

Article

Natural Earth Pigments as Geological Polymineal Systems: A Structured Approach to Provenance and Degradation Assessment in Heritage Science

Sebastiano Ettore Spoto ^{1,*}  and Roberta Somma ^{2,*} ¹ Dipartimento di Scienze della Terra, Università degli Studi di Firenze, 50121 Firenze, Italy² Dipartimento di Scienze Matematiche e Informatiche, Scienze Fisiche e Scienze della Terra, Università degli Studi di Messina, 98166 Messina, Italy

* Correspondence: sebastianoettore.spoto@unifi.it (S.E.S.); roberta.somma@unime.it (R.S.)

Abstract

Natural earth pigments are polymineal materials whose present composition may differ from the source because of selection, grinding, heating, mixing, binder incorporation, and later alteration. This study develops a structured method for separating these contributions and for stating which provenance and degradation claims are supported by available evidence. The material is represented by geological, technological, and degradation components linked to elemental, mineralogical, spectroscopic, microscopic, and colour observations. Provenance is expressed as probabilistic compatibility across source levels, whereas degradation assessment ranks classes of environmental exposure rather than reconstructing a unique history. Published data from Roussillon yellow ochres and commercial and historical green earths are re-analysed as proof-of-concept checks, and a literature-based Herculaneum application illustrates degradation reasoning. The ochre analysis shows that mineralogical and colour relations remain directionally stable under robustness checks and that wet sieving can shift measured composition without eliminating family-level compatibility. The green-earth analysis shows that commercial labels and bulk chemistry alone do not reliably establish celadonite–glauconite family membership, whereas structural and vibrational evidence provides stronger discrimination. The Herculaneum case demonstrates how thermal events and subsequent exposure can alter ochre states. The study provides a practical diagnostic sequence and defines the present limits of inference.

Keywords: natural earth pigments; yellow ochres; green earths; applied mineralogy; provenance; compositional data analysis; spectral unmixing; degradation assessment; probabilistic provenance; heritage diagnosis; inverse problems; heritage science



Academic Editor: Silvano Mignardi

Received: 14 May 2026

Revised: 5 August 2026

Accepted: 14 August 2026

Published: 17 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Natural earth pigments have a long history of artistic and technological use, but their familiar names do not identify fixed mineralogical materials. Standard reference works on artists' pigments and historical pigment terminology document the wide variation hidden within names such as ochre, sienna, umber, and green earth [1,2]. Recent heritage-science research has expanded this historical knowledge through spectroscopy, diffraction, microscopy, imaging, and geochemical analysis [3–7]. The resulting evidence shows that natural earth pigments are best understood as polymineal materials whose

properties depend on source geology, mineral proportions, particle size, preparation, and conservation history.

The analytical problem is not limited to identifying a colour-producing phase. An ochre may contain goethite, hematite, clay minerals, quartz, carbonates, manganese-bearing phases, and minor minerals in proportions that vary within and among deposits [4,5,8,9]. Green earths present a related problem because celadonite- and glauconite-bearing materials can overlap in bulk chemistry while retaining different structural and geological characteristics [10–12]. Colour is likewise a system property: phase proportions, crystallinity, particle-size distribution, refractive-index contrast, redox state, and binder or surface effects all influence the measured response [13,14].

Processing and subsequent alteration further separate the observed pigment from the geological raw material. Selection, levigation, grinding, sieving, heating, and mixing can change the phase proportions, particle-size distribution, and optical properties before application. After application, temperature, relative humidity, light, pollutants, salts, liquid-water chemistry, and wet–dry cycling may modify the material again [5,15–18]. Consequently, a present-day analysis records a combined geological, technological, and degradation history.

Heritage scientists already consider many of these effects when interpreting pigments. The contribution of this study is not the claim that processing or alteration has previously been ignored. Its purpose is to make the relationships explicit: to separate source, processing, degradation, and observation within one model; to state the assumptions needed for provenance claims; and to distinguish evidence of compatibility from evidence of unique attribution. This formal separation is intended to reduce deterministic interpretations when different material histories can produce similar measurements.

The study addresses two practical questions. First, which geological source classes remain compatible with the observed mineralogical and chemical evidence after processing and degradation are considered? Second, which broad environmental exposure regimes are compatible with the alteration markers observed today? In this article, environmental conditions means measurable or categorizable external conditions capable of modifying the pigment system, including temperature, relative humidity, light or ultraviolet dose, the pH of aqueous films or pore solutions, pollutant exposure, salt availability, and wet–dry cycling. An environmental exposure history is the time-dependent combination of those conditions, as reconstructed from available monitoring, documentary, stratigraphic, or material evidence [19]. The method does not reconstruct a unique day-by-day history; it ranks classes of exposure regimes that are consistent with the available evidence.

The article develops a material-state model, a probabilistic provenance module, and a mineral-family-specific degradation module. It then applies selected parts of the model to published data for Roussillon yellow ochres and green-earth pigments. The ochre re-analysis tests compositional treatment, phase–colour relations, and process-induced displacement. The green-earth re-analysis tests the difference between commercial or chemistry-led labels and structural mineral-family identification. A third, literature-based application to the House of the Bicentenary at Herculaneum shows how the degradation component maps a documented thermal conversion of yellow ochre to red. These applications identify both the practical value and the present limits of the proposed approach.

2. Materials and Methods: Scope and Analytical Model

2.1. Purpose and Scope

This study combines applied mineralogy, compositional statistics, colour physics, and inverse reasoning to support diagnostic interpretation. The full material state of a pigment cannot be measured directly; only partial observations are available from elemental, struc-

tural, spectroscopic, microscopic, and optical techniques. In this limited sense, the material state is latent: it is inferred from measurements rather than observed in its entirety. The model is designed to show which parts of that state can be constrained by a given analytical dataset and which remain uncertain.

The term natural earth pigments is used here for naturally occurring polymineral earth materials such as ochres, siennas, umbers, and green earths. Manufactured iron-oxide artist pigments historically marketed as Mars pigments are mentioned only to distinguish them from naturally occurring iron-rich earths and are excluded from the scope. Other synthetic and non-earth mineral pigments are also excluded unless explicitly discussed. This restriction is necessary because geological inheritance is a primary explanatory variable in the materials considered here.

Published studies provide the evidence base, and two compact re-analyses provide reproducible checks on selected parts of the model. The work does not present a new laboratory campaign or a fully calibrated prediction system. It provides a structured method for connecting geological source, processing, degradation, observation, and uncertainty. Table 1 summarizes the principal notation; every symbol is also defined where it first enters the argument.

Table 1. Core notation used throughout the article. The table gives the principal symbols needed to follow the argument; detailed definitions, units, and assumptions are provided locally at first use in the relevant sections.

Symbol	Meaning	Unit, Scale, or Domain
$\mathbf{S} = (\mathbf{G}, \mathbf{T}, \mathbf{D})$	Latent material state composed of geological, technological, and degradation components. The geological component records geological inheritance from the source deposit.	Abstract product-state space; components may be categorical, continuous, ordinal, normalized, or encoded according to the application.
$L_{\text{geo}}, \boldsymbol{\phi}, \mathbf{e}_{\text{geo}}, \mathbf{m}_{\text{acc}}, d_0$	Geological descriptors: lithology/deposit class, phase-fraction vector, geochemical signature, accessory phases, and initial particle-size distribution.	L_{geo} categorical; $\phi_i \geq 0$ and $\sum_i \phi_i = 1$; chemistry in wt%, ppm, or log-ratio coordinates; d_0 commonly in μm or mm.
$\lambda_{\text{lev}}, g_{\text{gr}}, \theta_{\text{heat}}, \tau_{\text{heat}}, \mathbf{M}_{\text{mix}}, b_{\text{int}}$	Technological descriptors: levigation, grinding, heating, mixing, and binder/interfacial state.	Normalized, ordinal, or measured variables; θ_{heat} in $^{\circ}\text{C}$ or K, τ_{heat} in time units, and $\mathbf{M}_{\text{mix}}$ as a closed mixing vector.
$\omega, \xi, \rho, s_D, \beta$	Degradation descriptors: hydration, oxidation/dehydroxylation, structural order, salt/weathering load, and pigment–binder–environment interaction.	Normalized, ordinal, categorical, or measured descriptors; s_D is distinct from optical scattering terms s_i and environmental salt forcing $S_{\text{salt}}(t)$.
$\mathbf{S}_{\text{raw}}, \mathbf{S}_{\text{app}}, \mathbf{S}_{\text{obs}}(t); \mathcal{P}_{\text{lev}}, \mathcal{P}_{\text{gr}}, \mathcal{P}_{\text{cal}}, \mathcal{P}_{\text{mix}}, \mathcal{P}_{\text{bind}}; \mathcal{A}_t$	Raw, applied, and observed material states; direct process operators; time-dependent degradation operator.	Material-state maps; t is the chosen exposure-time coordinate, expressed in physical or stratigraphic units as appropriate.
$\mathbf{O}, O_j, Q_j, \sigma_j, \delta_j, \Omega_j, \varepsilon_j$	Encoded observation vector and technique-specific observation model.	Heterogeneous analytical observables; σ_j and δ_j are lateral scale and probing depth, and Ω_j is sampled support.
$R_{\lambda}, F(R_{\lambda}), K_{\lambda}, S_{\lambda}, k_i, s_i, d_i, n_i$	Diffuse-reflectance and Kubelka–Munk optical quantities.	R_{λ} and $F(R_{\lambda})$ dimensionless; K_{λ} and S_{λ} have matching effective dimensions; d_i is length and n_i is dimensionless refractive-index contrast.
$\mathbf{x}, \text{clr}(\mathbf{x}), \mathbf{z} = \text{ilr}(\mathbf{x}), \mathbf{H}$	Closed composition, centred log-ratio coordinates, isometric log-ratio coordinates, and optional ilr basis matrix.	$\mathbf{x}$ is strictly positive and closed; $\mathbf{z} \in \mathbb{R}^{q-1}$; $\mathbf{H} \in \mathbb{R}^{(q-1) \times q}$ when used.
$Y_{p,\lambda}, \psi_k(\lambda), a_{pk}, K_{\text{end}}$	Spectral value, endmember spectrum, relative abundance, and number of endmembers.	Defined in the adopted spectral space; $a_{pk} \geq 0$ and typically $\sum_k a_{pk} = 1$; K_{end} is a positive integer.
$S_{\text{src}}, S_f = \pi(S_{\text{src}}), S_{\text{FY}}, U, \Delta, \Pi_f(S_f)$	Source hierarchy, source-family projection, French yellow-ochre demonstrator family, technological intervention, degradation state, and family-level posterior support.	Discrete or hierarchical source variables; U and Δ may be categorical, ordinal, or encoded; Π_f is normalized over source-family classes.

Table 1. Cont.

Symbol	Meaning	Unit, Scale, or Domain
$E(t), \mathbf{y}_r, \mathbf{Z}, h, \eta$	Environmental conditions, reduced diagnostic degradation state, alteration-marker vector, marker-generation operator, and residual.	$E(t)$ includes temperature, RH, pH, UV, gases, salt loading, and cycling; $\mathbf{y}_r$ and $\mathbf{Z}$ are encoded in mineral-family-specific coordinates.
$K_{\text{kao}}, G_{\text{goe}}, H_{\text{hem}};$ $C_{\text{label}}, N_{\text{family}}, N_{\text{label}}$	Roussillon compositional parts and green-earth label-concordance quantities.	Positive three-part ochre composition; C_{label} is dimensionless with $0 \leq C_{\text{label}} \leq 1$; N terms are integer counts.
$\mathbf{z}_{\text{obs}}, \mathbf{z}_{\text{src}}, \delta^*;$ $\mathbf{O}^{(G)}, \mathbf{O}^{(T)}, \mathbf{O}^{(D)}, \mathbf{I}_{GT}, \mathbf{I}_{GD}, \mathbf{I}_{TD}$	ilr coordinates and first-order perturbation terms; encoded geological, technological, degradation, and interaction signal components.	Vectors in the same ilr or encoded observable spaces; perturbations and interaction terms are interpretive first-order quantities, not directly separable physical observables.

Abbreviations: clr, centred log-ratio; FY, French yellow; ilr, isometric log-ratio; ppm, parts per million; RH, relative humidity; UV, ultraviolet; wt%, weight percent.

2.2. Direct and Inverse Questions

The analysis contains one forward question and two inverse questions. The forward question asks how a geological raw material changes after preparation, application, and environmental exposure. The first inverse question asks which source classes remain compatible with the observed pigment. The second asks which broad exposure regimes are capable of producing the alteration markers measured today. The same material system is thus followed forward when formation and alteration are described, and read backward when origin or exposure is inferred. The second inverse question does not assume that a unique environmental exposure history can be recovered; it compares the observed evidence with classes of plausible exposure conditions.

2.3. Material State and Processing History

A natural earth pigment is formalized as a material state

$$\mathbf{S} = (\mathbf{G}, \mathbf{T}, \mathbf{D}), \quad (1)$$

where $\mathbf{G}$ is the geological component, $\mathbf{T}$ is the technological component, and $\mathbf{D}$ is the degradation component. The term component is used here for consistency with the state-vector notation. Conceptually, however, $\mathbf{G}$ represents the geological inheritance of the pigment system: the deposit-related constraints carried into the cultural material before technological processing. These constraints include lithological setting, mineral phase assemblage, accessory minerals, geochemical signature, and inherited particle-size distribution.

