

基于 PLS-SEM 和 SHAP 模型的腾格里沙漠土壤水分影响因素分析

*

威海迪¹, 张定海¹, 张志山², 赵有益¹, 史战红¹

(1. 甘肃农业大学理学院数量生物研究中心, 甘肃 兰州 730070; 2. 中国科学院西北生态环境资源研究院沙坡头沙漠研究试验站, 甘肃 兰州 730000)

摘要:【目的】土壤水分对于干旱半干旱生态系统的结构和功能有重要影响, 其分布受地形、植被等多种因素影响。为制定有效的土壤水分保持策略, 针对沙丘不同微地貌类型和土壤水分特征建立固沙植被提供参考。【方法】以腾格里沙漠东南缘半固定沙丘 4 种微地貌类型为研究对象, 运用描述统计分析土壤水分在不同微地貌类型上的分布特征。采用偏最小二乘结构方程模型构建地形、灌木植被和草本植被对不同深度土壤水分的影响路径, 分析对土壤水分产生直接和间接影响的地形植被变量。利用基于随机森林和 XGBoost 算法的 SHAP 模型按照不同地形植被变量对土壤水分的贡献大小进行排序, 选择贡献程度大的变量分析对土壤水分的影响机制。【结果】1) 背风坡、丘底的土壤水分高于迎风坡和丘顶。表层和中层土壤水分分布的差异较大, 变化较为剧烈, 而深层土壤水分波动较小, 分布较为稳定。2) 地形对土壤水分存在显著的负面影响, 对草本植被存在显著的正面影响。草本植被和灌木植被直接对土壤水分存在显著的负面影响。且灌木植被对草本植被存在显著的负面影响, 从而间接影响土壤水分。3) 随机森林模型与基于随机森林和 XGBoost 的 SHAP 模型对不同地形植被变量贡献程度的排序存在差异, 但 3 种模型均表明草本盖度、坡向和高差对表层土壤水分的贡献程度较大, 坡向、草本盖度和生物量对中层土壤水分的贡献程度较大, 灌木盖度、坡向和草本盖度对深层土壤水分的贡献程度较大。4) 不同深度土壤水分均随草本盖度和坡向增加呈先下降后上升趋势, 表层土壤水分随高差增加呈先下降后上升趋势, 中层和深层土壤水分分别随生物量和灌木盖度增加呈先上升后下降趋势。【结论】在防风固沙实践中, 要针对不同微地貌类型上土壤水分特征合理建设固沙植被。在土壤水分更低的丘顶和迎风坡设置更加密集的草方格网和沙障。在表层和中层土壤水分变化较大的区域, 种植耐旱的草本植物和浅根灌木, 以快速覆盖地表, 减少水分蒸发, 在能够扎根到深层土壤的位置, 种植深根灌木, 利用深层稳定的水分, 提高植被的整体抗旱能力。根据坡向和高差合理进行植被布局, 在不利于保持土壤水分的高坡和阳坡应选择耐旱性更强的植物种植。在固沙植被管理过程中, 需要对植被进行适当的修枝和间伐, 减少植被体积和密度过大对土壤水分的消耗。

关键词: 半固定沙丘; 土壤水分; 地形-植被变量; PLS 结构方程模型; SHAP 模型

Analysis of Factors Affecting Soil Moisture in the Tengger Desert Based on PLS-SEM and SHAP Models

Abstract: 【Objective】 Soil moisture has a significant impact on the structure and function of arid and semi-arid ecosystems, and its distribution is influenced by various factors such as topography and vegetation. To formulate effective soil moisture conservation strategies, it is essential to provide references for stabilizing sand vegetation based on the different microtopographic types of dunes and the characteristics of soil moisture. 【Method】 Taking four microtopographic types of semi-fixed dunes on the southeastern edge of the Tengger Desert as the research object, descriptive statistics are used to analyze the distribution characteristics of soil moisture in different microtopographic types. A partial least squares structural equation model (PLS-SEM) is used to construct the influence paths of topography, shrub vegetation, and herbaceous vegetation on soil moisture at different depths, analyzing the topographical and vegetation variables that have direct and indirect impacts on soil moisture. The SHAP model, based on random forest and XGBoost algorithms, is used to rank the contributions of different topographical and vegetation variables to soil moisture, and the variables with the highest contributions are selected to analyze the mechanisms affecting soil moisture. 【Result】 1) The soil moisture on the leeward slope and at the dune bottom is higher than on the windward slope and dune top. There is a significant difference in the distribution of soil moisture between the surface and middle layers, with more dramatic changes, whereas the deep soil moisture fluctuates less and is more stable. 2)

基金项目: 国家自然科学基金项目(42361016); 甘肃省科技创新计划-“西部之光”人才计划项目(22JR9KA032); 甘肃省科技计划项目(21JR7RA831); 重点研发计划-农业领域(23YFNA0036)。

*张定海为通讯作者。

Topography has a significant negative impact on soil moisture and a significant positive impact on herbaceous vegetation. Herbaceous vegetation and shrub vegetation have a significant direct negative impact on soil moisture. Additionally, shrub vegetation has a significant negative impact on herbaceous vegetation, thereby indirectly affecting soil moisture. 3) There are differences in the ranking of the contribution levels of different topographical and vegetation variables to soil moisture between the random forest model and the SHAP model based on random forest and XGBoost algorithms. However, all three models indicate that herbaceous coverage, slope aspect, and height difference have a significant contribution to surface soil moisture; slope aspect, herbaceous coverage, and biomass have a significant contribution to middle layer soil moisture; and shrub coverage, slope aspect, and herbaceous coverage have a significant contribution to deep soil moisture. 4) Soil moisture at different depths shows a trend of initially decreasing and then increasing with the increase in herbaceous coverage and slope aspect. Surface soil moisture shows a trend of initially decreasing and then increasing with the increase in height difference. Middle layer and deep soil moisture show a trend of initially increasing and then decreasing with the increase in biomass and shrub coverage, respectively. 【Conclusion】 In windbreak and sand fixation practices, it is essential to reasonably establish sand-fixing vegetation based on the soil moisture characteristics of different microtopographic types. On dune tops and windward slopes, where soil moisture is lower, denser grass grids and sand barriers should be set up. In areas where surface and middle layer soil moisture vary greatly, drought-resistant herbaceous plants and shallow-rooted shrubs should be planted to quickly cover the surface and reduce moisture evaporation. In locations where roots can reach deep soil, deep-rooted shrubs should be planted to utilize stable deep soil moisture and improve the overall drought resistance of the vegetation. Vegetation should be arranged reasonably according to slope aspect and elevation difference. On high slopes and sunny slopes, which are less favorable for retaining soil moisture, more drought-tolerant plants should be selected for planting. In the management of sand-fixing vegetation, appropriate pruning and thinning are needed to reduce excessive vegetation volume and density, thereby minimizing the consumption of soil moisture.

