

SPATIAL STRUCTURE AND TAXONOMIC DIVERSITY OF THE GENUS *ROSA* L. IN THE TARBAGATAI MOUNTAINS OF EASTERN KAZAKHSTAN BASED ON GEOREFERENCED OBSERVATIONS

¹A.B. Myrzagaliyeva , ¹Y. Zhiger* , ¹S.A. Kauanov , ¹M.Zh. Sharipova , ¹A.E. Orazov 

¹Astana International University, Astana, Kazakhstan

*e-mail: ertaizhiger@gmail.com

A.B. Myrzagaliyeva – D.Sc., Professor of the Higher School of Natural Sciences, Astana International University, Astana, Kazakhstan, e-mail: an.myrzagaliyeva@gmail.com, <https://orcid.org/0000-0003-3604-4315>

Y. Zhiger – PhD Student, Astana International University, Astana, Kazakhstan, e-mail: ertaizhiger@gmail.com, <https://orcid.org/0009-0006-0273-8561>

S.A. Kauanov – Master student, Junior Researcher at the NatureLab Research Laboratory, Astana International University, Astana, Kazakhstan, e-mail: sultan.kaunov96@gmail.com, <https://orcid.org/0009-0002-8526-1877>

M.Zh. Sharipova – PhD, Associate Professor, Deputy Dean of the Higher School of Natural Sciences, Astana International University, Astana, Kazakhstan, e-mail: moldir_zhumagul@aiu.edu.kz, <https://orcid.org/0000-0002-9467-1483>

A.E. Orazov – PhD, Associate Professor, Head of the NatureLab Research Laboratory, Astana International University, Astana, Kazakhstan, e-mail: orazov1994aidyn@gmail.com, <https://orcid.org/0000-0003-1390-9507>

Abstract. The mountain systems of Central Asia are important centres of floristic diversity; however, spatially representative and taxonomically verified data on the distribution of many wild taxa remain fragmentary. This study aimed to quantitatively assess the taxonomic structure, spatial distribution, degree of clustering, documented local species richness, and quality of a long-term array of georeferenced observations of the genus *Rosa* L. in the Tarbagatai mountain system of eastern Kazakhstan. The initial sample included 127 observations from 2019-2026; after taxonomic filtering, the analysis included 118 *Rosa* records corresponding to 109 unique coordinates. A total of 61 records (51.7%) were identified to the species level, while 57 records (48.3%) remained at the *Rosa* sp. level. The dataset includes *R. spinosissima* (33 observations), *R. laxa* (13), *R. platyacantha* (10), *R. alberti* (4), and *R. canina* (1). Ninety-six records were positionally accurate to within 100 m; after duplicate removal, they corresponded to 88 unique locations. Nearest neighbour analysis revealed strong spatial aggregation of high-precision locations: the observed mean nearest neighbour distance was 0.75 km, compared to a mean expected value of 3.02 km across 999 runs of complete spatial randomness (empirical $p = 0.001$). With grid thinning, the 96 high-precision observations corresponded to 32 occupied 1 x 1 km cells and 26 occupied 2 x 2 km cells, indicating significant route and observer bias. Species-identified, high-precision records occupied 14 5 x 5 km cells; the maximum documented richness, equal to four species, was observed in two cells in the central part of the studied spatial cluster. Only 15 records (12.7%) had Research Grade status, while 103 (87.3%) were classified as Needs ID. These results characterise the structure of the documented array of finds, rather than the demographic structure of populations, and provide a reproducible spatial basis for further expert verification, standardised monitoring, and subsequent modelling of species distributions with mandatory control for spatial sampling bias.

Keywords: *Rosa*; Rosaceae; Tarbagatai; East Kazakhstan; georeferenced observations; spatial structure; nearest neighbour analysis; spatial thinning; iNaturalist; taxonomic uncertainty.

ГЕОГРАФИЯЛЫҚ КООРДИНАТТАРМЕН БАЙЛАНЫСТЫРЫЛҒАН БАҚЫЛАУ ДЕРЕКТЕРІ НЕГІЗІНДЕ ШЫҒЫС ҚАЗАҚСТАНДАҒЫ ТАРБАҒАТАЙ ТАУЛАРЫНДА *ROSA L.* ТУЫСЫНЫҢ КЕҢІСТІКТІК ҚҰРЫЛЫМЫ МЕН ТАКСОНОМИЯЛЫҚ АЛУАНТҮРЛІЛІГІ

¹А.Б. Мырзағалиева, ¹Е. Жігер*, ¹С.А. Кауанов, ¹М.Ж. Шарипова, ¹А.Е. Оразов

¹Астана халықаралық университеті, Астана, Қазақстан

*e-mail: ertaizhiger@gmail.com

А.Б. Мырзағалиева – б.ғ.д., профессор, Астана халықаралық университетінің Жаратылыстану ғылымдарының жоғары мектебінің профессоры, Астана қ., Қазақстан, e-mail: an.myrzagaliyeva@gmail.com, <https://orcid.org/0000-0003-3604-4315>

Е. Жігер – Докторант, Астана халықаралық университеті, Астана қ., Қазақстан, e-mail: ertaizhiger@gmail.com, <https://orcid.org/0009-0006-0273-8561>

С.А. Кауанов – магистрант, Астана халықаралық университетінің «NatureLab» ғылыми зерттеу зертханасының кіші ғылыми қызметкері, Астана қ., Қазақстан, e-mail: sultan.kaunov96@gmail.com, <https://orcid.org/0009-0002-8526-1877>

М.Ж. Шарипова – PhD, қауымдастырылған профессор, Астана халықаралық университетінің Жаратылыстану ғылымдарының жоғары мектебінің декан орынбасары, Астана қ., Қазақстан, e-mail: moldir_zhumagul@aiu.edu.kz, <https://orcid.org/0000-0002-9467-1483>

А.Е. Оразов – PhD, қауымдастырылған профессор, Астана халықаралық университетінің «NatureLab» ғылыми зерттеу зертханасының меңгерушісі, Астана қ., Қазақстан, e-mail: orazov1994aidyn@gmail.com, <https://orcid.org/0000-0003-1390-9507>

Аңдатпа. Орталық Азияның тау жүйелері флоралық алуантүрліліктің маңызды орталықтары болып табылады. Алайда кеңістіктік жағынан репрезентативті және таксономиялық тексерілген деректер әлі де толық емес. Бұл зерттеудің мақсаты — Шығыс Қазақстандағы Тарбағатай тау жүйесінде *Rosa L.* туысына қатысты көпжылдық, географиялық координаттармен байланыстырылған бақылау деректері негізінде таксономиялық құрылымын, кеңістіктік таралуын, шоғырлану дәрежесін, құжатталған жергілікті байлығын және сапасын сандық тұрғыдан бағалау.

Бастапқы деректер жиынтығы 2019-2026 жылдары тіркелген 127 бақылауды қамтыды. Таксономиялық сүзгіден кейін талдауға 109 бірегей координатаға сәйкес келетін *Rosa* туысының 118 жазбасы енгізілді. Олардың 61-і (51,7%) түр деңгейінде анықталған, ал 57-сі (48,3%) *Rosa sp.* деңгейінде қалды. Деректер жиынтығына *R. spinosissima* (33 бақылау), *R. laxa* (13), *R. platyacantha* (10), *R. alberti* (4) және *R. canina* (1) түрлері кірді. Орналасу дәлдігі 100 м-ден аспаған 96 жазба талдауға жоғары дәлдік координаталары ретінде алынды; дубликаттар алынып тасталғаннан кейін 88 бірегей орынға сәйкес келді.

