

Intrinsic shape analysis in archaeology: A case study on ancient sundials

MARTIN HANIK*, Freie Universität Berlin, Germany

BENJAMIN DUCKE, German Archaeological Institute, Germany

HANS-CHRISTIAN HEGE, Zuse Institute Berlin, Germany

FRIEDERIKE FLESS, German Archaeological Institute, Germany

CHRISTOPH VON TYCOWICZ, Freie Universität Berlin, Germany

The fact that the physical shapes of man-made objects are subject to overlapping influences—such as technological, economic, geographic, and stylistic progressions—holds great information potential. On the other hand, it is also a major analytical challenge to uncover these overlapping trends and to disentangle them in an *unbiased* way. This paper explores a novel mathematical approach to extract archaeological insights from ensembles of similar artifact shapes. We show that by considering all shape information in a *find collection*, it is possible to identify shape patterns that would be difficult to discern by considering the artifacts individually or by classifying shapes into predefined archaeological types and analyzing the associated distinguishing characteristics.

Recently, series of high-resolution digital representations of artifacts have become available. Such data sets enable the application of extremely sensitive and flexible methods of shape analysis. We explore this potential on a set of 3D models of ancient Greek and Roman sundials, with the aim of providing alternatives to the traditional archaeological method of “trend extraction by ordination” (typology).

In the proposed approach, each 3D shape is represented as a point in a *shape space*—a high-dimensional, curved, non-Euclidean space. Proper consideration of its mathematical properties reduces bias in data analysis and thus improves analytical power. By performing regression in shape space, we find that for Roman sundials, the bend of the shadow-receiving surface of the sundials changes with the latitude of the location. This suggests that, apart from the inscribed hour lines, also a sundial’s shape was adjusted to the place of installation. As an example of more advanced inference, we use the identified trend to infer the latitude at which a sundial, whose location of installation is unknown, was placed.

We also derive a novel method for differentiated morphological trend assertion, building upon and extending the theory of geometric statistics and shape analysis. Specifically, we present a regression-based method for statistical normalization of shapes that serves as a means of disentangling *parameter-dependent effects (trends)* and *unexplained variability*. In addition, we show that this approach is robust to noise in the digital reconstructions of the artifact shapes.

CCS Concepts: • **Applied computing** → **Archaeology**; • **Computing methodologies** → *Shape analysis*; Modeling methodologies; • **Mathematics of computing** → Regression analysis; Differential calculus.

Additional Key Words and Phrases: ancient sundials, shape analysis, geometric statistics, statistical normalization

*Corresponding author.

Authors’ addresses: Martin Hanik, hanik@zib.de, Freie Universität Berlin, Kaiserswerther Str. 16–18, Berlin, Germany, 14195; Benjamin Duce, benjamin.ducke@dainst.de, German Archaeological Institute, Podbielskiallee 69–71, Berlin, Germany, 14195; Hans-Christian Hege, hege@zib.de, Zuse Institute Berlin, Takustr. 7, Berlin, Germany, 14195; Friederike Fless, praesidentin@dainst.de, German Archaeological Institute, Podbielskiallee 69–71, Berlin, Germany, 14195; Christoph von Tycowicz, tycowicz@zib.de, Freie Universität Berlin, Kaiserswerther Str. 16–18, Berlin, Germany, 14195.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2023 Association for Computing Machinery.

Manuscript submitted to ACM

Manuscript submitted to ACM

1

ACM Reference Format:

Martin Hanik, Benjamin Ducke, Hans-Christian Hege, Friederike Fless, and Christoph von Tycowicz. 2023. Intrinsic shape analysis in archaeology: A case study on ancient sundials. *ACM J. Comput. Cult. Herit.* 1, 1 (May 2023), 28 pages. <https://doi.org/XXXXXXX.XXXXXX>

1 INTRODUCTION

Within the discipline of archaeology, artifacts, i.e., the physical remains of tools, weapons, clothing, adornments and other man-made objects, represent the largest and most diverse source of evidence. Notwithstanding the importance of ornamentation, use-wear traces, material composition, etc. the aspect of shape has long served as the primary means of establishing structure in both time (*chronology*) and space (*chorology*) within the vast universe of known artifacts [40]. In this section, we introduce the basic premises of a novel mathematical approach to unravel the manifold sources of information contained in the shapes of artifacts.

1.1 Shape and archaeology

Indeed, shape has always been fundamental to the analysis of archaeological artifacts. While this can also be said about the objects of interest of related disciplines, such as paleontology, geology and biology, the “human dimension” of artifact creation adds complex layers of technological, economic, artistic and other social components to the physical manifestations of shapes. As a consequence, archaeology—a discipline that has its roots equally in the prospering natural sciences and antiquarian movement of the 18th century [48, Ch. 1]—traditionally uses the method of *typology* to impose a more or less well-defined order, called *typological sequence* [48, Ch. 4], on a series of artifacts. It is hypothesized to isolate the dominant morphological trends and provide a basis for extraction and further, informal analysis of the unexplained residual trends: e.g., geographic detrending through spatial ordering facilitates the detection of chronological trends; additional temporal ordering may reveal social stratification. There is a large body of published research on the process of typological ordering (often called *seriation* in archaeology and *ordination* in ecology [42]), including computational methods of varying complexity (see, e.g., Refs. [38, 39, 51]). The combinatorial complexity of seriation implies that optimal results are computationally infeasible even for a relatively small number of artifacts, whereas archaeological excavations often yield tens of thousands of artifacts. Consequently, no single procedure is undisputed and manual grouping and ordering based on subjective mixtures of features (artistic/stylistic, technological, functional, etc.) is still common, as are simplified and abstract graphical representations (drawings) of objects. These discretizations, of both concepts and representations, into archaeological “types”, i.e., idealized forms that serve as representatives of classes of objects in a particular spatio-temporal context, are not free of arbitrariness, and they imply a loss of information.

The resulting limitations are becoming more apparent as new image-based technologies provide economical means for digitally capturing and representing shapes with full 3D detail and high accuracy. Our contribution aims to use innovative mathematical approaches to enable the transition from type-based analysis of archaeological object shapes to a geometric, purely data-driven analysis. We focus on knowledge inference from a whole *set* of artifact shapes in order to demonstrate the potential of modern digital repositories. We use the example of ancient sundial surfaces from Italy and Greece. These are well suited to our approach because their physical shape must necessarily reflect at least one controllable trend, that of the (intended) geographic location of use, in order to function properly. We ignore *a priori* typological categories and base our analysis solely on the shapes themselves and their derived geometric properties. We do, however, make use of geographical (object find spots) and, as part of an application example, cultural (Greek, Roman) contextual information provided along with the 3D data, to drive our analysis. The 3D models of the sundials

were generated by an image-based 3D reconstruction process called “Structure from Motion” and “Multi-view Stereo” (SfM/MVS), followed by surface reconstruction and optional texturing to enhance visual detail, but without effecting geometric properties. This process will not be discussed in any detail here, as it is well known also in archaeology (see, e.g., Refs. [17, 25]). It is important to note that the quality of the resulting data¹ is more than sufficient for the presented approach, and that the actual scale of the objects does not matter.

1.2 Shape and mathematical shape space

The starting point for our approach is a shape space: Although each artifact is embedded in 3D Euclidean space (with coordinates resulting from physical data acquisition), the set of all its surface boundary points can be considered as a single point in a high-dimensional, curved space; this is the *shape space*². For an introduction see, e.g., the textbooks [16, 37]. The shape space is key to understanding fundamental advantages of our approach over more traditional statistical methods.

The basic challenge is to separate the influence of variables of interest on the shapes of the artifacts from noise (which includes data acquisition and modeling errors as well as the influence of unexplained or irrelevant variance). This task, of course, falls into the classic domain of multivariate statistics, which provides a wealth of methods for identifying key trends in data as well as ordering and grouping them according to robust criteria. At its core, the highly developed multivariate apparatus is based on the analysis of the variance of the input data—measured in Euclidean distances. Methods such as principal components analysis, clustering analysis or correspondence analysis (a method that has gained much popularity in archaeology [13, 26]) all address the problem of sorting data into groups (alternatively: along components, factors, etc.) with minimal internal (“in group”) and maximal external (“between groups”) difference (“variance”). While the assumption of “Euclideanness” simplifies the underlying mathematics and provides robust results, there is increasing evidence that it limits the ability to capture variations in shapes [6, 37, 56].

But even if the assumptions of Euclideanness and global orthogonality are relaxed, the fundamental issue remains that the resulting groups (components, factors, etc.) represent synthetic variables, determined by purely formal criteria, that have no immediately meaningful relations with any trends that were identified *a priori*. Using conventional multivariate statistics, it is very hard to isolate and identify subtle contributions (e.g., by various geographic, social or technological processes) in such complex physical manifestations as the shapes of artifacts or architectural elements. These problems can be alleviated by analyzing shapes directly, without grossly simplistic transformations and the associated loss of information. When applied to geometric objects, this requires a notion of space that provides a complete structural characterization of their shape along with an appropriate measure of (dis)similarity that allows subtle and individual trends to be assessed and isolated. Although such a treatment leads to high-dimensional and nonlinear problems, efficient implementations can be obtained employing a recent type of shape space [56] that provides closed-form expressions for intrinsic operations and, thus, simple and fast algorithms. To give the interested reader an easy way to test the proposed methods and to ensure the reproducibility of our results, we provide a Python implementation as part of the Morphomatics library [2]; notebooks with our experiments can be found online under <https://github.com/morphomatics/SundialAnalysis>.

By utilizing curved shape spaces, we leave multivariate statistics in its traditional form. Remarkably, employing differential geometric concepts, we can obtain generalized statistical tools [27, 45] that take the rich geometric structure

¹The 3D data sets of sundials used have been published as open data by the Topoi research cluster (see App. A).

²More precisely, there are several (known) ways to encode the surface boundaries as points in generally different high dimensional, curved spaces, which are all called shape spaces; which one to choose depends on the application at hand.

of shape space into account (i.e., dispense with linearization), while being compatible with the Euclidean counterparts (i.e., agree with the original definition for the special case of flat vector spaces). The transition to curved spaces with the analytical methods available there allows for more flexible derivations and higher analytical sensitivity than multivariate statistics with its dependence on global Euclideaness [3, 34, 56].

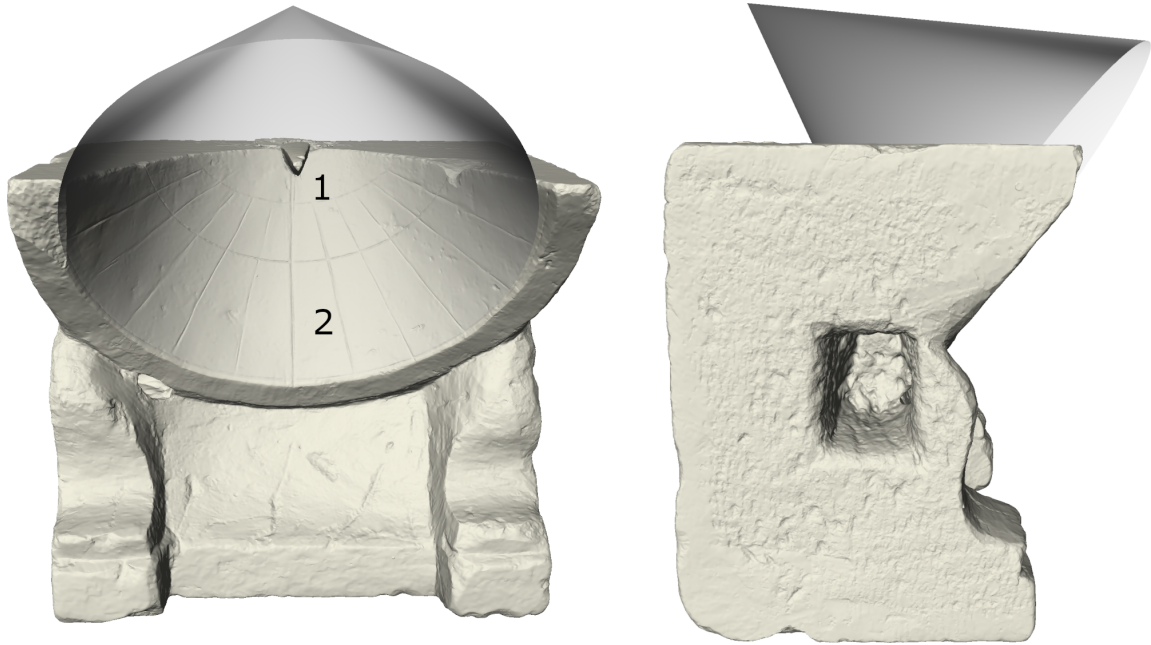


Fig. 1. Front (left) and side (right) view of a conical sundial with gnomon (1) and shadow surface (2). In order to craft the shadow surface, part of a cone was cut from the stone, as indicated by the gray cone in the picture.