The geological component is defined as

$$\mathbf{G} = \{L_{\text{geo}}, \boldsymbol{\phi}, \mathbf{e}_{\text{geo}}, \mathbf{m}_{\text{acc}}, d_0\}, \quad (2)$$

where L_{geo} is lithological setting or deposit type; $\boldsymbol{\phi} = (\phi_1, \phi_2, \dots, \phi_n)$ is the mineral phase-fraction vector; n is the number of mineral phases considered in the assemblage; $\mathbf{e}_{\text{geo}}$ is the geochemical signature; $\mathbf{m}_{\text{acc}}$ is the set of accessory minerals; and d_0 is the original particle-size distribution. In this sense, $\mathbf{G}$ should not be read as a static provenance label, but as the geological part of the present material state: it records what the pigment system carries forward from its depositional and lithological history. The phase-fraction vector $\boldsymbol{\phi}$ just defined in Equation (2) is constrained by the closure condition

$$\sum_{i=1}^n \phi_i = 1, \quad (3)$$

where each ϕ_i is the non-negative relative fraction of phase i within the selected assemblage. For log-ratio operations later in the paper, the relevant compositional parts are closed and strictly positive; zeros or below-detection-limit entries in other datasets require treatment before applying the transforms. Equation (3) is part of the definition of ϕ , not a separate geological descriptor.

The technological component is defined as

$$\mathbf{T} = \{\lambda_{\text{lev}}, g_{\text{gr}}, \theta_{\text{heat}}, \tau_{\text{heat}}, \mathbf{M}_{\text{mix}}, b_{\text{int}}\}, \quad (4)$$

where λ_{lev} is selection or levigation intensity, g_{gr} is the grinding state, θ_{heat} is treatment temperature, τ_{heat} is treatment duration, $\mathbf{M}_{\text{mix}}$ is the mixing vector with extenders or other pigments, and b_{int} denotes the binder or interfacial environment. These entries may be continuous, ordinal, or categorical depending on the available information; the notation records their role in the material history without requiring all of them to be numerically calibrated in the present study.

The degradation component is defined as

$$\mathbf{D} = \{\omega, \xi, \rho, s_D, \beta\}, \quad (5)$$

where ω is hydration state, ξ is oxidation or dehydroxylation state, ρ is crystallinity or structural order, s_D is the material-state descriptor for salt or environmental weathering load, and β is pigment–binder–environment interaction. The subscript in s_D is used to distinguish this degradation-state variable from the phase-specific scattering functions s_i in Equation (13) and from the time-dependent salt forcing $S_{\text{salt}}(t)$ in Equation (19).

The direct formation pathway is written from left to right as a sequence of material-state transformations. Process operators are denoted by action labels to avoid confusion with the state variables in Equations (1)–(5):

$$\begin{aligned} \mathbf{S}_{\text{raw}} &\xrightarrow{\mathcal{P}_{\text{lev}}} \mathbf{S}_{\text{lev}} \xrightarrow{\mathcal{P}_{\text{gr}}} \mathbf{S}_{\text{gr}} \xrightarrow{\mathcal{P}_{\text{cal}}} \mathbf{S}_{\text{cal}} \\ &\xrightarrow{\mathcal{P}_{\text{mix}}} \mathbf{S}_{\text{mix}} \xrightarrow{\mathcal{P}_{\text{bind}}} \mathbf{S}_{\text{app}}. \end{aligned} \quad (6)$$

Here $\mathbf{S}_{\text{raw}}$ is the raw geological material state; $\mathcal{P}_{\text{lev}}$, $\mathcal{P}_{\text{gr}}$, $\mathcal{P}_{\text{cal}}$, $\mathcal{P}_{\text{mix}}$, and $\mathcal{P}_{\text{bind}}$ denote selection/levigation, grinding or size reduction, calcination or thermal treatment, mixing, and binder incorporation, respectively; $\mathbf{S}_{\text{lev}}$, $\mathbf{S}_{\text{gr}}$, $\mathbf{S}_{\text{cal}}$, and $\mathbf{S}_{\text{mix}}$ are intermediate material states; and $\mathbf{S}_{\text{app}}$ is the applied pigment state. The present observed state is then

$$\mathbf{S}_{\text{obs}}(t) = \mathcal{A}_t(\mathbf{S}_{\text{app}}), \quad (7)$$

where $\mathcal{A}_t$ is a time-dependent degradation operator and $\mathbf{S}_{\text{obs}}(t)$ is the observed state at time t . Equations (6) and (7) emphasize that the measured pigment is the outcome of coupled technological and post-application histories.

2.4. Analytical Observations and Colour Response

Analytical methods do not observe $\mathbf{S}$ directly; they provide partial projections of it. The observation model is written as

$$\mathbf{O} = \mathcal{Q}(\mathbf{S}) + \boldsymbol{\varepsilon}, \quad (8)$$

where $\mathbf{O}$ is a vector of encoded observables, $\mathcal{Q}$ is the observation operator determined by the analytical method, and $\boldsymbol{\varepsilon}$ is a residual vector in the same encoded observable space as $\mathbf{O}$. The entries of $\mathbf{O}$ are heterogeneous physical measurements; they are grouped as a feature

vector for inference rather than treated as quantities with a single common unit. Operationally, the observable vector may include outputs from X-ray fluorescence (XRF), portable X-ray fluorescence (pXRF), X-ray diffraction (XRD), Raman spectroscopy, Fourier-transform infrared spectroscopy (FTIR), scanning electron microscopy with energy-dispersive spectroscopy (SEM-EDS), hyperspectral imaging (HSI), and International Commission on Illumination (CIE) $L^*a^*b^*$ colorimetry:

$$\mathbf{O} = \{O_{\text{XRF}}, O_{\text{pXRF}}, O_{\text{XRD}}, O_{\text{Raman}}, O_{\text{FTIR}}, O_{\text{SEM-EDS}}, O_{\text{HSI}}, O_{\text{CIELAB}}\}. \quad (9)$$

Here O_{CIELAB} denotes the CIE $L^*a^*b^*$ colour-coordinate output, where L^* denotes lightness and a^* and b^* denote the red–green and yellow–blue chromatic axes, respectively. Unless otherwise stated, wavelength λ is reported in nm, diffraction angle 2θ in degrees, wavenumber $\tilde{\nu}$ in cm^{-1} , stratigraphic coordinate z or probing depth in μm , and elemental abundances in wt% or at% as specified by the analytical method. The pXRF entry is separated from laboratory XRF only to mark the different in situ geometry and sampling support; both belong to the same elemental-analysis family. Equation (9) nevertheless requires an important analytical qualification: these techniques interrogate different material volumes, depths, lateral scales, and sample supports. For that reason, the observation operator is not only method-dependent but also scale-dependent. A more explicit notation is

$$O_j = \mathcal{Q}_j(\mathbf{S}; \sigma_j, \delta_j, \Omega_j) + \varepsilon_j, \quad (10)$$

where j indexes the analytical technique, σ_j denotes characteristic lateral scale, δ_j effective probing depth, Ω_j the measurement support or sampled area/volume associated with technique j , and ε_j the technique-specific residual or measurement uncertainty. The scale and depth parameters are normally reported in μm or mm according to the method. In practice, portable XRF, Raman spot analysis, micro-FTIR, SEM-EDS, hyperspectral imaging, and colorimetry represent different sampling supports rather than commensurate views of one identical voxel. The present notation defines a common inferential language while preserving the distinction among measurement scales.

The visible colour of an earth pigment is treated as the response of a heterogeneous absorbing–scattering particulate medium. At the level of diffuse reflectance, the present treatment adopts the Kubelka–Munk relation

$$F(R_\lambda) = \frac{(1 - R_\lambda)^2}{2R_\lambda} = \frac{K_\lambda}{S_\lambda}, \quad (11)$$

where $F(R_\lambda)$ is the Kubelka–Munk remission function, R_λ is reflectance at wavelength λ , K_λ is the effective absorption coefficient, and S_λ is the effective scattering coefficient. In the Kubelka–Munk ratio, K_λ and S_λ have identical dimensions, commonly expressed as inverse length for homogeneous layers, whereas R_λ and $F(R_\lambda)$ are dimensionless. For the absorption term in a polymineral pigment, let $k_i(\lambda, \xi_i, \rho_i)$ be the phase-specific absorption contribution of phase i , where ξ_i and ρ_i are the redox/dehydroxylation and structural-order states of that phase. Let $k_{\text{int}}(\lambda, \beta)$ be an interfacial absorption contribution controlled by the pigment–binder–environment interaction parameter β . The terms k_i and k_{int} have the same effective dimensions as K_λ ; ξ_i , ρ_i , and β are state descriptors rather than directly comparable physical units. The absorption coefficient is then written as

$$K_\lambda = \sum_{i=1}^n \phi_i k_i(\lambda, \xi_i, \rho_i) + k_{\text{int}}(\lambda, \beta), \quad (12)$$

For the scattering term, let $s_i(\lambda, d_i, n_i)$ be the phase-specific scattering contribution of phase i , where d_i is its characteristic particle size, expressed as a length, and n_i represents the dimensionless refractive-index contrast relevant to that phase. Let $s_{\text{tex}}(\lambda, g_{\text{gr}})$ be the effective textural scattering contribution controlled by grinding, packing, surface roughness, and porosity. The terms s_i and s_{tex} have the same effective dimensions as S_λ . The scattering coefficient is then written as

$$S_\lambda = \sum_{i=1}^n \phi_i s_i(\lambda, d_i, n_i) + s_{\text{tex}}(\lambda, g_{\text{gr}}). \quad (13)$$

Accordingly, colour is treated as an emergent property of the full absorption–scattering balance of the system rather than as the direct manifestation of a single phase [13,14,20]. Equations (11)–(13) provide a shared descriptive vocabulary for that balance. Their interpretive value is highest when samples behave as optically thick, diffuse, weakly stratified media, and lower in thin films, layered paint systems, intimate nonlinear mixtures, binder-rich media, rough or porous surfaces, and situations where substrate reflectance contributes strongly to the measured signal. These equations are most reliable when treated as disciplined approximations whose physical validity is checked against stratigraphy, support, and sampling conditions.

2.5. Probabilistic Source Attribution

The provenance problem begins with the fact that geochemical measurements are compositional. For a composition $\mathbf{x} = (x_1, \dots, x_q)$ living in the simplex, the centred log-ratio (clr) transform is

$$\text{clr}(\mathbf{x})_i = \ln\left(\frac{x_i}{\text{gm}(\mathbf{x})}\right), \quad i = 1, \dots, q, \quad (14)$$

where q is the number of compositional parts and $\text{gm}(\mathbf{x}) = (\prod_{i=1}^q x_i)^{1/q}$ is the geometric mean. The notation gm is used here to avoid confusion with the grinding state g_{gr} in Equation (4). Because clr coordinates are singular, inferential work is better performed in orthonormal isometric log-ratio (ilr) space,

$$\mathbf{z} = \text{ilr}(\mathbf{x}) \in \mathbb{R}^{q-1}. \quad (15)$$

Here, $\mathbb{R}^{q-1}$ denotes the $q - 1$ dimensional real coordinate space generated by the ilr transform. Equivalently, the same transform can be expressed as $\mathbf{z} = \mathbf{H} \text{clr}(\mathbf{x})$, where $\mathbf{H} \in \mathbb{R}^{(q-1) \times q}$ is an orthonormal log-contrast basis whose rows sum to zero. All log-ratio transforms in this article are applied after closure of strictly positive compositions. In the two published-data checks, the parts entering the transforms are strictly positive; no zero replacement is applied. Applications involving structural zeros, rounded zeros, or below-detection-limit entries require an explicitly declared replacement or censoring strategy before clr or ilr transformation. The move to clr or ilr coordinates is not a technical refinement added at the end of the workflow; it is the condition for treating source chemistry in a way that respects the geometry of mixtures [21–27].

Because pigments are rarely pure, mineralogical and spectral observations are also treated as mixed signals. A generic unmixing relation is written as

$$Y_{p,\lambda} = \sum_{k=1}^{K_{\text{end}}} a_{pk} \psi_k(\lambda) + \varepsilon_{p,\lambda}, \quad a_{pk} \geq 0, \quad \sum_{k=1}^{K_{\text{end}}} a_{pk} = 1, \quad (16)$$

where $Y_{p,\lambda}$ is the observed spectral intensity or transformed spectral value at pixel or sample p and wavelength λ ; K_{end} is the number of spectral endmembers; $\psi_k(\lambda)$ is the spectrum of endmember k ; a_{pk} is the relative abundance of endmember k at pixel or sample p ; and $\varepsilon_{p,\lambda}$ is the spectral residual. The symbol ψ_k is used for endmember spectra to avoid confusion with the scattering terms s_i in Equation (13). Equation (16) is used here as a local or transformed mixing model. Strict linearity in raw reflectance space is a special case rather than a general rule for pigment mixtures. In cultural-heritage hyperspectral applications, the choice between reflectance and $-\log R$ hyperspaces, as well as the choice of mixing model, can strongly affect performance. Tensor-decomposition approaches further allow the extraction of endmembers, concentrations, and temporal fading behaviour from higher-order datasets [20,28–30].

