Journal of Geoscience and Earth Observation
http://dx.doi.org/10.62762/JGEO.2026.171666

RESEARCH ARTICLE

EWLR - A New Method for Interpolating Elevation-Driven Variables: Annual Rainfall in Erbil Governorate

Azad Rasul[1,2,*]

[1] Department of Geography, Faculty of Arts, Soran University, Soran 44008, Iraq
[2] Department of Forestry, College of Agricultural Engineering Sciences, Salahaddin University-Erbil, Erbil 44002, Iraq

Abstract

Accurate spatial estimation of rainfall is critical for hydrological modeling, water resource management, and agricultural planning—particularly in mountainous and semi-arid regions with sparse monitoring networks. This study presents an Enhanced Elevation-Weighted Local Regression (EWLR) model to generate a high-resolution (30 m) annual rainfall surface for Erbil Governorate, northern Iraq. The EWLR model integrates distance weighting, elevation similarity weighting, and orographic enhancement within a locally weighted regression framework. Average annual rainfall, derived from rainy seasons spanning 1997–1998 to 2024–2025 across 19 meteorological stations, along with a 30 m resolution digital elevation model (DEM), were used to construct and validate the model. Hyperparameters were optimized via Leave-One-Out Cross-Validation (LOOCV), and performance was benchmarked against conventional methods including Inverse Distance Weighting (IDW), Kriging, Thin-Plate Spline, and Radial Basis Function interpolation. Results indicate that EWLR outperforms all benchmarks, achieving $R^2 = 0.797$, RMSE $= 120.9$ mm, and MAE $= 87.46$ mm. Rainfall shows a strong positive correlation with elevation ($r = 0.907$, $p < 0.001$), increasing nearly fivefold from lowland plains (∼270 mm) to mountainous areas (>1,300 mm). The final high-resolution rainfall map captures orographic effects reasonably well, providing a physically consistent, statistically robust dataset suitable for hydrological, climatic, and environmental modeling in data-sparse mountainous regions. The methodology offers a reproducible, elevation-centric framework adaptable to other elevation-driven variables (e.g., temperature lapse rates or snow accumulation) and complex terrains with limited observations.

Keywords: rainfall interpolation, Elevation-Weighted Local Regression (EWLR), orographic enhancement, Kurdistan Region, mountainous terrain, spatial precipitation mapping, hydrological modeling.

1 Introduction

Precipitation is a fundamental climatic parameter that shapes hydrological processes, agricultural productivity, and water resource management—especially in semi-arid and

Submitted: 12 June 2026
Revised: 14 July 2026
Accepted: 03 September 2026
Published: 24 September 2026

Vol. 1, **No.** 2, 2026.
10.62762/JGEO.2026.171666

***Corresponding authors:**
✉ Azad Rasul
azad.rasul@soran.edu.iq

Citation
Rasul, A. (2026). EWLR - A New Method for Interpolating Elevation-Driven Variables: Annual Rainfall in Erbil Governorate. *Journal of Geoscience and Earth Observation*, 1(2), 158–171.

mountainous regions where variability is pronounced [1]. In highly complex terrain, rainfall distribution is strongly influenced by topographic factors such as elevation, slope, and orographic uplift. For instance, topography exerts a significant influence on precipitation intensity and spatial distribution through orographic lifting of moist air masses [2]. In the context of northern Iraq within the foothills of the Zagros Mountains, rainfall exhibits marked spatial variability largely tied to elevation gradients [3]. Studies show that rainfall in the northern region of Iraq exceeds that in lower-lying areas, aligning with orographic influences in mountainous terrain [4]. Such variability is critical to capture when conducting water balance estimation, drought monitoring, land-use planning, and other environmental modeling tasks. Yet, ground-based meteorological station networks are often too sparse or unevenly distributed to represent fine-scale spatial variation—particularly in steep and inaccessible high-elevation zones.

Spatial interpolation techniques have long been employed to generate continuous rainfall surfaces from point observations. Common approaches include Inverse Distance Weighting (IDW), various forms of Kriging, and spline or radial basis function (RBF) interpolation. These methods can perform acceptably in flat or gently undulating terrain with dense station coverage. However, in mountainous regions their performance is often degraded due to the neglect of explicit terrain–rainfall relationships. For example, regression analyses reveal that precipitation–elevation correlation can be moderate to strong, but purely distance-based interpolation may not capture orographic enhancement well. Moreover, incorporation of elevation or terrain covariates has been shown to improve interpolation accuracy in data-sparse mountainous regions [5].

In Iraq more generally—and within the Kurdistan Region in particular—there remains a paucity of high-spatial-resolution rainfall surfaces that explicitly incorporate topographic effects. Many existing products rely on coarse grid resolutions (>1 km) or do not sufficiently account for orographic enhancement, limiting their use in detailed hydrological or ecological modeling in mountainous zones. Based on recent analyses of Iraqi rainfall distribution, it is evident that precipitation declines southward and that elevation remains a meaningful control [6].

To address these challenges, this study presents an Enhanced Elevation-Weighted Local Regression (EWLR) model for generating annual rainfall surfaces at 30 m spatial resolution for the Erbil Governorate region of northern Iraq. The model combines distance weighting, elevation similarity weighting, and orographic enhancement within a locally weighted regression framework. It is further optimized via Leave-One-Out Cross-Validation (LOOCV) to ensure robust performance across sparsely monitored terrain. While demonstrated here for rainfall interpolation, the EWLR framework is extensible to other elevation-dependent climatic variables, such as temperature, by adapting the orographic boosting factor to account for lapse rates or thermal gradients. The objectives of the research are to:

1. Develop and implement a high-resolution elevation-weighted interpolation model to estimate spatial rainfall distribution across Erbil Governorate, Kurdistan Region of Iraq.

2. Benchmark the EWLR model against conventional interpolation methods (IDW, Kriging, TPS, RBF).

3. Quantify the elevation–rainfall relationship and assess orographic influence using a high-resolution (30 m) digital elevation model (DEM).

4. Produce a validated, physically consistent rainfall map suitable for hydrological and climatic applications in mountainous terrain.

The novelty of the work rests in the explicit integration of topographic similarity and an adaptive orographic boosting factor within a locally weighted regression framework, paired with rigorous hyperparameter optimization and high-resolution raster output in a region that has lacked such detailed modeling. By doing so, this approach offers a computationally efficient and physically interpretable alternative to more conventional interpolators and provides rainfall estimates that reflect known elevation–rainfall gradients—while its modular design supports adaptation for variables like temperature in future extensions. Ultimately, the outcomes of this study enhance our understanding of spatial rainfall variability in northern Iraq and provide a reproducible geospatial framework applicable to other mountainous regions with sparse station networks. The high-resolution rainfall surface generated here supports fine-scale environmental and hydrological modeling efforts—including runoff estimation, drought monitoring, and ecosystem assessments—and lays groundwork for broader climatic variable mapping, such as temperature surfaces for agro-meteorological planning.

2 Methodology

2.1 Study Area

The study was conducted in the Erbil Governorate, Kurdistan Region of Iraq, encompassing approximately 15,000 km^2 within the northwestern Zagros mountain foothills, bounded by latitudes 35.5°–37.0° N and longitudes 43.5°–44.9° E (Figure 1). This region exemplifies a transitional landscape, featuring undulating alluvial plains in the southwest—ideal for agriculture and urbanization—contrasting with steep, dissected highlands in the northeast, where rugged anticlines and valleys dominate the terrain [7]. Elevations range from ∼260 m above sea level (a.s.l.) in the fertile low-lying plains along the Little Zab River to ∼3,600 m a.s.l. in the northeastern peaks of the High Folded Zone, part of the Simply Folded Belt of the Zagros orogen, shaped by Miocene–Pliocene compression and subsequent fluvial incision [8].

