Examining the Relationship Between Tatara Iron Production and Grassland Distribution in Western Japan: An Open Geospatial Approach to Historical Landscape Analysis

Ayaka Onohara [1], Takatora Bito [2], Nobusuke Iwasaki [2]

[1] Faculty of International Studies, Hokusei Gakuen University, Hokkaido, Japan - onohara@hokusei.ac.jp
[2] Faculty of Agriculture, Tottori University, Tottori, Japan - niwasaki@tottori-u.ac.jp

Keywords: Tatara iron production, Grassland distribution, Landscape memory, DBSCAN, Open geospatial workflow

ABSTRACT

In western Japan, traditional tatara iron production has long been associated, both in historical narratives and in early ecological work, with the persistence of grassland landscapes. This view was stated by Itow (1962), but the inferred link has rarely been examined with spatially explicit data. We examine the spatial relationship between tatara iron production sites and grassland distribution in Tottori Prefecture using an integrated open-source geospatial workflow. The two data layers are temporally offset by design: tatara sites spanning from the medieval period to the mid-Meiji era, while the earliest spatially explicit municipality-level grassland record—the 1950 World Agricultural Census—captures the “landscape memory” of those activities roughly half a century after the cessation of large-scale tatara. Census tables were digitised using AI-assisted, sub-page table extraction. Tatara site locations (N = 562) were compiled from the Tottori Prefecture cultural-heritage WebGIS. Geology, basin structure, and DEM-derived terrain were processed with GeoPandas, rasterio, scipy, and rasterstats; OLS regression, model comparison, and DBSCAN clustering were implemented in statsmodels and scikit-learn. Tatara sites form well-defined production districts in the granitic Hino River basin, but grassland is most strongly predicted by elevation, with tatara density adding only modest marginal information. The count × slope interaction is significantly negative, indicating that the apparent tatara–grassland association weakens sharply in steep terrain. We interpret these patterns as evidence that both phenomena occupy the same mountain environmental niche, consistent with a niche-construction reading rather than a direct landscape-transformation interpretation.

1. Introduction

Grasslands historically occupied a substantially larger portion of the Japanese landscape than they do today (Arizono, 1995). In western Japan, traditional *tatara* (鑪) iron production—a pre-industrial smelting technology using iron sand and charcoal in clay furnaces—has frequently been associated with the persistence of these open landscapes. Historical narratives often portray grasslands as by-products of the massive forest clearance and charcoal extraction that iron production required, together with the *kanna-nagashi* (鉄穴流) iron-sand-washing process that disturbed hillslopes in granitic terrain.

Early ecological work reinforced this interpretation. Itow (1962) noted a spatial correspondence between the distribution of tatara production areas and grassland occurrence in the Chūgoku region, an observation that has since been widely cited as evidence of a direct causal relationship. Despite the persistence of this interpretation, it has rarely been subjected to rigorous quantitative examination using spatially explicit data. In particular, the possibility that both tatara production and grassland landscapes were independently shaped by shared environmental conditions has received little systematic attention.

The available historical data require us to confront an intrinsic temporal mismatch. Tatara production in western Japan has roots extending to the medieval period, though large-scale commercial operations were concentrated in the late Edo to Meiji era; by the early twentieth century the last traditional furnaces had closed. The earliest nationwide spatially explicit, municipality-level enumeration of grassland in Japan, however, is the 1950 World Agricultural Census (MAF, 1955). This study therefore treats 1950 grassland not as a contemporary of tatara operation, but as a “landscape memory” indicator—the spatial residue of mountain land-use regimes that overlapped with and outlasted the active tatara period, recorded shortly after the postwar land reform and before the massive 1960s conversion of mountain grassland into coniferous plantation. The corresponding interpretive horizon is one of long-term landscape signatures rather than direct temporal causation

This study has three connected aims. The first is substantive: to test whether the spatial association between tatara sites and grassland persists once basic environmental controls—geology, basin structure, and terrain—are taken into account. The second is theoretical: to assess whether the resulting pattern is more consistent with a direct landscape-transformation narrative or with a shared-niche interpretation grounded in cultural-evolutionary and niche-construction thinking. The third is methodological: to demonstrate that the entire workflow—from digitising mid-twentieth-century printed statistics to fitting spatial regression and clustering models—can be assembled from FOSS4G components in a fully reproducible, FAIR-aligned manner.

