scientific reports

Check for updates

OPEN

Spatio-temporal analysis of litterfall load in the lower reaches of Qarqan and Tarim rivers using BP neural networks

Junyu Xu[1,2,3], Anwar Eziz[2,3]✉, Alishir Kurban[2,3], Ümüt Halik[1]✉, Zhiwen Shi[2], Saif Ullah[2,3], Gift Donu Fidelis[2,3], Yingdong Ma[1], Ziwargul Kibir[1,2], Toqeer Ahmed[2,4], Tim Van de Voorde[5], Adil Hujashim[6] & Hossein Azadi[2,3,5]

Litterfall load is crucial in maintaining ecosystem health, controlling wildfires, and estimating carbon stock in arid regions. However, there is a lack of spatiotemporal analysis of litterfall in arid riparian forests. This study aims to estimate Litterfall load using a BP neural network based on vegetation indices from Landsat 5 and 8 satellite images, litterfall inventory data, slope, and distance to major river tributaries. It also aims to analyze the spatiotemporal distribution pattern of litter in the research area by estimating and analyzing the spatiotemporal pattern of litterfall along the desert riparian forests of the lower Qarqan and Tarim Rivers from 2001 to 2021. The results show that the initiation of the ecological water transfer project has facilitated the decomposition of litterfall, leading to an initial decline. Subsequently, the vegetation gradually recovered, leading to an increase in leaf litter input. Since 2001, litterfall initially decreased until reaching its lowest value of 4.39×10^9 kg in 2005, followed by a subsequent increase, reaching its highest value of 12.5×10^9 kg in 2021. The study concludes that ecological water conveyance promotes both the decomposition and increase of litterfall. Initially, it accelerates litterfall decomposition, while later stages foster an increase in Litterfall load. Meanwhile, due to the ecological water transfer project and the higher vegetation cover along the Tarim River compared to the Qarqan River, the Tarim River basin experiences higher average Litterfall load and variation.

Keywords Litterfall load, BP neural network, Arid region, Desert riparian forest

Litterfall, defined as the organic material originating from the Canopy foliage and deposited onto the soil surface, is a pivotal link within the production-decomposition cycle of organic matter. This biomass transfer plays a crucial role in nutrient cycling, particularly in carbon cycling and overall ecosystem health, by acting as an intermediary nutrient carrier between plants and soil[1,2]. Moreover, as an integral component of the soil surface layer, litterfall fosters the soil microbiome by providing a nutrient-rich food source, thereby enhancing soil organic matter content, nutrient availability, and moisture level. Therefore, it enhances optimal conditions for vegetation growth[3–5]. Due to its significance, litterfall serves as a critical parameter for modeling ecosystem productivity and assessing ecosystem health, particularly in arid regions. Understanding the responses of forest litterfall to global changes across various ecosystem scales is essential in both ecological research and forestry practice. Litterfall decomposition by soil microorganisms releases CO_2, adding significantly to soil carbon emissions[6]. This annual variation in litterfall production, therefore, becomes a critical factor influencing the global carbon cycle and climate dynamics. Additionally, the annual litterfall production significantly impacts fuel accumulation in forests, particularly in arid regions. Consequently, litterfall serves as a proxy index for evaluating the wildfire risk in these regions. Given the critical role of litterfall load (LTFL hereafter), it is

[1]College of Ecology and Environment, Xinjiang University, Urumqi 830046, Xinjiang, China. [2]State Key Laboratory of Desert and Oasis Ecology, Key Laboratory of Ecological Safety and Sustainable Development in Arid Lands, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, 818 South Beijing Road, Urumqi 830011, China. [3]Sino-Belgian Joint Laboratory for Geo-Information, Urumqi 830011, China. [4]Centre for Climate Research and Development (CCRD), COMSATS University Islamabad, Tarlai Kalan, Park Road, Islamabad 45550, Pakistan. [5]Department of Geography, Ghent University, Krijgslaan 281 S8, 9000 Ghent, Belgium. [6]Forestry and Grassland Bureau of Ruoqiang County, Ruoqiang 841800, China. ✉email: anwareziz@ms.xjb.ac.cn; halik@xju.edu.cn

imperative to comprehensively understand its temporal and spatial patterns to manage ecosystems and mitigate associated risks effectively.

Given the sensitivity of arid regions to climate change and human activities[7], it is of paramount importance to comprehend spatiotemporal variations of LTFL. This understanding is key to assessing the influence of human activities and climate change on arid land ecosystems. Moreover, it provides insights into the potential contributions to atmospheric greenhouse gas fluctuations and the restoration of damaged ecosystems. Desert riparian forests, an important vegetation type in arid regions, are the main contributors to LTFL[8]. By studying litterfall in arid regions, in-depth exploration of the carbon sequestration capacity and conversion efficiency of regional ecosystems can be achieved while also enhancing the accuracy of forest fire prediction.

The Tarim and Qarqan rivers are major river tributaries in the Tarim basin, around which there is important riparian vegetation that serves as a windbreak zone and harbors biodiversity in arid regions. Intensive human activities have significantly altered vegetation cover and structure along the lower reaches of the Tarim and Qarqan Rivers, putting the ecosystem and local socioeconomic activities at risk. An Ecological Water Conveyance (EWC) project has been implemented since 2000 to avoid continuous deterioration of the situation. Since then, vegetation has shown partial recovery with water delivery to the lower Tarim River and its terminal lakes. Given its typical arid conditions, research on litterfall variation in this area can provide valuable insights into ecosystem restoration efforts and the mitigation of fire hazards in similar arid zones. Besides, it can provide helpful information about improving the efficacy of the ecosystem restoration project implemented in the area.

At present, commonly used indicators for estimating LTFL include breast height diameter, tree height, forest age, density, canopy closure, etc[8]. A substantial amount of research has been dedicated to the study of LTFL and its decomposition[4,9]. This includes the analysis of influencing factors[10] and the construction of the LTFL prediction models[11–13]. For instance, Yuanlong et al.[14] used backward stepwise linear regression analysis to establish a model for estimating combustible material load based on stand factors and various surface LTFLs in Dahinggan Mountains spruce forests. Haiqing et al.[15] employed ridge estimation to predict LTFL in Dahinggan Mountains coniferous forests using TM remote sensing images and stand factor data. Later, Changhong et al.[11] improved the accuracy of LTFL prediction models for Dahinggan Mountains larch forests using the Backpropagation (BP) network method. Wu et al.[16] determined the main influencing factors of surface dead combustible material in the Gannan region by constructing a structural equation model between surface combustible material load and various influencing factors. Yang et al.[17] estimated the total LTFL in the Tarim River Basin by calculating the average LTFL per plant. Although the studies have deepened our understanding of the spatiotemporal variation of LTFL across the global and national scale, due to the lack of ground truth data of LTFL and differences in sampling scheme, there are significant variations among the results, especially in arid lands. For instance, according to Ngangyo-Heya et al.[18], the average LTFL in the arid region of Mexico is 1345 $kg.ha^{-1}$, while a study conducted by Verma et al.[19] showed litter levels lower than 20% of the former. In addition, current research mainly relies on on-site data as influencing factors for estimation. Due to the difficulty in obtaining some of these data, it is difficult to sequentially estimate the LTFL of the entire region, and the spatiotemporal variation patterns are also relatively vague. Most importantly, compared to humid areas with abundant vegetation, there is currently limited research on the litter load in arid areas, especially in forests along the Tarim and Qarqan rivers, which necessitates the comprehensive study of LTFL spatiotemporal variation in this region as a representative of arid riparian forest LTFL.

The Backpropagation (BP) neural network is a popular machine learning model, known for simplicity and adaptability in regression and classification tasks. Its intuitive backpropagation algorithm allows for efficient weight adjustment, enabling automatic learning and feature extraction from complex data without human intervention. The BP neural network is robust against neuron failures due to its distributed architecture, maintaining overall stability. Additionally, it is less sensitive to hyperparameters, reducing the risk of overfitting compared to other algorithms[20]. Given its effectiveness in modeling non-linear relationships and its adaptability to diverse data types, the BP neural network is particularly suitable for ecological and environmental studies[58]. Its universal approximation capability allows it to accurately model complex, non-linear dynamics inherent in these fields. Therefore, this study employs the BP neural network to estimate litterfall load and analyze spatiotemporal variations using field observations and remote sensing data from the lower reaches of the Tarim and Qarqan River catchments.

The aim of this study is to construct a BP neural network model using vegetation indices, litter inventory data, terrain slope, and distance from major river tributaries from Landsat 5 and 8 satellite imagery to estimate litter load in riparian forests in arid areas. The spatiotemporal distribution patterns of litter in riparian forests downstream of the Tarim and Qarqan rivers from 2001 to 2021 will be analyzed to reveal the temporal and spatial variation patterns of litter load. We assume that the implementation of EWC projects will affect the decomposition and accumulation process of litter, initially promoting litter decomposition and leading to a decrease in load. Subsequently, as vegetation recovers, litter input will gradually increase. In addition, it is expected that due to the high vegetation coverage, the average litter load and volatility in the Tarim River Basin will be higher than those in the Qarqan River Basin.

Specifically, the study tried to answer (1) what is the total amount of LTFL in the study area? (2) what are the spatial patterns of LTFL and its dynamics in the study area? Finally, the study aims to understand the change in LTFL since the EWC and to provide LTFL data and theoretical support for ecological restoration fire risk prevention in arid regions.

Materials and method

Study area

The study area is situated on the southeastern edge of the Taklamakan Desert in Xinjiang, China. It includes the lower reaches of the Tarim River and parts of the Qarqan River watershed (Fig. 1). The region has a warm temperate

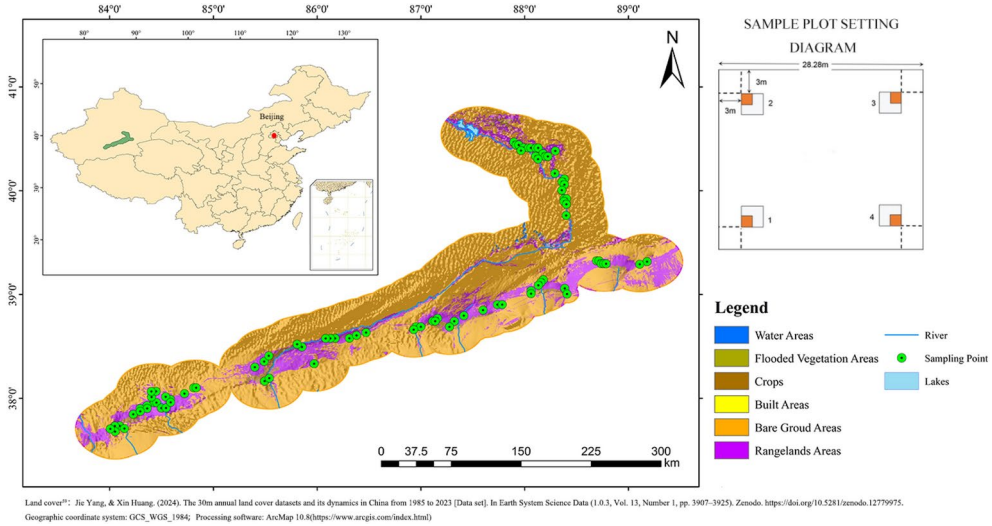

Fig. 1. Schematic diagram of study area showing sampling points.

continental desert climate with an average annual temperature of 10.8 ± 0.7 °C, around 3000 h of sunshine per year, a solar radiation of 1740 kW·h/m^2, evapotranspiration of 2750 mm, and low average annual precipitation ranging from 36 to 61 mm. Total water resources in the basin amount to 4.34×10^{10} m^3, predominantly sourced from glacier meltwater for both rivers[21–23]. The riparian forest primarily comprises *Populus euphratica*, *Elaeagnus angustifolia*, *Tamarix ramsissima*, *Lycium ruthenicum*, *Halimodendron halodendron*, *Phragmites australis*, *Alhagi sparsifolia*, *Apocynum Veneto*, *Karelinia caspia*, and *Glycyrrhiza inflata*[24].