Climatically, Erbil Governorate lies in a semi-arid to dry-summer Mediterranean regime (Köppen BSk/BSh in the plains grading to Csa in the highlands), with annual precipitation varying sharply from 250–400 mm in the plains to over 1,000 mm in the highlands due to orographic enhancement of winter westerly moist air masses from the Mediterranean. The resulting hydroclimatic gradients render the region highly sensitive to interannual precipitation deficits, with recurrent drought episodes over recent decades documented to have substantially degraded vegetation cover and elevated land surface temperatures across the Kurdistan Region [9]. Mean annual temperatures range from 10–15°C in winter to 30–40°C in summer, supporting diverse land uses: irrigated croplands and orchards in the valleys, transitioning to sparse oak woodlands and rangelands on slopes, and alpine meadows in the upper reaches. Hydrologically, the governorate drains via the Greater and Little Zab Rivers into the Tigris, with episodic flash floods underscoring the need for high-resolution precipitation mapping amid climate variability and sparse monitoring [7, 10]. This topographic and climatic heterogeneity renders Erbil a compelling testbed for elevation-aware spatial interpolation, addressing gaps in hydrological modeling for water-scarce, mountainous semi-arid environments.

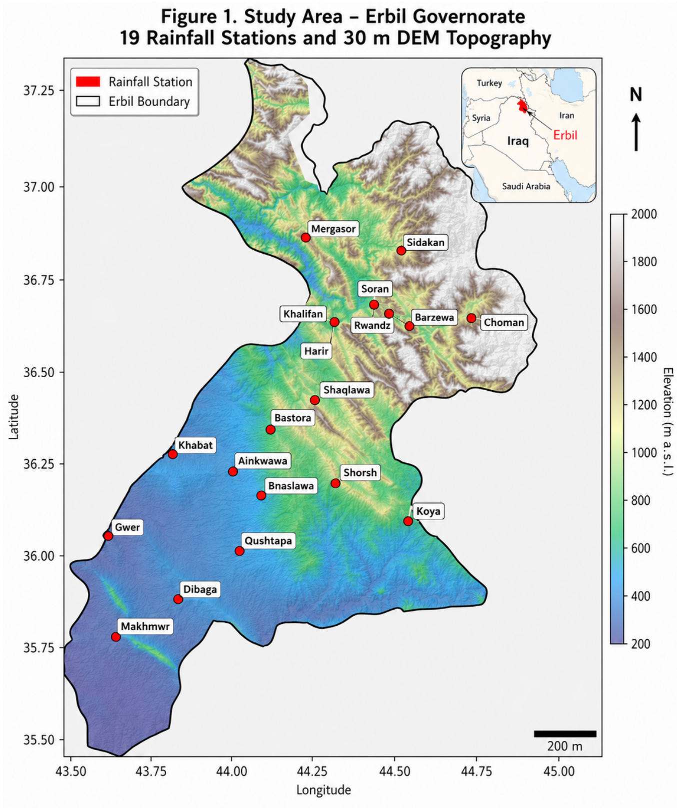

Figure 1. Overview of the Erbil Governorate study area, including the 19 rainfall stations and 30 m DEM-derived topography.

2.2 Data Sources

Average annual rainfall, derived from rainy seasons spanning 1997–1998 to 2024–2025, was compiled from 19 meteorological stations with complete records, operated by the Kurdistan Region Meteorological Directorate. Station locations were georeferenced and reprojected from WGS84 (EPSG:4326) to UTM Zone 38N (EPSG:32638) for distance computations. UTM extents ranged from easting 364,796–489,343 m to northing 3,959,937–4,077,351 m, ensuring comprehensive coverage across the topographic gradient.

Field-verified station elevations ranged from 264 m a.s.l. (Makhmwr) to 1,187 m a.s.l. (Mergasor), with annual rainfall varying from 257 mm (Makhmwr) to 1,384 mm (Mergasor), highlighting strong orographic influence (Table 1). The elevation distribution exhibited moderate sampling bias toward mid-elevations and a critical gap above 1,200 m a.s.l. (0 stations), justifying terrain-explicit interpolation methods:

- 0–300 m: 2 stations (Makhmwr, Khabat)
- 300–600 m: 5 stations (Qushtapa, Ainkawa, Gwer,

Bnaslawa, Dibaga)

- 600–900 m: 8 stations (Shorsh, Harir, Khalifan, Soran, Rwandz, Bastora, Barzewa, Koya)
- 900–1,200 m: 4 stations (Shaqlawa, Choman, Mergasor, Sidakan)
- >1,200 m: 0 stations

Table 1. Characteristics of the rainfall stations in Erbil Governorate.

Station	Long. (°E)	Lat. (°N)	Elev. (m)	Rain. (mm)
Qushtapa	44.0319	36.0039	397	350
Ainkawa	44.0119	36.2206	420	414
Khabat	43.67039	36.2758	286	339
Bnaslawa	44.1124	36.15444	501	381
Shorsh	44.386	36.1786	720	566
Shaqlawa	44.3214	36.4056	961	810
Harir	44.355	36.5475	714	591
Khalifan	44.4006	36.6039	676	768
Soran	44.5428	36.6581	663	701
Rwandz	44.5339	36.6169	802	736
Choman	44.8808	36.6278	1103	796
Sidakan	44.6711	36.7983	1006	824
Mergasor	44.3008	36.84	1187	1384
Makhmwr	43.57887	35.77511	264	257
Bastora	44.16148	36.33709	622	457
Dibaga	43.808	35.8747	341	299
Gwer	43.499	36.048	310	269
Barzewa	44.611	36.626	625	700
Koya	44.617	36.083	603	622

2.2.1 *Digital Elevation Model (DEM)*

A 30 m resolution Digital Elevation Model (DEM) was derived from the ALOS World 3D — 30 m (AW3D30) dataset, clipped to the Erbil Governorate study domain. The DEM was reprojected from its native geographic coordinates to WGS84 / UTM Zone 38N (EPSG:32638) to facilitate distance-based computations. NoData values were masked, yielding 11,285,644 valid pixels spanning an elevation range of 0–3,597 m a.s.l. All spatial operations were implemented using `rasterio` and `NumPy` in Python 3.12.13, ensuring efficient handling of the ∼11.3 million-pixel grid for interpolation.

2.3 Enhanced Elevation-Weighted Local Regression (EWLR)

An EWLR model was developed to generate a continuous 30 m rainfall surface, building on locally weighted regression frameworks tailored for orographically complex terrains [11]. The method integrates distance weighting, elevation similarity weighting, and orographic enhancement within a local linear regression framework.

For each target pixel (x_t, y_t, z_t), the predicted rainfall $\hat{r}_t$ is computed as

$$\hat{r}_t = \beta(z_t)\, z_t + \alpha, \tag{1}$$

where β and α are derived from weighted least squares using the 5 nearest stations ($n_{st} = 5$) with the highest combined weight,

$$w_i = \exp\left(-\frac{d_i}{h_d}\right) \exp\left(-\frac{|z_i - z_t|}{h_e}\right) b_i, \tag{2}$$

with:

- d_i: Euclidean distance (m) in UTM Zone 38N;
- $h_d = 15{,}000$ m: distance decay parameter;
- $h_e = 200$ m: elevation similarity scale;
- $b_i = \min\left(1 + 0.4 \cdot \max\left(0, \frac{z_i - z_t}{500}\right), 2.0\right)$: orographic boost favoring upslope stations;
- z_i: elevation (m) of weather station i; z_t: elevation (m) of target pixel t;
- `max_dist` = 100,000 m: maximum search radius beyond which a target pixel receives no valid station within range, triggering the gap-filling procedure described in Section 2.6.

Slope β is constrained adaptively:

- Minimum: 0.1 mm m^{-1} (0.3 mm m^{-1} if $z_t >$ 1,000 m);
- Maximum: 1.5 mm m^{-1} (2.25 mm m^{-1} if $z_t >$ 1,000 m).