关于土壤水分分布规律及地形-植被因子相互作用机理的研究有助于建立水文过程和生态过程之间的定量关系 (Wang et al., 2019)。土壤水分在影响植被群落建立和演替方面起着重要的作用, 同时也是土壤-植物-大气系统水文循环的关键环节。干旱半干旱沙区植物生长所需的水分主要依赖土壤水分供给 (Asbjornsen et al., 2011)。温度、降水、太阳辐射等气象因素以及土壤质地、地形和植被等对土壤水分呈现出高度的非线性影响 (Western et al., 2004; D'Odorico et al., 2010)。在干旱半干旱沙区中, 土壤水分除了随着土壤深度而垂直变化外, 也会随植被结构的水平变化而变化。已有研究表明, 植物-土壤之间的相互作用对土壤水分动态会产生关键影响 (Anderson et al., 2003; D'Odorico et al., 2010; Liu et al., 2019; 郭艳菊等, 2022)。植物冠层的截流和蒸腾可以调节土壤水分输入, 对表层土壤水分的极端波动有缓冲作用 (Asbjornsen et al., 2004; Potts et al., 2010)。茎流和优先流可以沿着根系通道进入更深的土层, 进一步促进土壤水分的空间异质性 (Li et al., 2009)。另外, 干旱半干旱沙区土壤贫瘠、植被稀少、地形地貌复杂, 地形对土壤水分的影响尤为重要。

不同微地貌类型上, 地形和植被的差异会影响土壤水分的分布。不同地形特征通过影响降水再分配和太阳辐射量, 直接导致土壤水分含量和分布不同 (Moeslund et al., 2013)。坡度影响入渗和径流等水文过程, 受重力影响陡坡处的水分横向向下移动比平坦处更快, 水分入渗随之减少, 导致土壤水分较低 (Qiu et al., 2001)。坡向影响太阳辐射和蒸散量, 导致不同坡向的土壤水分存在差异, 北半球通常东北坡的土壤水分高于西南坡 (Perring, 1959)。风也对土壤水分的分布起重要作用, 与迎风坡相比, 背风坡由于避风而蒸发量较低、露水沉积量较大, 土壤水分通常更高 (Biederman et al., 2011)。在我国, 沙漠生态系统是干旱半干旱生态系统的重要组成部分, 然而对不同微地貌类型上地形和植被特征对土壤水分影响的定量研究较少。因此, 本研究选取腾格里沙漠东南缘沙坡头地区的半固定沙丘进行研究, 分析沙丘丘顶、丘底、背风坡和迎风坡 4 种微地貌类型上 0 cm~300 cm 深度的土壤水分特征, 并探讨地形和植被因素对土壤水分的显著影响, 以解释地形-植被因素对土壤水分的影响机制。

1 材料与方法

1.1 研究区概况

研究区位于宁夏回族自治区中卫市腾格里沙漠东南部长流水地区的半固定沙丘。该地区位于 37°30'~40°10' N, 102°20'~105°55'E 之间, 年平均温度约 10.0 °C, 冬季最低气温可达 -25.1 °C, 夏季最高气温可达 38.1 °C。年平均降雨量约 186.2 mm, 全年约 80% 的降水事件发生在 6 月到 9 月, 大约在 64% 的单次降水过程中降水强度低于 0.5 mm/h。降水是该地区植被生长的主要水分来源, 研究区地下水埋深在 50 m~80 m 之

间,无法被植物利用(张世才等,2023)。该地区年平均风速约为 2.9 m.s-1,年均沙暴天数约为 59 天,年蒸发量 3000 mm(李新荣等,2014;孙琰蕙等,2022)。研究区内分布着大量密集的格状沙丘链与新月形沙丘链,土壤基质主要为流沙,土壤类型主要为原生灰钙土和风成沙土。固沙灌木与草本植物形成了该地区的固沙植被群落,其中优势灌木主要有柠条(*Caragana korshinskii*)、花棒(*Hedysarum scoparium*)等,草本植物主要有茵陈蒿(*Artemisia capillaris*)和小画眉草(*Eragrostis minor*)等。

1.2 试验设计

2019 年 7 月在对腾格里沙漠东南边缘长流水地区的天然固沙植被全面考察后,我们设置了一块试验样地(37°27'26"N, 104°46'12"E, 海拔 1573 m)。该样地大小为 40 m×180 m,包含迎风坡、背风坡、丘底和丘顶四种微地貌类型(图 1)。具体的试验方法如下:(1)划分样地:将样地划分成 10 行 37 列,每个小样方的大小为 4 m×4 m。(2)土壤水分数据获取:选取 3、6 和 9 行的平行对应位置做为采样点,在地势起伏较大的沙丘处增加采样点,总共设置 72 个土壤水分采样点,如图 1 所示,并用土钻法进行采样。将 0-300 cm 的土层划分成 18 层,将 50 cm 以上土层划分为 0 cm~5 cm、5 cm~15 cm、15 cm~25 cm、25 cm~35 cm、35 cm~50 cm,在每层中分别采样 1 次,对于深度超过 50 cm 的土层向下每隔 20 cm 采样 1 次。将采集到的土壤样本放入已知质量 m_0 的铝盒中带回实验室,利用电子天平测定铝盒和土壤的总质量 m_1 后将铝盒和土壤一起放入 105°C 烘箱中烘干 24 小时,再次测定铝盒和土壤的总质量 m_2 ,土壤水分计算公式为 $(m_1 - m_2) / (m_1 - m_0) \times 100\%$ 。(3)草本数据获取:选取第 5 行样地,考虑到沙漠地区的植被通常较为分散和稀疏,利用 1 m×1 m 的样方框调查该行每个样方中草本植物的盖度和多度。采用“全株收获法”将该行样地中的所有草本植物割除,装入已编号的纸袋中带回实验室并置于 65°C 的烘箱中烘干至恒定重量,使用电子天平称重并记录生物量数据。(4)灌木数据获取:用标记牌对样地中的所有灌木进行标记,利用钢卷尺测量每株灌木的株高以及东西和南北 2 个方向上的冠幅。在样地中的灌木植被下布设凋落物收集框,收集框直接接触地面,面积为 50 cm×50 cm,将收集到的凋落物置于 65°C 烘箱中烘干至恒重,使用电子天平称重得到凋落物数据。(5)地形数据获取:利用 RTK 测量仪获取每个样方的顶点坐标,通过数字高程模型计算每个样方的坡度、坡向和高差数据,具体计算方法如下:

(1) 高差的计算

将利用 RTK 测得各定位点的大地坐标系数据向直角坐标系数据进行转化,并将每 1 个点的高差用该点的海拔高度与样地最低点的海拔高度之差表示。将每个样方 4 个顶点的高差分别用 z_a 、 z_b 、 z_c 和 z_d 表示,则该样方的高差为 z_a 、 z_b 、 z_c 、 z_d 的平均值。

(2) 坡度的计算

设原点坐标为 (x_0, y_0, z_0) , 首先计算每个样方的单位法向量 $\vec{n}_{i,j}$, $\vec{n}_{i,j}$ 为 xoy 平面的基本向量 $\vec{a}_{i,j}$ 和 $\vec{b}_{i,j}$ 的向量积, 计算公式如下:

$$\vec{n}_{i,j} = \vec{a}_{i,j} \times \vec{b}_{i,j}$$

其中, $\vec{a}_{i,j}$ 和 $\vec{b}_{i,j}$ 的计算公式如下:

$$\begin{cases} \overrightarrow{a_{i,j}} = \overrightarrow{P_{i+1,j+1}} - \overrightarrow{P_{i,j}} = (\Delta x, \Delta y, z_{i+1,j+1} - z_{i,j}) \\ \overrightarrow{b_{i,j}} = \overrightarrow{P_{i,j+1}} - \overrightarrow{P_{i+1,j}} = (-\Delta x, \Delta y, z_{i,j+1} - z_{i+1,j}) \end{cases}$$

其中, Δx 、 Δy 表示样方的单位长度; $\overrightarrow{P_{i,j}}$ 表示标准向量, 计算公式如下:

$$\overrightarrow{P_{i,j}} = [x_0 + (i-1)\Delta x, x_0 + (j-1)\Delta y, z_{i,j}]$$