Ең жақын көрші әдісімен жүргізілген талдау орналасуы жоғары дәлдікпен тіркелген орындардың айқын кеңістіктік шоғырлануын көрсетті: ең жақын көршіге дейінгі бақыланған орташа қашықтық 0,75 км болды, ал толық кездейсоқ кеңістіктік үлестірімнің 999 іске қосылуында күтілетін орташа қашықтық 3,02 км құрады (эмпирикалық $p = 0,001$). Тор ұяшықтары бойынша кеңістіктік сирету нәтижесінде жоғары дәлдіктегі 96 бақылау өлшемі 1 x 1 км болатын 32 ұяшыққа және 2 x 2 км болатын 26 ұяшыққа сәйкес келді. Бұл бақылау маршруттары мен бақылаушыларға байланысты тіркеудің елеулі біркелкісіздігін көрсетеді. Түр деңгейінде анықталған, жоғары дәлдікті жазбалар өлшемі 5 x 5 км болатын 14 ұяшықта тіркелді. Құжатталған ең жоғары байлық төрт түрге тең болып, зерттелген кеңістіктік шоғырдың орталық бөлігіндегі екі ұяшықта байқалды.

Тек 15 жазба (12,7%) «Зерттеу деңгейі» (Research Grade) мәртебесіне ие болды, ал 103 жазба (87,3%) «Анықтауды қажет етеді» (Needs ID) санатына жатқызылды. Алынған нәтижелер популяциялардың демографиялық құрылымын емес, құжатталған тіркелімдер жиынтығының құрылымын сипаттайды. Олар әрі қарай сараптамалық тексеру, стандартталған мониторинг

және кеңістіктік тіркеу біржақтылығын міндетті түрде ескере отырып, түрлердің таралуын модельдеу үшін қайта өндіруге болатын кеңістіктік негіз қалыптастырады.

Клт сөздер: *Rosa*; Rosaceae; Тарбағатай; Шығыс Қазақстан; географиялық координаттармен байланыстырылған бақылаулар; кеңістіктік құрылым; ең жақын көрші әдісін талдау; кеңістіктік сирету; iNaturalist; таксономиялық белгісіздік.

ПРОСТРАНСТВЕННАЯ СТРУКТУРА И ТАКСОНОМИЧЕСКОЕ РАЗНООБРАЗИЕ РОДА *ROSA* L. В ГОРАХ ТАРБАГАТАЯ ВОСТОЧНОГО КАЗАХСТАНА НА ОСНОВЕ ГЕОПРИВЯЗАННЫХ ДАННЫХ НАБЛЮДЕНИЙ

¹А.Б. Мырзагалиева, ¹Е. Жигер*, ¹С.А. Кауанов, ¹М.Ж. Шарипова, ¹А.Е. Оразов

¹Международный университет Астана, г. Астана, Казахстан

*e-mail: ertaizhiger@gmail.com

Мырзагалиева А.Б. – д.б.н., профессор, профессор Высшей школы естественных наук Международного университета Астана, г. Астана, Казахстан, e-mail: an.myrzagaliyeva@gmail.com, <https://orcid.org/0000-0003-3604-4315>

Жигер Е. – Докторант Международного университета Астана, г. Астана, Казахстан, e-mail: ertaizhiger@gmail.com, <https://orcid.org/0009-0006-0273-8561>

Кауанов С.А. – магистрант, младший научный сотрудник научно-исследовательской лаборатории «NatureLab» Международного университета Астана, г. Астана, Казахстан, e-mail: sultan.kaunov96@gmail.com, <https://orcid.org/0009-0002-8526-1877>

Шарипова М.Ж. – PhD, ассоциированный профессор, заместитель декана Высшей школы естественных наук Международного университета Астана, г. Астана, Казахстан, e-mail: moldir_zhumagul@aiu.edu.kz, <https://orcid.org/0000-0002-9467-1483>

Оразов А.Е. – PhD, ассоциированный профессор, заведующий научно-исследовательской лабораторией «NatureLab» Международного университета Астана, г. Астана, Казахстан, e-mail: orazov1994aidyn@gmail.com, <https://orcid.org/0000-0003-1390-9507>

Аннотация. Горные системы Центральной Азии являются важными центрами флористического разнообразия, однако пространственно репрезентативные и таксономически проверенные данные о распространении многих дикорастущих таксонов остаются фрагментарными. Цель исследования — количественно оценить таксономическую структуру, пространственное распределение, степень кластеризации, документированное локальное видовое богатство и качество многолетнего массива геопривязанных наблюдений рода *Rosa* L. в горной системе Тарбагатай Восточного Казахстана.

Исходная выборка включала 127 наблюдений за 2019-2026 годы. После таксономической фильтрации в анализ вошли 118 записей рода *Rosa*, соответствующих 109 уникальным координатам. До видового уровня было определено 61 запись (51,7%), тогда как 57 записей (48,3%) остались на уровне *Rosa* sp. В наборе данных представлены *R. spinosissima* (33 наблюдения), *R. laxa* (13), *R. platyacantha* (10), *R. alberti* (4) и *R. canina* (1). Для 96 записей погрешность определения местоположения не превышала 100 м; после удаления дубликатов они соответствовали 88 уникальным местоположениям.

Анализ ближайшего соседа выявил выраженную пространственную агрегацию местоположений с высокой точностью геопривязки: наблюдаемое среднее расстояние до ближайшего соседа составило 0,75 км, тогда как среднее ожидаемое значение по результатам 999 реализаций модели полной пространственной случайности — 3,02 км (эмпирическое $p = 0,001$). При пространственном прореживании по сетке 96 высокоточных наблюдений соответствовали 32 занятым ячейкам размером 1 x 1 км и 26 занятым ячейкам размером 2 x 2 км, что указывает на существенное смещение выборки, связанное с маршрутами обследования и активностью наблюдателей. Записи, определённые до видового уровня и имеющие высокую точность геопривязки, были представлены в 14 ячейках размером 5 x 5 км. Максимальное

документированное видовое богатство составило четыре вида и было отмечено в двух ячейках центральной части исследуемого пространственного кластера.

Лишь 15 записей (12,7%) имели статус «Исследовательский уровень» (Research Grade), тогда как 103 записи (87,3%) относились к категории «Требуется определение» (Needs ID). Полученные результаты характеризуют структуру документированного массива находок, а не демографическую структуру популяций, и формируют воспроизводимую пространственную основу для дальнейшей экспертной проверки, стандартизированного мониторинга и последующего моделирования распространения видов с обязательным учётом пространственного смещения выборки.

Ключевые слова: *Rosa*; Rosaceae; Тарбагатай; Восточный Казахстан; геопривязанные наблюдения; пространственная структура; анализ ближайшего соседа; пространственное прореживание; iNaturalist; таксономическая неопределённость.

Introduction. The mountainous regions of Central Asia are characterised by high topographic heterogeneity and pronounced floristic differentiation. Elevation gradients, slope aspect, and spatial heterogeneity of moisture and soil conditions create a mosaic of habitats that supports diversity at the ecosystem, species, and population levels (Tayjanov et.al., 2017). In such landscapes, the spatial structure of plant finds is determined by both ecological gradients and the configuration of available habitats, making georeferenced data particularly valuable for analysing taxon distribution.

