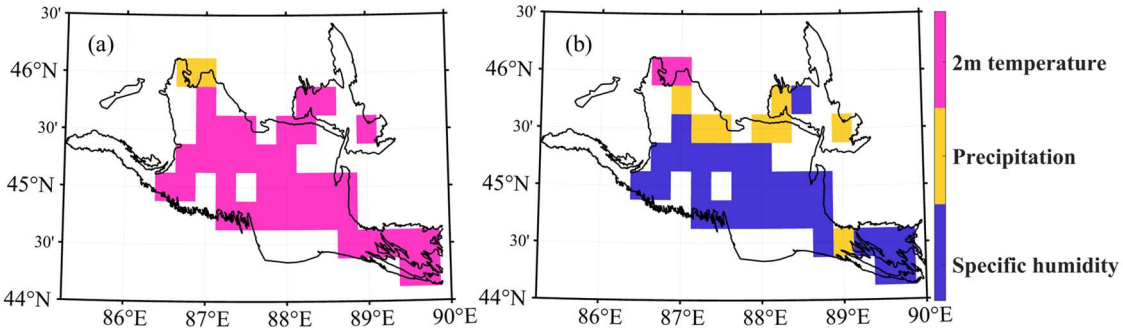


Fig. 4-17 Areas with significant changes in both BSC and climatic factors from 2000 to 2018. (a) Primary climatic drivers; (b) Secondary climatic drivers.

Fig. 4-18 presents the partial correlation analysis of BSC and climate factors with time lag, based on the model described in Section 4.3. The partial correlation coefficients between BSC and time-lagged climate factors over 19 years are calculated at the grid scale with a monthly time unit, and a significance test is performed.

The response of BSC to specific humidity exhibits spatial variability, with both positive

and negative correlations. The partial correlation coefficients between BSC and specific humidity, while controlling for other conditions, range from -0.2 to 0.3, with an average partial correlation coefficient of 0.16. Notably, in the northwestern desert, two adjacent grids display high correlations, yet one is positive and the other negative. The spatial distribution of partial correlation coefficients for temperature and solar radiation in these grids resembles that of specific humidity. In particular, in grids where solar radiation is positively correlated with BSC, specific humidity and temperature tend to show negative correlations. When analyzed alongside elevation maps (Fig. 3-3a), these grids coincide with areas exhibiting large changes in elevation slope (e.g., the tops and bottoms of dunes or the edges of mountain ranges), where shadowing and shading may promote water evaporation and retention (D. Zhang et al., 2024).

Wind speed demonstrates an approximately equal split between positive and negative correlations with BSC across most parts of the desert. A negative coefficient suggests that higher wind speeds cause greater damage to the surface crust. Meanwhile, temperature exhibits a positive correlation with BSC in most areas, except for the desert hinterland, where a negative correlation is observed. This may be attributed to the hot and arid conditions in the hinterland, where BSC growth is stressed due to increased potential evapotranspiration driven by enhanced solar radiation (Belnap et al., 2001).

The correlation between precipitation and BSC in the study area is generally low, fluctuating around zero, with a negative correlation observed along the southern edge of the desert. This suggests that higher temperatures may lead to increased surface evaporation and greater water depletion, which is not sufficiently compensated by precipitation. Conversely, in grids exhibiting a positive correlation, BSC growth is limited by decreasing precipitation and promoted by increasing precipitation. The low overall correlation may also be explained by the wetting effects surrounding BSC on rainy days and the low vegetation cover, which buffer BSC variation from climatic fluctuations (Belnap et al., 2001).

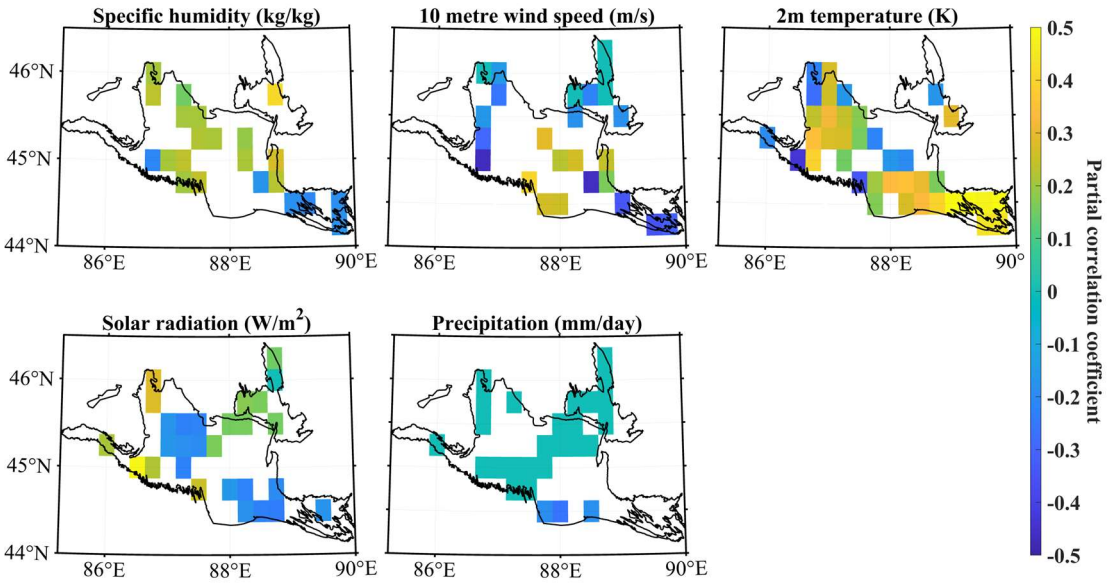


Fig. 4-18 Partial correlation analysis of BSC and climate factors with time lag.

To evaluate the effect of climatic factors on BSC variability using the fixed-variable model, the climate factors with the largest regression coefficients and the signs of these coefficients are analyzed. The results are presented in Fig. 4-19. The color bar labels indicate climate factors with time lags: *q* represents specific humidity, *si10* denotes 10-meter wind speed, *t2m* corresponds to 2-meter temperature, *ssrdc* refers to surface solar radiation downward (clear-sky conditions), and *tp* indicates total precipitation. Subscript numbers denote time lags, where zero indicates no lag. Across the desert region, temperature without a time lag emerges as the dominant climatic factor influencing BSC variability, exhibiting a consistently positive effect, which aligns with the findings in Section 4.1.1. In contrast, for temperature with a one-month lag, the regression coefficient is negative in the central desert hinterland, suggesting that prolonged high temperatures continue to suppress BSC growth. This finding reinforces the idea that temperature acts as a major limiting factor for BSC development. The second most influential factor is solar radiation, which primarily affects the sandy regions in the northeastern desert. The third key factor is specific humidity, which predominantly influences the grasslands in the central desert and some bare soil areas in the northeast.

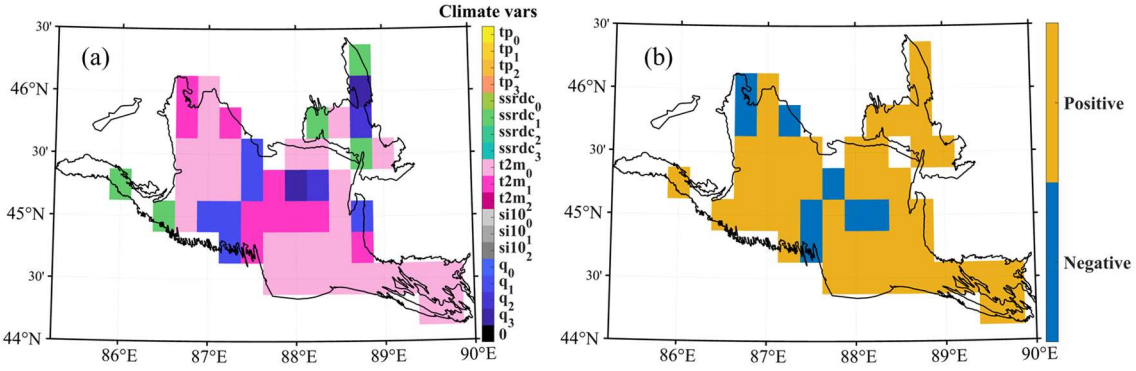


Fig. 4-19 (a) The most important variables in the fixed-variable model;
(b) the positivity or negativity of their contribution.

A similar approach as described above is used to identify the climate factors with the largest (Fig. 4-20a) and second-largest (Fig. 4-20b) regression coefficients in the machine learning regression model. The longer the color bar, the greater the spatial influence of that climate factor across the desert. As winter temperatures in the Gurbantunggut Desert generally fall below freezing, making BSC growth nearly impossible, the analysis focuses on the regression results from March to November.

For March and April, solar radiation is the dominant climate variable, contributing the largest share of regression coefficients. This is because early spring remains cold and requires the continuous influence of solar radiation. This also explains the observed time lag in solar radiation's dominant effect during these months. Additionally, sufficient solar radiation facilitates BSC photosynthesis. As for the second-largest climatic factor, specific humidity has the highest share, suggesting that early spring is not yet overly dry, and the moisture in the air promotes BSC growth. Overall, solar radiation and specific humidity are the key climatic factors influencing BSC variation in March and April.

A similar pattern is observed in May, October, and November, where no single climatic factor is overwhelmingly dominant. In May, solar radiation, specific humidity, and temperature all have high regression coefficients, and their spatial distributions appear

more sporadic. In October and November, solar radiation, temperature, specific humidity, and wind speed are unevenly distributed, indicating that the desert's climatic conditions are more complex during these months. Additionally, BSC peaks in October, which results from the combined effects of multiple climatic variables.

From June to September, temperature plays a significantly larger role, particularly in July, when almost the entire desert's BSC variation is influenced by temperature. Secondary variables are primarily specific humidity, with some localized areas also affected by wind speed. This pattern highlights the critical role of temperature in BSC variation from late spring to early autumn.

Additionally, the influence of precipitation and wind speed remains spatially limited throughout these months. This is because the Gurbantunggut Desert experiences low and highly sporadic precipitation, while wind speeds rarely exceed 6 m/s, which is insufficient to reach the friction velocity required to disturb BSC surfaces.

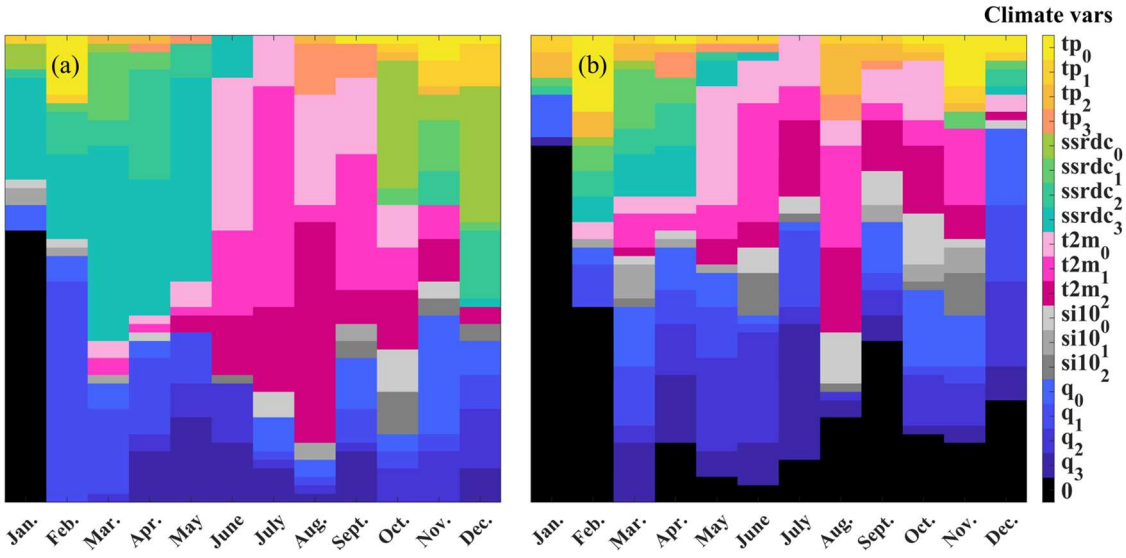


Fig. 4-20 The climate variables with the largest (a) and second largest (b) effects on machine learning regression models in different grid points in different months.

The spatial distribution of climate variables exerting the largest and second-largest influences in the machine learning regression model is presented in Fig. 4-21. In

February, the dominant climatic factor in the bare sand regions of the western and northeastern deserts is specific humidity with no lag. In contrast, in the grassland areas of the central and southwestern desert, precipitation with no time lag and solar radiation with a three-month lag serve as the primary drivers. From March to May, solar radiation with a two- to three-month lag emerges as the primary driver in the southern and western desert, while temperature and solar radiation with shorter time lags act as secondary drivers. In other parts of the desert, the climatic influences are more complex. Between June and September, temperature dominates as both the primary and secondary driver across most of the desert, though with varying time lags. The effects of precipitation and specific humidity are scattered and spatially variable. From October to December, the distribution of climatic influences becomes increasingly heterogeneous, though broad patterns can be identified: solar radiation and temperature in October, specific humidity and temperature in November, and solar radiation and specific humidity in December.

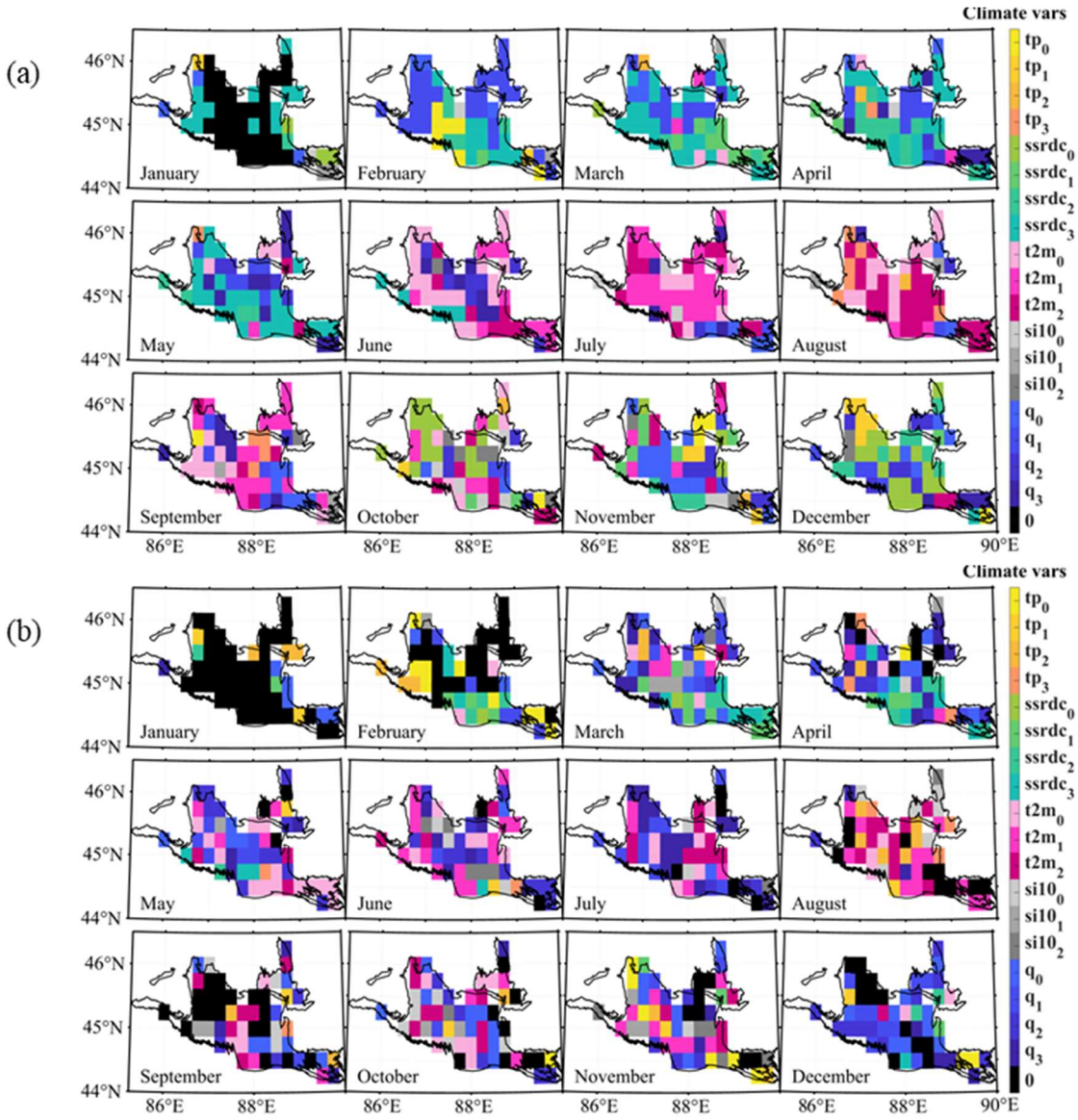


Fig. 4-21 The spatial distribution of climate variables with the largest (a) and second largest (b) effects on machine learning regression model.

4.6 Conclusion

This chapter first analyzes the time-lag and partial correlations between BSC and climate variables. It then introduces data normalization and bias correction methods. The selection of variables for the model is conducted using adjusted R-squared, BIC, and Mallows' Cp. Finally, a fixed-variable model and a machine learning model are developed. The main climatic factors affecting BSC are examined based on the regression coefficients of the model, and the reliability of the model on other datasets is assessed. The main findings are as follows:

1) Correlation analysis reveals that in some areas, BSC exhibits a stronger correlation with time-lagged climate factors when the time lag is considered, with specific humidity being the most pronounced. The BSC response to specific humidity typically lags by one to two months, and in principle, the time lag between BSC and climatic variables does not exceed three months. BSC responds rapidly to temperature, with a correlation coefficient of 0.7. Similarly, BSC responds quickly to precipitation, though the correlation coefficient is relatively low (0.46), with significant correlations primarily observed in the east and south.