其中, i 、 j 分别表示网格的长和宽; $z_{i,j}$ 为 $\overrightarrow{P_{i,j}}$ 在 z 轴的坐标。

坡度 (Q_{slop}) 定义为 $\overrightarrow{n_{i,j}}$ 与 z 轴间的夹角, 计算公式如下:

$$\begin{cases} u = \frac{z_a - z_b}{dis} = \frac{\sqrt{2}(z_a - z_b)}{2\Delta xy} \\ v = \frac{z_c - z_d}{dis} = \frac{\sqrt{2}(z_c - z_d)}{2\Delta xy} \\ Q_{slop} = \tan^{-1} \sqrt{u^2 + v^2} \end{cases}$$

其中, dis 为网格对角线的长度; Δxy 为网格的基本单元长度; u 为点 z_b 的倾斜角, v 为点 z_d 的倾斜角。

(3) 坡向的计算

坡向 (Q_{dir}) 定义为样方的单位法向量 $\overrightarrow{n_{i,j}}$ 在 xoy 平面上的投影与 x 轴的夹角, 并定义 x 轴指向的方向为南。当网格为正方形时, 样方的 Q_{dir} 定义如下:

$$\theta = \tan^{-1} \left(\frac{v}{u} \right) + \pi$$

当 $\mu < 0$ 时, 若 $v=0$, 那么 $\theta=0$, 此时 Q_{dir} 表示正南方向; 若 $v < 0$, 那么 $0 < \theta < 90^\circ$, 此时 Q_{dir} 表示西偏南方向; 若 $v > 0$, 那么 $-90^\circ < \theta < 0$, 此时 Q_{dir} 表示东偏南方向。

当 $\mu > 0$ 时, 若 $v=0$, 那么 $\theta=180^\circ$, 此时 Q_{dir} 表示正北方向; 若 $v < 0$, 那么 $90^\circ < \theta < 180^\circ$, 此时 Q_{dir} 表示北偏西方向; 若 $v > 0$, 那么 $180^\circ < \theta < 270^\circ$, 此时 Q_{dir} 表示北偏东方向。

当 $\mu=0$ 时, 若 $v < 0$, 那么 $\theta=90^\circ$, 此时 Q_{dir} 表示正西方向; 若 $v > 0$, 那么 $\theta=270^\circ$, 此时 Q_{dir} 表示正东方向。

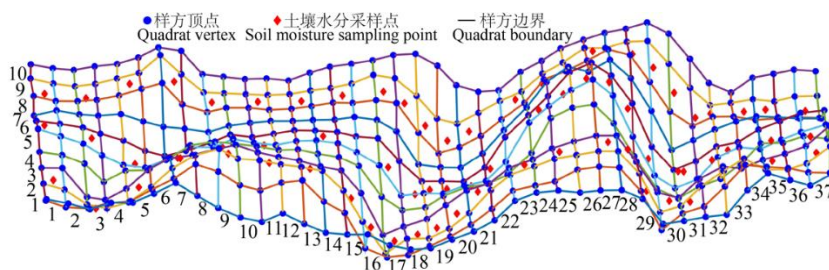


图1 样地示意图

Fig. 1 Sample plot diagram

1.3 研究方法

1.3.1 土壤水分划分依据 研究区草本植物根系主要分布在 0 cm~40 cm 深度的土层中, 约 80%的固沙灌木根系分布在 40 cm~200 cm 深度土层, 约 10%的灌木根系分布在 200 cm~300 cm 的土层范围内 (孙琰蕙等, 2022)。本研究将土层深度划分为表、中、深 3 层, 其中表层为 40 cm 以上土层, 中层为 40 cm 到 200 cm 深度的土层, 深层为 200 cm 到 300 cm 深度的土层, 分别对每层的土壤水分含量和影响因素进行分析。

1.3.2 偏最小二乘结构方程模型 偏最小二乘结构方程模型 (PLS-SEM) 能够通过建模对多个因变量和自变量之间的复杂关系进行估计。它克服了多元回归、逻辑回归等多变量分析方法只能分析一条因果关系链的局限性, 因为这些方法要求所有变量都可被观测且所有变量都无测量误差 (Hair *et al.*, 2021)。结构方程模型的主要优点是可以将不能被直接观测, 需要通过多种指标进行衡量的潜变量纳入模型, 从而能够测量复杂的模型关系并考虑指标中固有的测量误差。PLS-SEM 的重点在于最小化残差, 以最大限度提高内生潜变量的解释方差来估计模型间的关系。它由内部模型和外部模型组成, 可以有效处理小样本量和复杂模型。此外, 作为一种非参数方法, PLS-SEM 不需要进行分布假设 (Hair *et al.*, 2017)。

1.3.3 机器学习算法 随机森林基于 bagging 算法原理, 利用 Bootstrap 抽样从原始的训练样本集中随机抽取若干子集, 每个子集训练一颗决策树。在回归任务中取每棵决策树预测结果的平均值为最终预测值, 具有良好的抗过拟合能力和高准确性。XGBoost 是一种基于梯度提升的增强算法, 迭代地训练多个弱学习器, 每个新的学习器对前一个学习器的残差进行优化从而提高模型性能, 引入了 L_1 和 L_2 正则化, 防止过拟合, 增强了模型的泛化能力。SHAP 模型为机器学习模型的预测结果提供了可解释性, 基于计算每个特征边际贡献的 Shapley 值理论, 即计算该特征加入模型前后对预测结果的变化。SHAP 模型适用于线性回归、决策树、神经网络等任何类型的机器学习模型, 通过为每个特征分配一个 SHAP 值, 可以直观理解每个特征对预测结果的影响。

2 结果与分析

2.1 4种微地貌类型上土壤水分的分布特征

在不同微地貌类型上, 表层土壤水分的大小顺序为丘底 (1.21 ± 0.50)% > 丘顶 (0.81 ± 0.71)% > 背风坡 (0.71 ± 0.44)% > 迎风坡 (0.34 ± 0.19)%。背风坡、丘顶和迎风坡的土壤水分出现一个相对明显的峰值, 呈右偏态分布, 土壤水分出现较低值的概率更大, 丘底呈现出形态差异不明显的双峰分布 (图 2 (a))。中层 40 cm~200cm 的土壤水分的大小顺序为背风坡 (2.25 ± 1.18)% > 丘底 (2.23 ± 0.56)% > 丘顶 (1.57 ± 0.71)% > 迎风坡 (0.86 ± 0.35)%。迎风坡和丘底的土壤水分呈现明显较为集中的单峰分布, 丘顶的土壤水分分布出现两个峰值且呈现两端小中间大的分布, 显示其上土壤水分分布存在分化现象, 且出现极高和极低值的概率较小。背风坡上土壤水分的双峰分布不明显且分布具有延展性, 其上土壤水分的空间分布差异较大 (图 2 (b))。对深层土壤来说, 土壤水分含量在 4 种微地貌类型上的大小顺序为背风坡 (2.43 ± 0.88)% > 丘顶 (2.01 ± 0.54)% > 丘底 (1.98 ± 0.45)% > 迎风坡 (1.09 ± 0.55)%, 而迎风坡上的土壤水分呈现明显的右偏分布, 土壤水分出现较小值的概率更高 (图 2 (c))。不同深度土壤水分在 4 种微地貌类型上分布的变异系数 (表 1) 表明, 表层和中层土壤水分分布的差异较大, 变化较为剧烈, 而深层土壤水分波动较小, 分布较为稳定。丘底土壤水分在所有微地貌类型中的差异最小。