Tarbagatai is one of the key mountain systems of eastern Kazakhstan and represents a transitional area between the Central Asian, Siberian, and intracontinental floristic complexes. Published data on the region's flora vary in scope and methodology, but consistently point to high species richness and pronounced vertical zonation. The southern macroslope successively features desert-steppe, steppe, shrub, subalpine, and alpine belts (Orazov, A et.al., 2021; Myrzagaliyeva, A et.al., 2024). The rugged topography creates additional ecological heterogeneity: foothill hilly areas, lowlands, midlands, valleys, and gorges create contrasting microclimatic and edaphic conditions (Myrzagaliyeva, A et.al., 2024).

The genus *Rosa* L. includes ecologically and economically significant representatives of the family Rosaceae. Wild *Rosa* species in Kazakhstan are important components of natural communities, sources of genetic resources, and objects of medicinal, food, and breeding use. Field studies indicate interspecific differences in altitudinal and edaphic confinement and significant ecological plasticity of individual taxa (Danilova, AN et.al., 2025; Myrzagaliyeva, A et.al., 2026). At the same time, the genus taxonomy remains complex due to high morphological variability, hybridisation, and partial overlap of diagnostic features. Therefore, identification based on a limited set of photographs requires caution, and in controversial cases, it is advisable to combine morphological and molecular data (Yermagambetova, M et.al., 2025).

In recent years, botanical studies of wild and resource-significant plants of Kazakhstan have increasingly relied on integrating morphological, molecular phylogenetic, genomic, ecological, and spatial data. This approach was applied in a comprehensive taxonomic revision of *Gagea alberti*, a study of the chloroplast genome of the Kazakh apricot *Prunus armeniaca*, an assessment of the morphological and genetic resources of wild relatives of almond, and a floristic-anatomical analysis of *Crataegus ambigua* populations (Kubentayev, S et.al., 2025; Imanbayeva, A et.al., 2025). Taken together, these studies highlight the potential of combining independent lines of evidence to study taxonomically complex, ecologically flexible, and conservation-significant plants, which is particularly relevant for the spatial analysis of representatives of the genus *Rosa* in the mountain ecosystems of eastern Kazakhstan.

Comparable regional studies show that comparing spatially dispersed localities with habitat parameters and functional traits significantly improves understanding of the ecological and population characteristics of rare and wild species. Thus, for *Eranthis longistipitata* in the Western Tien Shan, interlocal differences in the phytochemical profile and growing conditions were revealed (Aimenova, Z. et.al., 2023). At the same time, an analysis of *Dactylorhiza incarnata* populations in the Kalbinsky Altai demonstrated the informative value of a joint assessment of the ecological optimum, plant

community structure, and anthropogenic load (Danilova, AN. et.al.,2021). These examples further substantiate the need to combine georeferenced *Rosa* finds with independent ecological, population, and functional measurements.

Citizen science platforms and digital biodiversity databases, including iNaturalist, are significantly expanding the volume of available georeferenced observations, especially for hard-to-reach areas. However, such data are generated independently of a unified sampling design. They are subject to several types of bias: uneven observational effort, concentration of finds along roads and routes, repeat visits to accessible sites, and taxonomic identification errors. For plants, an additional limitation is the lack of a full set of diagnostic features in photographs. Therefore, using such data in spatial ecology requires simultaneous assessment of the geographic accuracy, taxonomic reliability, and spatial dependence of observations (López-Guillén. et.al.,2021).

Despite individual floristic and ecological studies of Tarbagatai, the spatial structure of long-term georeferenced *Rosa* records and the influence of observer bias on their interpretation remain poorly characterised quantitatively. The present study aimed to comprehensively evaluate the long-term array of observations of the genus *Rosa* in the Tarbagatai mountain system of eastern Kazakhstan. In this study: (1) the taxonomic and temporal structures of the data were characterized; (2) the documented spatial coverage of individual species was assessed; (3) the deviation of the distribution of localities from complete spatial randomness was tested; (4) the sensitivity of the results to sample clustering by grid thinning was quantified; (5) the documented local species richness was determined; (6) the spatial precision, degree of taxonomic verification, and concentration of observational effort were assessed; and (7) the results were interpreted in the context of the orography and altitudinal zonation of Tarbagatai. The working hypothesis was that the observed spatial pattern would be significantly heterogeneous due to the combined effect of the ecological mosaic nature of the mountain landscape and the uneven survey of the territory.

Materials and methods.

2.1. Study area and relief features

The studied localities are located within the Tarbagatai mountain system and adjacent territories of eastern Kazakhstan, approximately between 46.984–47.434° N and 80.986–82.848° E. The region is characterised by a sharply continental climate, pronounced seasonality, and a combination of foothill, low-mountain, and mid-mountain relief forms. According to published data, the maximum absolute altitudes of Tarbagatai reach approximately 3000–3100 m, and the average annual precipitation in different parts of the region varies from approximately 177 to 405 mm (Myrzagaliyeva, A. et.al.,2024).

For the southern macroslope, a sequential change in altitudinal zones has been described: desert-steppe - approximately 500-700 m, steppe - 700(600)-1000(1200) m, shrub - 1000(1200)-1700(1800) m, subalpine - 1700(1800)-2400 m and alpine - 2400-3100 m (Orazov, A. et.al.,2021). In the middle part of Tarbagatai, shrub communities with the participation of *R. spinosissima* were previously noted from approximately 800-900 to 1350-1450 m above sea level (Orazov, A. et.al.,2021). Such pronounced vertical and orographic differentiation creates an ecological basis for the spatial separation of wild rose habitats.

The original dataset did not contain a standardised field for absolute elevation or other topographic covariates. Therefore, in this analysis, topography was used only as geographic and ecological context, based on locality locations and the published vertical zonation scheme. Quantitative conclusions about the influence of elevation, slope, or aspect on species distribution were not drawn, as the relevant parameters were not extracted from the digital elevation model for each observation point.

2.2. Initial data array and pre-processing

The original CSV file contained 127 observation records. After excluding records not belonging to the genus *Rosa*, the analytical sample consisted of 118 observations from 2019–2026. We used available metadata for each record, including observation ID, date, observer name, geographic coordinates, positional accuracy, taxonomic identification, quality status, and a link to the original record. All included observations were designated as uncultivated. Exactly matching coordinates

were included in the original dataset description but were removed during analysis steps that required spatially unique localities.

Figure 1 presents the general sequence of data preparation and analytical procedures. The temporal structure was estimated from the full analytical array ($n = 118$), while detailed spatial procedures were performed after filtering by positional accuracy (≤ 100 m) and removing duplicate coordinates.