Field data

To gather ground truth data for LTFL[59], a comprehensive field investigation was carried out along the lower Tarim and Qarqan rivers from January to February 2022. Adhering to the protocol specified in the 'Implementation Rules for Forest Combustible Standard Ground Survey in Xinjiang Uygur Autonomous Region', a total of 75 sample plots, each measuring 28.28 × 28.28 m, were designated. Within each sample plot, four sampling quadrats for herbaceous layer were positioned diagonally (Fig. 1). Based on existing forest resource survey data and high-resolution remote sensing images, analyze the characteristics and distribution of combustible materials on the forest surface in the study area. Vegetation zones, forest origin, age group, canopy closure, coverage, and dominant tree species (groups) were considered when selecting the sampling plot locations. During the investigation, dead branches, leaves, fruits, and dead grass were collected, and their fresh weight was measured. Subsequently, the collected samples were oven-dried at 65 °C for 24 h in the laboratory until a constant weight was attained. The dry weight of the samples was then determined, allowing for the calculation of the dry-to-fresh weight ratio (R), which was used to calculate the total dry weight (*TDW*) of LTFL in each sampling plot together with the total fresh weight (*TFW*) using the following formula:

$$TDW = R \times TFW \tag{1}$$

In addition, LTFL per hectare was calculated by dividing the total dry mass of the sampling plots by the total area of the sampling plots. This value was then converted into tons per hectare.

Remote sensing data

Given that most vegetation is located within a 20 km buffer zone surrounding the riverbed, we established a 30 km buffer zone using ArcMap 11.1 to delineate the boundaries of the study area. From 2001 to 2013, Landsat 5 data (Landsat 5 Surface Reflectance Tier 1, 'LANDSAT/LT05/C01 /T1 SR') were used to calculate the vegetation indices, while from 2014 to 2021 Landsat 8 data (Landsat 8 Surface Reflectance Tier 1, 'LANDSAT/LC08/C01/T1_SR') were utilized. As the sampling time for this study was concentrated between January and February 2022, the range of remote sensing images we selected was also narrowed down to this period to ensure that the information reflected in the remote sensing images better matches the research content to provide better training models and analysis results. Using Google Earth Engine (GEE), we obtained the data that has been completed with radiometric measurements and geometric corrections after masking the land satellite data with the boundary of the research area. Subsequently, the data were resampled to 28.28 × 28.28 m spatial resolution using bicubic interpolation to match the size of the sampling plots. Thirteen vegetation indices (Table 1)[25–37], closely related to vegetation biomass and LTFL, were calculated in GEE. These specific vegetation indices characterize plant physiological status, health status, soil moisture, and other environmental factors by analyzing surface reflectance spectral features. Each index has its unique sensitivity and applicability, and the purpose of selecting these vegetation indices is to provide more accurate and multidimensional remote

Abbreviation	Vegetation index	Formula	Purpose
RVI	Ratio vegetation index	$RVI = \frac{R}{NIR}$	Assess vegetation coverage and health status
TVDI	Temperature vegetation dryness index	$TVDI = \frac{T_c - T_{min}}{T_{max} - T_{min}}$	Assess the relationship between surface vegetation and drought
EVI	Enhanced vegetation index	$EVI = 2.5\,(NIR - R)$	Consider atmospheric and soil effects, provide more accurate vegetation information
GEMI	Global environment monitoring index	$GEMI = \frac{\mu(1-0.25\mu)-(R-0.125)}{R}$	Used for atmospheric correction to improve vegetation detection accuracy
MSAVI	Modified soil-adjusted vegetation index	$MSAVI = \frac{1}{2}(2(NIR+1) - \sqrt{(2NIR+1)^2 - 8(NIR-R)})$	Reduce soil impact by considering soil reflection, more accurately assess vegetation
ARVI	Atmospherically resistant vegetation index	$ARVI = \frac{NIR-RB}{NIR+RB}$	Reduce atmospheric impact, better assess vegetation status
EXG	Excess green index	$EXG = 2G - R - B$	Evaluate vegetation condition, sensitive to vegetation changes
GVI	Green vegetation index	$GVI = -0.283G - 0.66R + 0.577NIR + 0.388MIR$	Estimate the health and coverage of vegetation
OSAVI	Optimized soil-adjusted vegetation index	$OSAVI = \frac{1.16\times(NIR-RED)}{NIR+RED+0.16}$	Consider vegetation sparsity based on soil correction, suitable for low-coverage vegetation areas
PRI	Photochemical reflectance index	$PRI = \frac{R_{ref}-R_{531}}{R_{ref}+R_{531}}$	Evaluate the efficiency of plant chlorophyll photosynthesis reflect vegetation health
RI	Redness index	$RI = \frac{R-G}{R+G}$	Assess the red characteristics of vegetation
VEG	Vegetation index	$VEG = \frac{G}{RaB\times(1-a)}$, $\alpha = 0.667$	Evaluate vegetation coverage and health
DFI	Dryness fraction index	$DFI = 100 \times \left(1 - \frac{SWIR2}{SWIR1}\right) \times \left(\frac{R}{NIR}\right)$	Detect the degree of surface drought

Table 1. Selected vegetation index information.

sensing data support for the comprehensive assessment of litter load in riparian forests in arid areas. These indices were chosen to comprehensively evaluate the litter load of riparian forests in arid areas, as they cover multiple vegetation physiological parameters and are adapted to the conditions of low vegetation cover and soil background disturbance in arid regions. In addition, these indices can improve the accuracy of remote sensing data through anti atmospheric interference indices such as ARVI and EVI. Additionally, the MODIS land surface temperature (LST) product (MOD11_L2)[38] was used to extract the mean night time and day time annual LST in GEE.

Note: T_c, T_{min} and T_{max} are the maximum, minimum, and average values of temperature; R, G, B, NIR,MIR, SWIR are red, green, blue, near-infrared, middle infrared and shortwave infrared bands respectively; μ is a constant; MSS, Rref is the senser; R_{531} is the band ratio index.

Elevation and slope could have a potential impact on vegetation composition and distribution, leading to variations in litterfall distribution[39]. Therefore, the 12.5m spatial resolution SRTM Digital Elevation Model (DEM) from the NASA Earth Science Data website (https://nasadaacs.eos.nasa.gov/) was used to calculate the slope in ArcMap 11.2[40]. Moreover, the distance to a water body serves as a significant proxy index for the groundwater gradient, influencing the vegetation cover and its distribution[39]. To account for this, we generated a raster data set representing the distribution of surface water bodies. This data set was derived from the global surface water area data product[41]. We then calculated the distance to the nearest water body using the Euclidean distance method in ArcMap 11.2. Prior to model construction, data for the independent variables were extracted for each sampling plot in ArcMap 11.2 using the shapefile of the sampling points. Given that a single sampling plot may not align with an individual pixel within the raster data of independent variables, the values of the eight surrounding pixels were averaged to obtain a more accurate value for each variable at each sampling point, thereby reducing uncertainty. Lastly, before implementing the BP neural network for litterfall estimation, all data were resampled to a 1km spatial resolution using bicubic interpolation in ArcMap 11.2. This step was taken to ensure alignment in the spatial resolution of predictor variables.

Methodology

Principal component analysis (PCA)

The common biological indicators used for estimating LTFL (landscape tree functional level) include diameter at breast height, tree height, forest age, density, and canopy closure. However, due to limited data availability on these features, their application has been restricted to the analysis of LTFL at the sampling plot scale, making it challenging, if not possible, to extend to a larger area. In order to cover the features mentioned above in our study, 13 vegetation indices that are widely used in arid region research and are sensitive to changes in vegetation cover, water stress, and other environmental factors were included based on vegetation coverage, types, and phenology in the Qerchen and Tarim river riparian zones along with the relevant studies on vegetation indices in arid areas. These vegetation indices, along with the daily average temperature, nighttime average temperature of the sampling month, distance from the river, and slope, were used as predictor variables in constructing the LTFL prediction model. These vegetation indices and temperature data are derived from satellite imagery, which provides data across various wavelengths and combinations of spectral bands. However, due to the potential overlap among these variables, multicollinearity could adversely affect the model's performance. To address this issue, we applied principal component analysis (PCA) to reduce multicollinearity and enhance the model's robustness. Additionally, by extracting the principal features from the data, PCA helps eliminate noise and highlight the most critical information, which is crucial for improving the efficiency, stability, and accuracy of neural network training by reducing the convergence time. The dimensionality-reduced data not only reduces the risk of overfitting but also enhances the model's generalization ability on new data, making the BP neural network more efficient and reliable in handling complex nonlinear problems. Vegetation indices were normalized using the standard min–max technique to a range of 0 to 1 before PCA analysis. The first three principal components (Vegetation PC 1, 2, 3) were selected based on an 85% cumulative percentage of Eigenvalue. The scores of each vegetation index in these principal components are presented in Table 2. These three PCAs were used as predictor variables during the model-building process.

