



HAL
open science

What does matter? Machine Learning models to define archaeological patterns of site occupations and resource use in Kurdistan (Iraq) through a diachronic approach

Mathias Bellat, Benjamin Glissmann, Tobias Rentschler, Sconzo Paola, Peter Pfläzner, Thomas Scholten

► To cite this version:

Mathias Bellat, Benjamin Glissmann, Tobias Rentschler, Sconzo Paola, Peter Pfläzner, et al.. What does matter? Machine Learning models to define archaeological patterns of site occupations and resource use in Kurdistan (Iraq) through a diachronic approach. AFEQ - CNF INQUA international conference 14 - 18 Mars 2022 : “ Q13 - Palaeoclimate changes, landscape evolution and human societies: from sedimentary basins to industrial landscapes ”, Mar 2022, Strasbourg, France. , 2022. <hal-03551422>

HAL Id: hal-03551422

<https://hal.science/hal-03551422v1>

Submitted on 31 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

Machine Learning models to define archaeological patterns of site occupations and resource use in Kurdistan (Iraq) through a diachronic approach : What does matter?

MATHIAS BELLAT^{1,2}, BENJAMIN GLISSMANN^{1,3}, TOBIAS RENTSCHLER^{1,2}, PAOLA SCONZO³, PETER PFÄLZNER^{1,3}, THOMAS SCHOLTEN^{1,2,4}

¹ SFB 1070 ResourceCultures, Eberhard Karls University Tübingen, Germany
² Chair of Soil Science and Geomorphology, Department of Geosciences, Eberhard Karls University Tübingen, Germany
³ Institute for Ancient Near Eastern Studies (IANES), Eberhard Karls Universität Tübingen, Germany
⁴ Cluster of Excellence Machine Learning „New Perspectives for Science“, Eberhard Karls Universität Tübingen, Germany

