Reconstructing Holocene ENSO Variability through the Harmonization of Heterogeneous Proxy Records using a Gaussian Process Regression Framework

Kaya de Jong (s3962431)

Supervisors:
Dr. S.D. Fajardo Bernal & Dr. A. Saxena

Bachelor Thesis– Data Science and Artificial Intelligence

Leiden Institute of Advanced Computer Science (LIACS)
liacs.leidenuniv.nl

July 17, 2026

Abstract

The El Niño-Southern Oscillation (ENSO) is one of the dominant sources of interannual climate variability and influences climate systems worldwide. Understanding how ENSO has changed over time is important for improving knowledge of long-term climate variability. Since instrumental observations only cover a limited time period, paleoclimate proxy records are required to investigate long-term changes in ENSO behavior. However, substantial differences in temporal coverage, sampling density and proxy type make it hard to combine heterogeneous proxy archives into a coherent reconstruction.

This study investigates whether heterogeneous ENSO-sensitive proxy records can be harmonized into a coherent long-term reconstruction of ENSO variability. Multiple proxy datasets, including lake sediments, reconstructed sea surface temperatures, tree rings, corals, speleothems and biomarker records, were preprocessed with chronology and signal standardization, variability extraction and temporal binning. A composite signal was constructed and Gaussian Process Regression (GPR) was used to generate a continuous reconstruction with uncertainty estimations.

The resulting reconstruction shows moderate agreement with a Niño 3.4 reconstruction, achieving a Pearson correlation of 0.692 and a Spearman correlation of 0.593. The reconstruction also suggests that ENSO variability was not constant during the last 2000 years, with enhanced variability between 1500 and 1000 years BP. Sensitivity analyses show that reconstruction performance remains robust to changes in preprocessing and kernel configurations but remains strongly influenced by the selected proxy records.

These findings suggest that heterogeneous ENSO-sensitive proxy records can be harmonized into a coherent reconstruction of ENSO variability. While the resulting reconstruction remains sensitive to the selection of proxies and the uncertainty that paleoclimatic data brings, the proposed framework demonstrates that meaningful ENSO-related variability can be extracted by combining diverse proxy archives.

Contents

1 Introduction

Climate variability has been a persistent feature of the Earth system throughout both the present and the past. One of the dominant sources of such variability is the El Niño-Southern Oscillation (ENSO), a coupled ocean-atmosphere climate phenomenon originating in the Pacific Ocean that represents one of the most important drivers of interannual climate variability worldwide [MZG06, TAK+18]. ENSO influences precipitation, drought occurrence, temperature variability and extreme weather events across large parts of the world [LFZ22, AANR04]. ENSO variability has also been linked to impacts on ecosystems, fisheries and vegetation [RFCMMC+26, Le23].

Because ENSO has influenced the Earth's climate for thousands of years, understanding its long-term behavior is important for improving knowledge of past climate variability and the evolution of tropical climate systems. It may also provide insight into how climate variability has influenced environmental conditions, ecosystems, and human societies throughout history. However, reconstructing long-term ENSO variability does have its challenges. Instrumental climate observations only cover a relatively limited time span of ENSO activity [LXC+11]. In addition, ENSO variability itself is complex and influenced by multiple oceanic and atmospheric processes [TAK+18].

To prolong ENSO reconstruction beyond the instrumental period, paleoclimate proxy records such as tree-rings, corals, speleothems and sediment records are commonly used. Proxy records are indirect climate archives that preserve environmental information related to past climate conditions [Bra99]. Nevertheless, individual records differ substantially in their temporal coverage, sampling resolution, uncertainty and climate interpretation. These differences make direct comparison and integration of heterogeneous proxy records challenging.

When reconstructing past ENSO-related climate variability, relying on a single proxy record can provide an incomplete representation of long-term ENSO behavior, as individual archives only capture localized or partial representations of ENSO-related variability. Combining multiple heterogeneous proxy records may improve the robustness of the reconstruction by extracting a shared ENSO-related variability signal across different proxy types and geographical regions. However, this is challenging since proxy records are heterogeneous and there is currently no standardized harmonization framework for consistently combining ENSO-sensitive proxy records into a single reconstruction.

These challenges motivate the need for a systematic approach that can integrate heterogeneous proxy records despite their differences in structure, resolution and climatic sensitivity. This study therefore investigates whether such records can be harmonized and combined into a coherent long-term reconstruction of ENSO variability.

To investigate this, multiple heterogeneous proxy records are preprocessed and harmonized before being combined into a composite ENSO-related variability signal. Due to limitations in the availability in high-quality proxy records with sufficient temporal coverage, this study focuses primarily on the last 2000 years BP, where proxy availability and overlap are high. Gaussian Process Regression (GPR) is then applied to transform the noisy and irregular composite proxy signal into a continuous reconstruction with quantified uncertainty. The objective of this study is therefore to evaluate whether shared ENSO-related variability can be extracted from heterogeneous proxy records using this methodology to produce a meaningful long-term reconstruction of ENSO variability.

1.1 Research Questions

The main research Question that is addressed in this study is:

How can heterogeneous Holocene ENSO-sensitive proxy records be harmonized to produce a coherent long-term reconstruction of ENSO variability?

The objective of this research is to evaluate whether heterogeneous ENSO-sensitive proxy records can be systematically harmonized to extract a shared ENSO-related variability signal. By combining multiple proxy records with different characteristics, this study aims to assess whether a coherent long-term reconstruction of ENSO variability can be produced despite differences in temporal coverage, sampling resolution and uncertainty.

It is hypothesized that shared ENSO-related variability can be extracted from heterogeneous proxy records through preprocessing, harmonization and composite reconstruction techniques. However, the resulting reconstruction of ENSO variability is expected to remain sensitive to proxy selection and sampling resolution. In particular, periods with limited proxy overlap are expected to show greater reconstruction uncertainty, as fewer observations are available to constrain the Gaussian Process Regression model. Furthermore, individual proxy records are expected to contribute differently to the reconstruction because of differences in climatic sensitivity, temporal coverage and data availability.

In addition to the main research question, this study addresses the following subquestions:

1. **To what extent do heterogeneous Holocene ENSO-sensitive proxy records differ in their representation of long-term ENSO-related variability?**
2. **What are the strengths and limitations of Gaussian Process Regression when applied to heterogeneous Holocene ENSO reconstruction time series based on sensitivity analysis of kernel choice, length scale and parameter optimization?**

The first subquestion focuses on discussing and identifying the differences between the heterogeneous proxy records and how these differences influence the representation of ENSO-related variability. The second sub question evaluates the applicability of Gaussian Process Regression for transforming noisy and irregular proxy signals into a more continuous reconstruction with quantified uncertainty. To assess its strength and limitations, the influence of kernel choice, length scale and parameter optimization on reconstruction performance is systematically examined.

1.2 Thesis overview

The remainder of the thesis is structured as follows: Chapter 2 introduces the theoretical background related to ENSO, Paleoclimate Proxy Records and Gaussian Process Regression. Chapter 3 discusses related work on ENSO reconstruction, multi-proxy reconstruction approaches and Gaussian Processes in Paleoclimate reconstructions. Chapter 4 present the heterogeneous proxy datasets used throughout this study. Chapter 5 describes the preprocessing, harmonization, composite construction, and reconstruction methodology. Chapter 6 presents the reconstruction results and evaluation. Chapter 7 discusses the findings, limitation and implications of the study, followed by the conclusions and future work in Chapter 8

2 Background

This chapter provides the theoretical background necessary for understanding the framework developed in this thesis. First, the El Niño Southern Oscillation (ENSO) and its climatic significance are discussed. Afterwards, paleoclimate proxy records and their heterogeneity are introduced, followed by an overview of Gaussian Process Regression (GPR) and its relevance for the reconstruction.

2.1 El Niño Southern Oscillation

The El Niño Southern Oscillation (ENSO) is a climate phenomenon involving coupled atmospheric and oceanic processes occurring within the tropical Pacific Ocean [TAK+18]. The ENSO system is associated with alternating warm and cool sea surface temperature anomalies, commonly referred to as El Niño and La Niña. These phases influence atmospheric circulation patterns and contribute substantially to interannual climate variability on a global scale.

As discussed in Chapter 1, ENSO has widespread climatic and environmental impacts across the globe. These impacts motivate efforts to understand how ENSO variability changes through time and can be reconstructed over longer timescales.

However, studying long-term ENSO variability directly over a broad timeframe remains difficult. Modern ENSO observations primarily rely on instrumental measurements such as sea surface temperature records, atmospheric pressure observations and trade wind variability. Despite this, these instrumental records are often too short to characterize the natural variability of ENSO [LXC+11]. This limits their ability to reconstruct or assess ENSO variability over longer climatic timescales such as the past several millennia. Additionally, ENSO behavior is highly variable and irregular through time, making it difficult for short instrumental records to fully capture the long-term nature of ENSO variability. Therefore, instrumental records alone are not suitable for reconstructing long-term ENSO variability over extended timescales.

2.2 Paleoclimate Proxy Records

To overcome the limitations of instrumental observations, paleoclimate proxy records are commonly used to reconstruct past climate variability over a longer timescale. Paleoclimate proxy records are indirect climate archives that preserve information about past environmental and climate conditions [Bra99]. Unlike instrumental observations, proxy records do not directly measure climate variables. Instead, climate information is collected from physical, chemical or biological characteristics preserved within natural archives.

Paleoclimate proxy records exist in many different forms. Common examples include tree rings, where variations in the width or density of the ring can reflect climate conditions such as temperature and humidity. Coral records contain geochemical signals related to sea surface temperature variability, while speleothems preserve isotopic information related to temperature and water chemistry [Bra99]. Other examples include marine sediments, which preserve environmental deposits that can reflect the past oceanic and climatic conditions, and ice cores, whose geochemistry and physical properties contain information about past atmospheric composition, temperature variability, and sea level changes [Bra99]. These proxy records therefore preserve climate information from many different environmental systems and can provide insight into past climate variability over timescales far beyond instrumental observations.

Many paleoclimate proxy records are sensitive to climate changes associated with ENSO variability. Consequently, ENSO-related changes in sea surface temperature, atmospheric circulation patterns and hydroclimatic conditions can be preserved within these proxy archives, making them valuable sources for reconstructing past ENSO variability.

However, paleoclimate proxy records also present significant challenges when attempting to extract a coherent climate signal from multiple proxy records. First of all, paleoclimate proxy records are highly heterogeneous, as different proxy types preserve climatic information through different physical, chemical, and biological processes. The heterogeneity makes it challenging to directly compare and combine climatic signals extracted from different natural archives. Secondly, the temporal resolution of proxy records can differ substantially, meaning that proxy archives do not always preserve climatic information over the same timescale or over the same periods.

Additionally, proxy records can vary substantially in sampling density and temporal continuity. Some proxy archives contain annually resolved observations, while others only preserve climatic information at irregular intervals. Furthermore, many proxy datasets contain missing observations or large temporal gaps, complicating direct comparison and alignment between proxy records. Finally, proxy records are often both influenced by global and local climatic conditions. As a result, even similar proxy types may preserve different climatic signals depending on their geographic location and environment. Together, these challenges complicate the extraction of a coherent ENSO-related signal from multiple proxy archives. As a result, preprocessing and harmonization approaches are often required to make heterogeneous proxy datasets more comparable and suitable for combined climatic reconstructions.

Within paleoclimate studies, harmonization generally refers to the process of transforming heterogeneous proxy datasets into a more consistent and comparable form [PDFB21]. This can include standardizing temporal resolution, handling missing values, reducing long-term trends, aligning chronologies and normalizing their values. Harmonization is often necessary because proxy records may differ substantially in their characteristics. For example, chronological inconsistencies can arise due to dating uncertainties or differences in age models, making direct temporal alignment difficult. In addition, temporal resolution may vary considerably, with some proxy records providing annual observations while others represent multi-year or decadal averages, requiring resampling to a common temporal framework. Proxy records may also differ in their measurement scales and units, requiring standardization before meaningful comparisons can be made. Furthermore, different proxy archives may respond differently to the same climate process, resulting in differences in signal direction and climate sensitivities. Addressing these differences helps make heterogeneous proxy datasets more comparable and suitable for combined climatic reconstruction and signal extraction. After preprocessing and harmonization, a continuous climatic signal and estimated long-term climate variability can be extracted from heterogeneous proxy datasets using statistical reconstruction methods. The statistical reconstruction method used in this thesis is Gaussian Process Regression.

2.3 Gaussian Process Regression

Before discussing Gaussian Process Regression (GPR), it is useful to first introduce Gaussian Processes. A Gaussian Process is a stochastic process defined by a mean function and a covariance function [WR06]. A Gaussian Process is commonly represented as:

$$f(x) \sim GP(m(x), k(x, x'))$$

Here, $f(x)$ represents the function being modeled, which is assumed to be distributed according to a Gaussian Process defined by a mean function $m(x)$ and a covariance function $k(x, x')$ [WR06].

The mean function $m(x)$ represents the expected value of the Gaussian Process at each input location and defines the prior average behavior of the modeled signal.