1.3 Problem, analysis task

In antiquity, sundials were mainly used for the measurement of time [50]. While the sun is above the horizon, sundials indicate the time by the position of the shadow of a so-called *gnomon* on a shadow-receiving surface; see Fig. 1. In this article, we adopt the term *shadow surface* for the latter. Shape-wise, at least four types of sundials can be distinguished [49]: conical, spherical, cylindrical and planar sundials—determined solely by the shape of the shadow surface. In Figs. 1 and 2, we display a conical and a spherical sundial. Independently from the type, the geographical latitude had to be considered during construction to ensure proper functioning. In particular, it is well known that the *hour lines*, which were inscribed into the shadow surface to read off time, were adapted to the latitude of the location of installation [35, 36]. An obvious question then is whether other parts have also been changed: By identifying a latitude-dependent trend in the data, we show that, at least for the spherical type, it is very likely that the *shape* of the shadow surfaces was also adapted by the craftsman. We demonstrate an application of this trend by inferring the latitude of the place of installation of a sundial with uncertain geographical location of use. Furthermore, if one wants to compare shapes of shadow surfaces from different geographical locations, e.g., with respect to differences in construction principles, the shapes should be *normalized* to latitude; otherwise, one might misinterpret differences

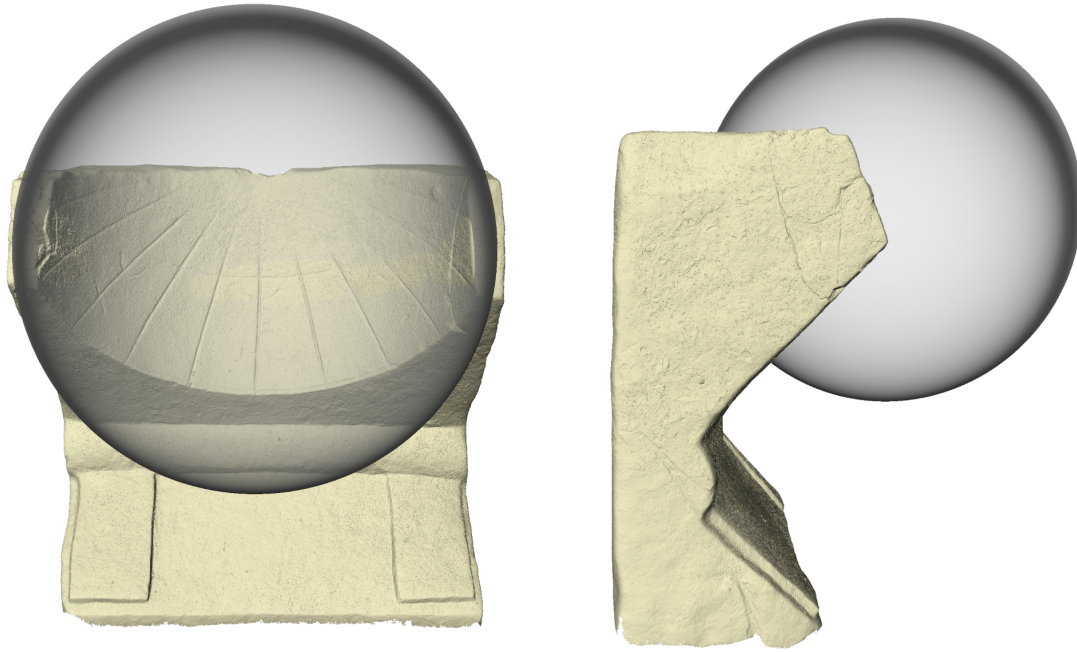


Fig. 2. Front (left) and side (right) view of a spherical sundial with missing gnomon. In order to craft the shadow surface, part of a round ball was cut from the stone, as indicated by the gray sphere.

that actually are “imposed” on the craftsman by the latitude of the intended place of use. In this work we develop the necessary methodology to perform this normalization. We then use it for an illustrative analysis, comparing the mean shapes of shadow surfaces from ancient Greek and Roman sundials. With more (reliable) data, such an analysis could—unbiased from latitudinal effects—reveal differences in construction principles between the groups.

To sum up, our contributions are the following. We

- demonstrate the use of statistical tools for shape data in archaeology,
- identify a latitude-driven shape trend for the shadow surface of ancient Roman sundials,
- use the trend to approximate the latitude of the place of installation of a sundial for which the latter is uncertain,
- present a new method for the analysis of parameter-dependent shape data that allows separating different influencing variables such as time and location.

The article is structured as follows: In Sec. 2, we describe the method. This includes the data and the pre-processing we performed as well as the shape space and the mathematical methods that we used. We then discuss the results in Sec. 3 and give a short conclusion in Sec. 5.

2 METHOD

In this section, we describe the analysis pipeline that we applied to analyze the shapes of the shadow surfaces. After presenting the data and the performed preprocessing, we describe the mathematical shape representation that was used,

and discuss statistical methods (in particular, generalizations of the mean and linear regression) that allow us to study sets of shapes. Furthermore, we introduce a novel method for normalizing shapes with respect to a given parameter and exemplify its use for studying residual shape variations. Since the normalization has not been published before, we give the mathematical details; further mathematical background is detailed in the Appendix.

2.1 Data and data preprocessing

Today, about 500 ancient Greek or Roman sundials are known and 3D models of many of them can be found in the repository of the Topoi Excellence of Cluster [24]. They were reconstructed in a Structure from Motion/Multiview Stereo (SfM/MVS) procedure, as detailed in Ref. [22]. For this study, we used spherical sundials from Greece and the Italian peninsula and considered all available models with well preserved shadow surface; their IDs and sites are given in App. A.

As for data preparation, we extracted the shadow surfaces with the software Amira [54]. We manually selected the shadow surfaces and corrected triangles near the new boundary that had extreme angles. Furthermore, the resolution of the surface meshes was reduced to about 20 *k* faces using the *quadric edge collapse decimation scheme* implemented in the free software MeshLab [12]. An example of a segmented shadow surface can be seen in Fig. 3; it is shown after extraction and simplification in Fig. 4.

To perform statistics on shapes we rely on a group-wise correspondence, i.e., consistent point-to-point relationships of the meshes. There are many approaches to establish such correspondences, e.g., non-rigid registration [11, 55] and functional map-based [43] approaches. We obtained one by choosing a shadow surface as reference and registering it with a harmonic map to all other surfaces through the following two steps: First, we registered the boundaries by using 3 landmarks. While the two geodesically farthest points were automatically approximated as the global minimizer and maximizer of the surface’s Fiedler vector (the eigenvector corresponding to the second smallest eigenvalue of the surface’s Laplace-Beltrami operator), the third landmark was manually selected as the lowest point of the gnomon hole. (The gnomon itself was missing at all but one of the sundials under study.) Then, we established the correspondence of the interior part using the *discrete harmonic map for disc-like topology* as described in Ref. [10]. Finally, effects of rotation and scale were removed by a group-wise Procrustes alignment of the meshes [23]. The processed meshes are available online; see [30].

Note that on many sundials the upper left and right corners are only poorly preserved, showing a varying degree of decay. Thus, the correspondences of the worn edges (which are rather small in comparison to the rest of the sundial) are only approximate. In the later analysis we therefore concentrated on the other areas (that are conserved very well).

2.2 Shape spaces

Informally in mathematics, a “shape” is the set of all the geometric properties of an object that are invariant under similarity transformations, i.e., under translations, rotations and scalings. To analyze sets of shapes one needs the notion of shape spaces, i.e., spaces in which shapes naturally “live”, that is, in which each point represents a full geometric shape. (For a good references on the analysis of individual shapes see Ref. [9].) There are different approaches to construct such shape spaces and to utilize them, depending on application characteristics. A classical example is Kendall shape space [16, 37], which relies on landmark configurations. An overview of shape spaces whose elements are diffeomorphisms (i.e., differentiable, bijective maps with differentiable inverses) can be found in [6]—with more in-depth discussions, e.g., in Refs. [53, 58]. Skeletal models are discussed in Refs. [47, 52]. Finally, physically-based spaces were investigated in Refs. [3, 31, 56]. Almost all of the above shape spaces have an inherent non-Euclidean structure. Luckily,

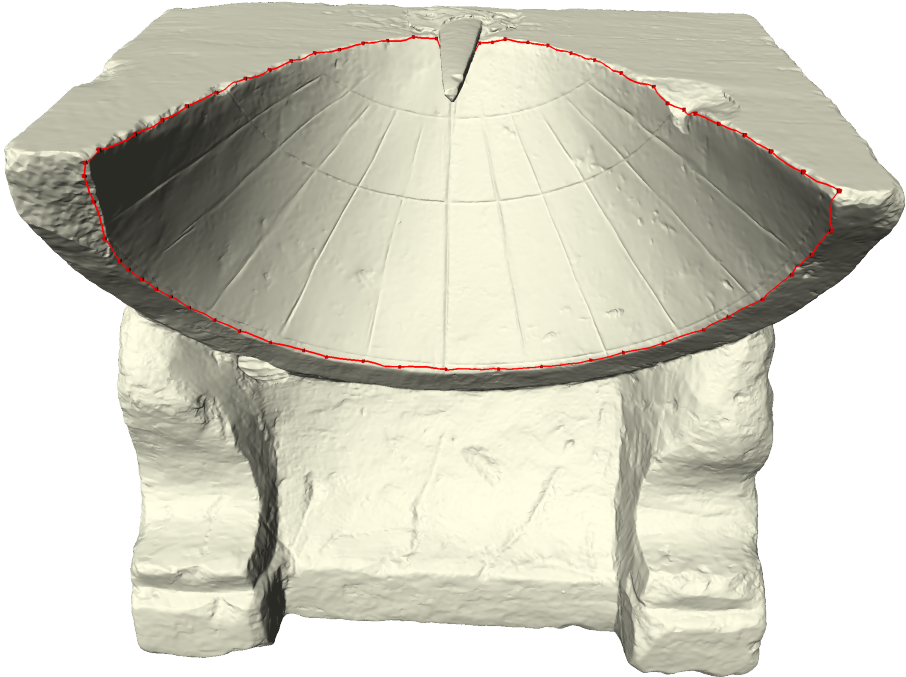


Fig. 3. Manually segmented shadow surface

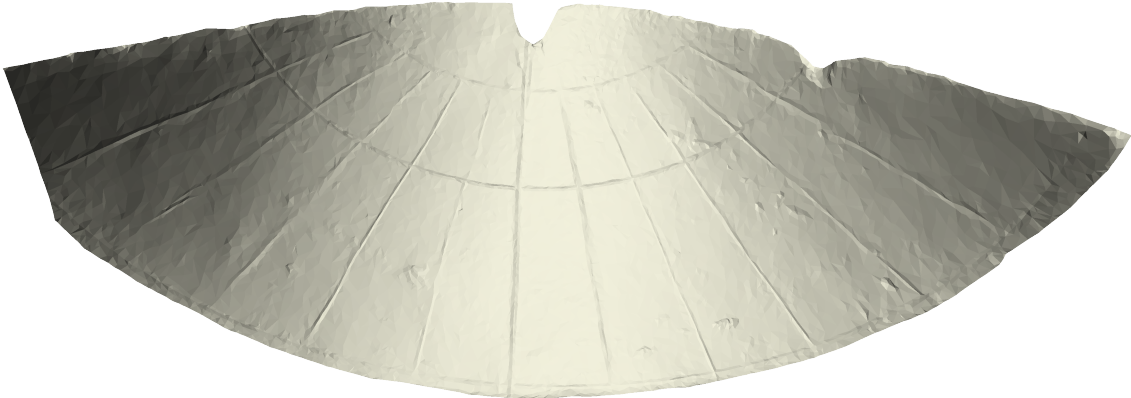


Fig. 4. Extracted shadow surface

techniques from the mathematical field of Riemannian geometry (see App. C for an introduction) can either be directly applied (since shape spaces are so-called manifolds) or have been successfully extended to them. This fact is utilized by geometric statistics (see Sec. 2.3) to generalize concepts from multivariate statistics for the analysis of shapes.

Before proceeding, we would like to emphasize that the methods we present in this article are not limited to any particular choice of shape representation; we require only that the shapes are modeled as elements of a Riemannian manifold, i.e., a manifold with a defined way of measuring angles and distances.

For our case study, we chose the shape space of *differential coordinates* [56], whose mathematical details are summarized in App. B. This space can be used when the objects under study are digitally given as triangular meshes in correspondence, the latter being a semantically motivated one-to-one relation of the (necessarily same number of) vertices.

2.3 Statistics on manifolds

Statistical analysis is the de-facto standard for sets of data. To apply it in non-Euclidean manifolds a generalization of methods from multivariate statistics is required. The first steps in this direction were taken a long time ago, e.g., by Fréchet, but recently there have been many advances. With increasing computational power, it is now possible to incorporate the nonlinear structure inherent in many types of data into our models of reality and still perform the computations in a reasonable amount of time. As a result, there has been and continues to be an increasing amount of research that has led to a new field often referred to as *geometric statistics*. As we will develop a new statistical tool for the analysis of the shadow surface shapes in Sec. 2.4, we introduce the notions from geometric statistics that are necessary to define it in the following.³ Mathematical background can be found in App. C. Since our new method works in any Riemannian manifold, we formulate it in general terms. As part of our study of the shadow surfaces in Sec. 3 we apply it in the shape space of differential coordinates.