2. Study Area

Tottori Prefecture occupies the northern flank of the Chūgoku Mountains and was historically one of the most active tatara iron production regions in Japan. Its geology is dominated by granitic and granodioritic formations in the west and centre, with distinct Quaternary volcanic deposits surrounding the Daisen volcano in the northwest. Three major river basins—the Sendai, Tenjin, and Hino river systems—organise the mountainous landscape into hydrologically distinct units. The Hino River basin, which drains the granitic mountains of the south-west, is the historical core of tatara production within the prefecture.

The combination of geological contrast (granite versus volcanic), strong basin segmentation, and a steep topographic gradient from the Sea of Japan coast to the Chūgoku divide makes Tottori a natural laboratory for examining whether environmental context shapes the co-distribution of iron production and grassland land use. At the spatial unit used in this study—the 1950 municipality boundaries—the prefecture comprises 55 administrative units. Figure 1 shows the prefecture with 1950 grassland area, dominant basin boundaries, and the spatial distribution of the N = 562 tatara sites used in the analysis. Municipality polygons are shaded by 1950 grassland area (quantile classes); dashed lines mark basin boundaries; black dots are tatara iron-production sites.

3. Data

The analysis integrates three classes of data: historical grassland statistics, tatara iron production site locations, and environmental geospatial layers.

3.1 Grassland Area from the 1950 Agricultural Census

Grassland area at the municipality level was derived from the 1950 World Agricultural Census of Japan, which provides the earliest spatially comprehensive record of pasture and managed grassland (hereafter "grassland") in Japan. The variable specifically enumerated is privately or commonly held mowing/grazing land—the mountain semi-natural grassland category historically maintained for fodder, manure, and seasonal grazing under *iriai*（入会）common-rights tenure—rather than improved sown pasture in the modern agronomic sense.

The choice of 1950 deserves explicit comment, because tatara operations had largely ceased several decades earlier, and 1950 itself sits in a particular postwar interval. Three considerations justify this dataset as a reasonable, indeed near-optimal, proxy for the "landscape memory" of the tatara-era mountain land-use regime within the constraints of available historical statistics. First, 1950 is the earliest year in which grassland area is enumerated at the municipality level on a fully comparable national basis; earlier prefectural reports record area only at coarser administrative scales. Second, 1950 precedes the postwar expansion of conifer plantations; afforestation policies shifted from cutover-land restoration to the active conversion of broadleaved forest and open land after the mid-1950s (Matsushita, 2015), driving a rapid reduction in meadows and pastures across western Japan from the early 1960s onwards (Ohta et al., 2022). Census records from 1960, 1965, or 1975 therefore capture progressively transformed landscapes in which the residue of pre-industrial land use is increasingly diluted. Third, iriai common-land tenure—which had collectively managed mountain grasslands for fodder, manure, and fuel since well before the Meiji era—remained functionally intact in 1950; the postwar substitution of fossil fuels and chemical fertilizers subsequently removed the economic rationale for maintaining these commons from the 1960s.

We therefore treat 1950 grassland as a *long-term landscape signature*—the integrated outcome of mountain land-use regimes that overlapped with the active tatara period and persisted into the early postwar—rather than a contemporary covariate of tatara operation.

Because the original records exist only as printed statistical tables, a structured digitisation protocol was developed. Sub-page regions containing the relevant columns were isolated, and AI-assisted table extraction was performed using Claude Sonnet 4.5 (Anthropic, 2025) via Web interface, converted each cropped region into structured CSV records. The extracted records were then manually verified against the original printed tables. Working at sub-page rather than whole-page scale was an explicit design choice: when an entire page was passed to the LLM-based extractor, transcription errors and column-misalignment problems were considerably more frequent than when a tightly cropped area containing only a few rows was supplied. All numeric values were cross-checked against the printed tables. The resulting dataset covers all municipalities in Tottori Prefecture as they existed in 1950.

3.2 Tatara Site Locations

Tatara iron production site locations were compiled from the Tottori Prefecture WebGIS cultural heritage database, *Tottori Web Map* (Tottori Prefectural Government, 2020). Raw spatial features were extracted via browser developer tools, parsed with Python, and converted to a clean point dataset including site name, classification, and coordinates. Points were aggregated to 1950-era municipalities using a spatial join in QGIS 3.40 (QGIS Development Team, 2024). After de-duplication and harmonisation, N = 562 tatara-related sites were retained.