2) The corrected climate model demonstrates strong simulation capabilities for temperature, solar radiation, relative humidity, wind speed, and precipitation in the Gurbantunggut Desert. The calibrated CMIP6 model effectively simulates precipitation, temperature, and solar radiation, meeting the study's requirements. Following bias correction, the discrepancies between climate model outputs and observed data are significantly reduced, leading to a notable improvement in the simulation performance of each climate model.

3) BSC-climate models are constructed using two approaches, both incorporating time lag effects. The first approach is based on fixed climate factors, while the second approach slides over the time series to dynamically select more appropriate climate factors at different time points. Multiple regression analysis is applied to both models, and statistical parameters are used for evaluation. The results indicate that the two approaches explain approximately 40% and 75% of BSC variability, respectively.

4) Climatic factors account for 70% of BSC variation, indicating that climate change significantly affects most areas where BSC changes occur. The remaining unexplained BSC variation may be attributed to other factors, such

as atmospheric CO_2 levels and nitrogen deposition (Nemani et al., 2003; Zhu et al., 2016), human activities, and natural disturbances (Lehmann et al., 2014; Peng & Li, 2018).

5) Due to spatial and temporal variability, it is challenging to derive a consistent conclusion regarding the influence of climatic factors on BSC solely from model coefficients. This variability arises because the dominant climatic factors and their time lags differ across grid points and months. However, in general, the key climatic factors influencing BSC are relative humidity, temperature, and solar radiation.

5 Application of the BSC-Climate model

In the previous chapters, a fixed-variable regression model and a machine learning regression model were developed, and the reliability of the AWI, MPI, and MIROC scenario data in CMIP6 for these two models was validated. In this chapter, these two models are applied in two ways. The first application utilizes paleoclimate data from CMIP6 to investigate and compare the temporal and spatial distribution of BSC in the Gurbantunggut Desert during the Last Glacial Maximum (LGM, 21 ka) and the Mid-Holocene (6 ka). The second application extends the model to the Atacama Desert in Chile, where modern climate data from the Atacama Desert are used to predict the temporal and spatial distribution of BSC and compare the results with NDVI data.

5.1 Paleo Gurbantunggut Desert

This section first analyzes the spatiotemporal distribution of climate variables in the Gurbantunggut Desert during the Last Glacial Maximum (LGM) and the Mid-Holocene. Then, data from three CMIP6 climate scenarios are applied to the fixed-variable regression model and the machine learning regression model to derive the BSC distribution in the Gurbantunggut Desert under paleoclimatic conditions. Finally, the distribution and variations of BSC across different paleoclimatic periods are examined by comparing the differences in climatic factors between the two periods and integrating the discussion of key climatic drivers from the previous chapters.

5.1.1 Spatiotemporal variability of climatological processes

From the discussions in the previous chapters, it is evident that the main climatic factors influencing the distribution of BSC in the Gurbantunggut Desert are moisture, temperature, and solar radiation. Therefore, the spatial distribution of these three environmental variables is analyzed.

Fig. 5-1 presents the seasonal spatial distribution of specific humidity. In both

paleoclimatic periods, specific humidity gradually increases from northeast to southwest, which is inverse to the elevation gradient in the Gurbantunggut Desert (Fig. 3-3a), meaning that humidity levels are lower at higher elevations. Regarding seasonal variations, in both paleoclimate periods, temperature is highest in summer, similar in spring and autumn, and lowest in winter. Across different climate scenarios, the AWI scenario simulates the highest specific humidity, followed by MPI, while MIROC exhibits the lowest values. For both periods, specific humidity during the LGM is lower than that of the Mid-Holocene, particularly in summer, with a difference of approximately 0.001 kg/kg. This suggests that the Mid-Holocene was a wetter period, which aligns with the findings of Xu et al. (2023), based on research on the southern desert margin profile, and Liu et al. (2020), based on Mg/Ca analysis from Baluk Cave.

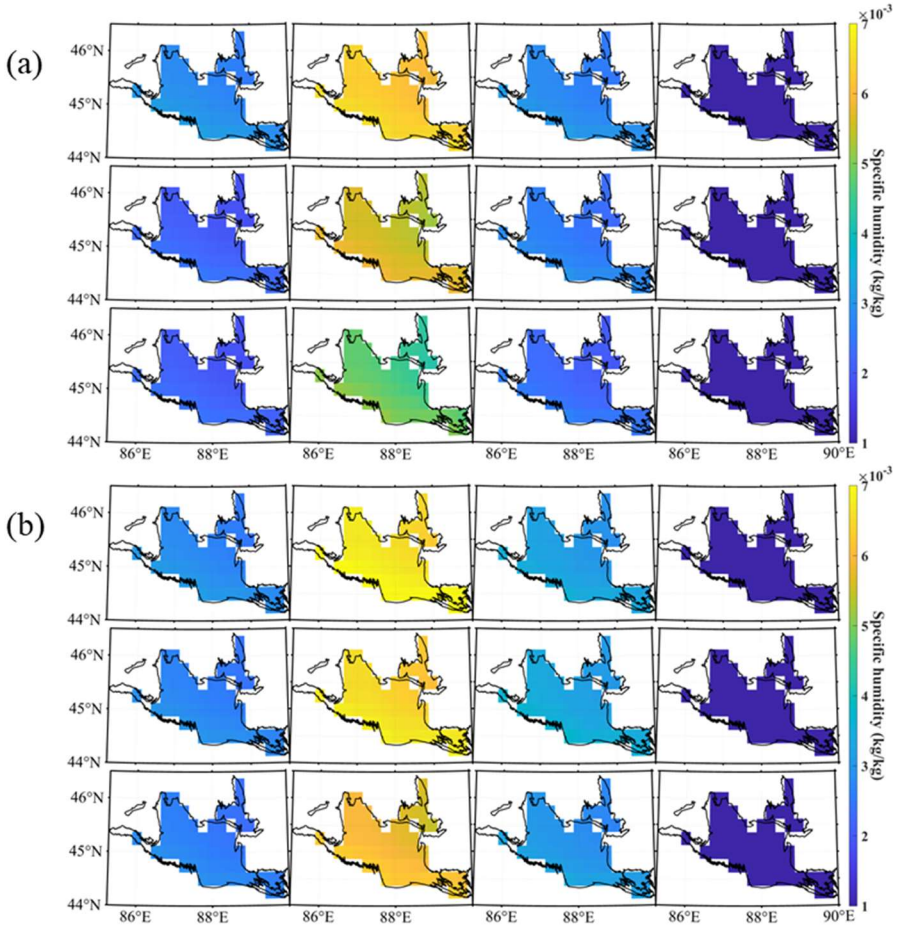


Fig. 5-1 Spatial distribution of specific humidity in Gurbantunggut Desert during LGM (a) and Mid-Holocene (b).

Fig. 5-2 presents the spatial distribution of temperature throughout the seasonal cycle. In both paleoclimatic periods, temperature gradually increases from northeast to southwest, which is inverse to the elevation gradient in the Gurbantunggut Desert (Fig. 3-3a), meaning that higher elevations correspond to lower temperatures. Across different seasons, temperature is highest in summer and lowest in winter in both periods. In the LGM, spring temperatures are higher than those in autumn, whereas in the Mid-Holocene, spring and autumn temperatures are similar. The simulated temperature values from different climate scenarios are generally consistent. In both paleoclimatic periods, summer temperatures during the LGM are significantly lower than those in the Mid-Holocene, with a difference of nearly 10 K. The summer temperature in the Mid-Holocene is approximately 290 K, aligning with the findings of Zhao et al. (2017).

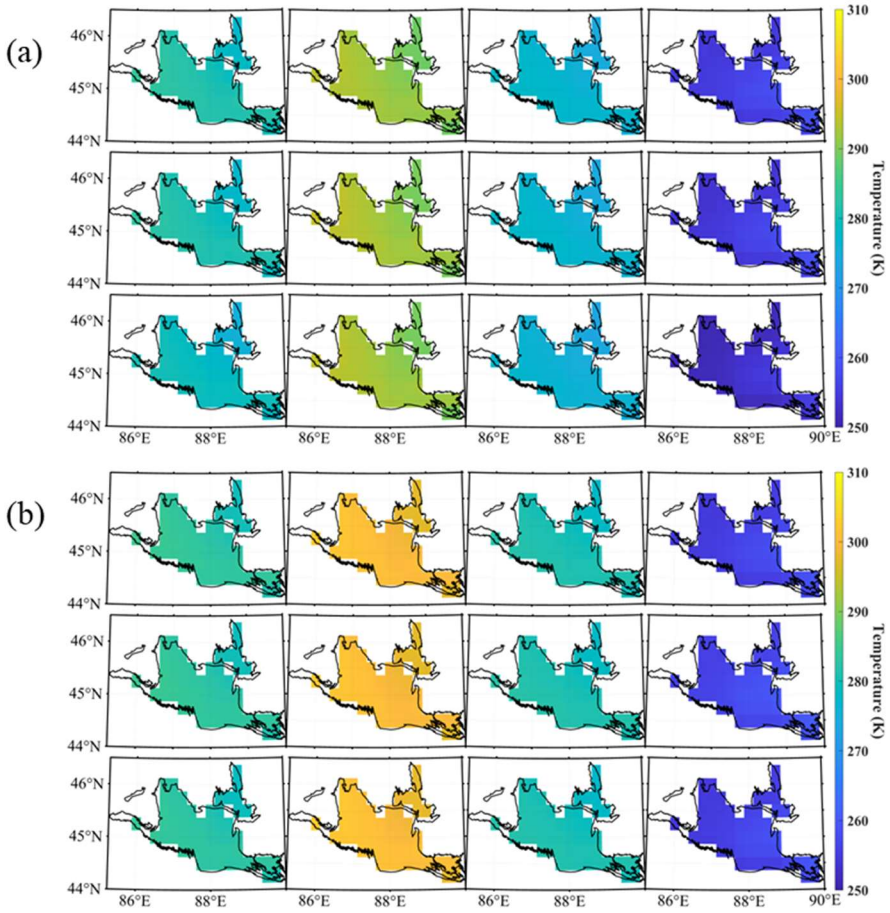


Fig. 5-2 Spatial distribution of temperature in Gurbantunggut Desert during LGM (a) and Mid-Holocene (b).

Fig. 5-3 presents the spatial distribution of solar radiation. In both paleoclimatic periods, the distribution of solar radiation across the desert is relatively uniform, with no significant spatial variations. Across different seasons, in both periods, solar radiation is highest in summer, followed by spring, then autumn, and lowest in winter. Regarding different climate scenarios, the AWI and MPI values are relatively similar, whereas MIROC exhibits higher values, particularly in summer. For both periods, solar radiation is slightly lower in the LGM than in the Mid-Holocene, with a more pronounced difference in summer. In the Mid-Holocene, summer solar radiation is approximately 350 W/m², which differs significantly from the findings of Xu et al. (2023), who reported that summer solar radiation at 45°N could reach 500 W/m². However, the winter solar radiation, at approximately 120 W/m², aligns with the results of Xu et al. (2023).

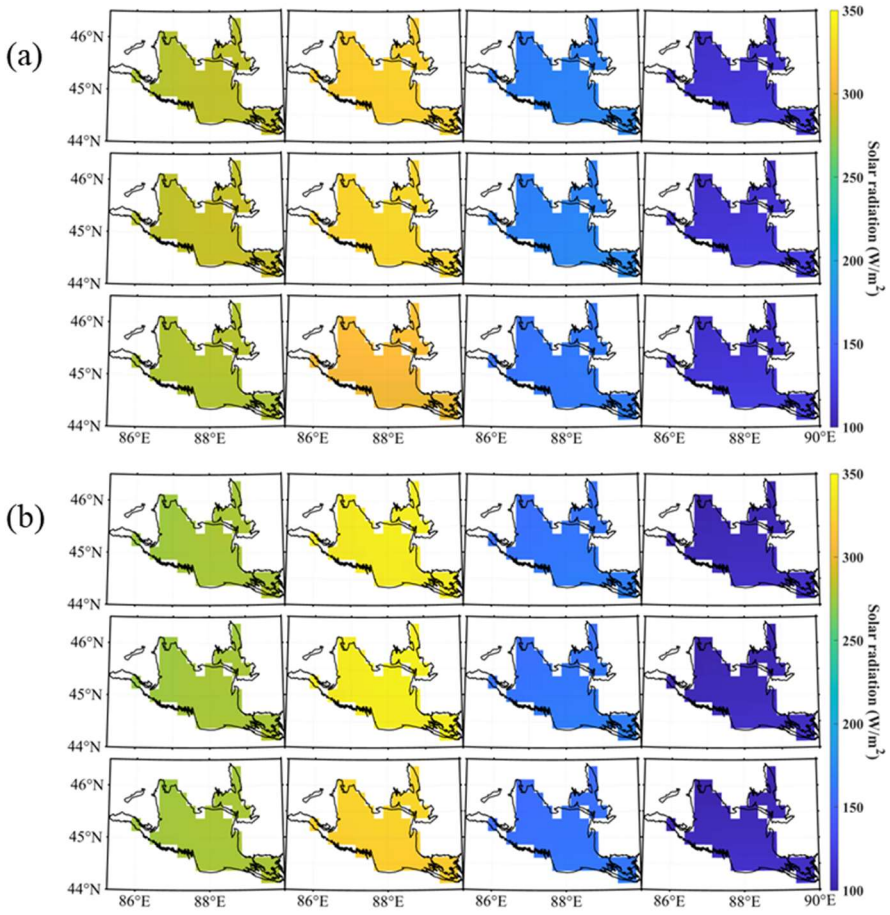


Fig. 5-3 Spatial distribution of solar radiation in Gurbantunggut Desert during LGM (a) and Mid-Holocene (b).

Fig. 5-4 presents the annual cycle of climate variables across three CMIP6 scenarios. Specific humidity follows a unimodal distribution in interannual variability, peaking in July, with the MIROC scenario exhibiting the lowest values during the same period. In the LGM, the annual cycle of the AWI scenario closely resembles that of the Mid-Holocene, whereas the MPI and MIROC scenarios display significantly lower values.

Wind speed follows a less pronounced bimodal distribution over the annual cycle, with peaks in May and September. A minor trough occurs in July, except for the MIROC values in the Mid-Holocene. The mean wind speed in the Mid-Holocene is approximately 1 m/s lower than in the LGM.

The annual cycles of temperature and solar radiation exhibit similar patterns. However, temperature is lower in the LGM than in the Mid-Holocene from May to October, whereas mean solar radiation levels remain comparable between the two periods.

The annual cycle of precipitation is highly irregular. Apart from the winter months, which have low rainfall, two troughs are observed: one in May and another in September, with notably low values in September. Precipitation peaks occur in April, July, and October, with the highest value recorded in April, reaching approximately 0.57 mm/day.

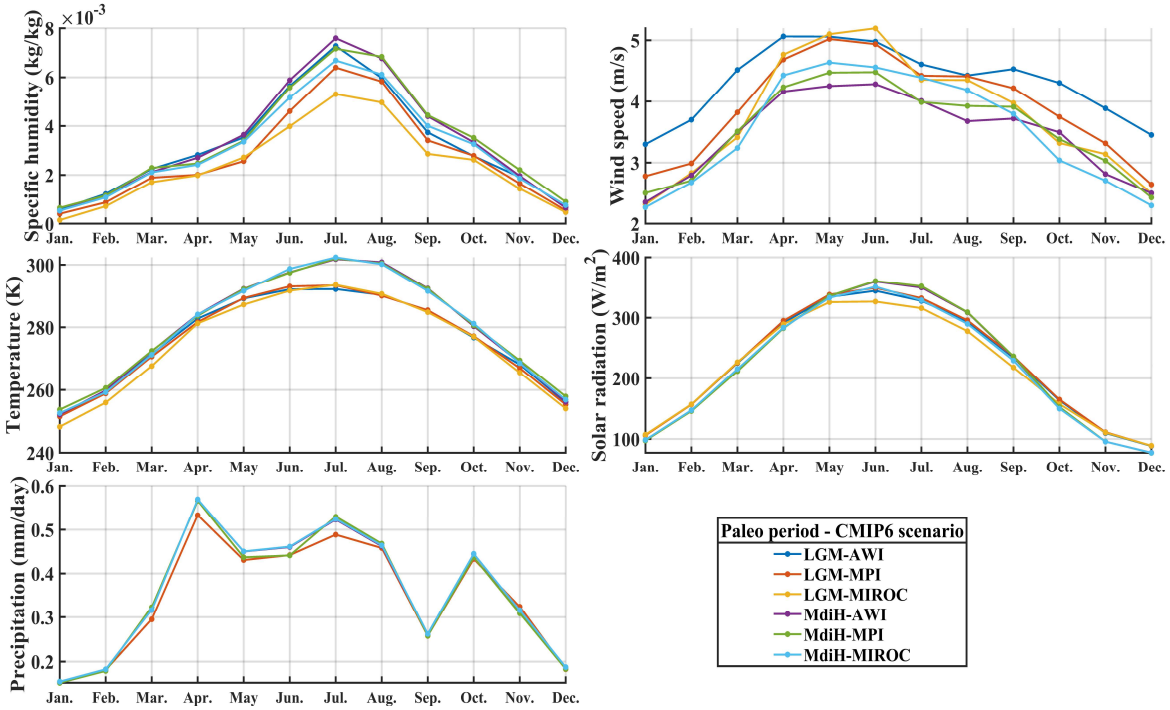


Fig. 5-4 Temporal distribution of climate variables in Gurbantunggut Desert in LGM and Hid-Holocene.