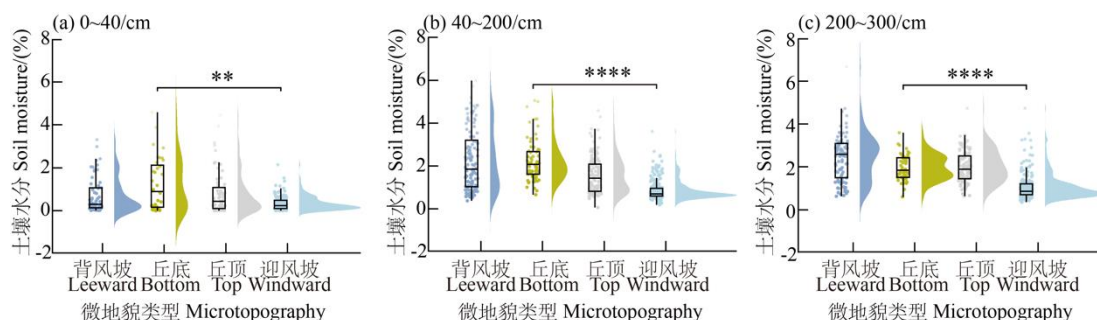


图 2 不同深度土壤水分在 4 种微地貌类型上的分布差异

Fig. 2 Differences in the distribution of soil moisture at different depths across four microtopographic types

表 1 不同微地貌类型上的表、中、深层土壤水分平均含量与变异系数

Tab. 1 Average content and coefficient of variation of surface, middle, and deep soil moisture on different microtopographic types

微地貌 Microtopography	0~40 cm		40~200 cm		200~300 cm	
	平均土壤水分 A	变异系数 Coeff	平均土壤水分 A	变异系数 Coeffi	平均土壤水分 A	变异系数 Coeffi
	verage soil mois	icient of variati	verage soil mois	cient of variatio	verage soil mois	cient of variatio
	ture (%)	on	ture (%)	n	ture (%)	n
背风坡 Leeward	0.71	0.62	2.25	0.52	2.43	0.36
丘底 Bottom	1.21	0.41	2.23	0.25	1.98	0.23
丘顶 Top	0.81	0.87	1.57	0.46	2.01	0.27
迎风坡 Windward	0.33	0.57	0.86	0.41	1.09	0.51

2.2 地形-植被因子对土壤水分的影响路径

将地形、灌木植被、草本植被和土壤水分视为潜在变量，9 个地形和植被指标以及表、中、深层土壤水分视为观测变量构建结构方程模型。利用最小二乘法(PLS)计算每条路径的路径系数并进行显著性检验，以探究对土壤水分的影响路径。研究假设地形既对土壤水分产生直接影响，又通过影响草本和灌木植被的分布格局对土壤水分产生间接影响，并假设灌木植被和草本植被之间存在相互作用间接影响土壤水分，且二者都能对土壤水分产生直接影响。

首先对外部模型进行检验，克伦巴赫 α 系数(Cronbach's α)通常用于衡量一组测量项之间的相关性或一致性，以确定它们是否在测量相同的概念或特征。它的取值范围从 0 到 1，越接近 1 表示测量项之间的一致性越高。通常，一个 α 系数高于 0.7 被认为是可接受的内部一致性水平，而高于 0.8 则被认为是较高的一致性水平(Tavakol *et al.*, 2011)。DG. rho 系数通常用于评估通过 PLS 构建的模型对外部指标的预测能力，即评估模型预测值与实际观测值之间的相关性。DG. rho 系数的取值范围通常为 0 到 1 之间。值越接近 1 表示模型对外部指标的预测能力越好，即预测值与实际观测值之间的相关性越高，通常 DG. rho 系数高于 0.7 被认为是可以接受的(Mehmetoglu, 2012)。第一特征根反映了外部模型的总预测能力，第一特征根越大表明模型对外部指标预测残差的方差较小，这意味着模型对外部指标的整体预测能力越强，能够较好地解释外部指标变异。第二特征根通常用于评估模型的过拟合情况(Dijkstra *et al.*, 2011)。第二特征根较小表明模型对外部指标的预测能力相对较好，并且模型没有过度拟合数据。检验结果(表 2)表明地形、土壤水分的 C. alpha 系数分别为 0.89 和 0.79，超过 0.7，4 个潜变量的 DG. rho 系数均达到 0.8 以上，且第一特征根均大于 1，第二特征根均小于 1，可以认为外部模型有效性较好。

表 2 外部模型有效性检验

Tab. 2 External model validity test

潜变量 Latent variable	C. alpha 系数	DG. rho 系数	第一特征根	第二特征根
	C. alpha coeffi	DG. rho coeffi	First eigenval	Second eigenvalu
	cient	cient	ue	e
地形 Topography	0.89	0.92	2.58	0.53
灌木植被 Shrub	0.61	0.84	1.43	0.56
草本植被 Herbaceous	0.66	0.80	2.09	0.99
土壤水分 Soil moisture	0.79	0.88	2.14	0.70

接下来通过计算模型的 GoF 值检验内部模型的解释能力，计算公式为

$$GoF = \sqrt{communality \times R^2}$$

其中, *communality* 表示观测变量被潜变量解释的平均程度, *GoF* 的取值在 0-1 之间。当 $GoF < 0.25$ 时, 模型解释能力较弱; 当 $0.25 < GoF < 0.36$ 时解释能力中等; 当 $GoF > 0.36$ 时模型解释能力较好 (Wetzels et al., 2009), 通过计算得到 $\overline{communality} = 0.56$, $\overline{R^2} = 0.61$, *GoF* 值为 0.58, 因此可以相信该模型的解释能力。

影响路径 (图 3) 结果表明地形对草本植被存在显著的正面影响, 对土壤水分存在显著的负面影响。草本植被和灌木植被对土壤水分都存在显著的负向影响, 可能是由于植被生长过程中对土壤水分的消耗。同时, 灌木植被对草本植被具有显著的负面影响, 可能是由于灌木植被与草本植被之间存在水分和空间等的竞争关系, 因此灌木植被会对草本植被的分布和数量产生影响, 从而间接影响土壤水分。对土壤水分影响的效应程度最大的为地形 ($\beta = -0.636$), 其次为草本植被 ($\beta = -0.329$), 灌木植被影响的效应程度最小 ($\beta = -0.168$)。其中, 地形中的坡度和高差对土壤水分的影响最显著 ($P < 0.01$), 草本植被中的草本盖度和草本多度对土壤水分的影响最显著 ($P < 0.001$), 灌木植被中的灌木盖度对土壤水分的影响最显著 ($P < 0.001$)。