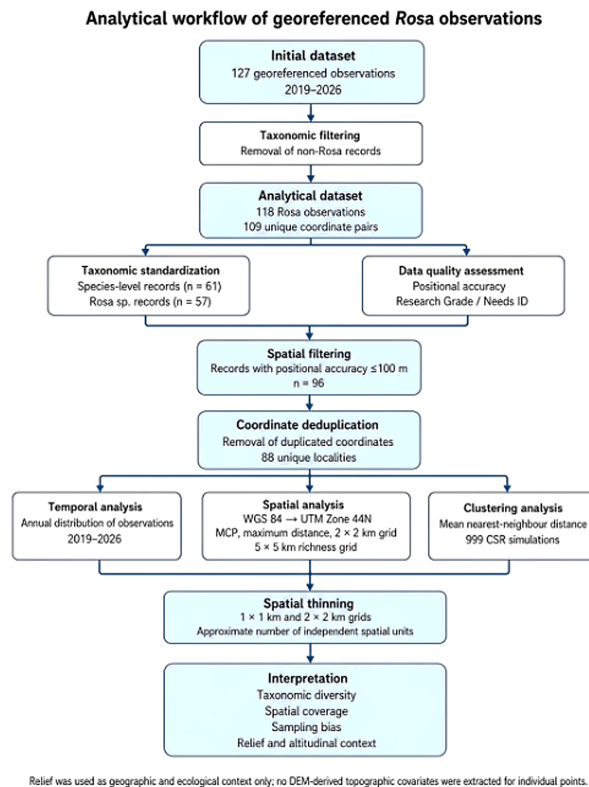


Figure 1. Analytical scheme for processing georeferenced observations of the genus *Rosa*

2.3. Taxonomic standardisation

At the species level, the dataset included *Rosa spinosissima* L., *R. laxa* Retz., *R. platyacantha* Schrenk, *R. alberti* Regel, and *R. canina* L. Records without species identification were retained in the *Rosa* sp. category and were not allocated between species based on spatial proximity or associated indirect features. This conservative approach minimises the risk of artificially inflating species representation in the context of the genus's complex taxonomy, where reliable identification requires a combined assessment of characters of leaves, stipules, spines, pedicels, sepals, flowers, and fruits (Yermagambetova, M. et.al.,2025).

2.4 Data quality assessment

Data quality was assessed using two independent components: taxonomic resolution and spatial accuracy. Taxonomic resolution was characterised by the proportion of records identified to species level, the proportion of *Rosa* sp. records, and the distribution of iNaturalist Research Grade and Needs ID statuses. Spatial accuracy was assessed using the positional_accuracy field; records with a stated error of ≤ 100 m were used for detailed spatial calculations. The proportion of observations with an error of ≤ 10 m and the median positional error were also calculated.

2.5. Time analysis

The number of observations was summarised by calendar year, separately considering records identified to the species level and records of *Rosa* sp. Because the dataset combines targeted expeditionary and opportunistic observations and is not based on a constant, standardised effort, interannual differences were interpreted solely as changes in documentation intensity, rather than changes in population abundance, density, or occurrence.

2.6. Spatial analysis

Geographic coordinates are initially presented in the WGS 84 system. To calculate distances and areas, records with a positional accuracy of ≤ 100 m were projected into UTM Zone 44N (EPSG:32644). For each species-defined taxon, the number of high-precision records and spatially unique localities, the maximum pairwise geographic distance, the area of the minimum convex polygon (MCP), and the number of occupied 2×2 km grid cells were calculated. To estimate documented local species richness, we used a 5×5 km grid and counted the number of different species-defined *Rosa* taxa in each occupied cell. We used MCP and grid metrics only as descriptive characteristics of the spatial coverage of observations. We did not interpret them as formal estimates of Extent of Occurrence or Area of Occupancy according to IUCN criteria.

2.7. Analysis of spatial clustering and sensitivity to selection bias

Spatial clustering was assessed using high-precision records (positional error ≤ 100 m) after removing completely overlapping coordinates. For the *Rosa* assemblage and for species with a sufficient number of unique localities, the mean nearest-neighbour distance (NN) was calculated. We compared the observed value with the null distribution from 999 realisations of complete spatial randomness (CSR), in which a fixed number of points were randomly placed within the MCP of the corresponding sample. The empirical p was determined from the position of the observed NN value relative to the simulated distribution. To assess the robustness of the inferences to spatial dependence, grid thinning was also performed, and the number of occupied 1×1 and 2×2 km cells was counted as an approximate number of spatially independent observation units. This approach allows estimation of the degree of pseudoreplication caused by repeated findings within the same local areas (Clark, P. J. et.al., 1954; Aiello-Lammens et.al., 2015).

2.8. Statistical processing

For categorical indicators, we calculated absolute frequencies and proportions. Because the data were collected opportunistically, taxon registration frequencies were not interpreted as indicators of relative abundance or population density. We used two-sided empirical significance levels for Monte Carlo procedures. Formal interspecies parametric comparisons of spatial metrics were not performed, as after precision filtering, removal of duplicate coordinates, and control for spatial dependence, the effective sample size for individual species was limited. All statistical results were interpreted as properties of the array of georeferenced findings, rather than as direct estimates of population demographic structure.

2.9. Using generative artificial intelligence

Generative artificial intelligence was used exclusively for language editing, structural standardisation of the draft text, and consistency checking of wording. It was not used to generate the source data, taxonomically identify observations, or perform statistical calculations. All numerical results were derived from the original dataset, and the authors reviewed and approved the final formulations, interpretations, and conclusions.

Results and discussion.

3.1. General structure and taxonomic composition of data

After taxonomic filtering, the analytical sample included 118 *Rosa* observations, corresponding to 109 unique coordinate pairs. Sixty-one records (51.7%) were identified to the species level, while 57 records (48.3%) were identified only to the genus level. Thus, the current verified dataset contains five species-identified taxa, but almost half of the observations remain taxonomically unresolved.

Table 1. Taxonomic structure of georeferenced observations of the genus *Rosa*

Taxon	Number of observations	Share, %
<i>Rosa spinosissima</i>	33	28.0
<i>Rosa laxa</i>	13	11.0
<i>Rosa platyacantha</i>	10	8.5
<i>Rosa alberti</i>	4	3.4

Taxon	Number of observations	Share, %
<i>Rosa canina</i>	1	0.8
<i>Rosa sp.</i>	57	48.3

R. spinosissima had the largest number of records—33 observations, accounting for 28.0% of the total sample and 54.1% of the records identified to species level. It was followed by *R. laxa* (13 observations), *R. platyacantha* (10), *R. alberti* (4), and *R. canina* (1). The high proportion of *Rosa sp.* limits the interpretation of the current species count as a complete estimate of the genus's taxonomic diversity in the studied part of Tarbagatai.

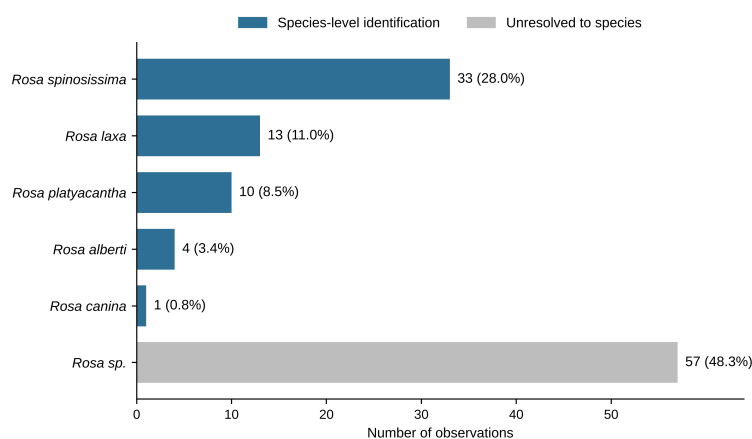


Figure 2. Taxonomic structure of the array of geo-referenced observations of the genus *Rosa*

3.2. Temporal structure of observations

The temporal distribution of observations was markedly uneven. Only isolated records were available from 2019 to 2022, while the bulk of the dataset was compiled in 2023, 2024, and 2026. The maximum number of observations was recorded in 2026—53 records, of which 27 were identified to the species level, and 26 remained at the *Rosa sp.* level.

