ABSTRACT
We present the methodological and technical issues of a research on ancient vegetation modelling in the Russian steppes combining remote sensing and palynological data. This research is framed in an international project developed by Russian and Spanish archaeologists from the Russian Academy of Sciences and the Spanish National Research Council (CSIC) and centred on the study of the Bronze Age mining complex of Kargaly (Orenburg, Russia).

INTRODUCTION
Kargaly is a vast copper mineral field located in the Orenburg region, in the forest steppe of the middle Ural river basin (Figure 1).

First mining exploitations go back to Early Bronze Age (EBA), in the middle of the 4th millennium BC, although mining and metallurgical works reach their maximum intensity in Late Bronze Age (LBA). Since the end of this period, mining activity completely stops and does not resume till Modern Ages, with the annexation of this region by the Russian Empire in 1740 AD, and fully decays in the beginning of the 20th century.

Since 1993, the prehistoric mining-metallurgical complex of Kargaly has been the subject of research by a Russian-Spanish team managed by Dr. E. N. Chernykh (Institute of Archaeology, Russian Academy of Sciences, Moscow) and formed by researchers from the Russian Academy of Sciences, the Spanish National Research Council (CSIC), the National Archaeological Museum in Madrid, and the Polytechnic University, Mª Isabel Martínez Navarrete being the Principal Investigator of the Spanish projects. During the course of research, several mining districts have been surveyed and mapped and three Bronze Age settlements and several Kurgan cemeteries have been excavated (1,2,3,4,5). At the same time, paleoenvironmental and landscape archaeology research has been carried out, which frames the work shown in this paper.

The study of Kargaly landscape and its transformations since Bronze Age until present time is fundamental for the historical understanding of the Kargaly complex, its development through Bronze
Age and its collapse at the end of this period. According to their relevance, two aspects are emphasized:

1. Mining and metallurgy are activities that have a great impact on environment, mainly in relation to forest resources. Changes in the distribution of these resources between Bronze Age and present time might constitute a good proxy to evaluate the scale of mining works and the existence or not of a local metallurgy.

2. Archaeological data show that agriculture was not practiced in this region till a very later period (6). This raises some questions about the subsistence economy of the Bronze Age mining communities that settled the area.

Our research has focused on the first topic up to now (7) being subsistence practices the subject of current works.

![Figure 1: Kargaly location in the Eurasian steppe close to the Ural river.](image)

**METHODS**

**Paleopalynology**

Paleopalynology is a basic discipline to produce knowledge about vegetation characteristics of past landscapes (8,9,10). Through a normalized methodology including fieldwork and laboratory tasks, it recovers and studies ancient pollens, spores and non-pollen palynomorphs in order to elaborate a picture about the presence and quantity of vegetation taxa existing in the environment. During fieldwork, soil samples containing pollens are taken from sediments, both in archaeological sites and natural contexts such as peatbogs and marshes. In the laboratory, pollens are extracted from their soil matrix and taxonomically identified and counted through microscopic inspection. Observations are tabulated and summarized in a histogram (palynogram) representing identified botanic taxa and their frequency into the sample.
Conventional paleopalynological studies use to infer ancient vegetation cover features directly from palynograms, qualifying the meaning of observed frequencies by means of known characteristics of pollen production, dispersion and deposition of each botanical taxa. In recent years, new approaches have been developed to interpret palynograms in a more systematic way through numerical landscape models (11,12). These models combine paleopalynological information with modern pollen rain data and various geographic features, both regarding vegetation (e.g. vegetation cover distribution or land uses) and other relevant environmental variables (e.g. topography, prevailing winds and so on). In this way, paleopalynological data are interpreted in light of identified patterns about the relationship between pollen deposition and the characteristics of the surrounding landscape. For producing vegetation cover information, some of these models use remote sensing imagery analysed through digital image processing methods (13,14,15), as it is our case.

Available data about vegetation distribution in Bronze Age are very limited and restricted to evidence provided by fossil pollen taken from archaeological sediments and natural deposits excavated during fieldwork. In order to interpret this information according to the aims of our project, we have proposed a methodology based on palynological spectra, taken by subsoil drill sampling, versus modern pollen rain patterns comparison, taken by surface soil samples (16). Understanding the relationship between these patterns and modern vegetation, will help us interpret observable differences between fossil and recent spectra as signs of changes in the vegetation cover. This requires a precise cartography about vegetation formations in the area, which is the aim of the remote sensing program shown in this paper.