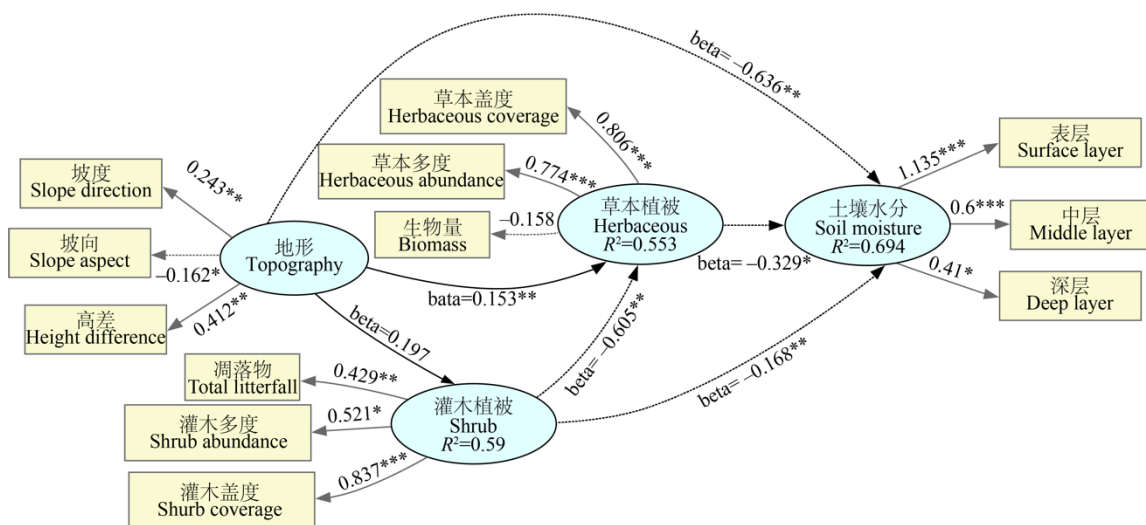


图 3 土壤水分与不同地形植被变量间的影响路径

Fig. 3 Pathways of influence between soil moisture and different topographic vegetation variables

2.3 基于 SHAP 模型的土壤水分影响因素分析

采用随机森林和 XGBoost 两种机器学习算法构建模型, 采用网格搜索和 5 折交叉验证对模型的超参数进行调优, 调优后的参数如表 3、4 所示。随机森林模型对表层、中层和深层土壤水分拟合的 R^2 分别为 0.84、0.70 和 0.83, XGBoost 模型对表层、中层和深层土壤水分拟合的 R^2 分别为 0.68、0.61、0.59, 随机森林模型的拟合优度更高。

表 3 随机森林模型最优参数

Tab. 3 Optimal parameters of random forest

参数 Parameter	取值 Value
n_estimators	49
min_samples_split	6
min_samples_leaf	1
random_state	42

表 4 XGBoost 模型最优参数

Tab. 4 Optimal parameters of XGBoost model

参数 Parameter	取值 Value
n_estimators	87
learning_rate	0.1

subsample	0.05
colsample_bytree	0.8
max_depth	5
reg_alpha	0
reg_lambda	1

分别使用基于随机森林和 XGBoost 的 SHAP 模型根据变量对土壤水分的重要性进行排序, SHAP 值的绝对值大小反映了变量对土壤水分影响的强弱, 绝对值越大变量对土壤水分的影响越强。SHAP 值为正表明变量对土壤水分变动具有正面影响。从变量取值大小来看, 变量值高呈现红色, 变量值低呈现蓝色。由于随机森林模型的拟合优度更高, 选择随机森林模型对变量的重要性进行排序, 综合图 4、图 5 的排序结果表明各变量对不同深度土壤水分影响的重要性并不完全相同, 但综合排序结果得到: 草本盖度、坡向和高差对表层土壤水分的贡献程度较大; 坡向、草本盖度和生物量对中层土壤水分的贡献程度较大; 灌木盖度、坡向和草本盖度对深层土壤水分的贡献程度较大, 其中生物量和灌木盖度对土壤水分的影响相似, 变量取值较高对土壤水分产生负面影响; 坡向和高差对土壤水分的影响相似, 变量取值较高对土壤水分产生正面影响。草本盖度呈现出的影响能力较为复杂, 但变量取值较高时对土壤水分呈现负面影响。