3.3 Environmental Geospatial Datasets

Geological composition was derived from the 1:200,000 Seamless Geological Map of Japan (Geological Survey of Japan, 2025). River basin membership was based on the national hydrological mesh dataset (MLIT, 2009). Terrain variables were derived from the GSI Elevation Tile DEM in Zoom Level 13 （EPSG:6673, JGD2011 / Japan Plane Rectangular CS V) (Geospatial Information Authority of Japan, 2026). Elevation,

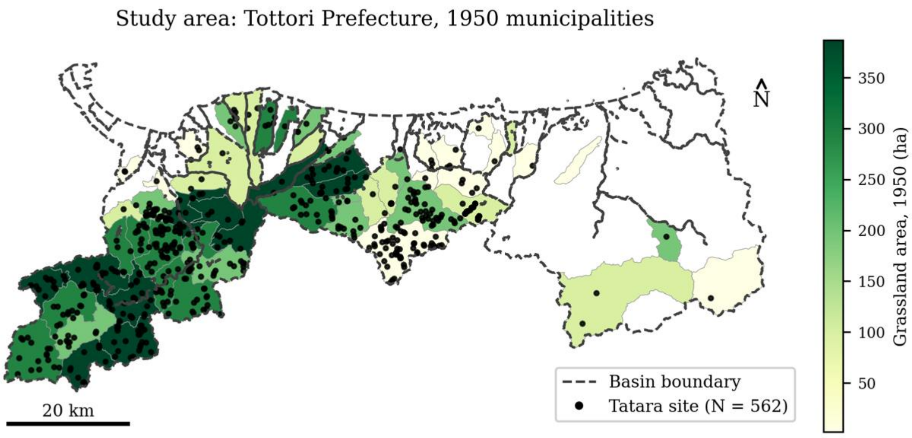

Figure 1. Study area

slope, and curvature were summarised per 1950 municipality; the computational procedure is detailed in Section 4.1. Geological diversity was summarised as a normalised Shannon entropy over rock-type categories (geology_entropy), together with explicit shares of igneous, sedimentary, and metamorphic substrate. Basin entropy and a dominant-basin variable (basin_main) were derived analogously.

4. Methods

4.1 Open Geospatial Workflow

Spatial data processing and analysis used Python-based open-source libraries. GeoPandas (Jordahl et al., 2020) handled vector data integration, spatial joins, and overlay operations between DBSCAN cluster hulls and municipality polygons. Slope and curvature were computed from the DEM by Gaussian smoothing and finite-difference partial derivatives, using scipy (Virtanen et al., 2020) gaussian_filter and NumPy (Harris et al., 2020) gradient, with slope expressed in degrees and curvature as the Laplacian of elevation; raster I/O used rasterio (Gillies et al., 2013). Zonal statistics (mean and standard deviation of elevation, slope, and curvature) were aggregated to 1950 municipalities with rasterstats (Perry, 2019). Vector data were managed in FlatGeobuf formats. QGIS 3.40 (QGIS Development Team, 2024) was used for cartographic checks and initial site-to-municipality spatial joins. Statistical modelling was implemented using statsmodels (Seabold and Perktold, 2010) and scikit-learn (Pedregosa et al., 2011). The full pipeline is implemented as a single Python script (tatara_grassland_analysis.py) in the project repository.

4.2 Statistical Modelling

The substantive analysis proceeded in three stages, with the 1950 municipality ($N = 55$) as the unit of analysis throughout.

4.2.1 Stage 1 – Environmental baseline: Ordinary least squares (OLS) regression was used to assess the relationship between environmental variables and municipality-level grassland area. Candidate predictors included geological entropy, igneous/sedimentary/metamorphic shares, basin entropy and dominant basin membership, and DEM-derived mean and standard deviation of elevation, slope, and curvature. AIC and adjusted R^2 guided the selection of a parsimonious environmental baseline.

4.2.2 Stage 2 – Marginal and conditional contribution of tatara: The tatara site count per municipality (count) was added to the environmental baseline, and stepwise comparison assessed whether iron-production density carried explanatory power beyond shared environmental constraints. To probe the conditional structure of the relationship, two interaction terms were considered: count × slope_mean (terrain-conditional adaptation) and count × geology_entropy (geology-conditional adaptation). The integrated final specification retained both. Predictors entering interactions were inspected for multicollinearity; condition numbers and the stability of coefficient signs across specifications are reported. We explicitly note that OLS estimates are vulnerable to unobserved confounders—plausible candidates include historical population density, iriai tenure intensity, and local microclimate (winter snow cover, precipitation gradient)—and we therefore treat the recovered coefficients as conditional associations rather than as causal effects.