Building and evaluation of BP neural network model

Model selection for estimating the LTFL using vegetation indices typically includes multivariable linear regression and various machine learning models. While multivariable linear regression is favored for its interpretability and broad applicability, it is limited in capturing nonlinear relationships. In contrast, machine learning models excel at modeling complex, nonlinear relationships, making them more suitable for current research. Given the complexity of vegetation growth, distribution, and response to environmental factors in arid environments—including nonlinear responses to moisture and temperature, spatial heterogeneity of vegetation cover, and diverse soil conditions—the relationship between LTFL and vegetation indices in these regions is distinctly nonlinear. Therefore, machine learning approaches are more appropriate for modeling this relationship. To test this assumption, we first applied multivariable linear regression, which yielded poor model performance ($R^2 = 0.031$, $p = 0.262$), suggesting a nonlinear relationship between LTFL and the influencing factors. Furthermore, normality tests on the model residuals, including the Omnibus (10.989, $p = 0.004$) and Jarque–Bera (11.065, $p = 0.004$) tests, indicated that the residuals were not normally distributed. These results support the need for models, such as neural networks, that are better suited to capturing nonlinear patterns in the data.

The neural network comprises interconnected neurons that collaborate to solve complex nonlinear problems. In addition, it is adaptable to various data types and can effectively handle large datasets. Additionally, its hierarchical structure allows for efficient learning and representation of intricate patterns in the data[42].

Vegetation index	Vegetation PC 1	Vegetation PC 2	Vegetation PC 3
RVI	0.955	−0.086	0.223
TVDI	−0.337	0.721	0.524
EVI	0.464	0.382	−0.484
GEMI	0.751	−0.619	−0.147
MSAVI	0.961	−0.09	0.219
ARVI	0.897	0.381	0.052
EXG	−0.739	−0.566	−0.029
GVI	−0.691	−0.627	−0.181
OSAVI	0.959	−0.088	0.223
PRI	0.09	0.893	−0.19
RI	−0.194	0.718	−0.521
VEG	−0.792	0.37	0.443
DFI	0.879	−0.066	−0.058

Table 2. Principal component analysis (PCA) results of vegetation indices: this table presents the component loadings for each vegetation index across the first three principal components. The values indicate the correlation between each index and the principal components, highlighting which indices contribute most significantly to each component. This analysis aids in understanding the underlying patterns in our vegetation data and informed our selection of indices for the LTFL estimation model.

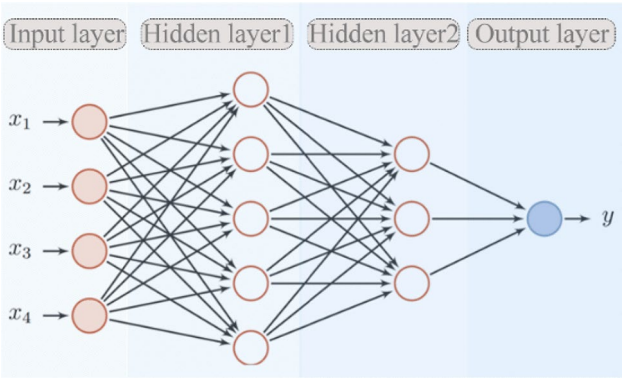

Fig. 2. BP neural network structure schematic diagram.

Among neural network algorithms[43], the Backpropagation (BP) neural network is notable for its ability to fine-tune connection weights through error backpropagation, achieving high precision and accuracy without significant overfitting[44]. Its simplicity, interpretability, and versatility make it suitable for both regression and classification tasks, while its architecture, with multiple neurons, ensures robustness against partial neuron failure. These attributes allow the BP neural network to effectively learn from complex datasets and automatically extract features without human intervention[43]. Given these advantages and the nonlinear relationships between vegetation litterfall and its predictor variables, this study employs the BP neural network to develop an LTFL estimation model for the study area.

The BP neural network is a machine learning algorithm that operates as follows: the first layer acts as the input layer and transmits signals to neurons in the subsequent layer, known as the hidden layer. Multiple hidden layers may exist where output of the previous layer will feed the next layers in a sequence. The last layer is the output layer, which sends stimulated signals to the external environment after traversing multiple layers[42]. (Fig. 2).

Mini batch stochastic gradient descent (Mini Batch SGD) provides higher computational efficiency by training with small batches of data in each iteration, making it particularly suitable for large-scale datasets. Compared with the full batch method, Mini batch SGD reduces memory requirements and avoids resource consumption issues. Meanwhile, by frequently updating weights, it can approach the optimal solution faster and improve training speed. Its main advantages include better gradient estimation and smoother gradient updates, avoiding the gradient oscillation problem in stochastic gradient descent (SGD). Mini Batch SGD combines the stability of batch gradient descent with the efficiency of SGD, enabling fast iterative updates while maintaining convergence stability, reducing the risk of getting stuck in local optima and oscillations. In contrast, other commonly used methods such as full batch gradient descent have high computational costs, large memory requirements, and slow training speeds; Although stochastic gradient descent is fast, the gradient is unstable and

prone to oscillation; The momentum method, although helpful in accelerating convergence, relies on complex hyperparameter tuning. Based on this, Mini batch SGD technology was selected for this study.

In estimating LTFL in the study area, the three principal components serve as input variables, while LTFL at the corresponding pixel locations is the output variable for modeling. After the input variables were identified, seven input and one output layer neuron were determined using Mini-batch stochastic gradient descents (Mini-batch SGD)[45] as the network training method. The sigmoid function is used to activate the input layer neuron. In addition, the network's initial weights were randomly initialized to values between − 0.5 and 0.5. The learning rate was determined using the momentum method[46]. Determining the number of neurons in the input and output layers is critical, with the number of hidden layer neurons affecting the model's nonlinear ability and generalization error[20]. While increasing the number of hidden layer neurons can enhance output accuracy, it also increases model complexity and training time[20]. The number of hidden layers was determined according to the following formula through the Mini-Batch SGD network training function:

$$\begin{cases} \sum_{i=0}^{n} C_M^i > K, i \le M \\ \sum_{i=0}^{n} C_M^i = 0, i > M \end{cases} \tag{2}$$

$$M = \sqrt{n+m} + a, 0 \le a \le 10 \tag{3}$$

$$M = log_2 n \tag{4}$$

where *k, M, n, m,* and *a* are the set size of input data, the number of hidden layer and input layer nodes, the number of network output layer nodes, and a constant value, respectively.

The regression model's performance was evaluated with the coefficient of determination (R^2) and root mean square error (RMSE). To achieve this, 30% of the data were randomly selected as a test set using the bootstrap sampling technique, while the remaining data was used as a train set. The training process was configured to stop when the number of rounds with no improvement in error reaches 100. It was found that the model performed the best when the number of hidden layers and nodes were set to 3 and 25, respectively. The transfer function from the input to the output layer was set as the identity function, while default values were used for the other parameters.

Finally, prediction results were classified into low (0–85 g/m^2), medium (85–184 g/m^2), and high (184 g/m^2 ~) LTFL sections using natural breaks methods to analyze the change in the spatial distribution of the LTFL.

Results

Model evaluation results

The evaluation results showed that during the training phase, the BP neural network achieved the lowest error rate and a higher R^2 of 0.96 with three hidden layers and 25 neurons in each hidden layer (Table 3). In addition, R^2 was 0.86, calculated based on the test data (Fig. 3). This relatively small difference in R^2 based on the test and train sets indicates a relatively robust model with little overfit in the training period.

Model assumption testing is critical in statistical modeling and machine learning, including BP neural networks, to ensure reliable, interpretable, and generalizable results. For BP neural networks, key assumptions to validate include the independence of observations, absence of multicollinearity among input features, and normality of residuals for efficient backpropagation and convergence. Common checks involve testing for linearity, homoscedasticity of errors, and absence of outliers, using methods like residual plots, variance inflation factor (VIF) analysis, and statistical tests. Additionally, verifying the network's capacity to learn patterns without overfitting is essential.

After conducting assumption testing for the BP neural network model, it was confirmed that the relationship between LTFL and predictor variables might be nonlinear, as evident from residual vs. predictor variables (Fig. 4). This necessitates the use of algorithms that can better capture the nonlinear complex associations, showcasing the model's adaptability. Additionally, multicollinearity is minimal among predictor variables with VIF values below the threshold of 5 (Table 4). The model's assumptions of homoscedasticity and normality are largely met

Hidden layer	Hidden layer neurons	R^2	RMSE	Mean absolute percentage error
3	25	0.9552	0.2644	4.6771
2	25	0.949	0.2824	4.6744
3	30	0.9429	0.2987	5.668
4	20	0.9308	0.3288	6.3055
4	25	0.9289	0.3334	6.1528
2	30	0.9252	0.3419	6.2438
4	30	0.9239	0.3442	6.1627
2	20	0.921	0.3513	6.055
3	20	0.9112	0.3726	6.9981

Table 3. Training results for different hidden layers and neurons.

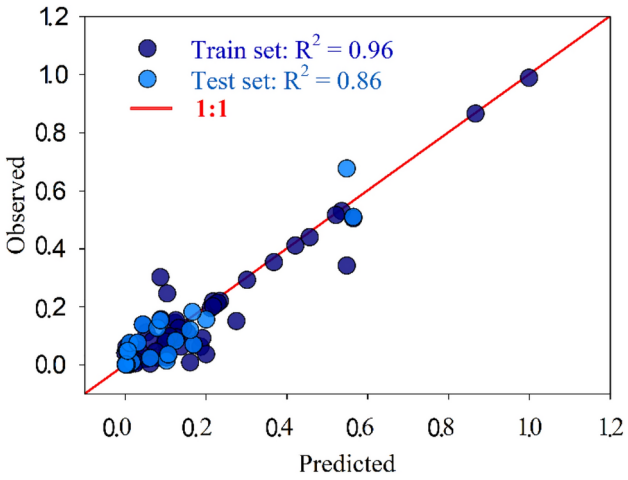

Fig. 3. BP neural network prediction results vs observation after standardized from 0 to 1.

(Fig. 4). No significant outliers were identified, suggesting that the model assumptions are reasonably met for reliable predictions. Furthermore, the 95% confidence interval calculated through 4999 times of bootstrapping technique showed relatively tight confidence intervals containing observed values, indicating the relatively small uncertainty of the prediction results of the model. Finally, upon conducting a sensitivity analysis, This research discovered that the distance to the major river channel had the most significant impact on the model's performance, while the other variables had minimal influence. This was demonstrated by Olden's methods[47] for sensitivity analysis, highlighting the robustness of our neural network model for estimating LTFL (Fig. 4).

Spatial pattern of litterfall

To analyze the overall change in LTFL in the study area over the past two decades and its association with the amount of water delivered to the study area through an ecological restoration project in the lower Tarim River zone, the total amount of LTFL was calculated based on estimated results using a trained BP neural network model. The changing trend of LTFL was then compared with the change in the amount of runoff water through an EWC, as shown in Fig. 5.