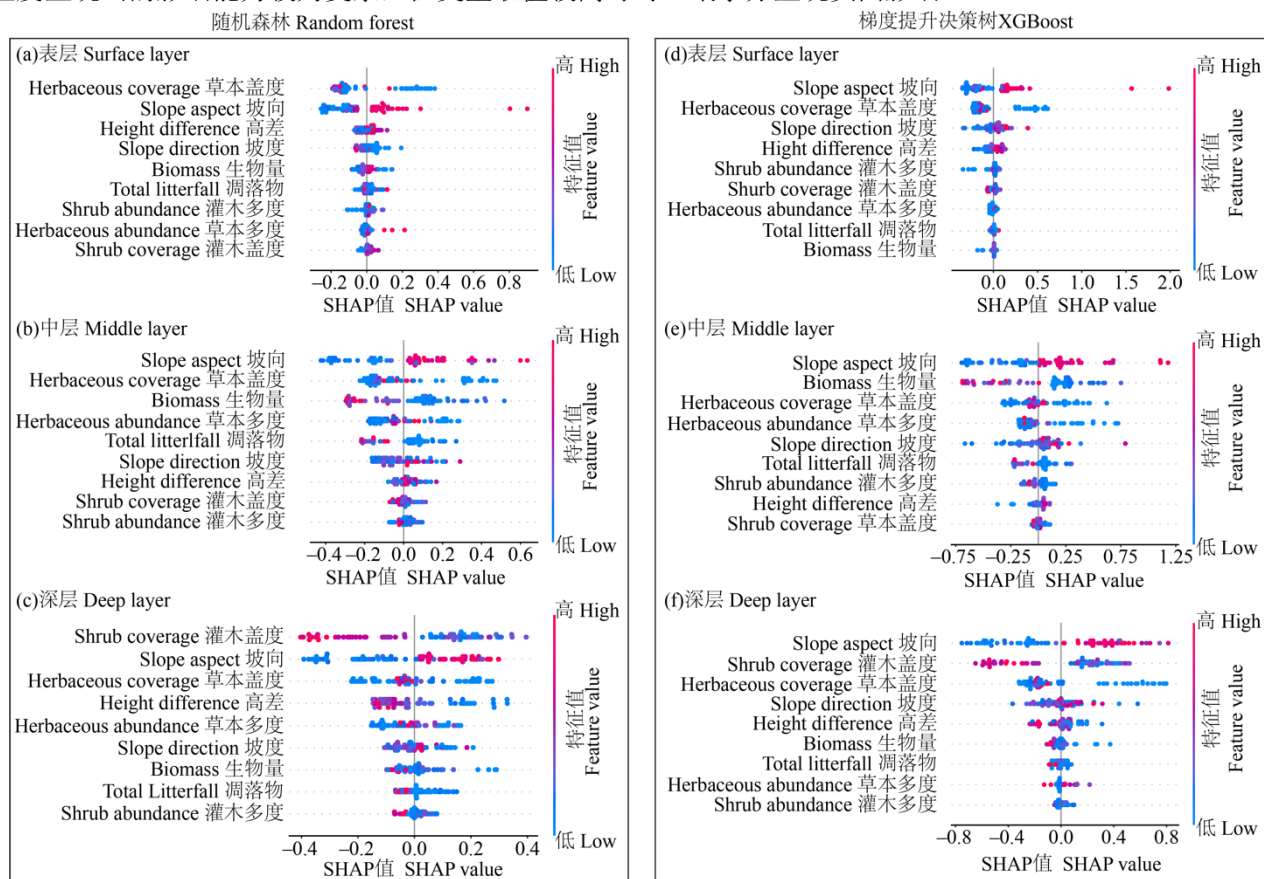


图 4 基于随机森林和 XGBoost 的 SHAP 解释图

Fig. 4 SHAP explanation plots based on random forest and XGBoost

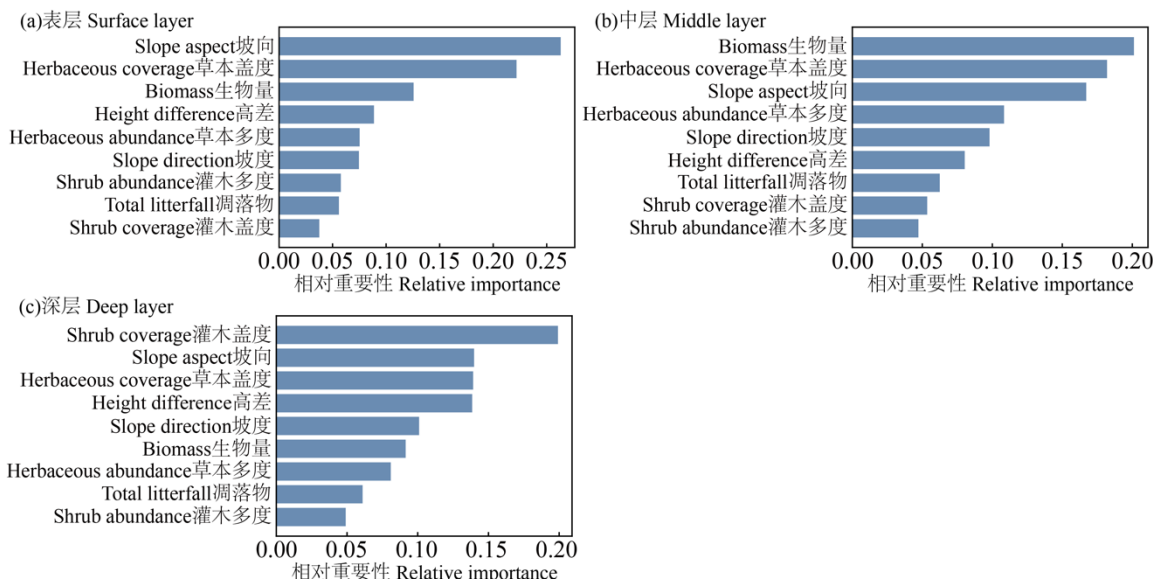


图 5 随机森林模型变量重要性排序

Fig. 5 Variable importance ranking in random forest model

图 6 进一步展示了变量对土壤水分的影响规律。横坐标表示变量的取值，左侧纵坐标表示该变量的 SHAP 值，右侧纵坐标表示与该变量交互作用最大的变量的取值。草本盖度在 0 附近时，对土壤水分既可能产生正面影响也可能产生负面影响，这可能是由于草本盖度较低时的影响作用不明显。随着草本盖度增加，对土壤水分表现出负面影响，当草本盖度超过 40% 时对土壤水分表现出正面影响，坡向低于 200 rad 时对表层土壤水分产生负面影响，高于 200 rad 时对表层土壤水分产生正面影响；低于 100 rad 时对中层和深层土壤水分产生负面影响，高于 100 时产生正面影响。高差低于 3 m 时对表层土壤水分产生负面影响，高于 3 m 时产生正面影响。生物量低于 30 g 时对中层土壤水分产生正面影响，高于 30 g 时产生负面影响。灌木盖度低于 40% 时对深层土壤水分产生正面影响，高于 40% 时产生负面影响，且随灌木盖度增加负面影响增大。也就是说，不同深度土壤水分均随草本盖度和坡向增加呈先下降后上升趋势，表层土壤水分随高差增加呈先下降后上升趋势，中层和深层土壤水分分别随生物量和灌木盖度增加呈先上升后下降趋势。