Since the Mid-Holocene, increased winter solar radiation has led to higher evaporation rates in upwind regions, such as the Mediterranean (Chen et al., 2008, 2019). The melting of ice in high-latitude areas of the Northern Hemisphere, the rising sea surface temperature in the North Atlantic, and the increased evaporation have provided sufficient water vapor for the Gurbantunggut Desert (Chen et al., 2016; Dyke, 2004; Praetorius et al., 2008; Jin et al., 2012). On the other hand, the weakening of summer solar radiation has caused a southward shift in subtropical high pressure, leading to reduced evaporation within the desert and an increase in atmospheric moisture content. Additionally, the solar radiation gradient between mid- and high latitudes has gradually increased, contributing to a corresponding intensification of the westerlies (Chen et al., 2016, 2019). Therefore, the combination of abundant water vapor sources and strong westerly winds has resulted in higher precipitation in the study area. This aligns with the findings of this study, which indicate that specific humidity was higher, and wind speed was lower in the Mid-Holocene compared to the LGM.

5.1.2 Fixed-variables model

Fig. 5-5 illustrates the application of the fixed-variable model in paleoclimatology. During the LGM, spring BSC is primarily distributed in the western and southeastern regions of the Gurbantunggut Desert, with fewer occurrences in the north. In summer, BSC is mainly concentrated along the southeastern edge of the desert. In autumn, BSC is more evenly distributed across the desert, with a relatively higher abundance in the west. In winter, BSC levels are very low, with only a few occurrences in the west and along the southeastern edge. Across different climate scenarios, the AWI and MPI predictions are highly similar, while the MIROC predictions are generally lower. However, MIROC exhibits higher values in localized areas, particularly along the southwestern edge and central east during summer, and along the southern edge during autumn.

During the Mid-Holocene, spring BSC is primarily distributed in the southeastern region of the Gurbantunggut Desert, with fewer occurrences in the north. In summer, BSC levels are generally low and are mainly concentrated in the southern part of the desert and along the southeastern edge. The autumn BSC distribution is similar to that of summer, but with higher values. In winter, BSC is mainly distributed in the southern and western parts of the desert. The spring and winter BSC predictions are relatively consistent across different climate scenarios. In summer, the MIROC prediction yields the highest values, followed by AWI, while MPI has the lowest values. In autumn, the AWI and MPI predictions are similar, whereas MIROC predicts significantly higher values in the southern region of the desert.

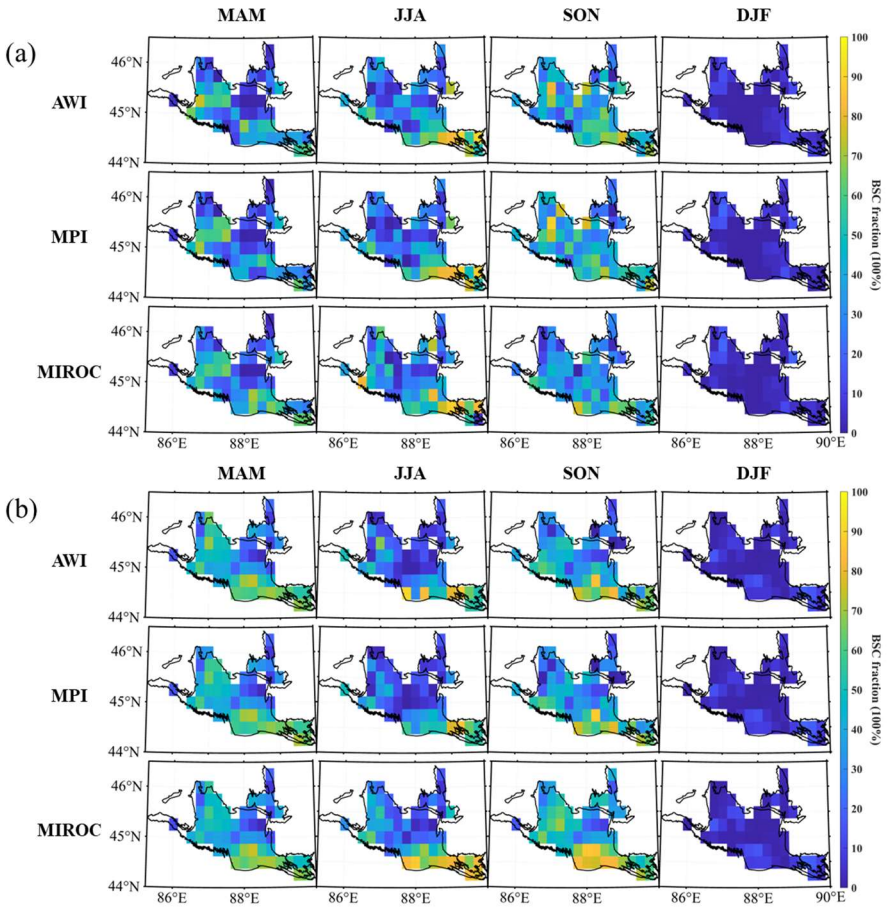


Fig. 5-5 Seasonal spatial distribution of BSC fraction predicted by the fixed-variables model in the LGM (a) and mid-Holocene (b).

Fig. 5-6a illustrates the difference in BSC between the LGM and the Mid-Holocene. In spring, the BSC fraction in the LGM is lower than in the Mid-Holocene for most grids, whereas the opposite trend is observed in summer. In autumn, the results vary across different climate scenarios. In the AWI and MPI scenarios, the BSC fraction is higher in the LGM for most parts of the desert, whereas the MIROC scenario shows the opposite trend. In winter, the difference in BSC is close to zero.

Fig. 5-6b presents the annual cycle of BSC across different scenarios and paleoclimatic periods. The BSC annual cycle exhibits bimodal characteristics, with peaks in April and October. Within the same climate scenario, the BSC fraction in the LGM is consistently lower than in the Mid-Holocene, aligning with the analysis of Fig. 5-5.

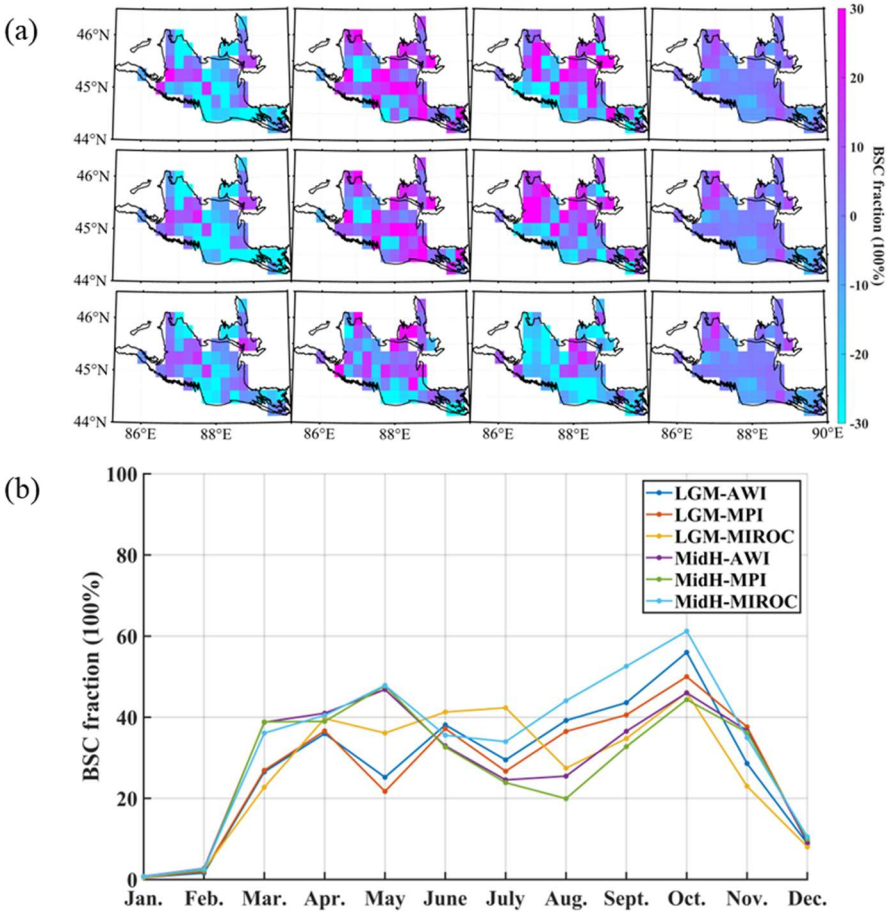


Fig. 5-6 Comparison of the distribution of BSC between the two paleo-periods: (a) Spatial distribution of the difference (LGM minors mid-Holocene) in BSC fraction; (b) Annual cycle of BSC fraction.

5.1.3 Machine learning model

Fig. 5-7 illustrates the application of the machine learning model in paleoclimatology. During the Last Glacial Maximum (LGM), spring BSC is primarily distributed in the southeastern region of the Gurbantunggut Desert, with fewer occurrences in the northern part. In summer, BSC is mainly concentrated along the southwestern and southeastern edges of the desert. In autumn, BSC is evenly distributed across the desert, with relatively higher concentrations in the southeastern region. In winter, BSC levels are very low, and the specific locations of BSC occurrence vary across different scenarios. Across different climate scenarios, the MPI and MIROC predictions are highly similar, whereas the AWI predictions yield lower values, particularly in the southern part of the desert during summer. MPI exhibits higher values in the

northwestern part of the desert in autumn, while AWI shows higher values in the west-central region in winter. MIROC predicts higher values in the eastern part of the desert and along the southeastern edge during winter.

During the Mid-Holocene, spring BSC is primarily distributed in the southeastern region of the Gurbantunggut Desert, with fewer occurrences in the north. In summer, BSC is mainly concentrated along the southern edge of the desert. In autumn, BSC is evenly distributed, with relatively higher concentrations in the central-southern and central-western regions. In winter, BSC is also present in small but relatively evenly distributed patches across the desert. The BSC distribution patterns remain similar across different climate scenarios.

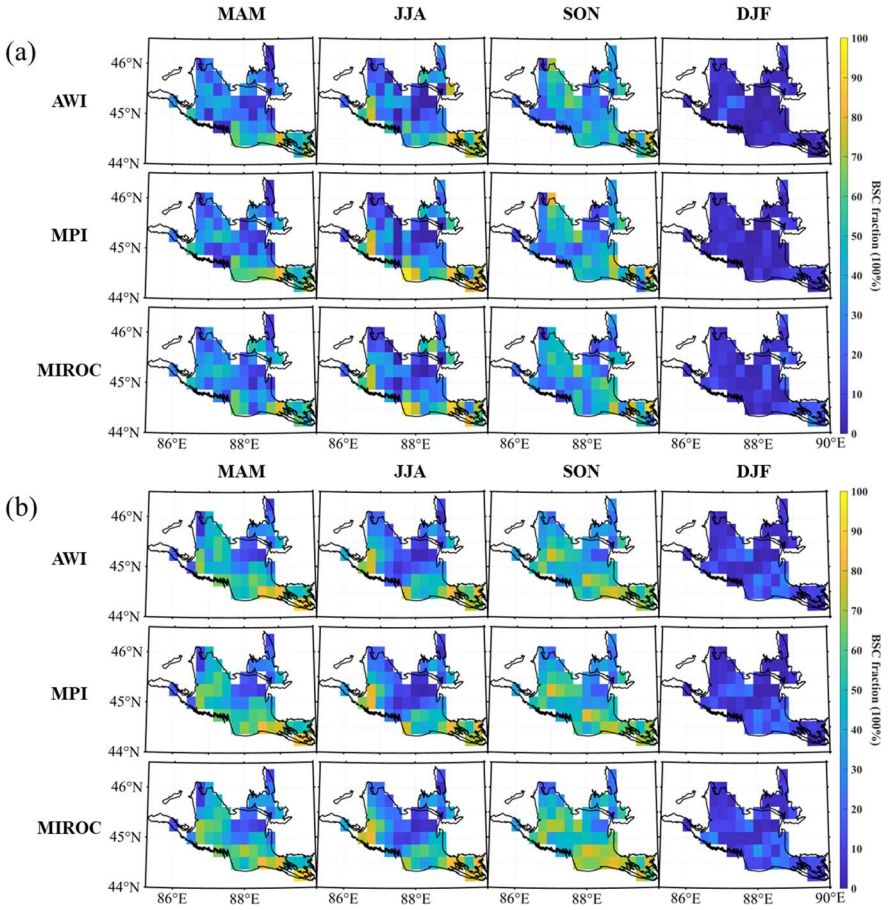


Fig. 5-7 Seasonal spatial distribution of BSC fraction predicted by the machine learning model in the LGM (a) and mid-Holocene (b).

Fig. 5-8a illustrates the difference in BSC between the LGM and the Mid-Holocene. In spring, the BSC fraction in the LGM is lower than in the Mid-Holocene across most grid cells. In summer, higher LGM values are primarily observed in the central and northern regions of the desert. In autumn, BSC values are higher in the southern desert during the LGM, whereas the northern desert exhibits higher values in the Mid-Holocene. In winter, the difference in BSC is close to zero.

Fig. 5-8b presents the annual cycle of BSC across different climate scenarios and paleoclimatic periods. The BSC annual cycle features a distinct peak in October. In May, the Mid-Holocene BSC fraction exhibits a small peak, whereas the LGM BSC fraction reaches a trough.

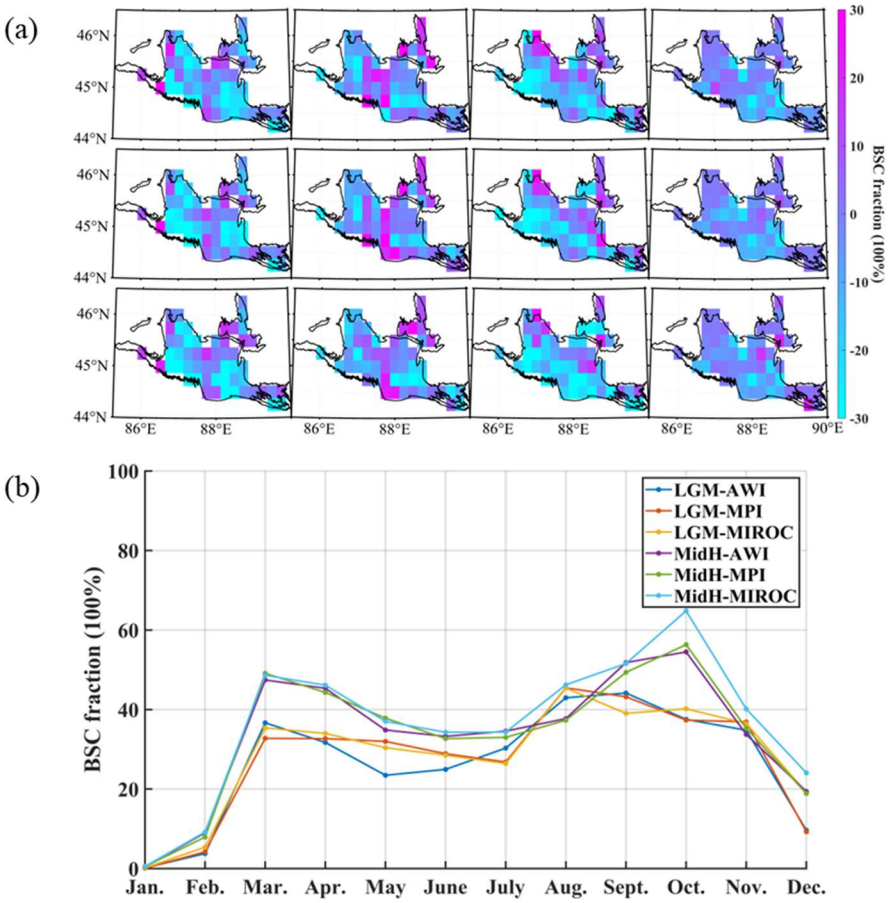


Fig. 5-8 Comparison of the distribution of BSC between the two paleo-periods: (a) Spatial distribution of the difference (LGM minors mid-Holocene) in BSC fraction; (b) Annual cycle of BSC fraction.

5.2 Atacama Desert

In this section, the generalisability of the model is tested by applying the fixed-variable regression model and the machine learning regression model to the Atacama Desert. The Atacama Desert is located in the Southern Hemisphere, where the seasons are reversed compared to the Northern Hemisphere. Therefore, in subsequent research, only the month is used to describe the spatiotemporal distribution.

5.2.1 Spatiotemporal variability of climate variables

Fig. 5-9 presents the temporal distribution of climate variables in the Atacama Desert, derived from the ERA5 dataset. The annual cycles of specific humidity, temperature, solar radiation, and evapotranspiration exhibit a consistent pattern, with lower values mid-year and higher values at the beginning and end of the year. Among these variables, specific humidity and temperature follow the same trend throughout the year, peaking in February and reaching their lowest levels in July. This pattern indicates that an increase (or decrease) in solar radiation and temperature leads to a corresponding increase (or decrease) in evapotranspiration, which subsequently results in a rise (or decline) in atmospheric water vapor and an associated change in specific humidity. Additionally, since water vapor is a greenhouse gas, its increase (or decrease) further amplifies temperature variations. Wind speeds remain low throughout the year, reaching their lowest levels in June and July at less than 1 m/s, and peaking in January and December, yet still remaining below 2 m/s. In contrast, the annual cycle of precipitation differs significantly from other climatic variables. This is due to the Atacama Desert's persistently high temperatures and minimal rainfall, with some years experiencing no precipitation at all. Consequently, the annual cycle is not a reliable reference. Appendix 7.5 provides station data on precipitation in the Atacama Desert for 2018, further illustrating the scarcity and unpredictability of rainfall in this extreme environment.