Proposed methodology is based on the Best Modern Analogue approach (17,18), developed by archaeological palynology in recent years. This approach consists of isolating present palynological spectra that are qualitative and quantitatively similar to fossil pollen spectra and analysing the environment of the former assuming a similar environment for the latter. Incorporating remote sensing to the BMA methodology provides a powerful tool for representing and analysing the environment. In this research we try to promote a synergy between palynology and remote sensing by means of statistical models, as it has been done by other authors, although usually at a larger scale (cf. 13-18).

Next, we briefly describe the general research design and the works carried out for the elaboration of a land cover cartography through satellite image processing. Finally, we discuss some of the results of the analysis of the palynological data in relation to the problem of changes in composition and distribution of forest resources between Bronze Age and present time.

Research design

In order to feed the analytical landscape model that will be used for testing the paleoenvironmental hypotheses, two basic sources of information have been used: palynological spectra and land cover cartography.

Palynological spectra are elaborated through the study of palynological samples taken at different locations in the research area. They provide information about botanical taxa existing in the surroundings of each sample point. Each taxon is represented by a variable containing its relative frequency (expressed through percentage) with regard to the whole sum of pollen grains in the sample.

Two types of palynological samples have been analysed: fossil samples and present pollen rain samples. Fossil paleopalynological samples are taken from ancient natural deposits and provide information about past vegetation. They are chrono-referenced through relative or numeric dating methods (e.g. radiocarbon). Modern pollen rain samples are taken on the ground surface level and provide information about modern episodes of pollen deposition. In order to offer a complete view of pollen rain in the whole study area, a representative set of samples must be taken. This has been achieved by means of a sampling strategy, combining both a selective and a systematic approach.

Paleopalynological samples have been taken at two natural deposits (22 samples at Novienki peat and 17 at Gorny mining trench) (19,20). These samples have been dated through the radiocarbon
method, applied both to pollen grain concentrations and paleosols. Chronology ranges since 4th millennium BC to present time.

Pollen rain data have been obtained by means of 76 surface samples, 55 of them resulting from a systematic sampling strategy and the other 21 selectively located according to botanical criteria (7,21) (Figure 2). Systematic sampling was based on two adjoining transects, each one defined around one of the paleopalynological samples. Each transect was divided into square sampling units 500 m on a side. A 10% of these units was randomly selected in each transect. In each selected unit a sample was taken in the upper 10 cm of the topsoil.

Figure 2: Pollen rain and palynological samples distribution

Regarding land cover cartography, its function consists of being compared statistically with pollen rain data for exploring how and to what degree these represent the quantitative presence and spatial distribution of different botanical taxa through landscape. Due to the lack of enough detailed pre-existing cartography, it has been generated using remote sensing imagery combined with ancillary information, such as a DEM, by means of digital image classification methods. Image classification has been supervised using existing knowledge about vegetation in Kargaly region, both botanical literature (22,23) and fieldwork observations.

The process of information analysis has the following stages (Figure 3):

1. Quantification of the presence of pollen taxa in each sample, through the elaboration of relative frequency variables.
2. Elaboration of land cover cartography using image classification methods. Primary data are remote sensing images and ancillary data, such as a DEM.
3. Comparative analysis of pollen rain and land cover cartography by means of statistical methods. This allows to elaborate a model of how pollen rain represents vegetation presence.
4. Finally, comparison of fossil pollen data with this model, for estimating which is the most likely kind of vegetation landscape that these data represent.
Remote sensing

Current landscape was to be modeled through remote sensing techniques\(^1\), in order to obtain an analytical document displaying the vegetation distribution responsible for the pollen rain samples gathered during fieldwork campaigns.

To achieve a correct land cover classification of the study area we had to relay on very heterogeneous data. A Landsat TM 5 image accounted for terrain conditions at the moment of field data gathering, but 30 m spatial resolution revealed too poor for the characterization of certain pollen species dispersal, and inadequate to delineate the narrow forest spots typical in Kargaly. Furthermore, there is no cartography of the area at an appropriate scale, so an ASTER image was acquired to enable the building of a DEM, as well as to improve spatial resolution, via data merging.

Although the Landsat scene was collected on September 1994 and the ASTER one on August 2002, data merging was considered feasible thanks to the continuities that link both landscapes. In spite of ownership and crop changes, land subdivision remained practically unchanged since soviet agrarian colonization.

The ASTER image was acquired at Level 1B, which means that it had been already geometrically corrected via orbital parameters; the TM had been geometrically corrected using accurate GCPs, consistent with pollen data samples. The ASTER image had to be co-registered to the TM image, a process that yielded a RMSE of 11 m.