4.2.3 Stage 3 – Diagnostics and spatial structure: Residuals from the final OLS model were inspected for normality (Jarque–Bera, skewness, kurtosis). Because municipality-level grassland and tatara counts are likely to be spatially autocorrelated—both through shared environmental gradients and through possible cultural spillover among neighbouring valleys—the spatial structure of the data is addressed in two complementary ways. First, the OLS specification absorbs the dominant component of spatial structure indirectly through its environmental predictors (elevation, slope, basin), which themselves carry strong spatial autocorrelation. Second, and more directly, the DBSCAN clustering described in Section 4.3 is used as an explicit, non-parametric summary of the residual spatial organisation of tatara sites, allowing comparison between regression-based and cluster-based views of the same pattern. We acknowledge that a fully spatially explicit econometric model (e.g., a spatial-error or spatial-lag specification, or a hierarchical Bayesian count model with neighbourhood random effects) would be the natural next step at a basin or multi-prefecture scale; the present study lays the methodological groundwork for that extension. As a supplementary discrimination check, a logistic regression was trained to classify “high-grassland” municipalities (top quartile of grassland area) using the full feature set; performance was evaluated under 5-fold stratified cross-validation (ROC-AUC, accuracy, F_1) inside a scikit-learn Pipeline combining median/most-frequent imputation, standardisation, and one-hot encoding.

4.3 Spatial Clustering of Tatara Sites

To complement the area-based regression analysis, DBSCAN clustering was applied to the tatara point pattern with parameters $\varepsilon = 3000$ m and minPts = 5. Points were projected to JGD2011 / Japan Plane Rectangular CS IX for the clustering, as in the reference implementation. The parameter pair was chosen to recover clusters at the scale of historical production districts—typically groups of adjacent valleys—rather than at the scale of individual furnace sites or of the prefecture as a whole. Resulting clusters were characterised by their environmental context by intersecting cluster convex hulls with municipality polygons (GeoPandas overlay) and aggregating per-municipality variables, allowing direct comparison with the regression-based view.

4.4 Reproducibility

The complete analysis code, dependency list, and documentation are openly available on GitHub at https://github.com/aonoa68/tatara-grassland-tottori, with a version-tagged snapshot archived on Zenodo (Onohara, 2026). All software components used—QGIS, GeoPandas, rasterio, scipy, rasterstats, statsmodels, scikit-learn—are released under OSI-approved open-source licences. Source datasets are publicly available from the agencies cited in Section 3.

5. Results

5.1 Spatial Clustering of Tatara Sites

DBSCAN identified four clusters and a small noise component in the 562-site point pattern. Two clusters dominate: 314 sites in the upper Hino River basin (cluster 3) and 206 sites in adjacent areas of the same basin (cluster 0). Two further clusters of 9 sites each sit on the basin margins, and 24 sites are flagged as noise (Table 1).

Cluster ID	Number of tatara sites	Number of municipalities intersecting cluster convex hull	Mean elevation (m)	Slope (°)	Mean municipality-level grassland area (ha)	Mean tatara site count per municipality
3	314	23	434.6	19	121.8	14.5
0	206	25	292	17.9	50.9	11.7
1	9	7	197.7	12.6	50	3
2	9	3	195.3	10	28	3.3
−1 (noise)	24	–	–	–	–	–

Table 1. DBSCAN clusters of tatara sites and their environmental context. "Sites" is the raw count of tatara points; remaining columns are means computed over municipalities intersecting each cluster's convex hull.

The two principal clusters are not merely spatial groupings; their environmental signatures differ in interpretable ways. Cluster 3 occupies higher elevations (mean ≈ 435 m), steeper terrain (mean slope ≈ 19°), and is associated with higher municipality-level grassland extent (mean ≈ 122 ha). Cluster 0 sits at lower elevations (mean ≈ 292 m), on gentler slopes (mean ≈ 18°), and with less than half the grassland extent (mean ≈ 51 ha), at a similar tatara density. The smaller clusters fall at still lower elevations and gentler slopes. Geologically, both major clusters fall predominantly on granitic substrate.