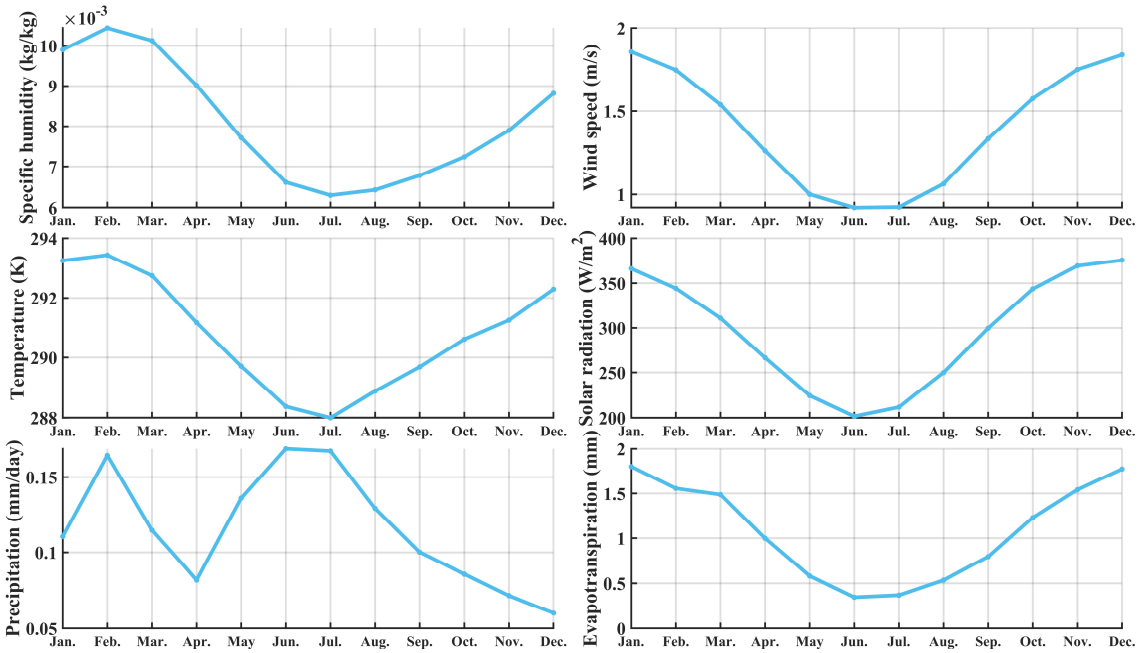


Fig. 5-9 Annual cycle distribution of climate variables in the Atacama Desert.

Fig. 5-10 presents the spatial distribution of climate variables in the Atacama Desert, based on the ERA5 dataset. Areas near the coastline exhibit higher specific humidity and lower temperatures, while the central region of the desert experiences stronger winds and higher precipitation levels. Solar radiation is relatively uniform across the desert, whereas evapotranspiration is higher in the northern region.

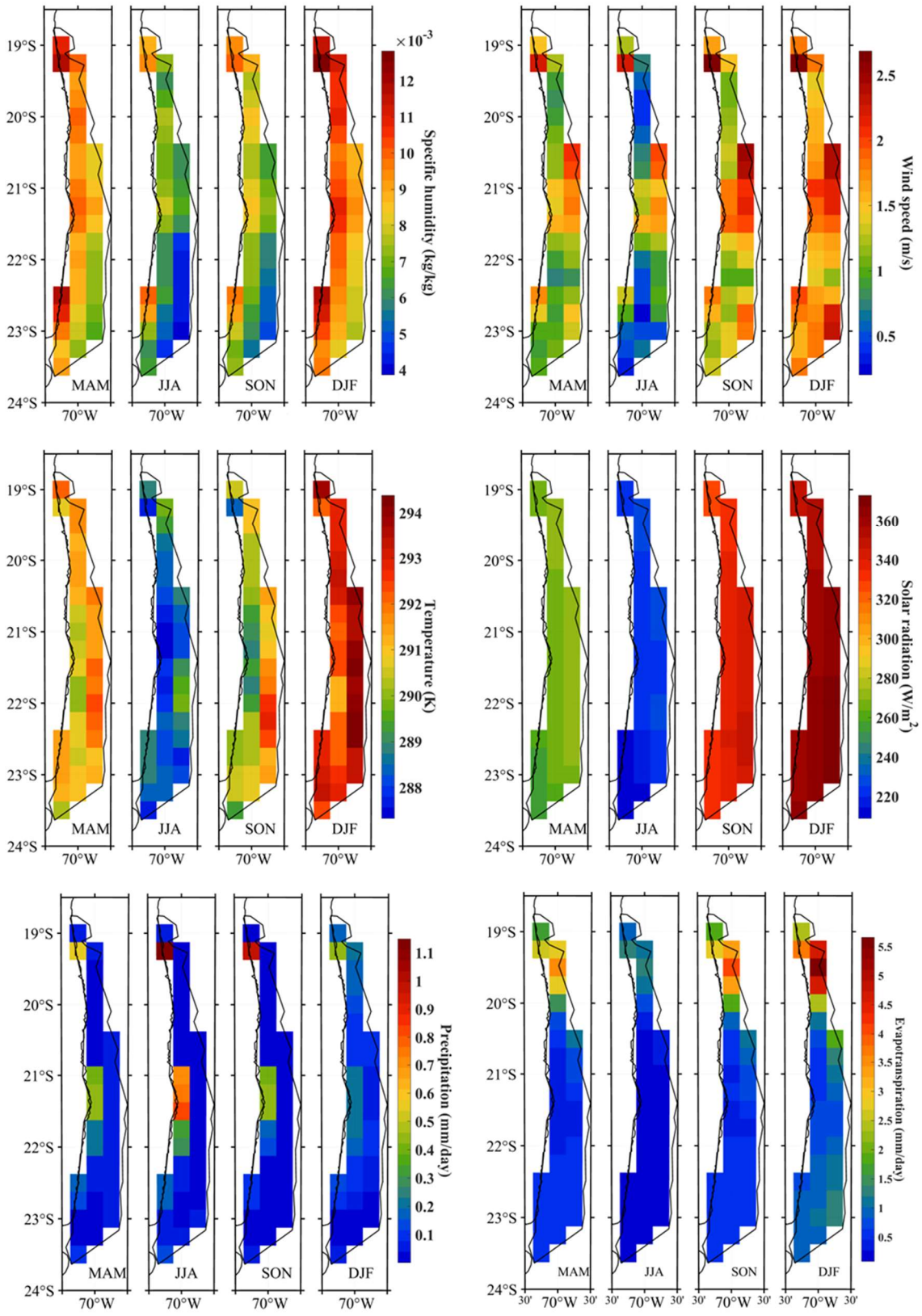


Fig. 5-10 Spatial distribution of climate variables in the Atacama Desert from 2000 to 2018.

5.2.2 Model selection

The model is trained separately for each grid point when modelling Gurbantunggut Desert data. As a result, each grid point is assigned a different regression model. In the fixed-variable model, only the regression coefficients differ, whereas in the machine learning regression model, both the selected variables and regression coefficients vary for each grid point. Therefore, to apply the model trained in the Gurbantunggut Desert to the Atacama Desert, selecting an appropriate regression model is crucial.

The Gurbantunggut Desert consists of 55 grid points, while the Atacama Desert has 35 grid points. The objective of this section is to select a suitable prediction model for each Atacama Desert grid point by comparing the climate data of both deserts. The primary method used is the k-nearest neighbour (KNN) algorithm (Cover & Hart, 1967). The model corresponding to the nearest climate variable lattice point, identified through this method, is applied as the prediction model for the Atacama Desert. The results are presented in Fig. 5-11.

Fig. 5-11a shows the selected month for applying the model to the Atacama Desert. The horizontal axis represents the months in the Atacama Desert, while the numbers inside the rectangular frames indicate the corresponding months in the Gurbantunggut Desert. The vertical axis represents the number of grid points assigned to each month. For example, the climatic conditions in January in the Atacama Desert (x-axis) are most similar to those in June in the Gurbantunggut Desert (indicated by the number 6 in the rectangular frame), reflecting the seasonal differences between the Northern and Southern Hemispheres. Multiple numbers appearing in the stacked bars indicate that the climate at each of the 35 grid points in the Atacama Desert is individually compared. For instance, in March, the climate of 28 grid points in the Atacama Desert is similar to May in the Gurbantunggut Desert, whereas the remaining 7 grid points align more closely with August.

Fig. 5-11b illustrates the selected grid points for applying the model to the Atacama Desert. The horizontal axis represents the months in the Atacama Desert, while the numbers inside the rectangular frames indicate the corresponding grid points in the Gurbantunggut Desert. The vertical axis represents the number of grid points assigned to each month. The colour coding differentiates land cover types: green represents grassland, while beige denotes barren land in the Gurbantunggut Desert. The grid points are associated with specific months. For example, in December, 22 grid points in the Atacama Desert correspond to the first grid point in the Gurbantunggut Desert in June. Additionally, 4, 1, 3, 1, 1, and 3 grid points correspond to the 4th, 11th, 16th, 20th, 39th, and 45th grid points, respectively, with the corresponding month also being June (as indicated in Fig. 5-11a).

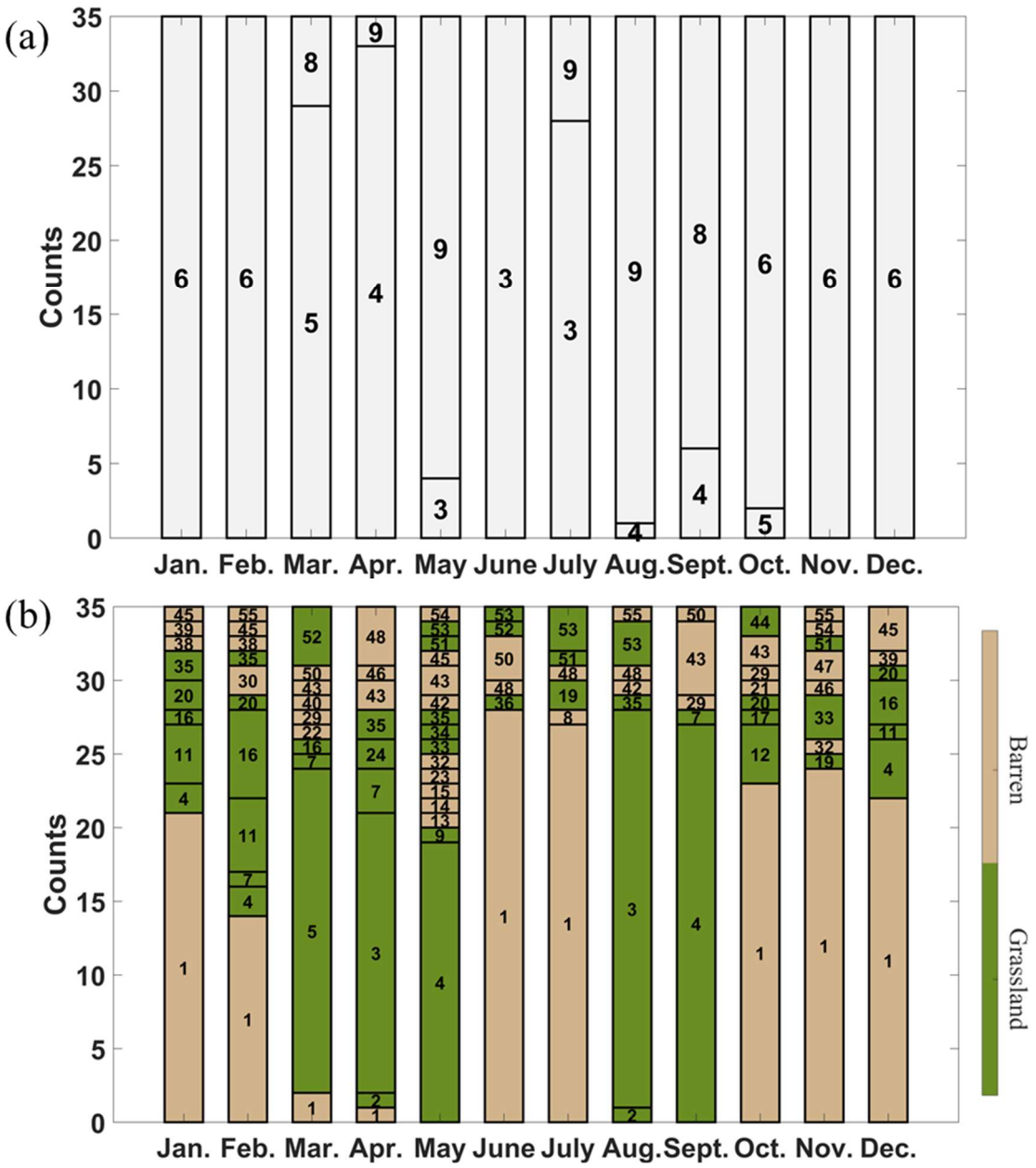


Fig. 5-11 (a) Month and (b) grid point corresponding to the most suitable model. Green represents grassland, while beige denotes barren land in the Gurbantunggut Desert.

Fig. 5-12 presents the spatial distribution corresponding to Fig. 5-11b. By comparing climatic conditions and considering the correspondence between months with similar climates in the Atacama Desert and the Gurbantunggut Desert (Fig. 5-11a), it is evident that most areas in the central and southern Atacama Desert during January, June, July, October, November, and December exhibit climatic similarities to the barren land in

the Gurbantunggut Desert. In February, the climate in central Atacama more closely resembles the grassland in the Gurbantunggut Desert, while the southern region remains closer to barren land. During the remaining months, the central and southern parts of the Atacama Desert share similar climatic conditions with the grassland in the Gurbantunggut Desert. The northern region of the Atacama Desert remains relatively stable, consistently corresponding to the grassland in the Gurbantunggut Desert.

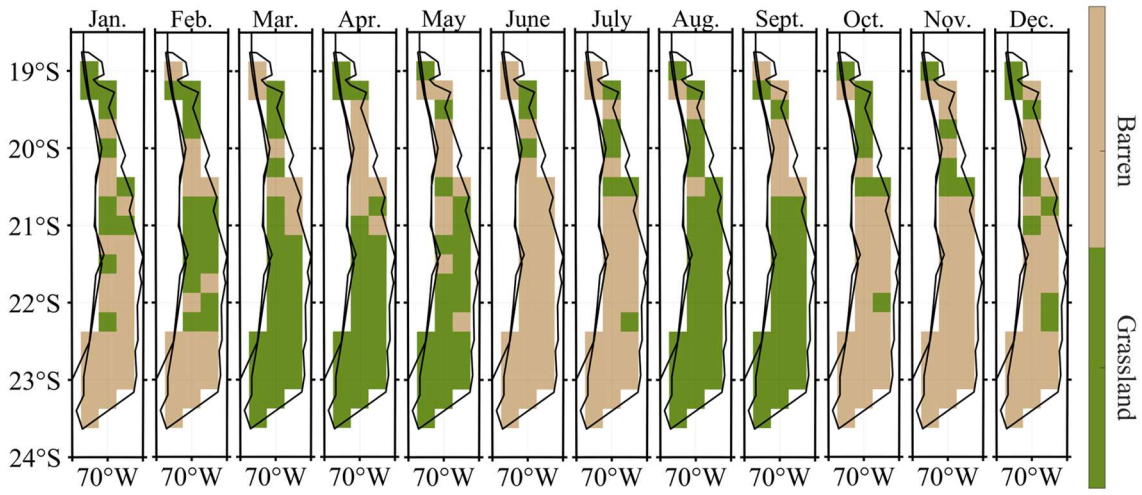


Fig. 5-12 Spatial distribution of grid point corresponding to the most suitable model. Green represents grassland, while beige denotes barren land in the Gurbantunggut Desert.

5.2.3 Results of model application

After selecting the appropriate model according to the previous section, the ERA5 climate data from 2000 to 2018 are used to predict the BSC fraction in the Atacama Desert. The results using the fixed-variables model and the machine learning model are shown in Fig. 5-13 and Fig. 5-14 respectively.

The two models produce similar predictions, with BSC primarily distributed in the northern part of the Atacama Desert, which aligns with the findings of Wang et al. (2024). The BSC fraction values are higher from January to April. However, unlike the fixed-variable model, the machine learning model's predictions indicate that BSC is also present in the central and southern parts of the Atacama Desert during January,

March, and September.