The Landsat image was merged with a panchromatic ASTER generated as an average of the three VNIR bands. We chose a method based on wavelet analysis, for it produces a higher spectral quality than other methods as PCA or HIS (24), a crucial point in order to perform a digital classification. The merged image was considered suitable according to the spectral correlation of the new image bands and the original one, as well as the ERGAS coefficient (25) (Table 1).

\(^1\) ERDAS Imagine was used for digital image processing, and ESRI ArcGIS software for ancillary data such as DEM.
A supervised classification was performed on the new merged image in order to identify three main land cover categories, subdivided according to the spectral characteristics of the vegetation in autumn:
1. Forest
   1.1. Headwaters forest
   1.2. River forest
   1.3. Scrubland
2. Steppe
   2.1. Dry
   2.2. Wet
3. Labor field
   3.1. Plowed fields
   3.2. Stubble
   3.3. Fallow
   3.4. Not harvested

Table 1: Spectral correlation of the new image bands and ERGAS coefficient.

<table>
<thead>
<tr>
<th>Band</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.97435174</td>
</tr>
<tr>
<td>2</td>
<td>0.97058677</td>
</tr>
<tr>
<td>3</td>
<td>0.97316609</td>
</tr>
<tr>
<td>4</td>
<td>0.9751107</td>
</tr>
<tr>
<td>5</td>
<td>0.97536803</td>
</tr>
<tr>
<td>6</td>
<td>0.97275086</td>
</tr>
<tr>
<td>ERGAS</td>
<td>1.57232994</td>
</tr>
</tbody>
</table>

A parallelepiped classification was performed with Mahalanobis distance used as overlapping rule. Afterwards forests were reclassified in river forests and headwaters forests, regarding their inclusion in first order water stream basins, according to Shreve’s method\(^3\) (26). A majority filter was applied inside parcels, so every pixel was assigned the most frequent value in its labor field, a process that required land subdivision digitizing.

Finally, the resulting image was evaluated using 62 ground control points gathered during field work. The special conditions of the field campaigns, that took place along three different years, prevented us from using all cover categories, so the image had to be recoded to the three main classes in order to perform the error matrix (Tables 2-4).

Table 2: Error matrix.

<table>
<thead>
<tr>
<th>Classified Data</th>
<th>Forest</th>
<th>Steppe</th>
<th>Labor fields</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Steppe</td>
<td>1</td>
<td>35</td>
<td>0</td>
</tr>
<tr>
<td>Labor fields</td>
<td>0</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>Column Total</td>
<td>5</td>
<td>36</td>
<td>21</td>
</tr>
</tbody>
</table>

Only two ground control points yielded a different result than the classified image, a steppe location classified as forest and a forest location classified as steppe, a situation that explains the coincidence of producers and users accuracy. Although, due to the special terms of the verification process, we must be careful examining the results, the classification may be regarded as a plausible land cover document of the study area.

\(^2\) The merging process does not alter significantly an image when the ERGAS coefficient is below 3.

\(^3\) Hydrological features (basins, streams) were obtained through ASTER DEM processing.
Table 3: Accuracy.

<table>
<thead>
<tr>
<th>Class Name</th>
<th>Reference Totals</th>
<th>Classified Totals</th>
<th>Number Correct</th>
<th>Producers Accuracy</th>
<th>Users Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>80,00%</td>
<td>80,00%</td>
</tr>
<tr>
<td>Steppe</td>
<td>36</td>
<td>36</td>
<td>35</td>
<td>97,22%</td>
<td>97,22%</td>
</tr>
<tr>
<td>Labor fields</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>100,00%</td>
<td>100,00%</td>
</tr>
<tr>
<td>Totals</td>
<td>62</td>
<td>62</td>
<td>60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Overall Classification Accuracy = 96.77%

Table 4: Kappa statistics.

<table>
<thead>
<tr>
<th>Class name</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest</td>
<td>0,7825</td>
</tr>
<tr>
<td>Steppe</td>
<td>0,9338</td>
</tr>
<tr>
<td>Labor fields</td>
<td>1,0000</td>
</tr>
<tr>
<td>Overall Kappa</td>
<td>0,9404</td>
</tr>
</tbody>
</table>

Figure 5: Digital classification results.