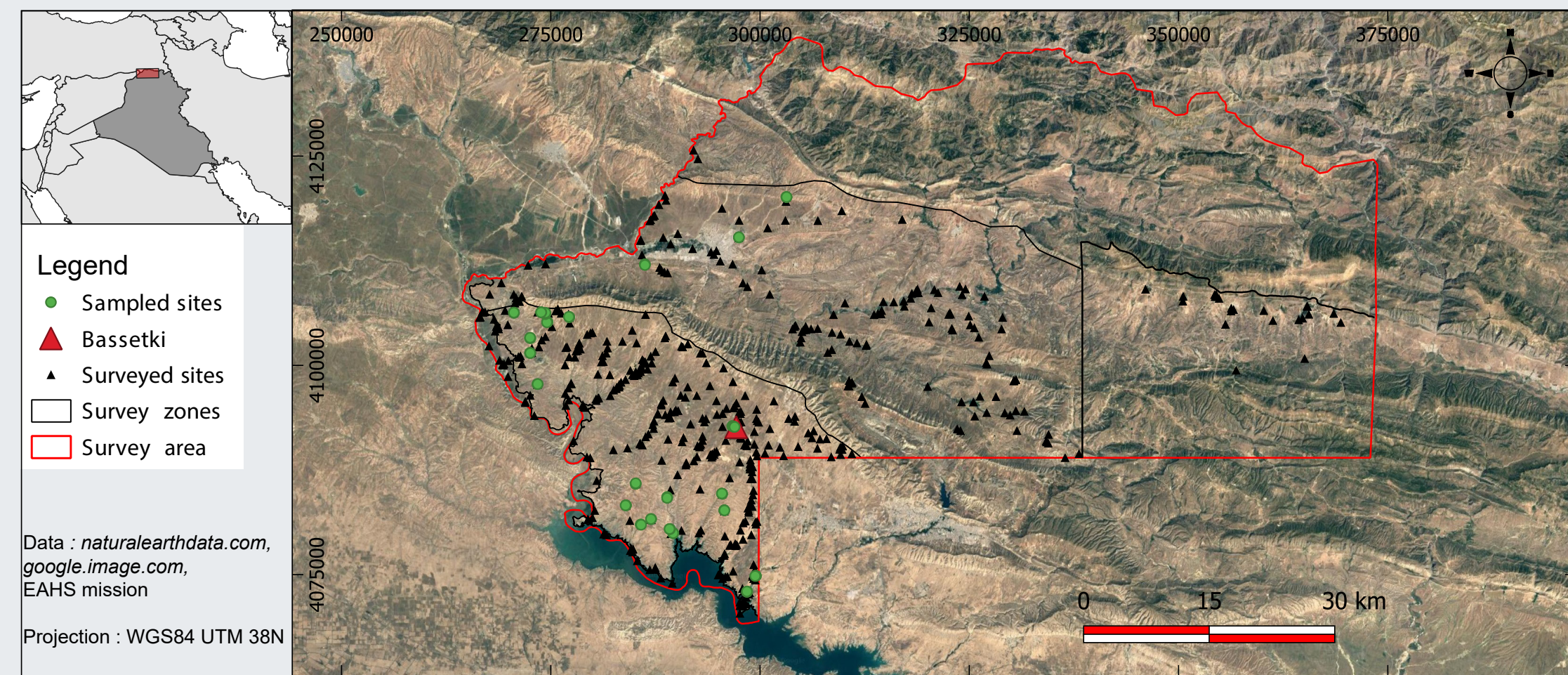


Fig.1 : Study area with sampled and survey sites

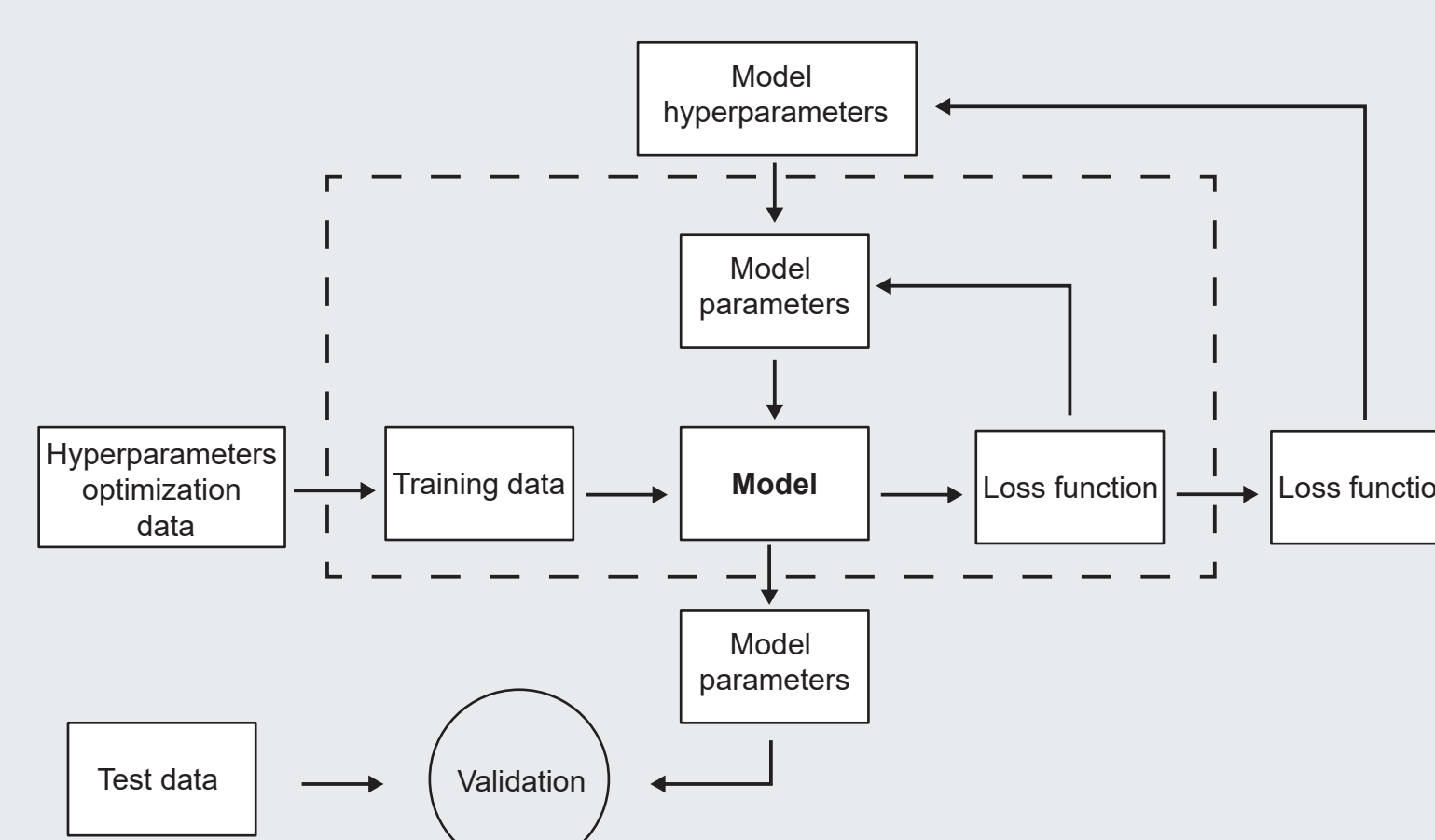


Fig.3 : Machine learning schematic function

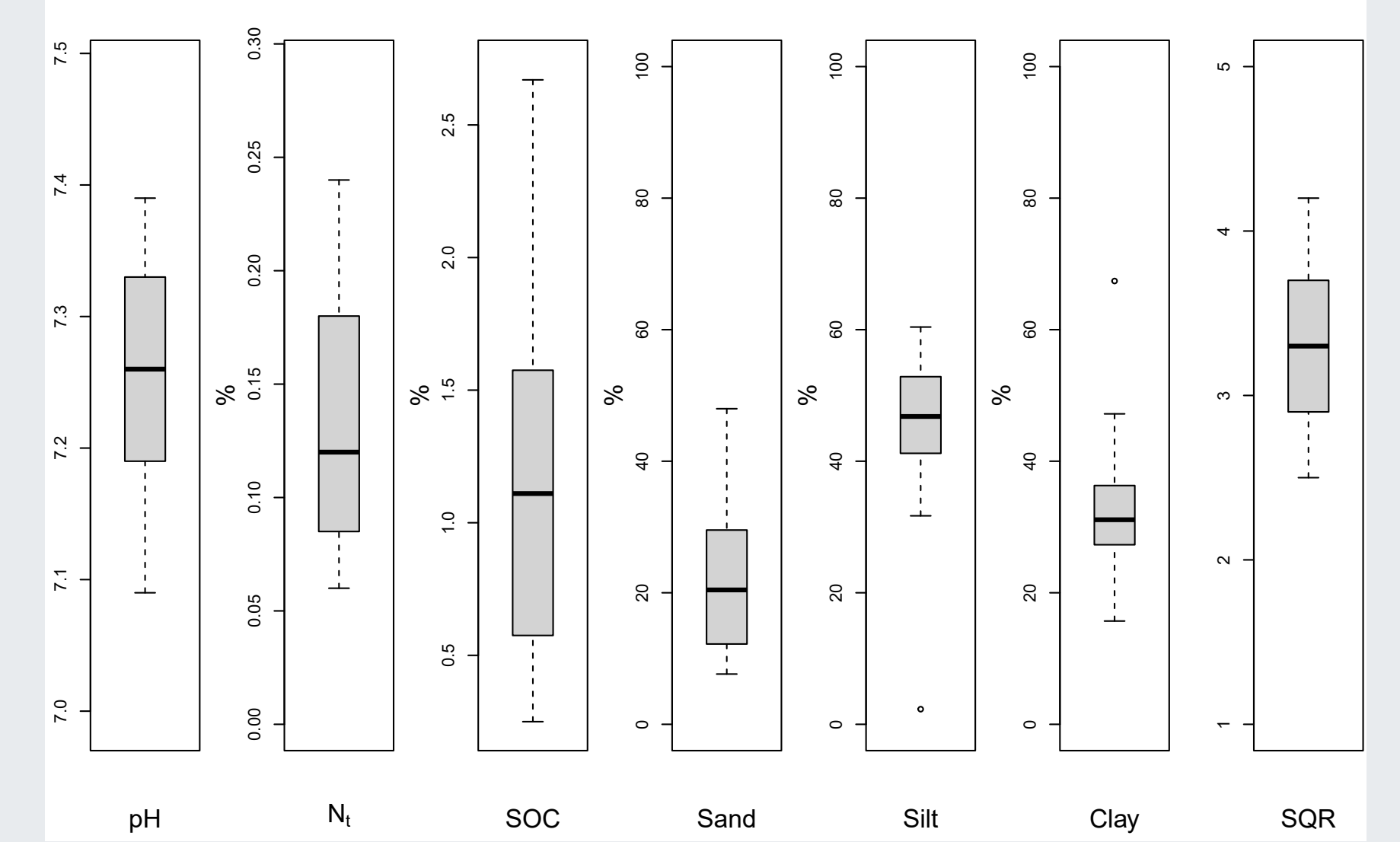


Fig.5 : Soil laboratory analysis