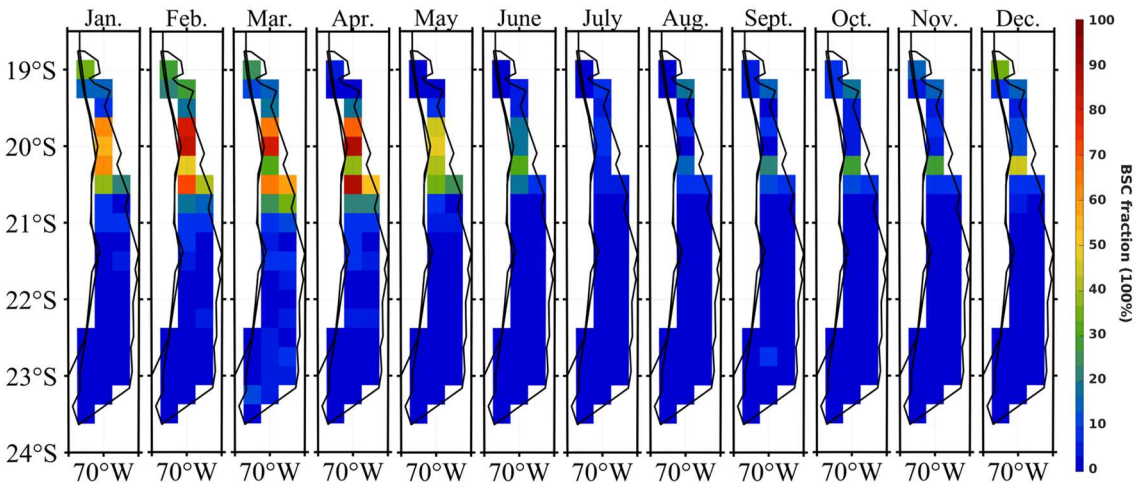


Fig. 5-13 BSC fraction predicted by the fixed-variables model in the Atacama Desert.

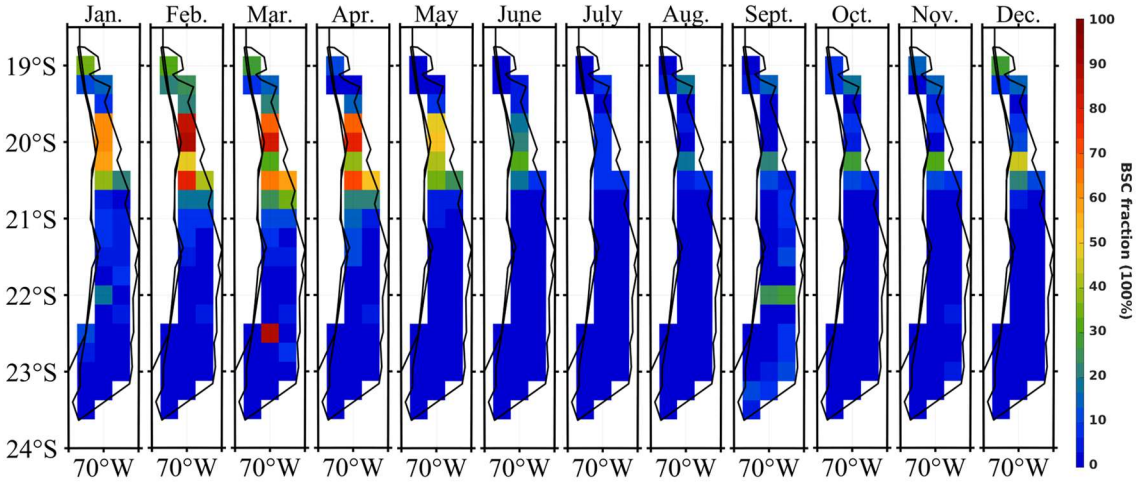


Fig. 5-14 BSC fraction predicted by the machine learning model in the Atacama Desert.

Fig. 5-15 presents the annual cycle of BSC fraction predicted by the fixed-variable model and the machine learning model in the Atacama Desert. The two models produced similar results, with relatively high values from January to April, peaking in March, while the lowest value occurs in June. The main difference between the two models is that the machine learning model predicts a small peak in September. When evaluated alongside Fig. 5-13 and Fig. 5-14, this peak appears to be influenced by the divergent predicted values around 22°S. Similarly, the March peak in the fixed-variable

model is lower than that in the machine learning model, likely due to the predicted values around 22.5°S.

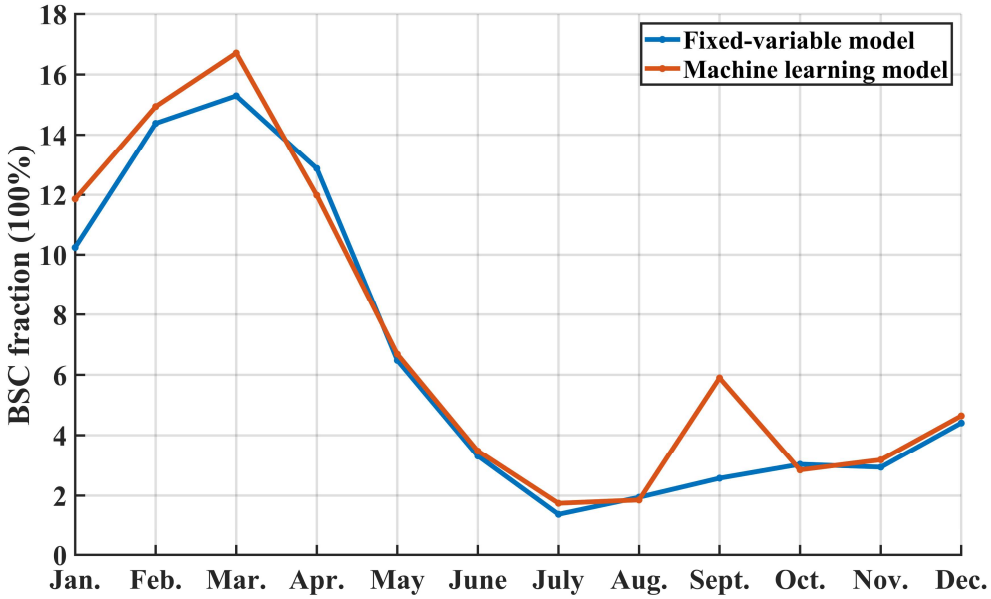


Fig. 5-15 Annual cycle of BSC fraction predicted by the fixed-variables model and the machine learning model in the Atacama Desert.

As there is no directly comparable data on BSC distribution in the Atacama Desert, NDVI is used as a reference indicator. Fig. 5-16 presents the spatial distribution and annual cycle of NDVI in the Atacama Desert. The NDVI values are higher in the northern region (approximately 0.2–0.5) and lower in the southern region (around 0.1–0.15). This pattern aligns with the findings in Fig. 5-13 and Fig. 5-14, which indicate a positive correlation between higher NDVI values and increased BSC coverage. Additionally, when compared with Fig. 5-15, the annual cycle of NDVI exhibits a generally consistent trend with the model predictions, with relatively high values from January to April, peaking in February.

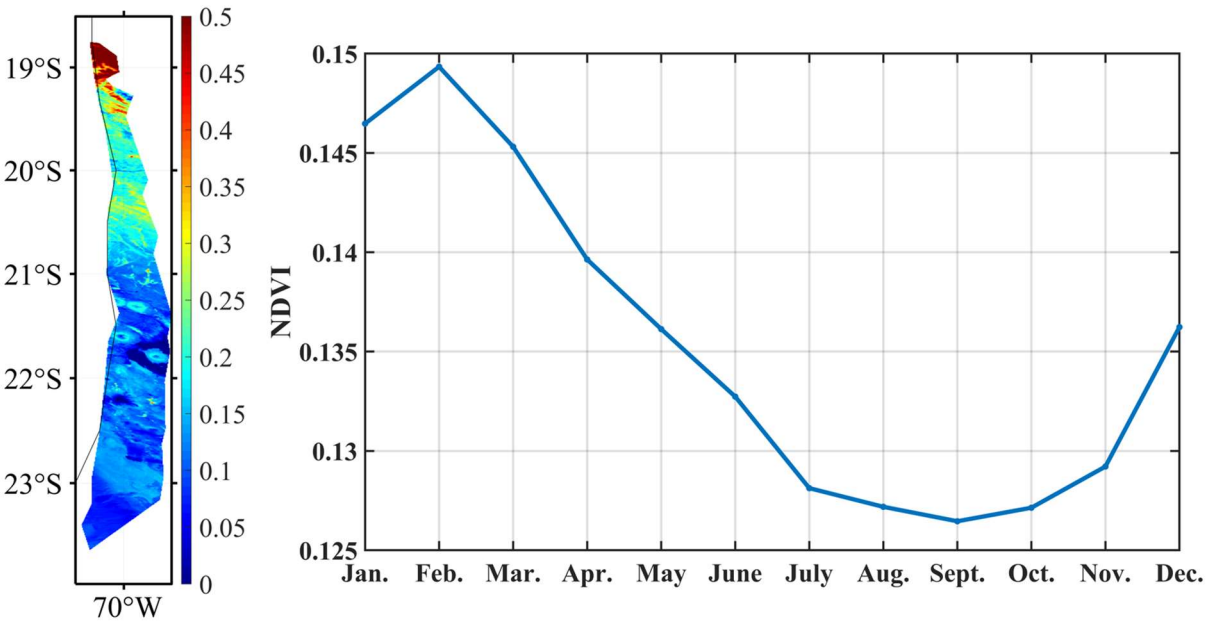


Fig. 5-16 Spatial distribution and annual cycle of NDVI in Atacama Desert from 2000 to 2018.

5.3 Conclusion

In this chapter, the fixed-variable model and the machine learning regression model are applied in two ways, as summarized below:

1) The spatiotemporal distribution of BSC in the Gurbantunggut Desert during the Last Glacial Maximum (LGM, 21 ka) and the Mid-Holocene (6 ka) is predicted using paleoclimate data from CMIP6. The results indicate that BSC is more abundant along the southern edge of the desert, and the BSC fraction is higher in the Mid-Holocene than in the LGM.

2) The temporal and spatial distribution of BSC in the Atacama Desert is predicted using modern climate data from the region. The results suggest that BSC in the Atacama Desert is primarily distributed in the northern part of the desert, with higher BSC fractions from January to April.

3) The predicted BSC distribution in the Atacama Desert is compared with NDVI

data, revealing consistent annual cycle variations. Assuming the NDVI value of BSC is approximately 0.2, the spatial distribution of NDVI closely aligns with the predicted BSC distribution.

6 Conclusion and outlook

6.1 Conclusion

A comprehensive understanding of the response of the biological soil crust to climate change requires BSC-climate models that represent the relevant processes in the atmosphere and on the land surface. In this study, a modelling system for BSC and climate drivers based on multi-datasets is developed in two approaches. The effects of climate variability on the long temporal and large spatial distribution of BSC are revealed by an improved BSC detection method and multiple linear regression. This model can be used to explain the dominant climatic factors associated with changes in BSC, the short- or long-term prediction of regional-scale distribution of BSC, the assessment of potential impacts of climate change on BSC availability and sustainable ecosystem development. Key outcomes include:

1) The long-term and large-scale distribution of biological soil crust (BSC) in the investigated area is determined through a four-step approach: (i) fusion of MODIS and Landsat 7 satellite data using the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) to obtain high spatial and temporal resolution multispectral data; (ii) calculation of BSCI and NDVI from the fused satellite spectral data; (iii) extraction of BSC in the study area based on BSCI thresholds derived from previous studies, with additional consideration of NDVI; and (iv) analysis of the extracted BSC data from multiple perspectives. The 19-year analysis reveals that, on a temporal scale, BSC variations exhibit interannual periodicity, with peaks in March and October and near-zero values in winter. On a spatial scale, BSC is predominantly distributed in the desert-oasis transition zone, with its presence becoming increasingly sparse toward the desert hinterland.

2) Lag correlation and partial correlation between BSC and climate variables are analyzed. In this study, five climatic variables—specific humidity, 10-meter

wind speed, 2-meter temperature, surface solar radiation, and total precipitation—along with their time lags, are used as independent variables. The results indicate that in some areas, BSC exhibits a stronger correlation with time-lagged climate factors when the time lag is considered, with this effect being most pronounced for specific humidity. The BSC response to specific humidity is typically delayed by 1 to 2 months, and in principle, the time lag between BSC and climatic variables does not exceed three months. BSC responds rapidly to temperature, with a correlation coefficient of 0.7, while its response to precipitation is also relatively fast but has a lower correlation coefficient of 0.46, with significantly correlated areas mainly located in the east and south. These correlation analyses provide a valuable reference for the selection of variables in subsequent modeling efforts.

3) Models of biological soil crust and climate factors are constructed using two approaches, both of which account for the influence of time lag. The first approach is based on fixed climate factors, while the second employs a sliding time series method to select the most appropriate climate factors at different time points. Multiple regression analysis is applied to both models, and statistical parameters are used for estimation. The results indicate that the two approaches explain approximately 40% and 75% of BSC variation, respectively. The primary climatic drivers influencing BSC distribution are temperature, specific humidity, and solar radiation.

4) The models are applied to paleoclimate scenarios (the Last Glacial Maximum and the Mid-Holocene) in the Gurbantunggut Desert and to historical climate conditions in the Atacama Desert. Changes in BSC distribution across different time periods are compared and analyzed. The results indicate that BSC is more abundant along the southern edge of the Gurbantunggut Desert, with the Mid-Holocene exhibiting a higher BSC fraction than the LGM. In the Atacama

Desert, BSC is primarily distributed in the northern region, with higher BSC fractions observed from January to April.

6.2 Outlook

Based on the results of this study, future research should focus on the following aspects:

1) Improving prediction accuracy through dataset selection can be achieved in two ways. The first approach is to incorporate multi-model ensemble (MME) data, which can help mitigate biases introduced by individual models. The second is downscaling, where higher-resolution data can better capture geographical and climatological features, leading to more accurate model predictions.

2) Enhancing model calibration by utilizing existing BSC distribution datasets from literature and measured data can improve regional adaptability and predictive accuracy. Future work should focus on refining the model using these datasets to ensure greater reliability across different environments.

3) Expanding the spatial and temporal scope of the model could provide broader insights. If computational resources and memory capacity allow, the model could be extended globally and applied to past or future climate scenarios. Unlike existing static models, the approach used in this study generates a continuous time series of BSC fraction. Scaling up the application of this model could contribute to the development of a more comprehensive dataset of BSC distributions.

7 Appendix

7.1 Kappa index

Assume that n samples are distributed into k^2 cells, here each sample is assigned to one of k categories in the map (usually the rows), and independently to one of the same k categories in the reference data set (usually the columns). Let n_{ij} denote the number of samples classified into category i ($i = 1, 2, \dots, k$) in the map and category j ($j = 1, 2, \dots, k$) in the reference data set (Fig. 7-1) (Congalton & Green, 2019).

		j = Columns (Reference)			
		1	**2**	**k**	**Row Total n_{i+}**
i = Rows (Classification)	**1**	n_{11}	n_{12}	n_{1k}	n_{1+}
	2	n_{21}	n_{22}	n_{2k}	n_{2+}
	k	n_{k1}	n_{k2}	n_{kk}	n_{k+}
	Column Total n_{+j}	n_{+1}	n_{+2}	n_{+k}	n

Fig. 7-1 Mathematical Example of an Error Matrix

Let p_{ij} denote the proportion of samples in the i_{th}, j_{th} cell, corresponding to n_{ij}, which

$$p_{ij} = \frac{n_{ij}}{n}$$

Then let p_{i+} and p_{+j} be defined by

$$p_{i+} = \sum_{j=1}^{k} p_{ij}$$

and

$$p_{+j} = \sum_{i=1}^{k} p_{ij}$$

Let

$$p_c = \sum_{i=1}^{k} p_{i+} p_{+j}$$

be the actual agreement, and

$$p_{+j} = \sum_{i=1}^{k} p_{ij}$$

be the chance agreement.

Assuming a multinomial sampling model, the maximum likelihood estimate of Kappa is given by

$$\widehat{K} = \frac{p_o - p_c}{1 - p_c} \tag{7-1}$$

7.2 Sen's Slope

Sen's slope is a non-parametric statistical estimator used to quantify the magnitude of trends in a time series. Unlike simple linear regression, Sen's method is robust to outliers and non-normally distributed data, making it efficient and resistant to measurement errors and anomalies. This robustness makes it well-suited for detecting long-term trends in environmental and climate-related datasets.

Sen's slope is defined as:

$$Sen's\ slope = Median\left\{\frac{x_j - x_i}{j - i} : i < j\right\} \tag{7-2}$$

A $1-\alpha$ confidence interval for Sen's slope can be calculated as (lower, upper), where:

$$N = C(n, 2),$$

$$k = se \cdot z_{crit},$$

$$lower = m_{\frac{N-2}{2}},$$

$$upper = m_{\frac{N+2}{2}+1}.$$

Here, N = the number of pairs of time series elements (x_i, x_j), where $i < j$ and se is the standard error for the Mann-Kendall Test. Also, m_h is the h^{th} smallest in the set $\{(x_j - x_i)/(j - i): i < j\}$ and z_{crit} is the $1 - \frac{\alpha}{2}$ critical value for the normal distribution.

7.3 Mann-Kendall Test

Mann-Kendall is a non-parametric test method. Compared with other parametric test methods, it does not require the sample to follow a certain distribution, is less disturbed by outliers, and is more suitable for ordinal variables. The Mann-Kendall test has been successfully used in a large number of studies related to hydrological and meteorological trend changes to determine the significance of trends in runoff, precipitation, climate, etc.

The Mann-Kendall Test is used to determine whether a time series has a monotonic upward or downward trend. It does not require that the data be normally distributed or linear. It does require that there is no autocorrelation.

The null hypothesis for this test is that there is no trend, and the alternative hypothesis is that there is a trend in the two-sided test or that there is an upward trend (or downward trend) in the one-sided test. For the time series $x_1, \ldots, x_n$, the MK Test uses the following statistic:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sign(x_j - x_i) \tag{7-3}$$

Note that if S > 0 then later observations in the time series tend to be larger than those that appear earlier in the time series, while the reverse is true if S < 0.