Let M be a manifold. Since linear combinations of its elements need not lie in M again, even the simplest statistical index, the mean, is *not* well-defined. A Riemannian metric, i.e., a smoothly varying inner product $\langle \cdot, \cdot \rangle_p$ on the tangent spaces $T_p M$, $p \in M$, yields a solution: It allows to measure angles between tangent vectors from the same tangent space and induces a distance between any two points in M . The resulting distance function, in turn, gives rise to a notion of mean (see, e.g., Ref. [45, Sec. 2.2]) and principal component analysis (PCA) (see, e.g., Refs. [21, 33]), as well as further essential statistical quantities and procedures.

Around every point $p \in M$ the manifold can locally be mapped one-to-one into the tangent space $T_p M$ under preservation of point-to-origin distances. The corresponding function is called Riemannian logarithm \log_p . Its inverse is the Riemannian exponential \exp_p . Because of this, $\log_p(q)$ can be interpreted as the difference between $q \in M$ and $p \in M$: It is the vector in $T_p M$ that points to q and is as long as the distance between them. A major difference between curved Riemannian manifolds and Euclidean space appears when one wants to compare vectors (and thus the above differences) from tangent spaces at different points: In manifolds, we must explicitly transport one to the other along a curve; this operation is called *parallel transport*. Whenever the space is curved, the result of the transport depends on the chosen path.

Geodesics are very useful objects in Riemannian geometry. Being generalized straight lines in M , they are also (locally) shortest paths. In particular⁴, they allow for the generalization of multivariate linear regression—a problem that can be understood as finding the best fitting straight line to a vector-space-valued data set. Analogously, in geodesic regression one tries to find the best fitting geodesic for manifold-valued data [19]. The latter method has mostly been used in biological and medical applications; however, in what follows we will show that it can be made fruitful for problems in archaeology as well. The mathematical details of geodesic regression can be found in App. C; with Fig. 22 it also yields a visualization of the underlying idea. In our study we used this method to regress the shapes of the spherical

³Given data in a manifold we will always assume that all data points can be pairwise connected by a unique shortest path. This is the case when the data is sufficiently localized in the shape space. For example, within every hemisphere of the 2 dimensional sphere any two points can be joined by a unique segment of a great arc. Thus, without mentioning it again, we always assume that the data is localized in such a neighborhood.

⁴In fact, geodesics play a major role in almost all concepts of geometric statistics, including the generalization of the mean and PCA, as they are the curves that realize the distance between points.

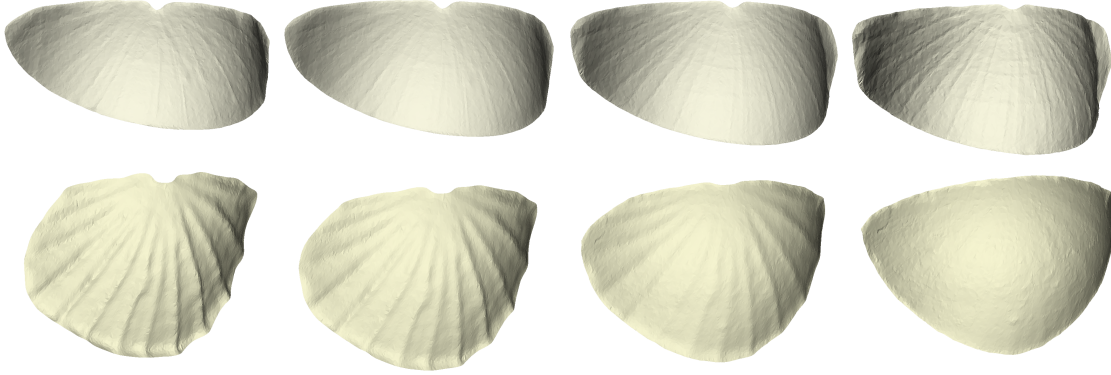


Fig. 5. Results of geodesic regression for shadow surfaces of spherical sundials w.r.t. latitude. The results for Roman and Greek sundials are shown in the top and bottom row, respectively. After calculating the trajectories in the space of differential coordinates, we sampled each curve at 4 points and computed the corresponding triangular meshes. The Roman geodesic models the trend for latitudes in $I_R = [40.7030, 43.3155]$ while the Greek one is defined on $I_G = [36.0917, 37.3900]$.

shadow surfaces from each group (Roman and Greek origin) against the latitude of their site using the differential coordinates space; the result for spherical sundials are shown in Fig. 5.

Finally, note that regression with generalized polynomials is also possible on manifolds [28]. We rely on geodesics in this article, but these higher order methods could be easily incorporated if a more complicated dependence on the parameter is expected.

2.4 Regression-based normalization of data

Let M be a Riemannian manifold (e.g., a shape space) and $I_i \subset \mathbb{R}$, $i = 1, \dots, k$, closed intervals with non-empty intersection, i.e., $\cap_{i=1}^k I_i \neq \emptyset$. The latter assumption guarantees a *common* parameter value that serves as a reference point for the normalization. We further assume that we are given k groups of data points in M , where each element is labeled with a parameter:

$$\left(q_j^{(i)}, t_j^{(i)} \right) \in M \times I_i, \quad i = 1, \dots, k, \quad j = 1, \dots, n_i.$$

We assume that within each group the variability of the points $q_j^{(i)}$ with respect to the parameters $t_j^{(i)}$ can be well approximated by a geodesic. Our goal is an inter-group comparison of the data without the confounding influence that is caused by different values of the parameters t_i . Therefore, we suggest to normalize the data in each group to the same value $t_0 \in \cap_{i=1}^k I_i$ using geodesic regression. Applying it separately to each group yields best-fitting geodesics $\gamma^{(i)} : I_i \rightarrow M$, $i = 1, \dots, k$. These geodesics can already provide valuable information about the data, as we will see when we compute them for sundials.

The difference of each point $q_j^{(i)}$ to its corresponding point $p_j^{(i)} := \gamma^{(i)}(t_j^{(i)})$ on $\gamma^{(i)}$ can be encoded as the tangent vector

$$v_j^{(i)} := \log_{p_j^{(i)}} \left(q_j^{(i)} \right) \in T_{p_j^{(i)}} M.$$

Using parallel transport, we can translate these vectors along $\gamma^{(i)}$ to $\gamma^{(i)}(t_0)$, resulting in vectors

$$w_j^{(i)} \in T_{\gamma^{(i)}(t_0)} M.$$

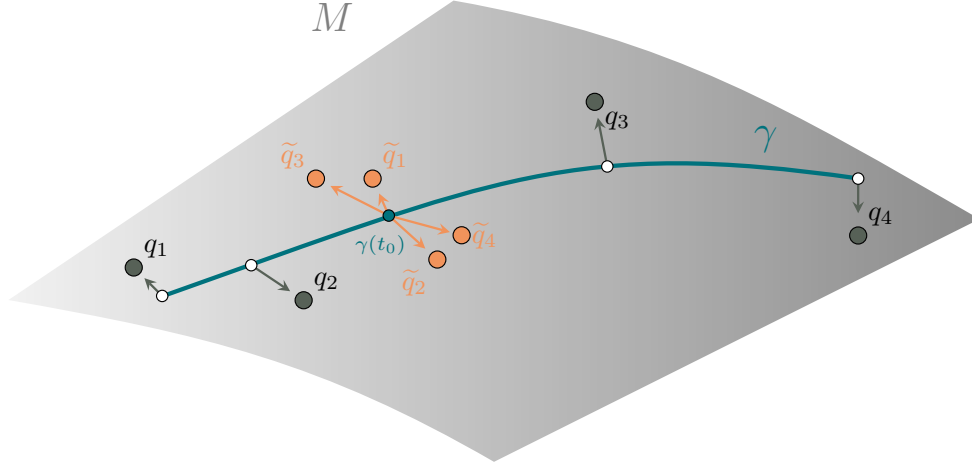


Fig. 6. Normalization w.r.t. some parameter t for a single data group (q_j, t_j) , $j = 1, 2, 3, 4$, in (a part of) a shape manifold M : Each point represents a complete shape. The curve γ is the result of geodesic regression w.r.t. t . The points $\gamma(t_j)$ are depicted in white, while the tangent vectors v_j are the gray arrows. Finally, the parallel translations w_j of the v_j in the tangent space $T_{\gamma(t_0)}M$ at $\gamma(t_0)$ are shown in orange, yielding the normalized data \tilde{q}_j .

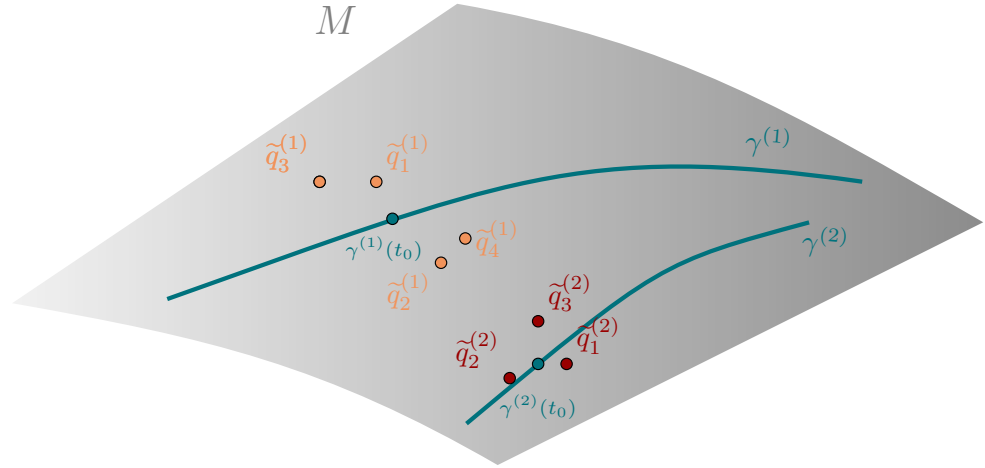


Fig. 7. Normalization results for two groups: Both geodesics γ_1 and γ_2 are defined on the same interval and the data is normalized at the same value $t = t_0$. The groups (orange and red) can now be compared without a bias caused by the influence of variable t .

They represent the differences of the data points to the geodesics *normalized at t_0* . Mapping them back to the manifold gives the normalized data points:

$$\tilde{q}_j^{(i)} := \exp_{\gamma^{(i)}(t_0)}(w_j^{(i)}).$$

Now, we can perform *inter-group comparison* at the normalized value t_0 . For example, we can compute the Fréchet mean and perform group tests for equality of them; for the latter see Refs. [41, Sec. 3.3] and [29]. The normalization process and its result for the data of two groups are visualized in Figs. 6 and 7. We also provide an example:

Example 2.1. Let M be a shape space with manifold structure like the one introduced in Sec. 2.2. Suppose that we are given two sets of knives, say from Greece and Spain, and let p_1^1, \dots, p_m^1 and p_1^2, \dots, p_n^2 be their encoded shapes in M , respectively. Further, suppose that we know their creation dates t_1^1, \dots, t_n^1 and t_1^2, \dots, t_n^2 , which lie between the years 0 and 300 AC.

Now, assuming a continuous evolution of knife shapes over time, we can apply our method to compare them after removing the chronological trend. For this, we set $I_1 = I_2 := [0, 300]$; then, the shapes $\tilde{q}_j^{(i)}$ approximate how the knives of both groups looked at any (fixed) time $t_0 \in [0, 300]$.

Note that geodesics can be extended in almost all practical scenarios (see Ref. [15, Ch. 7] for details). (In ordinary Euclidean space, this corresponds to continuing a straight line.) This allows—within limits that depend on the data—to extrapolate the trends γ^i so that they are defined on larger intervals. This can be very helpful when $\cap_{i=1}^k I_i = \emptyset$, but moderate extrapolation assures overlapping intervals. We used this property in our study to extend two intervals so that their intersection is non-empty.

2.5 Analysis pipeline

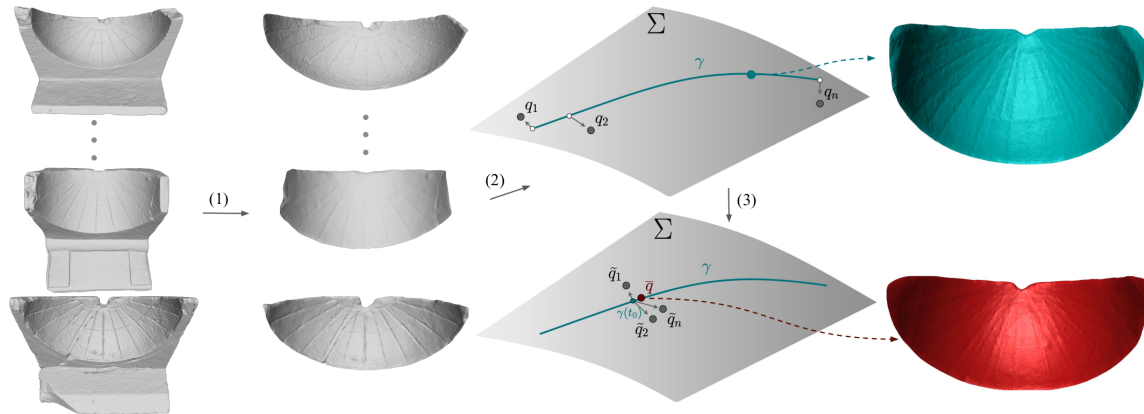


Fig. 8. Visualization of the analysis pipeline.