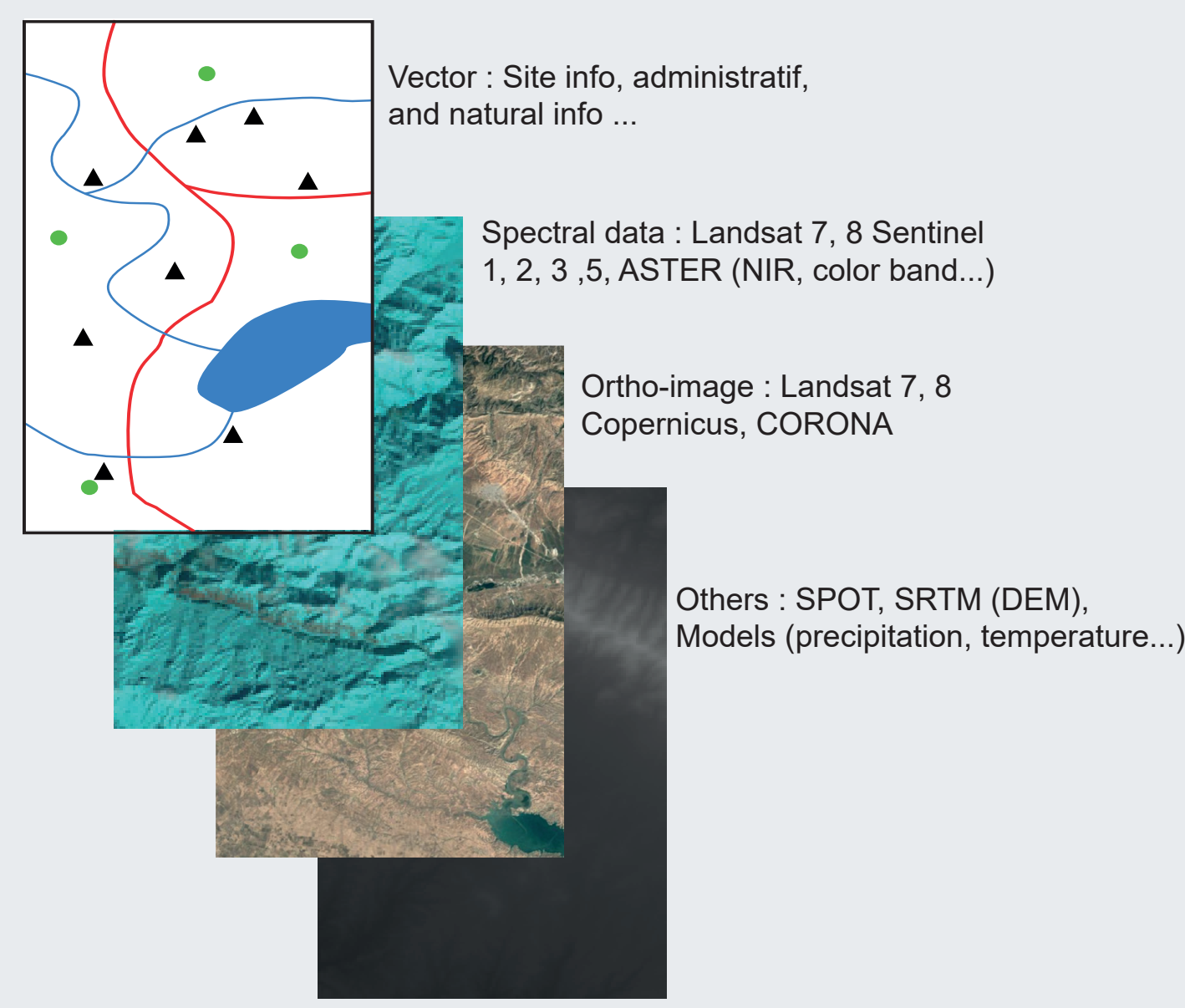


Fig.2 : Multiscale and multisource data

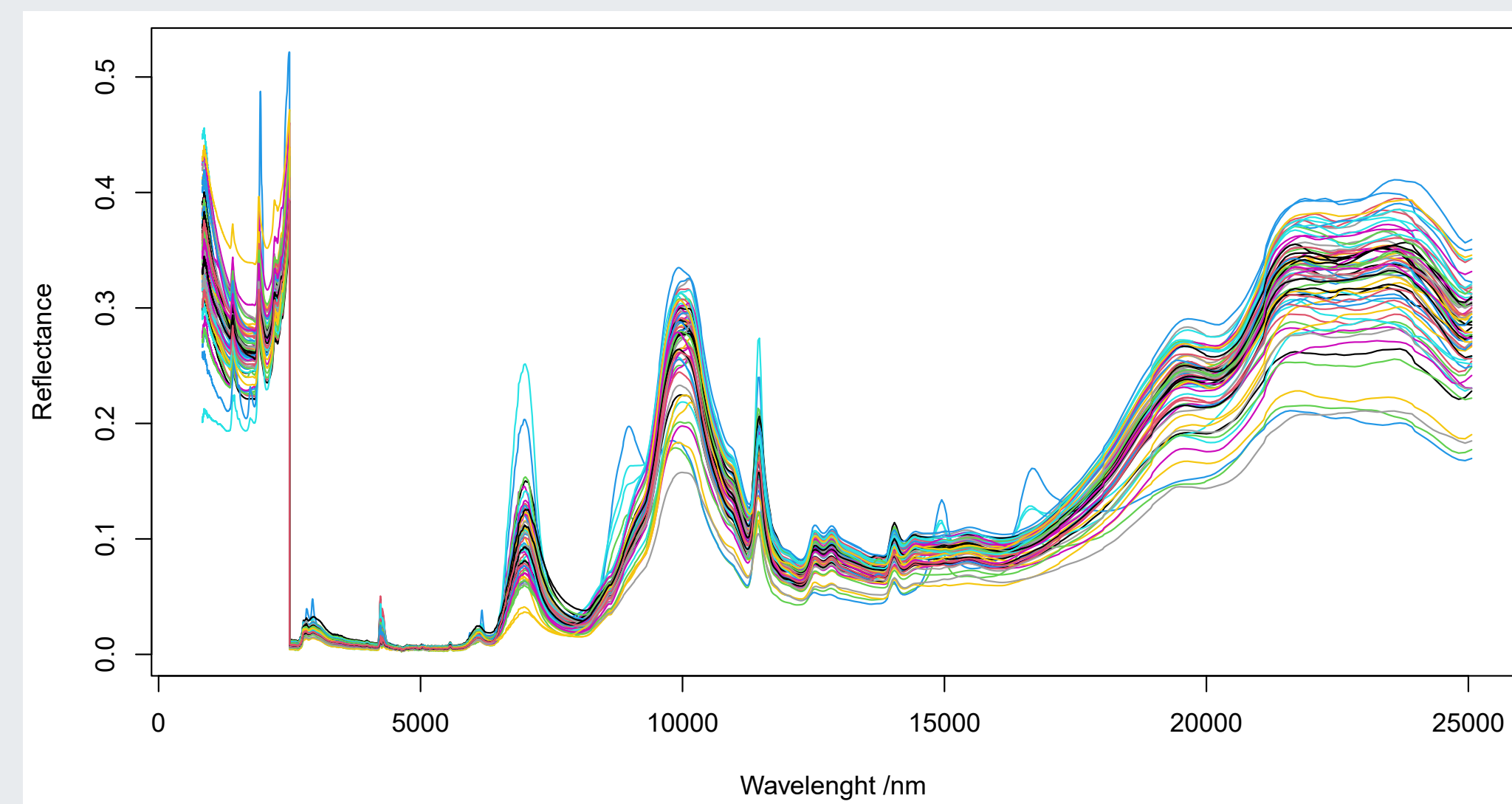


Fig.4 : NIR and MIR Spectrum of 66 samples

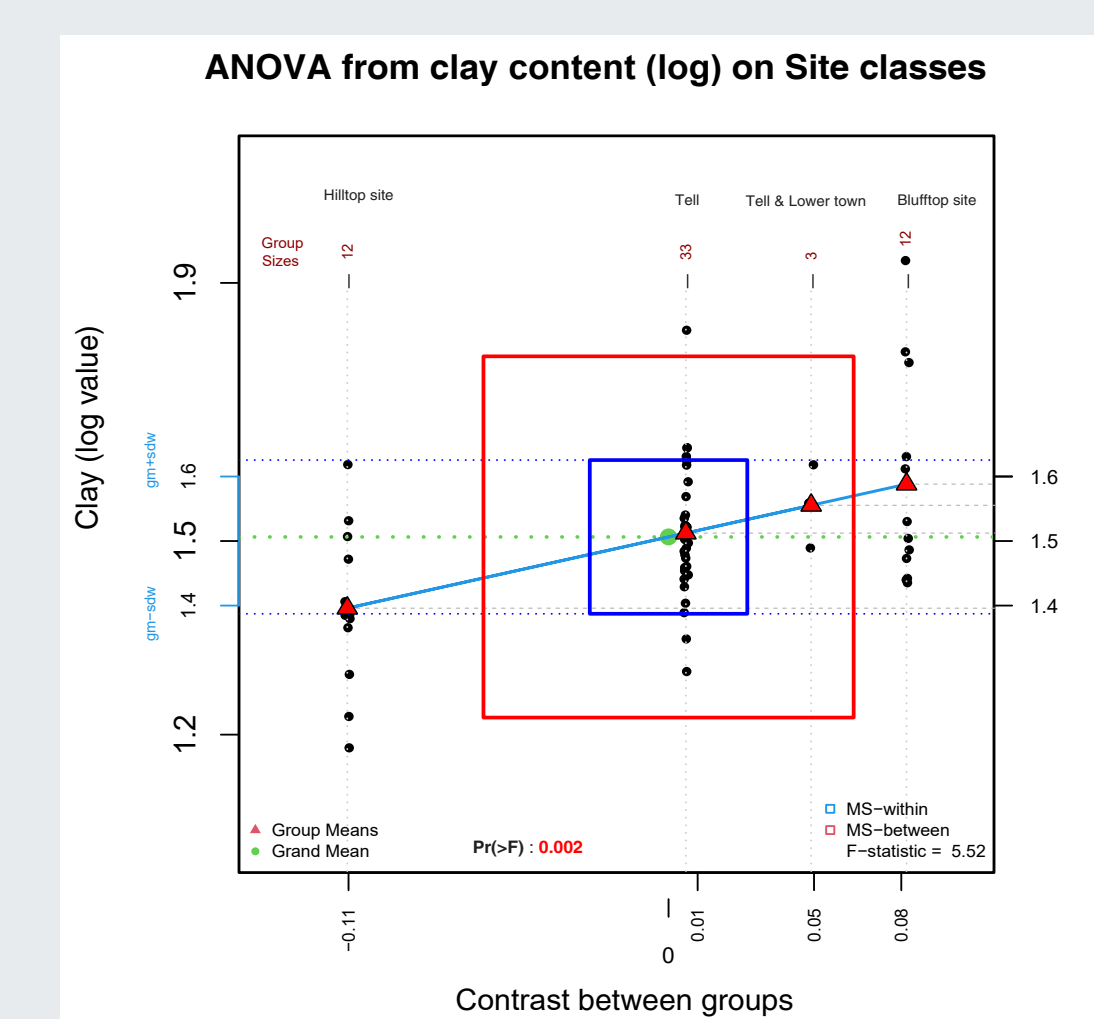


Fig.6 : ANOVA of clay content link to site type

1. INTRODUCTION

The analysis of sites occupation patterns, resource use and land management practices by ancient communities opens windows on their social and cultural organization. Based on the concept of „ResourceComplexes“ and „ResourceCultures“ (SCHADE et al. 2021), we aim to investigate the links between communities and their living environment and mobility.

Studies on sites occupations in the Mesopotamian area has been conducted since the '60s (ADAMS 1965). More recently remote sensing data have been used to distinguish new sites and channels (JOTHERI et al. 2020). Nonetheless, these previous studies were limited to Earth's surface signals and not including soil information and other objects like archaeological finds and geoarchives in terms of a multi-proxy analysis of sites occupations and resources use in Kurdistan (Iraq).