The variance of S is given by

$$var = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_t f_t(f_{t-1})(2f_t+5)\right] \tag{7-4}$$

where t varies over the set of tied ranks and f_t is the number of times (i.e., frequency) that the rank t appears.

The MK Test uses the following test statistic:

$$z = \begin{cases} \frac{S-1}{se}, & S > 0 \\ 0, & S = 0 \\ \frac{S+1}{se}, & S < 0 \end{cases} \tag{7-5}$$

where se is the square root of var. If there is no monotonic trend (the null hypothesis), then for a time series with more than 10 elements, $z \sim N_{(0,1)}$, i.e., z has a standard normal distribution.

7.4 Supplementary figures to Table 4-5

Fig. 7-2 as the supplement to Table 4-5, the monthly box plots for the three machine learning models on different evaluation metrics is shown. The evaluation indicators of the linear regression demonstrate superiority in comparison to both the linear kernel and the RBF kernel support vector regression (SVR) during the growing season.

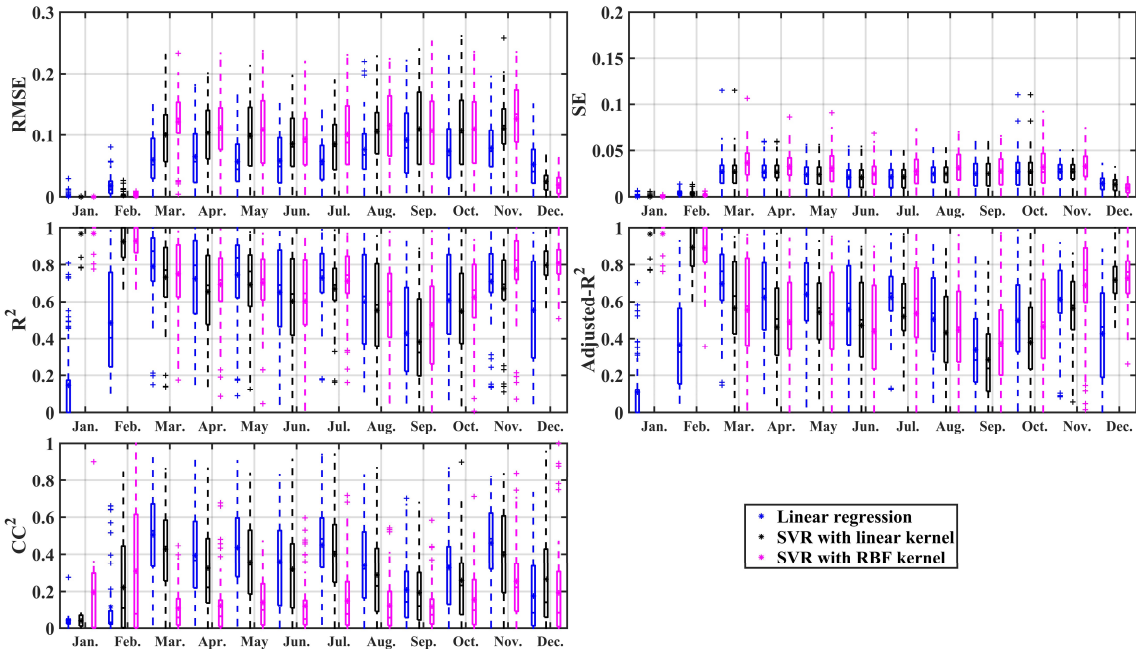


Fig. 7-2 Performance of three machine learning models in the study area. The blue, black and magenta boxes indicate the evaluation indicators for linear regression, linear kernel SVR and RBF kernel SVR, respectively.

7.5 Precipitation station data for the Atacama Desert in 2018

The stations data of Fig. 7-3 is from station C2 of SFB1211 "Earth Evolution at the dry Limit" (Hoffmeister, 2018). Stations 1.1, 1.2 and 1.3 are located in the middle of the desert, and stations 2.1, 2.2 and 2.3 are located in the north of the desert. As can be seen in the figure, even stations in proximity area to each other still have very different precipitation levels. Majority of the stations have very low monthly precipitation. There are many rainy months throughout the year at station 1.3, with higher precipitation in August and September. Station 2.2 has a higher precipitation event in November. This is a supplement to Fig. 5-9 and further illustrates the low and random precipitation in the Atacama Desert.

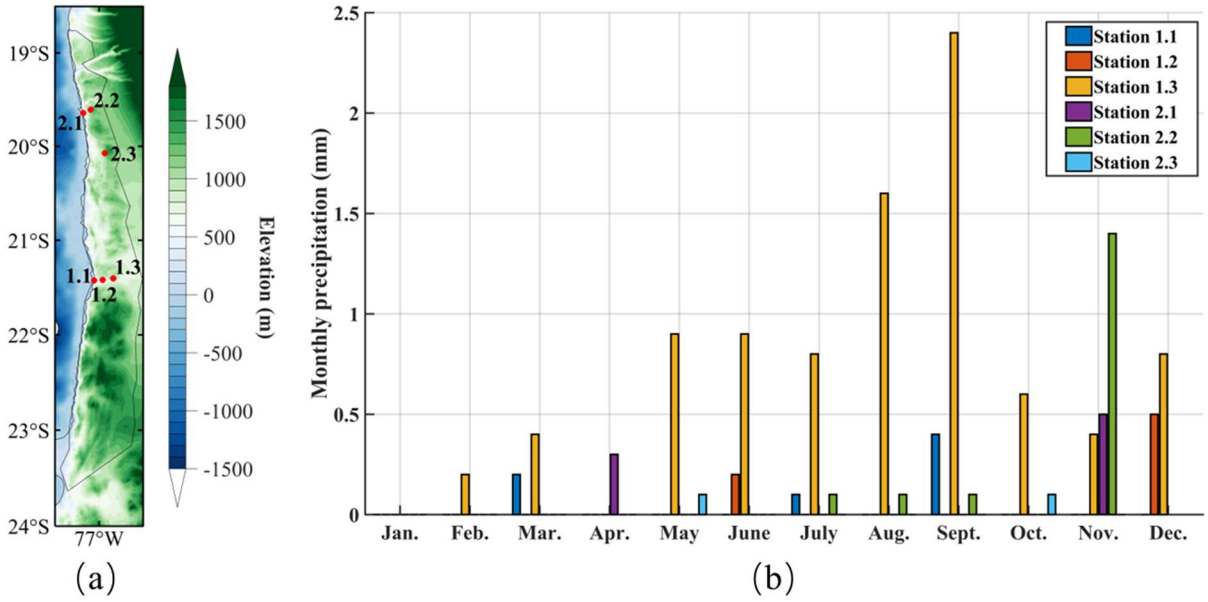


Fig. 7-3 Stations data of precipitation in the Atacama Desert in 2018. (a) Station location (red dot); (b) Monthly precipitation.

8 References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In Second International Symposium on Information Theory (pp. 267–281). Akademiai Kiado.

Akaike, H. (1974). A new look at the statistical model identification. IEEE transactions on automatic control, 19(6), 716-723.

Beaugendre, N., Issa, O. M., Choné, A., Cerdan, O., Desprats, J. F., Rajot, J. L., ... & Valentin, C. (2017). Developing a predictive environment-based model for mapping biological soil crust patterns at the local scale in the Sahel. Catena, 158, 250-265.

Belnap, J. (2002). Nitrogen fixation in biological soil crusts from southeast Utah, USA. Biology and fertility of soils, 35, 128-135.

Belnap, J. (2003). Biological soil crusts and wind erosion. In Biological soil crusts: structure, function, and management (pp. 339-347). Berlin, Heidelberg: Springer Berlin Heidelberg.

Belnap, J. (2006). The potential roles of biological soil crusts in dryland hydrologic cycles. Hydrological Processes: An International Journal, 20(15), 3159-3178.

Belnap, J., Büdel, B., & Lange, O. L. (2003). Biological soil crusts: characteristics and distribution. In Biological soil crusts: structure, function, and management (pp. 3-30). Berlin, Heidelberg: Springer Berlin Heidelberg.

Belnap, J., & Eldridge, D. (2003). Disturbance and recovery of biological soil crusts. In Biological soil crusts: structure, function, and management (pp. 363-383). Berlin, Heidelberg: Springer Berlin Heidelberg.

Belnap, J., Harper, K. T., & Warren, S. D. (1994). Surface disturbance of cryptobiotic soil crusts: nitrogenase activity, chlorophyll content, and chlorophyll degradation. Arid Land Research and Management, 8(1), 1-8.

Belnap, J., Kaltenecker, J. H., Rosentreter, R., Williams, J., Leonard, S., & Eldridge, D. (2001). Biological soil crusts: ecology and management. US Department of the Interior, Bureau of Land Management, National Science and Technology Center. Denver, Colorado, 110.

Belnap, J., & Lange, O. L. (2003). Structure and functioning of biological soil crusts: a synthesis. In Biological soil crusts: structure, function, and management (pp. 471-479). Berlin, Heidelberg: Springer Berlin Heidelberg.

Belnap, J., Miller, D. M., Bedford, D. R., & Phillips, S. L. (2014). Pedological and geological relationships with soil lichen and moss distribution in the eastern Mojave Desert, CA, USA. Journal of arid environments, 106, 45-57.

Belnap, J., Phillips, S. L., & Miller, M. E. (2004). Response of desert biological soil crusts to alterations in precipitation frequency. Oecologia, 141, 306-316.

Belnap, J., Phillips, S. L., & Smith, S. D. (2007). Dynamics of cover, UV-protective pigments, and quantum yield in biological soil crust communities of an undisturbed Mojave Desert shrubland. Flora-Morphology, Distribution, Functional Ecology of Plants, 202(8), 674-686.

Belnap, J., Weber, B., & Büdel, B. (2016). Biological soil crusts as an organizing principle in drylands. In Biological Soil Crusts: An Organizing Principle in Drylands (pp. 3-13). Springer International Publishing.

Billings, S. A., Schaeffer, S. M., & Evans, R. D. (2003). Nitrogen fixation by biological soil crusts and heterotrophic bacteria in an intact Mojave Desert ecosystem with elevated CO2 and added soil carbon. Soil Biology and Biochemistry, 35(5), 643-649.

Blanco-Sacristán, J., Panigada, C., Gentili, R., Tagliabue, G., Garzonio, R., Martín, M. P., ... & Rossini, M. (2021). UAV RGB, thermal infrared and multispectral imagery used to investigate the control of terrain on the spatial distribution of dryland biocrust. Earth Surface Processes and Landforms, 46(12), 2466-2484.

Blank, G. B., & Cameron, R. E. (1966). Desert algae-Soil crusts and diaphanous substrata as algal habitats (No. JPL-TR-32-971).

Bowker, M. A. (2007). Biological soil crust rehabilitation in theory and practice: an underexploited opportunity. Restoration Ecology, 15(1), 13-23.

Bowker, M. A., Belnap, J., Davidson, D. W., & Phillips, S. L. (2005). Evidence for micronutrient limitation of biological soil crusts: importance to arid-lands restoration. Ecological Applications, 15(6), 1941-1951.

Bowker, M. A., Belnap, J., Davidson, D. W., & Goldstein, H. (2006). Correlates of biological soil crust abundance across a continuum of spatial scales: support for a hierarchical conceptual model. Journal of Applied Ecology, 43(1), 152-163.

Bowker, M. A., & Belnap, J. (2008). A simple classification of soil types as habitats of biological soil crusts on the Colorado Plateau, USA. Journal of Vegetation Science, 19(6), 831-840.

Bowker, M. A., Maestre, F. T., & Escolar, C. (2010). Biological crusts as a model system for examining the biodiversity–ecosystem function relationship in soils. Soil Biology and Biochemistry, 42(3), 405-417.

Bowker, M. A., Maestre, F. T., Eldridge, D., Belnap, J., Castillo-Monroy, A., Escolar, C., & Soliveres, S. (2014). Biological soil crusts (biocrusts) as a model system in community, landscape and ecosystem ecology. Biodiversity and Conservation, 23, 1619-1637.

Bu, C., Zhao, Y., Hill, R. L., Zhao, C., Yang, Y., Zhang, P., & Wu, S. (2015). Wind erosion prevention characteristics and key influencing factors of bryophytic soil crusts. Plant and soil, 397, 163-174.

Bu, C., Zhang, P., Wang, C., Yang, Y., Shao, H., & Wu, S. (2016). Spatial distribution of biological soil crusts on the slope of the Chinese Loess Plateau based on canonical correspondence analysis. Catena, 137, 373-381.

Burgheimer, J., Wilske, B., Maseyk, K., Karnieli, A., Zaady, E., Yakir, D., & Kesselmeier, J. (2006a). Relationships between Normalized Difference Vegetation Index (NDVI) and carbon fluxes of biologic soil crusts assessed by ground measurements. Journal of Arid Environments, 64(4), 651-669.

Burgheimer, J., Wilske, B., Maseyk, K., Karnieli, A., Zaady, E., Yakir, D., & Kesselmeier, J. (2006b). Ground and space spectral measurements for assessing the semi-arid ecosystem phenology related to CO2 fluxes of biological soil crusts. Remote Sensing of Environment, 101(1), 1-12.

Büdel, B., Darienko, T., Deutschewitz, K., Dojani, S., Friedl, T., Mohr, K. I., & Weber, B. (2009). Southern African biological soil crusts are ubiquitous and highly diverse in drylands, being restricted by rainfall frequency. Microbial Ecology, 57(2), 229–247.

Byers, H. R. (1965). Elements of cloud physics. (No Title).

Castillo-Monroy, A. P., Maestre, F. T., Rey, A., Soliveres, S., & García-Palacios, P. (2011). Biological soil crust microsites are the main contributor to soil respiration in a semiarid ecosystem. Ecosystems, 14, 835-847.

Chamizo, S., Stevens, A., Cantón, Y., Miralles, I., Domingo, F., & Van Wesemael, B. (2012). Discriminating soil crust type, development stage and degree of disturbance in semiarid environments from their spectral characteristics. European Journal of Soil Science, 63(1), 42-53.

Chamizo, S., Belnap, J., Eldridge, D. J., Cantón, Y., & Issa, O. M. (2016). The role of biocrusts in arid land hydrology. In Biological soil crusts: An organizing principle in drylands (pp. 321–346). Springer International Publishing.

Chaudhary, V. B., Bowker, M. A., O'Dell, T. E., Grace, J. B., Redman, A. E., Rillig, M. C., & Johnson, N. C. (2009). Untangling the biological contributions to soil stability in semiarid shrublands. Ecological Applications, 19(1), 110-122.

Chávez, R. O., Moreira-Muñoz, A., Galleguillos, M., Olea, M., Aguayo, J., Latín, A., ... & Manríquez, H. (2019). GIMMS NDVI time series reveal the extent, duration, and intensity of "blooming desert" events in the hyper-arid Atacama Desert, Northern Chile. International Journal of Applied Earth Observation and Geoinformation, 76, 193-203.

Chen, Y. (2007). Study of effect of soil microorganism on biological soil crust in the treatment of desert soil. Inner Mongolia Normal University, Hohhot, Inner Mongolia, China.

Chen, F., Chen, J., Huang, W., Chen, S., Huang, X., Jin, L., ... & Feng, S. (2019). Westerlies Asia and monsoonal Asia: Spatiotemporal differences in climate change and possible mechanisms on decadal to sub-orbital timescales. Earth-science reviews, 192, 337-354.

Chen, F., Jia, J., Chen, J., Li, G., Zhang, X., Xie, H., ... & An, C. (2016). A persistent Holocene wetting trend in arid central Asia, with wettest conditions in the late Holocene, revealed by multi-proxy analyses of loess-paleosol sequences in Xinjiang, China. Quaternary Science Reviews, 146, 134-146.

Chen, Y., Li, W., Zhou, Z., & Liu, J. (2005). Ecological and environmental explanation of microbiotic crusts on sand dune scales in the Gurbantonggut Desert, Xinjiang. Progress in Natural Science, 15(12), 1089-1095.

Chen, L., Liu, Y., Li, D., Shen, Y., & Xie, Z. (2003). The research process/progress of desert algae and crust. Science Foundation in China, 17(2):90-93.

Chen, Y., Wang, Q., Li, W., & Ruan, X. (2007). Microbiotic crusts and their interrelations with environmental factors in the Gurbantonggut desert, western China. Environmental Geology, 52, 691-700.

Chen, F., Yu, Z., Yang, M., Ito, E., Wang, S., Madsen, D. B., ... & Wünnemann, B. (2008). Holocene moisture evolution in arid central Asia and its out-of-phase relationship with Asian monsoon history. Quaternary Science Reviews, 27(3-4), 351-364.

Chen, J., Zhang, M. Y., Wang, L., Shimazaki, H., & Tamura, M. (2005). A new index for mapping lichen-dominated biological soil crusts in desert areas. Remote Sensing of Environment, 96(2), 165-175.

Clarke, J., (2006). Antiquity of aridity in the Chilean Atacama Desert. Geomorphology 73:101–114.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and psychological measurement, 20(1), 37-46.

Congalton, R. G., & Green, K. (2019). Assessing the accuracy of remotely sensed data: Principles and practices (3rd ed.). CRC Press.

Conti, M. E., & Cecchetti, G. (2001). Biological monitoring: lichens as bioindicators of air pollution assessment—a review. Environmental pollution, 114(3), 471-492.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE transactions on information theory, 13(1), 21-27.