5.2 Environmental Drivers of Grassland Distribution

The environmental baseline explains a substantial share of the variance in municipality-level grassland area (terrain-only specification: AIC = 633.9, $R^2 = 0.459$). Across all specifications, mean elevation is by far the most consistent predictor ($\beta \approx 0.45$ in the full-controls model; $t = 5.34$, $p < 0.001$). Elevation standard deviation enters with a negative coefficient ($p = 0.003$), and slope mean is significant once elevation variance is controlled ($p = 0.007$). Higher-elevation municipalities—particularly those with relatively uniform mountainous topography—hold substantially more grassland, consistent with both the documented prevalence of mountain pasture in the uplands of Tottori and the ecological preference of semi-natural grassland communities for cool, well-drained mountain sites. Basin contrasts and geological entropy add comparatively little once terrain is accounted for.

5.3 Marginal and Conditional Effects of Tatara Density

Adding raw tatara site counts on top of the environmental baseline improves model fit only modestly (AIC = 634.7, $R^2 = 0.471$), and the main effect of count is not significant on its own ($p = 0.34$). Once a count × slope interaction is introduced, the picture changes substantially. In the interaction model, the main effect of tatara count becomes large and significant ($\beta = 17.80$, $p < 0.001$), and the interaction itself is significant and negative ($\beta = -0.79$, $p < 0.001$), with overall fit improving ($R^2 = 0.485$, adjusted $R^2 = 0.408$, $F = 6.31$, $p = 3.0 \times 10^{-5}$). Municipalities with many tatara sites in gently sloping terrain show a strong positive association with grassland area, while in steeper terrain the same activities are associated with markedly weaker grassland signals.

The integrated final specification (grass ~ geology_entropy + basin_entropy + elev_mean + slope_mean + curv_mean + count + count:geology_entropy + count:slope_mean) gives the same picture: $R^2 = 0.491$, adjusted $R^2 = 0.402$, $F = 5.54$, $p = 6.3 \times 10^{-5}$; elevation remains the dominant environmental predictor ($\beta = 0.25$, $p < 0.001$), the main effect of tatara count is positive and significant ($\beta = 14.99$,

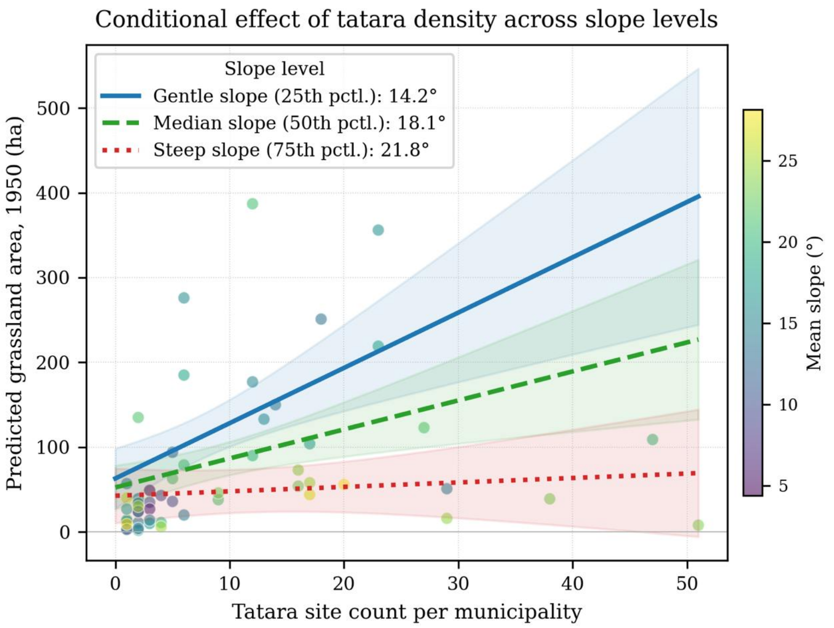

Figure 2. Conditional effect of tatara site count on 1950 grassland area at three representative slope levels (25th, 50th, and 75th percentiles of slope_mean).

$p = 0.014$), and the count × slope interaction is significant and negative ($\beta = -0.71$, $p = 0.005$). The competing count × geology_entropy interaction is not significant in the integrated model ($p = 0.45$), indicating that the conditional structure of the tatara effect is captured primarily by terrain rather than geology.