For pixels with $z_t > 800$ m, a final enhancement is applied:

$$\hat{r}_t \leftarrow \hat{r}_t \times \left(1 + 0.001 \times (z_t - 800) \times 0.4\right). \tag{3}$$

The complete EWLR pipeline—full-grid interpolation, gap filling, and post-processing for 11,285,644 pixels—completed in 27.2 s on a 13th Gen Intel Core i7-13700H (20 logical cores, 15.7 GB RAM) running Windows 11 and Python 3.12.13; the core parallelized interpolation kernel (Numba JIT) alone accounted for 4.0 s (∼2.8 million pixels s^{-1}), with the remainder spent on coordinate transformation and elevation-based gap filling.

2.4 Hyperparameter Optimization

Hyperparameters were optimized using Leave-One-Out Cross-Validation (LOOCV) across the 19 stations, evaluating cross-validated R^2 to

Table 2. Key metrics used for assessing model performance.

Metric	Description	Formula
Mean Absolute Error (MAE)	Average absolute deviation	$\frac{1}{n}\sum_{i=1}^{n} \lvert r_i - \hat{r}_i \rvert$
Root Mean Square Error (RMSE)	Quadratic mean of errors	$\sqrt{\frac{1}{n}\sum_{i=1}^{n} (r_i - \hat{r}_i)^2}$
Coefficient of Determination (R^2)	Proportion of variance explained	$1 - \frac{\sum_{i=1}^{n}(r_i - \hat{r}_i)^2}{\sum_{i=1}^{n}(r_i - \bar{r})^2}$
Bias	Systematic over/under-estimation	$\frac{1}{n}\sum_{i=1}^{n} (\hat{r}_i - r_i)$

balance bias–variance trade-offs in sparse-data interpolation [12]. The search grid included:

- $h_d \in \{15{,}000,\ 20{,}000,\ 25{,}000\}$ m;
- $h_e \in \{150,\ 200,\ 250,\ 300\}$ m;
- $n_{st} \in \{5,\ 6,\ 7\}$;
- Orographic factor $\in \{0.3,\ 0.4,\ 0.5\}$.

The combination $h_d = 15{,}000$ m, $h_e = 200$ m, $n_{st} = 5$, orographic factor $= 0.4$ maximized R^2 (0.797) and was adopted for final mapping. It should be noted that because the same 19-station dataset was used for both hyperparameter selection and performance reporting, the quoted metrics represent optimistic estimates; nested cross-validation or an independent holdout set would be required for unbiased performance assessment, which is a recognized limitation given the sparse station network.

Physically, h_d (distance decay) controls how quickly a station's influence fades with distance, reflecting the spatial footprint of local rainfall processes; h_e (elevation decay) governs how strongly elevation similarity is weighted, capturing the fact that stations at similar altitudes experience more comparable orographic conditions; the orographic boost factor upweights stations situated upslope of the target pixel, reflecting enhanced precipitation on windward terrain; and n_{st} (number of stations) balances local sensitivity against noise by setting how many nearby stations contribute to each local regression.

2.5 Benchmark Interpolation Methods

Five standard interpolation methods were implemented for benchmarking against EWLR, selected for their prevalence in spatial rainfall mapping:

- **Inverse Distance Weighting (IDW):** Power parameter $p = 2$, using the 12 nearest stations ($k = 12$).
- **Ordinary Kriging (OK):** Spherical variogram model with 6 lags, fitted via maximum likelihood estimation.
- **Universal Kriging (UK):** Linear trend model with spherical variogram (6 lags), fitted via maximum likelihood estimation.
- **Thin-Plate Spline (TPS):** Implemented using `scipy.interpolate.Rbf` with `function='thin_plate'`.
- **Multiquadric Radial Basis Function (RBF):** Implemented using `scipy.interpolate.Rbf` with `function='multiquadric'`.

All methods were applied to the same 19-station dataset and extrapolated to the full 30 m grid, employing chunked processing for computational efficiency in Python (e.g., via `scipy` and `pykrige` libraries). Benchmark parameters followed standard defaults and common practice in the interpolation literature: IDW used $k = 12$ nearest neighbors (a typical mid-range choice), and OK/UK variograms used a spherical model with 6 lags fitted via maximum likelihood, following conventional geostatistical practice; TPS and RBF used `scipy`'s default `thin_plate` and `multiquadric` kernel settings without additional tuning, ensuring a fair, untuned comparison against the optimized EWLR model.

2.6 Gap Filling and Post-Processing

Interpolated values were clipped to the observed range [257, 1,384] mm. Gaps (due to distance cutoff) were filled using elevation-binned mean rainfall from valid predictions, divided into 10 quantiles of target elevation. This ensured physically plausible values in data-sparse high-elevation zones.

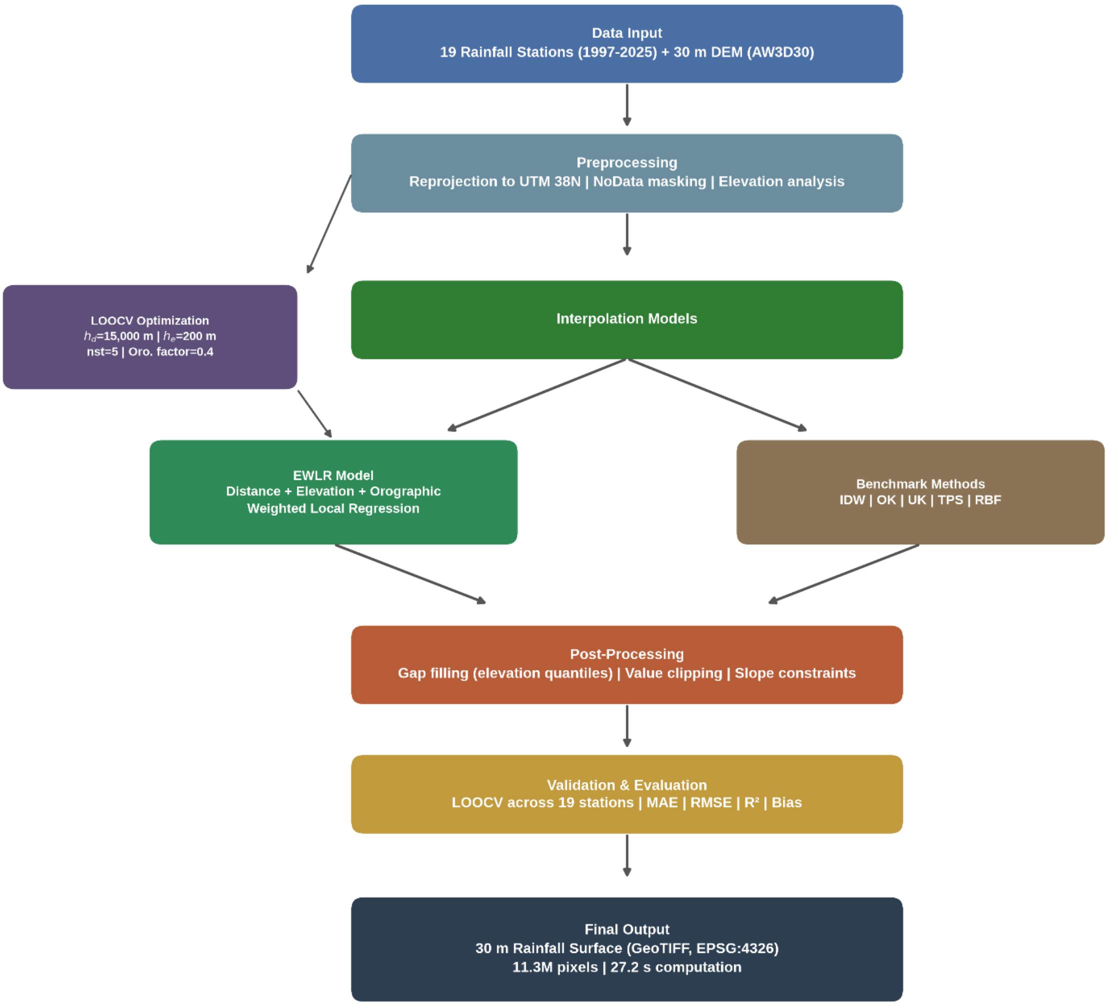

Figure 2. Workflow of the proposed Enhanced Elevation-Weighted Local Regression (EWLR) methodology for rainfall interpolation.