The findings reveal a downward trend in total litter until 2006, reaching a minimum of 4.7×10^9 kg, followed by a gradual rise, especially after 2018, culminating at 12.13×10^9 kg in 2021. Notably, fluctuations in water supplied through ecological conveyance closely mirrored variations in litter fallout throughout the study period (Fig. 5). The Pearson correlation coefficient, measuring the relationship between this water delivery and litter variation, was 0.664 ($p=0.001$), indicating a moderate positive correlation. Post-2005, the coefficient rose to 0.738 ($p=0.0007$), illustrating a strengthened correlation.

The spatiotemporal variation of LTFL was analyzed by creating annual distribution maps from 2001 to 2021 using predictions from the BP neural network model, as shown in Fig. 6. The analysis revealed consistent changes in both quantity and spatial distribution over the years. Therefore, the years 2001, 2005, 2010, 2015, and 2021 were selected as representatives for the analysis of the variation of LTFL throughout the study period.

According to the results, the LTFL values for the years 2001, 2005, 2010, 2015, and 2021 are 7.01×10^9 kg, 4.40×10^9 kg, 6.19×10^9 kg, 7.09×10^9 kg, and 12.5×10^9 kg. In addition, the mean value per square meter is 173.41 ± 92.06 g (for 2001), 108.79 ± 90.03 g (for 2005), 152.75 ± 83.86 g (for 2010), 175.25 ± 182.86 g (for 2015), and 310.04 ± 347.36 g (for 2021), respectively. Spatially, High LTFL areas are mainly in the central and northeastern regions, while lower values are found near the desert. The areas with high LTFL shift from northwest to southeast, then to northeast and southwest. Meanwhile, areas with moderate LTFL gradually concentrate to the northeast, and those with low loads move from northeast to southwest, gradually decreasing thereafter. Overall, the LTFL gradually decreases with increasing distance from the river channel (Fig. 6). In addition, the average LTFL in the Qarqan River buffer zone (QRBZ) is smaller than that of the Tarim River buffer zone (TRBZ), with values for TRBZ being 141.3 g/m^2 (2001), 127.83 g/m^2 (2005), 176.29 g/m^2 (2010), 229.03 g/m^2 (2015), and 376.17 g/m^2 (2021), and for QRBZ being 190.71 g/m^2 (2001), 103.35 g/m^2 (2005), 377.63 g/m^2 (2010), 147.28 g/m^2 (2015), and 289.65 g/m^2 (2021). Meanwhile, T-test was performed to determine the significance of the differences between the litterfall load of the Tarim River and the Qarchan River after confirming the data's normality using the Kolmogorov–Smirnov test (Fig. 7).

Interannual variation of LTFL

The total LTFL decreased from 7.01×10^9 kg in 2001 to 4.40×10^9 kg, followed by a significant increase to 12.5×10^{10} kg, particularly notable between 2015 and 2021. The changes in LTFL mainly occurred in areas with higher loads (Figs. 8, 9). Throughout the study period, LTFL in the QRBZ exhibited a similar pattern to the overall LTFL in the entire area, with a significant drop followed by a gradual increase. Conversely, LTFL in the TRBZ showed a gradual increase (Fig. 8).

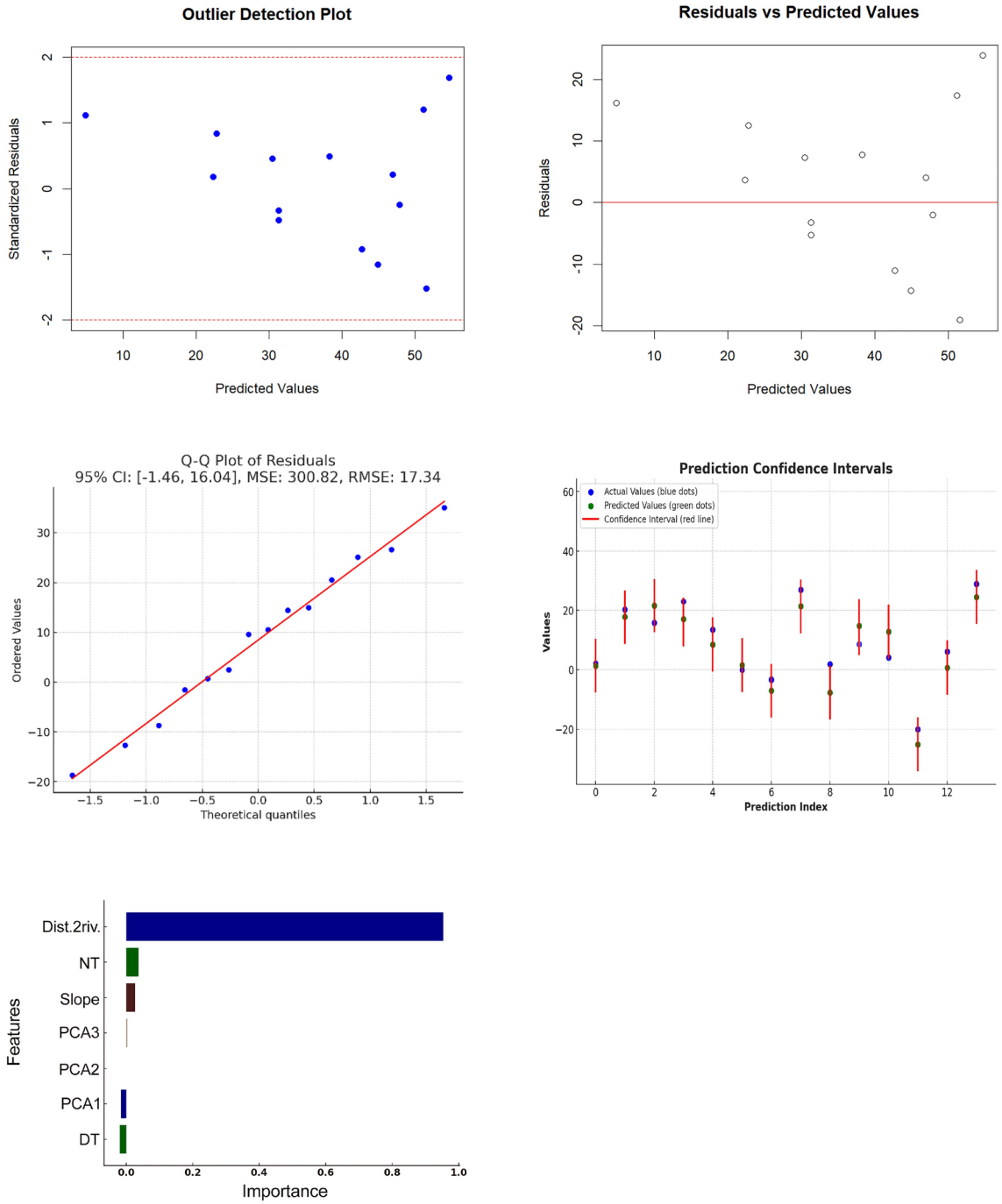

Fig. 4. Detection of prediction outliers, homoscedasticity, normality testing, 95% confidence interval, sensitivity analysis, assessment of multicollinearity among predictor variables, and Shapiro–Wilk normality test. Dist.2riv., NT, DT, PCA1, PCA2, PCA3, and VIF are distance to major river tributaries, nightly average land surface temperature, daily average land surface temperature, principal component 1, 2, and 3, and variables inflation factor (VIF) used for testing multicollinearity among predictor variables.

Additionally, During the study period, regions with high litterfall (LTFL) initially decreased but then steadily increased. However, TRBZ exhibited a significant increase until 2010, followed by a plateau for the remainder of the period. However, areas with medium LTFL initially decreased, followed by a significant recovery after 2015, during which the minimum value was observed for both TRBZ and QRBZ. The low LTFL showed a similar trend, with the most significant decline occurring between 2015 and 2021. Generally, the changes in litterfall are concentrated in the mutual transformation between areas with medium and low LTFL, with TRBZ showing a more pronounced change than QRBZ (Fig. 9).

Variables	VIF	Shapiro–Wilk normality test	
Dist to river	1.32	W	0.883
NT	1.26	p	0.000
DT	1.04		
PCA1	1.10		
PCA2	1.29		
PCA3	1.09		
Solope	1.16		

Table 4. Multicollinearity and normality test results. Dist to river, NT, DT, PCA1, PCA2, PCA3, and VIF are distance to major river tributaries, nightly average land surface temperature, daily average land surface temperature, principal component 1, 2, and 3, and variables inflation factor (VIF) used for testing multicollinearity among predictor variables.

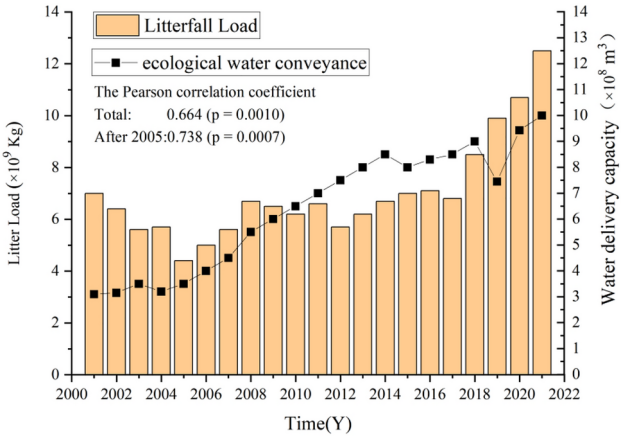

Fig. 5. Comparison of LTFL and water transfer volume from 2001 to 2021.

Discussion

Spatial distribution of LTFL in arid zone riparian forest

The global estimate for litterfall indicates that the highest amount of litterfall is found in the tropics, followed by temperate and subtropical regions[3]. Interestingly, the highest average litter load occurs in subtropical regions[3], suggesting that moderate precipitation and high temperatures support litterfall accumulation[48]. In comparison to the global average LTFL for deciduous forests (65.07–1078.26 g/m^2)[3], our results fall within the range of the second quartile (108.79–310.04 g/m^2) of the global average value despite the area's sparse vegetation[49]. This could be due to the study area being an arid region with little rainfall and high temperatures, where riparian vegetation is primarily supported by groundwater influenced by the Tarim and Qarqan river channels[49]. Vegetation in this area can produce a significant amount of dead leaves and branches, as plants experience periodic dry spells when water consumption in the upper stream-irrigated region increases[50]. Due to the lack of rainfall, microbial activity is hindered, and decomposition is slower compared to humid regions[51]. As a result, the LTFL is relatively high in our study area. The spatial distribution patterns of land surface temperature, LTFL, and river channel (Fig. 10) show similarity that also supports this explanation. Moreover, the Coefficient of Variability (CV) in our results ranges from 53.09 to 112.04%, which is higher than the global average CV of 53%[3]. This variability could be due to the large variability and unpredictability of river runoff[8] that affects groundwater depth, ultimately leading to huge variations in the production and decomposition of litterfall in the region.