We now discuss the analysis pipeline that we applied separately to the Roman and Greek groups; it is visualized in Fig. 8. Let n denote the number of sundials in the group. After digitizing the sundials in the form of triangle meshes, shadow surfaces were extracted and a group-wise correspondence was established (1). The shapes q_1, \dots, q_n were then encoded in the shape space Σ of differential coordinates (2). The statistical analysis of the shapes of the shadow surfaces was then performed in the manifold Σ ; we applied geodesic regression [19] w.r.t. latitude, which yielded the fitted geodesic γ (as depicted). Being a function that depends on latitude, we could evaluate γ at various latitudes. Any resulting shape represents a typical shadow surface at the given latitude. It can be transformed into a triangular mesh for visualization (the green shadow surface in Fig. 8).

The Roman geodesic already revealed interesting information: We used it to infer the latitude of the installation site of an unknown sundial (see next section).

In step (3), the novel normalization method was applied to normalize all shapes to a common latitude t_0 , resulting in points $\tilde{q}_1, \dots, \tilde{q}_n \in \Sigma$. The shapes of the shadow surfaces could now be further analyzed without the additional

variability introduced by the influence of latitude. From these normalized shapes, we computed the mean shape \bar{q} at t_0 . A visualization of this mean (the red shadow surface in Fig. 8) can again be obtained when we transform \bar{q} back into a triangular mesh.

3 RESULTS

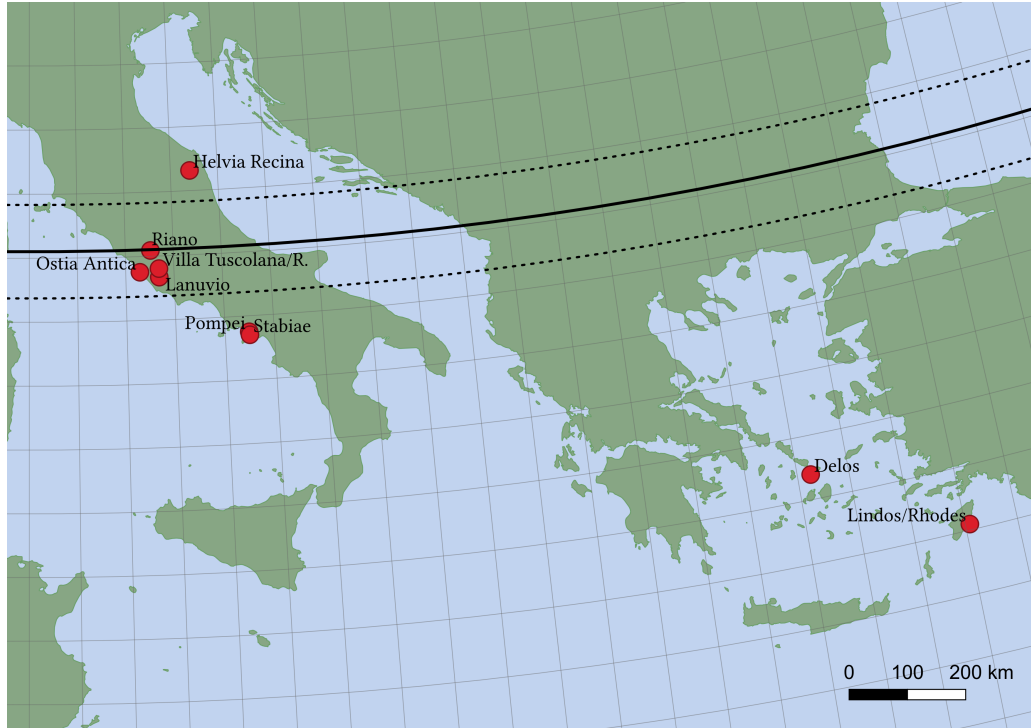


Fig. 9. Map of the relevant area: The sites of the used sundials are given by the red circles. The solid line indicates the latitude at which the sundial with Dialface ID 39 was probably installed. The dashed lines bound the region of uncertainty of this prediction as determined by the computed MAE from all sundials with known location of installation.

In this section, we discuss the results of our case study. First, we describe how we used geodesic regression to analyze the latitude dependence of ancient Roman sundials from the Italian peninsula and how we used the results to narrow down the possible location of installation of a sundial with uncertain site. Then, we show an illustrative application of the normalization method from Sec. 2.4 by comparing shadow surfaces from Greece with those from the Italian peninsula.

We used 10 Roman sundials from the Italian peninsula; their IDs as well as their sites—including longitude and latitude—can be found in App. A; the latter are also depicted in Fig. 9. We pre-processed them as described in Sec. 2.1.

Given the shapes in terms of differential coordinates (i.e., points in shape space), we performed geodesic regression with the latitude of the site as explanatory variable; the corresponding interval was determined by the latitudes of the most southern and northern site, i.e., $I_R = [40.7030, 43.3155]$. It seemed reasonable to assume that a change in latitude causes a geodetic drift in the data: in the range of latitudes we studied, shapes can be expected to change at a

constant rate in a fixed direction in shape space, to compensate for the latitude-dependent shift in sunlight angle.⁵ Equidistant samples of the resulting geodesics are depicted in the top row of Fig. 5 and Fig. 10. They show a “bending” of the shadow surface that increases with latitude; this is highlighted in Fig. 10.

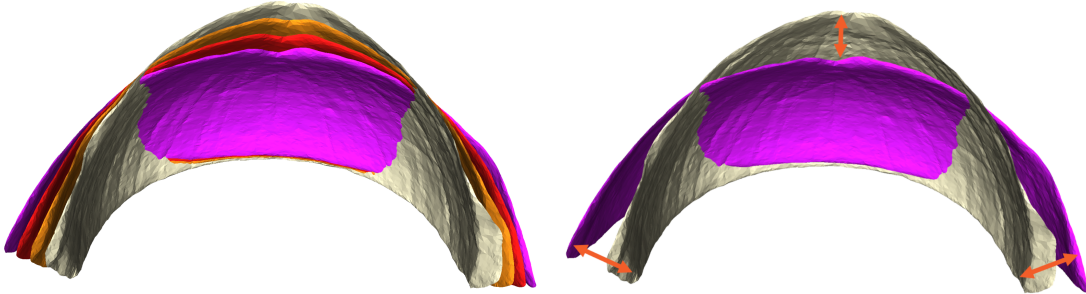


Fig. 10. Geodesic regression w.r.t. latitude over the interval $I_R = [40.7030, 43.3155]$ for shadow surfaces of Roman sundials. We evaluated the optimal geodesic at 4 equidistant points and show the top view of the resulting shadow surfaces, with colors beige, orange, red, and violet ranging (in this order) from the most northern to the most southern site. On the left, the shadow surfaces are overlaid to show the bending. The latter is also indicated by arrows on the right hand side, where only the shadow surfaces belonging to the most northern and southern latitude are shown.

The “bending” is an interesting effect: it suggests that the *form* of the whole shadow surface (and not only that of the hour lines) were adapted to the location of installation⁶. This fact can be used to place sundials of unknown or uncertain installation site: after projecting orthogonally [4, Sec. 3.3] the shape of the sundials’ shadow surface onto the regressed geodesic (which here mostly amounts to finding the most similarly bend shadow surface on the curve), we can “read off” the latitude at which the sundial was probably installed (see Fig. 11).

To demonstrate this, we selected a sundial from Italy (Dialface ID 39 in the Topoi database; see Table 1), which is now in a museum in Vatican City, but whose installation site is uncertain. (Fig. 12 shows a visualization of this sundial.) Projecting the shape of its shadow surface onto the Roman geodesic gave an approximate latitude of $t^* \approx 42.1029^\circ$, which is slightly north of the latitude of Rome. Thus, our model suggests that the sundial was once located at a site near the *solid* line shown in Fig. 9. In order to investigate the accuracy of this prediction we performed a leave-one-out cross validation test for the Roman sundials with *known* location: Each sundial was taken out, regression was computed using all other samples, and the latitude of the left-out sample was predicted. The *mean absolute error* (MAE) we thereby found was approximately 0.75° latitude (with standard deviation of 0.46°), which is about 80 kilometers in north-south direction.

We use the computed MAE as a measure of uncertainty for the prediction of the unknown sundial and depict the latitudes $t^* \pm 0.75^\circ$ as *dashed* lines in Fig. 9; it is likely that the location of installation was inside the resulting band. To compare our method against a traditional approach, we additionally applied *partial least squares regression* [57] on the Procrustes aligned coordinates of the shadow surface meshes in order to predict the corresponding latitude. For this, we used the “PLSRegression” module from scikitlearn 1.1.1. Setting the number of components to 1 yielded the best results. The MAE we thereby obtained was approximately 0.87° latitude (with standard deviation of 0.78°). This was about 0.12°

⁵To corroborate this, we also tested regression with higher-order curves in shape space, i.e., quadratic and cubic generalized Bézier curves, as explained in Refs. [28]. This overfitted the data, i.e., nearly interpolated the given shadow surfaces at their corresponding latitudes, *while showing unreasonable behavior in between*.

⁶It would be interesting to know how this was technically realized in production. However, we are not aware of any ancient source that provides information about this.

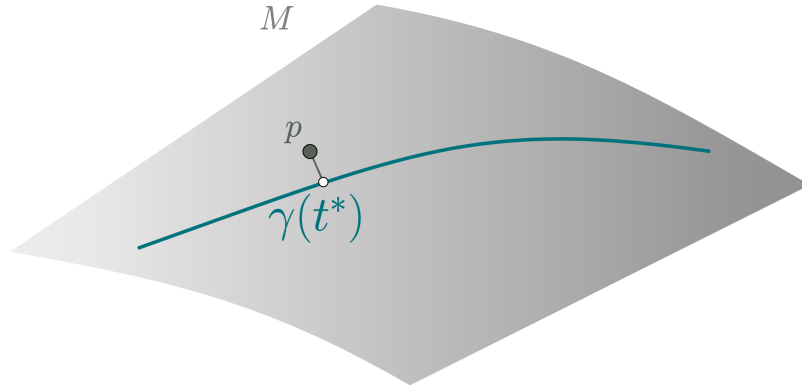


Fig. 11. Projection of an arbitrary point p in a manifold M onto a geodesic γ . The number t^* is the corresponding parameter of γ , i.e., the minimizer of $\mathcal{E}(t) := d(\gamma(t), p)$, where d is the Riemannian distance in M .

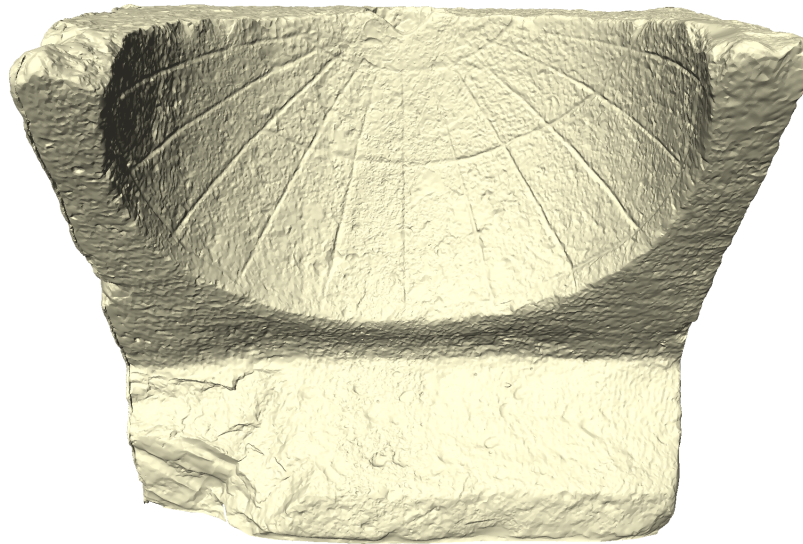


Fig. 12. 3D model of the sundial with dialface ID 39.

(about 12.32 km) larger than the MAE of our proposed method. Furthermore, the standard deviation was substantially higher when using PLS regression. The sundial-wise errors and MAEs of both methods are shown in Fig. 13.

We used the normalization method to compare the mean shape of shadow surfaces from the Italian peninsula with that of 3 samples from Greece; information on the latter can be found in Table 1 while their sites are also depicted in Fig. 9. Controlling for latitude before calculating the means correctly accounts for the fact that all Greek sundials were located further south than those from the Italian peninsula (an essential step at least for the Roman sundials).