Provenance is assessed across nested geological source levels while accounting for geochemical, mineralogical, technological, and degradation uncertainty. Let $S_{\text{src}} = (S_{\text{src}}^{(1)}, S_{\text{src}}^{(2)}, S_{\text{src}}^{(3)})$ denote the source hierarchy of province, deposit, and sub-facies; let S_f denote the family-level source or mineral-state information obtained as a coarse-grained projection of S_{src} , written explicitly as $S_f = \pi(S_{\text{src}})$ for a deterministic projection map π ; let U denote technological intervention; let $\mathbf{E}_{0:t_f}$ denote the environmental regime over the relevant interval; and let Δ denote degradation state. In the following equations, $\Pr(\cdot)$ denotes probability or posterior support, and proportionality signs leave normalizing constants implicit. The resulting probability model is

$$\Pr(S_{\text{src}}, U, \Delta, \mathbf{E}_{0:t_f} \mid \mathbf{O}) \propto \Pr(\mathbf{O} \mid S_{\text{src}}, U, \Delta) \Pr(\Delta \mid S_f, U, \mathbf{E}_{0:t_f}) \times \Pr(\mathbf{E}_{0:t_f}) \Pr(U \mid S_{\text{src}}) \Pr(S_{\text{src}}). \quad (17)$$

Equation (17) encodes the conditional-independence convention used in the directed acyclic graph: the direct effect of the environmental exposure history on the observation vector is mediated by the degradation state. Thus, within this factorization, $\mathbf{O} \perp\!\!\!\perp \mathbf{E}_{0:t_f} \mid (S_{\text{src}}, U, \Delta)$; applications requiring a direct environmental effect on the observation operator require explicit addition of the corresponding arrow and refactorization of the likelihood. When the environmental contribution is not carried explicitly at the provenance stage, the factor $\Pr(\Delta \mid S_f, U)$ is interpreted as the marginal obtained by integrating over plausible environmental regimes,

$$\Pr(\Delta \mid S_f, U) = \int \Pr(\Delta \mid S_f, U, \mathbf{E}_{0:t_f}) \Pr(\mathbf{E}_{0:t_f}) d\mathbf{E}_{0:t_f}, \quad (18)$$

rather than as dependence on technology alone. This formulation avoids one-to-one deterministic source assignment and defines the probabilistic structure of posterior support over source classes, weighted by geological plausibility, technological filtering, and degradation uncertainty. Consequently, provenance is treated as a posterior distribution over source support rather than as retrieval of a single label [31–34]. In the article, Equation (17) is used as a theoretical inferential structure: likelihood structure, priors, and calibration strategy require explicit specification for any operational implementation. The factorization also encodes a conditional-independence convention: conditional on S_{src} , U , and Δ , the observable vector $\mathbf{O}$ has no separate direct dependence on $\mathbf{E}_{0:t_f}$; environmental conditions enter the provenance model through its effect on Δ .

2.6. Degradation and Environmental-Regime Assessment

Degradation is described as a family of mineral-specific or mineral-family-specific state transitions driven by environmental conditions and filtered by technology. Let the environmental conditions vector be

$$\mathbf{E}(t) = \{\Theta(t), RH(t), pH(t), UV(t), O_2(t), SO_2(t), NO_x(t), S_{\text{salt}}(t), C_{\text{cyc}}(t)\}, \quad (19)$$

where t is time; $\Theta(t)$ is temperature, reported in °C or K; $RH(t)$ is relative humidity, reported in %; $pH(t)$ is dimensionless effective interfacial or pore-solution pH; $UV(t)$ is ultraviolet irradiance or cumulative dose, reported for example in W m^{-2} or J m^{-2} ; $O_2(t)$ is oxygen availability, expressed as oxygen concentration or partial pressure when measured directly, or as a categorical microenvironmental proxy when only qualitative exposure information is available; $SO_2(t)$ and $NO_x(t)$ are sulfur- and nitrogen-bearing pollutant forcings, reported as ppm, ppb, or $\mu\text{g m}^{-3}$ as appropriate; $S_{\text{salt}}(t)$ is salt load or solution availability, reported for example as wt% or mg cm^{-2} ; and $C_{\text{cyc}}(t)$ represents wet–dry or condensation–evaporation cycling, reported as number of cycles or cycles per unit time. Here, $pH(t)$ is understood as the effective pH of interfacial aqueous films, pore solutions, or condensed water layers when such liquid phases are present; it is not the pH of dry air. The notation $S_{\text{salt}}(t)$ and $C_{\text{cyc}}(t)$ is intentionally distinct from the material state $\mathbf{S}$ and from the scattering functions in Equation (13). For each mineral phase or family r , the state evolution is written as

$$\frac{d\mathbf{y}_r}{dt} = f_r(\mathbf{y}(t), \mathbf{E}(t), U; \boldsymbol{\theta}_r), \quad r = 1, \dots, K_{\text{fam}}, \quad (20)$$

where $\mathbf{y}_r$ is the reduced diagnostic state vector of mineral phase or family r , $\mathbf{y}(t)$ is the collection of all such reduced states at time t , K_{fam} is the number of mineral families considered in the degradation module, U summarizes technological preconditioning, f_r is the phase- or family-specific response function, and $\boldsymbol{\theta}_r$ is the corresponding vector of kinetic, thermodynamic, or empirical response parameters. For any calibrated or simulation-oriented use, Equation (20) requires an initial condition $\mathbf{y}_r(0) = \mathbf{y}_{r,0}$. This initial condition may be measured from an unexposed analogue, assigned from the pre-exposure state of the material, or represented by a prior distribution $\text{Pr}(\mathbf{y}_{r,0} | S_f, U)$ when the initial state is uncertain. The vector $\mathbf{y}_r$ is not intended as a complete thermodynamic state; it may include diagnostically useful quantities such as effective abundance, hydration index, structural-order proxy, spectral marker intensity, or colour coordinate. In operational models, each component is required either to retain its own physical unit or to be normalized before joint inference. Once the initial state is specified, the reduced state path may be written compactly as $\mathbf{y}_{0:t_f} = \Phi_{0:t_f}(\mathbf{y}_0, \mathbf{E}_{0:t_f}, U, \boldsymbol{\theta})$, where $\Phi_{0:t_f}$ denotes the solution operator induced by the family-specific response functions. The marker likelihood used below is induced by the composition of the marker-generation function with this state-evolution operator.

Present-day alteration markers are then written as

$$\mathbf{Z} = h(\mathbf{y}_{0:t_f}, U, S_f) + \boldsymbol{\eta}, \quad (21)$$

where $\mathbf{y}_{0:t_f}$ denotes the reduced diagnostic state path from initial time 0 to final observation time t_f , h is the marker-generation function, U is technological preconditioning, S_f is the family-level source or mineral-state information, and $\boldsymbol{\eta}$ is marker-level uncertainty in the same marker space as $\mathbf{Z}$. The marker vector may include secondary phases, spectral alteration bands, salt or crust indicators, colour shifts, roughness proxies, or other diagnostically interpretable degradation features.

The inverse problem is not to reconstruct an exact exposure trajectory but to estimate the posterior support for compatible environmental regimes. When phase-resolved evidence has already narrowed the source information to a mineral family, the conditioned inverse form is

$$\Pr(\mathbf{E}_{0:t_f}, U \mid \mathbf{Z}, S_f) \propto \Pr(\mathbf{Z} \mid \mathbf{E}_{0:t_f}, U, S_f) \Pr(\mathbf{E}_{0:t_f}) \Pr(U \mid S_f). \quad (22)$$

In Equation (22), $\Pr(\mathbf{Z} \mid \mathbf{E}_{0:t_f}, U, S_f)$ is the likelihood of observing the marker vector under a candidate environmental regime, technological state, and source family; $\Pr(\mathbf{E}_{0:t_f})$ is the prior support for environmental regimes; and $\Pr(U \mid S_f)$ is the family-conditioned prior support for technological intervention. The notation covers both discrete regime classes and continuous exposure trajectories; sums replace integrals when the regime space is discretized. When family-level support is already available from the provenance module, the preferred family-marginalized form is

$$\Pr(\mathbf{E}_{0:t_f}, U \mid \mathbf{Z}) = \sum_{S_f} \Pr(\mathbf{E}_{0:t_f}, U \mid \mathbf{Z}, S_f) \Pi_f(S_f), \quad (23)$$

where $\Pi_f(S_f)$ denotes family-level support supplied by the provenance module, typically the family marginal of the broader provenance posterior induced by Equation (17); the summation is written for discrete family classes, while an integral would be used for a continuous family space.

In practice, this inversion is ill posed. The inverse module is proposed as an identifiability-oriented analytical scheme: it targets probabilistic regimes—for example, humid-photochemical, acid-humid, dry-thermal, or salt-cycling conditions—rather than a uniquely identifiable day-by-day environmental exposure history. Equation (22) is intentionally structural rather than fully operational: it marks the structure of the inverse problem and the information that would be required for calibration, but it does not claim that such calibration has already been achieved here. That limitation follows from the amount of information preserved in altered materials and from the fact that different process histories can converge toward similar present-day markers [17,18,35–40].

The conditional structure implied by Equations (17) and (22) is summarized in the simple directed acyclic graph in Figure 1. The graph is not intended as a fully parameterized graphical model; rather, it makes the inferential assumptions visually explicit. The source hierarchy and its family-level reduction influence both the range of plausible technological interventions and the way later degradation can unfold; technology preconditions degradation; environmental regime drives degradation; and observables inherit partially separable contributions from source, process, and degradation, while alteration markers form the observed subset most directly used for regime-based inversion. Operationally, the intended reading is direct: source-family information may either be conditioned upon when phase-resolved evidence is available, or marginalized over when only partial observations are available.

2.7. Evidence Base and Application Cases

The analysis is organized around five recurrent evidence domains in the literature: (i) ochre provenance studies using geochemistry and multivariate statistics [21,31,32,41,42]; (ii) earth pigments characterized by mixed phase assemblages and technological signatures [9,15,16,43,44]; (iii) colour–composition relations in ferric systems [13,14]; (iv) phase differentiation in green-earth systems [10–12,45,46]; and (v) the preventive-conservation and degradation modelling literature linking environmental factors to damage mechanisms [17,18,47,48]. Two demonstrators then anchor the later discussion more explicitly.

The first is a French yellow-ochre family, used as a constrained forward–inverse demonstrator because Roussillon-type materials combine goethite-rich yellow earths with variable kaolinite, quartz, and hematite contents and are known to shift under fractionation and heating [5,13,43,44]. The second is the green-earth family, used as a lower-dimensional comparative case because celadonite and glauconite partly overlap chemically while remaining distinguishable through structural and vibrational criteria [10,12,45,46]. These cases serve as conceptual stress tests for the analytical argument rather than as an exhaustive review corpus. The overall material logic of the study is summarized in Figure 2, while the main inferential dependencies are made explicit in the directed acyclic graph in Figure 1. Decision guidance and robustness summaries are provided in the relevant later sections, where they are discussed alongside their interpretive limits.

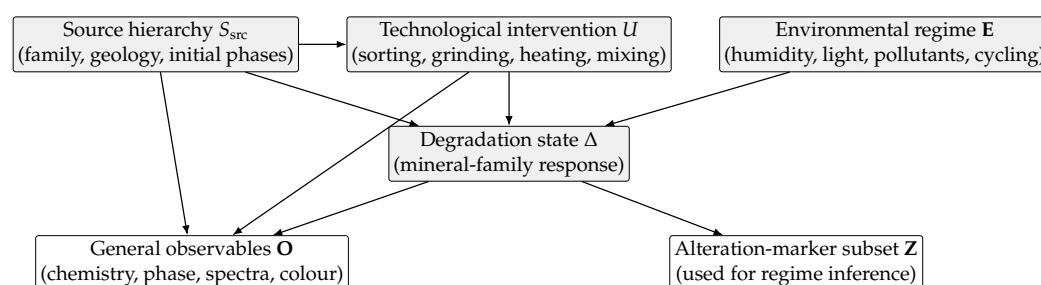


Figure 1. Minimal directed acyclic graph of the inferential dependencies assumed in the proposed analytical construction. The graph states only the causal ordering required by the argument: the source hierarchy S_{src} conditions technology and later degradation behaviour; environmental regime acts through degradation; and present observables contain combined source, process, and degradation signals, while alteration markers are the subset most directly used for regime-based environmental reasoning. Grey-filled boxes represent latent material, process, or environmental states; unfilled boxes represent observed analytical quantities; directed arrows denote the assumed conditional dependencies.

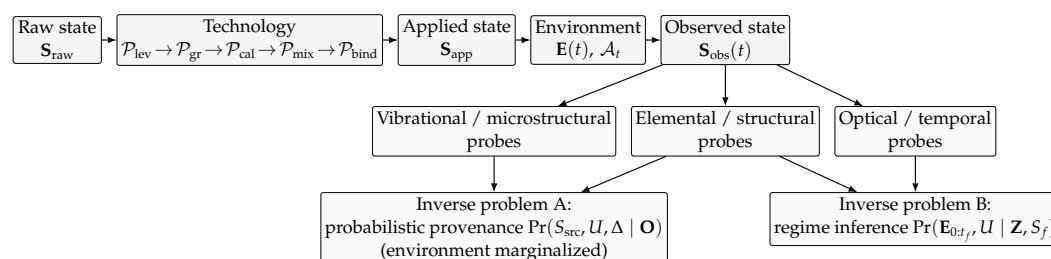


Figure 2. Schematic representation of the proposed analytical construction. S_{raw} is the raw geological material state; P_{lev} , P_{gr} , P_{cal} , P_{mix} , and P_{bind} denote selection/levigation, grinding, calcination, mixing, and binder incorporation; S_{app} is the applied pigment state; $S_{obs}(t)$ is the present observed state; $E(t)$ is environmental conditions; A_t is the degradation operator; O is the observable vector; S_{src} is source hierarchy; U is technological intervention; Δ is degradation state; S_f is source-family support; and Z is the alteration-marker vector. Arrows denote process or information flow; darker grey boxes identify material, technological, and environmental states, whereas lighter grey boxes identify analytical projections and inverse modules. The lower row represents measurable analytical projections whose uncertainty propagates into both inverse modules. Inverse problem A is shown in a compact environment-marginalized form; the expanded factorization is given in Equation (17). Inverse problem B is shown in its source-family-conditioned form; the corresponding family-marginalized form is given in Equation (23).