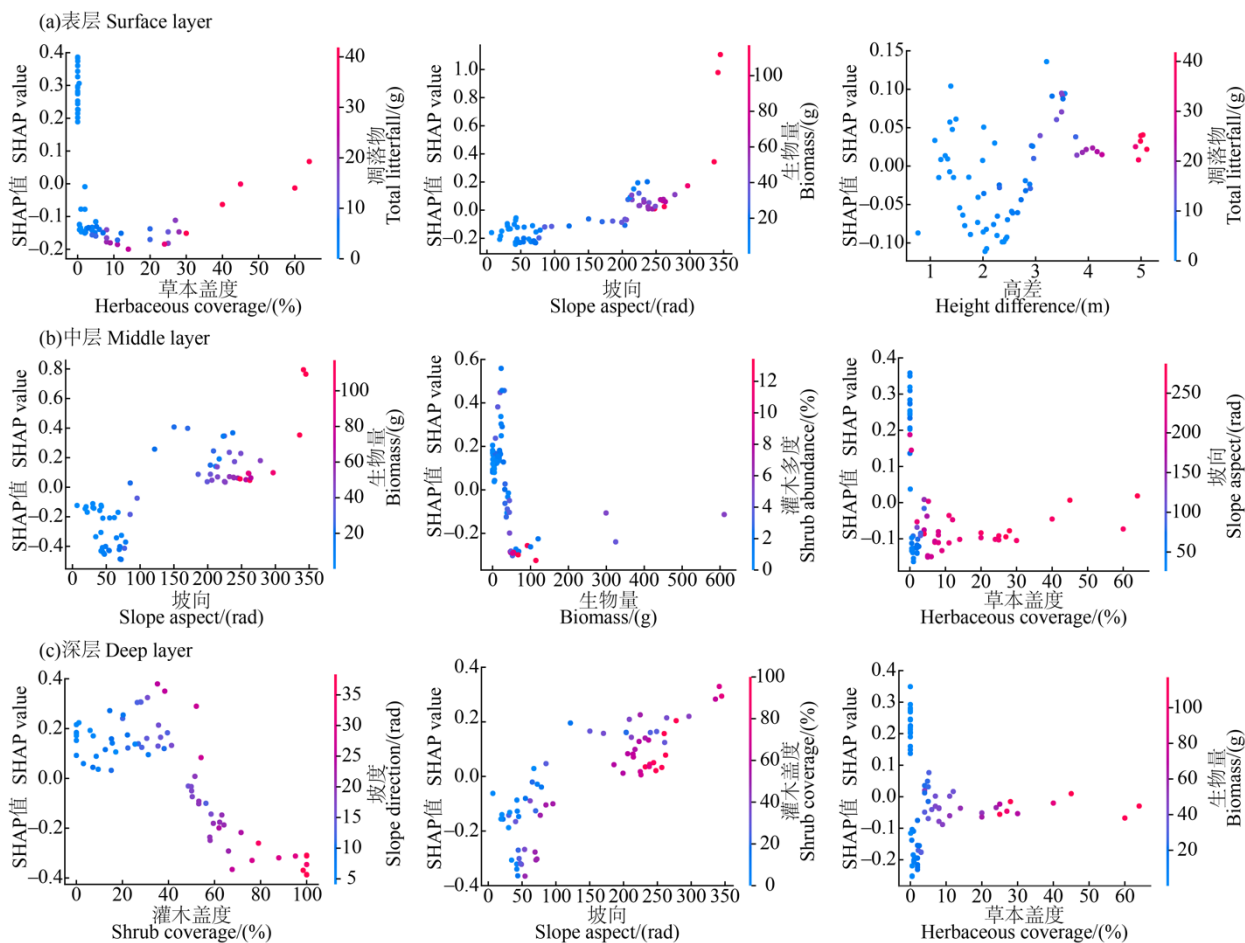


图 6 基于随机森林的 SHAP 特征依赖图

Fig. 6 SHAP Feature Dependence Plot Based on Random Forest

选取表层土壤水分绘制瀑布图（图 7），进一步解释了 4 种微地貌类型上单个样本土壤水分变动的影响因素。草本盖度、坡向和高差对迎风坡土壤水分影响较大，草本多度为 48%、坡向为 262.195 rad、高差为 5.109 m 时对迎风坡土壤水分产生正向影响，生物量为 141.212 g 时对迎风坡土壤水分产生负向影响。坡向、草本盖度和高差对丘底土壤水分影响较大，坡向为 277.839 rad、高差为 3.554 m、凋落物为 2.652 g 时对丘底土壤水分产生正面影响，草本盖度为 4% 时对丘底土壤水分产生负面影响。草本盖度、生物量和凋落物对背风坡土壤水分影响较大，草本盖度为 25%、凋落物为 16.724 g 时对背风坡土壤水分产生负面影响，生物量为 324.704 g、坡向为 247.875 rad 时对背风坡土壤水分产生正面影响。坡向、生物量和草本多度对丘底土壤水分影响较大，坡向为 345.117 rad、生物量为 141.212 g、草本多度为 92%、草本盖度为 64% 时对丘顶土壤水分产生正面影响。瀑布图反映了单个样本的影响结果，每个样本的变量重要性顺序并不相同，与总体的变量重要性顺序也不完全相同。结果进一步验证图 6：草本盖度较低（低于 40%）时对土壤水分产生负面影响（图 7（b, c）），草本盖度较高时对土壤水分产生正面影响（图 7（d）），坡向高于 200 时对土壤水分产生正面影响（图 7（a, b, c, d）），高差超过 3m 时对土壤水分产生正面影响（图 7（a, b））。

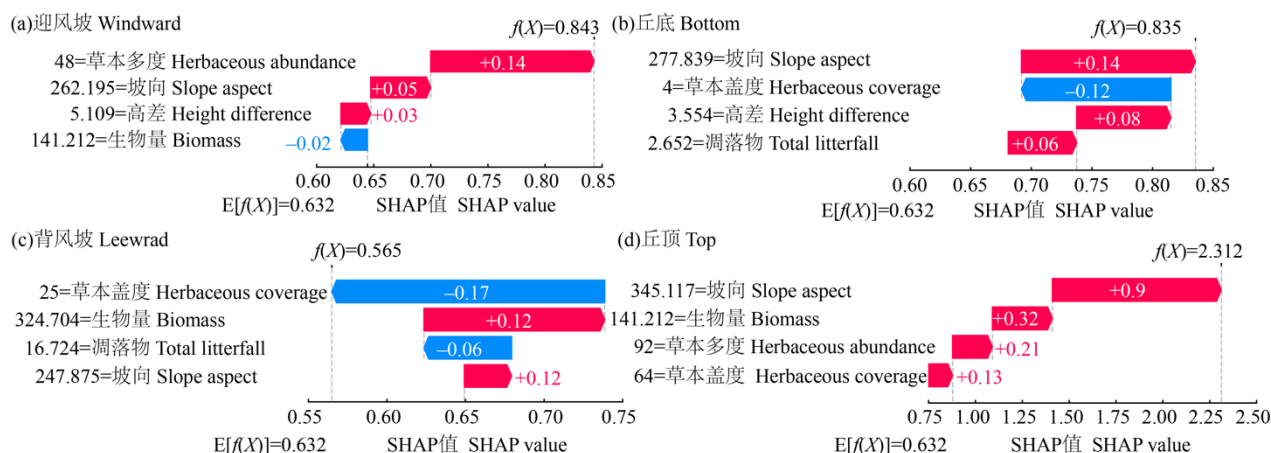


图 7 4 种微地貌上土壤水分影响瀑布图
Fig. 7 Waterfall plot of soil moisture impacts on four microtopographies

3 结论与建议

3.1 4 种微地貌类型上土壤水分的分布特征分析

对 4 种微地貌类型上土壤水分的统计分析得到结论：背风坡、丘底的土壤水分较迎风坡和丘顶更高背风坡土壤水分较迎风坡更高可能是由于受风的影响，迎风坡的土壤水分较背风坡蒸发更快。且由于风力和降水冲刷，迎风坡上土壤可能更为疏松，不利于土壤水分保持。而丘底的土壤水分较丘顶更高可能由于不同微地貌上形成的生物土壤结皮厚度不一致。生物土壤结皮的发育能够有效提高土壤持水能力，土壤生物结皮越厚的地方更容易保持更高的土壤水分（张志山等, 2007）。由于重力作用，丘底更有利于降水积聚，更有利于保持土壤水分（马风云等, 2006; 李新荣等, 2013）。

表层和深层土壤水分分布的差异较大，变化较为剧烈，而深层土壤水分波动较小，分布较为稳定。丘底土壤水分在所有微地貌类型中的差异最小。这可能是由于研究区内土壤质地疏松，表层土壤水分受到风沙作用的影响更明显。丘底既接受直接降水，还能够接受从周围高地流入的径流，且由于地势低洼，水分在丘底的停留时间较长，蒸发和径流的流失相对较少，有利于水分的积累和保存。

在防风固沙工作中，需要对半固定沙丘上的植被进行科学管理，栽植固沙植被并对天然植被进行恢复。针对由于风沙活动导致丘顶和迎风坡上流沙运动、土壤水分低、固沙植被难以存活的问题，可以在丘顶和迎风坡设置更加密集的草方格网和沙障，在丘底和背风坡设置较稀疏的草方格网固定沙面，防止半固定沙丘退化为流动沙丘。针对不同深度土壤水分的波动状况不同，可以构建多层次的植被结构，在表层和深层土壤水分变化较大的区域，种植柠条、沙棘等耐旱的草本植物和浅根灌木，以快速覆盖地表，减少水分蒸发和土壤侵蚀。在能够扎根到深层土壤的位置，种植梭梭等深根灌木，利用深层稳定的水分，提高植被的整体抗旱能力。

3.2 地形-植被变量影响不同深度土壤水分的原因分析

SHAP 模型的分析结果表明土壤水分随草本盖度、坡向和高差增加呈先下降后上升趋势，这可能是由于随着草本植物盖度增加，会对土壤表面产生一定的遮阴作用，减少太阳辐射直接照射到土壤表面，从而减少水分蒸发损失。但在初期阶段，这种遮阴效应可能还不够显著，土壤水分仍会因为植物蒸腾作用而减少。当草本植被盖度达到一定程度后，沙漠生态系统环境得到改善，水分保持能力增强，使得土壤水分回升。坡向增加表示由西南坡逐渐转向东北坡。西南坡受到更强烈的阳光直射，温度较高、蒸发量较大，而东北坡太阳辐射强度和温度相对较低，蒸发量小，土壤水分保留较好。在高差增加的初始阶段，土壤水分由于重力作用和地表径流迅速流失，导致地势高处土壤水分快速减少。随着高差继续增加，水流逐渐在地势低洼处汇集，使低洼处土壤水分升高。在不利于保持土壤水分的高坡和阳坡应选择耐旱性更强的植物种植。

土壤水分随灌木盖度和生物量增加呈先上升后下降趋势可能是由于灌木盖度和生物量增加提供遮阴效果，减少了直射到土壤表面的太阳辐射强度，降低了土壤水分蒸发速率（Ozkan *et al.*, 2017）。灌木的存在也降低了风蚀作用，从而进一步减少土壤水分的蒸发，土壤水分含量上升。当灌木盖度和生物量继续增加时，