The covariance function $k(x, x')$, commonly referred to as the kernel, represents the relationship between observations at different input locations. It determines how strongly observations are correlated with one another within the modeled signal. As a result, the covariance function controls important characteristics of the reconstruction, such as smoothness, variability, and the influence of observations on nearby predictions. Different kernels can be used depending on the characteristics of the modeled data. Common examples of kernels include the Radial Basis Function (RBF) kernel, which assumes smooth and continuous variability, the Matérn kernel, which allows for more irregular and less smooth variations within the reconstructed signal, and the Rational Quadratic (RQ) kernel, which can be interpreted as a combination of RBF kernels operating on different length scales.

All of these kernels contain a length-scale parameter, which determines the scale over which observations are considered related to one another. Larger length-scale values produce smoother reconstructions, whereas smaller values allow more localized variability. The Matérn kernel additionally contains a smoothness parameter (ν), which controls how rough the reconstructed signal is. Lower values of ν allow more irregular variability, whereas higher values produce smoother reconstructions. The RQ kernel contains an additional scale-mixture parameter (α), which controls the contribution of different length scales to the reconstruction. Smaller values of α allow greater variability across scales, whereas larger values causes the kernel to behave more similarly to the RBF kernel.

As can be seen, a Gaussian Process defines a probability distribution over functions that could describe the underlying data-generating process. Gaussian Process Regression uses this probabilistic framework together with observed data to estimate a reconstruction and its uncertainty [WR06]. Given a set of observed data points, GPR updates the prior Gaussian Process to obtain a posterior distribution conditioned on the observations. Predictions at unobserved locations are obtained by exploiting the correlations defined by the kernel, allowing nearby or statistically similar observations to influence the estimated value. The strength of this influence depends on the covariance between observations and the prediction location. Higher covariance values indicate a stronger relationship between observations and the prediction location, causing observations to have greater influence on the predicted value.

Because paleoclimate proxy datasets are often heterogeneous, noisy and irregularly sampled, GPR provides a suitable framework for reconstructing long-term variability from proxy records. The probabilistic nature of GPR also makes it possible to estimate uncertainty within the reconstruction, which is particularly important when working with paleoclimate datasets containing observational and chronological uncertainty.

3 Related Work

A substantial amount of research has already been conducted on reconstructing past ENSO variability using paleoclimate proxy records. Previous studies have applied a wide range of proxy archives, reconstruction methodologies and statistical techniques to investigate ENSO behavior across different timescales. This section will provide an overview of existing ENSO reconstruction studies, multi-proxy reconstruction approaches and the application of Gaussian Processes within paleoclimate studies.

3.1 Proxy-Based ENSO Reconstructions

A large number of studies have attempted to reconstruct past ENSO variability using paleoclimate proxy records. Because instrumental observations only cover a relatively short period, proxy archives are an important source of information over longer climate timescales. Various proxy types, including tree rings, coral records and speleothems, have been used to reconstruct past ENSO variability.

One example is the work of Li et al. [LXC+11], who reconstructed ENSO variability over the past 1,100 years using tree-ring based proxy records from the North American Drought Atlas. Their annually resolved reconstruction revealed substantial variability in ENSO amplitude through time. The results also suggested a quasi-periodic modulation occurring on timescales of approximately 50-90 years. The study demonstrated the potential of proxy records to extend ENSO reconstructions far beyond the instrumental record.

Similarly, Atwood and Sachs [AS14] reconstructed hydroclimatic variability in the Galápagos over the past 3,000 years using hydrogen isotopes from multiple biomarkers. Their study demonstrated that biomarker-based proxy records can preserve ENSO-related rainfall variability and provide insights into long-term climatic changes.

Bernal et al. [BLM+11] presented a Holocene climate reconstruction based on variations in $\delta^{18}O$ preserved within speleothem records from southwestern Mexico. Their reconstruction suggested that regional moisture variability was influenced by multiple large-scale climatic processes, including changes in ENSO-related variability. These findings highlight the potential of speleothem proxy archives for reconstructing long-term hydroclimatic variability over the Holocene.

Together, these studies demonstrate that a wide variety of paleoclimate proxy archives can be used to reconstruct past ENSO-related climate variability. Tree-rings, biomarkers and speleothems each preserve climatic information through different environmental processes and provide valuable insights into past climate conditions. However, these proxies differ in their climatic sensitivity, temporal resolution and geographic location. Consequently, individual reconstructions might be influence by proxy-specific uncertainties and regional environmental conditions. This raises the question of whether proxy records that differ substantially in their characteristic can be combined into a robust representation of long-term ENSO variability. These limitations have motivated the development of multi-proxy reconstruction approaches that combine information from multiple proxy archives.

3.2 Multi-Proxy Reconstruction and Harmonization

To address the limitations associated with individual proxy records, researchers have increasingly explored multi-proxy reconstruction frameworks that combine information from multiple proxy archives. By integrating different proxy records, these approaches seek to reduce the influence of proxy-specific uncertainties and improve the robustness of the climate reconstruction. An important question, however, is whether combining multiple proxy records actually results in a more reliable climate reconstruction than approaches based on individual proxy records.

Li et al. [LNA10] investigated the value of multi-proxy reconstructions by developing a Bayesian framework capable of integrating multiple proxy records with different temporal resolutions. Their framework succeeded in combining heterogeneous proxy records into a coherent climate reconstruction and demonstrated that incorporating information from multiple proxy archives can improve the quality and robustness of climate reconstructions compared to approaches based on individual proxy records.

More recently, Liu et al. [LMZZ24] developed a global multi-proxy ENSO reconstruction using tree-ring and speleothems $\delta^{18}O$ records. Their reconstruction demonstrated that combining multiple proxy archives can successfully capture long-term ENSO variability and showed high consistency with existing ENSO reconstructions. However, although the proxy archives differ, all selected records were based on the same isotopic signal ($\delta^{18}O$), resulting in a more homogeneous proxy network than studies that combine fundamentally different proxy signals.

Although both studies demonstrated the potential of multi-proxy reconstructions, they differ in the degree of proxy heterogeneity considerd. The study of Li focused on combining proxy records with different temporal resolution, whereas the study of Liu reconstructed ENSO using multiple archives that were based on the same isotopic signals. As a result, it remains unclear whether reconstruction frameworks can successfully integrate proxy records that differ fundamentally in both proxy type and temporal coverage.

As demonstrated by these studies, multi-proxy reconstructions offer several advantages for reconstructing long-term climate variability. But despite these advantages, combining multiple proxy records remains challenging. Different proxy archives often differ substantially in temporal resolution, sampling density, climatic sensitivity and the mechanisms through which climatic signal are preserved. These differences can complicate direct comparison and integration of heterogeneous proxy records, making it difficult to develop a consistent reconstruction framework.

Roberts et al. [RTH+19] highlighted the challenges associated with combining proxy records that differ in temporal resolution and sampling density. To address these issues, they proposed a reconstruction framework designed to account for unevenly sampled paleoclimate records, allowing heterogeneous proxy datasets to be more effectively integrated within a common reconstruction. Their work highlights the importance of reconciling differences between proxy records before they can be reliably combined.

Similarly, Pfalz et al. [PDFB21] emphasized the importance of harmonization when integrating heterogeneous paleoclimate datasets. Their study demonstrated that proxy records often differ substantially in structure, format, temporal resolution and sampling characteristics, making direct comparison difficult. To address these challenges, they transformed heterogeneous proxy datasets into a common framework through the use of standardization and synchronization procedures. Their harmonization approach showed that initially heterogeneous datasets could be made comparable, allowing different proxy records to be analyzed together.

Together, these studies demonstrate that differences in temporal resolution, data structure and proxy characteristics can complicate the combination of heterogeneous proxy records. However, they also show that, despite these challenges, harmonization procedures can improve the comparability of different proxy datasets and enable their integration within a common reconstruction framework, motivating the need for harmonization prior to the reconstruction. Consequently, careful attention must be given to the characteristics of the selected proxy records and procedures used to make them comparable prior to the reconstruction.

3.3 Gaussian Processes in Paleoclimate Reconstruction

Gaussian Process models have increasingly been applied within climate and paleoclimatology research due to their ability to model complex, nonlinear relationships while simultaneously quantifying uncertainty. For example, Tipton et al. [THN+19] applied multivariate Gaussian Process inverse prediction to reconstruct paleoclimate conditions from compositional proxy data. Their study demonstrated that Gaussian Process models can provide a flexible probabilistic framework for paleoclimate reconstructions while accounting for uncertainty within the inferred climate signal. Despite these applications, the use of Gaussian Process Regression for reconstructing ENSO variability from heterogeneous proxy records appears to be limited. Based on the literature reviewed in this thesis, no studies were identified that applied GPR as the primary reconstruction framework for integrating ENSO-sensitive proxy records into a continuous reconstructions. This methodological gap motivates the approach adopted in this thesis and described in section 5.

4 Datasets

Multiple heterogeneous proxy datasets were selected to represent long-term ENSO-related climate variability. The selected proxy records differ in proxy type, temporal coverage, sampling resolution, geographic location and climate interpretation. These differences make direct comparison between the datasets challenging.

Table 1: Overview of the proxy datasets considered in this study. Negative BP values indicate years after 1950 CE, where 0 BP is defined as 1950 CE.

Dataset	Proxy Type	Location	Temporal Span	Proxy Variable	Records
Laguna Pallcacocha	Lake sediments	Peru	48–14705 BP	PC1 XRF signal	7024
SST Niño 3.4	SST reconstruction	Pacific Ocean	850–1981 AD	SST anomaly	1132
Tree Rings	Tree rings	North America	900–2000 AD	21-year ENSO variance	1103
Fossil Coral	Coral	Pacific Ocean	4220–4397 BP	Coral signal	2118
Diablo Speleothem	Speleothem	Mexico	1278–10870 BP	$\delta^{18}O$	152
Galápagos Biomarkers	Sediment biomarkers	Galápagos	-51–2720 BP	δD bottom biomarker	61

Table 1 provides an overview of the proxy datasets considered in this study and their key characteristics. These datasets were selected because previous studies have linked them to ENSO-related climate variability [Bra99]. In addition, they represent a diverse range of proxy types, geographical regions and temporal resolutions, allowing ENSO-related signals to be examined across different environmental archives. All six datasets were initially included, and their suitability for the final reconstruction is evaluated in later sections.

The following subsections introduce the different proxy datasets considered within the study, including their proxy type, temporal coverage, and relevance for representing ENSO-related climate variability.

4.1 Laguna Pallcacocha

The Laguna Pallcacocha dataset consists of XRF-derived sediment measurements obtained from Laguna Pallcacocha, a lake located in the Andes mountain range of South America. The sediment records are associated with ENSO-related climate variability because El Niño events influence regional rainfall intensity, causing increased transport of sediment into the lake basin [MARM22]. Previous studies have concluded that the sediment deposition within Laguna Pallcacocha is closely linked to El Niño events. As a result, the sedimentary record preserves ENSO-related climate variability [HNdH+21].

One of the strengths of this dataset is its broad temporal coverage, spanning from approximately 48 to 14,705 years before present (BP). The extensive coverage makes the dataset suitable for reconstruction across a large portion of the Holocene period. Within this study, the first principal component (PC1) derived from the XRF measurements was used as the primary proxy variability signal.

4.2 Reconstructed SST-NINO3.4 anomalies

The second dataset consists of reconstructed sea surface temperature (SST) Niño 3.4 anomalies [DMS+22]. The Niño 3.4 index represents the spatial average of SST anomalies within the Niño

3.4 region of the tropical Pacific Ocean. ENSO strongly influences the sea surface temperatures within this region. El Niño events are associated with warmer-than-average conditions and La Niña events with cooler-than-average conditions. Consequently, the Niño 3.4 index is commonly used as an indicator of ENSO activity. The index therefore represents a climate signal containing ENSO-related variability.

The dataset spans the years 850 to 1981. A noticeable characteristic of the dataset is that each individual year in the temporal coverage contains a value, resulting in a continuous record. This continuous coverage makes it suitable for comparing and aligning ENSO-related variability patterns throughout its reconstruction period. Within this study, the reconstructed SST anomalies were used as the primary proxy variable signal.

4.3 Tree rings

The dataset consists of a tree-ring based ENSO proxy reconstruction derived from the North American Drought Atlas (NADA) [LXC+11]. ENSO influences the atmospheric circulation and rainfall variability, which impacts the drought variability in North America. Since tree-ring growth is sensitive to drought and precipitation conditions, the tree-ring records can indirectly preserve ENSO-related climate variability patterns.

The reconstructed dataset is a continuous annual-resolution record spanning the years 900 to 2000. Within the study, the 21-year running ENSO variance was used as the primary proxy signal, as it represents long-term changes in ENSO amplitude over time.

4.4 Fossil Coral

The fossil coral dataset consists of coral-based ENSO proxy records obtained from the central tropical Pacific Ocean [GCL+20]. Corals are sensitive to changes in climate conditions such as sea surface temperature and oceanographic variability. Since ENSO influences these climate conditions, coral records can preserve ENSO-related climate signals over time.