2.8. Reproducibility Note for the Published-Data Checks

All derived calculations for the two published-data checks were carried out in Python 3.11.8 using NumPy 1.24.0 and pandas 1.5.3; Matplotlib was used to generate Figure 3. For the Roussillon ochre check, the phase-fraction data and CIE-L*a*b* coordi-

nates were transcribed from the numerical values explicitly reported by Dubiel et al. [43]. The calculation uses tabulated published values rather than reprocessed raw spectra or digitized continuous curves. The three-part compositions were then transformed into ilr coordinates using the basis written later in Equation (26). Pearson correlations were calculated on the six bulk samples for which both phase and colour values were reported. In the present study, the expression *French-family centroid* denotes the arithmetic centroid of those six bulk Roussillon samples in ilr space. Formally, if $\bar{\mathbf{z}}_{\text{FY}}$ is that centroid, distances are calculated as $d(\mathbf{z}, \bar{\mathbf{z}}_{\text{FY}}) = \|\mathbf{z} - \bar{\mathbf{z}}_{\text{FY}}\|_2$. The family envelope, also called the dispersion envelope in the demonstrator section, is defined as the closed Euclidean ball centred at $\bar{\mathbf{z}}_{\text{FY}}$ with radius $r_{\text{FY}} = \max_i \|\mathbf{z}_i - \bar{\mathbf{z}}_{\text{FY}}\|_2$ over the six bulk Roussillon samples. Euclidean distances for the bulk state and two grain-size fractions of sample 9 were calculated relative to that centroid and envelope.

For the green-earth check, the published sample-level structural readings in Fanost et al. [46] were condensed into a categorical comparison table. The inclusive summary counts GE_Boh as family-positive but structurally ambiguous, because the published text places it close to the celadonite–glauconite family while also emphasizing mixed-layer disorder. A stricter sensitivity rule that counts GE_Boh as non-family lowers the descriptive concordance from $4/8 = 0.50$ to $3/8 = 0.375$; excluding the ambiguous case from the denominator gives $3/7 = 0.43$. The two derived comma-separated values (CSV) files are provided as Supplementary Materials and are sufficient to reproduce the reported ilr balances, correlations, centroid distances, and counting summaries. For the empirical checks reported here, Supplementary Files S3 and S4 report the robustness summaries and rule-sensitive concordance outputs, while Supplementary Files S5 and S6 regenerate Files S3 and S4 directly from Files S1 and S2 and regenerate the visualization shown in Figure 3. Supplementary File S5 is the Python script; Supplementary File S6 mirrors the same workflow in notebook form.

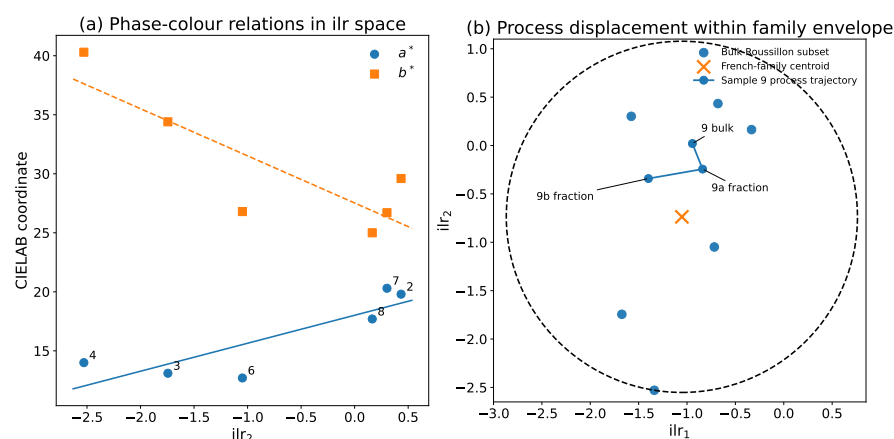


Figure 3. Author-generated visualization of a limited published-data check based on numerical values reported by Dubiel et al. for Roussillon ochres [43]. The original phase-fraction and CIE-L*a*b* values were reported in Ref. [43]; all log-ratio transformations, correlations, centroid and envelope calculations, process-trajectory representation, and graphical elements were produced by the authors of the present study. No published figure, image, or table layout is reproduced or adapted. (a) Light-blue circles represent a^* values and orange squares represent b^* values after the three-part phase composition is re-expressed in ilr space; the hematite balance ilr_2 remains positively associated with a^* and negatively associated with b^* in the six-sample subset for which colour coordinates were published. (b) Blue circles represent the bulk Roussillon subset, the cross marks the French-family centroid, and the connected line represents the bulk state and two grain-size fractions of published sample 9. The sample 9 trajectory shows that mechanical preparation shifts the observed composition while keeping the states within the French-family envelope. Abbreviations: CIE, International Commission on Illumination; ilr, isometric log-ratio.

Because the Article combines a conceptual construction with limited published-data checks, Table 2 states which components are conceptual, which are partly operationalized, and which still require future calibration.

Table 2. Current operational status of the analytical model and its three applications.

Component	Status in This Paper	What Is Already Supported Here	What Still Requires Future Calibration
State representation and observation model	Conceptually specified	Explicit state variables, operators, and scale-dependent observation logic	Cross-technique weighting, uncertainty propagation, and application-specific measurement design
Roussillon ochre check	Partly operationalized	ilr re-expression, phase-colour correlations, and process-displacement distances derived from published data	Larger datasets, uncertainty envelopes, and blinded source-class tests
Green-earth family check	Partly operationalized	Published structural assignments condensed into a family-discrimination check with sensitivity analysis	Broader historical and commercial corpora, plus formal classifier benchmarking
Herculaneum degradation application	Literature-mapped application	Source, technology, environmental event, degradation state, and multi-technique observations are explicitly connected	Quantitative kinetic calibration and uncertainty propagation require controlled reference data
Hierarchical provenance module	Conceptually specified; not yet fully calibrated	Posterior structure, class hierarchy, and confidence logic	Priors, likelihoods, training sets, and evaluation against larger reference collections
Regime-based degradation module	Conceptually specified; not yet fully calibrated	State-space formulation, forcing logic, and regime-based inverse reasoning	Mineral-family response libraries, controlled-ageing datasets, and calibrated inverse models

3. Results and Analytical Implications

3.1. From Historical Names to Material States

A first implication of the proposed method is conceptual and directly relevant to diagnostic practice: historical names such as ochre, sienna, umber, and green earth do not identify stable analytical objects. They are culturally meaningful, but materially under-resolved. Once translated into state variables, natural earth pigments appear as partially observed polymineral geological systems whose observable properties are shaped by geology, technology, and degradation. This shifts the analytical starting point from nominal identification toward state reconstruction: the relevant question becomes which material state is represented by the observed assemblage. A heated earth from one source may be materially closer to an unheated earth from another source than to the unprocessed material from its own deposit.

3.2. Implications for Probabilistic Source Attribution

A second implication concerns method rather than terminology. Provenance is not treated as a direct match between a sample and a deposit. Instead, the model estimates relative support for source classes defined at regional, deposit, and sub-facies levels. This shift has three methodological consequences. First, pigment raw materials inherit geological variability from deposits, weathering profiles, and sub-facies. Second, technological operations such as levigation and heating modify the observed signal before application. Third, degradation redistributes or weakens the source signal after application. Posterior source support requires conditioning not only on composition but also on plausible techno-

logical and degradation states. Equation (17) formalizes this point: provenance support is strongest when independent lines of evidence converge, and weakest when chemistry, phase assemblage, and context diverge.

This probabilistic reformulation is consistent with current archaeometric evidence. Linear discriminant analysis (LDA) can discriminate classes efficiently when source groups are defined carefully, while principal component analysis (PCA), clr transformation, or ilr transformation can improve the exploratory treatment of geochemical datasets affected by constant-sum constraints [21,22,32]. The model predicts that source confidence should increase only when geochemistry, phase assemblage, accessory minerals, and contextual geological plausibility converge. The main statistical layers involved in this step are summarized in Table 3.

Table 3. Statistical layers of the provenance module.

Layer	Primary Task	Typical Tools and Outputs
Compositional preprocessing	Respect the simplex structure of geochemical data and suppress spurious correlations	clr or ilr transforms, robust handling of below-detection values, PCA or LDA in transformed space; outputs are geometry-consistent multivariate summaries.
Phase and spectral separation	Estimate phase-related endmembers and their relative abundances in mixed systems	Raman/XRD pairing, hyperspectral unmixing, parallel factor analysis (PARAFAC) or related tensor decompositions; outputs are endmembers, abundances, and sometimes fading rates.
Hierarchical source inference	Estimate posterior support for province, deposit, and sub-facies classes under technological and degradation uncertainty	Bayesian classification, mixture models, discriminant models, expert priors; outputs are compatibility classes rather than deterministic matches.

Abbreviations: clr, centred log-ratio; ilr, isometric log-ratio; LDA, linear discriminant analysis; PARAFAC, parallel factor analysis; PCA, principal component analysis; XRD, X-ray diffraction.

3.3. Partial Separation of Geological, Technological, and Degradation Signals

The analytical construction predicts that observables are mixtures of three signal families and their interactions. This is expressed as

$$\mathbf{O} \approx \mathbf{O}^{(G)} + \mathbf{O}^{(T)} + \mathbf{O}^{(D)} + \mathbf{I}_{GT} + \mathbf{I}_{GD} + \mathbf{I}_{TD} + \boldsymbol{\varepsilon}, \quad (24)$$

where $\mathbf{O}^{(G)}$, $\mathbf{O}^{(T)}$, and $\mathbf{O}^{(D)}$ are the geological, technological, and degradation contributions to the observable vector; $\mathbf{I}_{GT}$, $\mathbf{I}_{GD}$, and $\mathbf{I}_{TD}$ are pairwise interaction terms between geological–technological, geological–degradation, and technological–degradation signals; and $\boldsymbol{\varepsilon}$ is residual measurement noise plus unmodelled structure. Equation (24) is a first-order interpretive decomposition in encoded observable space, not an exact physical partition of independently measurable contributions. It provides the structural answer to the question of phase separation by stating that the measured signal is a composite of geological, technological, and degradation contributions. Statistical analysis can provide only a partial disentangling of signal families across data domains: chemistry, mineralogy, spectroscopy, texture, and colour.

This distinction has practical implications. Elemental data can preserve regional source fingerprints but may not resolve mineral phase. Vibrational and diffraction data may discriminate phase but not prove provenance. Optical data may be extremely sensitive to technological processing and photochemical change. The model implies that phase separation is a multilevel inference task combining compositional transforms, spectral unmixing, process-aware interpretation, and awareness of the instrumental limitations of portable analytical setups [8,20,28,29,49]. The three signal families and their typical observables are summarized in Table 4.

Table 4. Three-signal matrix proposed for the diagnosis of natural earth pigments.

Signal Family	Core Variables	Typical Processes	Representative Observables and Implications
Geological signal	Deposit type L_{geo} ; phase-fraction vector ϕ ; accessory minerals $\mathbf{m}_{\text{acc}}$; geochemical background $\mathbf{e}_{\text{geo}}$	Primary mineral association, weathering-profile inheritance, natural textural heterogeneity	XRD phase assemblage, Raman band identity, accessory-mineral markers, multielement fingerprints; strongest direct relevance for source compatibility.
Technological signal	Levigation λ_{lev} ; grinding g_{gr} ; heating (θ_{heat} , τ_{heat}); mixing $\mathbf{M}_{\text{mix}}$; binder/interfacial term b_{int}	Fractionation, particle-size redistribution, dehydroxylation, oxidation, dilution, interfacial modification	Grain-size changes, hematite/goethite rebalance, modified spectral intensity, altered colour coordinates; may mimic or obscure geological signal.
Degradation signal	Hydration ω ; redox state ξ ; structural order ρ ; weathering load s_D ; interface term β	Rehydration, salt damage, secondary phase formation, recrystallization, surface hydroxylation, pollutant interaction	FTIR hydration signatures, broadened Raman bands, secondary sulfates or salts, microscopy evidence of decay, colour drift; reduces provenance confidence but increases sensitivity to environmental regimes.

3.4. Mineral-Specific Degradation Responses

A fourth implication is that degradation is described more rigorously at the level of mineral families than at the level of a generic pigment label. Ferric systems, green-earth systems, carbonates, sulfates, and accessory phases each define different response spaces under environmental conditions. For ferric systems, chromatic change may result from phase redistribution, surface hydroxylation, dehydration, and particle-size effects [13,14,50,51]. For carbonate and sulfate-bearing systems, humidity, acidic gases, and solution chemistry may drive dissolution, recrystallization, and hydration–dehydration sequences, with phase selection strongly dependent on temperature, salinity, and water activity [36,52]. For light-sensitive minerals, ultraviolet exposure may induce measurable chromatic drift [53].

This means that degradation cannot be modelled as one scalar index. It is more appropriately treated as a vector-valued response surface over mineral families, environmental conditions, and technological preconditioning. The presence of heating or calcination in the technological history is especially important because it may generate altered precursor states that later influence hydration, wetting, or spectral response. The mineral-family response space used in the present study is summarized in Table 5.

Table 5. Mineral-family response space proposed for the degradation module.

Mineral Family	Dominant Forcing Variables	Indicative Alteration Markers
Fe oxides/oxyhydroxides	RH , pH, thermal history, UV dose, oxygen availability	Chromatic drift, redistribution of hematite/goethite signatures, surface-hydroxyl changes, modified wetting behaviour, band broadening.
Fe-bearing clays/green earths	RH , solution chemistry, cation exchange, thermal history	Variations in vibrational bands, mixed-layer behaviour, altered spectral slopes, confusion between celadonic and glauconitic signatures.
Carbonates	Acidic gases, liquid water, condensation, salt solutions	Dissolution, recrystallization, roughness increase, gypsum or mixed sulfate formation in polluted humid conditions.
Sulfates	Temperature, water activity, salinity, wet–dry cycles	Hydration–dehydration shifts among gypsum, bassanite, and anhydrite; crusts, efflorescences, hygroscopic behaviour.
Accessory phases	Variable; often governed by microenvironmental chemistry and technological inheritance	Context-specific markers that may strengthen or weaken provenance inference and help identify modification mechanisms.

3.5. Compatible Environmental Regimes Rather than a Unique History

A fifth implication follows from the ill-posed nature of Equation (22). At this point, the model becomes more conservative and materially realistic. Present alteration markers cannot usually reconstruct a unique environmental trajectory. However, they can constrain the set of compatible environmental regimes. The inverse module suggests four broad inferential classes for environmental exposure history: (i) humid-photochemical regimes, where light and moisture jointly dominate; (ii) acid-humid regimes, where pollutants and moisture promote secondary sulfate or corrosive products; (iii) dry-thermal regimes, where dehydration and heating dominate; and (iv) cyclic salt-driven regimes, where repeated wet–dry transitions and saline solutions generate crystallization stress and secondary phase reorganization.

