Waterways, Topography and Iron Age Settlement Location in North Zealand, Denmark

A GIS-Based Multivariate Research Framework

Per Palmkvist Knudsen

Research and methodological assistance: OpenAI GPT-5.6 Sol

WORKING PAPER - PRE-ANALYSIS VERSION FOR EXPERT REVIEW

19 August 2026

Status. The hypotheses and analytical framework are specified before the principal GIS analysis is executed. No new archaeological site, predictive result, or statistically demonstrated relationship is claimed in this document.

Purpose. This version is circulated for methodological criticism by archaeologists, landscape archaeologists, GIS specialists, geographers and palaeoenvironmental researchers before model fitting and prediction.

Abstract

This working paper proposes a GIS-based investigation of Iron Age settlement location in North Zealand, Denmark. Rather than treating elevation, agricultural suitability or proximity to water as independent explanations, settlements are conceptualised as positions within a multidimensional landscape combining topography, hydrology, agricultural resources and accessibility. The central hypothesis is that Iron Age settlements may occupy recurring landscape configurations reflecting both agricultural requirements and access to contemporary terrestrial and aquatic movement corridors. The proposed analysis combines archaeological site data, digital terrain models, reconstructed palaeohydrology, multivariate analysis, clustering and archaeological predictive modelling. Established GIS predictive modelling demonstrates the feasibility of probability mapping (Espa et al., 2006), while later work highlights true-absence problems, sampling bias and spatial autocorrelation (Yaworsky et al., 2020; Comer et al., 2023; Boemke et al., 2023). The project is at the pre-analysis stage.

1. Research question

Can the geographical distribution of known Iron Age settlements in North Zealand be explained by recurring combinations of elevation, relative topographic position, hydrology, palaeo-shoreline proximity, agricultural suitability and landscape accessibility? A second question follows: if such combinations exist, can they identify areas environmentally similar to known settlement landscapes but where Iron Age settlements have not yet been documented?

2. Conceptual starting point

The modern landscape is strongly conditioned by roads, drainage, urbanisation and present coastlines. These features cannot simply be projected backwards onto the Iron Age. Settlement location should therefore be examined within reconstructed contemporary landscapes rather than only against modern geography. Danish evidence also shows that settlement and wetland activity can occur in close association; the Aldersro complex in eastern Jutland is a documented example of a Pre-Roman Iron Age settlement juxtaposed with a wetland activity area (Roborg & Lovschal, 2022).

3. Central and competing hypotheses

H0 - Null model. Known settlement locations do not differ systematically from environmentally available locations after accounting for archaeological observation and preservation.

H1 - Hydrological-accessibility model. Settlement locations are disproportionately associated with palaeo-waterways, wetlands, lakes or coastal connections.

H2 - Agricultural model. Settlement distribution is predominantly explained by cultivable soils, drainage and agricultural catchments.

H3 - Topographic model. Relative elevation, slope and local prominence explain settlement location more effectively than hydrological accessibility.

H4 - Integrated landscape model. Settlement location reflects recurring combinations of agricultural, topographical and accessibility variables rather than one dominant environmental factor.

4. Water as connection, not only constraint

The working hypothesis gives particular attention to water. For prehistoric communities, lakes, coastlines, streams, wetlands and navigable water systems may in some circumstances have formed elements of mobility networks rather than merely obstacles. The analytically relevant quantity may therefore be travel cost and connectivity rather than Euclidean distance to water. This proposition is not assumed to be true. Its independent explanatory value must be tested after topography, agricultural suitability, chronology and archaeological discovery processes are taken into account.

5. Chronology and palaeolandscape

'Iron Age' is too broad a category for the final model. Sites should be assigned to narrower archaeological phases wherever dating quality permits, and analyses should be repeated by phase. Modern hydrology is likewise an imperfect proxy. Drainage, wetland reclamation, channel modification, forestry, agriculture, urbanisation and coastal change must be considered. Where reconstruction is uncertain, alternative defensible palaeolandscape scenarios should be modelled and sensitivity reported.

6. Proposed variables

Topography: absolute elevation, local relative elevation, slope, aspect, terrain ruggedness and local prominence.
Hydrology: distance to reconstructed freshwater, wetlands and palaeo-shorelines; catchment position; potential hydrological connectivity.
Agricultural landscape: soil characteristics, drainage, cultivable terrain and estimated agricultural catchment.
Accessibility: terrestrial least-cost paths, potential aquatic routes, combined land/water cost surfaces and connectivity with neighbouring settlements.
Archaeological context: chronological phase, site type, dating confidence, excavation/survey method, preservation and observation intensity.

7. Multivariate and cluster analysis

Each archaeological site will be represented as a multidimensional environmental vector. Continuous variables will be standardised. Principal component analysis may first be used to examine covariance and dimensionality. Multiple clustering approaches should then be compared rather than relying on a single algorithm. The purpose is to test whether settlements form reproducible landscape types. If stable configurations emerge, the analysis can be inverted to identify other parts of North Zealand with comparable combinations. Such output must be described as a landscape-suitability or archaeological-probability surface, not as a map of undiscovered settlements. GIS-based probability mapping has a substantial methodological precedent (Espa et al., 2006).

8. Presence, absence and model choice

An unregistered settlement is not equivalent to a demonstrated archaeological absence. Yaworsky et al. (2020) compared generalized linear models, generalized additive models, maximum entropy and random forests, and emphasised that the frequent lack of true-absence data can violate assumptions of several conventional approaches. Their results make presence-background or pseudo-absence approaches, including MaxEnt, particularly relevant candidates for comparison in the present project.