Compared to the other datasets, the fossil coral dataset has a relatively limited temporal coverage, only spanning from 4220 to 4397 years BP. However, despite this restricted time interval, the dataset contains 2118 observations, providing a substantially higher sampling density than most of the other proxy records considered in this study. The dataset remains valuable because it covers a period of the Holocene for which relatively few proxy records are available. In addition, corals are directly influenced by ENSO-related changes, making them particularly relevant for ENSO reconstructions. Within this study, the coral signal was used as the primary proxy variability signal.

4.5 Diablo Speleothems

The speleothems dataset contains stable isotope data from stalagmite CBD-2, collected from Cueva del Diablo in Mexico [BLM+11]. Speleothems are sensitive to changes in precipitation and moisture variability, which influence the geochemical composition of the stalagmite formations. Because ENSO influences the regional rainfall and moisture transport, variations in stalagmite isotopes can preserve ENSO-related climate variability signals.

The dataset has a broad temporal coverage, spanning 1278 to 10870 years BP. However, the dataset has a limited sampling density, containing only 152 records through this period. Despite this limitation, the broad temporal coverage makes the dataset valuable for reconstruction analyses across a large portion of the Holocene. Within the study, the $\delta^{18}O$ (VPDB) isotope ratio of calcium carbonate was used as the primary proxy variability signal.

4.6 Galápagos Island

The final dataset contains biomarker concentrations from a sediment core near the Galápagos island [AS14]. The dataset is based on hydrogen isotope variability δD preserved within sedimentary biomarkers. Since ENSO strongly influences precipitation variability and hydrological conditions near the Galápagos islands, variation in biomarker isotope composition can preserve ENSO-related variability signals through time.

The temporal coverage of the dataset spans -51 to 2720 years BP. Similar to the Diablo Speleothems dataset, the biomarker record has a relatively limited sampling density, containing only 61 records. Despite this limitation, the proxy dataset remains a valuable addition for representing ENSO-related hydrologic variability within the tropical Pacific region. Within the study, the δD bottom biomarker was used as the primary proxy variability signal. This proxy was selected over the δD dinoflagellate biomarker because it provided a more continuous record with fewer missing values throughout its temporal coverage.

Together, the heterogeneous proxy datasets provide a broad representation of ENSO-related climate variability across different proxy types, temporal scales and geographical regions. Because the datasets differ strongly in structure and characteristics, preprocessing and harmonization are necessary before they can be directly compared and combined.

5 Methodology

Now that the proxy records and their characteristics have been introduced, the methodological framework used to harmonize the heterogeneous proxy records and reconstruct the ENSO-like variability is described. Figure 1 provides an overview of the framework applied throughout the thesis.

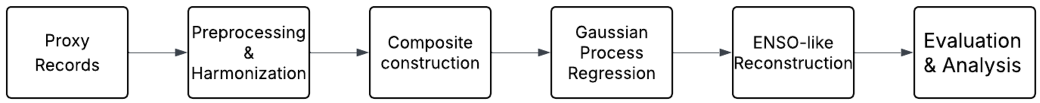

Figure 1: Overview of the methodological framework used to reconstruct ENSO-like variability from heterogeneous proxy records

The framework consists of several stages. First, the proxy records were preprocessed and harmonized to account for differences in temporal resolution and proxy-specific variability. Afterwards, a composite construction method was applied to extract a shared climate variability signal from the processed proxy records. Gaussian Process Regression (GPR) was then used to model a continuous ENSO-like reconstruction with uncertainty quantified. Finally, the resulting reconstructions were evaluated using statistical metrics and sensitivity analyses.

The full framework was implemented using the programming language Python [Pyt26]. The following subsections describe each stage of the framework in more detail.

5.1 Preprocessing & Harmonization

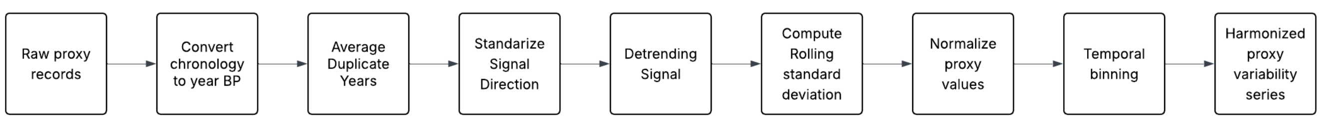

Figure 2: Overview of the workflow of the preprocessing of proxy records into a Harmonized ENSO variability series

Before the proxy records can be compared and combined, each dataset must first be preprocessed and harmonized individually. Since the proxies originate from different sources and differ in scale, temporal resolution and data structure, preprocessing is required to make them comparable and suitable for combination.

Figure 2 illustrates how the raw proxy records are converted into a harmonized proxy variability series. The workflow was designed to reduce the differences between the heterogeneous proxy datasets, allowing the records to be transformed into a consistent and comparable form for the ENSO variability analysis. However, it should be noted that not every preprocessing step was applied to each proxy record. Some proxy datasets were already partially aligned with the target harmonized structure, making certain preprocessing procedures unnecessary. The following subsections discuss each preprocessing and harmonization step in more detail, including its methodological purpose and role within the framework.

5.1.1 Chronology standardization

The first preprocessing steps focus on standardizing the chronology of the proxy records. First, the chronological time variable is converted to the years Before Present (BP) timescale, which is widely used in paleoclimatic research. Using a shared temporal framework allows the proxy records to be aligned and combined consistently, making this an essential harmonization step within the preprocessing workflow.

Next, duplicate years within individual proxies are averaged. Due to inconsistencies in proxy sampling, some proxy datasets contain multiple observations for the same year, often representing values from different months. Averaging these duplicate entries ensures that each proxy record contains a single value per year, resulting in a more consistent temporal structure.

5.1.2 Signal Standardization

Following chronology standardization, the signal itself will be standardized. Since different proxy datasets may represent ENSO variability in opposite directions, the direction of the signal is standardized such that higher values represent stronger than average ENSO-related activity, while lower values represent weaker than average activity. To proxy records were higher proxy values were associated with reduced ENSO activity, the signal was multiplied by -1 to ensure a consistent interpretation across all datasets. This adjustment was applied to the Diablo Speleothem and Laguna Pallcacocha proxy archives.

Ensuring the same signal direction enables comparison and alignment of different proxies by giving positive and negative values the same interpretation across all datasets.

Proxy datasets may contain long-term underlying trends that influence the observed signals without being immediately visible. These hidden trends influence the long-term behavior of proxy records and can obscure the short-term climate variability that is relevant for the reconstruction of ENSO. By detrending the individual proxy datasets, the long-term background signal is reduced, allowing the analysis to focus more on the ENSO-related climate variability.

It is important to extract the ENSO-related signal from the heterogeneous proxy records. Rather than using the standardized proxy values directly, this study focuses on changes in proxy variability through time. ENSO itself is primarily characterized as a mode of climate variability [HMPMS+20], and many paleoclimate studies are concerned with changes in ENSO strength and amplitude rather than changes in mean proxy values. Therefore, changes in variability are more relevant to the objectives of this study than absolute proxy values, which may be strongly influenced by proxy-specific characteristics and local environmental conditions.

To capture this variability, the rolling standard deviation of the detrended signal was calculated. The rolling standard deviation is a statistical measure that quantifies fluctuations within a moving temporal window and therefore provides an estimate of how strongly the proxy signal varies through time. Periods with higher rolling standard deviation indicate greater variability and potentially stronger ENSO activity, whereas lower values indicate reduced variability. This makes the rolling standard deviation particularly suitable for investigating long-term changes in ENSO-related variability.

A temporal window of 50 years was applied to calculate the rolling standard deviation. This window size captures multi-decadal changes in ENSO-related variability while reducing the effect of short-term fluctuations that may not represent persistent climate behavior. It should be noted that the rolling standard deviation with this temporal window does not reconstruct variability from individual ENSO events, which typically occur every 2–7 years. Instead, the rolling window smooths multiple ENSO cycles and captures the long-term changes in the amplitude of ENSO-related variability.

The rolling standard deviation can therefore be used to identify periods of relatively high or low ENSO-related variability throughout the reconstruction period. Applying the rolling standard deviation transforms the proxy records into a comparable representation of climate variability. This allows the heterogeneous datasets to be aligned and used for the extraction of a shared ENSO-related signal.

The resulting rolling variability signal will finally be standardized using z-scores. Z-score standardization transforms data by subtracting the mean and dividing by the standard deviation. This results in the distribution of the data having a mean of 0 and a standard deviation of 1. This preprocessing step places the different proxy datasets, that differ in units and scale, within a common comparable scale, allowing comparison and combination of the datasets.

5.1.3 Temporal Standardization

The last standardization focuses on the temporal structure of the proxy datasets. Individual proxy records differ in both temporal coverage and sampling resolution, resulting in irregular alignment between datasets. To reduce these inconsistencies and improve comparability, the proxy records were aggregated into 50-year temporal bins. A 50-year bin size was selected as a compromise between increasing proxy overlap and preserving temporal variability. The temporal interval reduces the influence of short-term noise and irregular sampling while retaining the longer-term ENSO-related variability patterns relevant to the reconstruction. Temporal binning reduces the influence of irregular sampling by combining observations into consistent temporal intervals, which helps reduce temporal gaps and sampling differences between proxy records. During temporal binning, missing observation were removed from the data where necessary to ensure that the resulting bins contain valid proxy values. This improved the temporal harmonization and comparability of the heterogeneous proxy datasets.

5.1.4 Proxy Selection and Temporal Coverage

After preprocessing, the proxy datasets were compared within a combined overview plot to assess their temporal coverage and overlap. The selected proxy records do not cover the full Holocene period, resulting in differences in data availability through time. To improve the comparability and consistency of the reconstruction framework, a restricted temporal interval with sufficient proxy overlap was selected.

Within the selected temporal coverage, pairwise correlations between the preprocessed proxy records were calculated to evaluate the degree of shared variability and consistency between heterogeneous proxy datasets. The correlation analysis was used to assess the level of agreement between proxy records, justify the harmonization approach and identify proxy records that show relatively weak relationships with the other datasets.

5.2 Composite Construction

After preprocessing and harmonization, the derived proxy variability records were combined into composite signals. A composite signal represents the integration of multiple proxy records into a single shared representation of climate variability. Multiple approaches can be used to construct such composite signals, each differing in how each proxy contributes to the final reconstruction.

One approach is the mean composite. The mean composite is constructed by averaging the harmonized proxy records across the selected datasets. In this approach, each proxy contributes equally to the final composite, allowing the reconstruction to represent the average shared climate variability between the different proxies.

Another composite construction approach is the median composite. Instead of averaging the proxy records, the median value is calculated across the different proxies for each temporal interval. Compared to the mean composite, the median composite is less sensitive to extreme proxy values and outliers, making it potentially a more robust representation of shared climate variability.

The final composite method that will be used is the weighted composite. Unlike the previous methods, the weighted composite assigns different weights to individual proxy records according to predefined weighting criteria. In this study, inverse variance weighting was applied, meaning that proxy records with lower variance contribute more to the final composite than proxy records with higher variance. This approach reduces the influence of highly variable records that may contain proxy-specific noise or variability unrelated to the ENSO-related signal.

The resulting composite signals were subsequently used as input to the Gaussian Process Regression framework.

5.3 Gaussian Process Regression

In this study, Gaussian Process Regression (GPR) was used to transform the composite signal into a continuous ENSO-like variability reconstruction. This is a suitable method for this reconstruction framework because paleoclimate records are often heterogeneous, noisy and irregularly distributed. GPR models the temporal relationship between observations, allowing the reconstruction of smooth climate variability patterns. In addition, GPR estimates the uncertainty of the reconstruction, making it suitable for subsequent analysis and evaluation of the reconstruction framework. The Gaussian Process Regression model was implemented using the *GaussianProcessRegressor* framework from the scikit-learn library [PVG+11]. The behavior of this model is controlled through some parameters including *alpha*, *normalize_y*, *n_restarts_optimizer* and *kernel*.

The *alpha* parameter specifies how much observational noise the model assumes is present in the proxy records. This helps improve numerical stability during model fitting while allowing the model to account for noise and uncertainty in the observations. The *normalize_y* parameter determines whether the target reconstruction signal is normalized before model fitting. In this study, *normalize_y* was set to *True*, ensuring that the Gaussian Process model operates on a standardized target distribution during optimization. The *n_restarts_optimizer* parameter controls the amount of times the optimizer is restarted. The restarts improve the convergence of the kernel hyperparameters and to reduce the likelihood that it is optimizing towards local minima during its optimization. For this study, the parameter is set to 10 restarts.

The last parameter used is *kernel*. The kernel is one of the most important components of the Gaussian Process algorithm, as it determines how the temporal relationships between observations are modeled within the reconstruction framework. Different covariance kernels influence the smoothness, variability structure and flexibility of the signal in different ways. The following subsection discusses the three kernels that were evaluated in this study: Matérn, Radial Basis Function (RBF) and Rational Quadratic kernels. All kernels were implemented using the *sklearn.gaussian_process.kernels* library from Scikit-learn [PVG+11].