Gaps arise where target pixels exceed the maximum search distance from any station, a common occurrence in the high-elevation zones outside the station network's spatial extent. Because no stations exist above 1,187 m, elevation-binned quantiles—rather than direct distance-based interpolation—were used to extend physically plausible estimates up to the DEM's maximum elevation of 3,597 m, leveraging the observed elevation–rainfall relationship instead of spatial proximity.

2.7 Model Evaluation and Validation

Model performance was assessed using LOOCV on the 19 stations. Key metrics included MAE, RMSE, R^2, and Bias (Table 2).

Spatial outputs were saved as GeoTIFFs in WGS84 (EPSG:4326) with 30 m resolution and float32 precision (Figure 2). Zonal statistics and elevation–rainfall correlations were computed across the full grid.

3 Results

3.1 Hyperparameter Optimization via LOOCV

Table 3. Parameter grid for LOOCV hyperparameter optimization in the EWLR model.

Parameter	Description	Tested Values
h_d	Horizontal distance decay scale	[15,000; 20,000; 25,000] m
h_e	Elevation decay scale	[150; 200; 250; 300] m
n_{st}	Number of nearest stations	[5; 6; 7]
Orographic factor	Orographic enhancement multiplier	[0.3; 0.4; 0.5]

A grid search via LOOCV was conducted to optimize the EWLR model's key parameters, using cross-validated R^2 as the primary criterion to balance fit and generalizability across the 19 stations. The parameter grid tested $3 \times 4 \times 3 \times 3 = 108$ combinations (Table 3).

For each combination, predictions were generated for the held-out station using the remaining 18, subject to physical constraints (e.g., maximum slope 1.5 mm m^{-1}, rising to 2.25 mm m^{-1} above 1,000 m). Secondary evaluation used MAE.

The optimal configuration—h_d = 15,000 m, h_e = 200 m, n_{st} = 5, orographic factor = 0.4—maximized R^2 at 0.797 (MAE = 87.46 mm), outperforming non-elevation baselines (R^2 = 0.50–0.55; see Table 4) by roughly 45–59% in relative R^2 and yielding stable, monotonic orographic trends without overfitting. This set was applied for all subsequent interpolations and raster generation (Figure 3).

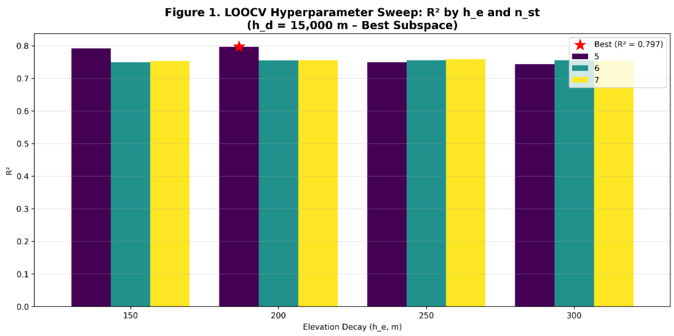

Figure 3. LOOCV hyperparameter sweep showing R^2 as a function of elevation decay (h_e) and number of stations (n_{st}), with h_d fixed at 15,000 m.

3.2 Comparative Evaluation of Interpolation Methods

Using the optimized EWLR parameters (h_d = 15,000 m, h_e = 200 m, n_{st} = 5), a full LOOCV benchmarked the model against five conventional techniques: IDW, OK, UK, TPS, and RBF. Performance was quantified using four metrics: MAE, RMSE, R^2, and Bias. Results are summarized in Table 4 and visualized in Figure 4 (scatter plots of observed vs. predicted annual rainfall).

Table 4. LOOCV performance metrics for rainfall interpolation methods.

Method	MAE (mm)	RMSE (mm)	R^2	Bias (mm)
EWLR	**87.46**	**120.90**	**0.797**	**+18.46**
IDW	106.98	180.05	0.549	−16.85
OK	108.31	188.80	0.504	−25.10
UK	NaN	NaN	NaN	NaN
TPS	113.15	185.39	0.522	−31.21
RBF	141.93	181.11	0.544	−11.77

EWLR outperformed all benchmarks, attaining the highest R^2 (0.797) and lowest RMSE (120.9 mm), with MAE (87.46 mm) ~18% better than IDW—the next closest. This underscores EWLR's enhanced predictive fidelity and explanatory power for orographic gradients. IDW offered moderate accuracy but underpredicted high-elevation rainfall (negative bias). OK and UK underperformed; UK failed to converge due to variogram instability with the sparse 19-station dataset, a known limitation of trend-surface variogram fitting at low sample sizes that made reliable semivariogram estimation infeasible. TPS and RBF generated overly smoothed surfaces, inadequately resolving steep terrain-driven contrasts. EWLR's modest positive bias (+18.46 mm) aligns physically, slightly overestimating mid-elevations while reasonably reproducing highland peaks. Collectively, EWLR delivered the most consistent, terrain-responsive rainfall estimates.

Note: The Universal Kriging (UK) model did not converge because of the limited sample size (19 rainfall stations); therefore, no prediction results are shown for this method.

3.3 Full-Grid Interpolation and Raster Outputs

Four 30 m resolution rainfall surfaces were generated across the full valid DEM domain (~11.3 million pixels) via chunked processing for memory efficiency. Visual inspection of Figure 5 reveals distinct methodological signatures across the interpolated rainfall surfaces. The EWLR raster exhibits sharp, monotonic rainfall escalation along elevation contours, reasonably reproducing orographic uplift with peaks exceeding 1,300 mm near Mergasor, all without introducing artifacts. In contrast, the IDW surface displays moderate gradients but systematically underestimates rainfall in the highlands, especially in the northeast of the study area, with deficits of approximately 200–300 mm relative to observations and subtle influences from station clustering. The TPS interpolation generates pronounced bull's-eye rings around high-rainfall stations, such as Mergasor, which amplify local anomalies and distort broader regional patterns. Meanwhile, the RBF method imposes uniform smoothing that attenuates topographic variability, resulting in implausibly flat mid-elevation zones around 400–600 mm. Overall, EWLR yields the most geophysically plausible surface, aligning closely with station data and terrain controls while mitigating the biases evident in the baseline methods.