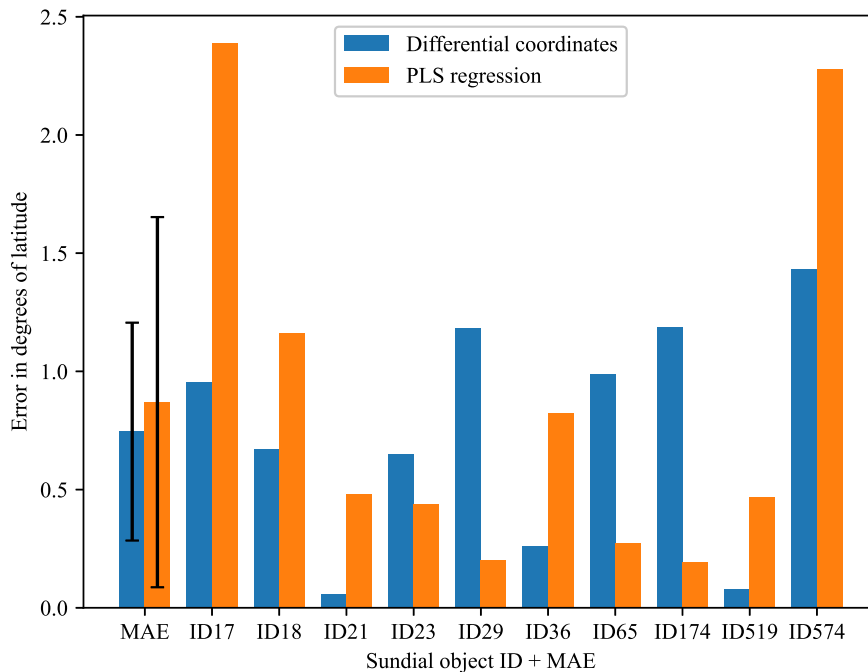


Fig. 13. Sundial-wise and mean absolute errors obtained during latitude prediction. “Differential coordinates” refers to the proposed method; “PLS regression” stands for the results of the method that is based on partial least squares regression. MEAs are depicted with standard deviations.

Because the experiment could only be performed with 3 Greek samples, it should be seen as an illustrative example. Nevertheless, we discuss possible research questions that our experiment could answer in the discussion section.

We performed geodesic regression with respect to latitude for the Greek sundials over the underlying interval $I_G = [36.0917, 37.3900]$; samples from the resulting geodesic are shown in Fig. 14. The shadow surface hardly bends with changing latitude contrasting our earlier result for Roman sundials. Since $I_R \cap I_G = \emptyset$, we extrapolated both the Roman and the Greek geodesic such that they were defined on the intervals $\tilde{I}_R := [38.5, 43.3155]$ and $\tilde{I}_G := [36.0917, 39]$, respectively. This allowed us to normalize both data sets at different latitudes and, in each case, to compute the mean shape of the normalized shadow surfaces as described in Sec. 2.4; the results are shown in Fig. 15. Note that the means “bend” like the geodesics, which is the effect of the normalization.

Widening the window for comparison by further extrapolation was of no avail due to increased artifact formation in the normalized data. The appearance of artifacts can be attributed to initial noise (occurring mainly at the corners, as explained in Sec. 2.1) that is amplified too much. First signs of such artifacts can be seen in Fig. 15; they begin to form at the corners of the Greek mean at 39° latitude.

In Fig. 16, we show both means after normalizing at 38.5° latitude. When overlaid they look quite different—in particular their curvatures. (The artifacts at the corners are again due to extrapolation of noise.) In order to quantify

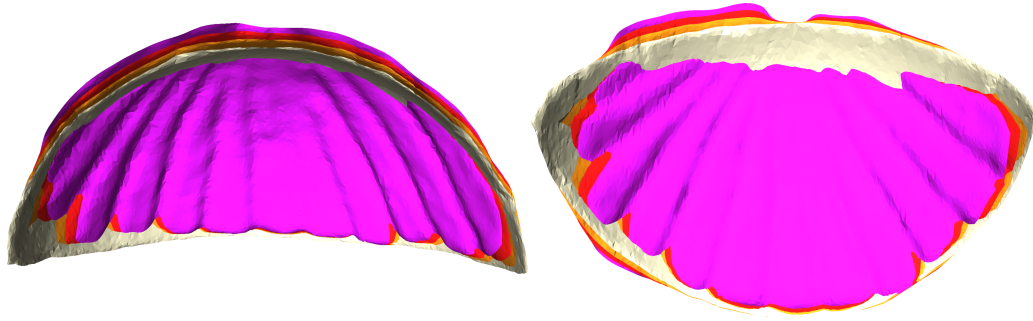


Fig. 14. Geodesic regression w.r.t. latitude for shadow surfaces of sundials from Greece $I_G = [36.0917, 37.3900]$; the top view is on the left while the front view is on the right. We evaluated the resulting optimal geodesic at 4 equidistant points, with colors beige, orange, red, and violet ranging (in this order) from the most northern to the most southern site.



Fig. 15. Roman (left) and Greek (right) means of shadow surfaces of spherical sundials after normalizing the Roman data at 40° (red), 41° (orange) and 42° (beige) latitude and the Greek data at 37° (red), 38° (orange) and 39° (beige) latitude, respectively.

this, we computed the spheres that fit the shadow surfaces best⁷; the results can be seen in Fig. 17. Both were different in size, the Greek sphere’s radius measuring 66% of the Roman. (Since size is normalized in our study only relative notions are of interest.) The fact that the “bending” is different influences the comparison of the means. As shown, at 38.5° latitude they are differently curved; because of the stronger bending in the Roman group the forms would converge, though, if we used this trend to compare them further north.

In order to compare the non-linear shape space with a Euclidean alternative (which is difficult as no statistically significant quantitative results can be expected because of the small sample size) we investigated how the Mahalanobis distance between the mean shapes of the Roman and Greek groups adapts to the normalization. After mapping the Roman *samples* and the Greek *mean* to the tangent space at the Roman *mean*, we computed the Mahalanobis distance⁸ of the Greek mean to the distribution of the Roman samples. (We chose this direction because of the numbers of samples.) This was done for both the original data and the one normalized at 38.5° . Afterwards, the same computations were performed using the standard Euclidean shape space from [14]. For both shape spaces the distance increased

⁷More precisely, we computed the center $c^* \in \mathbb{R}^3$ and radius $r^* > 0$ of the sphere with smallest sum of squared distances to the vertices of the shadow surface. Denoting the usual Euclidean norm by $\|\cdot\|$, it can be seen that the distance of a point $p \in \mathbb{R}^3$ to a sphere with center c and radius r is equal to $\tau(c, r; p) := \|\|p - c\| - r\|$. We thus computed $(c^*, r^*) := \min_{(c,r) \in \mathbb{R}^3 \times \mathbb{R}_{>0}} \sum_i \tau(c, r, v_i)^2$, where the v_i are the vertices of the mesh of the shadow surface. The “L-BFGS-B” method was used for minimization. The problem is discussed in depth in Ref. [8].

⁸See Ref. [44, Sec. 7] for the Mahalanobis distance in manifolds. As the dimension of the differential coordinate space is (much) higher than the number of samples, the covariance matrix was not invertible. Therefore, as is often done, we substituted the inverse of the covariance matrix of the Roman samples by its pseudoinverse to compute the distance.

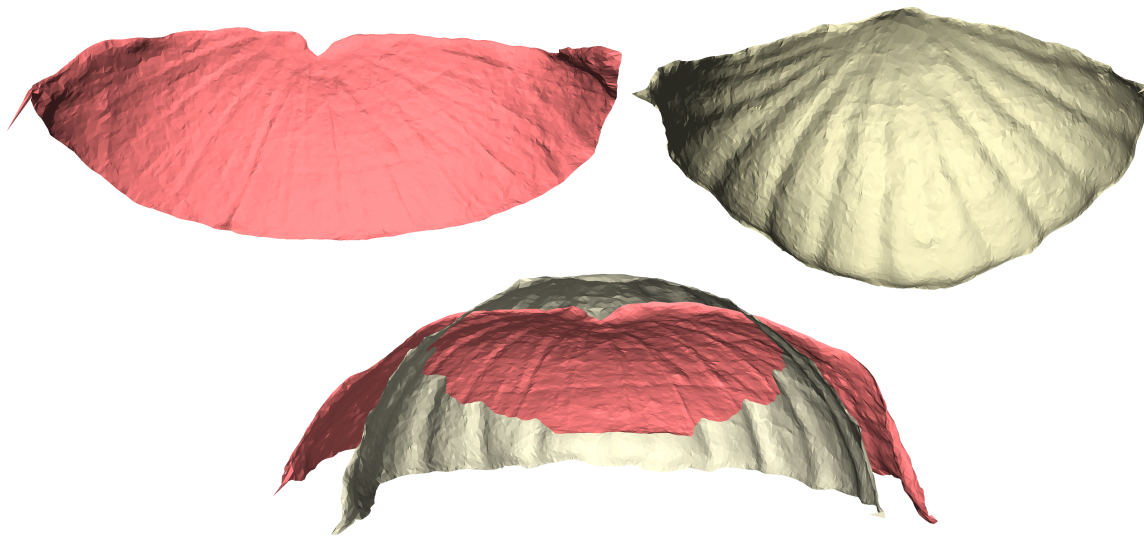


Fig. 16. Comparison of the Roman (red) and Greek (beige) mean shapes after normalizing the respective samples with the proposed method at 38.5° . They are shown both individually (top row) and overlaid (bottom row).

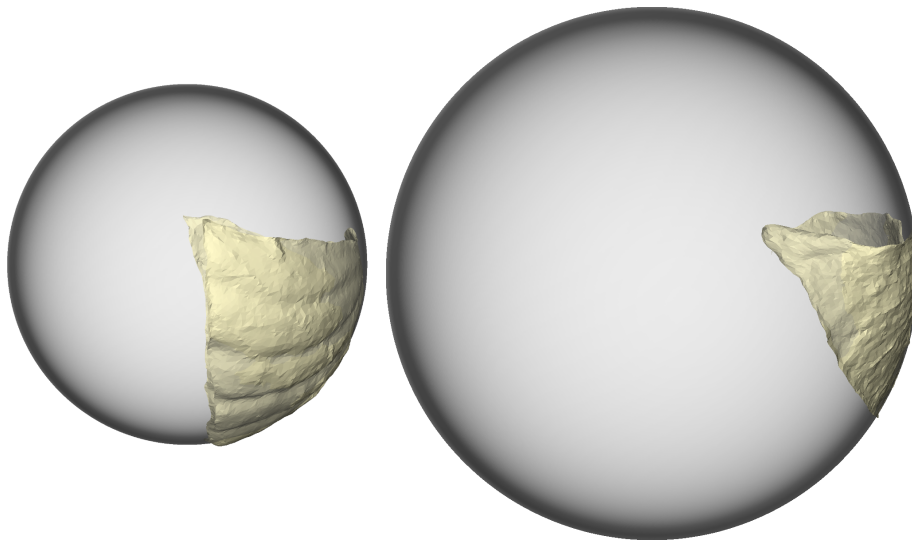


Fig. 17. Spheres fitted to the Roman (left) and Greek (right) mean of the data that was normalized at 38.5° latitude. Center and radius of the spheres were chosen such that the average squared distances to the vertices of the triangle meshes were minimal.

when the samples were normalized. This makes sense because the in-group variability (here within the Roman group) reduces through the normalization (cf. Figs. 6 and 7). As the Mahalanobis distance weights the difference between the means (i.e., the logarithm) against the inverse of the observed covariance (matrix) in the Roman group, a decrease of the latter leads to a larger value. Interestingly, while the distance for the nonlinear space increased by a factor of 1.73,

it was multiplied by 5.1×10^5 for the Euclidean model.⁹ (Both numbers are rounded.) That is, the latter loses a high amount of variability while normalizing. We think that the differential coordinates conserve the finer details, whereas the Euclidean model tends to focus more on global features. But we believe that for the sundials smaller details are of high interest, in particular, when additional characteristics like the hour lines would be investigated as part of further analysis.

We also examined to which extent the results depend on the resolution of the meshes by simplifying them. More precisely, we reduced the number of faces of each of the shadow surfaces (further) to 10k, 2k and 1k triangles with Meshlab’s “quadric edge collapse” and computed normalized means for both the Roman and Greek group. As can be seen in Figs. 18, 20, and 19, the means were effectively unchanged even with a poor resolution of only 1,000 faces.

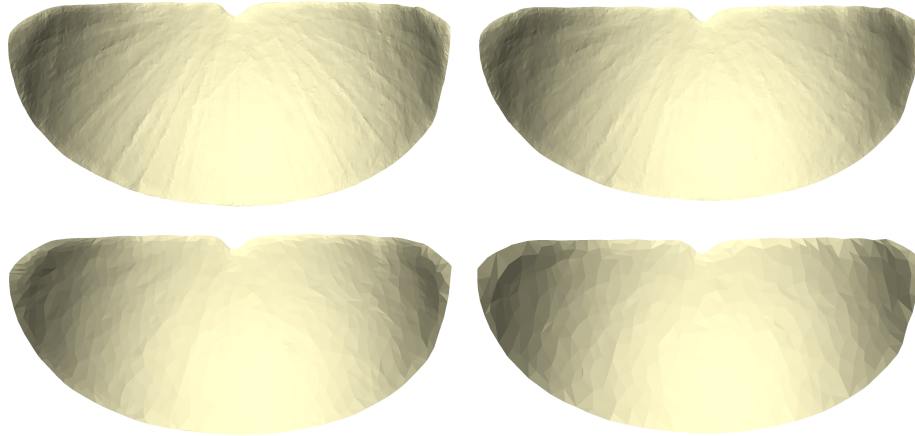


Fig. 18. Mean shapes of the normalized (at 41°) sundials from the Italian peninsula for different mesh resolutions. Results are shown for resolutions of 20k (upper left), 10k (upper right), 2k (lower left) and 1k (lower right) faces.