Here, we try to combine both and use a machine learning approach for remote sensing data, in situ soils sampling and data from the previous survey and complete it with statistics and other complementary analysis.

2. MATERIEL AND METHODS

STUDY AREA

- Iraqi Kurdistan in the area of Dohuk (≈ 4 000 km²).
- Over 500 sites have been identified (Fig. 1) and 22 sites sampled.
- Additional data have been collected through remote sensing or other free available datasets (Fig. 2).

METHODS

- Soils have been analyzed with classical wet chemistry methods, and additional predicted data from near and mid infrared spectra NIR, MIR (Fig. 4).
- Machine learning models (Fig. 3) are used to predicted soil properties from environmental data (precipitation, temperature...), and clustering or classifying archaeological sites data.
- Statistics is used to find any correlation between the various data and the type of occupation on the site.

3. RESULTS

- The soils analysis (Fig. 5) points out soils lightly acid (high pH: 7.2 – 7.35) with low nutrient contents and dominated by a sandy-loam, silty-loam and clayed-loam texture. This indicates a low capacity for plant growth despite a high-water holding capacity (HAZELTON & MURPHY 2016).
- Several maps (precipitation, temperature...) have been produced by a predictive model on machine learning (Random Forests).
- Statistics points out a correlation between the clay content and the type of site (Fig. 6). There seem to be a significant difference (95% confidence level) between the clay content in the Hilltop site and both Tell and Blufftop sites (Tab. 1). This correlation points out future possibility in machine learning classification or clustering models.
- Different sampling strategies have been done with statistics forecasting for preparing future missions.

| Site type relation | Difference | Lower | Uper | p. adjust |
|--------------------------------|------------|---------|---------|---------------|
| Hilltop - Blufftop | -0.1922 | -0.3210 | -0.0635 | 0.0012 |
| Tell - Blufftop | -0.0759 | -0.1822 | 0.0305 | 0.2442 |
| Tell and lower town - Blufftop | -0.0328 | -0.2363 | 0.1708 | 0.9737 |
| Tell - Hilltop | 0.1164 | 0.0101 | 0.2227 | 0.0266 |
| Tell and lower town - Hilltop | 0.1594 | -0.0441 | 0.3630 | 0.1742 |
| Tell - Tell and lower town | 0.0431 | -0.1471 | 0.2332 | 0.9317 |

Tab.1 : Tukey HSD Table on the relation between clay mesurment and site type

4. DISCUSSION AND CONCLUSION

- Machine learning (ML) combined with statistics show a clear relation between settlement pattern and soils properties for clayed soil.
- We assume that ML has a great potential to analyze and better understand sites occupations and resources use. However, multi-proxy and multi-scalar studies need large data and a wide spectrum of statistical methods, but not necessarily at high costs.
- In a next step, the application of this approach to settlement patterns and sites determination will be tested.

ACKNOWLEDGMENTS

We thank the German Research Foundation (DFG) for supporting this research through the Collaborative Research Centre SFB 1070 RESOURCECULTURES (subprojects Z, S and B07).

CONTACT

*Mathias Bellat
mathias.bellat@uni-tuebingen.de



FURTHER READINGS

R. M.-C. Adams (1965), *Land Behind Baghdad: A History of Settlement on the Diyala Plains*, University of Chicago Press, Chicago. J. Jotheri, M. Allen (2020), "Recognition of Ancient Channels and Archaeological Sites in the Mesopotamian Floodplain Using Satellite Imagery and Digital Topography" In : D. Lawrence, M. Altaweel, G. Philip (Ed.), *New Agendas in Remote Sensing and Landscape Archaeology in the Near East Studies in Honour of Tony J. Wilkinson*, Archaeopress, Oxford, pp. 283-305. T. Shade, B. Schweizer, S. Teuber et al. (dir.) (2021), *Exploring resources: on cultural, spatial and temporal dimensions of ResourceCultures*, Tübingen University Press, Tübingen. P. Hazelton, B.W. Murphy, (2016), *Interpreting Soil Test Results What Do All the Numbers Mean?*, 3th Edition, CSIRO Publishing, Clayton South.