5.3.1 Kernels

The kernel structure used in this study can be represented by:

$$k(x, x') = C \cdot k_{base}(x, x') + W$$

where C represents the Constant Kernel component, k_{base} represents the primary covariance kernel, and W represents the White Kernel component used to model observational noise. Within this study, the primary covariance kernels evaluated were the Matérn, Radial Basis Function (RBF), and Rational Quadratic kernels.

The Constant Kernel defines the overall size of the covariance values within the Gaussian Process model and can be combined with different covariance kernels through addition or multiplication. Within the reconstructed framework, the Constant Kernel allows the variability scale of the primary covariance kernel to be adjusted during model optimization. The White Kernel is used to model the noise within the proxy datasets. By incorporating this kernel into the covariance structure, the Gaussian Process model is able to account for uncertainty and variability that may not represent the underlying ENSO-related signal. This improves the robustness and stability of the framework when working with noisy proxy records. The Radial Basis Function (RBF) kernel assumes smooth and continuous relationship within the reconstruction framework. This makes the kernel useful for reconstructing gradual variability patterns by enforcing a strong similarity between nearby observations over time. However, the strong smoothness that the RBF kernel assumes could result in oversmoothed variability patterns within the reconstruction.

Another important covariance kernel is the Matérn kernel. This kernel closely resembles the RBF kernel but allows more flexibility in the smoothness of the reconstructed signal. As a result, the Matérn kernel is more suitable for modeling irregular and inconsistent variability patterns that may be present within heterogeneous proxy records.

The final kernel evaluated in this study is the Rational Quadratic kernel. This covariance kernel is capable of modeling variability across multiple temporal scales simultaneously. This makes the kernel suitable in representing both short-term and long-term fluctuations in climate variability patterns.

These three covariance kernels were evaluated to investigate how different assumptions about temporal relationships influence the resulting ENSO-like reconstruction.

5.4 Evaluation & Analysis

To evaluate the reconstruction framework, the generated ENSO-like reconstructions were compared against a reference reconstruction dataset. For this purpose, the 700 year ENSO Niño 3.4 Index Reconstruction [fEIN24] was selected as a comparative benchmark for assessing similarities in reconstructed variability patterns. Although the Niño 3.4 reconstruction is a reconstructed climate record rather than an official ground truth, it provides a useful reference for evaluating the temporal behavior and variability of the generated reconstructions. To improve comparability between the reference dataset and the generated ENSO-like reconstruction, the same preprocessing framework was applied on the Niño 3.4 dataset.

5.4.1 Statistical methods

To evaluate similarities between the generated and reference reconstruction, several statistical methods were used. These statistical methods were selected to assess different aspects of reconstruction similarity, variability structure and uncertainty.

One of these methods was the Pearson correlation coefficient, which measures the linear relationship between the reconstructions. This metric evaluates how well the variability amplitudes and temporal fluctuations align between the reconstructed ENSO-like signals.

Another statistical method used was the Spearman correlation coefficient, which measures the monotonic relationship between the generated and reference reconstruction based on ranked observations. Instead of comparing the exact reconstruction values directly, the Spearman correlation evaluates whether relative increases and decreases within the reconstructions follow similar temporal patterns.

The Root Mean Square Error (RMSE) was also used to quantify the average difference between the generated and reference reconstructions. RMSE measures the magnitude of the reconstruction error by calculating the average deviation between corresponding reconstruction values. A lower RMSE value indicates a greater similarity between the generated reconstruction and the reference reconstruction.

The last statistical method evaluated was the mean uncertainty. The mean uncertainty represents the average uncertainty estimated by the Gaussian Process Regression model throughout the reconstruction period. This metric was used to assess the overall confidence and stability of the generated model.

In addition to the statistical evaluation, the reconstructed ENSO-like variability signals were visually analyzed using reconstruction plots. The baseline reconstruction was compared with the reference Niño 3.4 reconstruction over the overlapping time interval to assess the overall reconstruction structure and uncertainty estimates. Furthermore, reconstruction plots with different methodological configurations were visually compared to evaluate the influence of different methodological choices.

5.4.2 Analysis Framework

To evaluate how different methodological choices influence the reconstruction, several analyses were performed. These analyses focus on the effects of temporal selection, Gaussian Process model configuration, proxy selection and temporal binning on reconstruction similarity, variability structure and reconstruction uncertainty.

The temporal analysis focuses on evaluating how different rolling standard deviation window sizes influence the extracted ENSO-like signal. The rolling standard deviation window is an important component in the reconstruction framework since it controls the temporal scale at which the climate variability is extracted, directly influencing the smoothness and variability structure of the reconstructed signal. Multiple rolling standard deviation windows, including 21-, 25-, 40- and 50-year intervals, were evaluated during this analysis.

The GPR analysis focuses on evaluating how different covariance kernels influence the resulting reconstruction. For this purpose, the RBF, Matérn and Rational Quadratic kernels were used in the analysis. Different kernel length scales were also evaluated on values of 0.3, 0.5 and 1.0. The length scale of a covariance kernel directly influences the temporal smoothness of the reconstructed signal. Smaller length scales directly influence the temporal smoothness, whereas larger length scales produce a more smoother reconstructions. In addition, fixed and optimized kernel configurations were compared to assess how parameter flexibility influences reconstruction behavior and uncertainty.

The proxy analysis focuses on evaluating how different composition approaches influence the variability signal. For this purpose, mean, median and weighted composite methods were compared within the reconstruction framework. In addition to comparing different composite methods, a leave-one-out proxy analysis was performed to evaluate the sensitivity of the reconstruction framework to individual proxy records.

The temporal binning analysis evaluates how different temporal aggregation intervals influence the reconstruction framework. Bin sizes of 25, 40 and 50 years were tested to assess how the temporal aggregation affects proxy overlap, reconstruction smoothness, variability structure and statistical similarity to the reference reconstruction.

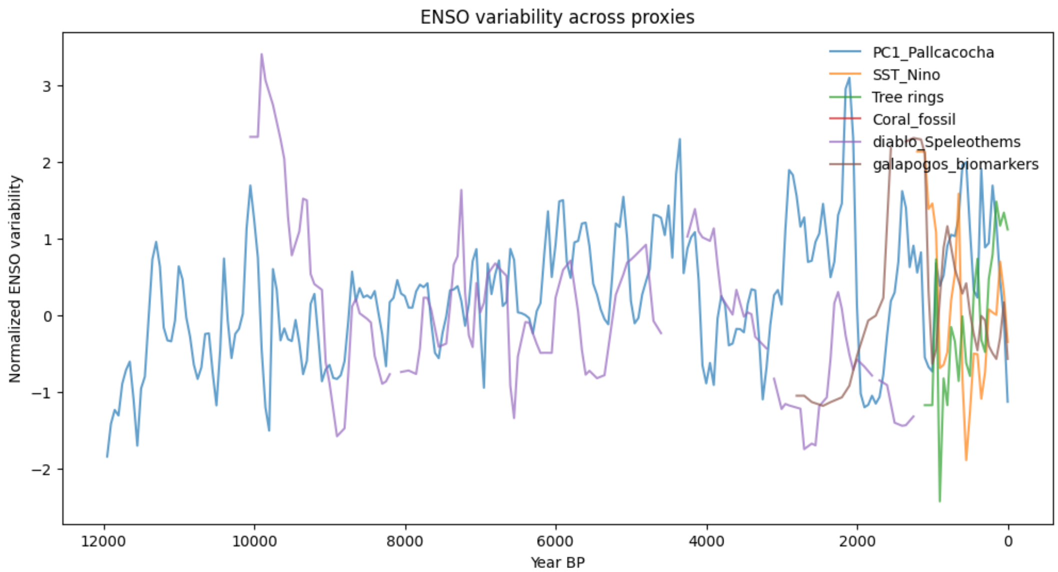

Figure 3: Reconstruction of preprocessed ENSO-related proxy records in the timeframe 12000-0 BP.

6 Results and Analysis

To begin the reconstruction analysis, an initial baseline reconstruction configuration is established using the preprocessing and harmonization methodology described in chapter 5. This baseline reconstruction serves as a reference configuration for evaluating the behavior of the harmonized proxy records and the resulting ENSO-like variability signal.

The baseline preprocessing configuration applies a rolling standard deviation window of 50 years, with the minimum number of required observation within each rolling window set to 50% of the rolling window size on each proxy. The tree-ring proxy record is an exception within this process, since the dataset already provides a processed ENSO variability signal through a running ENSO variance reconstruction.

Figure 3 shows the processed and standardized proxy records across their available temporal coverage. Although some proxy records extend far into the Holocene, differences in temporal overlap and sampling density remain visible. Proxy overlap becomes considerably stronger within the last approximately 2000 years BP, where a larger number of heterogeneous proxy records coexist. Because the robustness of the harmonization depends on sufficient overlap between proxy records, the subsequent reconstruction analysis primarily focuses on the later interval where proxy availability and comparability are strongest.

It is also noticeable that the Coral proxy record is not visible in the combined preprocessed reconstruction showed in figure 3. This is caused by the relatively limited temporal coverage of the proxy record compared to the other datasets. Because of its limited contribution to the combined reconstruction framework and its distinct temporal coverage, the Coral proxy record was excluded from the subsequent reconstruction analysis.

Figure 4 shows a combined preprocessed reconstruction focused on the found temporal window. Although the overlap between different proxy records within the same temporal window remains limited, the later Holocene interval shows a stronger coexistence between heterogeneous proxy

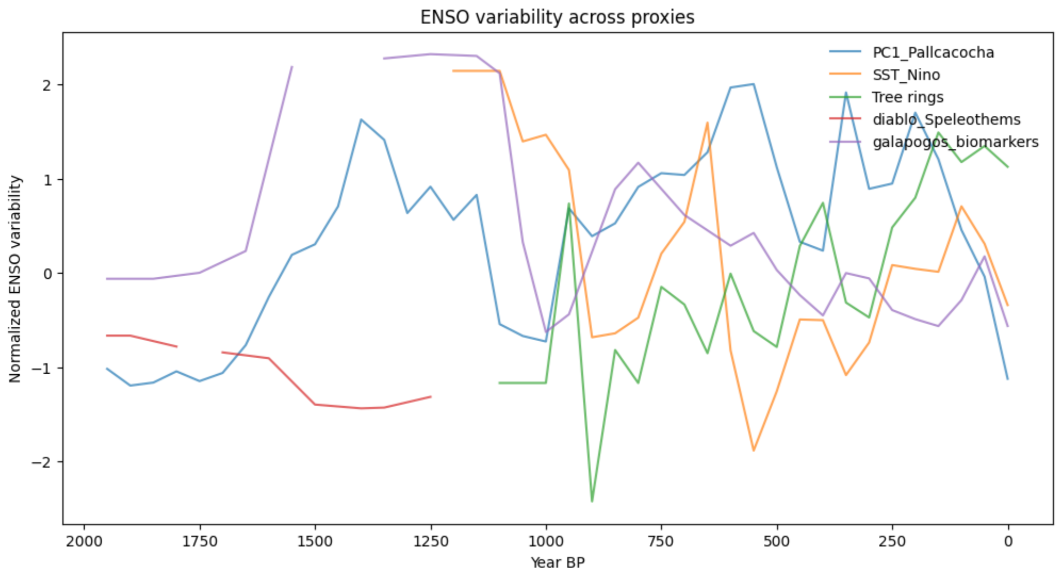

Figure 4: Reconstruction of preprocessed ENSO-related proxy records in the timeframe 2000-0 BP.

datasets compared to a full-Holocene overview. Differences between variability structure, temporal continuity and amplitude between the proxy records remain visible, further highlighting the heterogeneous of the proxy datasets.

	PC1_Pallcacocha	SST_Nino	Tree rings	Diablo_Speleothems	Galapogos_biomarkers
PC1_Pallcacocha	1.00	-0.44	0.05	-0.95	0.18
SST_Nino	-0.44	1.00	-0.02	N/A	0.47
Tree rings	0.05	-0.02	1.00	N/A	-0.55
Diablo_Speleothems	-0.95	N/A	N/A	1.00	-0.95
Galapogos_biomarkers	0.18	0.47	-0.55	-0.95	1.00

Table 2: Pairwise correlations between the preprocessed proxy records within the selected temporal interval 2000-0 BP.

To further evaluate consistency between the proxy records within the selected temporal interval, pairwise correlations between the records were calculated. Table 2 presents the resulting correlation values.

Instantly noticeable are the results associated with the Diablo Speleothems proxy records. The proxy record shows a strong negative correlation with other proxy datasets, including PC1 Pallcacocha and Galapogos biomarkers ($r = -0.95$). The speleothems proxy also has a limited overlap within the selected temporal interval, resulting in missing correlation values with the other proxy records. These results do not necessarily indicate that the Diablo proxy records lack ENSO-related climate variability. Although the Diablo proxy record spans more than 8000 years of ENSO-related variability, the chosen temporal time-frame only overlaps with a small portion of 722 years, which means that the calculated correlations are based on these overlapping observations. In addition, the overlapping interval does not coincide with the temporal coverage of other proxy datasets, limiting the contribution of the Diablo speleothems proxy record within the harmonization framework. Because the correlation analysis and harmonization process rely on sufficient temporal overlap and comparability between proxy records, the Diablo_Speleothems proxy record was excluded from the subsequent reconstruction analysis.