Crucil, G., & Van Oost, K. (2021). Towards mapping of soil crust using multispectral imaging. Sensors, 21(5), 1850.

Dillon, M. O., & Hoffmann, A. E. (1997). Lomas formations of the Atacama Desert, northern Chile. Centres of plant diversity, a guide and strategy for their conservation, 3, 528-535.

Dong, W., & Meng, J. (2018). Review of spatiotemporal fusion model of remote sensing data. Remote Sensing for Land & Resources, 30(2), 1–11.

Draper, N. R., & Smith, H. (1998). Applied regression analysis (Vol. 326). John Wiley & Sons.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support vector regression machines. Advances in neural information processing systems, 9.

Duan, Z., Wang, G., Xiao, H., & Dong, Z. (2003). Abiotic soil crust formation on Dunesin an extremely arid environment: a 43-year sequential study. Arid Land Research and Management, 17(1), 43-54.

Dunai, T. J., López, G. A. G., & Juez-Larré, J. (2005). Oligocene–Miocene age of aridity in the Atacama Desert revealed by exposure dating of erosion-sensitive landforms. Geology, 33(4), 321-324.

Durham, R. A., Doherty, K. D., Antoninka, A. J., Ramsey, P. W., & Bowker, M. A. (2018). Insolation and disturbance history drive biocrust biodiversity in Western Montana rangelands. Plant and Soil, 430, 151-169.

Dyke, A. S. (2004). An outline of North American deglaciation with emphasis on central and northern Canada. Developments in quaternary sciences, 2, 373-424.

Eisele, A., Lau, I., Hewson, R., Carter, D., Wheaton, B., Ong, C., ... & Kaufmann, H. (2012). Applicability of the thermal infrared spectral region for the prediction of soil properties across semi-arid agricultural landscapes. Remote Sensing, 4(11), 3265-3286.

Eisele, A., Chabrillat, S., Hecker, C., Hewson, R., Lau, I. C., Rogass, C., ... & Kaufmann, H. (2015). Advantages using the thermal infrared (TIR) to detect and quantify semi-arid soil properties. Remote Sensing of Environment, 163, 296-311.

Elbert, W., Weber, B., Burrows, S., Steinkamp, J., Büdel, B., Andreae, M. O., & Pöschl, U. (2012). Contribution of cryptogamic covers to the global cycles of carbon and nitrogen. Nature Geoscience, 5(7), 459-462.

Eshel G., Araus V., Undurraga S., Soto D. C., Moraga C., Montecinos A., Gutiérrez R. A. (2021). Plant ecological genomics at the limits of life in the Atacama Desert. Proceedings of the National Academy of Sciences, 118(46), e2101177118.

Eldridge, D. J. (1996). Distribution and floristics of terricolous lichens in soil crusts in arid and semi-arid New South Wales, Australia. Australian Journal of Botany, 44(5), 581-599.

Eldridge, D. J. (1999). Distribution and floristics of moss-and lichen-dominated soil crusts in a patterned Callitris glaucophylla woodland in eastern Australia. Acta Oecologica, 20(3), 159-170.

Eldridge, D. J. (2003). Biological soil crusts and water relations in Australian deserts. In Biological soil crusts: Structure, function, and management (pp. 315-325). Berlin, Heidelberg: Springer Berlin Heidelberg.

Eldridge, D. J., & Greene, R. S. B. (1994). Microbiotic soil crusts-a review of their roles in soil and ecological processes in the rangelands of Australia. Soil Research, 32(3), 389-415.

Eldridge, D., & Leys, J. (2003). Exploring some relationships between biological soil crusts, soil aggregation and wind erosion. Journal of arid environments, 53(4), 457-466.

Eldridge, D. J., Reed, S., Travers, S. K., Bowker, M. A., Maestre, F. T., Ding, J., Havrilla, C., Rodríguez-Caballero, E., Weber, B., & Zaady, E. (2020). The pervasive and multifaceted influence of biocrusts on water in the world's drylands. Global Change Biology, 26(10), 6003–6014.

Eldridge, D. J., & Tozer, M. E. (1997). Environmental factors relating to the distribution of terricolous bryophytes and lichens in semi-arid eastern Australia. The Bryologist, 100(1), 28–39.

Ericksen G. (1983) The Chilean nitrate deposits. Am Sci 71:366–374.

Escribano, P., Palacios-Orueta, A., Oyonarte, C., & Chabrillat, S. (2010). Spectral properties and sources of variability of ecosystem components in a Mediterranean semiarid environment. Journal of Arid Environments, 74(9), 1041-1051.

Ezekiel, M. (1930). Methods of correlation analysis.

Feng, X., Bu, C., Hao, H., Yang, Y., & Zhang, G. (2015). Research on Biological Soil Crust Extraction by Spectral Analysis in Mu Us Desert, China. Journal of Natural Resources, 30(6), 1024-1034.

Follmann G. (2008). Two new crustaceous soil lichens (Arthoniales) from the Chilean Atacama Desert, South America. Herzogia 21:25–39.

Friedmann, E. I., & Galun, M. (1974). Desert algae, lichens and fungi. Desert biology, 2, 165-212.

Fritsch, F. E. (1907). The role of algal growth in the colonization of new ground and in the determination of scenery. Geographical Journal, 531-548.

Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. IEEE Transactions on Geoscience and Remote sensing, 44(8), 2207-2218.

Garcia-Pichel, F., Johnson, S. L., Youngkin, D., & Belnap, J. (2003). Small-scale vertical distribution of bacterial biomass and diversity in biological soil crusts from arid lands in the Colorado Plateau. Microbial Ecology, 46, 312-321.

Gevaert, C. M., & García-Haro, F. J. (2015). A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. Remote sensing of Environment, 156, 34-44.

Green, G. M. (1986). Use of SIR-A and Landsat MSS data in mapping shrub and intershrub vegetation at Koonamore, South Australia. Photogrammetric engineering and remote sensing, 52(5), 659-670.

Guo, C., Ning, N., Guo, H., Tian, Y., Bao, A., & De Maeyer, P. (2024). Does ERA5-Land Effectively Capture Extreme Precipitation in the Yellow River Basin?. Atmosphere, 15(10), 1254.

Guo, J., Xu, J., & Yan, C. (2012). A Review on Effects of Biological Soil Crusts on Soil Water Cycle. Journal of Inner Mongolia Normal University (Natural Science Edition, Chinese), 41(1), 99–104.

Hui, R., Liu, L., Xie, M., & Yang, H. (2019). Variation in snow cover drives differences in soil properties and microbial biomass of BSC in the Gurbantunggut Desert—3 years of snow manipulations. Ecohydrology, 12(6), e2118.

Huang, B., & Song, H. (2012). Spatiotemporal reflectance fusion via sparse representation. IEEE Transactions on Geoscience and Remote Sensing, 50(10), 3707-3716.

Hilker, T., Wulder, M. A., Coops, N. C., Linke, J., McDermid, G., Masek, J. G., ... & White, J. C. (2009). A new data fusion model for high spatial-and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. Remote Sensing of Environment, 113(8), 1613-1627.

Hill, J., Udelhoven, T., Schutt, B., & Yair, A. (1998). Differentiating biological soil crusts in a sandy arid ecosystem based on multi-and hyperspectral remote sensing data. In EARSeL Workshop on Imaging Spectroscopy (pp. 427-436).

Hill, J., Udelhoven, T., Jarmer, T., & Yair, A. (2008). Land cover in the Nizzana sandy arid ecosystem. Mapping surface properties with multi-spectral remote sensing data. In Arid dune ecosystems: the Nizzana sands in the Negev desert (pp. 157-172). Berlin, Heidelberg: Springer Berlin Heidelberg.

Hoffmeister, D. (2018). Meteorological and soil measurements of the permanent weather stations in the Atacama Desert, Chile. CRC1211 Database (CRC1211DB). DOI: 10.5880/CRC1211DB.1.

Houston, J., & Hartley, A. J. (2003). The central Andean west-slope rainshadow and its potential contribution to the origin of hyper-aridity in the Atacama Desert. International Journal of Climatology: A Journal of the Royal Meteorological Society, 23(12), 1453-1464.

Hu C., Liu Y. (2000). Species composition and distribution of algae in semi desert algal crusts[J]. Chinese Journal of Applied Ecology, 11(2): 62-66.

Jacobs, A. F., Heusinkveld, B. G., & Berkowicz, S. M. (1999). Dew deposition and drying in a desert system: a simple simulation model. Journal of Arid Environments, 42(3), 211-222.

Jacobs, A. F., Heusinkveld, B. G., & Berkowicz, S. M. (2000). Dew measurements along a longitudinal sand dune transect, Negev Desert, Israel. International Journal of Biometeorology, 43, 184-190.

Jia, R., Chen, N., Yu, K., Wang, Y., Zhang, L., Shao, Y., Xu, X., Zhou, H., Huang, L., & Shao, M. (2019). High rainfall frequency promotes the dominance of biocrust under low annual rainfall. Plant and Soil, 435(1-2), 257–275.

Jiang, C., Wu, L., Liu, D., & Wang, S. (2019). Dynamic monitoring of eco-environmental quality in arid desert areas by remote sensing: A case study of the Gurbantunggut Desert, China. Chinese Journal of Applied Ecology, 30(3), 877–883.

Jin, L., Chen, F., Morrill, C., Otto-Bliesner, B. L., & Rosenbloom, N. (2012). Causes of early Holocene desertification in arid central Asia. Climate Dynamics, 38, 1577-1591.

Karnieli, A. (1997). Development and implementation of spectral crust index over dune sands. International Journal of Remote Sensing, 18(6), 1207-1220.

Karnieli, A., Gabai, A., Ichoku, C., Zaady, E., & Shachak, M. (2002). Temporal dynamics of soil and vegetation spectral responses in a semi-arid environment. International Journal of Remote Sensing, 23(19), 4073-4087.

Karnieli, A., Kidron, G. J., Glaesser, C., & Ben-Dor, E. (1999). Spectral characteristics of cyanobacteria soil crust in semiarid environments. Remote Sensing of Environment, 69(1), 67-75.

Karnieli, A., & Sarafis, V. (1996). Reflectance spectrophotometry of cyanobacteria within soil crusts—a diagnostic tool. International Journal of Remote Sensing, 17(8), 1609-1615.

Karnieli, A., Shachak, M., Tsoar, H., Zaady, E., Kaufman, Y., Danin, A., & Porter, W. (1996). The effect of microphytes on the spectral reflectance of vegetation in semiarid regions. Remote Sensing of Environment, 57(2), 88-96.

Karnieli, A., & Tsoar, H. (1995). Spectral reflectance of biogenic crust developed on desert dune sand along the Israel-Egypt border. Remote Sensing, 16(2), 369-374.

Kidron, G. J. (2019a). The enigmatic absence of cyanobacterial biocrusts from the Namib fog belt: Do dew and fog hold the key? Flora, 257, 151416.

Kidron, G. J. (2019b). Biocrust research: A critical view on eight common hydrological-related paradigms and dubious theses. Ecohydrology, 12(2), e2061.

Kidron, G. J., Herrnstadt, I., & Barzilay, E. (2002). The role of dew as a moisture source for sand microbiotic crusts in the Negev Desert, Israel. Journal of Arid Environments, 52(4), 517-533.

Kidron, G. J., Vonshak, A., Dor, I., Barinova, S., & Abeliovich, A. (2010). Properties and spatial distribution of microbiotic crusts in the Negev Desert, Israel. Catena, 82(2), 92-101.

Lan, S., Wu, L., Zhang, D., & Hu, C. (2015). Analysis of environmental factors determining development and succession in biological soil crusts. Science of the Total Environment, 538, 492-499.

Li, C., Fu, B., Wang, S., Liu, Y., Liu, Y., Feng, X., Wang, Y., Wang, Y., Liu, G., & Lyu, Y. (2021). Drivers and impacts of changes in China's drylands. Nature Reviews Earth & Environment, 2(12), 858–873.

Li, X., He, M., Zerbe, S., Li, X., & Liu, L. (2010). Micro-geomorphology determines community structure of biological soil crusts at small scales. Earth Surface Processes and Landforms, 35(8), 932-940.

Li X. Jia Y., Long L., Wang X., Zhang J, (2001). Advances in microbiotic soil crust research and its ecological significance in arid and semiarid regions. Journal of Desert Research, 21(1): 4-11.

Li, J. Y., Jin, X. Y., Zhang, X. C., Chen, L., Liu, J. L., Zhang, H. M., ... & Jin, D. (2020). Comparative metagenomics of two distinct biological soil crusts in the Tengger Desert, China. Soil Biology and Biochemistry, 140, 107637.

Li, X., Song, G., Hui, R., & Wang, Z. (2017). Precipitation and topsoil attributes determine the species diversity and distribution patterns of crustal communities in desert ecosystems. Plant and Soil, 420, 163-175.

Li, X., Zhao, Y., & Zhang, Y. (2009). A study of biological soil crusts: recent development, trend and prospect. Advances in earth science, 24(1), 11-24.

Liao, C., Wang, J., Pritchard, I., Liu, J., & Shang, J. (2017). A spatio-temporal data fusion model for generating NDVI time series in heterogeneous regions. Remote Sensing, 9(11), 1125.

Liu, X., Liu, J., Shen, C. C., Yang, Y., Chen, J., Chen, S., ... & Chen, F. (2020). Inconsistency between records of δ18O and trace element ratios from stalagmites: Evidence for increasing mid–late Holocene moisture in arid central Asia. The Holocene, 30(3), 369-379.

Liu, M., Yang, W., Zhu, X., Chen, J., Chen, X., Yang, L., & Helmer, E. H. (2019). An Improved Flexible Spatiotemporal DAta Fusion (IFSDAF) method for producing high spatiotemporal resolution normalized difference vegetation index time series. Remote sensing of environment, 227, 74-89.

Liu, F., Zhang, G., Yang, H., Gao, J., Cui, Q., & Li, B. (2014). Effect of wind direction and Artemisia ordosica on the distribution pattern of biological soil crust. Science of Soil and Water Conservation (04),100-105.

Liu, C., Zhou, X., Lu, Y., Zhang, S., Liu, X., Zhang, Y., & Zhang, L. (2023). Effects of biological soil crusts on vertical distribution of soil nitrogen pools in the Gurbantunggut Desert. Acta Ecologica Sinica, 43(12), 5005–5016.

Lu, X., Zhang, K., & Li, R. (2007). The Study on Main Factor About Influence of the Living Beings form Covers in Agriculture and Animal Husbandry of the North Interlocks. Research of Soil and Water Conservation, 14(6), 1–4.

Maestre, F. T., Bowker, M. A., Cantón, Y., Castillo-Monroy, A. P., Cortina, J., Escolar, C., ... & Martínez, I. (2011). Ecology and functional roles of biological soil crusts in semi-arid ecosystems of Spain. Journal of arid environments, 75(12), 1282-1291.

Maestre, F. T., & Cortina, J. (2002). Spatial patterns of surface soil properties and vegetation in a Mediterranean semi-arid steppe. Plant and soil, 241, 279-291.

Maestre, F.T., Escolar, C., de Guevara, M.L., Quero, J.L., Lázaro, R., Delgado-Baquerizo, M., Ochoa, V., Berdugo, M., Gozalo, B. and Gallardo, A. (2013). Changes in biocrust cover drive carbon cycle responses to climate change in drylands. Glob Change Biol, 19: 3835-3847.

Maestre, F. T., Quero, J. L., Gotelli, N. J., Escudero, A., Ochoa, V., Delgado-Baquerizo, M., ... & Zaady, E. (2012). Plant species richness and ecosystem multifunctionality in global drylands. Science, 335(6065), 214-218.

Mallows, C. L. (1973). Some Comments on C p. Technometrics, 15(4), 661-675.

Marsh, J., Nouvet, S., Sanborn, P., Addison, J., Sutherland, G. D., & Bull, G. (2006). Composition and function of biological soil crust communities along topographic gradients in grasslands of central interior British Columbia (Chilcotin) and southwestern Yukon (Kluane). Canadian Journal of Botany, 84(5), 717–736.

Meng, J., Bu, C., Zhang, X., & Sun, Z. (2011). Developmental Characteristics of Biological Soil Crust under Different Vegetation Types in Wind-water Erosion Crisscross Region, Northern Shaanxi Province, China. Journal of Northwest Forestry University, 26(4), 41–46.

Miller A., (1976) The climate of Chile. In: Schwerdfeger W (ed) World Survey of Climatology, vol 12 Climate of Central and South America. Elsevier. Amsterdam, pp 113–145.

McKay CP, Friedmann EI, Gómez-Silva B, Cáceres L, Andersen DT, Landheim R. (2003) Temperature and moisture conditions for life in the extreme arid region of the Atacama Desert: Four years of observations including the El Niño of 1997–98. Astrobiology 3:393–406.

Ochoa-Hueso, R., Hernandez, R. R., Pueyo, J. J., & Manrique, E. (2011). Spatial distribution and physiology of biological soil crusts from semi-arid central Spain are related to soil chemistry and shrub cover. Soil Biology and Biochemistry, 43(9), 1894-1901.

O'neill, A. L. (1994). Reflectance spectra of microphytic soil crusts in semi-arid Australia. Remote Sensing, 15(3), 675-681.

Panigada, C., Tagliabue, G., Zaady, E., Rozenstein, O., Garzonio, R., Di Mauro, B., ... & Rossini, M. (2019). A new approach for biocrust and vegetation monitoring in drylands using multi-temporal Sentinel-2 images. Progress in Physical Geography: Earth and Environment, 43(4), 496-520.