Using the optimized EWLR parameters identified through LOOCV, the final annual rainfall distribution for the Erbil Governorate was generated at a spatial resolution of 30 m (Figure 6). The resulting map reproduces the expected orographic rainfall gradient, with precipitation increasing from the low-elevation

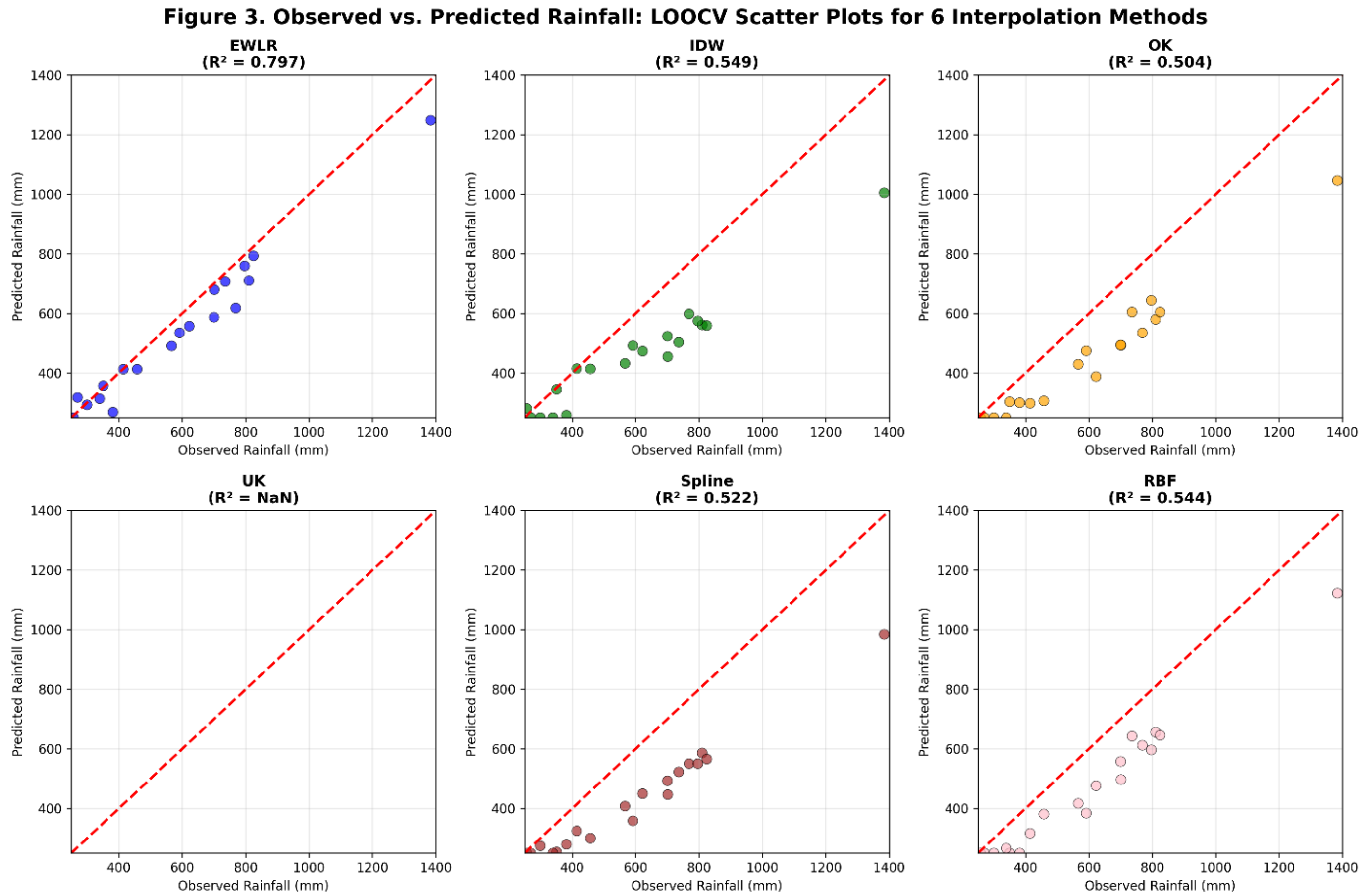

Figure 4. LOOCV scatter plots for six interpolation methods. EWLR achieves the strongest fit ($R^2 = 0.797$); UK failed to converge due to variogram instability with $n = 19$ stations.

southwestern plains toward the high-elevation northeastern Zagros Mountains. Areas above 1,200 m are indicated as extrapolation zones because no rainfall stations are available at these elevations.

3.4 Final Enhanced EWLR Model

The EWLR model was finalized using LOOCV-optimized parameters and physical constraints, ensuring stability and orographic fidelity. The adopted configuration—$h_d = 15{,}000$ m, $h_e = 200$ m, $n_{st} = 5$, `max_dist` $= 100{,}000$ m, `max_slope` $= 1.5$ mm m^{-1}, orographic factor $= 0.4$—delivered $R^2 = 0.797$ and RMSE $= 120.9$ mm, outperforming baselines by 45–59% in explanatory power.

Interpolation processed the $\sim$11.3 million-pixel DEM domain in 100,000-pixel chunks via Numba-accelerated parallelism. Spatial gaps from the `max_dist` cutoff were infilled using elevation-binned medians (10 bins, 200 m intervals), maintaining gradient integrity without introducing artifacts.

The resultant 30 m surface delineates rainfall from $\sim$270 mm in southwestern plains to $>$1,300 mm in northeastern highlands, robustly capturing topographic controls and aligning with station observations.

3.5 Spatial Patterns and Elevation–Rainfall Relationship

Global statistics of the enhanced EWLR rainfall surface are summarized in Table 5.

Table 5. Global statistics of the EWLR rainfall surface.

Statistic	Value
Minimum rainfall	257 mm
Maximum rainfall	1,384 mm
Mean $\pm$ SD	741.8 ± 434.3 mm
Elevation–rainfall correlation	$r = 0.907$ ($p < 0.001$)

Rainfall variability peaks (SD $\approx$ 185 mm) in the 900–1,200 m transition zone, attributable to slope-aspect effects and sparse high-elevation sampling. The surface highlights maxima ($>$1,300 mm) along northeastern Zagros ridges, mirroring known orographic uplift.