Looking at the pairwise correlations between the remaining proxy records, considerable differences in correlation strength remain visible. The strongest positive correlation is observed between SST Nino and Galapogos biomarkers ($r = 0.47$), while the strongest negative correlation is observed between Tree rings and Galapogos biomarkers ($r = -0.55$). Overall, these results suggest that the heterogeneous proxy records exhibit weak to moderate positive and negative relationship with one another.

These relatively weak to moderate correlations do not necessarily indicate that the preprocessing or harmonization methodology failed, but are expected within paleoclimatic proxy datasets. The proxy records differ substantially in geographic location, proxy type and climate sensitivity. In addition, paleoclimatic proxy records often contain environmental and climatic variability signals beyond ENSO-related variability alone, which can further influence direct agreement between heterogeneous proxy datasets. Furthermore, negative correlations are not necessarily problematic, since different proxy records may respond inversely to ENSO-related changes depending on their regional environmental conditions and seasonal sensitivity. Despite the observed differences in correlation behavior, the results still suggest the presence of shared ENSO-related variability across the heterogeneous proxy records. This supports the possibility of applying harmonization and reconstruction techniques to extract a shared ENSO-related variability signal.

Using the selected proxy records, a baseline ENSO-like variability composite was constructed. The baseline reconstruction uses the mean composite of the harmonized proxy records because it provides a straightforward representation of the shared variability structure between the different proxies while assigning equal weight to each proxy archive.

To transform the composite to a continuous reconstruction, Gaussian Process Regression (GPR) was applied using a baseline Matérn kernel combined with the ConstantKernel and WhiteKernel component. The baseline kernel configuration uses a length scale of 1.0 with fixed length scale bounds throughout the reconstruction process.

Figure 5 shows the baseline GPR reconstruction produced from the harmonized ENSO-like composite signal. The reconstructed signal generally follows the variability structure of the composite observations while smoothing noisy local variations within the composite.

The reconstruction shows a gradual increase in variability between approximately 2000 and 1250 years BP, followed by a gradual decline towards the year 400 BP and a renewed increase towards the present interval near 0 BP. From a climatic perspective, this pattern may indicate that ENSO-related variability was not constant throughout the reconstructed interval. The higher variability between approximately 1500 and 1100 years BP may suggest a period of enhanced ENSO-related activity, whereas the lower reconstructed values between approximately 750 and 400 years BP may indicate reduced ENSO-related variability. If these variations reflect changes in ENSO amplitude, they could imply that the influence of ENSO on regional climate systems varied through time. Despite noisy fluctuations within the composite observations, the GPR reconstruction produces a relatively continuous structure across the full interval.

The uncertainty interval remains relatively broad throughout the reconstruction, reflecting the heterogeneous nature of the proxy network and the limited number of the underlying proxy datasets. Increased uncertainty is especially visible after approximately 1000 years BP, where the uncertainty interval remains relatively broad throughout the later reconstruction interval.

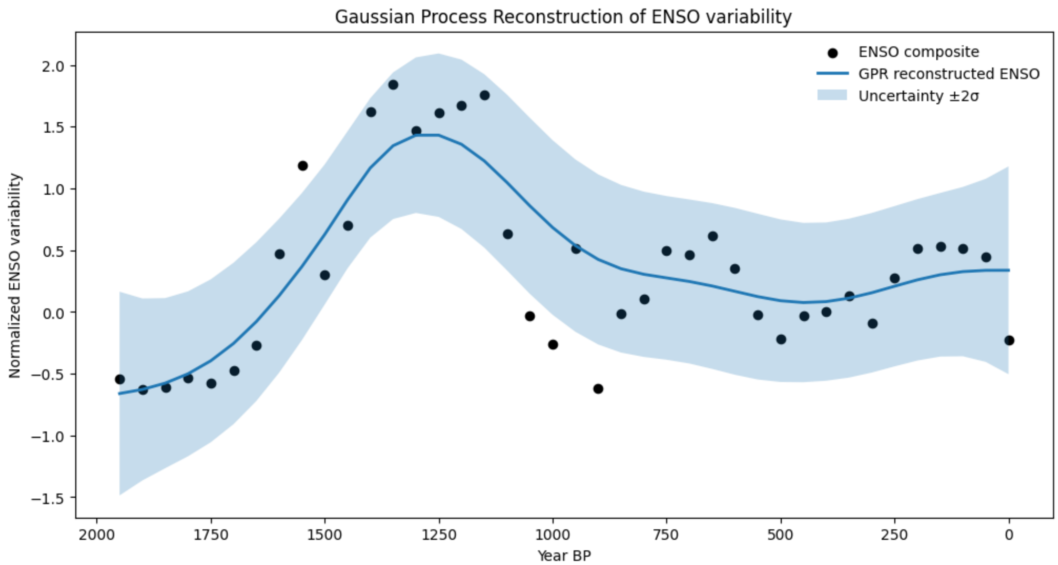

Figure 5: Gaussian Process Regression baseline reconstruction including uncertainty.

Additional uncertainty increases are visible near the reconstruction boundaries around 2000 and 0 years BP. Overall, the uncertainty estimation suggests that although the reconstruction preserves large-scale variability, reconstruction confidence remains limited within certain intervals due to the heterogeneous nature of the proxy datasets.

To conclude the baseline reconstruction analysis, the resulting ENSO-like variability reconstruction was compared with an independent 700 year Niño3.4 reconstruction over the overlapping interval. This overlapping interval focuses on the last 600 years BP. Although the overlapping interval represents only a small portion of the full reconstruction period, it represents the most suitable available reference interval for evaluating and comparing the reconstructed ENSO-like variability signal. The Niño 3.4 reconstruction used for validation is distinct from the SST proxy record included in the harmonized proxy network and was therefore treated as an independent reference dataset. Furthermore, the Nino3.4 reconstruction provides continuous yearly observations throughout the overlapping interval making it a suitable framework for comparison with the ENSO-like reconstructed variability signal. Although the overlapping interval does not represent the full reconstruction period, it provides the closest available reference framework for evaluating the reconstructed ENSO-like variability signal against an existing ENSO reconstruction.

Figure 6 compares the baseline GPR with the Nino 3.4 reconstruction in its overlapping interval. Observationally, direct visual comparison between both reconstructions is limited because of the relatively short overlapping interval. The Niño3.4 reconstruction exhibits a less smooth and continuous signal than the GPR reconstruction since the Gaussian Process Regression framework explicitly smooths irregular variability within the composite signal to produce a continuous reconstruction.

To further evaluate the agreement between both reconstructions, multiple statistical comparison metrics were calculated as shown in Table 3. The Pearson correlation of 0.692 shows a strong positive agreement, suggesting that the baseline reconstruction captures a comparable large-scale pattern of ENSO-related variability. The Spearman correlation of 0.593 suggests that there is a moderate similarity in the relative variability behavior of both reconstructions.

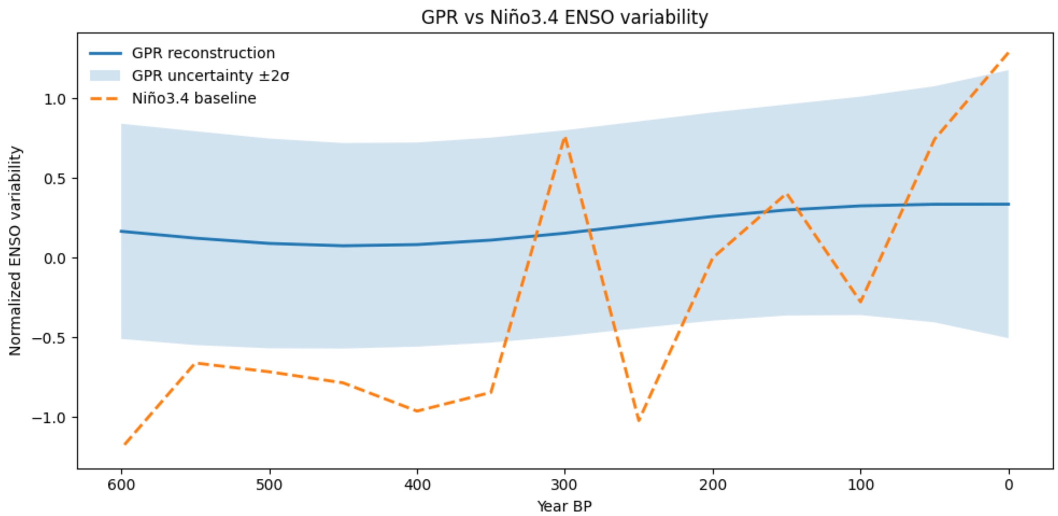

Figure 6: Gaussian Process Regression baseline reconstruction compared with the Niño 3.4 reconstruction.

Metric	Value
Pearson correlation	0.692
Spearman correlation	0.593
RMSE	0.844
Mean uncertainty	0.339

Table 3: Statistical comparison between the baseline GPR reconstruction and the Niño 3.4 reconstruction.

The RMSE of 0.844 reflects moderate differences between both reconstructions, which is expected given the heterogeneous proxy framework and relatively limited overlap interval. Finally, the mean uncertainty of 0.339 indicates moderate reconstruction uncertainty throughout the reconstruction period, reflecting its heterogeneous nature and limited overlap between the proxy datasets. Overall, the statistical metrics suggest moderate agreement between the baseline reconstruction and the Nino3.4 reconstruction. This can be expected, since the signal likely contains climatic and environmental variability beyond ENSO-related variability.

6.1 Temporal analysis

After establishing the baseline method, the temporal behavior of the reconstruction was further analyzed. Compared to the baseline reconstruction configuration, additional rolling variability windows of 40, 25 and 21 years were tested to evaluate their influence on the temporal resolution of the ENSO-like reconstruction variability signal.

Because the minimum number of required observations within each rolling standard deviation window is defined relative to the window size, changes in the rolling variability window also directly influence the minimum periods parameter in the preprocessing stage.

Figure 7 compares the GPR reconstructions produced using different rolling variability windows during preprocessing. Considerable differences between the reconstructions remain visible, specifi-

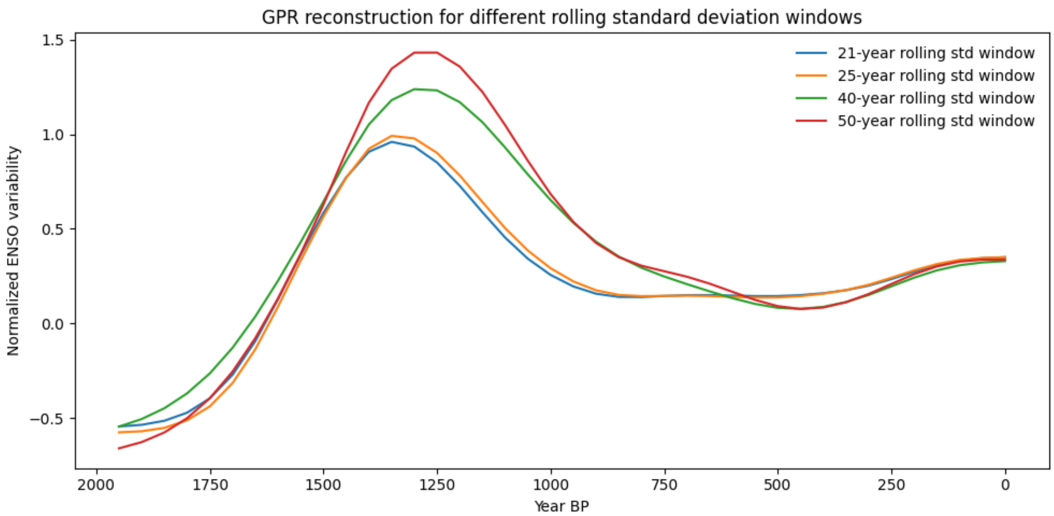

Figure 7: Gaussian Process Regression reconstructions with different rolling standard deviations

cally within the interval between approximately 1500 to 750 years BP. Larger rolling variability windows produce smoother reconstructions with stronger large-scale variability structures, while smaller rolling windows produce reconstructions with a more reduced amplitude. Another noticeable observation is that the reconstruction converge around approximately 600 year BP. After this point, the smaller rolling window exhibit greater local variability, whereas the larger rolling windows produce smoother and more stable reconstructions. This behavior is expected because smaller rolling windows retain more short-term fluctuations, while larger windows emphasize broader multi-decadal variability patterns.

These results suggest that the selected rolling variability window substantially influence the structure of the extracted ENSO-like variability signal in the reconstruction framework. Smaller variability windows increase the temporal sensitivity of the reconstruction and preserve more short-term variability, while larger variability windows emphasize broader long-term variability through stronger temporal smoothing.