Table 2. Annual structure of observations and level of taxonomic identification

Year	Total	Defined to the species level	<i>Rosa sp.</i>
2019	1	1	0
2020	1	0	1
2022	2	1	1
2023	22	15	7
2024	35	16	19
2025	4	1	3
2026	53	27	26
Total	118	61	57

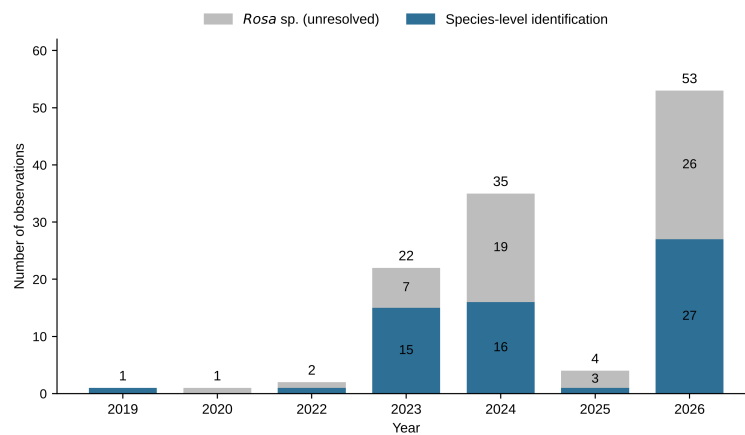


Figure 3. Interannual distribution of *Rosa* observations, divided into species-level records and *Rosa* sp. level records

The observed increase in records in recent years primarily reflects increased intensity of field and digital documentation. Because there is no standardised, ongoing observational effort, this increase cannot be interpreted as evidence of population growth or expansion. Similarly, the low number of records in 2019–2022 and 2025 reflects the structure of the available dataset, rather than the actual absence of taxa in the region.

3.3 Spatial distribution

All observations were located within 46.984–47.434° N and 80.986–82.848° E, forming a latitudinal spatial pattern generally consistent with the orographic orientation of the Tarbagatai Range and adjacent valleys. The distribution included both local clusters of repeated observations and distant localities, expanding the genus's documented spatial coverage. Overlaying the points on the relief shows a preferential concentration of observations within the mountain front, foothills, and accessible valleys; however, this visual association is not considered a quantitatively confirmed dependence on topography.

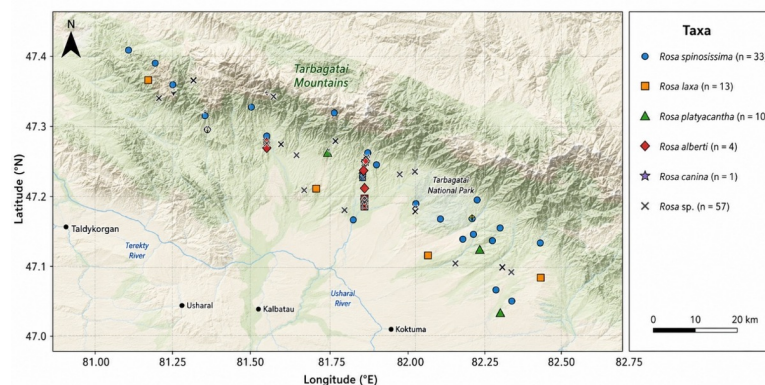


Figure 4. Spatial distribution of 118 georeferenced observations of the genus *Rosa* in the Tarbagatai mountain system against the background of the relief

Ninety-six observations with a positional accuracy of ≤ 100 m were used for spatial calculations. Of these, 44 records were identified to species level, and 52 belonged to *Rosa* sp. Among the species-identified taxa, the greatest documented spatial coverage was found for *R. spinosissima*.

Table 3. Descriptive indicators of the spatial coverage of species-defined *Rosa* records with a positional accuracy of ≤ 100 m.

View	Entries, n	Unique localities, n	Area of occupied 2 x 2 km cells, km ²	MCP, km ²	Max. distance, km
<i>Rosa spinosissima</i>	18	17	40	974.3	121.1
<i>Rosa laxa</i>	11	9	24	242.9	46.3
<i>Rosa platyacantha</i>	10	9	28	400.5	56.0
<i>Rosa alberti</i>	4	4	12	69.0	25.9
<i>Rosa canina</i>	1	1	4	---	---

For *R. spinosissima*, the maximum distance between high-precision localities was 121.1 km, and the MCP area was 974.3 km². The corresponding maximum distances were 56.0 km for *R. platyacantha*, 46.3 km for *R. laxa*, and 25.9 km for *R. alberti*. A single high-precision record of *R. canina* did not allow us to estimate the spatial extent or area of the MCP. These values characterise only the documented coverage of the available sample and are not estimates of the taxa's full regional range.

3.4 Spatial clustering and sensitivity to spatial sampling bias

Of the 96 high-precision observations, 88 corresponded to unique coordinates. For the *Rosa* set, the average distance to the nearest unique locality was 0.75 km, while the average value across 999 CSR realisations was 3.02 km (95% simulation interval 2.66–3.37 km). The observed NN distance was significantly smaller than expected under complete spatial randomness (empirical $p = 0.001$), indicating significant aggregation of the observation set. For individual species, no statistically significant deviation from the CSR was detected: for *R. spinosissima*, the observed/expected distance was 4.52/4.96 km ($p = 0.610$), for *R. laxa*, 3.79/3.44 km ($p = 0.652$), and for *R. platyacantha*, 2.62/4.34 km ($p = 0.070$). For the latter species, a tendency toward spatial aggregation was observed; however, given the number of localities available, it did not reach the accepted level of statistical significance.

Table 4. Nearest-neighbour scores and grid-thinning results for high-precision *Rosa* observations

Group	Unique localities, n	Observed NN, km	Expected NN (CSR), km	p	1 x 1 km cells	2 x 2 km cells
All Rosa	88	0.75	3.02	0.001	32	26
<i>R. spinosissima</i>	17	4.52	4.96	0.610	11	10
<i>R. laxa</i>	9	3.79	3.44	0.652	8	6
<i>R. platyacantha</i>	9	2.62	4.34	0.070	7	7

Grid thinning confirmed the high spatial redundancy of the original dataset. Ninety-six high-precision records corresponded to 32 occupied 1x1 km cells and 26 occupied 2x2 km cells. After thinning, 11 and 10 occupied cells remained for *R. spinosissima*, 8 and 6 for *R. laxa*, and 7 and 7 for *R. platyacantha*. Consequently, the nominal number of records significantly exceeds the approximate number of spatially independent observation units, which must be accounted for in subsequent statistical modelling.