A complementary observation arises from comparing the regression view with the DBSCAN clusters: the production district that does sit in high-elevation, grassland-rich terrain (cluster 3) is precisely the one where the broader environmental signal (elevation) is itself most favourable to grassland, while cluster 0 has comparable tatara density but markedly lower grassland extent. Figure 2 visualises the conditional structure of this relationship by plotting the model-predicted grassland area against tatara count at three representative slope levels: at gentle slopes (25th percentile of slope_mean), more tatara sites are associated with substantially more grassland; at steep slopes (75th percentile), the slope of the predicted line is markedly shallower, and the predicted line at the 75th percentile of slope stays consistently below the gentle-slope line over the observed range of count.

5.4 Supplementary Classification and Residual Diagnostics

A 5-fold cross-validated logistic regression discriminating high-grassland municipalities (top quartile) using the full feature set achieved a mean ROC-AUC of 0.79, accuracy of 0.73, and F_1 of 0.59, indicating moderate but non-trivial discriminative power consistent with the regression results.

Residuals of the final OLS model show clear deviations from normality (skewness = 0.93, kurtosis = 6.7, Jarque–Bera $p \approx 2 \times 10^{-9}$), reflecting the strongly right-skewed, non-negative distribution of municipality-level grassland area. The condition number of the design matrix is high ($\approx 3.6 \times 10^6$), driven by the joint inclusion of raw count and interaction terms on different scales; coefficient signs and p-values are nevertheless stable across specifications. The estimates are best interpreted as descriptive summaries of the conditional structure of the relationship rather than as a fully specified causal model.

6. Discussion

6.1 Revisiting the Itow (1962) Interpretation

Taken together, the results qualify rather than refute the long-standing narrative associating tatara iron production with grassland landscapes. Tatara sites are not randomly distributed: they form well-defined production districts concentrated in granitic, mid-to-high elevation parts of the Hino River basin. Grassland is, in turn, more abundant in the mountainous parts of the prefecture. The spatial correspondence observed by Itow (1962) and subsequent authors is therefore real at the scale of regional pattern.

The mechanism behind that correspondence, however, looks different from a direct landscape-transformation story. Grassland distribution is explained primarily by elevation, with tatara density adding modest marginal information mainly through its interaction with slope. The cleanest reading is that tatara production and grassland landscapes share a common environmental opportunity space, defined jointly by geology, basin structure, and terrain. Where these two opportunity spaces overlap—in moderately sloping mid-elevation terrain on granitic substrate—tatara may indeed have helped to maintain or extend open land; in steeper settings, the same level of iron-production activity does not translate into a comparable grassland footprint.

This interpretation is sensitive to the temporal mismatch between the two data layers, and that sensitivity cuts in a specific direction. The half-century gap between the cessation of large-scale tatara and the 1950 census means that any direct landscape-transformation effect of tatara should, if anything, be attenuated in our data by intervening land-use change. The fact that our analysis nevertheless recovers a clear environmental signal (elevation) and only a conditional cultural signal (the count × slope interaction) is therefore more conservative against a strong-causal reading of the Itow interpretation than against a shared-niche reading.

6.2 A Niche-Construction Reading

This pattern fits naturally into a niche-construction frame. Tatara can be read as a technological specialisation adapted to a particular ecological niche—granitic mountains with sufficient relief and water for iron-sand washing and a forest resource base capable of sustaining charcoal production. The cultural activity is shaped by the environment; in return, sustained operation likely modified vegetation, soil disturbance regimes, and land-management practices within that niche. From this perspective, the apparent correlation between industrial sites and grassland landscapes does not require a single, monocausal forest-clearance mechanism. It arises partly because both phenomena are anchored in the same environmental constraints, and partly because, in a subset of conditions, the cultural activity reinforces and stabilises the open-land state that the environment already favours.

6.3 Methodological Contribution: A FAIR FOSS4G Workflow

Beyond its substantive findings, the study illustrates that open geospatial workflows can support rigorous historical landscape analysis end-to-end. AI-assisted, sub-page digitisation of mid-twentieth-century statistical tables provides a tractable path from printed primary sources to analysis-ready CSVs, without the deterioration in accuracy typical of whole-page OCR on tabular material. GeoPandas, rasterio, scipy, and rasterstats together handle the geospatial side of the pipeline, including raster–vector integration and zonal statistics. FlatGeobuf, an openly specified binary format, serves as a durable on-disk format. Statistical modelling and clustering rely on statsmodels and scikit-learn. The complete pipeline is archived on Zenodo (Onohara, 2026), satisfying FAIR principles (Wilkinson et al., 2016): findable via a DOI, accessible without paywalls, interoperable through open formats, and reusable under an explicit open-source licence.