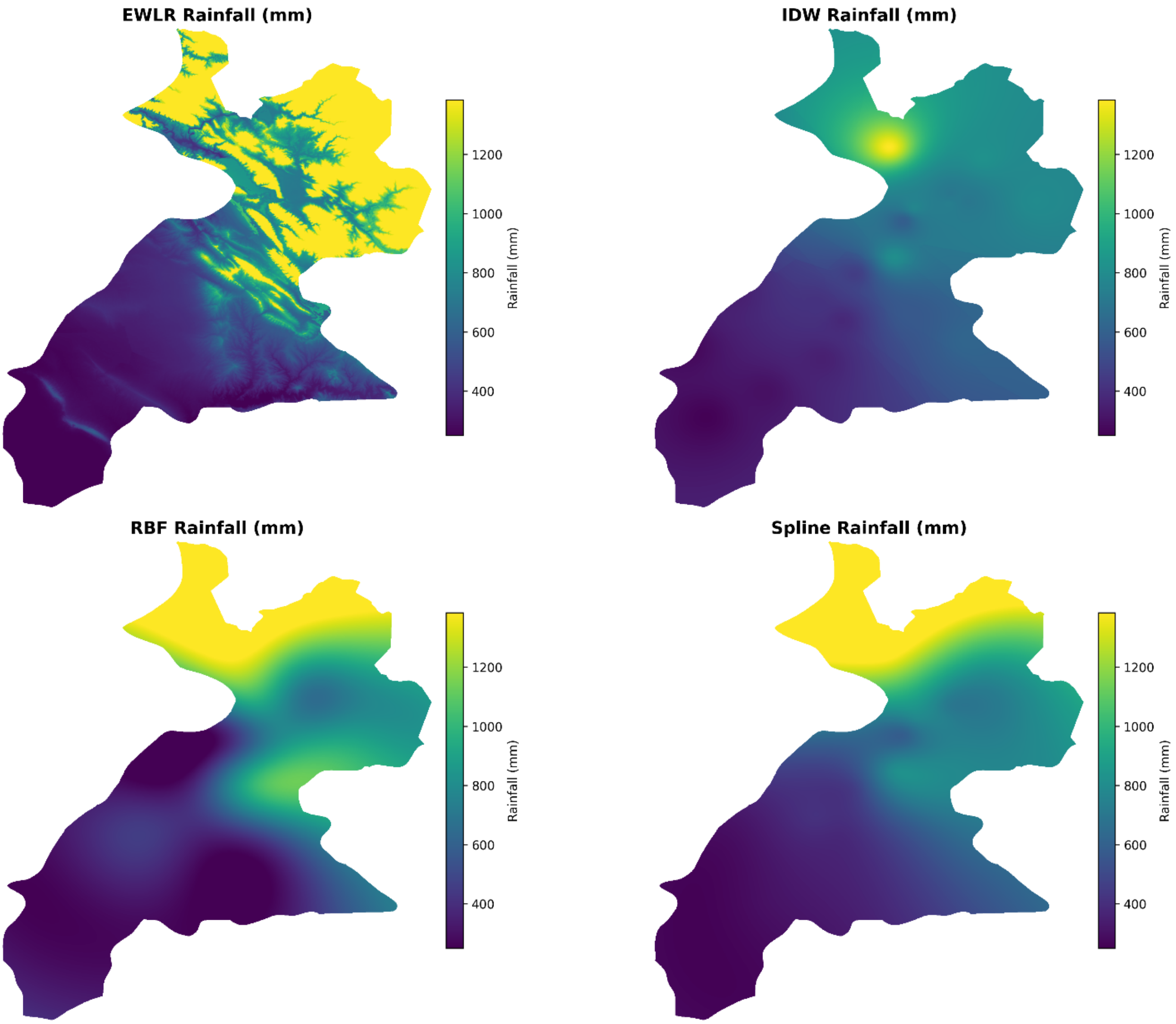

Figure 5. Side-by-side comparison of interpolated 30 m rainfall rasters from four methods. EWLR (top-left) captures orographic patterns most realistically; IDW (top-right) shows station artifacts; RBF and TPS (bottom) produce smoother but less topographic surfaces.

This very strong positive correlation ($r = 0.907$) affirms elevation as the primary driver of rainfall distribution in Erbil Governorate. Zonal statistics by elevation band (Table 6) further delineate the relationship, revealing a near-fivefold escalation from lowlands to highlands.

Note: estimates above ∼1,200 m extrapolate beyond the observed station elevation range and should be interpreted with caution.

Table 6. Zonal rainfall statistics by elevation band.

Elevation Band (m)	Mean Rainfall (mm)	Std. Dev. (mm)	% of Pixels
0–300	266.6	24.0	18.2%
300–600	396.2	88.0	31.5%
600–900	693.1	131.0	28.9%
900–1,200	1,074.4	185.4	15.1%
>1,200	1,360.1	68.9	6.3%

4 Discussion

4.1 Interpretation of Main Findings

The enhanced Elevation-Weighted Local Regression (EWLR) model reflected the pronounced dependence

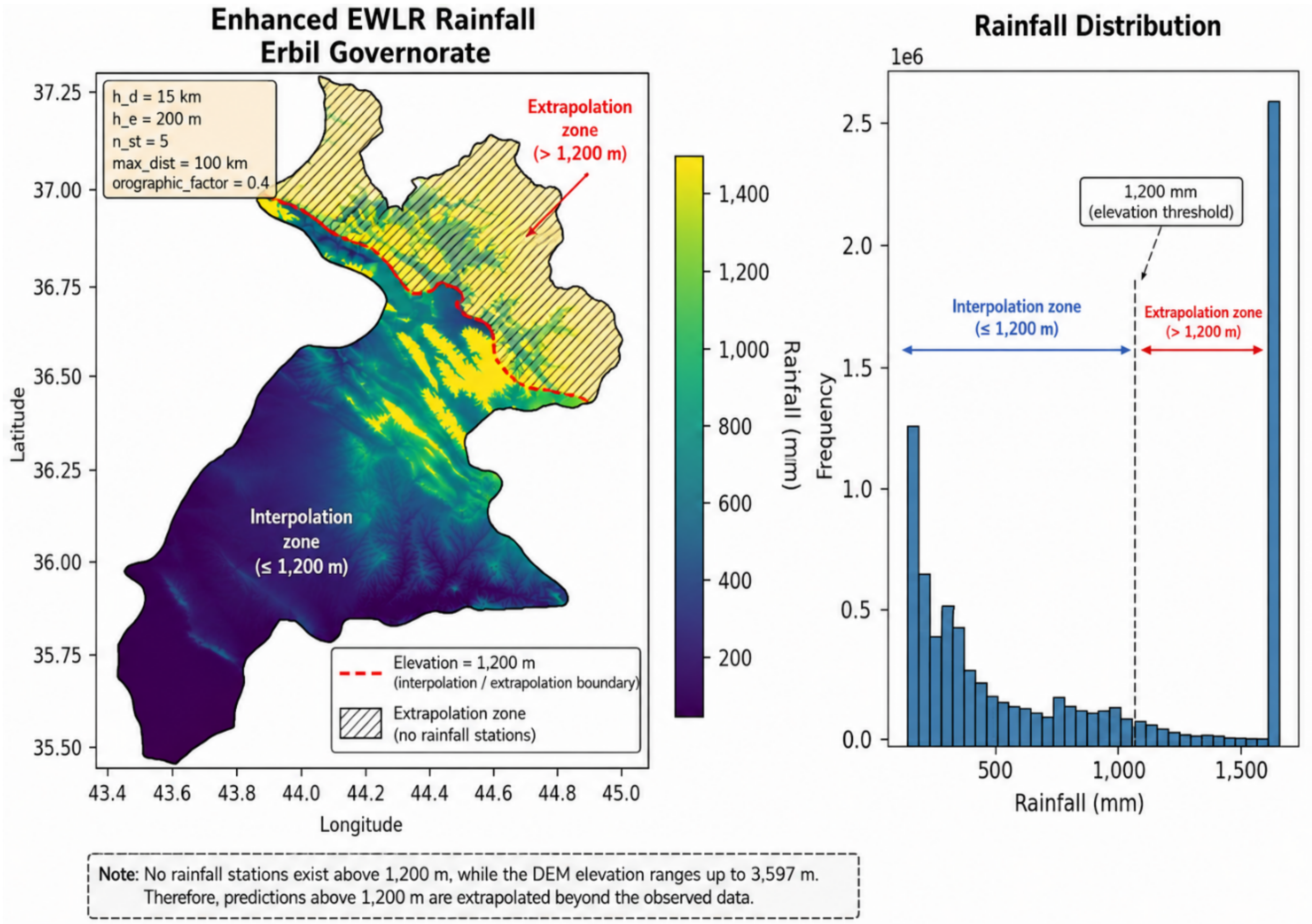

Figure 6. Final EWLR annual rainfall map (mm) for Erbil Governorate at 30 m resolution. The main panel illustrates spatial patterns with optimized parameters: $h_d = 15$ km, $h_e = 200$ m, orographic factor $= 0.4$. The inset histogram shows the distribution of $\sim$11.3 million valid pixels, peaking at 400–600 mm with a right-skewed tail reflecting orographic enhancement.

of rainfall on elevation across the Erbil Governorate, yielding a robust coefficient of determination ($R^2 = 0.797$) and a strong Pearson correlation ($r = 0.907$) between elevation and precipitation. These metrics underscore the model's ability to quantify topographic influences on hydrological patterns in this region. The findings affirm that orographic uplift, driven by the interaction of prevailing moist air masses with the rugged Zagros topography, serves as the primary mechanism governing rainfall distribution in the semi-arid mountainous terrains of northern Iraq [3]. This is evidenced by the observed fivefold escalation in annual precipitation—from approximately 270 mm in the low-lying plains (elevations $\sim$260–400 m above sea level) to over 1,300 mm in the highlands (elevations $>$900 m)—which mirrors the physical processes of forced ascent and condensation enhancement along the Zagros foothills, where synoptic-scale weather systems are amplified by terrain-induced convergence [13].

4.2 Comparison with Previous Studies

The elevation–rainfall relationship observed in this study aligns with findings from regional and global investigations into topographic controls on precipitation. For instance, incorporating elevation and orographic anisotropy into interpolation frameworks has yielded substantial enhancements in accuracy, particularly in capturing directional gradients over rugged terrains [14, 15]. Similarly, topographic variables, such as elevation and slope, account for a substantial portion of rainfall variance—often exceeding 80% in combined models—across the Zagros-adjacent landscapes of northern Iraq and western Iran, thereby validating the EWLR's emphasis on these predictors [16].