9. Principal methodological danger: discovery bias

The distribution of registered sites is not the original distribution of Iron Age settlement. Archaeological remains are disproportionately discovered where construction, infrastructure, quarrying or survey exposes them. Boemke et al. (2023), using 4,194 Upper and Final Palaeolithic occupations across Europe, found modern land use to be the strongest predictor of site presence in their analysis. Although the archaeological period differs, the methodological warning is directly relevant: a model can inadvertently learn where archaeology has been observed rather than where people preferred to live. Survey intensity, modern development and other discovery-related variables should therefore be modelled explicitly wherever suitable data can be obtained.

10. Spatial autocorrelation and validation

Nearby observations are not statistically independent by default. Comer et al. (2023) demonstrate that spatial autocorrelation and sampling strategy can substantially alter measured performance in archaeological predictive models. Randomly separating neighbouring cells between training and testing sets can produce misleadingly optimistic performance estimates. Validation should therefore include spatially separated training and test areas, leave-one-site-out tests, chronological or regional validation where feasible, and ultimately prospective validation in which predictions are frozen before later independent archaeological investigation.

11. Interpretation and falsification

The project does not begin with the conclusion that waterways determined settlement. It asks whether reconstructed hydrological accessibility adds explanatory information after agricultural suitability, topography, chronology and discovery bias are controlled. A negative result is scientifically informative. If water accessibility contributes little independent explanatory power, the waterways hypothesis should be rejected or substantially modified. The model must be capable of failing.

12. Current status and intended outputs

This is a pre-analysis framework. The complete archaeological GIS dataset has not yet been assembled and the principal statistical GIS analysis has not yet been executed. No new archaeological discovery or statistically demonstrated settlement-waterway relationship is claimed. Intended outputs include a reproducible geospatial database, period-specific environmental layers, comparison with appropriate background/control locations, multidimensional landscape classification, predictive or suitability surfaces with quantified uncertainty, explicit treatment of observation bias and spatial autocorrelation, and archived scripts and parameter choices sufficient for replication.

13. Authorship and use of artificial intelligence

Per Palmkvist Knudsen initiated the investigation, formulated and developed the archaeological and landscape questions, directed the iterative research process and is responsible for decisions concerning the investigation.

OpenAI GPT-5.6 Sol has been used as an artificial-intelligence research assistant. Its role has included helping formulate competing and falsifiable hypotheses, identifying methodological literature, structuring the proposed GIS and statistical analysis, identifying possible confounders, proposing validation strategies, and assisting in drafting and revision.

The AI system is not an archaeologist, does not possess professional archaeological qualifications, and is not an independent authority for archaeological authentication or interpretation. AI-generated suggestions and bibliographic information require verification against primary data and scholarly sources. Scientific responsibility, source verification, interpretation and decisions about publication remain with the human investigator.

14. Questions for expert review

We particularly invite criticism concerning: chronological classification of North Zealand Iron Age settlements; defensible reconstruction of palaeohydrology and palaeo-coastlines; archaeological discovery and survey bias; construction of background/control samples; modelling of waterborne accessibility; spatial autocorrelation; clustering and predictive-model selection; and validation design. The immediate objective is not confirmation, but attempted falsification before prediction.

References - verified for this circulation version

Boemke, B., Maier, A., Schmidt, I., Romer, W. & Lehmkuhl, F. (2023). Testing the representativity of Palaeolithic site distribution: The role of sampling bias in the European Upper and Final Palaeolithic record. *Quaternary Science Reviews*, 316, 108220. DOI: 10.1016/j.quascirev.2023.108220.

Comer, J. A., Comer, D. C., Dumitru, I. A., Priebe, C. E. & Patsolic, J. L. (2023). Sampling methods for archaeological predictive modeling: Spatial autocorrelation and model performance. *Journal of Archaeological Science: Reports*, 48, 103824. DOI: 10.1016/j.jasrep.2022.103824.

Espa, G., Benedetti, R., De Meo, A., Ricci, U. & Espa, S. (2006). GIS based models and estimation methods for the probability of archaeological site location. *Journal of Cultural Heritage*, 7(3), 147-155. DOI: 10.1016/j.culher.2006.06.001.

Roborg, A. S. & Lovschal, M. (2022). The Aldersro wetland-settlement complex: Deposition and mortuary practices in Pre-Roman Iron Age, Eastern Jutland, Denmark. *Danish Journal of Archaeology*, 10, 1-8. DOI: 10.7146/dja.v10i0.125120.

Yaworsky, P. M., Vernon, K. B., Spangler, J. D., Brewer, S. C. & Codding, B. F. (2020). Advancing predictive modeling in archaeology: An evaluation of regression and machine learning methods on the Grand Staircase-Escalante National Monument. *PLOS ONE*, 15(10), e0239424. DOI: 10.1371/journal.pone.0239424.

Suggested citation

Palmkvist Knudsen, P. (2026). *Waterways, Topography and Iron Age Settlement Location in North Zealand, Denmark: A GIS-Based Multivariate Research Framework.* Working paper, pre-analysis version, 19 August 2026. Research and methodological assistance by OpenAI GPT-5.6 Sol.

Reference note: This circulation version deliberately contains only references whose bibliographic details were checked against the journal/publisher record before PDF production.