Pearson, K. (1896). Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 187, 253–318.

Piao, S., Fang, J., Zhou, L., Ciais, P., & Zhu, B. (2006). Variations in satellite-derived phenology in China's temperate vegetation. Global change biology, 12(4), 672-685.

Pickard, J. (1986). Antarctic oasis: terrestrial environments and history of the Vestfold Hills. Academic Press.

Praetorius, S. K., McManus, J. F., Oppo, D. W., & Curry, W. B. (2008). Episodic reductions in bottom-water currents since the last ice age. Nature Geoscience, 1(7), 449-452.

Qian, Y., Wu, Z., Zhang, L., Zhao, R., Wang, X., & Li, Y. (2007). Spatial distribution characteristics of ephemeral plants in the Gurbantünggüt Desert. Chinese Science Bulletin, 52(19), 2299–2306.

Qin, Z., Li, W., Burgheimer, J., & Karnieli, A. (2006). Quantitative estimation of land cover structure in an arid region across the Israel–Egypt border using remote sensing data. Journal of arid environments, 66(2), 336-352.

Read, C. F., Duncan, D. H., Vesk, P. A., & Elith, J. (2014). Biocrust morphogroups provide an effective and rapid assessment tool for drylands. Journal of Applied Ecology, 51(6), 1740-1749.

Rivera-Aguilar, V., Montejano, G., Rodríguez-Zaragoza, S., & Durán-Díaz, A. (2006). Distribution and composition of cyanobacteria, mosses and lichens of the biological soil crusts of the Tehuacán Valley, Puebla, México. Journal of arid environments, 67(2), 208-225.

Rodríguez-Caballero, E., Escribano, P., & Cantón, Y. (2014). Advanced image processing methods as a tool to map and quantify different types of biological soil crust. ISPRS Journal of Photogrammetry and Remote Sensing, 90, 59-67.

Rodríguez-Caballero, E., Knerr, T., & Weber, B. (2015). Importance of biocrusts in dryland monitoring using spectral indices. Remote Sensing of Environment, 170, 32-39.

Root, H. T., & McCune, B. (2012). Regional patterns of biological soil crust lichen species composition related to vegetation, soils, and climate in Oregon, USA. Journal of Arid Environments, 79, 93-100.

Rozenstein, O., & Adamowski, J. (2017). A review of progress in identifying and characterizing biocrusts using proximal and remote sensing. International Journal of Applied Earth Observation and Geoinformation, 57, 245-255.

Rosentreter, R., & Belnap, J. (2003). Biological soil crusts of North America. In Biological soil crusts: structure, function, and management (pp. 31-50). Berlin, Heidelberg: Springer Berlin Heidelberg.

Rozenstein, O., & Karnieli, A. (2015). Identification and characterization of Biological Soil Crusts in a sand dune desert environment across Israel–Egypt border using LWIR emittance spectroscopy. Journal of Arid Environments, 112, 75-86.

Rundel, P. W., Dillon, M. O., Palma, B., Mooney, H. A., Gulmon, S. L., & Ehleringer, J. R. (1991). The phytogeography and ecology of the coastal Atacama and Peruvian deserts. Aliso: A Journal of Systematic and Floristic Botany, 13(1), 1-49.

Schieferstein, B., & Loris, K. (1992). Ecological investigations on lichen fields of the Central Namib: I. Distribution patterns and habitat conditions. Vegetatio, 98, 113-128.

Schwarz, G. (1978). Estimating the dimension of a model. The annals of statistics, 461-464.

Seitz, S., Nebel, M., Goebes, P., Käppeler, K., Schmidt, K., Shi, X., ... & Scholten, T. (2017). Bryophyte-dominated biological soil crusts mitigate soil erosion in an early successional Chinese subtropical forest. Biogeosciences, 14(24), 5775-5788.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. Statistics and computing, 14, 199-222.

Song, Y., & Hu, X. (2011). Estimation models of the aboveground biomass of Haloxylon ammodendron in different ecotypes of the Gurbantunggut Desert. Journal of Northwest Forestry University, 26(2), 31–37.

Song, H., & Huang, B. (2013). Spatiotemporal satellite image fusion through one-pair image learning. IEEE Transactions on Geoscience and Remote Sensing, 51(4), 1883–1896.

Song, H., Liu, Q., Wang, G., Hang, R., & Huang, B. (2018). Spatiotemporal satellite image fusion using deep convolutional neural networks. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(3), 821-829.

Spier, L., & Van Herk, C. M. (1997). Recent increase of Parmelia borreri in The Netherlands. The Lichenologist, 29(4), 390-393.

Stovall, M. S., Ganguli, A. C., Schallner, J. W., Faist, A. M., Yu, Q., & Pietrasiak, N. (2022). Can biological soil crusts be prominent landscape components in rangelands? A case study from New Mexico, USA. Geoderma, 410, 115658.

Su, Y., Chen, Y., Padilla, F., Zhang, Y., & Huang, G. (2020). The influence of biocrusts on the spatial pattern of soil bacterial communities: A case study at landscape and slope scales. Soil Biology and Biochemistry, 142, 107721.

Su, Y., Li, X., Zhao, X., Li, A., & Chen, Y. (2011). The Nitrogenase Activity of Biological Soil Crusts and Their Responses to Environmental Factors. Advances in earth science, 26(3), 332.

Sun, H., Ma, X., Liu, Y., Zhou, G., Ding, J., Lu, L., ... & Zhang, F. (2024). A new multi-angle method for estimating fractional biocrust coverage from Sentinel-2 data in arid areas. IEEE Transactions on Geoscience and Remote Sensing.

Tan, Z., Yue, P., Di, L., & Tang, J. (2018). Deriving high spatiotemporal remote sensing images using deep convolutional network. Remote Sensing, 10(7), 1066.

Trewartha, G. T. (1961). The earth's problem climates: Madison.

Van Herk, C. V., Aptroot, A., & Van Dobben, H. F. (2002). Long-term monitoring in the Netherlands suggests that lichens respond to global warming. The Lichenologist, 34(2), 141-154.

Vargas Castillo, R., Stanton, D., & Nelson, P. R. (2017). Aportes al conocimiento de la biota liquénica del oasis de neblina de Alto Patache, Desierto de Atacama. Revista de Geografía Norte Grande, (68), 49-64.

Wang, R. (2011). The study on factors affecting formation and development of biological soil crusts of loess region in north of Shaanxi province. Beijing Forestry University, Beijing, China.

Wang, Z., Hu, S., Li, H. (2018). Evapotranspiration characteristics of Haloxylon ammodendron community in interdune lowland at the southern edge of Gurbantunggut Desert[J]. Arid Land Geography 2018, 41(6): 1303-1309.

Wang, S., Ma, L., Yang, L., Ma, Y., Zhang, Y., Zhao, C., and Chen, N.: Advancing studies on global biocrust distribution, SOIL, 10, 763–778, https://doi.org/10.5194/soil-10-763-2024, 2024.

Wang, L., Qin S., Zhang Y., Wu B., Feng W., Liu J., ... & She W. (2017). Influence of biological soil crusts on soil moisture in Artemisia ordosica community in Mu Us Desert, northwestern China.

Wang, Z., Wu, B., Zhang, M., Zeng, H., Yang, L., Tian, F., ... & Wu, H. (2022). Indices enhance biological soil crust mapping in sandy and desert lands. Remote Sensing of Environment, 278, 113078.

Wang, B., Zha, T., Jia, X., Wu, B., Zhang, Y., and Qin, S. (2014). Soil moisture modifies the response of soil respiration to temperature in a desert shrub ecosystem, Biogeosciences, 11, 259–268.

Wang, X., Zhang, Y., & Jiang, J. (2007). Influence of grazing on the characteristics of longitudinal dune Surface in the southern part of Gurbantunggut Desert. Acta Geographica Sinica, 62(7): 698-706

Wang, X., Zhang, Y., Zhang, W., Han, Z. (2004). Wind tunnel experiment of biological crust effect on wind erodibility of sand surface in Gurbantunggut Desert, Xinjiang. Journal of Glaciology and Geocryology, 26(5), 632-638.

Wang, X., Zhao, C., Yang, R., & Li, J. (2015). Landscape pattern characteristics of desertification evolution in southern Gurbantunggut Desert. Journal Name, 38(6), 1213–1225.

Warren, S. D. (2001). Biological soil crusts and hydrology in North American deserts. In Biological soil crusts: structure, function, and management (pp. 327-337). Berlin, Heidelberg: Springer Berlin Heidelberg.

Weber, B., & Hill, J. (2016). Remote sensing of biological soil crusts at different scales. In Biological soil crusts: an organizing principle in drylands (pp. 215-234). Springer International Publishing.

Weber, B., Olehowski, C., Knerr, T., Hill, J., Deutschewitz, K., Wessels, D. C. J., ... & Büdel, B. (2008). A new approach for mapping of biological soil crusts in semidesert areas with hyperspectral imagery. Remote Sensing of Environment, 112(5), 2187-2201.

Wessels, DCJ & Van Vuuren, D. R. J. (1986). Landsat imagery-its possible use in mapping the distribution of major lichen communities in the Namib Desert, South West Africa. Madoqua, 1986(4), 369-373.

West, N. E. (1990). Structure and function of microphytic soil crusts in wildland ecosystems of arid to semi-arid regions. Advances in ecological research, 20, 179-223.

Williams, A. J., Buck, B. J., Soukup, D. A., & Merkler, D. J. (2013). Geomorphic controls on biological soil crust distribution: a conceptual model from the Mojave Desert (USA). Geomorphology, 195, 99-109.

Wu, M., Huang, W., Niu, Z., Wang, C., Li, W., & Yu, B. (2018). Validation of synthetic daily Landsat NDVI time series data generated by the improved spatial and temporal data fusion approach. Information Fusion, 40, 34-44.

Wu, M., Niu, Z., Zhao, Y., & Wang, C. (2012). A model for spatial and temporal data fusion. Journal of Infrared and Millimeter Waves, 31(1), 80–84.

Wu, Y., Gao, C., & Cheng, G. (2002). Ecological function of biological soil crusts. Chinese Journal of Ecology, (4), 41.

Wu. Y., & Yu, X. (2003). DistribLtion influencing factors and monitoring of biological soil crusts. Acta Genetica Sinica, 22(3), 38-42.

Xiong, H., Duan, J., Wang, Y., & Zhang, X. (2011). Effects of biological soil crust on water infiltration and redistribution in the Mu Us Sandland, Inner Mongolia, Northern China. Research of Soil and Water Conservation, 18(4), 82-87.

Xu, Z., Han, Y., Tam, C. Y., Yang, Z. L., & Fu, C. (2021). Bias-corrected CMIP6 global dataset for dynamical downscaling of the historical and future climate (1979–2100). Scientific Data, 8(1), 293.

Xu, Y., Liu, B., Sun, A., Wang, K., Li, D., & Zhao, H. (2023). Research progress of Holocene environmental evolution in the Gurbantunggut Desert and its surrounding areas. Arid Land Geography, 46(4), 550-562.

Yair, A. (2003). Effects of biological soil crusts on water redistribution in the Negev Desert, Israel: a case study in longitudinal dunes. Biological soil crusts: structure, function, and management, 303-314.

Yu, L., Hu, Z., Zhang, F., & Yang, K. (2022). Unmanned aerial vehicle image biological soil crust recognition based on UNet++. International Journal of Remote Sensing, 43(7), 2660-2676.

Zaady, E., Eldridge, D. J., & Bowker, M. A. (2016). Effects of local-scale disturbance on biocrusts. In Biological soil crusts: An organizing principle in drylands (pp. 429–451). Springer, Cham.

Zaady, E., Karnieli, A., & Shachak, M. (2007). Applying a field spectroscopy technique for assessing successional trends of biological soil crusts in a semi-arid environment. Journal of Arid Environments, 70(3), 463-477.

Zedda, L., Gröngröft, A., Schultz, M., Petersen, A., Mills, A., & Rambold, G. (2011). Distribution patterns of soil lichens across the principal biomes of southern Africa. Journal of Arid Environments, 75(2), 215-220.

Zhang, H. (2013). Distribution and Dynamic of Soil Moisture in Artemisia Ordosica Community in Mu Us Sandy Land. Chinese Academy of Forestry, Beijing, China.

Zhang, J. (2014). Distribution characteristics of biological soil crust for Artemisia ordosica community in Mu Us Sandy Land. Bulletin of Soil and Water Conservation, 34(3), 227–230.

Zhang, P. (2015). Spatial distribution characteristics of biological soil crusts and its impact factors and modeling on slope scale in small watershed of northern Shaanxi. Northwest A&F University, Xianyang, Shaanxi, China.

Zhang, L., & Chen, C. (2002). On the general characteristics of plant diversity of Gurbantunggut sandy desert. Chinese Science Bulletin, 22(11), 1923–1932.

Zhang, Y., Chen, J., Wang, X.., Pan, H.., Gu, Z., & Pan, B. (2005). The distribution patterns of biological soil crust in Gurbantunggut Desert. Acta Geographica Sinica, 60(1), 53-60.

Zhang, Y., Chen, J., Wang, L., Wang, X., & Gu, Z. H. (2007). The spatial distribution patterns of biological soil crusts in the Gurbantunggut Desert, Northern Xinjiang, China. Journal of Arid Environments, 68(4), 599-610.

Zhang, Z., Chen, Y., Xu, B., Huang, L., Tan, H., & Dong, X. (2015). Topographic differentiations of biological soil crusts and hydraulic properties in fixed sand dunes, Tengger Desert. Journal of Arid Land, 7, 205-215.

Zhang, Z., Li, X., Nowak, R., Wu, P., Gao, Y., Zhao, Y., ... & Jia, R. (2013). Effect of sand-stabilizing shrubs on soil respiration in a temperate desert. Plant and Soil, 367, 449-463.

Zhang, R., Liang, Y., Gao, Y., & Wang, Y. (2024). Distribution characteristics of soil crusts under two kinds of aerial seeding plants in the northeastern margin of Tengger Desert. Frontiers in Environmental Science, 12, 1423596.

Zhang, J., Liu, G., Xu, M., Xue, S., & Liu, J. (2008). Influencing factors of biological crust degradation in Hilly-gullied Region of Loess Plateau in China. Science of Soil and Water Conservation, 6(6), 14-20.

Zhang, Y., Pan, H., & Pan B. (2004). Distribution Characteristics of Biological Crust on Sand Dune Surface in Gurbantunggut Desert, Xinjiang. Journal of Water and Soil Conservation, 18(4), 61-66.

Zhang, J., Wang, X., Liu, L. & Xiao, H. (2000). Study on soil microbiotic crust and its influences on sand-fixing vegetation in arid desert region. Journal of Integrative Plant Biology, 42(9), 965.

Zhang, J., Wu, B., Jia, Z., Cui, L., Li, H., Zhang, S., ... & Gao, D. (2010). Pattern of biological soil crust and its driving factors under Artemisia ordosica in Mu Us Sandy land. Forest Research, 23(6), 866.

Zhang, J., Zhang, Y. M., Downing, A., Cheng, J. H., Zhou, X. B., & Zhang, B. C. (2009). The influence of biological soil crusts on dew deposition in Gurbantunggut Desert, Northwestern China. Journal of Hydrology, 379(3-4), 220-228.

Zhang, D., Zhao, Y., Qi, H., Shan, L., Chen, G., & Ning, T. (2024). Effects of Micro-Topography and Vegetation on Soil Moisture on Fixed Sand Dunes in Tengger Desert, China. Plants, 13(11), 1571.

Zhao, H., Guo, Y., Zhou, R., & Zhao, X. (2009). Effects of vegetation cover on physical and chemical properties of bio-crust and under-layer soil in Horqin Sand Land. Chinese Journal of Applied Ecology, 20(7), 1657-1663.

Zhao, J., An, C., Huang, Y., Morrill, C., & Chen, F. (2017). Contrasting early Holocene temperature variations between monsoonal East Asia and westerly dominated Central Asia. Quaternary Science Reviews, 178, 14-23.

Zhao, Y., Qin, N., Weber, B., Wu, D., & Zhao, W. (2014). Response of biological soil crusts to raindrop erosivity and underlying influences in the hilly Loess Plateau region, China. Biodiversity and Conservation, 23(7), 1669–1686.

Zhou, H., Gao, Y., Jia, X., Wang, M., Ding, J., Cheng, L., ... & Wu, B. (2020). Network analysis reveals the strengthening of microbial interaction in biological soil crust development in the Mu Us Sandy Land, northwestern China. Soil Biology and Biochemistry, 144, 107782.

Zhou, X., Ke, T., Li, S., Deng, S., An, X., Ma, X., ... & Chen, L. (2020). Induced biological soil crusts and soil properties varied between slope aspect, slope gradient and plant canopy in the Hobq desert of China. Catena, 190, 104559.

Zhou, H., Li, Y., Tang, Y., Zhou, B., & Xu, H. (2009). The characteristics of the snow-cover and snowmelt water storage in Gurbantunggut Desert.

Zhou, Y., Li, Y., Li, W., Li, F., & Xin, Q. (2022). Ecological responses to climate change and human activities in the arid and semi-arid regions of Xinjiang in China. Remote Sensing, 14(16), 3911.

Zhu, X., Chen, J., Gao, F., Chen, X., & Masek, J. G. (2010). An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. Remote Sensing of Environment, 114(11), 2610-2623.

Zhu, X., Helmer, E. H., Gao, F., Liu, D., Chen, J., & Lefsky, M. A. (2016). A flexible spatiotemporal method for fusing satellite images with different resolutions. Remote Sensing of Environment, 172, 165-177.