To further evaluate the influence of the rolling variability window on reconstruction performance, statistical comparison metrics were recalculated against the Niño 3.4 reconstruction. For consistency, the Niño 3.4 reconstruction was also preprocessed using the corresponding rolling variability window before comparison.

Rolling Window	Pearson	Spearman	RMSE	Mean Uncertainty
50	0.692	0.593	0.844	0.339
40	0.698	0.621	0.807	0.330
25	0.693	0.659	0.816	0.308
21	0.696	0.665	0.839	0.313

Table 4: Statistical comparison metrics for different rolling variability window.

Table 4 shows the statistical comparison metrics for different rolling variability window configurations. Overall, the statistical metrics remain relatively comparable between the different

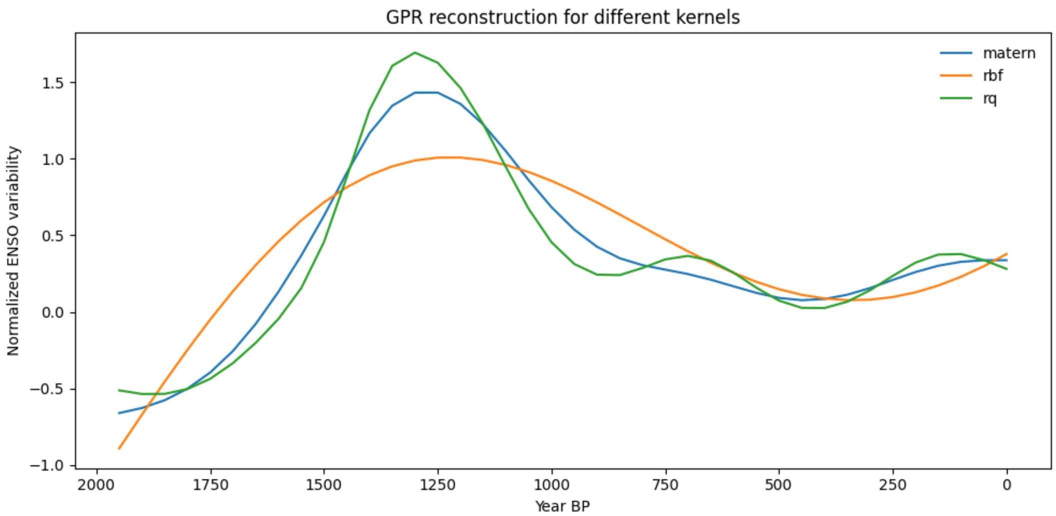

Figure 8: Gaussian Process Regression reconstructions with different kernels

preprocessing windows, indicating that the reconstruction framework remains relatively stable under moderate changes in the rolling variability window.

The 40-year rolling window produced both the lowest RMSE of 0.807 and the highest Pearson correlation of 0.698, suggesting the closest overall agreement with the Niño 3.4 reconstruction. In contrast, the 21-year rolling window produces the highest Spearman correlation of 0.665, indicating a slightly improved agreement in variability behavior between both structures. The 25-year rolling variability window produces the lowest mean uncertainty, although this configuration also results in a slightly higher RMSE than the other configurations. Overall, these results suggest that the selected rolling variability window influences different aspects of the reconstruction performance, although the overall statistical differences between the configurations remain relatively small.

6.2 GPR kernel and parameter analysis

After evaluating the influence of preprocessing-related temporal smoothing, the sensitivity of the reconstruction framework to the selected Gaussian Process Regression (GPR) kernel configuration was analyzed. Different covariance kernels and kernel parameter settings were tested to evaluate their influence on the reconstructed ENSO-like variability signal, reconstruction smoothness and agreement with the Niño 3.4 reference reconstruction.

The analysis compares the Matérn, Radial Basis Function (RBF), and Rational Quadratic (RQ) covariance kernels using multiple length scale configurations. In addition, both fixed and optimized kernel parameter configuration were evaluated to assess the influence of kernel flexibility on the resulting reconstruction behavior.

Figure 8 compares the GPR reconstructions produced using different covariance kernels while maintaining a fixed length scale configuration. Considerable differences between kernels can be seen. The RBF kernel produces the smoothest reconstruction, with a relatively reduced influence from short-term variability. In contrast, the RQ kernel generates stronger local variability and has a larger reconstruction amplitude, while the Matérn kernel produces a reconstruction that lies between both kernels. These results suggest that the selected covariance kernel substantially

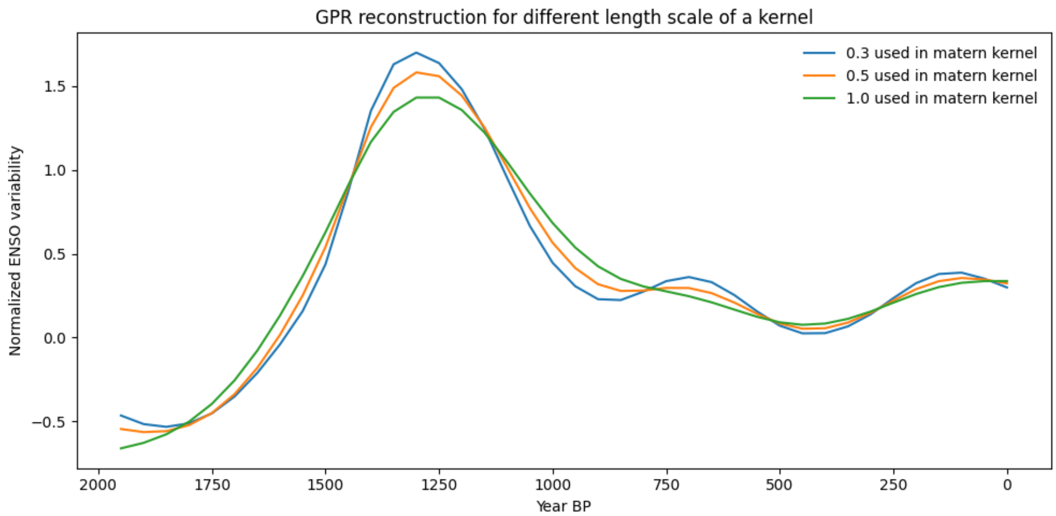

Figure 9: Gaussian process regression reconstruction with different length scales

influences the smoothness and variability pattern of the reconstructed ENSO-like signal.

Figure 9 compares the GPR reconstructions produced using the Matérn kernel with different length scale configurations. The overall large-scale reconstruction structure remains relatively similar across the different configurations, although noticeable differences in local variability amplitude can be observed. Smaller length scales produce stronger local variability and slightly larger reconstruction amplitudes, most noticeable between 1300 and 900 years BP. Larger length scales produce a reduced variability magnitude. These results suggest that the selected length scale primarily influences the sensitivity of the reconstruction to smaller local temporal variability rather than variability on a broader temporal scale.

The complete statistical comparison metrics for all kernel and parameter configurations are provided in the Appendix in table 6. Overall, there are some noticeable differences between the different configurations, with the strongest differences occurring within the RBF kernel. The fixed RBF kernel with a length scale of 0.5 produces the highest Pearson correlation of 0.761, suggesting the strongest statistical agreement with the Niño3.4 reconstruction among the tested configurations. However, the RBF kernel tends to produce higher uncertainty values than the other kernels, particularly within the fixed kernel configuration.

The Matérn kernel produces more moderate and stable statistical behavior across the tested parameter configurations, without the stronger statistical variability observed within the RBF kernel. The Rational Quadratic kernel also remains stable across the tested parameter configurations, with only moderate differences between the fixed and optimized configurations. Fixed kernel reconstructions produce stronger Pearson and Spearman correlation, although they also produce higher mean uncertainty values compared to the optimized configurations. Overall, these results suggest that kernel selection and parameter configuration substantially influence the reconstruction smoothness and local variability behavior. The statistical differences between most kernel configurations remain relatively small, suggesting that the reconstruction framework remains stable under changes in kernels and parameters when comparing to Niño3.4.

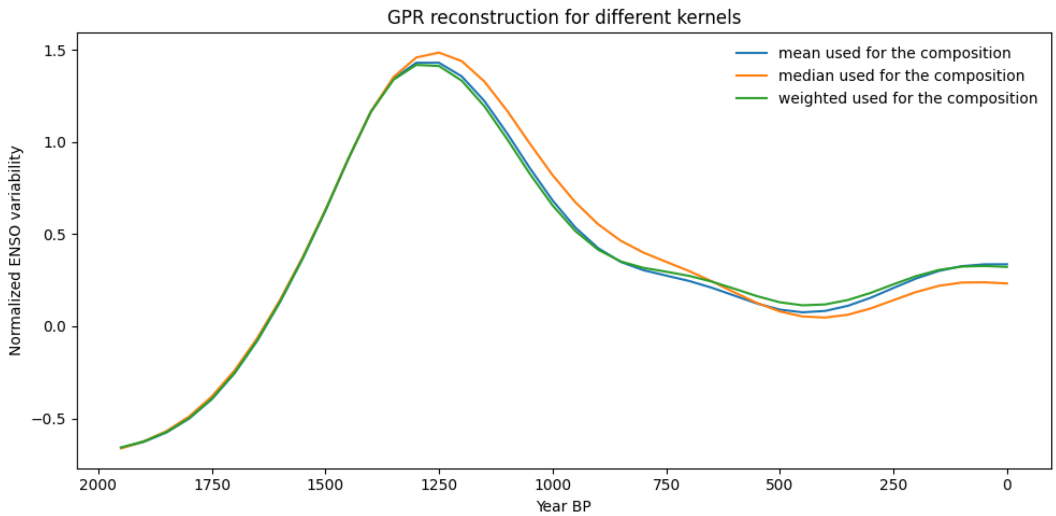

Figure 10: Gaussian process regression reconstruction produced by different composite methods

6.3 Proxy analysis

After evaluating the Gaussian Process Regression parameter selection, the sensitivity of proxy selection was analyzed. Different composite reconstruction methods, including the mean, median and weighted methods, were compared. In addition, individual proxy records were removed to evaluate how each proxy independently influences the ENSO-like variability signal.

Figure 10 shows a comparison between GPR reconstruction produced using different composition methods. Overall, the reconstructions remain similar across the different composition methods, with only minor differences in variability amplitude. The median composition produces a slightly stronger variability peak around 1260 year BP, while the mean and weighted composition remain nearly identical throughout the whole reconstruction interval. These results suggest that the selected composition method has limited influence on the overall ENSO-like reconstruction structure.

Figure 11 shows a comparison between GPR reconstructions produced after removing individual proxy datasets from the reconstruction framework. Considerable differences between the reconstructions can be observed, indicating that individual proxies have a strong influence on the reconstructed variability signal.

The removal of the Pallcacocha dataset shows the strongest structural changes in the reconstruction. The most notable differences can be observed between approximately 1600 and 1100 years BP, where the amplitude of the reconstruction is substantially larger compared to the other reconstructions. This indicates that the Pallcacocha dataset strongly constrains the magnitude of the reconstructed ENSO-like variability signal during this interval. Removing the SST Niño proxy alters the reconstruction between approximately 1200 and 800 years BP. Within this interval, the signal has reduced variability compared to other reconstructions, including a minimum around 1000 years BP. These results show that SST Niño proxy has strong influence on the reconstructed signal during that interval.

Removing the tree rings dataset produces moderate structural differences compared to other proxy removed reconstructions. However, the reconstruction displays stronger local variability and reduced smoothness throughout the later parts of the reconstruction, indicating that the tree rings

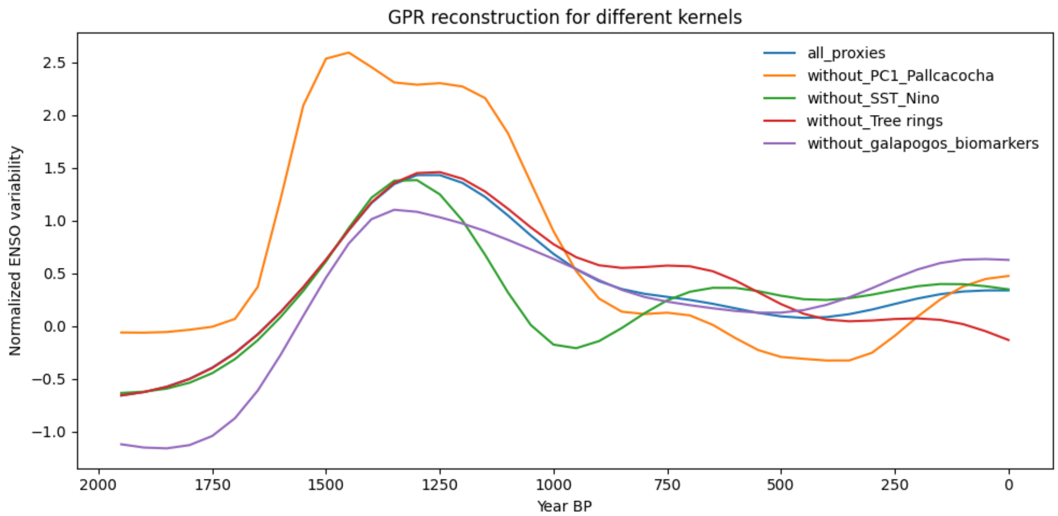

Figure 11: Gaussian process regression reconstruction produced after removing individual proxy records

dataset contributes to stabilizing short term variability throughout the reconstruction framework. In contrast, removing the Galápagos biomarkers dataset shows a reduced amplitude throughout the early parts of the reconstruction interval and generates smoother reconstruction behavior overall. This indicates that the Galapogos biomarkers proxy contributes to preserving local variability within the reconstructed ENSO-related signal.