This regime-based assessment avoids the unsupported assumption that a unique environmental trajectory can be recovered from incomplete present-day evidence. It is also more useful for heritage science because preventive-conservation decisions depend primarily on identifying active risk regimes and dominant forcings rather than reconstructing an exact historical microclimate with impossible precision [18,35,47,48,54,55].

3.6. Interpretive Confidence Classes

A final implication is that confidence classes can be defined for both source attribution and environmental-regime inference. Provenance can be classified as broad compatibility, constrained compatibility, regional attribution, or high-confidence source attribution. Environmental inversion can be classified as weakly constrained regime, partially constrained regime, dominant regime with secondary alternatives, or strongly supported regime. In both cases, confidence increases when independent observables converge and decreases when technological or degradation confounders are likely. Tables 4 and 5 gather the main interpretive matrices used in the discussion.

4. Published-Data Applications

4.1. Roussillon Yellow Ochres: Original Findings and Model Application

Dubiel et al. examined ochres from the Sentier des Ogres near Roussillon using Mössbauer spectroscopy and CIE-L*a*b* colourimetry [43]. Their principal result was that colour varied systematically with the relative amounts of hematite, goethite, and kaolinite: redness increased with the hematite contribution, yellowness was associated more strongly with goethite, and light-coloured phases modified chroma and lightness. The study also showed that wet-sieved grain-size fractions of the same sample could differ in phase proportions. Elias et al. independently demonstrated that ochre colour depends on the balance between ferric chromophores and lighter associated phases [13].

The present study retains those conclusions. It uses the published values to ask a different question: how should source compatibility be described when a known preparation step changes the measured composition? For this purpose, the relevant source object is a demonstrator family S_{FY} comprising goethite-rich French yellow earths with variable clay, quartz, and hematite contents. The notation S_{FY} is specific to this application and is distinct from the general family variable S_f used in the source-attribution model.

For this demonstrator, let S_{FY} denote the French yellow-ochre source family; let $\mathcal{P}_{mech}$ denote mechanical preparation; let $\mathcal{P}_{cal}$ denote thermal treatment or calcination; let $\mathcal{P}_{use}$ denote use-stage mixing or application; let $\mathcal{D}_E$ denote degradation under environmental regime E ; and let $\mathcal{Q}$ denote the observation operator. With these quantities defined, the forward chain can be written as

$$\mathbf{O} = \mathcal{Q}[\mathcal{D}_E(\mathcal{P}_{use}(\mathcal{P}_{cal}(\mathcal{P}_{mech}(S_{FY}))))] + \epsilon, \quad (25)$$

where $\mathbf{O}$ is the resulting observable vector and ε is residual observational uncertainty. Mechanical preparation is transformative rather than neutral. In the Roussillon case, each raw sample was separated into four grain-size fractions by wet sieving, and the colourimetric behaviour shifted with the relative abundance of hematite, goethite, and kaolinite [43]. Thermal treatment introduces a second non-uniqueness: goethite dehydrates to hematite around 260–300 °C [5,56], but mechanical grinding can also induce a topotactic goethite-to-hematite transformation [44]. The consequence is important for interpretation: an observed hematite-rich state need not record a hematite-rich source.

In inverse mode, the demonstrator is designed to return a restricted set of compatible histories rather than a single quarry. This limitation is methodological. Raman spectroscopy can separate ordered from disordered hematite, but it cannot by itself distinguish heated goethite from other disordered hematites generated by weathering or grinding [57]. The expected output is posterior support over a family of geological–technological–environmental exposure histories, for example “goethite-rich French-type yellow earth, mechanically fractionated, possibly thermally modified, subsequently affected by humid sulfate-rich conditions”. This form of inference is more cautious than direct naming and more faithful to the material problem.

4.2. Roussillon Re-Analysis in Compositional Space

A compact proof-of-concept re-analysis was carried out on numerical values reported by Dubiel et al. for Roussillon ochres [43]. No new measurements were produced. The purpose is narrower: the re-analysis examines whether two low-level claims of the proposed method are supported by published data. The first claim is that three-part phase-fraction data are more appropriately analysed in log-ratio space than as raw percentage vectors. The second is that mechanical preparation changes the observed bulk composition of a natural earth pigment, but need not erase its family-level compatibility.

For the six bulk samples for which Dubiel et al. reported both Mössbauer phase-fraction data and reflectance CIE-L*a*b* coordinates (samples 2, 3, 4, 6, 7, 8), the three-part composition is written as $\mathbf{x} = (K_{\text{kao}}, G_{\text{goe}}, H_{\text{hem}})$ after closure to the kaolinite–goethite–hematite subsystem. Here, K_{kao} , G_{goe} , and H_{hem} are the corresponding kaolinite, goethite, and hematite proportions. These three proportions were re-expressed in two explicit ilr balances,

$$\text{ilr}_1 = \sqrt{\frac{1}{2}} \ln\left(\frac{K_{\text{kao}}}{G_{\text{goe}}}\right), \quad \text{ilr}_2 = \sqrt{\frac{2}{3}} \ln\left(\frac{H_{\text{hem}}}{\sqrt{K_{\text{kao}} G_{\text{goe}}}}\right), \quad (26)$$

where the balances are defined for strictly positive values of K_{kao} , G_{goe} , and H_{hem} , the percentage abundances of kaolinite, goethite, and hematite reported by Dubiel et al. The two balances correspond to a sequential binary partition in which hematite is first contrasted with the kaolinite–goethite group and then kaolinite is contrasted with goethite within that group. Positive ilr_2 values indicate hematite-rich states relative to the kaolinite–goethite geometric mean, and more negative values indicate goethite-dominant yellow states. The derived dataset used in this compact check is reported in Table 6.

Within this small six-sample subset, the hematite balance ilr_2 shows the chromatic direction expected from the theoretical argument: it correlates positively with a^* (Pearson $r = 0.85$) and negatively with b^* ($r = -0.84$). These are not large-sample estimates and should not be over-interpreted; nevertheless, they indicate that the phase–colour logic of the proposed method remains visible after the data are moved into a compositionally valid coordinate system. Figure 3a visualizes these two relations.

The second check uses sample 9 from the Dubiel et al. dataset, because that sample is reported both as a bulk material and as two grain-size fractions produced by wet sieving (9, 9a, 9b). When re-expressed in ilr space, the three states define a monotonic process

trajectory: ilr_2 decreases from 0.02 to -0.34 as the relative hematite contribution falls and the goethite contribution rises. At the same time, the three states remain inside the dispersion envelope of the six published bulk Roussillon samples used to define the French-family centroid, using the closed-ball rule specified in Section 2.8. Their distances to that centroid in ilr space are 0.77, 0.54, and 0.53, all well below the observed maximum distance of 1.81 within the six-sample bulk subset itself. This is the kind of behaviour expected under the process-sensitive argument advanced here: mechanical preparation can rewrite the measured bulk composition and shift the optical balance, yet still preserve family-level compatibility at the family scale. Figure 3b shows the displacement, while Table 6 reports the derived coordinates.

Taken together, these small checks do not constitute full empirical confirmation of the inferential construction, and they do not convert the article into a data-rich predictive study. They do, however, operationalize two of its core claims with published measurements: first, that composition is most appropriately handled in log-ratio space; second, that bulk chemistry remains informative despite its sensitivity to processing. In this paper, this proof-of-concept re-analysis is meant to connect the conceptual construction with measurable behaviour, not to imply that full calibration has already been achieved.

Table 6. Derived dataset used for the first published-data check. Phase-fraction data (K_{kao} = kaolinite, G_{goe} = goethite, H_{hem} = hematite) and reflectance coordinates are taken from Dubiel et al. [43]; ilr coordinates were calculated for the present study. Asterisks denote process states of published sample 9.

State	K (%)	G (%)	H (%)	a^*	b^*	ilr_1	ilr_2
2	15.7	41.1	43.2	19.8	29.6	-0.68	0.43
3	8.3	88.5	3.2	13.1	34.4	-1.67	-1.74
4	12.9	85.6	1.5	14.0	40.3	-1.34	-2.53
6	23.7	65.4	10.9	12.7	26.8	-0.72	-1.05
7	6.8	63.2	30.0	20.3	26.7	-1.58	0.30
8	24.1	38.6	37.3	17.7	25.0	-0.33	0.16
9* bulk	14.7	55.9	29.4	—	—	-0.94	0.02
9a* frac.	17.8	58.3	23.9	—	—	-0.84	-0.24
9b* frac.	10.0	72.3	17.7	—	—	-1.40	-0.34

4.3. Green Earths: Original Findings and Diagnostic Question

The green-earth application addresses mineral-family identification rather than a colour-phase correlation. Previous studies showed that celadonite and evolved glauconite can overlap substantially in major-element composition, whereas XRD, FTIR, Raman, and selected near-infrared features retain more useful structural distinctions [10–12,45]. Fanost et al. compared geological reference materials, eight commercial pigments sold as “green earth”, and a historical Monte Baldo sample [46]. They concluded that the commercial name was not a reliable mineralogical assignment: some products were celadonitic or glauconitic, whereas others were dominated by gypsum, calcite, quartz, smectitic material, or organic colourants. They also showed that mobile XRF alone was insufficient for celadonite–glauconite discrimination in complex paint layers, while structural and vibrational measurements provided stronger evidence.

The present re-analysis condenses those published assignments into a transparent family-membership check. It asks whether the commercial label or chemistry-limited screening can substitute for phase-resolved evidence. The answer is deliberately restricted to the published subset: the label describes a market category, whereas structural measurements determine whether a sample belongs to the celadonite–glauconite mineral family.

4.4. Green-Earth Re-Analysis and Family Discrimination

A second published-data check can be built from the published multi-analytical dataset of Fanost et al., which includes two geological reference minerals, eight commercial pigments sold as “green earth”, and one sample from a historical Monte Baldo location [46]. The question here is different from the Roussillon ochre case. The target is not a colour-phase relation but family discrimination: can a commercial or chemistry-led description alone identify membership in the celadonite–glauconite system, or does reliable separation emerge only when structural markers are introduced?

Within the eight commercial samples, the published XRD-based assignments indicate that GE_Cyp and BGE_Cyp are celadonitic, GE_Rus is glauconitic, and GE_Boh is an ill-ordered borderline material whose d_{060} line at 0.1512 nm lies close both to the glauconitic threshold and to previously described interlayered celadonite/smectite material [46]. By contrast, GE_Ver, GE_Bren, GE_Nico, and GE_Sen do not show celadonite or glauconite reflections, while the historical GE_Bren_h sample is mainly smectitic [46]. For this small dataset-bound check, let N_{family} be the number of commercial-label samples assigned to the celadonite–glauconite family, let N_{label} be the total number of samples carrying the commercial label, and let C_{label} be the corresponding descriptive concordance. The commercial label “green earth” then gives

$$C_{\text{label}} = \frac{N_{\text{family}}}{N_{\text{label}}} = \frac{4}{8} = 0.50,$$

(27)

when GE_Boh is counted inclusively as family-positive but ambiguous. Under a stricter rule that counts GE_Boh as non-family because of its structural ambiguity, the same ratio falls to $3/8 = 0.375$; if the ambiguous case is excluded from the denominator, it becomes $3/7 = 0.43$. The breakdown is summarized in Table 7. The exercise is heuristic and dataset-bound, best read as descriptive concordance within the published subset rather than as a general performance estimate.

Table 7. Compact published-data check on green earths derived from Fanost et al. [46]. The table does not reproduce the full original dataset; it condenses the published assignments into a simple descriptor test showing that the commercial label “green earth” is much broader than actual celadonite–glauconite family membership. The inclusive count treats GE_Boh as family-positive but ambiguous; the stricter sensitivity rule is discussed in the text.

Published Sample Group	Sold as “Green Earth”	Published Structural Reading	Inference for Family Membership
GE_Cyp, BGE_Cyp	Yes	Celadonite-compatible XRD/NIR behaviour; better-resolved celadonitic signatures	Positive: celadonitic branch of the family
GE_Rus	Yes	Glauconite-compatible pattern	Positive: glauconitic branch of the family
GE_Boh	Yes	Borderline case; $d_{060} = 0.1512$ nm and ill-ordered mixed-layer behaviour	Positive but structurally ambiguous; illustrates non-uniqueness within the family
GE_Ver, GE_Bren, GE_Nico, GE_Sen	Yes	No celadonite/glauconite reflections; mixtures dominated by gypsum, calcite, quartz and/or other phases	Negative: commercial label does not imply true green-earth family membership
GE_Bren_h (historical site sample)	No commercial label	Mainly smectitic material from a historical location	Negative for the celadonite–glauconite family in the analysed state

Abbreviations: FTIR, Fourier-transform infrared spectroscopy; IR, infrared; NIR, near-infrared; pXRF, portable X-ray fluorescence; SEM-EDS, scanning electron microscopy with energy-dispersive spectroscopy; XRD, X-ray diffraction.

This result is not intended as an evaluation of commercial terminology; it serves as a diagnostic stress test for the proposed analytical construction. The same paper states explicitly that mobile XRF is ineffective for distinguishing celadonite from glauconite in complex paint layers because Al, Mg, and Na cannot be quantified accurately enough, whereas XRD and SEM–EDS with sampling, together with infrared/near-infrared (IR/NIR) spectroscopy in favourable in situ cases, provide discriminating structural information [46]. The larger mineralogical review by Singh et al. reaches the same conclusion from a different angle: the major-element fields of celadonite and evolved glauconite overlap substantially, while the decisive differences are better preserved in XRD and FTIR criteria [12]. This separation corresponds to the distinction advanced in this paper. Label-level or chemistry-limited descriptors remain too weak on their own, whereas phase-resolved and structurally informed observations recover a narrower and more defensible mineral-family decision.