4 DISCUSSION

The fact that the shape of the Roman shadow surfaces “bends” with latitude is a significant finding of this study. To the best of our knowledge it has never been observed before. It seems likely that the form of the whole shadow surface was adapted to the location of installation—a conclusion that is supported by the accuracy of the derivative method for latitude prediction. Our results thus add a piece to the puzzle how ancient Roman craftsmen created working sundials. Clearly, an interesting question for future research is whether similar bending can also be observed for sundials from locations outside of Italy.

Our results also indicate that the new method to identify the latitude of the location of installation is accurate; furthermore, they suggest that utilizing the non-Euclidean shape space structure for prediction is superior to the commonly employed PLS regression. The only other method for the prediction of a sundial’s working latitude we are aware of is the one from Pistellato, Traviglia, and Bergamasco [46]. Their method predicts the working latitude by utilizing the positions of the shadow surface’s inscribed hour lines. Knowledge of celestial mechanics, which they assume the craftsmen possessed, allows them to compute the latitude from these positions. While their method has

⁹We also investigated how the geometric distance between the means behaves under the normalization. It increased by a factor of approximately 1.86 and 1.64 for the differential coordinate space and the Euclidean shape space, respectively. Therefore, the extreme increase in Mahalanobis distance for the latter cannot be explained by diverging means.

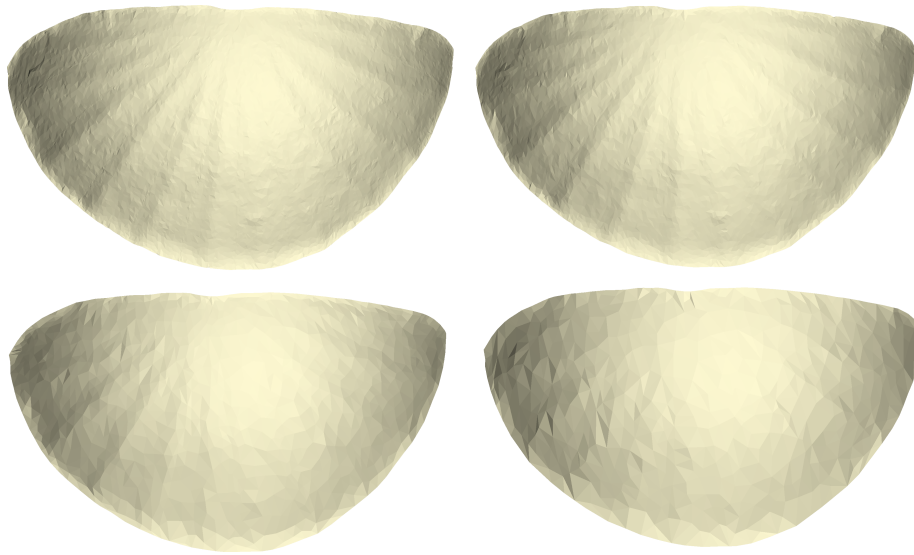


Fig. 19. Mean shapes of the normalized (at 37°) sundials from Greece for different mesh resolutions. Results are shown for resolutions of 20k (upper left), 10k (upper right), 2k (lower left) and 1k (lower right) faces.

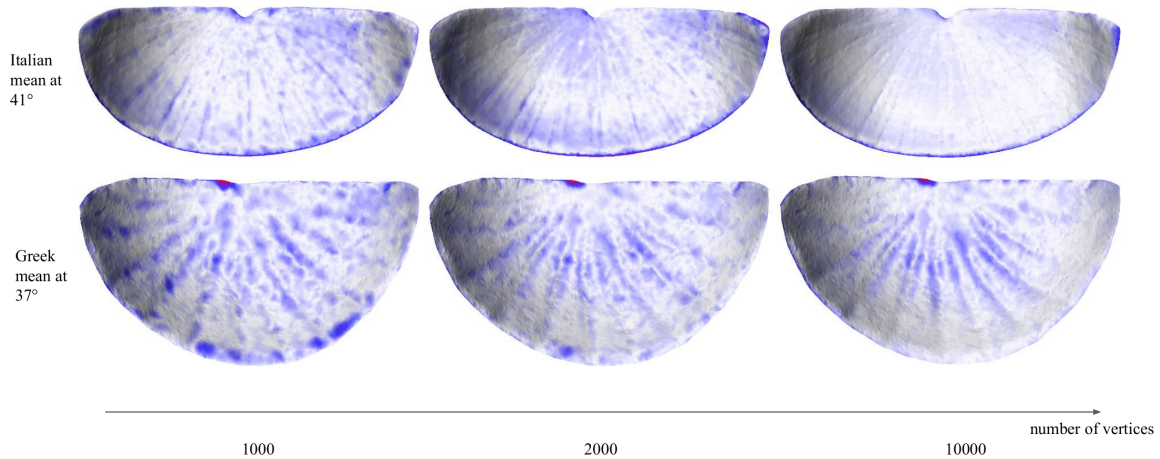



Fig. 20. Surface distance for different mesh resolutions when computing normalized means. The respective meshes of highest resolution are depicted with coloring (0.001 mm  5 mm) that encodes the distance to the corresponding mesh with lower resolution. Areas where the distance is smaller than 0.001 mm are white. To put this into context, the distance between the upper left and upper right corner were about 29 and 21 cm for the Italian and Greek mean, respectively.

the advantage that it does not need additional sundials for prediction, it requires high mesh resolutions for accurate computations. In contrast, the above results show that the newly presented method works well even at coarse resolutions. Moreover, being purely data-driven, it requires no assumptions on the interplay between celestial mechanics and sundial shape. On the other hand, since both approaches make use of different, non-overlapping features of the shadow surface's make-up, they can complement each other and even be applied together to utilize more of the available information.

A limiting factor of this study is the small number of samples that were available. Convincing statistical significance can only be achieved with more samples. This is a common situation in archaeology. While more ancient (digitized) sundials are available, they are too damaged to be used with current methods. New methods for statistical analysis of partial forms are needed to alleviate this problem; these could allow a greater number of existing ancient sundials to be used.

More data would also help to decide an interesting question raised by the proposed normalization method: whether craftsmen from Greece and the Italian peninsula, when building sundials, made different adaptations in order to account for changes in the angle of the sun. If backed up, the results of our experiment could indicate not only that the mean shape of a shadow surface depends on the latitude of the installation site in both regions, but also that the strength of dependence *varies among the groups*.

It could be, for example, that the Greek craftsmen mainly adjusted the hour lines, while their Roman counterparts also adjusted the shape of the shadow area. It seems very interesting to investigate this question further; however, it would require more data, especially from Greece.

Last but not least, we clearly showed that the proposed methods can be used even when only relatively poor digital representations are at hand.

5 CONCLUSION

Recent advances in image-based reconstruction (Structure from Motion and Multi-view Stereo) have made the acquisition of highly detailed 3D artifact models simple, fast and economical. The sundial models explored in the present study represent a rapidly growing volume of 3D research data. These data should be explored with mathematical approaches that take full advantage of the rich information they contain and avoid information losses such as those caused by *a priori* typological grouping. We demonstrated such a procedure combining shape space methods and statistics on manifolds, and showed that there is great potential to extract more information from groups of artifact shapes using such methods without prior assumptions. Conversely, methods such as those discussed here have the potential to generate new, less biased and less convoluted baselines for typological seriation, in which the ordering of data represents a better defined (e.g., geographical or temporal) trend. An obvious next step is thus to apply the proposed methods to other artifact collections and to explore their full potential in archaeology.

In this study, the proposed methods revealed a latitudinal dependency of the shape of shadow surfaces in sundials of the Italian peninsula. This sheds new light on the construction principles of sundials in ancient times. We were also able to show how this dependence can be used to accurately determine the latitude of a sundial's working location.

Shape spaces are a current research topic, including how best to deal with challenges that are particularly common in archaeological investigations, such as incompleteness and weathering of artifacts. The majority of the considered shape spaces exhibits a Riemannian structure and, hence, can be combined with the presented approach. Because of this flexibility, we expect the proposed methods will be of great importance in a wide range of application scenarios.

Finally, it is also worth noting that we introduced a new method that allows to control for confounding effects caused by parameter-dependent variations. This method has practical uses in archaeology, as we have demonstrated. However, as a generic method, it can also be applied to any other manifold-valued data. Thus, this work is a contribution not only to archaeology, but also to geometric statistics in general.

APPENDIX

A LIST OF SUNDIALS

The sundials used in the study are listed in Table 1 along with the latitude and longitude of their sites. They are part of the Topoi database. The triangle meshes of the whole sundials can be downloaded from <http://repository.edition-topoi.org/collection/BSDP>. The metadata can also be found there. The pre-processed meshes of the shadow surfaces are also available for download [30].

B SHAPE SPACE OF DIFFERENTIAL COORDINATES

In this section we recall the fundamental facts on the shape space of differential coordinates. The underlying idea is to model a shape as the derivative of the deformation that a common reference mesh has to undergo in order to coincide with the object. Being a derivative, this representation is invariant under translations and, when the objects are scaled to the same size and Procrustes-aligned [23] beforehand, effects of scale, rotation, and relative translation are also minimized. Conversely, to given differential coordinates a corresponding triangle mesh can be computed. This is used in many illustrations in this article.

Suppose that homogeneous objects are given in the form of triangle meshes $T_i \subset \mathbb{R}^3$ that are in correspondence, scaled to the same size and Procrustes-aligned [23]. Then, we can encode their shape in the space of differential coordinates as introduced in Ref. [56]. To this end, we view each mesh T_i as the result of a deformation ϕ_i of a common reference \bar{T} ; i.e., each $\phi_i : \bar{T} \rightarrow T_i$ is an orientation-preserving, simplicial isomorphism that yields a semantic correspondence. We compute \bar{T} in an iterative process as the mesh representation of the data's Fréchet mean as follows. To initialize, we choose one of the given geometries as reference, say T_1 , and encode the shapes of every T_i as explained below. Then, we compute their mean shape (see App. C) and use its mesh representation to update the reference. This procedure is repeated until convergence, or until the norm of the update does not decrease any further. The method has the advantage that it minimizes the bias that is introduced by a particular choice of reference.

From here on, everything works the same way for all meshes T_i . For this reason, we omit the index.

Let m be the number of faces of T (which does not depend on i because of the one-to-one correspondence) and denote the set of (real) 3-by-3 matrices with positive determinant by $\text{GL}^+(3)$. The Jacobian $D\phi$ of ϕ is a constant 3-by-3 matrix G_j with $\det(G_j) > 0$ on each face F_j of \bar{T} , i.e.,

$$D\phi|_{F_j} = G_j \in \text{GL}^+(3) \quad \text{for all } j = 1, \dots, m.$$

Hence, using the polar decomposition, we can find rotation matrices $R_j \in \text{SO}(3)$ and symmetric positive definite matrices $U_j \in \text{Sym}^+(3)$ such that

$$G_j = R_j U_j \quad \text{for all } j = 1, \dots, m.$$

Both have a physical interpretation since R_j and U_j can be viewed as the rotation and stretch, respectively, that F_j undergoes when \bar{T} is deformed by ϕ . We call

$$S := ((R_1, U_1), \dots, (R_m, U_m))$$

the *differential coordinates of T* . Thus, the *space of differential coordinates for triangle meshes with m faces* is

$$\Sigma_m := (\text{SO}(3) \times \text{Sym}^+(3))^m.$$

This is a manifold that can be endowed with a Riemannian product metric using the canonical metric of $\text{SO}(3)$ and the Log-Euclidean metric [5] of $\text{Sym}^+(3)$. Using it, allows for efficient statistics in the space; see App. C for a very brief introduction to this. Moreover, given differential coordinates $S \in \Sigma_m$ and a reference \bar{T} , the triangular mesh corresponding to S can be found very efficiently as the solution of a sparse linear system of equations [56].

C A SHORT INTRODUCTION TO GEOMETRIC STATISTICS IN RIEMANNIAN MANIFOLDS

In the following, whenever we say “smooth” we mean “infinitely often differentiable”.

A (smooth) manifold M of dimension n is a space that *locally* looks like n -dimensional Euclidean space. In particular, at each point $p \in M$ the tangent space T_pM is an n -dimensional vector space; but in contrast to Euclidean space, the tangent spaces cannot simply be identified with each other (think of the 2-dimensional sphere in \mathbb{R}^3 whose tangent spaces are rotated against each other). Thus, keeping track of the base point p is important. For more on this including the exact definitions and examples we refer to Ref. [15, Ch. 0].

We call M *Riemannian manifold* when it is endowed with a *Riemannian metric*, i.e., a smooth map $M \ni p \mapsto \langle \cdot, \cdot \rangle_p$ that assigns a scalar product to each tangent space T_pM . It allows to measure the length of a smooth curve $\alpha : I \rightarrow M$, $I \subset \mathbb{R}$, by defining

$$\text{length}(\alpha) := \int_I \langle \alpha'(t), \alpha'(t) \rangle_{\alpha(t)} dt,$$

where $\alpha' := \frac{d}{dt}\alpha$ is the tangent vector (field) of α . This notion can be used to introduce a distance on M . To this end, let Γ be the set of all curves that start in $p \in M$ and end in $q \in M$. The distance between p and q is then given by

$$d(p, q) := \inf_{\alpha \in \Gamma} \text{length}(\alpha).$$

The infimum is necessary because of the possibility that there is no shortest curve connecting p to q (think of a punctured plane in \mathbb{R}^3 and p, q such that the missing point is on the straight line connecting them in the plane).