Conventional interpolation techniques, including IDW and Kriging, frequently exhibit reduced efficacy in heterogeneous terrains owing to their inadequate handling of topographic covariates, resulting in smoothed gradients and elevated errors (e.g., RMSE increases >20% relative to elevation-augmented methods) [15]. The EWLR's superior performance (RMSE = 120.9 mm) relative to these baselines underscores the value of dual distance–elevation weighting alongside orographic adjustments, echoing global calls for terrain-explicit gridded products to support hydrological applications [17].

At the regional scale, studies on the Zagros foothills have documented pronounced rainfall gradients, with precipitation escalating markedly with elevation in the Erbil and Duhok governorates—often doubling or tripling from plains to peaks—due to orographic uplift from westerly flows [3, 18]. By leveraging 30 m resolution DEMs, this work extends these insights, addressing persistent gaps in high-fidelity, topographically informed rainfall surfaces for Iraq.

4.3 Orographic Influence and Model Behavior

The EWLR model's incorporation of a physically constrained slope parameter and an adaptive orographic factor (set at 0.4) promoted numerical stability while effectively preserving the region's pronounced rainfall gradients, which can exceed 5 mm per 100 m in steep topographic transitions. The observed slight positive bias (+18 mm) at mid-elevations (approximately 400–800 m) is attributable to inherent smoothing artifacts in locally weighted regression approaches, particularly across windward–leeward transition zones where asymmetric orographic forcing—such as enhanced uplift on windward slopes and rain-shadow desiccation on leeward sides—generates sharp precipitation discontinuities that global or local averaging tends to attenuate [19]. Despite this, the EWLR's raster outputs exhibit physically coherent elevational trends, with monotonic increases aligning to observed orographic enhancement rates, thereby outperforming conventional distance-only interpolation methods (e.g., inverse distance weighting), which systematically underrepresent and flatten these topographic signals in complex terrains [11, 15].

4.4 Uncertainty and Limitations

Despite the EWLR's strong performance, several limitations persist that could inform future refinements. First, the sparse distribution of rain gauges above 1,200 m a.s.l. hampers direct validation in the highest terrains, where orographic maxima prevail, thereby elevating interpolation uncertainties in these underobserved zones. Critically, because no stations exist above 1,187 m a.s.l. while the DEM extends to 3,597 m, all predictions above ∼1,200 m represent pure extrapolation beyond the observed elevation range; the reported zonal statistics for the >1,200 m band (Table 6) should therefore be interpreted with particular caution, and the gap-filling procedure applied in those zones introduces additional uncertainty not captured by the LOOCV metrics [20]. Comparable challenges are documented in other rugged landscapes, underscoring diminished reliability of spatial estimates amid low high-elevation gauge density [21]. Second, reliance on annual precipitation totals overlooks intra-annual dynamics; as demonstrated for Iraq, precipitation is highly concentrated seasonally, with the degree of concentration increasing from the mountainous northeast toward the arid south and southwest, and monthly-to-seasonal analyses are needed to capture the distinct contributions of winter frontal systems versus sporadic warm-season convection [22]. Lastly, while the EWLR relies solely on elevation as a predictor, geostatistical frameworks that treat topography as an exogenous variable within copula-based or kriging-with-external-drift approaches have demonstrated improved interpolation fidelity across a range of temporal scales [23]; extending such frameworks to additionally incorporate atmospheric drivers—such as wind direction, humidity, or synoptic circulation patterns—represents a promising avenue for further boosting predictive accuracy in regions subject to variable moisture regimes.

This optimism is consistent with reported magnitudes in comparable spatial prediction studies: ordinary cross-validation schemes similar to LOOCV have been shown to be up to 40% more optimistic than spatially-aware validation when training data are spatially clustered [24]. Given the Erbil network's clustering toward mid-elevations (Section 2.6), the reported R^2 (0.797) and MAE (87.46 mm) likely represent a best-case bound; true out-of-sample performance in the underrepresented high-elevation zones is plausibly notably worse than the LOOCV estimate suggests. An independent high-elevation holdout would be needed to quantify this precisely.

4.5 Implications for Hydrological and Environmental Modelling

The high-resolution rainfall map produced by the EWLR model has direct implications for regional water resource management, hydrological modelling, and ecological planning. Accurate rainfall surfaces are essential for runoff estimation, drought risk assessment, and soil erosion modelling—applications particularly relevant in the semi-arid and topographically complex environment of Kurdistan. By providing rainfall estimates at 30 m resolution, this study bridges a critical scale gap between local meteorological observations and coarse global precipitation products (e.g., CHIRPS, ERA5). The methodology can be readily extended to other mountainous regions in the Middle East where station density is low but DEM data are widely available.

4.6 Future Work

Priority directions include assimilating satellite-derived precipitation products to fill high-elevation data gaps; as demonstrated for Erbil, CHIRPS gridded data can serve as a validated alternative source for rainfall estimation in areas with sparse and unreliable gauge networks [25], while further integration of IMERG, GPM, and ERA5-Land products offers additional avenues for constraining orographic precipitation in data-scarce high-elevation zones. Expanding the in-situ gauge network at higher altitudes to better constrain orographic effects is also advisable [26], as is incorporating additional terrain derivatives such as slope aspect and curvature to capture micro-scale forcing [27]. A hybrid AI–geostatistical framework fusing ground and remote-sensing data [28] offers a longer-term path toward improved spatiotemporal integration.

5 Conclusion

This study investigated the spatial distribution of rainfall across the Erbil Governorate in northern Iraq, emphasizing the elevation–rainfall nexus through an enhanced suite of interpolation techniques, including the novel EWLR model. The analyses robustly demonstrated that topographic features dominate precipitation patterns, with annual rainfall escalating from approximately 270 mm in the southwestern plains to over 1,300 mm in the northeastern highlands—a gradient emblematic of orographic uplift in the Zagros foothills. Among evaluated methods, the EWLR outperformed traditional deterministic (e.g., IDW) and geostatistical (e.g., Kriging) approaches by integrating elevation similarity and adaptive orographic factors, yielding superior accuracy and physically coherent spatial surfaces.

The pronounced positive correlation ($r = 0.907$) between elevation and rainfall underscores orography as a pivotal driver, modulated by interactions with prevailing westerly flows and local microclimates that generate sharp intra-regional contrasts. These insights resonate with broader patterns in semi-arid mountainous domains, where terrain-induced forcing amplifies hydrological variability. Practically, these high-resolution (30 m) rainfall maps provide actionable tools for water resource management, precision agriculture, and flood mitigation in Erbil, enabling nuanced hydrological simulations amid escalating climate pressures. By addressing data sparsity via terrain-explicit modeling, this framework exemplifies scalable solutions for under-monitored landscapes.

Notwithstanding these advances, constraints persist: reliance on a modest station array may attenuate capture of sub-kilometer fluctuations, and exclusion of ancillary variables (e.g., slope aspect, land cover) curtails holistic process representation. Future work could integrate multi-decadal satellite archives (e.g., CHIRPS, GPM) and hybrid machine learning–geostatistical approaches to improve temporal fidelity and predictive accuracy. In sum, this work affirms elevation's primacy in sculpting Erbil's rainfall regime and introduces an interpretable, high-fidelity interpolation paradigm that addresses observational gaps in rugged terrain. The resultant geospatial assets lay a cornerstone for resilient environmental governance and extended climatological inquiry in the Kurdistan Region of Iraq and analogous regions. It should be noted that rainfall estimates above ~1,200 m extrapolate beyond the observed station network and warrant cautious interpretation.