Previous studies[52] on LTFL in desert regions reported an average LTFL value of 154 ± 173 g/m^2 in arid climates and 111 ± 140 g/m^2 under semi-arid climates. This is close to the results we obtained in some desert edge areas in the study area, which can prove that the prediction results of this model have certain reference value. Additionally, LTFL estimations in a study by Yang et al.[8] in the Tarim River vicinity reported significantly higher values than those in our study. This could be due to differences in the size of the study area and study procedures. Their study only selected the riverbank forest area within a few hundred meters of the river channel, which is less than 14.2% of our study area and includes areas with lower vegetation coverage. Furthermore, the average tree diameter in their sampling pool is higher than ours, which may explain the discrepancies. Previous research on litterfall dynamics[53] has demonstrated that as trees age, indicated by their diameter at breast height, the quantity of dead material they produce tends to increase. Additionally, the variation may be attributed to enhanced litter decomposition rates due to ecological water supply, resulting in lower LTFL estimates in our study than those derived from individual tree litterfall estimates.

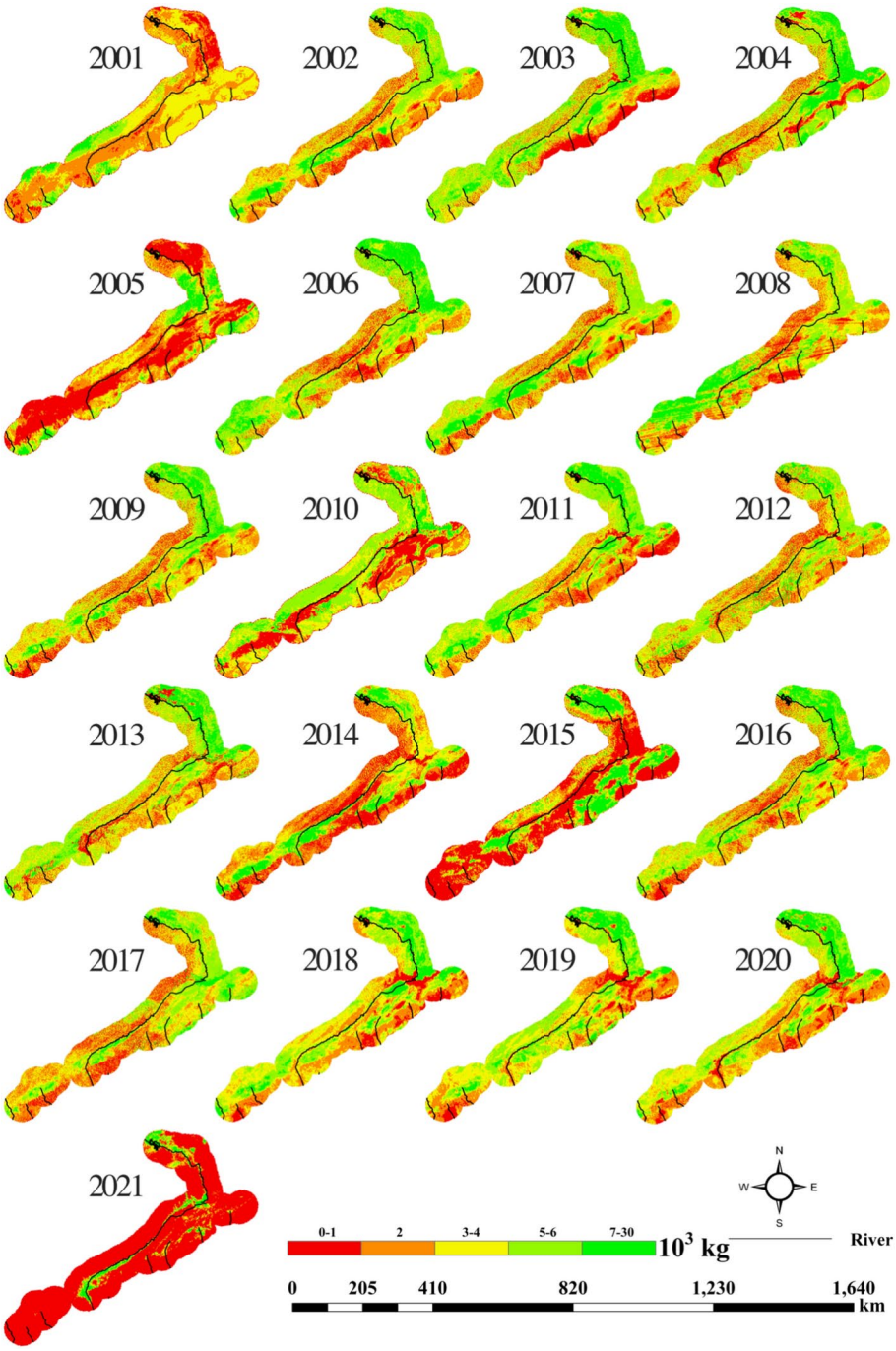

Fig. 6. The distribution map of the litterfall load (LTFL) from 2001 to 2021.

In the study area, the low to medium LTFL section accounts for more than half of the total study area. This is because the study area predominantly has sparse vegetation with many shrubs, resulting in lower litterfall production in most of the study area. In addition, as *Populus euphratica* is a major source of LTFL production, LTFL spatial distribution patterns are overwhelmingly similar with vegetation coverage and distribution, that is, medium to high LTFL mainly situated around the vicinity of riverbeds and lakes. The distribution pattern mentioned above was consistently observed throughout most of the years in our data. However, it is important to note that the spatial distribution pattern in 2005 deviated significantly from the general trends (Fig. 6). Notably, the smallest value of LTFL was recorded during this particular year over the study period. Furthermore, LTFL exhibited a slight increase while the water volume through the ecological water transfer to the region showed a relatively large increase. Based on these patterns, it is reasonable to assume that this anomaly can be attributed to a significant increase in the volume of water being conveyed downstream of the Tarim River, owing to an ecological restoration project. This increased water flow has likely fostered the decomposition of

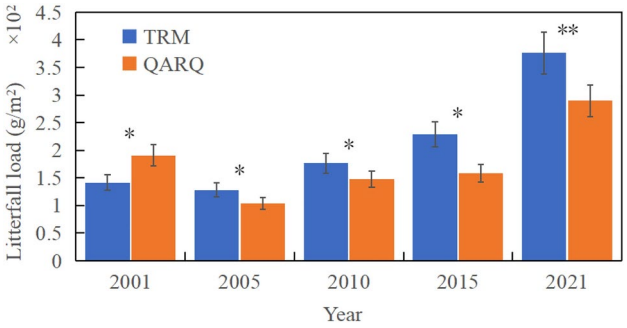

Fig. 7. comparison of LTFL between Tarim river (TRM) and Qarqan river (QARQ) buffer zone from 2001 to 2021.

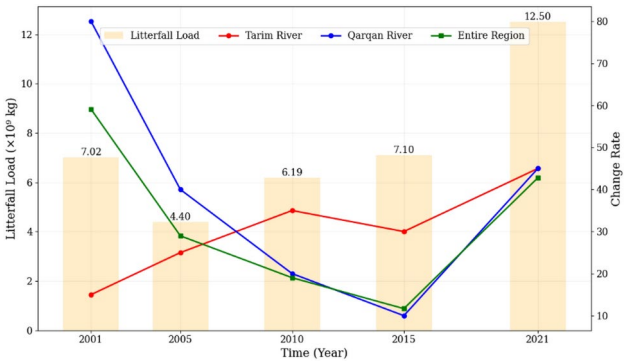

Fig. 8. LTFL variation during the study period.

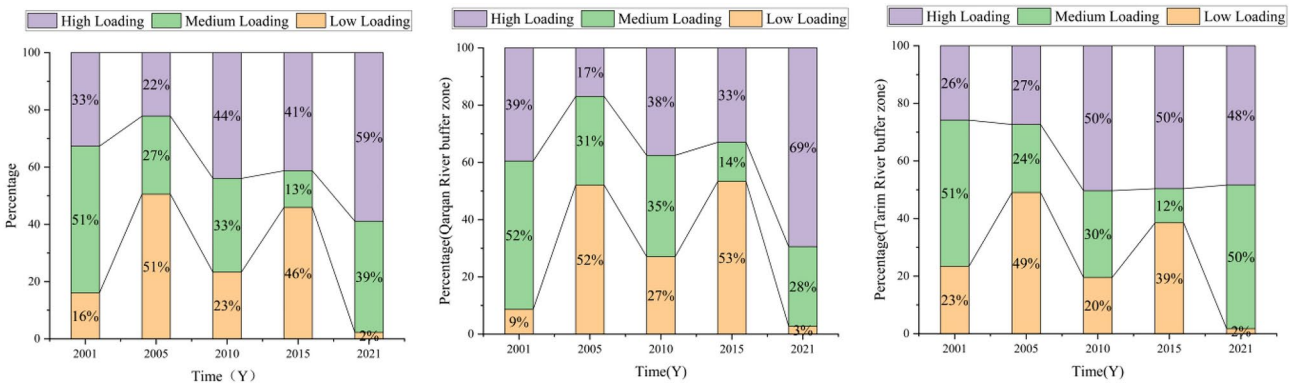

Fig. 9. The change of proportion of high, medium, and low LTFL sections of the whole study area, Tarim river buffer zone (TRBZ) and Qarqan river buffer zone (QRBZ) during the study period.

litter, particularly in regions proximate to the river channels, thereby resulting in discernible differences in distribution during this time period.

Furthermore, the average LTFL in the TRBZ is significantly higher than in the QRBZ. Especially, the proportion of medium LTFL areas is larger in TRBZ than in QRBZ, possibly because woody litter constitutes a significant portion of annual litterfall production (22–81%) in TRBZ. The dense vegetation and well-developed *populus euphoretic* forests downstream of the Tarim River benefit from the ecological water supply, explaining the higher LTFL. In contrast, the QRBZ, dominated by shrubs and annual herbs, shows lower stability and is more susceptible to changes in water supply.