RESULTS

Current Kargaly landscape is a typical forest steppe, representing the southern strip of the Great Eurasian Steppe. It is a region of gentle hills and gullies, with a predominance of herbaceous species and scarce forest patches across the water courses, both permanent or seasonal. In our current research stage, attention has been focused on these scarce forest resources, trying to determine whether their distribution during Bronze Age was significantly different to the present one. This is a crucial point to understand the historical development of the Kargaly mining complex, particularly its collapse by mid of the 2nd millennium BC.
Palynological data from two paleopalynological sequences taken in natural deposits provide a continuous coverage for the last five thousand years. One of these sequences (Gorny) is in the center of one of the mining areas, while the other (Novienky) is located at an area where there has never been mining activity. This contrast shall enable the measurement of mining activities effects on landscape dynamics.

The samples that form these sequences were aggregated in stratigraphic phases, having their absolute chronology established through a C14 direct dating of the pollen concentrations. This allowed for the organization of the deposit formation processes, according to the historical events established by means of archaeological works carried out in two sites close to the natural deposits.

This method has enabled the establishment of four phases:

1. EBA and MBA periods (4th and 3rd millennium BC): Beginning of the mining exploitation in the region.
2. LBA period, Srubnaya culture (mid 3rd to mid 2nd millennium BC): Maximum intensity of the prehistoric mining exploitation.
3. Nomadic settlement period (mid 2nd millennium BC – c. 1750 AD): It is a stage without mining activity nor permanent settlement.

The comparison between fossil pollen data from these phases and observed patterns of present pollen rain yields some significant results.

Characteristic palynological patterns have been established for every phase, studying pollen spectra changes along time in a diachronic approach, while in a synchronic approach, contemporary phases of both sequences have been compared to study landscape dynamics differences. These changes give way to an interpretation in terms of vegetation distribution shifts by fossil pollen spectra versus nowadays pollen rain comparison.

According to the BMA logic, we assume that likely analogies of palynological patterns of the four phases and nowadays pollen rain, might be interpreted as an evidence of analogies in the formation conditions of both palynological records, and therefore offer an approach to the distribution of past forest resources and their evolution through time.

The elaboration of a detailed forest cover cartography using the classification techniques described above and their qualification with the aid of ground truth, especially through the elaboration of systematic floristic inventories, enabled the linking of this palynological patterns with the vegetation formations according with their position in landscape topography.

The analogy degree between paleopalynological samples and nowadays pollen rain, has been established by means of a Principal Component Analysis, establishing an interpretative pattern based on the quantitative variation of the first four components in relation to the location of forest species. Thus we were able to define the spatial relationship between current forest distribution and its palynological correlate, setting at the same time the extent of the correlation between ancient and current samples.

This analysis yielded some relevant and, partly unexpected, results. An intense deforestation associated to phases 1 and 2 has been detected in the mining area located at the north of the study area. Although this outcome was expected for phase 2, that of the maximum intensity of prehistoric mining, it shows that the scale of exploitation in the early phases of the complex is larger than what could be inferred from the archaeological evidence. The end of prehistoric mining opens a course of forest recovery along phase 3, interrupted in the latter stage, marked by the resumption of copper mining activities in the 18th century.

The south part shows a different development, there was no mining activity but we notice an intense impact of pastoral practices during phase 3, involving a withdrawal of forest resources. Contradictory tendencies in deforestation processes on both areas help discard climate changes as a
possible explanation within global or regional processes, and display the weight of local anthropogenic factors in the formation of paleopalynological records.

Finally, we can observe an important variation in the composition of forest resources, consisting of practical extinction of oak (*Quercus robur*) as a consequence of anthropic impact during phase 4. This fact has relevant implications about the energetic potential of Kargaly forest resources and provides a key reference for the historical interpretation of this Bronze Age mining-metallurgical complex.

**CONCLUSIONS**

Apart from the specific results that we have discussed, research put forward in this paper shows that the proposed methodological design is feasible and relevant for Landscape Archaeology. Combining remote sensing technology and palynological research in an experimental statistical framework provides a powerful tool for evaluating landscape long-term changes, giving way to a new spatial approach to paleoenvironmental research.

**ACKNOWLEDGEMENTS**

The authors would like to thank E. N. Chernykh and the Institute of Archaeology (Russian Academy of Sciences, Moscow) team work, it is thanks to them that we could take part in the Kargaly project.

We would also like to thank all the Spaniards in charge of the palynological research: Pilar López García and Antonio López Sáez, from the Instituto de Historia (CSIC, Madrid); and to I. Zavala Morencos (Universidad Politécnica, Madrid) and A. Rodríguez Alcalde (CSIC, Madrid) in charge of the cartography and remote sensing project, along with Juan M. Vicent during the field work.

**REFERENCES**