6.4 Limitations and Threats to Validity

Several limitations qualify the strength of the inferences and point to natural extensions.

6.4.1 Temporal mismatch between the two layers: Tatara operations had largely ceased by the early twentieth century, whereas the grassland enumeration is from 1950. As argued in Section 3.1, we treat 1950 grassland as a long-term landscape signature rather than as a contemporaneous outcome of tatara activity; the analysis is correspondingly framed in terms of co-occurrence within a shared environmental niche rather than direct causation. The bias introduced by the gap is, as discussed in Section 6.1, plausibly conservative against a strong-causal reading: half a century of intervening land-use change should attenuate, not amplify, any persistent tatara-induced grassland signal. Extending the analysis to earlier cadastral records (where partial Meiji-era land-use surveys exist for some municipalities) and to a 1950–1960–1975 panel of the same census series would allow direct examination of how the relationship has decayed over time.

6.4.2 Unobserved confounders: OLS estimates are vulnerable to bias from variables we did not measure. The most plausible candidates are historical population density (denser settlements in narrow valleys may simultaneously have hosted tatara furnaces and maintained pasture for animals), iriai tenure intensity, and local microclimate (precipitation, winter snow cover, length of growing season). The reported coefficients should therefore be read as conditional associations within the available control set, not as causal effects. In particular, historical population density is a high-priority extension: code scaffolding for incorporating municipality-level population layers is already present in the analysis script.

6.4.3 Spatial autocorrelation and cultural diffusion: Tatara sites in the Chūgoku region are widely understood to have spread along apprenticeship and supply networks, which generates spatial spillovers that an OLS model on municipality units, which treats them as independent observations, does not capture. Our environmental predictors (elevation, slope, basin) absorb the dominant spatial gradients indirectly, and DBSCAN clustering provides an explicit non-parametric description of the residual spatial organisation of tatara sites. However, a full spatial econometric treatment—a spatial-lag or spatial-error specification on the area-based outcome, or a hierarchical Bayesian count model with conditional autoregressive random effects at the municipality level—remains a natural and important next step, particularly when extending the analysis to a multi-prefecture Chūgoku-wide setting.

6.4.4 Sample size and unit choice. The unit count (N = 55 municipalities) is small for elaborate regression specifications, especially when categorical basin and geology variables are included. Confidence intervals on individual coefficients are correspondingly wide. We have therefore emphasised effect signs and significance of stable terms across specifications rather than precise effect magnitudes. Aggregating into broader basin- or geology-defined units, or pooling Tottori with neighbouring prefectures (Shimane, Okayama, Hiroshima), would improve power, at the cost of merging units that may have differed in tenure and cultural history.

6.4.5 Geological scale. Geological substrate is summarised at the municipality level. Finer-scale distinctions—particularly between granitic and non-granitic substrate within the same municipality—are likely to refine the interpretation, especially when extending the analysis to neighbouring prefectures where the granite/non-granite contrast is more pronounced.

7. Conclusions

The widely cited spatial correspondence between tatara iron production and grassland distribution in western Japan is real, but its mechanism appears to be largely environmental rather than directly anthropogenic. In Tottori Prefecture, grassland distribution is dominated by elevation; tatara sites are strongly clustered in granitic mountains of the Hino River basin; and tatara density carries clear marginal explanatory power for grassland only in interaction with slope. Given the 50-year offset between the cessation of large-scale tatara and the 1950 grassland record, this pattern is more consistent with a shared-niche interpretation, in which both tatara siting and mountain grassland reflect a common environmental opportunity space, than with a direct landscape-transformation reading.

Methodologically, the work demonstrates that a fully open, FAIR-compliant geospatial workflow—built on QGIS, GeoPandas, rasterio, rasterstats, statsmodels, scikit-learn, and FlatGeobuf, combined with AI-assisted digitisation of printed historical statistics—is sufficient for the spatially explicit examination of long-standing landscape-history claims and is transferable to other historical industrial landscapes. The natural extensions—an explicit spatial econometric specification, the addition of historical population density, and a multi-prefecture Chūgoku-wide replication—are tractable on the same open-source stack.