3.5. Documented local species richness

Species-identified high-precision observations occupied 14 5x5 km grid cells. Eight cells recorded one species, four cells recorded two species, and two cells recorded local richness of up to four species. Both richest cells were located in the central part of the spatial cluster, approximately 47.20–47.24° N and 81.83° E. In one cell, *R. spinosissima*, *R. laxa*, *R. platyacantha*, and *R. alberti* were recorded together; in the second, *R. spinosissima*, *R. laxa*, *R. alberti*, and *R. canina*. These cells are priority sites for standardised field verification; however, their increased richness should not be separated from uneven survey intensity.

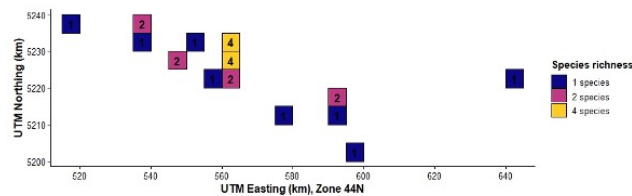


Figure 5. Documented local species richness of *Rosa* in occupied 5x5 km cells based on high-precision species-identifying observations. Numbers within cells correspond to the number of species recorded

3.6 Spatial accuracy, taxonomic verification and observational effort

Spatial accuracy was high for most of the dataset, with 85 observations (72.0%) having a reported error of ≤ 10 m and 96 (81.4%) having an error of ≤ 100 m; the median positional error was 4 m. Taxonomic verification, in contrast, remained limited, with only 15 records (12.7%) having Research Grade status, while 103 (87.3%) were Needs ID.

The main data-quality limitations were therefore not georeferencing, but taxonomic uncertainty and the concentration of observational effort. The two most active observers generated 107 of the 118 records (90.7%), indicating a significant imbalance in data sources and potentially amplifying route and spatial bias. Combined with the grid-thinning results, this confirms the need to control for spatial dependence and observer bias before any subsequent species distribution modelling.

3.7. Field photo illustrations of the morphological diversity of *Rosa*

To document morphological diversity, the article includes field photographs of five species-defined taxa from the analysed dataset (Figure 6): *Rosa laxa* Retz., *R. alberti* Regel, *R. canina* L., *R. spinosissima* L., and *R. platyacantha* Schrenk. The images cover the flowering and fruiting phases and demonstrate differences in the general appearance of flowers or hypanthia, leaflet structure, and shoot characteristics. For the genus *Rosa*, such comprehensive photo documentation is essential, as reliable field identification often requires the simultaneous assessment of several morphological traits.

Photographs served as a documentary supplement to georeferenced data and were not used as independent confirmation of each record's identification. For subsequent expert review of Needs ID observations and *Rosa* sp. records, standardised photo series are preferred, including the general appearance of the bush, spine character, stipules, both leaflet surfaces, pedicels, sepals, flowers, and mature fruits. When possible, link digital records to herbarium vouchers and specimens for molecular analysis.

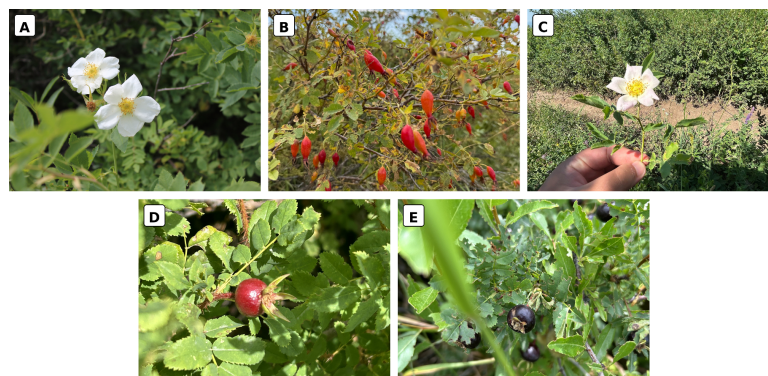


Figure 6. Field photographs of species of the genus *Rosa*: (A) *R. laxa* Retz.; (B) *R. alberti* Regel; (C) *R. canina* L.; (D) *R. spinosissima* L.; (E) *R. platyacantha* Schrenk

Discussion.

4.1. Diversity of the genus *Rosa* and the structure of observations

Analysis of the documented dataset confirmed the presence of five species-defined *Rosa* taxa: *R. spinosissima*, *R. laxa*, *R. platyacantha*, *R. alberti*, and *R. canina*. However, 48.3% of observations were identified only to the genus level, so the identified number of species should be interpreted as the minimum confirmed for the current sample, rather than as a comprehensive inventory of Tarbagatai *Rosa*. This taxonomic uncertainty directly impacts the assessment of local richness and species-specific spatial coverage.

The resulting data structure is consistent with published data on the significant role of Rosaceae in Tarbagatai vegetation. In a study of the associated flora of *Prunus ledebouriana*, the Rosaceae family was among the families with high species richness, and *R. spinosissima* was noted as a characteristic component of mid-mountain shrub communities (Orazov, A. et.al.,2021). At a broader regional level, the widespread distribution of *R. spinosissima* and *R. laxa* and their ecological plasticity were demonstrated for the Kazakh Altai (Danilova, AN. et.al.,2025). In the present study, these literature data are spatially supplemented as a georeferenced array of findings, although interpretation is limited by uneven observational effort.

4.2. Relief, altitudinal zonation and distribution of *Rosa*

The orographic heterogeneity of Tarbagatai provides important context for interpreting spatial differences in *Rosa*'s findings. Within a relatively small geographic area, foothill plains, dissected lowlands, intermontane valleys, river gorges, and mid-mountain slopes combine. This structure creates conjugate gradients of temperature, moisture content, snow cover duration, soil profile thickness, and solar radiation (Orazov, A. et.al.,2021; Myrzagaliyeva, A. et.al.,2024), potentially creating conditions for spatial segregation of the ecological niches of closely related species.

Of particular interest is the shrub and lower-middle mountain belt. According to the published scheme of southern Tarbagatai, shrub vegetation predominantly occupies altitudes of approximately 1000 (1200)–1700 (1800) m, whereas communities with a dominance or significant participation of *R. spinosissima* have been described from approximately 800–900 m up to 1350–1450 m (Orazov, A. et.al.,2021). This is consistent with the assumption that a significant portion of *Rosa*'s finds may be associated with the transition between mountain steppes and shrub communities, where increased mosaic relief, moisture differences, and heterogeneous soil conditions combine locally.

Slope exposure and position can further differentiate microhabitats. In a sharply continental climate, southern and southwestern slopes typically receive higher insolation and lose moisture more quickly, while northern exposures and lower slopes can maintain more mesophytic conditions longer. River valleys and gorges simultaneously function as linear relief elements and local zones of increased moisture availability. Consequently, the geographical overlap of different species does not preclude their ecological separation based on microrelief, exposure, and soil conditions.

The largest documented spatial range of *R. spinosissima* is consistent with data on this species' broad ecological range. In the Kazakh Altai, *R. spinosissima* has been recorded across a wider altitudinal range than several other wild roses (Danilova, AN. et.al.,2025). In Tarbagatai, its presence has previously been demonstrated in various shrub-steppe communities (Orazov, A. et.al.,2021). In the current sample, this species also had the largest maximum distance between high-precision localities and the largest MCP area. However, these indicators may reflect not only the ecological breadth of the species but also differences in its detectability and the spatial structure of field routes.