The published NIR case study on a green garland in Nicolas Poussin's *Bacchanales d'enfants* offers a compact in situ illustration of the same logic. Fanost et al. report that bands at 2258, 2303, and 2353 nm matched celadonite better than the glauconitic pattern at 2314 and 2363–2364 nm, allowing the painted green to be assigned to a celadonitic source family even without invasive sampling [46]. The significance of the NIR case is not that this technique resolves every green-earth problem; rather, structural or vibrational signatures provide stronger evidence than generic commercial labels and chemistry-limited screening when the family distinction is the analytical target.

4.5. Degradation-Oriented Application: Ochre Transformation at Herculaneum

The two numerical re-analyses above primarily test source-family and processing questions; neither dataset contains a controlled time series of degradation. To show how the degradation component applies to heritage evidence, the model is mapped here to the published study of the tablinum in the House of the Bicentenary at Herculaneum [15]. Secco et al. combined in situ pXRF with XRD, Raman spectroscopy, and electron microscopy to distinguish original red ochres from yellow ochres converted to red by heat associated with the AD 79 eruption. Their study reconstructed an original decorative scheme containing both red and yellow fields and showed that eruption-related temperatures modified part of the yellow palette through goethite-to-hematite conversion. It also documented later deterioration associated with water infiltration and secondary surface deposits.

Within the present notation, the source-family term S_f denotes an ochre family containing goethite- and hematite-bearing materials; U includes pigment selection, mixing, and wall-painting preparation; the environmental-condition vector $E(t)$ includes the short, high-temperature volcanic event and later moisture exposure; the degradation state Δ includes thermally converted ferric phases and secondary alteration products; and the observable vector O contains pXRF, XRD, Raman, SEM–EDS, stratigraphic, and colour evidence. The published interpretation is supported because several analytical domains converge on a process-consistent explanation. The case also demonstrates the limit of inversion: the present red colour alone cannot distinguish an originally red ochre from a thermally converted yellow ochre, and the full multi-technique evidence is required.

This application does not calibrate the dynamic model in Equation (20). Rather, it demonstrates how the model organizes documented degradation evidence and how a specific environmental event can be assessed against competing material histories. Quantitative calibration of the degradation module still requires controlled-ageing data or reference collections with known exposure histories.

Table 8 maps the three applications to the principal variables of the analytical model and states the inferential limits of each case.

Table 8. Mapping of the three published applications to the analytical model. The table distinguishes the evidence used, the components represented, the inference supported, and the principal unresolved limitation.

Application	Source, Process, and Degradation Terms	Published Observations Used	Supported Inference and Principal Limit
Roussillon yellow ochres	S_f : goethite-rich French ochre family; U : wet sieving and grain-size fractionation; Δ : not directly observed	Mössbauer phase fractions and CIE- $L^*a^*b^*$ colour coordinates	Processing shifts composition and colour while family compatibility may persist; no degradation history or unique quarry is inferred.
Green earths	S_f : celadonite–glauconite family; U : commercial formulation and extenders; Δ : weakly constrained	XRD, SEM–EDS, IR/NIR, and published sample assignments	Structural evidence is more diagnostic than label or bulk chemistry alone; exact geological source remains uncertain.
Herculaneum ochres	S_f : ferric ochre family; U : pigment preparation and wall-painting application; $E(t)$: eruption heat and later moisture; Δ : thermal conversion and secondary products	pXRF, XRD, Raman, SEM–EDS, stratigraphy, and colour	Convergent evidence distinguishes original red from thermally converted yellow ochre; the dynamic degradation model is illustrated but not quantitatively calibrated.

4.6. Resampling and Sensitivity Analyses of the Published-Data Checks

The two published-data demonstrators remain limited in scope, but their internal behaviour can be examined more explicitly through stress-testing of the derived datasets. To strengthen the empirical layer without overstating their evidential reach, three additional checks were carried out on the derived datasets supplied as Supplementary Materials: (i) nonparametric bootstrap resampling of the six bulk Roussillon samples; (ii) leave-one-out perturbation of the French-family centroid and dispersion envelope; and (iii) rule-sensitive concordance summaries for the green-earth subset. These operations do not convert the article into a calibrated predictive study. Their purpose is narrower: they assess whether the directional claims made by the analytical construction remain stable under internal perturbation of the compact published-data checks.

Bootstrap resampling was performed at the sample level on the six bulk observations, with replacement, for 20,000 replicates; reported intervals are percentile intervals computed from the bootstrap distribution. The exact permutation tests enumerate all $6!$ label permutations of the six-sample subset and evaluate two-sided tail probabilities for $|r|$ under the null of exchangeability. In the leave-one-out envelope check, each of the six bulk samples is omitted in turn, the centroid is recomputed from the remaining five samples, and the corresponding radius is set to the maximum within-family distance among those five samples. For the six bulk Roussillon samples, 20,000 bootstrap resamples preserve the expected sign structure of the transformed phase–colour relation. The correlation between ilr_2 and a^* remains positive in 98.4% of resamples, with a bootstrap median of 0.91 and a 95% percentile interval of $[0.56, 0.99]$. The correlation between ilr_2 and b^* remains negative in 92.9% of resamples, with a bootstrap median of -0.86 and a percentile interval of $[-0.99, 0.54]$. Leave-one-out omission of each of the six bulk samples preserves the same sign pattern in all six cases, with $r(ilr_2, a^*)$ ranging from 0.81 to 0.94 and $r(ilr_2, b^*)$ ranging from -0.93 to -0.64 . Exact permutation tests remain appropriately modest, with two-sided permutation p -values of 0.075 for a^* and 0.035 for b^* . These numbers do not externally validate a predictive model; they show that the directional logic of the transformed demonstrator is internally stable but still sample-limited.

The family-level argument around sample 9 and its fractions also survives a stronger internal perturbation. When the French-family centroid is recomputed in leave-one-out fashion from five of the six bulk samples, the bulk state and both grain-size fractions of sam-

ple 9 remain inside the corresponding family envelope in all 6/6 omission scenarios. Thus, the process displacement remains visible, while family compatibility is not attributable to a single fixed centroid definition.

The green-earth demonstrator can be strengthened in a parallel way by replacing a single summary value with an explicit sensitivity set. Under the inclusive rule that counts GE_Boh as family-positive but ambiguous, the descriptive concordance of the commercial label with structural family membership is $4/8 = 0.50$, with a Wilson 95% interval of $[0.22, 0.78]$. Under the stricter rule that counts GE_Boh as non-family, the concordance becomes $3/8 = 0.375$, with interval $[0.14, 0.69]$. Excluding GE_Boh from the denominator yields $3/7 = 0.43$, with interval $[0.16, 0.75]$. These wide intervals are themselves informative: they reinforce the article's central claim that a chemistry-limited or label-led descriptor is too fragile to stand in for structural family discrimination.

Taken together, these additional checks do not alter the evidential status of the paper. They strengthen its empirical footing within the restricted scope of published-data re-analysis. The Roussillon demonstrator now includes sign-stability, permutation, and leave-one-out evidence, while the green-earth demonstrator now reports explicit uncertainty and rule sensitivity rather than a single point summary. The derived outputs are supplied as Supplementary Files S3–S6, and the calculations can be reproduced directly from the accompanying Python script and companion notebook. In particular, Files S5 and S6 regenerate Files S3 and S4 directly from Files S1 and S2 and regenerate the visualization shown in Figure 3. The compact robustness summary is given in Table 9.

Table 9. Compact robustness summary for the two published-data demonstrators. The table does not claim external predictive validation; it records the stability of the illustrative checks under bootstrap, leave-one-out, permutation, and rule-sensitive perturbations.

Demonstrator	Perturbation	Main Result
Roussillon bulk subset	20,000 bootstrap resamples	$\text{sign}[r(\text{ilr}_2, a^*)] > 0$ in 98.4% of resamples; $\text{sign}[r(\text{ilr}_2, b^*)] < 0$ in 92.9% of resamples.
Roussillon bulk subset	Leave-one-out correlations	Sign pattern preserved in all 6/6 omissions; $r(\text{ilr}_2, a^*) \in [0.81, 0.94]$ and $r(\text{ilr}_2, b^*) \in [-0.93, -0.64]$.
Roussillon bulk subset	Exact permutation tests	Two-sided permutation $p = 0.075$ for a^* and $p = 0.035$ for b^* .
Sample 9 process trajectory	Leave-one-out family envelope	Bulk state and both fractions remain inside the family envelope in all 6/6 centroid-perturbation scenarios.
Green-earth subset	Inclusive concordance rule	$4/8 = 0.50$; Wilson 95% interval $[0.22, 0.78]$.
Green-earth subset	Stricter sensitivity rules	$3/8 = 0.375$ with interval $[0.14, 0.69]$; excluding GE_Boh gives $3/7 = 0.43$ with interval $[0.16, 0.75]$.

4.7. Bulk Chemistry Is a Conditional Descriptor, Not an Invariant Fingerprint

Because processing and ageing can alter bulk chemistry, its use in provenance assessment requires explicit qualification. A descriptor with high discriminatory potential can become misleading if it is interpreted prematurely as a source fingerprint. Bulk chemistry is not an invariant source marker; nevertheless, it remains informative when its stable components are evaluated after mineral-family separation, process screening, and matrix effects have been made explicit. In this restricted sense, bulk chemistry functions as a conditional descriptor of high interpretive value.

Let $\mathbf{x}$ denote a bulk compositional vector and let $\mathbf{z} = \text{ilr}(\mathbf{x})$ denote its isometric log-ratio coordinate representation. In those compositional coordinates, the observed chemistry can be written as a first-order perturbation expression,

$$\mathbf{z}_{\text{obs}} = \mathbf{z}_{\text{src}} + \delta_{\text{mech}} + \delta_{\text{therm}} + \delta_{\text{mix}} + \delta_{\text{deg}} + \delta_{\text{sub}} + \varepsilon, \quad (28)$$

where $\mathbf{z}_{\text{obs}}$ and $\mathbf{z}_{\text{src}}$ are observed and source compositions in ilr coordinates; δ_{mech} , δ_{therm} , δ_{mix} , δ_{deg} , and δ_{sub} are vector shifts caused by mechanical processing, thermal treatment, mixing, degradation, and substrate/matrix contribution; and ε is residual uncertainty in the same ilr coordinate space. The equation is not formulated as an exact conservation law in Euclidean space; it is a first-order interpretive decomposition written in ilr coordinates so that compositional perturbations can be discussed without violating simplex geometry. Like Equation (24), it should be read as a linearization in encoded coordinate space, rather than as a literal physical separation of independently measurable contributions. The Roussillon case already demonstrates the first of these effects: wet sieving changes phase fractions and shifts the effective bulk composition of the analysed fraction [43]. The second effect follows from chemistry and mineralogy together: thermal or mechanochemical goethite-to-hematite conversion changes the mineralogical host of Fe and, in mass-normalized observations, alters the composition actually measured [5,44]. Mauran et al. also showed that inter-equipment variability can make datasets appear incompatible until they are standardized, which is a reminder that chemistry can be rewritten by instrumentation as well as by material history [41]. Finally, Martín-Ramos et al. demonstrated that in pXRF studies of rock art, substrate-derived elements may drive much of the apparent statistical discrimination unless pigment and substrate variance are separated explicitly [21].

This argument does not imply that chemical data should be discarded. Once source families have been separated mineralogically, chemistry can retain substantial interpretive value. Chanteraud et al. showed that major elements may carry strong rock-type information when interpreted alongside petrography and geological context, and that geochemistry is most useful when the relevant elements are selected with reference to mineral phases and rock genesis [42]. In surface and porous systems, however, the observed chemistry is also an interfacial chemistry. Adsorption between goethite, quartz, and kaolinite depends strongly on pH and ionic strength [58]; hematite wetting depends on the surface hydroxylation state [50,51]; and salts or pollution-derived sulfates can add chemically important phases during ageing [36,39,52]. The resulting methodological implication is that bulk chemistry is unsuitable as a stand-alone provenance engine. Its interpretive value increases after mineral family, process history, substrate, and degradation effects have been screened as separate contributors. The contrast between the two demonstrators is summarized in Table 10.

Table 10. Two conceptual demonstrators used to clarify the forward–inverse logic of the analytical construction.

Case	Relatively Stable Descriptors	Process- or Environment-Sensitive Descriptors	Inverse Output That Is Realistically Defensible
French yellow ochres	Goethite-rich source family, associated kaolinite/quartz matrix, broad ferric-phase balance, regional geological plausibility	Grain-size fractionation, goethite-to-hematite conversion by heating or grinding, dilution by white phases, sulfate and salt additions, substrate contributions	A restricted family of compatible geological–technological–environmental exposure histories, not a unique quarry assignment.
Green earths	Membership in the celadonite–glauconite family, structural order, geological setting compatibility	Overlapping bulk chemistry, extenders such as gypsum or calcite, non-earth green mixtures, environmental alteration of phyllosilicate signatures	Identification of the most compatible green-earth family and geological setting, with explicit uncertainty about mixtures and commercial substitutions.

5. Discussion: Limits and Operational Pathways

5.1. Physical Limits of the Optical and Mixing Idealizations

The analysis uses Kubelka–Munk language and spectral-unmixing notation because they provide a shared vocabulary across colour physics and heritage diagnostics. This shared vocabulary does not imply unrestricted validity. In heritage applications, intimate nonlinear mixtures, thin paint films, layered systems, rough or porous surfaces, binder-rich media, and strong substrate contributions can all disrupt the idealized separation between absorption and scattering or the idealized additivity implied by unmixing relations. For this reason, Equations (11)–(16) are best understood as idealized local descriptions whose value depends on whether the physical situation is close enough to the assumptions built into them. When these conditions are not met, the analytical construction remains useful chiefly for identifying where the underlying simplifications are likely to fail.