Data Availability Statement

Meteorological station records are available from the corresponding author subject to permission from the Kurdistan Region Meteorological Directorate. AW3D30 elevation data are freely available from the Japan Aerospace Exploration Agency (JAXA) upon user registration at https://www.eorc.jaxa.jp/ALOS/en/dataset/aw3d30/.

Funding

This work was supported without any funding.

Conflicts of Interest

The author declares no conflicts of interest.

AI Use Statement

The author declares that Grok (developed by xAI) was used for language editing of the manuscript. The author has carefully reviewed, revised, and verified the AI-assisted output and takes full responsibility for the content of the manuscript.

Ethical Approval and Consent to Participate

Not applicable. This study did not involve human participants or animals; it relies solely on meteorological station records and publicly available digital elevation data.

References

[1] Houze Jr, R. A. (2012). Orographic effects on precipitating clouds. *Reviews of Geophysics, 50*(1). [CrossRef]

[2] Roe, G. H. (2005). Orographic precipitation. *Annual Review of Earth and Planetary Sciences, 33*(1), 645-671. [CrossRef]

[3] Alijani, B. (2008). Effect of the Zagros Mountains on the spatial distribution of precipitation. *Journal of Mountain Science, 5*(3), 218-231. [CrossRef]

[4] Aljaff, M. A. (2025). Trend and variability of rainfall in recent years in Iraq: A physical perspective. *International Journal of Science and Research Archive, 14*(2), 1118-1128. [CrossRef]

[5] Alsafadi, K., Bi, S., Bashir, B., Sharifi, E., Alsalman, A., Kumar, A., & Shahid, S. (2023). High-resolution precipitation modeling in complex terrains using hybrid interpolation techniques: Incorporating physiographic and MODIS cloud cover influences. *Remote Sensing, 15*(9), 2435. [CrossRef]

[6] Muter, S. A., Al-Jiboori, M. H., & Al-Timimi, Y. K. (2025). Assessment of spatial and temporal monthly rainfall trend over Iraq. *Baghdad Science Journal, 22*(3), 910-922. [CrossRef]

[7] Forti, L., Perego, A., Brandolini, F., Mariani, G. S., Zebari, M., Nicoll, K., ... & Zerboni, A. (2021). Geomorphology of the northwestern Kurdistan Region of Iraq: landscapes of the Zagros Mountains drained by the Tigris and Great Zab Rivers. *Journal of Maps, 17*(2), 225-236. [CrossRef]

[8] Doski, J. A., & McClay, K. (2022). Tectono-stratigraphic evolution, regional structure and fracture patterns of the Zagros fold-thrust belt in the Duhok region, Kurdistan, northern Iraq. *Tectonophysics, 838*, 229506. [CrossRef]

[9] Gaznayee, H. A. A., Al-Quraishi, A. M. F., Mahdi, K., & Ritsema, C. (2022). A geospatial approach for analysis of drought impacts on vegetation cover and land surface temperature in the Kurdistan Region of Iraq. *Water, 14*(6), 927. [CrossRef]

[10] Forti, L., Pelfini, M., Sissakian, V. K., & Zerboni, A. (2023). Settling the riverscape of Erbil (Kurdistan Region of Iraq): long-term human overprint on landforms and present-day geomorphological hazard. *Anthropocene, 44*, 100413. [CrossRef]

[11] Zandi, O., Nasseri, M., & Zahraie, B. (2023). A locally weighted linear ridge regression framework for spatial interpolation of monthly precipitation over an orographically complex area. *International Journal of Climatology, 43*(6), 2601-2622. [CrossRef]

[12] Goovaerts, P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of hydrology, 228*(1-2), 113-129. [CrossRef]

[13] Mustafa, A. M., Muhammed, H., & Szydłowski, M. (2019). Extreme rainfalls as a cause of urban flash floods: A case study of the Erbil-Kurdistan region of Iraq. *Journal of Water and Land Development, 42*, 113-132. [CrossRef]

[14] Dura, V., Evin, G., Favre, A. C., & Penot, D. (2025). Improving precipitation interpolation using anisotropic variograms derived from convection-permitting regional climate model simulations. *EGUsphere, 2025*, 1-25. [CrossRef]

[15] Workneh, H. T., Chen, X., Ma, Y., Bayable, E., & Dash, A. (2024). Comparison of IDW, Kriging and orographic based linear interpolations of rainfall in six rainfall regimes of Ethiopia. *Journal of Hydrology: Regional Studies, 52*, 101696. [CrossRef]

[16] Javari, M. (2017). Assessment of temperature and elevation controls on spatial variability of rainfall in Iran. *Atmosphere, 8*(3), 45. [CrossRef]

[17] Alsafadi, K., Mohammed, S., Mokhtar, A., Sharaf, M., & He, H. (2021). Fine-resolution precipitation mapping over Syria using local regression and spatial interpolation. *Atmospheric Research, 256*, 105524. [CrossRef]

[18] Hajani, E., Shajee, K., Kaleel, F., & Abdulhaq, H. (2022). Characteristics of changes in rainfall data in the Kurdistan Region, Iraq. *Arabian Journal of Geosciences, 15*(6), 509. [CrossRef]

[19] Gottardi, F., Obled, C., Gailhard, J., & Paquet, E. (2012). Statistical reanalysis of precipitation fields based on ground network data and weather patterns: Application over French mountains. *Journal of Hydrology, 432*, 154-167. [CrossRef]

[20] Bertoncini, A., & Pomeroy, J. W. (2025). Quantifying spatiotemporal and elevational precipitation gauge network uncertainty in the Canadian Rockies. *Hydrology and Earth System Sciences, 29*(4), 983-1000. [CrossRef]

[21] Xiaojun, G., Peng, C., Xingchang, C., Yong, L., Ju, Z., & Yuqing, S. (2021). Spatial uncertainty of rainfall and its impact on hydrological hazard forecasting in a small semiarid mountainous watershed. *Journal of Hydrology, 595*, 126049. [CrossRef]

[22] Bandar, Q. K. A., & Muslih, K. D. (2025). Spatial assessment of precipitation concentration and seasonality in Iraq. *Theoretical and Applied Climatology, 156*(2), 134. [CrossRef]

[23] Bárdossy, A., & Pegram, G. (2013). Interpolation of precipitation under topographic influence at different time scales. *Water Resources Research, 49*(8), 4545-4565. [CrossRef]

[24] Pohjankukka, J., Pahikkala, T., Nevalainen, P., & Heikkonen, J. (2017). Estimating the prediction performance of spatial models via spatial k-fold cross validation. *International Journal of Geographical Information Science, 31*(10), 2001-2019. [CrossRef]

[25] Aziz, A. A., Omer, P. A., & Kak Ahmed Khayyat, H. A. (2026). Assessing The Accuracy of Satellite Rainfall Data for Erbil city. *Passer Journal of Basic and Applied Sciences, 8*(1), 146-176. [CrossRef]

[26] Morsy, M., Taghizadeh-Mehrjardi, R., Michaelides, S., Scholten, T., Dietrich, P., & Schmidt, K. (2021). Optimization of rain gauge networks for arid regions based on remote sensing data. *Remote Sensing, 13*(21), 4243. [CrossRef]

[27] Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., ... & Pasteris, P. P. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology: a Journal of the Royal Meteorological Society, 28*(15), 2031-2064. [CrossRef]

[28] Sekulić, A., Kilibarda, M., Heuvelink, G. B., Nikolić, M., & Bajat, B. (2020). Random forest spatial interpolation. *Remote Sensing, 12*(10), 1687. [CrossRef]

Azad Rasul is a Senior Assistant Professor of Remote Sensing at Soran University and Salahaddin University-Erbil, Iraq, with over 18 years of experience. He holds a PhD from the University of Leicester (2016) and has authored 80+ publications with over 12,000 citations in journals including The Lancet. He is a GBD Collaborator at the University of Washington and has taught over 103,000 students worldwide through online courses. (Email: azad.rasul@soran.edu.iq)