Quantitative verification of topographic differentiation requires linking each independent locality to a digital elevation model and extracting absolute elevation, slope, aspect, and derived elevation indices. Only after this procedure can the elevation-specific niches of species be statistically compared and the topographic signal separated from geographic and observer bias. Therefore, the conclusions of this study regarding the role of relief are considered ecologically valid hypotheses that require direct quantitative verification.

4.3 Spatial clustering, route bias, and effective sample size

One of the most striking results is the pronounced clustering of the *Rosa* dataset as a whole: the average distance to the nearest unique locality was approximately four times smaller than that

expected under random distribution. However, this pattern should not be interpreted as direct evidence of biological population aggregation. For opportunistic data, the observed spatial structure results from a combination of true habitat mosaicism, route accessibility, repeat visits to convenient sites, and differences in observer activity. The fact that two observers accounted for over 90% of the records, and thinning reduced the 96 high-precision observations to 26 occupied 2x2 km cells, indicates a significant contribution from spatial sampling bias.

At the level of individual species, statistically significant deviations from the CSR were not confirmed, which is likely due to the limited number of spatially independent localities. *R. spinosissima* is particularly illustrative: 33 original records, after precision filtering and thinning, corresponded to only 10 occupied 2×2 km cells. Therefore, using the nominal number of records as the effective sample size would overestimate the independence of the data. For potential distribution models, including MaxEnt, this pseudoreplication can increase the risk of overfitting and overestimating predictive ability; therefore, spatial filtering, independent taxonomic validation, and expansion of the locality network are necessary before modelling (Boria, R.A. et.al., 2014; Orazov, A. et.al., 2025).

4.4 Species-specific spatial patterns

Rosa spinosissima was the most frequently recorded and spatially widespread species; the distance between extreme high-precision localities exceeded 120 km. This pattern is consistent with the species' wide ecological range, but may also reflect its relatively high detectability and the directionality of its field routes. Therefore, the spatial breadth of the documented array should not be directly equated with the width of its ecological niche.

Rosa laxa and *R. platyacantha* were characterised by intermediate spatial coverage. The MCP area of *R. platyacantha* exceeded that of *R. laxa*, despite a smaller number of initial observations, due to the more dispersed location of the extreme localities. This result emphasises that the number of registrations and the spatial extent of the sample describe different properties of the data and should be considered together.

The four high-precision localities of *R. alberti* formed a relatively compact spatial pattern. However, such a small sample size precludes interpretation as evidence of a narrow regional or ecological niche. Differences in edaphic conditions between *R. alberti* and *R. spinosissima* have previously been demonstrated for eastern Kazakhstan (Myrzagaliyeva, A. et.al., 2026), making further comparisons of topographic and soil parameters at independently confirmed localities promising.

This single, highly accurate record of *R. canina* is purely indicative. Until independent taxonomic verification and the acquisition of additional localities, it cannot serve as a basis for concluding the species' spatial structure or regional distribution.

4.5. Documented areas of increased local species richness

Two 5x5 km cells each contained four species-defined *Rosa* taxa and represented the highest values of documented local richness in the current dataset. However, these cells are better viewed not as proven diversity hotspots, but as priority sites for independent field verification. The concentration of several taxa within a limited space provides a convenient basis for comparative study of their ecological preferences, but may also be a consequence of increased survey intensity. To verify the biological nature of the identified pattern, it is necessary to standardise comparisons of elevation, aspect, slope, indices of relief ruggedness and position, soil properties, plant community structure, and equally distributed observation effort within these cells.

4.6 Taxonomic uncertainty and iNaturalist data quality

The high proportion of Needs ID records and observations identified only to the genus level is a key limitation of taxonomic interpretation. This situation is typical for morphologically complex plant groups: even a georeferenced record with a photograph does not always contain sufficient features for reliable species identification (López-Guillén. et.al., 2024). Consequently, a digital record and coordinates are no substitute for expert review, especially when using data for species-specific spatial analysis.

Standardised recording of diagnostic traits is critical for *Rosa*. Field photographs should include the general appearance of the bush; the nature and arrangement of thorns; and the upper and

lower surfaces of leaflets, stipules, pedicels, sepals, flowers, and ripe fruits. Whenever possible, link digital observations to herbarium vouchers and molecular samples. Plastome studies of Kazakhstani *Rosa* specimens demonstrate that molecular data can significantly improve resolution in complex taxonomic cases (Yermagambetova, M. et.al.,2025). The field photographs presented in Figure 5 demonstrate the diagnostic value of recording different phenological stages but do not replace a full revision of each record.

The high positional accuracy of most observations increases the value of taxonomically uncertain records: after expert revision, they can significantly increase the number of confirmed localities of individual species, change estimates of their spatial coverage, and improve the robustness of subsequent ecological analyses.

4.7. Limitations of the study

The study is based on an opportunistic dataset generated without a single protocol and with unequal observational effort. Therefore, the frequency of registration in the database does not reflect the relative abundance, density, or probability of occurrence of populations. The pronounced clustering of the overall sample and the significant reduction in the number of units after spatial thinning indicate that the effective size of a spatially independent sample is significantly smaller than the nominal number of records. Any subsequent SDM/MaxEnt analyses on these data require prior control for spatial bias, taxonomic verification, and, preferably, external spatial validation.

The second limitation is related to taxonomic incompleteness: almost half of the records are not identified to species level, and most observations do not have Research Grade status. The third limitation is the lack of standardised topographic covariates in the original dataset, including elevation, slope, and aspect. Therefore, this study discusses the role of topography as ecological context rather than as a statistically established cause-and-effect relationship. Direct testing of topographic differentiation requires extracting DEM parameters for independent localities and then comparing them across species.

Finally, MCP and the area represented by occupied cells are sensitive to route configuration and extreme observations. These metrics describe the current spatial coverage of the data and should not be interpreted as formal conservation range assessments without further verification of survey completeness and the actual status of populations.

4.8. Prospects for monitoring, environmental analysis and conservation

High-precision coordinates provide the basis for revisiting localities, establishing permanent monitoring sites, and targeted sampling for morphological, genetic, and ecological studies. The primary objective is an expert review of the 57 *Rosa* sp. records, as well as confirmation of isolated and rare species registrations. Such a review can improve the dataset's taxonomic resolution and increase the effective number of localities for subsequent spatial analysis.

An optimal design for further research should integrate georeferenced photographic documentation, herbarium vouchers, expert taxonomic identification, and independent environmental measurements. For each confirmed locality, it is advisable to extract absolute elevation, slope, aspect, Terrain Ruggedness Index, and Topographic Position Index from the digital elevation model and combine these parameters with climate and soil variables. With a sufficient number of spatially independent data points, such an array will enable multivariate analysis of ecological space, nonparametric interspecies comparisons, quantitative assessment of niche overlap, and species distribution models with spatial cross-validation.

For Tarbagatai, where altitudinal changes occur along relatively short spatial gradients, integrating taxonomically confirmed findings with topographic and edaphic data is particularly promising. This will enable a transition from digital spatial pattern descriptions to testable hypotheses about species distribution mechanisms, as well as the identification of priority monitoring sites and repeat surveys in the face of climate and anthropogenic change.