5.2. Scale, Depth, and Support in the Observation Model

A related limit concerns commensurability. Portable XRF, XRD, Raman, FTIR, SEM–EDS, hyperspectral imaging, and colorimetry do not sample the same lateral area, the same depth, or the same kind of material support. A portable in situ spectrum, a microsample analysed by SEM–EDS, and a hyperspectral map pixel cannot be treated as if they referred to one analytically homogeneous point. This distinction is especially relevant in stratified, porous, or substrate-sensitive systems. The observation model proposed here is best read as an integration scheme across partial and scale-dependent views, rather than as a claim that all measurements are directly interchangeable.

5.3. Interpretive Consequences of Partial Observation

Once this scale dependence is acknowledged, pigment identity is no longer a matter of lexical recognition but of reconstructing a present material state from incomplete observables. For heritage science, the point is important because earth pigments have long been described through historically inherited labels that hide the mineralogical complexity of the actual material [4,34]. Provenance consequently becomes more constrained methodologically, but also more faithful to the complexity of the material system. Recent ochre studies show that source definition is methodologically consequential: grouping or separating lithologies within a weathering profile may alter attribution performance, and the success of multivariate methods depends strongly on how classes are defined [32]. At the same time, recent pXRF-based studies on rock art demonstrate that clr transformation and supervised models can improve classification performance when the compositional nature of the data is respected [21]. These studies support probabilistic source attribution across nested geological levels rather than direct matching.

The same logic applies to separability. The analyst does not observe geology, technology, and degradation independently; their superposition is what is observed analytically. PCA is useful only when interpreted in compositional space and with awareness that some components may still reflect technological or degradation filtering rather than geology alone [22,23]. Spectral unmixing is likewise conditional on the adopted mixing model, the spectral domain, and the quality of endmembers [20,30]. The two demonstrators address this issue through complementary examples. French yellow ochres illustrate how one source family can generate analytically similar end states after fractionation, heating, mixing, and ageing [43,44,57]. Green earths illustrate the converse inference problem: even when the number of candidate minerals is smaller, chemistry alone may overlap too strongly to be decisive, and structural information becomes the primary discriminator [12,46].

A similar caution applies to degradation theory. Inorganic pigments rarely respond to single agents in isolation; climate variables, pollutants, and display conditions interact

in mineral-dependent ways [17,53]. Hematite and goethite are not merely colour labels; they are also phases with different hydration, surface, and optical behaviour. The surface state of hematite, for example, changes with pH and hydroxylation, which affects wetting and the chemistry of its interaction with water-bearing environments [50,51]. For this reason, inverse degradation reasoning is formulated as regime inference rather than as exact reconstruction of exposure history. The proposed inferential treatment is conservative in this respect: it extracts compatible environmental classes from present markers rather than over-claiming a unique history [18,35,37,38,47,48].

5.4. What the Formalization Adds to Current Practice

The individual factors represented in the model are familiar to practising heritage scientists. Mineral assemblage, preparation, substrate contribution, alteration, and instrumental support are routinely considered in well-designed studies. The added value of the present formalization lies in making their relationships and assumptions explicit. It prevents a source-compatible chemical pattern from being treated automatically as a unique provenance result; it requires process and degradation alternatives to be considered before attribution; and it links the strength of the final claim to the analytical support actually available. The model is consequently a reporting and reasoning tool rather than a replacement for expert mineralogical interpretation.

5.5. Practical Diagnostic Workflow

Table 2 distinguishes between components specified conceptually, components partly implemented in the published-data applications, and components that still require calibration. Table 11 summarizes the claims that remain defensible under four common analytical conditions. It is decision guidance, not a calibrated accuracy table.

For a heritage case study, the model can be used through the following sequence:

1. Define the material question and the spatial or stratigraphic scale of the object;
2. Identify the mineral phases and accessory minerals as far as the evidence allows;
3. Record indicators of processing, including grain-size selection, heating, mixing, and binder incorporation;
4. Separate pigment, substrate, restoration, and degradation contributions before interpreting bulk chemistry;
5. Analyse compositional data in log-ratio coordinates and treat spectra as mixtures when appropriate;
6. Report source attribution as a graded compatibility statement and report degradation as support for broad exposure regimes, with explicit residual uncertainty.

This sequence translates the model into a practical diagnostic procedure. It also states when a conclusion should remain limited to screening-level compatibility rather than being presented as a unique source or exposure history.

This decision sequence does not eliminate interpretive complexity; it specifies the logic by which competing explanations are evaluated. The proposed method does not imply certainty; it distinguishes material, inferential, and observational sources of uncertainty. It also remains theoretical in important respects. Future operationalization will require curated geological reference collections, experimentally transformed analogues, and degradation datasets with known forcing histories [59]. Provenance models still need well-described training sets, careful treatment of below-detection values, and transparent prior choices, while inverse degradation models still need mineral-family response libraries built from controlled ageing experiments rather than retrospective interpretation. Standardised reporting of particle-size distributions, accessory phases, and treatment histories also remains uneven in pigment studies [6,34]. The proposed model is most directly applicable to natural

earth pigments and less directly applicable to complex modern synthetic systems, although parts of the same degradation logic can later be extended to them [60].

Table 11. Decision-oriented inferential guidance under four common analytical conditions. The table is a synthesis derived from the analytical construction and cited literature; it is not a statistically calibrated accuracy map.

Analytical Condition	What Can Be Inferred with Reasonable Confidence	What Remains Weak or Ambiguous	Recommended Claim Style
In situ pXRF only	Broad chemical compatibility classes, detection of ferric versus non-ferric systems, coarse substrate or contamination flags, and screening-level anomalies when reference data are available.	Mineral-family assignment, celadonite–glaucanite discrimination, unique provenance, and regime ranking remain weak because phase information and depth control are limited.	“Chemically compatible with”, “screening-level indication”, or “not inconsistent with” rather than direct identification.
Raman + XRD on microsamples	Phase-family assignment, detection of heating-related ferric transformations, structurally resolved celadonite/glaucanite discrimination, and constrained source compatibility when geology is already narrowed.	Whole-surface representativity, fine spatial heterogeneity, and exact environmental exposure history remain uncertain because the support is local and microsampled.	“Phase-resolved compatibility” or “process-consistent evidence” rather than unique source attribution.
HSI + FTIR	Spatial mapping of pigment families, mixture domains, hydration/binder signatures, and candidate alteration regimes across surfaces or motifs.	Source attribution on its own remains limited, and some phase assignments still require confirmatory microsampling or diffraction evidence.	“Spatially resolved candidate families” and “regime-sensitive markers” with explicit uncertainty about source class.
Multi-technique phase-resolved analysis	Strongest support for hierarchical provenance, process screening, and ranked environmental regimes by combining chemistry, phase, texture, and spatial information.	Under this condition, unique quarry assignment and exact environmental reconstruction remain rarely defensible; posterior support remains class-based and regime-based.	“Posterior support over source classes” and “compatible environmental regimes” with confidence levels.

6. Conclusions

Natural earth pigments are geological polymineral materials whose present state reflects source geology, preparation, application, and subsequent alteration. The model developed here separates those contributions into geological, technological, and degradation components and links them to the partial observations produced by heritage-science methods. Its purpose is to improve the wording and evidential basis of provenance and degradation claims, not to replace mineralogical expertise or to claim unique reconstruction from incomplete data.

Three applications clarify the practical scope. The Roussillon re-analysis confirms that phase–colour relations remain visible after compositionally appropriate transformation and shows that wet sieving can shift measured composition without necessarily removing family-level compatibility. The green-earth re-analysis shows that commercial labels and bulk chemistry alone do not reliably establish celadonite–glaucanite family membership; structural and vibrational evidence provide the stronger basis for classification. The Herculaneum application shows how multi-technique evidence can distinguish original red ochre from yellow ochre converted to red by eruption-related heating, while also demonstrating why colour alone is insufficient.

The numerical checks are small and provide internal consistency tests rather than external predictive validation. The degradation application is literature-based and does not calibrate the dynamic model. Future work requires larger geological reference collections,

controlled processing and ageing experiments, explicit priors and likelihoods, and independent tests on heritage objects. Within these limits, the study provides a practical sequence for separating source, process, substrate, and alteration contributions and for reporting source compatibility and environmental-regime support with appropriate uncertainty.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/heritage9080322/s1>, File S1 (Dubiel re-analysis dataset, CSV), File S2 (green-earth published-data check, CSV), File S3 (robustness summaries for the published-data checks, CSV), File S4 (green-earth concordance sensitivity summaries, CSV), File S5 (Python script regenerating Files S3 and S4 directly from Files S1 and S2 and regenerating the visualization shown in Figure 3), and File S6 (companion notebook reproducing the same outputs in executable form).

Author Contributions: Conceptualization, S.E.S.; methodology, S.E.S. and R.S.; formal analysis, S.E.S. and R.S.; writing—original draft preparation, S.E.S.; writing—review and editing, S.E.S. and R.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The study is based on published measurements and does not include a new laboratory campaign. Derived datasets and reproducibility materials constructed from numerical values reported in published sources are provided with this article as Supplementary Materials: File S1 (Dubiel re-analysis dataset, CSV), File S2 (green-earth published-data check, CSV), File S3 (robustness summaries for the published-data checks, CSV), File S4 (green-earth concordance sensitivity summaries, CSV), File S5 (Python script regenerating Files S3 and S4 directly from Files S1 and S2 and regenerating the visualization shown in Figure 3), and File S6 (companion notebook reproducing the same outputs in executable form). The ochre-derived values are also summarized in Table 6, the internal robustness summaries are reported in Table 9, and the corresponding visualization is shown in Figure 3. The calculation environment and counting rules are described in the reproducibility note in Section 2.8.

Acknowledgments: ChatGPT (OpenAI) was used only for language refinement and draft preparation. The authors reviewed and edited the AI-assisted text as needed and take full responsibility for the content of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

CIE	International Commission on Illumination
clr	centred log-ratio
CSV	comma-separated values
FTIR	Fourier-transform infrared spectroscopy
HSI	hyperspectral imaging
ilr	isometric log-ratio
IR	infrared
LDA	linear discriminant analysis
NIR	near-infrared
PARAFAC	parallel factor analysis
PCA	principal component analysis
pXRF	portable X-ray fluorescence
RH	relative humidity
SEM-EDS	scanning electron microscopy with energy-dispersive spectroscopy
UV	ultraviolet

XRD	X-ray diffraction
XRF	X-ray fluorescence

References

- Feller, R.L. (Ed.) *Artists' Pigments: A Handbook of Their History and Characteristics*; National Gallery of Art: Washington, DC, USA; Cambridge University Press: Cambridge, UK, 1986; Volume 1.
- Eastaugh, N.; Walsh, V.; Chaplin, T.; Siddall, R. *Pigment Compendium: A Dictionary of Historical Pigments*; Elsevier Butterworth-Heinemann: Oxford, UK, 2004; ISBN 978-0-7506-5749-5.
- Harth, A. The Study of Pigments in Cultural Heritage: A Review Using Machine Learning. *Heritage* **2024**, *7*, 3664–3695. [[CrossRef](#)]
- Siddall, R. Mineral Pigments in Archaeology: Their Analysis and the Range of Available Materials. *Minerals* **2018**, *8*, 201. [[CrossRef](#)]
- Dayet, L. Invasive and Non-Invasive Analyses of Ochre and Iron-Based Pigment Raw Materials: A Methodological Perspective. *Minerals* **2021**, *11*, 210. [[CrossRef](#)]
- Ma, C.; Dou, H.; Zhao, Z.; Qiu, X.; Li, H.; Wang, X. Review of In-Situ Non- and Micro-Destructive Techniques for Pigment Analysis in Architectural Heritage. *npj Herit. Sci.* **2025**, *13*, 222. [[CrossRef](#)]
- Vezzoli, L. Geological Contributions to the History of the Artist's Iron-Based Natural Earth Pigments and the Case Study of Terra d'Ombra (Umbra). *Colorants* **2026**, *5*, 2. [[CrossRef](#)]
- Cortea, I.M.; Ghervase, L.; Rădvan, R.; Seritan, G. Assessment of Easily Accessible Spectroscopic Techniques Coupled with Multivariate Analysis for the Qualitative Characterization and Differentiation of Earth Pigments of Various Provenance. *Minerals* **2022**, *12*, 755. [[CrossRef](#)]
- Marketou, A.K.; Kouzeli, K.; Facorellis, Y. Colourful Earth: Iron-Containing Pigments from the Hellenistic Pigment Production Site of the Ancient Agora of Kos (Greece). *J. Archaeol. Sci. Rep.* **2019**, *26*, 101876. [[CrossRef](#)]
- Moretto, L.M.; Orsega, E.F.; Mazzocchin, G.A. Spectroscopic Methods for the Analysis of Celadonite and Glauconite in Roman Green Wall Paintings. *J. Cult. Herit.* **2011**, *12*, 384–391. [[CrossRef](#)]
- Pérez-Rodríguez, J.L.; Jiménez de Haro, M.C.; Sigüenza, B.; Martínez-Blanes, J.M. Green Pigments of Roman Mural Paintings from Seville Alcazar. *Appl. Clay Sci.* **2015**, *116–117*, 211–219. [[CrossRef](#)]
- Singh, P.; Banerjee, S.; Roy Choudhury, T.; Bhattacharya, S.; Pande, K. Distinguishing Celadonite from Glauconite for Environmental Interpretations: A Review. *J. Palaeogeogr.* **2023**, *12*, 179–194. [[CrossRef](#)]
- Elias, M.; Chartier, C.; Prévot, G.; Garay, H.; Vignaud, C. The Colour of Ochres Explained by Their Composition. *Mater. Sci. Eng. B* **2006**, *127*, 70–80. [[CrossRef](#)]
- Dos Santos, I.F.S.; Edwards, H.G.M.; de Faria, D.L.A. Hematite Colour Revisited: Particle Size and Electronic Transitions. *Spectrochim. Acta A* **2024**, *310*, 123810. [[CrossRef](#)] [[PubMed](#)]
- Secco, M.; Rainer, L.; Graves, K.; Heginbotham, A.; Artioli, G.; Piqué, F.; Angelini, I. Ochre-Based Pigments in the Tablinum of the House of the Bicentenary (Herculaneum, Italy) between Decorative Technology and Natural Disasters. *Minerals* **2021**, *11*, 67. [[CrossRef](#)]
- Gomes, H.; Collado, H.; Garcês, S.; Lattao, V.; Nicoli, M.; Eftekhari, N.; Marrocchino, E.; Rosina, P. Archeometrical Characterization of Rock Art Pigments from Puerto Roque Open-Air Rock Art Shelter (Valencia de Alcántara, Extremadura, Spain). *Heritage* **2024**, *7*, 1123–1139. [[CrossRef](#)]
- Coccatto, A.; Moens, L.; Vandenabeele, P. On the Stability of Mediaeval Inorganic Pigments: A Literature Review of the Effect of Climate, Material Selection, Biological Activity, Analysis and Conservation Treatments. *Herit. Sci.* **2017**, *5*, 12. [[CrossRef](#)]
- Frasca, F.; Verticchio, E.; Bosco, E.; Kuka, E.; Lee, D.S.-H.; Andersen, C.K.; Bertolin, C.; Siani, A.M. Assessing Microclimate Thresholds for Heritage Preventive Conservation to Achieve Sustainable and Energy Efficiency Goals in a Changing Climate. *Sci. Rep.* **2024**, *14*, 18707. [[CrossRef](#)] [[PubMed](#)]
- Camuffo, D. *Microclimate for Cultural Heritage: Measurement, Risk Assessment, Conservation, Restoration, and Maintenance of Indoor and Outdoor Monuments*, 3rd ed.; Elsevier: Amsterdam, The Netherlands, 2019; ISBN 978-0-444-64106-9. [[CrossRef](#)]
- Valero, E.M.; Martínez-Domingo, M.Á.; López-Bal-domero, A.B.; López-Montes, A.; Abad-Muñoz, D.; Vílchez-Quero, J.L. Unmixing and Pigment Identification Using Visible and Short-Wavelength Infrared: Reflectance vs. Logarithm Reflectance Hyperspaces. *J. Cult. Herit.* **2023**, *64*, 290–300. [[CrossRef](#)]
- Martín-Ramos, P.; Cuchí-Oterino, J.A.; Bea-Martínez, M. Revisiting Albarracín Rock Art Through Multivariate pXRF Analysis of White, Black, and Red Pigments. *Minerals* **2025**, *15*, 1328. [[CrossRef](#)]
- Pacífico, L.R.; Guarino, A.; Iannone, A.; Albanese, S. Accounting for the Compositional Nature of Geochemical Data to Improve the Interpretation of Their Univariate and Multivariate Spatial Patterns: A Case Study from the Campania Region (Italy). *Geosciences* **2025**, *15*, 20. [[CrossRef](#)]
- Vermeesch, P. Exploratory Analysis of Provenance Data Using R and the Provenance Package. *Minerals* **2019**, *9*, 193. [[CrossRef](#)]
- Aitchison, J. *The Statistical Analysis of Compositional Data*; Chapman & Hall: London, UK, 1986.