Prior to 2001, data before the start of ecological water transfer showed that Moran's I was 0.91, indicating a highly concentrated spatial distribution of litter load and strong positive spatial autocorrelation, with significant accumulation of litter in certain areas (Fig. 11). The data from 2002 to 2021 shows that after the start of ecological water transfer, Moran's I statistic fluctuates and still maintains a strong positive spatial autocorrelation. Especially

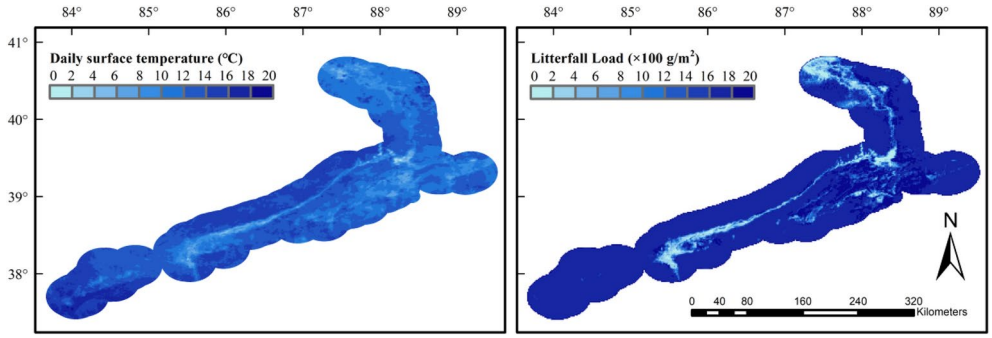

Fig.10. Comparison of litter load distribution and temperature in the study area.

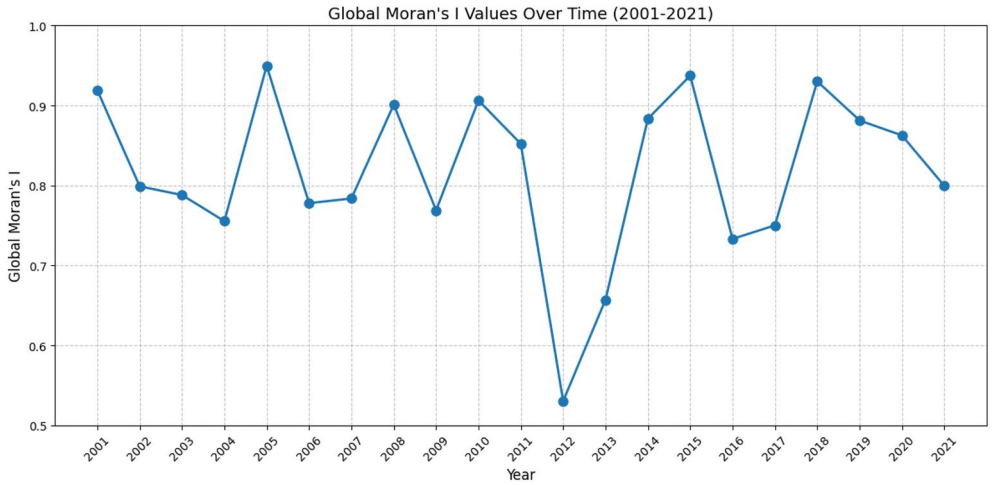

Fig.11. Global spatial autocorrelation analysis of LTFL from 2001 to 2021.

in 2005, Moran's I reached a peak of 0.94, indicating that ecological water transport significantly enhanced the agglomeration effect in certain regions. In 2012, Moran's I dropped to its lowest value of 0.53, which may reflect the uneven or lagging water delivery effect in some regions, leading to a weakening of spatial clustering effect.

Between 2001 and 2021, spatial autocorrelation remained at a high level, with Moran's I significantly higher than 0.75 in most years, indicating that the spatial distribution of litter load is still concentrated, and the accumulation of litter in certain areas continues to be high. Especially from 2016 to 2021, Moran's I index gradually rebounded, indicating that the long-term effects of ecological water transport are gradually becoming apparent, and vegetation in more areas is being restored and forming new aggregation effects. In 2018 and 2019, Moran's I reached 0.93 and 0.88, respectively, reflecting the sustained positive impact of water transport on vegetation in these years. The initiation of ecological water supply significantly enhanced vegetation cover and litter load in certain areas over the years, especially reaching its peak in 2005, indicating that the introduction of water resources may have effectively promoted ecological restoration in some arid regions. Although the overall effect is positive, the spatial autocorrelation in some years (such as 2012) has significantly decreased, which may be related to uneven water supply, timeliness issues, or external factors. In the long run, after 2016, spatial autocorrelation gradually increased, indicating that the sustained effect of ecological water transport made the spatial aggregation of litter load more apparent (Fig. 12).

Interannual variation of LTFL in arid desert riparian forests

In the past 20 years, LTFL in the study area decreased significantly from 2001 to 2005, then gradually increased from 2006 to 2021, with a rapid rise from 2015 onwards. The declining trend of global average LTFL from 2001 to 2010[3] aligns with the variation trend of LTFL in our study, possibly due to wetting combined with the warming trend of the global climate, especially in arid regions[54]. In addition, before the initiation of the EWC project in 2000, the lower reaches of the Tarim River received almost no water due to intensive irrigation in the mid and upper sections of the river. This led to an increase in the production of a relatively large amount of LTFL accumulation until the initial stages of the EWC project while hindering microbial activity responsible for slow or minimal litter decomposition. Consequently, significant amounts of undecomposed litter accumulated beneath *Populus euphratica* trees. Field experiments have also shown that flooding disturbances promote the breakdown of *Populus euphratica* leaf litter[55]. Thus, EWC accelerated litter decomposition, leading to a significantly higher

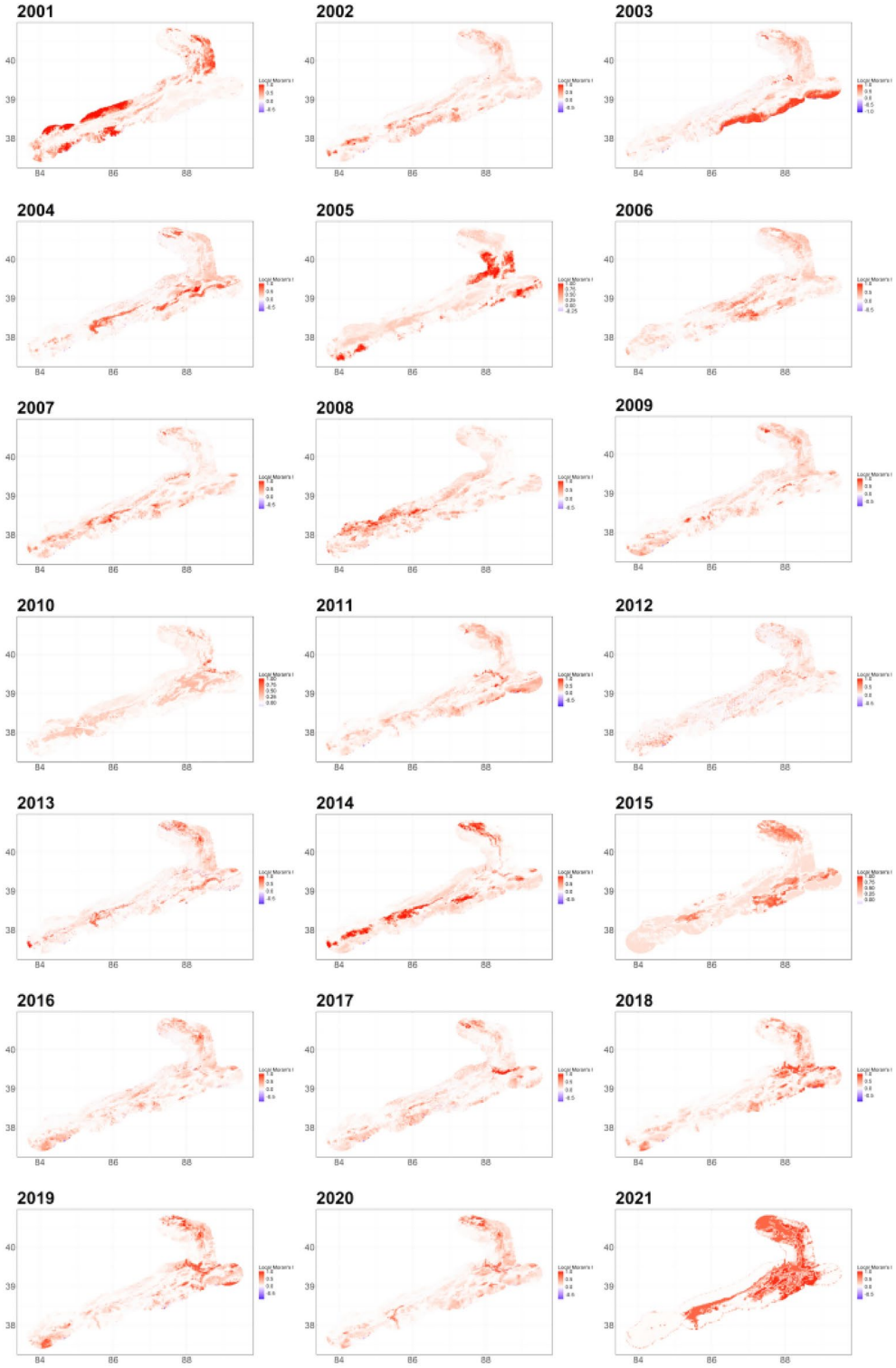

Fig.12. Local spatial autocorrelation analysis of LTFL from 2001 to 2021.

rate of LTFL change in the study area compared to the global average from 2001 to 2005[56]. From 2015 to 2021, significant increase in the amount of water delivered to TBRZ has led to a marked increase in vegetation cover in the lower reaches of the Tarim River, thereby driving the increase in LTFL until the ecosystem reaches a relatively stable state.

During the last 20 years, the rate of change in LTFL in the QRBZ was higher than that in the TRBZ in the first and last five years, while the mid-term rate of change was higher in the TRBZ than in QRBZ. This difference may be attributed to the varying responsiveness of regions to changes in ecological water supply. The TRBZ, being the

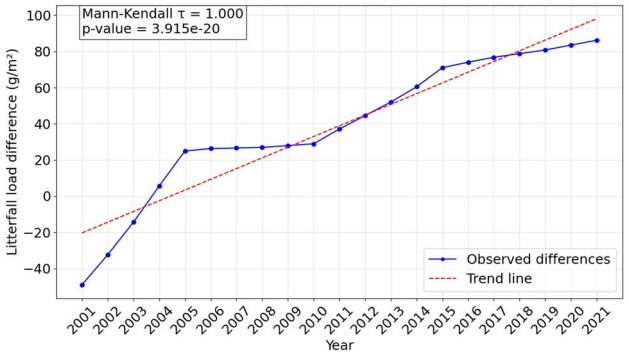

Fig. 13. LTFL difference (D-value) between Tarim river and Qarqan river from 2001 to 2021.

primary target of ecological water supply, exhibits higher vegetation recovery intensity compared to the QRBZ. Thus, the sensitivity of LTFL in the Tarim River basin is stronger, while the Qarqan River basin, dominated by shrubs, is less stable and more prone to interference from changes in water supply.