Conclusion. A multi-year set of georeferenced observations allowed us to quantitatively characterise the documented spatial structure of *Rosa* in the Tarbagatai mountain system. The analysis included 118 observations, of which five species-defined taxa were confirmed; *R. spinosissima* was the most frequently recorded and spatially widespread species. A pronounced spatial aggregation was

revealed for the set of high-precision unique localities: the average distance to the nearest neighbour was 0.75 km, compared to the expected 3.02 km in CSR modelling ($p = 0.001$).

However, spatial thinning revealed that the 96 high-precision observations corresponded to only 32 occupied 1x1 km cells and 26 2x2 km cells; for *R. spinosissima*, 10 cells were retained at the 2x2 km scale. Therefore, the original number of registrations significantly overestimates the number of spatially independent units and should not be used as a direct equivalent of the effective sample size. Two 5x5 km cells contained four *Rosa* species each and represent priority sites for standardised field verification, but cannot be considered confirmed centres of biodiversity without additional controls for observational effort.

The key limitations of the dataset are the high proportion of taxonomically ambiguous records and the significant concentration of observational effort. Topography and altitudinal zonation are ecologically plausible explanations for some of the spatial heterogeneity, but this requires direct verification using DEM parameters and standardised field data.

Expert verification of 57 *Rosa* sp. records, expansion of the network of independent localities, and integration of topographic, soil, and climatic covariates will enable a transition from descriptive findings to a more statistically rigorous assessment of ecological niches and potential species distributions. In its current form, the dataset is best used as a spatial basis for targeted monitoring, repeated field verification, and planning of future *Rosa* studies in Tarbagatai.

Funding. This research was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan, grant No. AP22687057, “Assessment of the Current State of Shrubs in Eastern Kazakhstan Based on the Bioecological and Genetic Characteristics of *Rosa* L. Populations”.

References

- Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2015). spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, 38, 541–545. <https://doi.org/10.1111/ecog.01132>
- Aimenova, Z. E., Matchanov, A. D., Esanov, R. S., Sumbembayev, A. A., Duissebeyev, S. E., Dzhumanov, S. D., & Smailov, B. M. (2023). Phytochemical profile of *Eranthis longistipitata* Regel from three study sites in the Kazakhstan part of the Western Tien Shan. *Biodiversitas*, 24, 6031–6038. <https://doi.org/10.13057/biodiv/d241124>
- Boria, R. A., Olson, L. E., Goodman, S. M., & Anderson, R. P. (2014). Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, 275, 73–77. <https://doi.org/10.1016/j.ecolmodel.2013.12.012>
- Clark, P. J., & Evans, F. C. (1954). Distance to nearest neighbour as a measure of spatial relationships in populations. *Ecology*, 35, 445–453. <https://doi.org/10.2307/1931034>
- Danilova, A. N., & Sumbembayev, A. A. (2021). The status of the *Dactylorhiza incarnata* populations in the Kalba Altai, Kazakhstan. *Biodiversitas*, 22, 3180–3195. <https://doi.org/10.13057/biodiv/d220812>
- Danilova, A. N., Vdovina, T. A., Kotukhov, Y. A., Anufriyeva, O. A., Vinokurov, A. A., Isakova, E. A., Lagus, O. A., & Sumbembayev, A. A. (2025). Study of the geographical distribution, ecological–biological characteristics, and economic value of *Rosa acicularis* Lindl., *Rosa laxa* Retz., and *Rosa spinosissima* L. (Rosaceae) in Kazakhstan's part of the Altai Mountains. *Diversity*, 17, Article 441. <https://doi.org/10.3390/d17070441>
- Imanbayeva, A., Ishmuratova, M., Duisenova, N., Sagyndykova, M., Orazov, A., & Tuyakova, A. (2025). Floristic and anatomical diversity of *Crataegus ambigua* C. A. Mey. ex A. K. Becker populations in different areas of the arid Mangystau region (Kazakhstan). *Forests*, 16, Article 585. <https://doi.org/10.3390/f16040585>
- Imanbayeva, A., Zharassova, D., Duisenova, N., Orazov, A., Tolep, N., & Tlepiyeva, G. (2025). Unravelling the chloroplast genome of the Kazakh apricot (*Prunus armeniaca* L.) through MinION long-read sequencing. *Plants*, 14, Article 638. <https://doi.org/10.3390/plants14050638>
- iNaturalist. (2026). Georeferenced observation export used in this study [Data set]. Retrieved August 27, 2026, from <https://www.inaturalist.org/>
- Kubentayeva, S., Levichev, I., Nobis, M., Baasanmunkh, S., Klichowska, E., Orazov, A., Alibekov, D., Kubentayeva, B., & Choi, H. J. (2025). Integrative taxonomic revision concerning phylogenetic relationships, morphological variation, distribution, ecology and conservation of *Gagea alberti* (Liliaceae) within its natural range. *Botanical Studies*, 66, Article 16. <https://doi.org/10.1186/s40529-025-00465-z>
- López-Guillén, E., Herrera, I., Bensid, B., Gómez-Bellver, C., Ibáñez, N., Jiménez-Mejías, P., Mairal, M., Mena-García, L., Nualart, N., Ujtes-Mascó, M., & López-Pujol, J. (2024). Strengths and challenges of using iNaturalist in plant research with focus on data quality. *Diversity*, 16, Article 42. <https://doi.org/10.3390/d16010042>
- Myrzagaliyeva, A., Irsaliyev, S., Tustubayeva, S., Samarkhanov, T., Orazov, A., & Alemseitova, Z. (2024). Natural resources of *Rhaponticum carthamoides* in the Tarbagatai State National Nature Park. *Diversity*, 16, Article 676. <https://doi.org/10.3390/d16110676>
- Myrzagaliyeva, A., Kauanov, S., Tustubayeva, S., Irsaliyev, S., Sharipova, M., & Orazov, A. (2026). Edaphic niche differentiation exceeds field ecophysiological divergence among wild *Rosa* species across the mountain systems of Eastern Kazakhstan. *Biology*, 15, Article 1247. <https://doi.org/10.3390/biology15151247>
- Orazov, A., Myrzagaliyeva, A., Tustubayeva, S., Irsaliyev, S., Sailaubekova, A., & Turalin, B. (2025). Valuable morphological traits and genetic resources of wild almond relatives in Kazakhstan. *ES Food & Agroforestry*, 20, Article 1577. <https://doi.org/10.30919/faf1577>
- Orazov, A., Tustubayeva, S., Alemseitova, J., Mukhitdinov, N., Myrzagaliyeva, A., Turuspekov, Y., & Sramko, G. (2021). Flora accompanying *Prunus ledebouriana* (Schltdl.) Y. Y. Yao in the Tarbagatai State National Park in Kazakhstan. *International Journal of Biology and Chemistry*, 14, 21–34. <https://doi.org/10.26577/ijbch.2021.v14.i1.02>

- Tayjanov, K., Mamadalieva, N. Z., & Wink, M. (2017). Diversity of the mountain flora of Central Asia with emphasis on alkaloid-producing plants. *Diversity*, 9, Article 11. <https://doi.org/10.3390/d9010011>
- Yermagambetova, M., Imanbayeva, A., Ishmuratova, M., Sumbembayev, A., & Almerikova, S. (2025). Characterisation of the four *Rosa* L. species from Kazakhstan based on complete plastomes and nuclear ribosomal internal transcribed spacer (ITS) sequences. *Genes*, 16, Article 852. <https://doi.org/10.3390/genes16080852>