Acknowledgements

Part of this work was presented in an undergraduate thesis submitted to Tottori University in 2025. The authors thank the members of the Socio-Ecological Landscape Science course for their valuable advice. This work was supported by JSPS KAKENHI Grant Number JP25K21880 and the Environment Research and Technology Development Fund (JPMEERF25S12421) of the Environmental Restoration and Conservation Agency of Japan, funded by the Ministry of the Environment, Japan.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the authors used generative AI and AI-assisted technologies, including Claude Sonnet 4.5 (Anthropic), to assist with translation of portions of the manuscript, improvement of readability and language, and drafting of selected code segments. After using these tools, the authors reviewed, tested, revised, and edited the outputs as needed and take full responsibility for the content of the published article.

References

Anthropic, 2025. Claude Sonnet 4.5. Anthropic. www.anthropic.com/claude (10 December 2025).

Arizono S., 1995. 1.2 Land Use in Japan circa 1850 (pp.4-5), Nishikawa O., Himiyama Y. et al. eds., *Atlas - Environmental Change in Modern Japan* (in Japanese), Asakura Publishing Co. Ltd., Tokyo.

Geological Survey of Japan, AIST, 2025. 1:200,000 Seamless Geological Map of Japan V2. National Institute of Advanced Industrial Science and Technology, Tsukuba. gbank.gsj.jp/seamless/ (10 December 2025).

Geospatial Information Authority of Japan, 2026. GSI Elevation tile. Tsukuba. maps.gsi.go.jp/development/ichiran.html#dem (10 December 2025).

Gillies, S., et al., 2013. Rasterio: geospatial raster I/O for Python programmers. Mapbox, github.com/rasterio/rasterio (10 December 2025).

Harris, C. R., Millman, K. J., van der Walt, S. J., et al., 2020: Array programming with NumPy. *Nature*, 585(7825), 357–362. doi.org/10.1038/s41586-020-2649-2.

Itow, S., 1962: Grassland vegetation in uplands of western Honshu, Japan: Part I. Distribution of grassland. *Japanese Journal of Ecology*, 12(4), 123–129.

Jordahl, K., Van den Bossche, J., Fleischmann, M., et al., 2020. geopandas/geopandas: v0.8.1. doi.org/10.5281/zenodo.3946761.

Matsushita, K., 2015. Japanese forestation policies during the 20 years following World War II. In: Zlatic, M. (Ed.), *Precious Forests - Precious Earth*. IntechOpen, London, Chapter 4. doi.org/10.5772/61268.

Ministry of Agriculture and Forestry, 1955. *World Agricultural Census of Japan, 1950*. Agricultural Statistics Association, Tokyo.

MLIT, 2009. National Land Numerical Information: hydrological mesh dataset (W07). Ministry of Land, Infrastructure, Transport and Tourism, Tokyo. nlftp.mlit.go.jp/ksj/ (10 December 2025).

Ohta, Y., Tsutsumi, M., Watanabe, S., et al., 2022: Changes in the distribution of grassland in the twentieth century in the Chūgoku region of western Japan. *Landscape and Ecological Engineering*, 18(1), 125–130. doi.org/10.1007/s11355-021-00477-4.

Onohara, A., 2026. aonoa68/tatara-grassland-tottori: v1.0. Zenodo. doi.org/10.5281/zenodo.19042339.

Pedregosa, F., Varoquaux, G., Gramfort, A., et al., 2011: Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

Perry, M., 2019. rasterstats: zonal statistics and raster sampling. github.com/perrygeo/python-rasterstats (10 December 2025).

QGIS Development Team, 2024. QGIS Geographic Information System, Version 3.40. Open Source Geospatial Foundation. qgis.org.

Seabold, S., Perktold, J., 2010. statsmodels: econometric and statistical modeling with Python. *Proceedings of the 9th Python in Science Conference*, 92–96.

Tottori Prefectural Government, 2020. Tottori Prefecture Geographic Information System: Tottori Web Map. Regional Development Promotion Division, Tottori. www2.wagmap.jp/pref-tottori/Portal (10 December 2025).

Virtanen, P., Gommers, R., Oliphant, T. E., et al., 2020: SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261–272. doi.org/10.1038/s41592-019-0686-2.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., et al., 2016: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. doi.org/10.1038/sdata.2016.18.