Every Riemannian manifold comes with a notion of a *covariant derivative* along curves. Given a vector field X along a curve α in M (i.e., a map that smoothly assigns to each point on α a vector in the associated tangent space), the covariant derivative allows to differentiate X along α ; we denote the resulting vector field along α by $\frac{DX}{dt}$.

The covariant derivative induces the notion of “straightness”. More precisely, *geodesics*—i.e., generalized straight lines—can be defined as curves $\gamma : [0, 1] \rightarrow M$ whose acceleration vanishes; that is, $\frac{D\gamma'}{dt}$ is the zero vector field. An important fact is that every point in M has a so-called convex neighborhood U . Each pair $p, q \in U$ can be connected by a unique, length-minimizing geodesic $[0, 1] \ni t \mapsto \gamma(t; p, q)$ that does not leave U . In particular, geodesics are also shortest paths in U . (The latter need not be true on a larger scale.) In this work, we always assume that the given data lies in such a convex neighborhood. Then, γ is also differentiable with respect to its start and end point. Explicit formulas of these differentials involve the Riemannian curvature tensor R (which intuitively measures local deviation from flat space; see [15, Ch. 4]).

The covariant derivative also allows to translate vectors along curves. This is important, e.g., when we want to compare vectors from different tangent spaces. For every pair $p, q \in U$ and $v \in T_pM$ there is a unique vector field X along the geodesic γ from p to q that solves the differential equation $\frac{DX}{dt} = 0$ with initial value $X(p) = v$. We call $X(q) \in T_qM$ the *parallel transport* of v to q along γ as the vector field does not undergo any intrinsic changes. (Extrinsically—i.e., viewed from outside of the manifold—there might be a change however; this can be seen, e.g., by considering the velocity vector of a *straight-driving* car from the viewpoint of an astronaut in the ISS.)

Another important map is the *Riemannian exponential*. Let $X \in T_p M$ such that there is a geodesic $[0, 1] \ni t \mapsto \gamma(t; p, q)$ in U with $X = \gamma'(0; p, q)$. The exponential map at p is then defined by $\exp_p(X) := q$. Its inverse is the *Riemannian logarithm* \log_p . In particular, we find $\log_p(q) = \gamma'(0; p, q)$ and

$$d(p, q) = \|\log_p(q)\| = \sqrt{\langle \log_p(q), \log_p(q) \rangle_p}.$$

Because of this and the fact that it is parallel to the shortest path from p to q (i.e., “points” to q), $\log_p(q)$ can be seen as the “difference vector” between q and p .

In the differential coordinate space, the exponential and logarithm map as well as parallel transport can be computed using explicit formulas: Because of the product structure, they can be calculated component-wise for every element of $\text{SO}(3)$ and $\text{Sym}^+(3)$; formulas for these manifolds are derived in Refs. [5, 18], respectively.

Amongst others, the tools gathered so far allow to generalize the notion of mean and linear regression to a Riemannian manifold M in the form of the Fréchet mean and geodesic regression, respectively. Given data points q_1, \dots, q_n in a

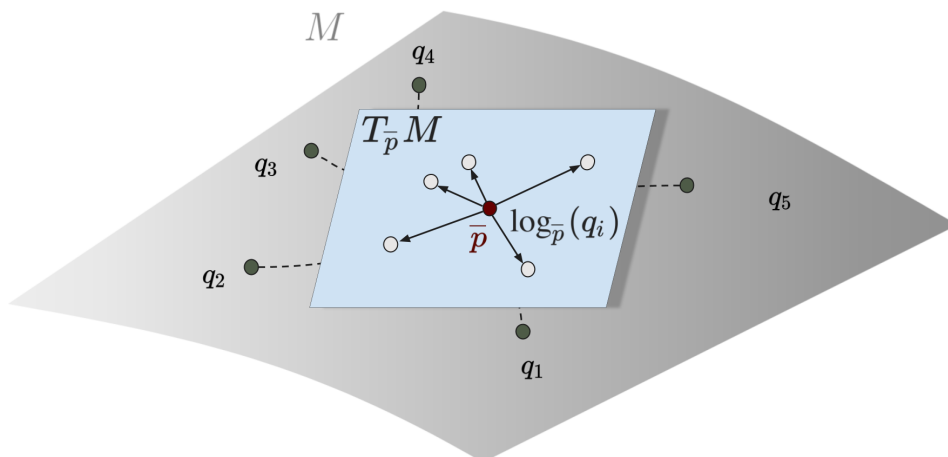


Fig. 21. Fréchet mean of q_1, \dots, q_5 in a Riemannian manifold M . After mapping the points into the tangent space $T_{\bar{p}} M$ at the Fréchet mean \bar{p} with $\log_{\bar{p}}$, the resulting tangent vectors $\log_{\bar{p}}(q_i)$, $i = 1, \dots, 5$ sum to zero. The broken lines are the geodesics from the mean to the data points.

convex neighborhood $U \subseteq M$, their *Fréchet mean* $\bar{p} \in U$ is defined implicitly by

$$\sum_{i=1}^n \log_{\bar{p}}(q_i) = 0.$$

That is, we look for a point $\bar{p} \in U$ in whose tangent space the usual optimality condition of the multivariate mean holds—namely, that the sum of difference vectors to the data points vanishes. This idea is visualized in Fig. 21. Importantly, Fréchet means can be computed efficiently using an iterative scheme. (For more on Fréchet means including a discussion of their existence, we refer the reader to Ref. [20, Sect. 2.2].)

Given n data points $(q_i, t_i) \in U \times \mathbb{R}$ together with parameter values, *geodesic regression* [19] assumes that the q_i are realizations of an M -valued random variable Q according to the model

$$Q(t) := \exp_{\gamma(t; p, q)}(\epsilon), \tag{1}$$

where ϵ is tangent-vector-valued noise and $p, q \in U$. The goal is to find the geodesic γ in (1). A good approximation in many cases is the *least-squares estimator*. Since geodesics in U can be parametrized by their start and endpoint, it is the minimizer $(\hat{p}, \hat{q}) \in U \times U$ of the *sum-of-squared error*

$$\mathcal{E}(p, q) := \sum_{i=1}^n d(\gamma(t_i; p, q), q_i)^2.$$

In the spaces that are relevant to us, (\hat{p}, \hat{q}) can be found efficiently using optimization methods. In particular, in the space of differential coordinates (which, as a product of symmetric spaces, is also a symmetric space [32]) an explicit formula for the intrinsic gradient of \mathcal{E} is available [7, 28], so that we can use Riemannian gradient descent [1] to compute the minimizer. A visualization of geodesic regression with respect to 4 data points is depicted in Fig. 22.

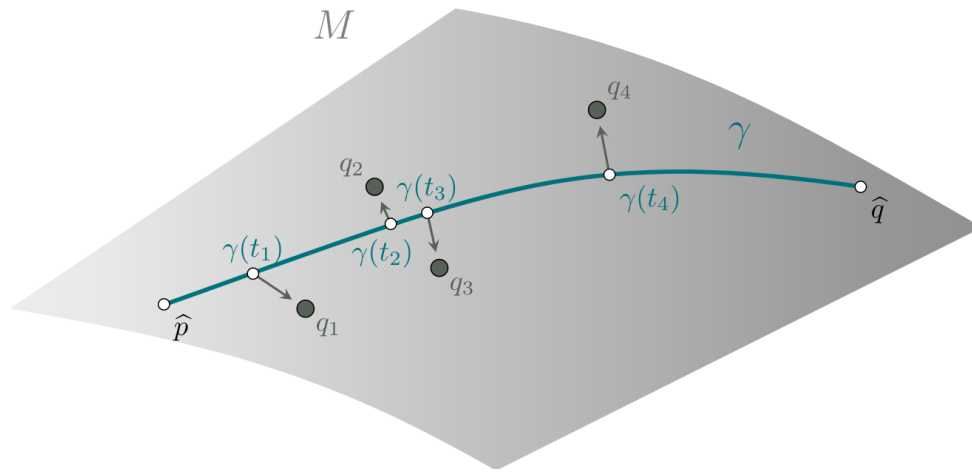


Fig. 22. Geodesic regression in (a part of) a Riemannian manifold M w.r.t. data $(q_1, t_1), \dots, (q_4, t_4)$. The geodesic γ is the least-squares estimator with start and end point \hat{p} and \hat{q} , respectively. The gray arrows are the vectors $\log_{\gamma(t_i)}(q_i) \in T_{\gamma(t_i)}M$.

ACKNOWLEDGMENTS

M. Hanik is funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy through MATH+, the Berlin Mathematics Research Center, under Grant No.: EXC-2046/1 – project ID: 390685689. This work was also supported by the Bundesministerium für Bildung und Forschung (BMBF) through BIFOLD, the Berlin Institute for the Foundations of Learning and Data under Grant Nos.: 1IS18025A and 01IS18037A.

Last but not least, we thank the anonymous reviewers for their insightful comments and helpful suggestions.

REFERENCES

- [1] P.-A. Absil, R. Mahony, and Rodolphe Sepulchre. 2007. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, NJ, USA. <https://doi.org/10.1515/9781400830244>
- [2] Felix Ambellan, Martin Hanik, and Christoph von Tycowicz. 2021. *Morphomatics: Geometric morphometrics in non-Euclidean shape spaces*. <https://doi.org/10.12752/8544> <https://morphomatics.github.io/>.

- [3] Felix Ambellan, Stefan Zachow, and Christoph von Tycowicz. 2019. A surface-theoretic approach for statistical shape modeling. In *Proc. Medical Image Computing and Computer Assisted Intervention, Part IV*. Springer, Berlin, 21–29. https://doi.org/10.1007/978-3-030-32251-9_3
- [4] Felix Ambellan, Stefan Zachow, and Christoph von Tycowicz. 2021. Geodesic B-score for improved assessment of knee osteoarthritis. In *Information Processing in Medical Imaging – 27th International Conference, IPMI 2021*. Springer, Berlin, 177–188. https://doi.org/10.1007/978-3-030-78191-0_14
- [5] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. 2007. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.* 29, 1 (2007), 328–347. <https://doi.org/10.1137/050637996>
- [6] Martin Bauer, Martins Bruveris, and Peter W. Michor. 2014. Overview of the geometries of shape spaces and diffeomorphism groups. *J. Math. Imaging Vis.* 50 (2014), 60–97. <https://doi.org/10.1007/s10851-013-0490-z>
- [7] Ronny Bergmann and Pierre-Yves Goussens. 2018. A variational model for data fitting on manifolds by minimizing the acceleration of a Bézier curve. *Front. Appl. Math. Stat.* 4 (2018), 1–16. <https://doi.org/10.3389/fams.2018.00059>
- [8] Mark Berman and David Culpin. 1986. The statistical behaviour of some least squares estimators of the centre and radius of a circle. *J. R. Stat. Soc. Series B Stat. Methodol.* 48, 2 (1986), 183–196. <https://doi.org/10.1111/j.2517-6161.1986.tb01401.x>
- [9] Silvia Biasotti, Bianca Falcidieno, Daniela Giorgi, and Michela Spagnuolo. 2014. Mathematical tools for shape analysis and description. *Synthesis Lectures on Computer Graphics and Animation* 6, 2 (2014), 1–138. <https://doi.org/10.1007/978-3-031-79558-9>
- [10] Alan D. Brett and Chris J. Taylor. 2000. Construction of 3D shape models of femoral articular cartilage using harmonic maps. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Berlin, 1205–1214. https://doi.org/10.1007/978-3-540-40899-4_129
- [11] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. 2008. *Numerical geometry of non-rigid shapes*. Springer, New York. <https://doi.org/10.1007/978-0-387-73301-2>
- [12] Paolo Cignoni, Marco Callieri, Massimiliano Corsini, Matteo Dellepiane, Fabio Ganovelli, and Guido Ranzuglia. 2008. MeshLab: an Open-Source Mesh Processing Tool. In *Eurographics Italian Chapter Conference*, Vittorio Scarano, Rosario De Chiara, and Ugo Erra (Eds.). The Eurographics Association, Salerno, Italy, 129–136. <https://doi.org/10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2008/129-136>
- [13] Hilary E. M. Cool and Mike J. Baxter. 1995. Finds from the fortress: artefacts, buildings and correspondence analysis. In *Computer Applications and Quantitative Methods in Archaeology 1993 (BAR International Series, 598)*. Archaeopress, Oxford, UK, 177–182. https://proceedings.caaconference.org/files/1993/25_Cool_Baxter_CAA_1993.pdf
- [14] Timothy F. Cootes, Chris J. Taylor, David H. Cooper, and Jim R. Graham. 1995. Active shape models-their training and application. *Comput. Vis. Image Underst.* 61, 1 (1995), 38–59. <https://doi.org/10.1006/cviu.1995.1004>
- [15] Manfredo P. do Carmo. 1992. *Riemannian Geometry* (2 ed.). Birkhäuser Boston, Cambridge, MA, USA.
- [16] Ian L. Dryden and Kanti V. Mardia. 2016. *Statistical Shape Analysis, with Applications in R* (2 ed.). John Wiley and Sons, Chichester, UK. <https://doi.org/10.1002/9781119072492>
- [17] Benjamin Duce. 2018. *The Encyclopedia of Archaeological Sciences*. American Cancer Society, Atlanta, Chapter Multi-View Stereo, 1–4. <https://doi.org/10.1002/9781119188230.saseas0398>
- [18] Alan Edelman, Tomás A. Arias, and Steven T. Smith. 1998. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.* 20, 2 (1998), 303–353. <https://doi.org/10.1137/S0895479895290954>
- [19] P. Thomas Fletcher. 2013. Geodesic regression and the theory of least squares on Riemannian manifolds. *Int. J. Comput. Vis.* 105, 2 (2013), 171–185. <https://doi.org/10.1007/s11263-012-0591-y>
- [20] P. Thomas Fletcher. 2020. 2 - Statistics on manifolds. In *Riemannian Geometric Statistics in Medical Image Analysis*. Academic Press, London, 39–74. <https://doi.org/10.1016/B978-0-12-814725-2.00009-1>
- [21] P. Thomas Fletcher, Conglin Lu, Stephen M. Pizer, and Sarang Joshi. 2004. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Trans. Med. Imaging* 23, 8 (2004), 995–1005. <https://doi.org/10.1109/TMI.2004.831793>
- [22] Bernhard Fritsch, Elisabeth Rinner, and Gerd Graßhoff. 2013. 3D Models of ancient sundials: a comparison. *Int. J. Herit. Digit. Era* 2, 3 (2013), 361–373. <https://doi.org/10.1260/2047-4970.2.3.361>
- [23] Colin Goodall. 1991. Procrustes methods in the statistical analysis of shape. *J. R. Stat. Soc. Series B Stat. Methodol.* 53, 2 (1991), 285–321. <https://doi.org/10.1111/j.2517-6161.1991.tb01825.x>
- [24] Gerd Graßhoff, Elisabeth Rinner, Karlheinz Schaldach, Bernhard Fritsch, Liba Taub, Jessica Sum, and Florian Kotschka. 2016. Ancient Sundials. <https://doi.org/10.17171/1-1-Edition-Topoi>
- [25] Susie Green. 2018. *The Encyclopedia of Archaeological Sciences*. American Cancer Society, Atlanta, Chapter Structure from motion, 1–3. <https://doi.org/10.1002/9781119188230.saseas0559>
- [26] Michael J. Greenacre. 1984. *Theory and Applications of Correspondence Analysis*. Academic Press, London. <https://doi.org/10.5771/0943-7444-1985-2-107-1>
- [27] Philipp Grohs, Martin Holler, and Andreas Weinmann (Eds.). 2020. *Handbook of Variational Methods for Nonlinear Geometric Data*. Springer, Berlin. <https://doi.org/10.1007/978-3-030-31351-7>
- [28] Martin Hanik, Hans-Christian Hege, Anja Hennemuth, and Christoph von Tycowicz. 2020. Nonlinear regression on manifolds for shape analysis using intrinsic Bézier splines. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*. Springer, Berlin, 617–626. https://doi.org/10.1007/978-3-030-59719-1_60
- [29] Martin Hanik, Hans-Christian Hege, and Christoph von Tycowicz. 2020. Bi-invariant two-sample tests in Lie groups for shape analysis. In *Shape in Medical Imaging*. Springer, Berlin, 44–54. https://doi.org/10.1007/978-3-030-61056-2_4