Appendix Table 7 presents the statistical comparison metrics for all possible combinations of composition method and proxy-removal configurations. One of the most noticeable observations is the negative Pearson and Spearman values that occur when the tree rings proxy is removed. This indicates that the tree rings proxy strongly influences the direction and variability structure within the overlapping comparison interval with the 3.4 Niño reconstruction. The exact reason why removing the tree-rings proxy produces such dramatic reduction in correlation with the Niño 3.4 reconstruction remains unclear. While the continuous temporal coverage of the tree-rings record may contribute to its strong influence within the composite, the current analysis does not allow a definite explanation for this. This result highlights the sensitivity of the reconstruction framework to individual proxy records and suggests that some proxies may contribute disproportionately to agreement with the reference reconstruction.

Another observation is the reduction in Pearson correlation and increase in RMSE whenever the SST Niño proxy is removed. This indicates that there is weaker similarity in variability structure and larger absolute differences between the reconstruction and the Niño 3.4 reconstruction. These results suggest that the SST Niño proxy plays an important role in creating statistical agreement between these reconstructions.

It should be noted that the SST Niño proxy and the Niño 3.4 reference reconstruction both represent ENSO-related sea-surface temperature variability. As a result, some degree of agreement between the reconstruction and reference dataset can be expected. However, removing the SST Niño proxy does not eliminate the agreement entirely. Although Pearson correlation decreases substantially and RMSE increases, positive correlations remain for the mean and weighted composite methods. This suggests that the observed agreement with the Niño 3.4 reconstruction is not solely

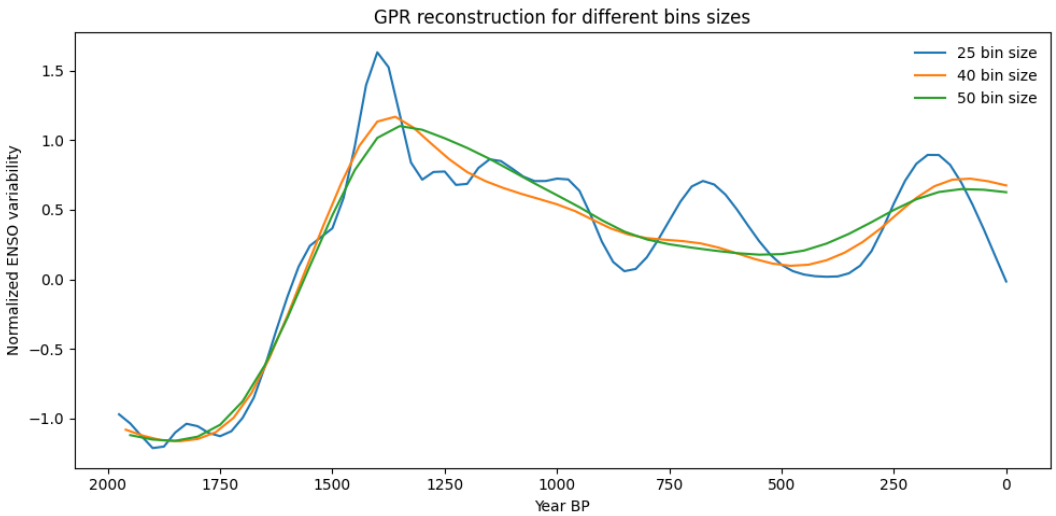

Figure 12: Gaussian Process regression reconstructions produced using different temporal binning sizes during preprocessing

driven by the SST Niño proxy and that the remaining proxy records still capture part of the shared ENSO-related variability signal.

Removing the Pallcacocha proxy results in a significantly lower RMSE and mean uncertainty when comparing to the Niño 3.4 dataset. This suggests that the Pallcacocha proxy strongly influences the variability magnitude of the reconstruction, although it does not improve statistical agreement with the Niño 3.4 dataset. Removing the Galápagos biomarker proxy produces the highest Pearson values for all composite methods, although it increases the RMSE compared to the baseline reconstruction. This indicates that the overall variability structure remains similar, while variability magnitude becomes more different.

The median composite method also appears less stable, producing two negative correlation values together with relatively high RMSE and mean uncertainty across multiple configurations. The mean and weighted composite methods generally display more consistent behavior across multiple proxy-removal configurations compared to the median composite method.

Overall, these results suggest that the reconstruction framework remains sensitive to the inclusion and exclusion of individual proxy datasets, indicating that certain proxy records have substantial influence on the reconstructed ENSO-like variability signal.

6.4 Binning analysis

The final sensitivity analysis evaluates the influence of temporal binning sizes on the reconstructed ENSO-like signal. Different temporal binning configurations were applied during the preprocessing phase to investigate how they influence the preservation of local variability in the reconstruction framework.

Figure 12 shows GPR reconstructions produced using different temporal binning values. Lower temporal binning values produce reconstruction with stronger local variability and larger amplitude fluctuations, which can result in a less smooth reconstruction. Higher temporal binning values, however, generates reconstructions with smaller amplitude fluctuations and weaker representation of local variability, resulting in a smoother reconstruction.

Bin Size	Pearson	Spearman	RMSE	Mean Uncertainty
50	0.692	0.593	0.860	0.352
40	0.516	0.459	0.887	0.328
25	0.331	0.248	0.930	0.266

Table 5: Statistical comparison metrics for different temporal binning configurations.

Table 5 shows the statistical comparison metrics when comparing the reconstructions to Niño 3.4. A similar pattern can be observed within the statistical metrics. Smaller temporal binning values produce lower Pearson and Spearman correlation together with higher RMSE values, whereas the opposite pattern can be observed from larger temporal binning values. This is likely the result of the stronger preservation of local variability and the reduced smoothing associated with smaller bin sizes. These results suggest that the reconstruction framework more consistently preserves long-term variability, while smaller temporal binning values increase the sensitivity of the framework to local variability.

7 Discussion

The main objective of this study was to investigate how heterogeneous Holocene ENSO-sensitive proxy records can be harmonized into a coherent long-term reconstruction of ENSO variability. The results demonstrate that despite differences in proxy type, temporal coverage, sampling density and climatic sensitivity, it is possible to extract a shared ENSO-related variability signal through a combination of preprocessing, composite construction and Gaussian Process Regression (GPR).

This is supported by the statistical comparison between the GPR reconstruction and the Niño3,4 reconstruction, which produced a Pearson correlation of 0.692 and a Spearman correlation of 0.593. These values indicate moderate agreement between the reconstruction. This suggests that the framework was able to capture meaningful aspects of long-term ENSO variability. Although the observed correlations do not indicate a near-perfect correspondence between the reconstructions, such a result is not necessarily expected when working with heterogeneous paleoclimate reconstructions. The reconstructed ENSO variability signal is derived from multiple proxy archives that differ in proxy type, geographic location and climatic sensitivity. As a result, each record captures only part of the ENSO-related climatic variability. Furthermore, the proxies are influenced by climatic and environmental processes other than ENSO, which cannot be completely removed using preprocessing and harmonization. Consequently, some disagreement between the reconstruction is expected.

The proxy-removal analyses further support the ability of the framework to produce a coherent reconstruction. Differences between the mean, the median and weighted composite methods were relatively small when all proxies were included, suggesting that the overall reconstruction is not strongly depended on the composite method. This indicates that the harmonization framework is reasonably robust to variations in composite construction.

In contrast, the removal of individual proxies produce substantially larger changes in reconstruction performance. For example, removing the Galápagos biomarker proxy increases the Pearson correlation of the mean composite from 0.692 to 0.718, while removing the tree-rings proxy results in a negative correlation score for all composite reconstruction methods. These results suggest that reconstruction quality is sensitive to the proxy selection used for harmonization.

The substantial reduction in correlation and increase in RMSE after SST Niño removal indicates that this proxy contributes strongly to the observed agreement with the Niño 3.4 reconstruction. However, the remaining positive correlations suggest that agreement is not solely driven by the SST Niño proxy and that other proxy records also contribute to the shared ENSO-related variability signal.

The varying impact of individual proxies on reconstruction performance highlights the heterogeneous nature of the selected proxy records. This heterogeneity is evident in several aspects of the dataset. The different proxy data have different temporal coverage and sampling densities, impacting how well they capture climate signals such as ENSO. Furthermore, after preprocessing, weak-to-moderate positive and negative pairwise correlations can be observed between proxy pairs, indicating that ENSO is not represented uniformly across proxy archives. This is likely the result of differences in regional climate as well as proxy-specific sensitivities. While some proxies may only respond to ENSO-related changes in sea-surface temperature, others may record changes in atmospheric circulation, rainfall or other environmental processes that are ENSO-related.

Overall, these analyses demonstrate that the selected proxy records differ substantially in their representation of long-term ENSO-related variability. These differences arise from variations in temporal coverage, sampling density, climatic sensitivity, and the environmental processes recorded by each proxy.

While these results provide insight into the heterogeneity of the selected proxy records, they also raise the question of how effectively Gaussian Process Regression helps transform heterogeneous and irregular proxy signals into a coherent reconstruction. The reconstruction results therefore provide an opportunity to evaluate both the strengths and limitations of GPR within the context of long-term ENSO reconstructions.

As illustrated by the baseline GPR reconstruction, one of the main strengths of GPR is its ability to reconstruct a continuous signal from sparse and irregular observations. This is valuable in paleoclimate research, where proxies often have gaps and differ in temporal resolution, as illustrated in figure 4. GPR also provides uncertainty estimates alongside the reconstruction. This makes it possible to assess the confidence associated with different parts of the reconstructed signal. Finally, the kernel, parameter and input analysis showed that the reconstruction is relatively robust to changes in rolling-window size and composite method, suggesting that GPR produces reasonably stable results under moderate methodological variations.

Although GPR demonstrated several advantages for reconstructing ENSO-related variability from heterogeneous proxy records, the results also highlight a number of important limitations. As shown in figure 8, the choice of kernel can substantially influence the shape of the reconstructed signal, particularly at local timescales. Different kernels produce different degrees of smoothness and local variability, which can lead to different interpretations of shorter-term fluctuations within the reconstruction. A second limitation is that reconstruction performance is almost always dependent on the quality of the proxy records. The proxy-removal analysis showed that changes in proxy selection can substantially alter reconstruction performance, indicating that GPR cannot fully compensate for differences in proxy characteristics.

Despite these limitations, the moderate agreement between the GPR reconstruction and the Niño3.4 reconstruction suggests that GPR remains capable of extracting meaningful ENSO-related variability from heterogeneous proxy datasets. The results indicate that GPR provides a useful balance between reconstruction continuity, uncertainty quantification and robustness, making it a suitable tool for long-term paleoclimate reconstructions.

7.1 Connecting to Literature

The findings of this study are broadly consistent with previous research on multi-proxy climate reconstructions. Li et al. [LNA10] demonstrated that heterogeneous proxy records can be combined into a coherent climate reconstruction through a Bayesian framework. Similarly, the results of this study support the idea that meaningful ENSO-related variability can be extracted from heterogeneous proxy datasets despite substantial differences between proxy records. The results also support the findings of Liu et al. [LMZZ24], who demonstrated that is was possible to reconstruct an ENSO signal using multiple proxy records.

The challenges encountered during the reconstruction process are also consistent with the observations of Roberts et al. [RTH+19] and Pfalz et al. [PDFB21]. Both studies emphasized the difficulties associated with heterogeneous proxy datasets and highlighted the importance of harmonization procedures. The strong sensitivity of reconstruction performance to proxy selection observed in this study further supports their conclusion that differences between proxy records remain an important challenge when trying to create an multi-proxy climate reconstruction.

The findings of this study can also be considered in the context of the broader debate on Holocene

ENSO evolution. Previous studies have suggested that ENSO variability was reduced during parts of the early and middle Holocene and subsequently intensified towards the late Holocene [LLZC18]. Although the reconstruction in this study covers only the last 2000 years of the Holocene and therefore cannot assess ENSO variability during the early and middle Holocene, it does indicate that ENSO-related variability was not constant through time. In particular, the reconstruction shows enhanced variability between approximately 1500 and 1000 years BP, followed by a decline and a renewed increase towards the present. These observations are broadly consistent with the hypothesis that ENSO variability intensified during parts of the late Holocene. However, because of the reconstruction uncertainty and the limited temporal coverage, these findings should be interpreted cautiously and should not be considered conclusive evidence for long-term Holocene ENSO evolution. Nevertheless, although the primary objective of this study was methodological rather than climatological, the agreement with previous interpretations of late-Holocene ENSO variability provides additional confidence that the proposed harmonization framework is capable of capturing meaningful ENSO-related variability.