Annual differences in litterfall load between Tarim and Qarqan riparian zones were calculated and analyzed to assess the spatiotemporal variations in litterfall production between these two regions (Fig. 13). In 2001, the litter load in the Qarqan River was significantly higher than that in the Tarim River, with a difference of about 49.4 g/m^2, indicating better vegetation conditions in the Qarqan River area. However, by 2005, the litter load in the Tarim River began to exceed that in the Qarqan River, with a difference of 24.5 g/m^2. By 2010, the advantage of the Tarim River had further expanded, with a gap of 29.0 g/m^2. In 2015, the litter load in the Tarim River increased significantly, with a gap of 70.7 g/m^2. By 2021, the difference in litter load in the Tarim River reached its highest point of 86.5 g/m^2.

During this period, a large number of vegetation restoration projects were implemented in the Tarim River Basin, such as windbreak and sand fixation forests and ecological restoration. The introduction of plants and improvement of the ecological environment significantly promoted vegetation growth and the accumulation of litter. In addition, climate change and water resource management are also important factors contributing to the increase in litter load. Through artificial irrigation and water conveyance projects, vegetation in the Tarim River Basin has been provided with better growth conditions, and litter accumulation has accelerated. In contrast, the Qarqan River Basin has stable natural vegetation, has not undergone large-scale artificial intervention, and relies less on artificial water resource scheduling, resulting in a slower increase in litter load. The differences in vegetation types between the two river basins also have an impact on litter load. The Tarim River basin has introduced more drought tolerant plants with shorter litter cycles and higher yields, while the perennial drought tolerant plants in the Qarqan River basin have led to relatively slow accumulation of litter.

A comparison of LTFL distribution changes between TRBZ and QRBZ revealed that TRBZ had more high-load areas in the early study stages, likely due to the prevalence of deciduous forests and litterfall accumulation hindered by adverse environmental conditions. We developed a BP neural network model using first-hand litterfall observation data to estimate LTFL in both regions from 2000 to 2021. By integrating established sampling standards, forest resource survey data, and high-resolution remote sensing images, we analyzed the characteristics and distribution of forest surface combustibles. Despite only having 75 data points, the model effectively estimated LTFL, which is crucial in arid regions where data collection is often limited. Although the model performed well, using time-series data could help reduce prediction uncertainties, particularly for historical periods, as we relied on 2021 data to estimate LTFL for other years of interest. To better reflect annual changes in litter and minimize seasonal influences, we selected indices based on images obtained from the same month each year, averaging all images from that month to reduce acquisition errors. However, using the same model for LTFL estimation across multiple years, especially historical data, may introduce errors due to changing LTFL loads over the study period. Nevertheless, the small overfitting observed in the 2021 model suggests it can estimate LTFL in different years if the relationship between dependent and predictor variables remains stable. Previous studies indicate that at a macro scale, LTFL load is driven by a combination of precipitation and temperature[56]. At the macro to small scale, as in this study, LTFL is driven by local factors such as local terrain, soil, and vegetation structure[57]. Although the influencing factors are different at different scales, the relationship between LTFL and driving factors is likely to be relatively stable, which means our estimation for historical periods could give us a relatively reliable understanding of the spatiotemporal dynamics of LTFL in the study area. In addition, we performed Kendal's change detection test to test the potential changes in temperature and precipitation that may alter the LTFL spatiotemporal dynamics in the study area and adjust the weight of the factors incorporated in the model. This test is used to detect whether there is a significant monotonic trend in time series data, providing background information about climate driving factors for the LTFL model and helping to determine the weight of variables in the model and the reliability of estimation results. The results showed that temperature and precipitation changes were not statistically significant during the study period (Fig. 14), potentially reaffirming the reliability of our estimation. However, the lack of direct inventory data before 2021 still prevents LTFL prediction results of the historical periods from being used in management practices. Therefore, by setting a fixed observation station of LTFL in the study area based on our LTFL estimation, long-term observation of LTFL can give us a more in-depth mechanistic explanation of the

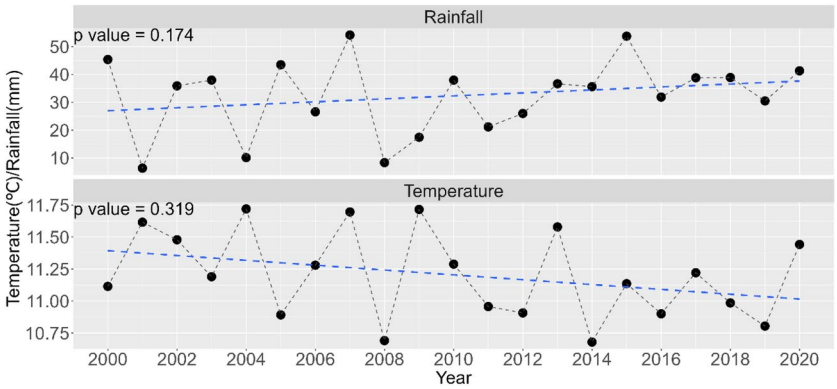

Fig. 14. The change in precipitation and temperature of Qarqan county during 2000–2020 (The P-value represents the significance test result of Mankendal trend analysis. Although there are fluctuations in annual average temperature and annual precipitation during the study period, the trend is not significant, indicating that climate factors are relatively stable.)

spatiotemporal dynamics of LTFL, which can be useful for targeted restoration of damaged riparian ecosystems and wildfire hazard management in policy front.

Conclusion

This study used the BP neural network model based on LTFL inventory data to estimate litter load and analyze its spatiotemporal variation in the riparian forest ecosystem of TRBZ and QRBZ. The results showed that the LTFL in the study was within the mid-range of the global average for deciduous forests, and it was mainly concentrated around the river channels. Additionally, the LTFL increased from the west to the east, possibly due to the EWC. The vegetation in the east of the study area recovered well due to increased surface runoff from the EWC, promoting the accumulation of litter. Over the past 20 years, the LTFL decreased first and then increased. In the early stages of the EWC, the litter decomposition process accelerated, potentially driven by increased microbial activities from improved soil conditions with the EWC, causing a decrease in the LTFL. Since 2005, vegetation coverage has significantly increased, causing an increase in the litter input and further promoting an increase in the LTFL. By estimating the litter load and spatiotemporal distribution patterns of riparian forests in arid areas, this study provides scientific basis for ecosystem management and wildfire risk. The load of litter is related to vegetation growth and decomposition processes, revealing the impact of EWC projects on litter and assisting in water resource optimization and vegetation restoration. Meanwhile, as an important component of surface fuel, excessive accumulation of litter increases the risk of wildfires. By monitoring its changing trends, research provides key information for wildfire risk management, helps control surface fuels, and reduces the probability of wildfire occurrence. Combining remote sensing technology with BP neural network models, this study provides precise tools for ecological management and fire prevention decision-making.

Policy advice

In wildfire risk management, focusing on areas characterized by high Litter Total Fuel Load (LTFL) is imperative. Recognized as a principal fuel source for wildfires, regions exhibiting elevated LTFL levels are susceptible to ignition under conducive moisture and temperature conditions. Hence, establishing long-term observation sites for wildfire surveillance is advocated, particularly within the eastern sector of the study area where LTFL concentrations are notably high. This strategy is especially pertinent during autumn, marked by diminished humidity levels and constrained litter moisture content. Furthermore, implementing targeted EWC warrants consideration. LTFL, integral to soil composition, serves as a significant contributor to organic carbon input. Under favorable soil moisture conditions conducive to microbial activity, heightened litter decomposition rates ensue, thereby augmenting soil organic carbon levels. Additionally, litter's provision of essential nutrients fosters microbial species diversity, thereby enhancing the structural integrity and bolstering the resilience and stability of desert riparian ecosystems. Hence, our research underscores the efficacy of directing water resources towards regions with elevated LTFL concentrations. Such an approach ensures that EWC initiatives yield maximal ecological benefits, surpassing the outcome of indiscriminate water dispersion across the landscape.

Data availability

The field data used in our study comes from the forest combustible material standard plot survey project in Xinjiang Uygur Autonomous Region. All field sampling data is currently preserved by the Forestry and Grassland Bureau of Ruoqiang County. Due to the fact that the original data is considered government property, in order to maintain the integrity and confidentiality of the data, it is not publicly available. If there is a reasonable request, please contact the corresponding author.

Received: 30 April 2024; Accepted: 5 December 2024
Published online: 07 January 2025