25. Egozcue, J.J.; Pawlowsky-Glahn, V.; Mateu-Figueras, G.; Barceló-Vidal, C. Isometric Logratio Transformations for Compositional Data Analysis. *Math. Geol.* **2003**, *35*, 279–300. [\[CrossRef\]](#)
26. Martín-Fernández, J.A.; Barceló-Vidal, C.; Pawlowsky-Glahn, V. Dealing with Zeros and Missing Values in Compositional Data Sets Using Nonparametric Imputation. *Math. Geol.* **2003**, *35*, 253–278. [\[CrossRef\]](#)
27. Pawlowsky-Glahn, V.; Egozcue, J.J.; Tolosana-Delgado, R. *Modeling and Analysis of Compositional Data*; John Wiley & Sons: Chichester, UK, 2015. [\[CrossRef\]](#)
28. Ciortan, I.M.; Poulsson, T.G.; George, S.; Hardeberg, J.Y. Tensor Decomposition for Painting Analysis. Part 1: Pigment Characterization. *Herit. Sci.* **2023**, *11*, 76. [\[CrossRef\]](#)
29. Ciortan, I.M.; Poulsson, T.G.; George, S.; Hardeberg, J.Y. Tensor Decomposition for Painting Analysis. Part 2: Spatio-Temporal Simulation. *Herit. Sci.* **2023**, *11*, 84. [\[CrossRef\]](#)
30. Liu, J.; Li, T.; Yu, Z.; Lyu, S.; Fu, Q.; Yu, X. Hyperspectral Identification of Mineral Pigments in Thangka Paintings for Cultural Heritage Conservation. *npj Herit. Sci.* **2025**, *13*, 649. [\[CrossRef\]](#)
31. Attard Montalto, N.; Shortland, A.J.; Rogers, K.D. The Provenancing of Ochres from the Neolithic Temple Period in Malta. *J. Archaeol. Sci.* **2012**, *39*, 1094–1102. [\[CrossRef\]](#)
32. Dayet, L. Defining Iron-Bearing Rock Sources in Compositional Data Statistical Analysis for Provenance Studies: A Case Study from South Africa. *J. Archaeol. Sci. Rep.* **2023**, *49*, 104010. [\[CrossRef\]](#)
33. Augustsson, C.; Aehnelt, M.; Olivarius, M.; Voigt, T.; Gaupp, R.; Hilse, U. Provenance from the Geochemical Composition of Terrestrial Clastic Deposits—A Review with Case Study from the Intracontinental Permo-Triassic of European Pangea. *Sediment. Geol.* **2023**, *456*, 106496. [\[CrossRef\]](#)
34. Marketou, A.; Rodler, A. Ochres as Earth Pigments in Hellenistic and Roman Polychromy: State of the Art and Future Directions on Material Origins and Trade. *J. Archaeol. Sci.* **2025**, *180*, 106276. [\[CrossRef\]](#)
35. Fabbri, K.; Bonora, A. Two New Indices for Preventive Conservation of the Cultural Heritage: Predicted Risk of Damage and Heritage Microclimate Risk. *J. Cult. Herit.* **2021**, *47*, 208–217. [\[CrossRef\]](#)
36. Li, W.; Bai, X.; Pan, Z. Air Pollution and Its Impact on Stone Heritage Sites: Mechanisms and Mathematical Models. *npj Herit. Sci.* **2025**, *13*, 247. [\[CrossRef\]](#)
37. Ye, Z.-S.; Chen, N. The Inverse Gaussian Process as a Degradation Model. *Technometrics* **2014**, *56*, 302–311. [\[CrossRef\]](#)
38. Voigt, T.; Kohlhase, M.; Nelles, O. Incremental DoE and Modeling Methodology with Gaussian Process Regression: An Industrially Applicable Approach to Incorporate Expert Knowledge. *Mathematics* **2021**, *9*, 2479. [\[CrossRef\]](#)
39. Charola, A.E. Salts in the Deterioration of Porous Materials: An Overview. *J. Am. Inst. Conserv.* **2000**, *39*, 327–343. [\[CrossRef\]](#)
40. Oguchi, C.T.; Yu, S. A Review of Theoretical Salt Weathering Studies for Stone Heritage. *Prog. Earth Planet. Sci.* **2021**, *8*, 32. [\[CrossRef\]](#)
41. Maura, G.; Caron, B.; Beck, L.; Salomon, H.; Pichon, L.; Lemasson, Q.; Chalmin, E.; MacDonald, B.L. Standardization Procedure to Provide a Unified Multi-Method Elemental Compositional Dataset, Application to Ferruginous Colouring Matters from Namibia. *J. Archaeol. Sci. Rep.* **2022**, *43*, 103454. [\[CrossRef\]](#)
42. Chanteraud, C.; Salomon, H.; Chalmin, E.; Lemasson, Q.; Pacheco, C.; Pichon, L.; Goemaere, E.; Noûs, C.; MacDonald, B.L. Why Is It So Difficult to Work on Geochemical Composition? Supervised Geochemical Composition Data Processing to Study Colouring Iron Oxide-Rich Rocks in Archaeological Contexts. *J. Archaeol. Sci.* **2025**, *183*, 106277. [\[CrossRef\]](#)
43. Dubiel, S.M.; Cieślak, J.; Tarasiuk, J.; Nizioł, J. Relationship between the Colour of Ochre from Roussillon and the Content of Iron-Bearing Minerals. *Appl. Clay Sci.* **2011**, *51*, 54–60. [\[CrossRef\]](#)
44. González, G.; Sagarzazu, A.; Villalba, R. Study of the Mechano-Chemical Transformation of Goethite to Hematite by TEM and XRD. *Mater. Res. Bull.* **2000**, *35*, 2295–2308. [\[CrossRef\]](#)
45. Ospitali, F.; Bersani, D.; Di Leonardo, G.; Lottici, P.P. “Green Earths”: Vibrational and Elemental Characterization of Glauconites, Celadonites and Historical Pigments. *J. Raman Spectrosc.* **2008**, *39*, 1066–1073. [\[CrossRef\]](#)
46. Fanost, A.; Gimat, A.; de Viguerie, L.; Martinetto, P.; Giot, A.-C.; Clémancey, M.; Blondin, G.; Gaslain, F.; Glanville, H.; Walter, P.; et al. Revisiting the Identification of Commercial and Historical Green Earth Pigments. *Colloids Surf. A Physicochem. Eng. Asp.* **2020**, *584*, 124035. [\[CrossRef\]](#)
47. Lucchi, E. Review of Preventive Conservation in Museum Buildings. *J. Cult. Herit.* **2018**, *29*, 180–193. [\[CrossRef\]](#)
48. Varas-Muriel, M.J.; Fort, R.; Martínez-Garrido, M.I.; Zornoza-Indart, A.; López-Arce, P. Microclimatic Monitoring in an Historic Church Fitted with Modern Heating: Implications for the Preventive Conservation of Its Cultural Heritage. *Build. Environ.* **2018**, *145*, 290–307. [\[CrossRef\]](#)
49. Shugar, A.N. Portable X-ray Fluorescence and Archaeology: Limitations of the Instrument and Suggested Methods To Achieve Desired Results. In *Archaeological Chemistry VIII*; Armitage, R.A., Burton, J.H., Eds.; ACS Symposium Series; American Chemical Society: Washington, DC, USA, 2013; pp. 173–193. [\[CrossRef\]](#)
50. Shrimali, K.; Jin, J.; Vaziri Hassas, B.; Wang, X.; Miller, J.D. The Surface State of Hematite and Its Wetting Characteristics. *J. Colloid Interface Sci.* **2016**, *477*, 16–24. [\[CrossRef\]](#) [\[PubMed\]](#)

51. Barrón, V.; Torrent, J. Surface Hydroxyl Configuration of Various Crystal Faces of Hematite and Goethite. *J. Colloid Interface Sci.* **1996**, *177*, 407–410. [[CrossRef](#)]
52. Reigl, S.; King, H.E.; Stawski, T.M.; Benning, L.G. Revisiting the Roles of Salinity, Temperature and Water Activity in Calcium Sulfate Crystallisation. *CrystEngComm* **2022**, *24*, 1808–1822. [[CrossRef](#)]
53. Sitzia, F.; Moita, P.; Bottura Scardina, S.; Lisci, C. Colour Stability of Light-Sensitive Minerals Under UVA (340 nm) Irradiation: Implications for the Conservation of Cultural Heritage and Museum Display Conditions. *Minerals* **2025**, *15*, 999. [[CrossRef](#)]
54. Nastou, M.P.P.; Zerefos, S.C. Effects of Climate Change on Open Air Heritage: A Review and the Situation in the Region of Mediterranean. *Herit. Sci.* **2024**, *12*, 367. [[CrossRef](#)]
55. Visone, F.; Abate, N.; Sileo, M.; Masini, N. Assessing Stone Material Recession of Cultural Heritage: New Approach Based on Satellite-Based Rainfall Data and Dose-Response Functions—Case of UNESCO Site of Matera. *Remote Sens.* **2025**, *17*, 1443. [[CrossRef](#)]
56. Cavallo, G.; Fontana, F.; Gialanella, S.; Gonzato, F.; Riccardi, M.P.; Zorzin, R.; Peresani, M. Heat Treatment of Mineral Pigment During the Upper Palaeolithic in North-East Italy. *Archaeometry* **2018**, *60*, 1045–1061. [[CrossRef](#)]
57. de Faria, D.L.A.; Lopes, F.N. Heated Goethite and Natural Hematite: Can Raman Spectroscopy Be Used to Differentiate Them? *Vib. Spectrosc.* **2007**, *45*, 117–121. [[CrossRef](#)]
58. Goldberg, M.C.; Weiner, E.R.; Boymel, P.M. Adsorption of Goethite onto Quartz and Kaolinite. *J. Chem. Soc., Faraday Trans. 1* **1984**, *80*, 1491–1498. [[CrossRef](#)]
59. Salomon, H.; Chanteraud, C.; Chassin de Kergommeaux, A.; Monney, J.; Pradeau, J.-V.; Goemaere, E.; Coquinot, Y.; Chalmin, E. A Geological Collection and Methodology for Tracing the Provenance of Palaeolithic Colouring Materials. *J. Lithic Stud.* **2021**, *8*, 1–38. [[CrossRef](#)]
60. Costantini, R.; Nodari, L.; La Nasa, J.; Tomasin, P. Relative Humidity, Light, and Inorganic Fillers: Defining Different Roles on the Ageing of Modern Oil Paints through Infrared Spectroscopy. *Microchem. J.* **2025**, *208*, 112534. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.