- [30] Martin Hanik and Christoph von Tycowicz. 2022. Triangle meshes of shadow-receiving surfaces of ancient sundials. <https://doi.org/10.12752/8425>
- [31] Behrend Heeren, Martin Rumpf, Peter Schröder, Max Wardetzky, and Benedikt Wirth. 2014. Exploring the geometry of the space of shells. *Comput. Graph. Forum* 33, 5 (2014), 247–256. <https://doi.org/10.1111/cgf.12450>
- [32] Sigurdur Helgason. 2001. *Differential Geometry, Lie Groups, and Symmetric Spaces*. Graduate Studies in Mathematics, Vol. 34. American Mathematical Society, Providence, RI, USA. <https://doi.org/10.1090/gsm/034>
- [33] Stephan F. Huckemann and Benjamin Eltzner. 2020. *Handbook of Variational Methods for Nonlinear Geometric Data*. Springer, Berlin, Chapter Statistical methods generalizing principal component analysis to non-Euclidean spaces, 317–338. https://doi.org/10.1007/978-3-030-31351-7_10
- [34] Stephan F. Huckemann, Thomas Hotz, and Axel Munk. 2010. Intrinsic shape analysis: geodesic PCA for Riemannian manifolds modulo isometric Lie group actions. *Stat. Sin.* 20, 1 (2010), 1–58. <https://www.jstor.org/stable/24308976>
- [35] Manfred Hüttig. 2000. The conical sundial from Thyrreion—reconstruction and error analysis of a displaced antique sundial. *Arch. Hist. Exact Sci.* 55, 2 (2000), 163–176. <https://doi.org/10.1007/s004070000023>
- [36] Alexander Jones. 2019. *Down to the Hour: Short Time in the Ancient Mediterranean and Near East*. Brill, Leiden, Netherlands, Chapter Greco-Roman sundials: precision and displacement, 125–157. https://doi.org/10.1163/9789004416291_006
- [37] David G. Kendall, Dennis Barden, Thomas K. Carne, and Huiling Le. 2009. *Shape and Shape Theory*. Wiley, Hoboken, NJ, USA. <https://doi.org/10.1002/9780470317006>
- [38] R. R. Laxton and J. Restorick. 1989. Seriation by similarity and consistency. In *CAA 89. Computer Applications and Quantitative Methods in Archaeology*. Number 48 in BAR International Series. Oxbow Books, Oxford, UK, 215–225. https://proceedings.caaconference.org/files/1989/21_Laxton_Restorick_CAA_1989.pdf
- [39] Torsten Madsen. 1989. Seriation and multivariate statistics. In *CAA 89. Computer Applications and Quantitative Methods in Archaeology*. Number 48 in BAR International Series. Oxbow Books, Oxford, UK, 205–214. https://proceedings.caaconference.org/files/1989/20_Madsen_CAA_1989.pdf
- [40] Oscar Montelius. 1903. *Die typologische Methode*. Self-published by the author, Stockholm.
- [41] Prasanna Muralidharan and P. Thomas Fletcher. 2012. Sasaki metrics for analysis of longitudinal data on manifolds. In *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, Piscataway, NJ, USA, 1027–1034. <https://doi.org/10.1109/CVPR.2012.6247780>
- [42] Michael J. O’Brien and R. Lee Lyman. 1998. *Seriation, Stratigraphy, and Index Fossils. The Backbone of Archaeological Dating*. Springer, Berlin. <https://doi.org/10.1007/b110115>
- [43] Maks Ovsjanikov, Etienne Corman, Michael Bronstein, Emanuele Rodolà, Mirela Ben-Chen, Leonidas Guibas, Frederic Chazal, and Alex Bronstein. 2016. Computing and processing correspondences with functional maps. In *SIGGRAPH ASIA 2016 Courses*. ACM, New York, NY, USA, 1–60. <https://doi.org/10.1145/3084873.3084877>
- [44] Xavier Pennec. 2006. Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements. *J. Math. Imaging Vision* 25 (2006), 127–154. <https://doi.org/10.1007/s10851-006-6228-4>
- [45] Xavier Pennec, Stefan Sommer, and P. Thomas Fletcher (Eds.). 2020. *Riemannian Geometric Statistics in Medical Image Analysis*. Academic Press, London. <https://doi.org/10.1016/C2017-0-01561-6>
- [46] Mara Pistellato, Arianna Traviglia, and Filippo Bergamasco. 2020. Geolocating time: digitisation and reverse engineering of a roman sundial. In *European Conference on Computer Vision*. Springer, Berlin, 143–158. https://doi.org/10.1007/978-3-030-66096-3_11
- [47] Stephen M. Pizer, Junpyo Hong, Jared Vicory, Zhiyuan Liu, J. S. Marron, Hyo young Choi, James Damon, Sungkyu Jung, Beatriz Paniagua, Jörn Schulz, Ankur Sharma, Liyun Tu, and Jiyao Wang. 2020. 6 - Object shape representation via skeletal models (s-reps) and statistical analysis. In *Riemannian Geometric Statistics in Medical Image Analysis*. Academic Press, London, 233–271. <https://doi.org/10.1016/B978-0-12-814725-2.00014-5>
- [48] Colin Renfrew and Paul G. Bahn. 2019. *Archaeology: Theories, Methods and Practice* (8 ed.). Thames and Hudson, London.
- [49] Elisabeth Rinner, Bernhard Fritsch, and Gerd Graßhoff. 2013. Die unvollendete Sonnenuhr von der Agora der Italiker auf Delos. *eTopoi. J. Anc. Studies* 2 (2013), 111–130. <https://doi.org/10.17171/4-2-4>
- [50] Karlheinz Schaldach. 2016. *Time and Cosmos in Greco-Roman Antiquity*. Princeton University Press, Princeton, NJ, USA, Chapter Measuring the hours: sundials, water clocks, and portable sundials, 63–93. https://doi.org/10.1163/9789004416291_002
- [51] Andy Scott. 1992. A parametric approach to seriation. In *Computing the Past, CAA 92. Computer Applications and Quantitative Methods in Archaeology*. Aarhus University Press, Aarhus, Denmark, 317–324. https://proceedings.caaconference.org/files/1992/34_Scott_CAA_1992.pdf
- [52] Kaleem Siddiqi and Stephen M. Pizer (Eds.). 2008. *Medial representations: mathematics, algorithms and applications*. Springer, Berlin. <https://doi.org/10.1007/978-1-4020-8658-8>
- [53] Anuj Srivastava and Eric P. Klassen. 2016. *Functional and Shape Data Analysis*. Springer, Berlin. <https://doi.org/10.1007/978-1-4939-4020-2>
- [54] Detlev Stalling, Malte Westerhoff, and Hans-Christian Hege. 2005. Amira: a highly interactive system for visual data analysis. In *The Visualization Handbook*. Academic Press, London, Chapter 38, 749–767. <https://doi.org/10.1016/B978-012387582-2/50040-X>
- [55] Gary KL Tam, Zhi-Quan Cheng, Yu-Kun Lai, Frank C Langbein, Yonghuai Liu, David Marshall, Ralph R Martin, Xian-Fang Sun, and Paul L Rosin. 2012. Registration of 3D point clouds and meshes: A survey from rigid to nonrigid. *IEEE Trans. Vis. Comput. Graph.* 19, 7 (2012), 1199–1217. <https://doi.org/10.1109/TVCG.2012.310>
- [56] Christoph von Tycowicz, Felix Ambellan, Anirban Mukhopadhyay, and Stefan Zachow. 2018. An efficient Riemannian statistical shape model using differential coordinates. *Med. Image Anal.* 43 (2018), 1–9. <https://doi.org/10.1016/j.media.2017.09.004>
- [57] Jacob A. Wegelin. 2000. *A Survey of Partial Least Squares (PLS) Methods, with Emphasis on the Two-Block Case*. Technical Report 371. University of Washington, Department of Statistics. <https://stat.uw.edu/sites/default/files/files/reports/2000/tr371.pdf>

- [58] Laurent Younes. 2019. *Shapes and Diffeomorphisms*. Springer, Berlin. <https://doi.org/10.1007/978-3-662-58496-5>

Table 1. Spherical sundials used in the study

Object ID	Dialface ID	Latitude	Longitude	Site	Dating	Type (Edition Topoi)
Italy						
574	623	43.3155	13.4082	Helvia Recina	N/A	sphere, spherical
17	17	42.0913	12.5231	Riano	N/A	sphere, spherical-cut dials—central gnomon point
36	35	41.8034	12.6890	Villa Tuscolana/Villa Rufinella ⁵ (near Rome)	N/A	sphere, spherical-cut dials—central gnomon point
65	62	41.7561	12.2927	Ostia Antica (near Rome)	N/A	sphere, spherical-cut/ quarter dis-shaped
18	18	41.6700	12.6900	Lanuvio	N/A	sphere, spherical-cut dials—central gnomon point
21	21	40.7503	14.4871	Pompei	before 79 CE	sphere, spherical-cut dials—central gnomon point
23	23	40.7503	14.4871	Pompei	before 79 CE	sphere, spherical-cut dials—central gnomon point
174	173	40.7503	14.4871	Pompei	N/A	sphere, spherical-hemispherical —central gnomon point
29	29	40.7503	14.4871	Pompei	before 79 CE	sphere, spherical-cut dials—central gnomon point
519	559	40.7030	14.4988	Stabiae	N/A	sphere, spherical
40	39	N/A	N/A	N/A ⁶	N/A	sphere, spherical-cut dials—central gnomon point
Greece						
76	73	37.3900	25.2600	Delos	N/A	sphere, spherical- hemispherical—central gnomon point
77	74	37.3900	25.2600	Delos	N/A	sphere, spherical-transposed hemispherical
546	583	36.0917	28.0881	Lindos on Rhodes ⁷	hellenistic	sphere, spherical

⁵This site is not 100% certain; the only reference comes from a dissertation written in 1764.

⁶Although in the data base the site is stated to be Vatican City, we did not find evidence for this. Therefore, we consider the site to be uncertain.

⁷The longitude and latitude of this sundial are wrong in the database. The references stated there confirm that the sundial is from Lindos. We have used the correct values, specified in the table.