References

1. Caspers, H. Biogeochemistry of a forested ecosystem. *Int. Rev. Hydrobiol.* **63**, 772–772 (2010).
2. Zedaker, S. M. Forest ecosystems: Concepts and management. *For. Sci.* **3**, 841–842 (1986).
3. Shen, G., Chen, D., Wu, Y., Liu, L. & Liu, C. Spatial patterns and estimates of global forest litterfall. *Ecosphere* https://doi.org/10.1002/ecs2.2587 (2019).
4. Staelens, J. et al. Spatio-temporal litterfall dynamics in a 60-year-old mixed deciduous forest. *Ann. For. Sci.* **68**, 89–98. https://doi.org/10.1007/s13595-011-0010-5 (2011).
5. Starr, M., Lindroos, A.-J. & Ukonmaanaho, L. Weathering release rates of base cations from soils within a boreal forested catchment: variation and comparison to deposition, litterfall and leaching fluxes. *Environ. Earth Sci.* **72**, 5101–5111. https://doi.org/10.1007/s12665-014-3381-8 (2014).
6. Huo, S. et al. Characteristics of dissolved organic nitrogen (DON) in the surface water of Beijing Olympic Forest Park. *Environ. Earth Sci.* **71**, 4021–4028. https://doi.org/10.1007/s12665-013-2785-1 (2014).
7. Eggleston, H. S. In *2006 Guidelines for National Greenhouse Gas Inventories* (eds Eggleston, H. S. et al.) (IGES, 2006).
8. Yang, C. et al. Streamflow abrupt change and the driving factors in glacierized basins of Tarim Basin, Northwest China. *Adv. Clim. Change Res.* **15**(1), 75–89 (2024).
9. Alvarez, J. A., Villagra, P. E., Rossi, B. E. & Cesca, E. M. Spatial and temporal litterfall heterogeneity generated by woody species in the Central Monte desert. *Plant Ecol.* **205**, 295–303 (2009).
10. Guo, Y. et al. Composition and spatio-temporal dynamics of litter fall in a northern tropical karst seasonal rainforest in Nonggang, Guangxi, southern China. *Biodivers. Sci.* **25**, 265–274 (2017).
11. He, C., Huang, Q., Shen, S. & Wang, F. Forest fuel loading estimates based on a back propagation neutral network. *Tsinghua Univ. Sci. Technol.* **51**, 230–233. https://doi.org/10.16511/j.cnki.qhdxxb.2011.02.018 (2011).
12. Ottmar, R. D. et al. An overview of the fuel characteristic classification system—quantifying, classifying, and creating fuelbeds for resource planning. *Can. J. For. Res.* **37**(12), 2383–2393 (2007).
13. Chuvieco, E. et al. Remote sensing of large wildfires: in the European Mediterranean Basin. In *Integrated Fire Risk Mapping* 61–100 (Springer, 1999).
14. S Yanlong, ZMH Haiquan. Model of surface combustibles in camphor pine forests in the Greater Khingan Mountains region. *J. Northeast For. Univ.* (2005).
15. Haiqing, H., Sisheng, L. B. L., Zhenshi, W., Zhangwen, S. & Shujing, W. The prediction of moisture content of surface ground fuel of typical forest stand in Daxing' anling mountains. *J. Cent. S. Univ. For. Technol.* **38**, 1–9. https://doi.org/10.14067/j.cnki.1673-923x.2018.11.001 (2018).
16. Qingyun, W. & Keane, R. E. Relationship between surface dead fuel loadings and environmental factors in southern Jiangxi, China. *Chin. J. Appl. Ecol.* **33**, 1539–1546. https://doi.org/10.13287/j.1001-9332.202206.021 (2022).
17. Yang, Y., Zhou, H., Ye, Z. & Zhu, C. Estimation of *Populus euphratica* forest leaf litterfall and time variation of nutrient in leaf litter during decomposition along the main channel of the Tarim River, China. *Water* **13**, 2514 (2021).
18. Ngangyo-Heya, M. et al. Foliar biomass production and litterfall pattern of five timber species in forest plantations of semi-arid lands of the Northeastern Mexico. *Bot. Sci.* **95**, 295–305. https://doi.org/10.17129/botsci.770 (2017).
19. Verma, A. et al. Litter production and litter dynamics in different agroforestry systems in the arid western region of India. *Biol. Agric. Hortic.* **38**, 40–60. https://doi.org/10.1080/01448765.2021.1971110 (2022).
20. Feisi. *Neural Network Theory and MATLAB 7 Implementation*. Neural Network Theory and MATLAB 7 Implementation. (2005).
21. Fan, Z. L., Xu, H. L., Zhang, P. & Zhao, X. The Qarqan River in Xinjiang and its water resources utilization. *Arid Zone Res.* **31**, 20–26. https://doi.org/10.13866/j.azr.2014.01.003 (2014).
22. Wang, Y. *Study on Variations of Runoff and its Components in Three Source Streams of Tarim River* (Xi'an University of Technology, 2022).
23. Jikai, G. *Responses of vegetation coverage to climate change and human action activities in the Tarim River Basin* (Beijing Forestry University, 2016).
24. Jin, Q. R. et al. Assessment of fire risk level of non-growing season in the lower reaches of Tarim River using remotely sensed data. *Anhui Agric. Sci.* **51**, 101–105 (2023).
25. Cao, X., Chen, J., Matsushita, B. & Imura, H. Developing a MODIS-based index to discriminate dead fuel from photosynthetic vegetation and soil background in the Asian steppe area. *Int. J. Remote Sens.* **31**, 1589–1604 (2010).
26. Gamon, J. A., Penuelas, J. & Field, C. B. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* **41**, 35–44 (1992).
27. Huete, A. & Escadafal, R. Assessment of biophysical soil properties through spectral decomposition techniques. *Remote Sens. Environ.* **35**, 149–159 (1991).
28. Kaufman, Y. J. & Tanre, D. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE Trans. Geosci. Remote Sens.* **30**, 261–270 (1992).
29. Kauth, R. J., & Thomas, G. S. The tasselled cap--a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In LARS symposia, 159 (1976).
30. Meyer, G. E., Mehta, T., Kocher, M. F., Mortensen, D. A. & Samal, A. Textural imaging and discriminant analysis for distinguishing weeds for spot spraying. *Trans. ASAE* **41**, 1189–1197 (1998).
31. Miura, T., Huete, A. R. & Yoshioka, H. Evaluation of sensor calibration uncertainties on vegetation indices for MODIS. *IEEE Trans. Geosci. Remote Sens.* **38**, 1399–1409 (2000).
32. Novozhilov, G. N., Dav'Ydov, O. V., Mazurov, K. V., Dudochkin, N. A. & Mikhalov, N. M. The vegetative index of Kerdo as an indication of primary adaptation to hot climate conditions. *Voenno-meditsinskii Zhurnal* **8**, 68–69 (1969).
33. Pearson, R. L., & Miller, L. D . Remote mapping of standing crop biomass for estimation of the productivity of the shortgrass prairie, Pawnee National Grasslands, Colorado(1972).
34. Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. A modified soil adjusted vegetation index. Remote sensing of environment, 48(2), 119-126 (1994).
35. Rondeaux, G., Steven, M. & Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **55**, 95–107 (1996).
36. Sandholt, I., Rasmussen, K. & Andersen, J. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. *Remote Sens. Environ.* **79**, 213–224 (2002).
37. Verstraete, B. P. M. GEMI: a non-linear index to monitor global vegetation from satellites. *Vegetation* **101**, 15–20 (1992).
38. Wan Z, Hook S, Hulley G. NASA EOSDIS Land Processes Distributed Active Archive Center. (2021).
39. Zhang, H. et al. High-resolution vegetation mapping using eXtreme gradient boosting based on extensive features. *Remote Sens.* https://doi.org/10.3390/rs11121505 (2019).
40. Farr, T. G. et al. The shuttle radar topography mission. *Rev. Geophys.* https://doi.org/10.1029/2005RG000183 (2007).
41. Pekel, J.-F., Cottam, A., Gorelick, N. & Belward, A. S. High-resolution mapping of global surface water and its long-term changes. *Nature* https://doi.org/10.1038/nature20584 (2016).
42. Li, M. Comprehensive review of backpropagation neural networks. *Acad. J. Sci. Technol.* **9**, 150–154 (2024).
43. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**(6088), 533–536 (1986).
44. Roelofs R et al. A meta-analysis of overfitting in machine learning. *Adv. Neural Inf. Process. Syst.* **32** (2019).
45. Khirirat S, Feyzmahdavian HR, Johansson M. In *56th Annual IEEE Conference on Decision and Control (CDC)*. IEEE. (2017).

46. Kingma, D.P., & Ba, J. Adam: A Method for Stochastic Optimization. CoRR, abs/1412.6980, (2014).
47. Olden, J. D. & Jackson, D. A. Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. *Ecol. Model.* **154**(1–2), 135–150 (2002).
48. Ahirwal, J. et al. Forests litter dynamics and environmental patterns in the Indian Himalayan region. *For. Ecol. Manag.* **499**, 119612 (2021).
49. Jiang, N., Zhang, Q., Zhang, S., Zhao, X. & Cheng, H. Spatial and temporal evolutions of vegetation coverage in the Tarim River Basin and their responses to phenology. *Catena* **217**, 106489. https://doi.org/10.1016/j.catena.2022.106489 (2022).
50. Song, J., Betz, F., Aishan, T., Halik, U. & Abliz, A. Impact of water supply on the restoration of the severely damaged riparian plants along the Tarim River in Xinjiang, Northwest China. *Ecol. Indic.* https://doi.org/10.1016/j.ecolind.2024.111570 (2024).
51. Liu, X. et al. Climatic drivers of litterfall production and its components in two subtropical forests in South China: A 14-year observation. *Agric. For. Meteorol.* **344**, 109798. https://doi.org/10.1016/j.agrformet.2023.109798 (2024).
52. Neumann, M. et al. Dynamics of necromass in woody Australian ecosystems. *Ecosphere* https://doi.org/10.1002/ecs2.3693 (2021).
53. Chen, H. Y. H., Brant, A. N., Seedre, M., Brassard, B. W. & Taylor, A. R. The contribution of litterfall to net primary production during secondary succession in the Boreal Forest. *Ecosystems* **20**, 830–844. https://doi.org/10.1007/s10021-016-0063-2 (2017).
54. Zhang, D., Zhang, L., Yang, J. & Feng, G. The impact of temperature and precipitation variation on drought in China in last 50 years. *Acta Phys. Sin.* **59**, 655–663. https://doi.org/10.7498/aps.59.655 (2010).
55. Yang, Y., Zhou, H., Ye, Z., Zhu, C. & Chen, Y. Effects of transient flooding on leaf litter decomposition: a case study of *Populus euphratica* leaf in an arid area. *Int. J. Agric. Biol.* **22**, 1385–1392 (2019).
56. Zhang, H., Yuan, W., Dong, W. & Liu, S. Seasonal patterns of litterfall in forest ecosystem worldwide. *Ecol. Complex.* **20**, 240–247. https://doi.org/10.1016/j.ecocom.2014.01.003 (2014).
57. Morffi-Mestre, H. et al. Multiple factors influence seasonal and interannual litterfall production in a tropical dry forest in Mexico. *Forests* https://doi.org/10.3390/f11121241 (2020).
58. Özesmi, S. L., Tan, C. O. & Özesmi, U. Methodological issues in building, training, and testing artificial neural networks in ecological applications. *Ecol. Model.* **195**, 83–93 (2006).
59. Yang, J. & Huang, X. The 30 m annual land cover datasets and its dynamics in China from 1985 to 2023. *Earth Syst. Sci. Data* **13**(1), 3907–3925 (2024).

Acknowledgements

The authors extend their sincere gratitude to their esteemed colleagues, Nuria Abduwayit and Abduaini Maimaiti, for their invaluable assistance during the fieldwork. Furthermore, the authors wish to express profound appreciation to the personnel at the Xinjiang Institute of Ecological Geography for their unwavering technical support and selfless dedication to the collection of field data for this project. Additionally, the authors extend their heartfelt thanks to my colleagues for them invaluable assistance in enhancing the language of the manuscript. Finally, the authors would like to convey their gratitude to the staff of the Forestry Department in Qiemo (Charchen) and Ruoqiang (Charqiliq) for their selfless contributions during the sample collection phase of this research endeavor.

Author contributions

X.J. : field investigation, conceptualization and writing of the initial draft paper, drawing figures. A. E. designing and writing sections of the study. A.K. and U.H. reviewed and edited the paper, M.Y., Z.K., A.H. and S.Z. helped complete the data collection and processing of this study, while T.A., T.V. and H.A. helped complete the final review of the article and provided suggestions, S.U. and G.F. participated in data collection and processing during the article revision phase.

Funding

This work was jointly funded by the National Natural Science Foundation of China (Grant No. 32071655; 32260285), Tianchi talent (young scientist) fund (E335030101), Chinese Academy of Sciences President's International Fellowship Initiative (PIFI, 2021VCA0004, 2024PVA0101, 2024PVB0064). and the Project for Cultivating High-Level Talent of Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences (Grant No. E450030101).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-82435-2.

Correspondence and requests for materials should be addressed to A.E. or Ü.H.

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

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

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

© The Author(s) 2024