植被之间的竞争加剧,尤其是在水资源有限的沙漠环境中,导致土壤水分减少 (Breshears *et al.*, 1999)。在未来的植物固沙和生态恢复过程中,需要进行适当的修枝和间伐,减少植被体积和密度过大对土壤水分的消耗。

参 考 文 献

- Anderson M C, Kustas W P, Norman J M. 2003. Upscaling and downscaling: A regional view of the soil-plant-atmosphere continuum. *Agronomy Journal*, 95(6): 1408-1423.
- Asbjornsen H, Ashton M S, Vogt D J, et al. 2004. Effects of habitat fragmentation on the buffering capacity of edge environments in a seasonally dry tropical oak forest ecosystem in Oaxaca, Mexico [J]. *Agriculture, Ecosystems & Environment*, 103(3): 481-495.
- Asbjornsen H, Goldsmith G R, Alvarado-barrientos M S, et al. 2011. Ecohydrological advances and applications in plant-water relations research: a review. *Journal of Plant Ecology*, 4(1-2): 3-22.
- Biederman L A, Whisenant S G. 2011. Using mounds to create microtopography alters plant community development early in restoration. *Restoration Ecology*, 19(101): 53-61.
- Breshears D D, Barnes F J. 1999. Interrelationships between plant functional types and soil moisture heterogeneity for semiarid landscapes within the grassland/forest continuum: a unified conceptual model. *Landscape Ecology*, 14: 465-478.
- Dijkstra T K, Henseler J. 2011. Linear indices in nonlinear structural equation models: Best fitting proper indices and other composites. *Quality & Quantity*, 45: 1505-1518.
- D'Odorico P, Laio F, Porporato A, et al. 2010. Ecohydrology of terrestrial ecosystems. *BioScience*, 60(11): 898-907.
- 郭艳菊, 马晓静, 许爱云, 等. 2022. 宁夏东部风沙区沙化草地土壤水分和植被的空间特征. *生态学报*, 42(4): 1571-1581.
- (Guo Y J, Ma X J, Xu A Y, et al. 2022. Spatial characteristics of soil moisture and vegetation in desertification grassland in eastern Ningxia sand-wind area, China. *Acta Ecologica Sinica*, 42(4): 1571-1581. [in Chinese])
- Hair Jr J F, Hult G T M, Ringle C M, et al. 2021. Partial least squares structural equation modeling (PLS-SEM) using R: A workbook. Springer Nature.
- Hair Jr J F, Matthews L M, Matthews R L, et al. 2017. PLS-SEM or CB-SEM: updated guidelines on which method to use. *International Journal of Multivariate Data Analysis*, 1(2): 107-123.
- Liu Y, Cui Z, Huang Z, et al. 2019. Influence of soil moisture and plant roots on the soil infiltration capacity at different stages in arid grasslands of China. *Catena*, 182: 104147.
- Li X Y, Yang Z P, Lin Y T, et al. 2009. Connecting ecohydrology and hydropedology in desert shrubs: Stemflow as a source of preferential flow in soils. *Hydrol Earth Syst Sci*, 13(7): 1133-1144.
- 李新荣, 张志山, 黄磊, 等. 2013. 我国沙区人工植被系统生态-水文过程和互馈机理研究评述. *科学通报*, 58(Z1): 397-410.
- (Li X R, Zhang Z S, Huang L, et al. 2013. Review of the ecohydrological processes and feedback mechanisms controlling sand-binding vegetation systems in sandy desert regions of China. *Chinese Sci Bull*, 58(Z1): 397-410. [in Chinese])
- 李新荣, 张志山, 谭会娟, 等. 2014. 我国北方风沙危害区生态重建与恢复: 腾格里沙漠土壤水分与植被承载力的探讨. *中国科学: 生命科学*, 44(3): 257-266.
- (Li X R, Zhang Z S, Tan H J, et al. 2014. Ecological restoration and recovery in the wind-blown sand hazard areas of northern China: relationship between soil water and carrying capacity for vegetation in the Tengger Desert. *Sci China Life Sci*, 44(3): 257-266. [in Chinese])
- 马风云, 李新荣, 张景光, 等. 2006. 沙坡头人工固沙植被土壤水分空间异质性. *应用生态学报*, 17(5): 789-795.
- (Ma F Y, Li X R, Zhang J G, et al. 2006. Spatial heterogeneity of soil moisture in Shapotou sand-fixing artificial vegetation area. *Chinese Journal of Applied Ecology*, 17(5): 789-795. [in Chinese])
- Mehmetoglu M. 2012. Partial least squares approach to structural equation modeling for tourism research. *Advances in Hospitality and Leisure*. Emerald Group Publishing Limited.
- Moeslund J E, Arge L, Bøcher P K, et al. 2013. Topography as a driver of local terrestrial vascular plant diversity patterns. *Nordic Journal of Botany*, 31(2): 129-144.
- Ozkan U, Gökbulak F. 2017. Effect of vegetation change from forest to herbaceous vegetation cover on soil moisture and temperature regimes and soil water chemistry. *Catena*, 149: 158-166.
- Perring F. 1959. Topographical gradients of chalk grassland. *The Journal of Ecology*, 447-481.
- Potts D L, Scott R L, Bayram S, et al. 2010. Woody plants modulate the temporal dynamics of soil moisture in a semi-arid

d mesquite savanna. *Ecohydrology*, 3(1): 20-27.

Qiu Y, Fu B, Wang J, et al. 2001. Soil moisture variation in relation to topography and land use in a hillslope catchment of the Loess Plateau, China. *Journal of Hydrology*, 240(3): 243-263.

孙琰蕙, 张定海, 张志山. 2022. 腾格里沙漠不同类型沙丘土壤水分含量与地形-植被因子关系研究. *干旱区地理*, 45(5): 1570-1578.

(Sun Y H, Zhang D H, Zhang Z S. 2022. Relationship between soil moisture content and topography-vegetation factors in different types of dunes in the Tengger Desert. *Arid Land Geography*, 45(5): 1570-1578. [in Chinese])

Tavakol M, Dennick R. 2011. Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2: 53.

Wang C, Fu B, Zhang L, et al. 2019. Soil moisture-plant interactions: an ecohydrological review. *Journal of Soils and Sediments*, 19(1): 1-9.

Western A W, Zhou S L, Grayson R B, et al. 2004. Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes. *Journal of Hydrology*, 286(1-4): 113-134.

Wetzels M, Odekerken-Schröder G, Van Oppen C. 2009. Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *MIS Quarterly*, 177-195.

张世才, 王慧娟, 张定海, 等. 2023. 腾格里沙漠东南缘 3 种沙丘 4 种微地貌上土壤水分与地形-植被因子之间的关系. *甘肃农业大学学报*, 58(3): 160-168, 179.

(Zhang S C, Wang H J, Zhang D H, et al. 2023. Relationship between soil moisture and topography-vegetation factors on three types of sand dunes and four types of microgeomorphology in the southeastern margin of Tengger Desert in northern China. *Journal of Gansu Agricultural University*, 58(3): 160-168+179. [in Chinese])

张志山, 何明珠, 谭会娟, 等. 2007. 沙漠人工植被区生物结皮类土壤的蒸发特性——以沙坡头沙漠研究试验站为例. *土壤学报*, 44(3): 404-410.

(Zhang Z S, He M Z, Tan H J, et al. 2007. Evaporation from soils covered with biological crusts in revegetated desert: A case study in Shapotou desert research and experiment station. *Acta Pedologica Sinica*, 44(3): 404-410. [in Chinese])