7.2 Limitations

Most of the limitations of this study are related to the availability and characteristics of the selected proxy records. Although a large number of paleoclimate proxy archives exist, only a subset provided sufficient temporal coverage, resolution, and sampling density for this thesis. Furthermore, the scope of the study had to be limited to the period from 2000 and 0 years BP because there were insufficient high-quality proxy records covering the entire Holocene. As a result, the reconstruction was based on a relatively small proxy network, which may have influenced the sensitivity and robustness of the final reconstruction.

Another limitation is the absence of a true observational ground truth for validating long-term ENSO variability. Instrumental observations cover only a relatively short period, making direct validation of the reconstruction across the Holocene impossible. Consequently, reconstruction performance was evaluated against the Niño3.4 reconstruction rather than against direct observations. However, the Niño3.4 dataset is by itself also a reconstruction and therefore contains uncertainties and methodological assumptions. The reported statistical comparisons therefore reflect agreement with the reference reconstruction rather than with the true historical ENSO signal.

An additional limitation concerns chronological uncertainty. Paleoclimate proxy records are associated with dating uncertainties that may affect the precise temporal alignment among the datasets. Although temporal binning and harmonization reduce the impact of these uncertainties, chronological errors could remain and may influence the timing of the variability. Consequently, some differences between proxy records may reflect dating uncertainty rather than climatic variability alone.

Finally, the geographical distribution of the selected proxy records introduces a potential source of bias. With the exception of the SST Niño3.4 dataset, the selected proxy archives primarily originate from North and South America, resulting in stronger representation of ENSO-related variability from the eastern Pacific region. Consequently, the reconstruction may be more representative of ENSO variability in the eastern Pacific than across the whole pacific ocean.

8 Conclusions

This thesis set out to investigate how heterogeneous ENSO-sensitive proxy records can be harmonized to produce a coherent long-term reconstruction of ENSO variability. To achieve this, a reconstruction framework was developed consisting of preprocessing, composite reconstruction and Gaussian Process Regression. The results demonstrated that, despite substantial differences in proxy type, temporal coverage, sampling density and climate sensitivity, ENSO-related variability can be extracted from heterogeneous proxy datasets.

At the same time, the results revealed substantial differences between the selected proxy records. The proxy-removal analysis and pairwise correlation tests showed that individual proxy records varied considerably in their representation of ENSO-related variability, reflecting differences in climate sensitivity, temporal coverage and location-specific climate responses. The results also indicate that Gaussian Process Regression works effectively for constructing a continuous reconstruction from sparse observations. In addition, the method provides uncertainty estimates that help quantify confidence in different parts of the reconstruction. However, the reconstruction remains sensitive to changes in kernel choice and quality of the proxy archives.

This study contributes to research on multi-proxy climate reconstruction by demonstrating that heterogeneous proxy archives can be transformed through a common framework and combined into a coherent reconstruction of ENSO variability. Unlike studies that focus on homogeneous proxy archives, this research integrates proxy records that differ substantially in proxy type. Overall, the findings suggest that the successful integration of heterogeneous proxy records depends not only on the reconstruction method itself, but also on careful harmonization and proxy selection. While limitations regarding proxy availability and validation remain, the proposed framework provides a foundation for future research on long-term reconstructions of ENSO variability using diverse paleoclimate proxy archives.

8.1 Future work

Several opportunities exist to further improve the proposed reconstruction framework. First, the proxy network could be expanded by adding additional ENSO sensitive proxy network with sufficient temporal coverage and resolution. A larger and more geographically diverse collection of proxy archives would improve the spatial representation of ENSO variability while reducing the influence of individual proxy records on the final reconstruction.

Second, Future studies could evaluate the framework against multiple existing ENSO reconstructions instead of relying on a single reference reconstruction. Such comparison would provide a better assessment of the robustness of the methodology and reduce dependence on the assumption gained from only one reference dataset.

Finally, future research could investigate the applicability of the proposed reconstruction framework to other climate phenomena and proxy networks. Applying the methodology to different paleoclimate reconstructions would help assess its generalizability and demonstrate whether the framework can be used beyond ENSO reconstruction.

References

[AANR04] ED Andrews, Ronald C Antweiler, Paul J Neiman, and F Martin Ralph. Influence of enso on flood frequency along the california coast. *Journal of Climate*, 17(2):337–348, 2004.

[AS14] Alyssa R Atwood and Julian P Sachs. Separating itcz-and enso-related rainfall changes in the galápagos over the last 3 kyr using d/h ratios of multiple lipid biomarkers. *Earth and Planetary Science Letters*, 404:408–419, 2014.

[BLM+11] J. P. Bernal, M. Lachniet, M. McCulloch, G. Mortimer, P. Morales, and E. Cienfuegos. A speleothem record of holocene climate variability from southwestern mexico. *Quaternary Research*, 75(1):104–113, 2011.

[Bra99] Raymond S Bradley. *Paleoclimatology: reconstructing climates of the Quaternary*, volume 68. Elsevier, 1999.

[DMS+22] Thierry Delcroix, SLL Michel, D Swingedouw, B Malaizé, A-L Daniau, R Abarca-del Rio, T Caley, and A-M Sémah. Clarifying the role of enso on easter island precipitation changes: Potential environmental implications for the last millennium. *Paleoceanography and Paleoclimatology*, 37(12):e2022PA004514, 2022.

[fEIN24] National Centers for Environmental Information (NCEI). Noaa/wds paleoclimatology - 700 year el niño/southern oscillation (enso) niÑo3.4 index reconstruction, 2024.

[GCL+20] Pamela R Grothe, Kim M Cobb, Giovanni Liguori, Emanuele Di Lorenzo, Antonietta Capotondi, Yanbin Lu, Hai Cheng, R Lawrence Edwards, John R Southon, Guaciara M Santos, et al. Enhanced el niño–southern oscillation variability in recent decades. *Geophysical Research Letters*, 47(7):e2019GL083906, 2020.

[HMPMS+20] Armand Hernández, Celia Martin-Puertas, Paola Moffa-Sánchez, Eduardo Moreno-Chamarro, Pablo Ortega, Simon Blockley, Kim M Cobb, Laia Comas-Bru, Santiago Giralt, Hugues Goosse, et al. Modes of climate variability: Synthesis and review of proxy-based reconstructions through the holocene. *Earth-Science Reviews*, 209:103286, 2020.

[HNdH+21] Kimberley Hagemans, Kees Nooren, Tjalling de Haas, Mario Córdova, Rick Hennekam, Martin CA Stekelenburg, Donald T Rodbell, Hans Middelkoop, and Timme H Donders. Patterns of alluvial deposition in andean lake consistent with enso trigger. *Quaternary Science Reviews*, 259:106900, 2021.

[Le23] Thanh Le. Increased impact of the el niño–southern oscillation on global vegetation under future warming environment. *Scientific Reports*, 13(1):14459, 2023.

[LFZ22] Aifeng Lv, Lei Fan, and Wenxiang Zhang. Impact of enso events on droughts in china. *Atmosphere*, 13(11):1764, 2022.

[LLZC18] Zhengyao Lu, Zhengyu Liu, Jiang Zhu, and Kim M Cobb. A review of paleo el niño-southern oscillation. *Atmosphere*, 9(4):130, 2018.

[LMZZ24] Yujia Liu, Wenmin Man, Tianjun Zhou, and Meng Zuo. Global multiproxy enso reconstruction over the past millennium. *Journal of Geophysical Research: Atmospheres*, 129(10):e2023JD040491, 2024.

[LNA10] Bo Li, Douglas W Nychka, and Caspar M Ammann. The value of multiproxy reconstruction of past climate. *Journal of the American Statistical Association*, 105(491):883–895, 2010.

[LXC+11] J. Li, S.-P. Xie, E. R. Cook, G. Huang, R. D. D'Arrigo, F. Liu, J. Ma, and X.-T. Zheng. Interdecadal modulation of el niño amplitude during the past millennium. *Nature Climate Change*, 1(2):114–118, 2011.

[MARM22] Samuel Z Mark, Mark B Abbott, Donald T Rodbell, and Christopher M Moy. Xrf analysis of laguna pallcacocha sediments yields new insights into holocene el niño development. *Earth and planetary science letters*, 593:117657, 2022.

[MZG06] Michael J McPhaden, Stephen E Zebiak, and Michael H Glantz. Enso as an integrating concept in earth science. *science*, 314(5806):1740–1745, 2006.

[PDFB21] Gregor Pfalz, Bernhard Diekmann, Johann-Christoph Freytag, and Boris K Biskaborn. Harmonizing heterogeneous multi-proxy data from lake systems. *Computers & Geosciences*, 153:104791, 2021.

[PVG+11] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.

[Pyt26] Python Software Foundation. Python programming language, 2026.

[RFCMMC+26] Belén Rodríguez-Fonseca, Elena Calvo-Miguélez, Lucia Montoya-Carramolino, Regina R Rodrigues, Irene Polo, Marta Martín-Rey, Teresa Losada, Jorge López-Parages, Iñigo Gómara, David Rivas, et al. Enso impacts on marine ecosystems and fisheries in the tropical and south atlantic. *Nature Reviews Earth & Environment*, 7(1):43–59, 2026.

[RTH+19] Jason Roberts, CR Tozer, M Ho, AS Kiem, Tessa Vance, Lenneke Jong, Felicity McCormack, and Tasman van Ommen. Reconciling unevenly sampled paleoclimate proxies: a gaussian kernel correlation multiproxy reconstruction. 2019.

[TAK+18] Axel Timmermann, Soon-Il An, Jong-Seong Kug, Fei-Fei Jin, Wenju Cai, Antonietta Capotondi, Kim M Cobb, Matthieu Lengaigne, Michael J McPhaden, Malte F Stuecker, et al. El niño–southern oscillation complexity. *Nature*, 559(7715):535–545, 2018.

[THN+19] John R Tipton, Mevin B Hooten, Connor Nolan, Robert K Booth, and Jason McLachlan. Predicting paleoclimate from compositional data using multivariate gaussian process inverse prediction. *The Annals of Applied Statistics*, 13(4):2363–2388, 2019.

[WR06] Christopher KI Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA, 2006.

A Additional configuration results

Kernel	Length Scale	Fixed	Pearson	Spearman	RMSE	Mean Uncertainty
Matérn	0.3	True	0.517	0.473	0.855	0.313
Matérn	0.3	False	0.431	0.407	0.863	0.306
Matérn	0.5	True	0.611	0.522	0.848	0.324
Matérn	0.5	False	0.431	0.407	0.863	0.306
Matérn	1.0	True	0.692	0.593	0.844	0.339
Matérn	1.0	False	0.431	0.407	0.863	0.306
RBF	0.3	True	0.755	0.621	0.851	0.338
RBF	0.3	False	0.494	0.484	0.860	0.317
RBF	0.5	True	0.761	0.610	0.816	0.359
RBF	0.5	False	0.494	0.484	0.860	0.317
RBF	1.0	True	0.509	0.352	0.851	0.386
RBF	1.0	False	0.494	0.484	0.860	0.317
RQ	0.3	True	0.636	0.549	0.851	0.328
RQ	0.3	False	0.494	0.484	0.860	0.317
RQ	0.5	True	0.524	0.484	0.855	0.321
RQ	0.5	False	0.494	0.484	0.860	0.317
RQ	1.0	True	0.497	0.484	0.856	0.319
RQ	1.0	False	0.494	0.484	0.860	0.317

Table 6: Complete statistical comparison metrics for different GPR kernel and parameter configurations.

Composite	N	Removed Proxy	Pearson	Spearman	RMSE	Mean Uncertainty
Mean	4	NAN	0.692	0.593	0.844	0.339
Mean	3	PC1_Pallcacocha	0.691	0.593	0.652	0.297
Mean	3	SST_Nino	0.412	0.418	0.955	0.398
Mean	3	Tree_rings	-0.668	-0.610	0.947	0.340
Mean	3	galapogos_biomarkers	0.718	0.588	0.899	0.354
Median	4	NAN	0.565	0.544	0.841	0.359
Median	3	PC1_Pallcacocha	0.603	0.544	0.682	0.313
Median	3	SST_Nino	-0.605	-0.341	0.985	0.450
Median	3	Tree_rings	-0.454	-0.264	0.895	0.359
Median	3	galapogos_biomarkers	0.688	0.577	0.906	0.366
Weighted	4	NAN	0.666	0.555	0.866	0.338
Weighted	3	PC1_Pallcacocha	0.687	0.582	0.666	0.295
Weighted	3	SST_Nino	0.261	0.220	0.959	0.396
Weighted	3	Tree_rings	-0.705	-0.698	0.971	0.337
Weighted	3	galapogos_biomarkers	0.704	0.555	0.936	0.352

Table 7: Statistical comparison metrics for different proxy removal configurations and composite reconstruction methods.