

Temperature affinities of dinocysts in the Southern Ocean during the Pliocene - Pleistocene compared to present time

Name: Florenca Kropman

Student number: 7324669

Institution: Utrecht University

Faculty: Geosciences

Supervisor: Dr. Peter Bijl

Second assessor: Dr. Lena Thöle

Master: Earth, Life and Climate

Course: GEO MSc Thesis Earth Sciences programmes

Version: Final version

Date: 25-1-2025



**Universiteit
Utrecht**

Table of content:

1. Abstract	3
2. Introduction.....	4
2.1 Dinoflagellate cysts	4
2.2 Literature thus far and research aims	5
3. Materials.....	8
3.1 Oceanographic setting	8
3.2 Modern data.....	9
3.3 Drill core locations	9
4. Methodology	11
4.1 Reproducing Thöle et al. (2023) in Python.....	11
4.2 Prepping paleo data	12
4.3 Adding paleo samples to cluster analysis.....	12
4.4 Individual dinocyst species analysis	12
5. Results.....	13
5.1 Cluster analysis.....	13
5.1.1 Oceanographic conditions.....	13
5.1.2 Cluster temperature compared to proxy records	16
5.2 Individual dinocyst species analysis	18
6. Discussion	20
6.1 Temperature reconstruction comparison to proxy temperature.....	20
6.1.1 North of Subtropical front.....	20
6.1.2 Between Subtropical front and Subantarctic front	20
6.1.3 South of Polar Front.....	20
6.2 Other oceanographic features	22
6.2.1 Salinity.....	22
6.2.2 Sea ice presence	22
6.2.3 Nutrients	22
6.3 Broader implications	22
7. Conclusion.....	24
8. Bibliography.....	25
9. Appendix	29
Appendix A: Boxplots PD	29
Appendix B: Boxplots ODP 1168.....	31
Appendix C: Boxplots ODP 1171.....	32
Appendix D: Boxplots ODP 1172	33
Appendix E: Boxplots MD11-3353	34
Appendix F: Boxplots U1475.....	35
Appendix G: Boxplots AS05-10.....	36
Appendix H: Cluster against Age	37
Appendix I: Individual dinocyst species analysis: relative abundance against temperature	38

1. Abstract

In this study, a quantitative analysis is conducted to study the Southern Ocean temperature affinities of dinocysts during the Pliocene and Pleistocene in comparison to the present day. This is done to investigate whether dinocyst species maintain consistent environmental preferences over time, thereby assessing their reliability as a proxy for temperature. Two statistical analyses were performed in Python: the k-means clustering method and the individual dinocyst species analysis. The Python script was partly adapted from the RStudio model developed by Thöle et al. (2023), which was originally designed for modern dinocyst data. Six downcore datasets, covering a total time period from 5.33 Ma until present, are distributed across the Southern Ocean both latitudinally and longitudinally, were analysed. Although results from the k-means clustering analysis vary per core, overall there is consistency in dinocyst cluster-based and proxy-based SST reconstructions. However, the agreement between temperature reconstructions diminishes substantially when dinocyst species diversity in the drilling core dataset is low. To improve the dinocyst-based proxy for temperature, better spatial coverage of modern Southern Ocean surface sediment samples and more tests from various settings and regions around Antarctica. Furthermore, the individual dinocyst species analysis highlights the dominant role of *Brigantedinium* spp. in shaping cluster patterns, with its relative abundance varying across different latitudes between the past and the present. This study validates this quantitative approach as a promising method for examining dinocyst temperature affinities through time.

2. Introduction

Understanding the Earth's climate system is more pressing than ever, now that a global warming acceleration is detected (Hansen & Sato, 2020), and the rapidly changing modern climate has challenged scientists to study past records of climates and rapid changes (Hatté & Schwartz, 2003). To better understand the Earth's present climate and its possible future changes, it is vital to gain a deeper understanding of paleoclimates.

Since direct measurements of paleoclimates are impossible, proxies are used. A proxy is a measurable quantity, that serves as an indirect indicator of a specific environmental parameter that cannot be measured directly (Wefer et al., 1999), like the isotopic composition of oxygen in ocean water to reconstruct the ice cap volume for instance temperature. Proxies inherently come with uncertainties, error margins and limitations, and some environmental features do not even have a proxy yet like surface water salinity (Wefer et al., 1999). Given the scarcity of reliable proxies, developing potential proxies and their properties is of great importance.

The Southern Ocean is particularly important for understanding paleoclimates and the modern climate dynamics. It connects the world's ocean basins and plays a crucial role in the thermohaline circulation by linking the shallow branches (Rintoul, 2011). Additionally, it functions as a major carbon sink area, meaning it is crucial for the global carbon cycle and climate since it is essential in regulating atmospheric CO₂ concentrations (McElroy, 1983). Furthermore, the Southern Ocean, in particular the strong Antarctic Circumpolar Current (ACC), prevent warm waters to reach Antarctica, leading to a colder Antarctica (Wang et al., 2024). Proxy data yielding information on the past behaviour of the Southern Ocean is crucial for a better understanding of past climate systems (Crosta et al., 2022).

2.1 Dinoflagellate cysts

Dinoflagellate cysts (dinocysts) are one type of proxy that can be used to reconstruct palaeoceanographic features due to their occurrence across a wide range of oceanographic conditions (de Vernal et al., 2013). Dinoflagellates (dinos) are protists ranging in size from 0.01 to 2.0 mm, with species exhibiting substantial variation in morphology, nutrition and habitat (Spector, 2012), which makes them reflective of a wide range of environmental conditions. For instance, dinos are commonly found in both marine and freshwater environments (Taylor et al., 2008).

Dinocysts are valuable proxies for reconstructing past oceanographic parameters, for instance temperatures, salinity and nitrate concentrations (Radi & de Vernal, 2008), as oceanographic factors affect the morphology of the dinocyst species by affecting the cyst formation processes (Bravo & Figueroa, 2014). The composition of dinocyst assemblages is strongly shaped by oceanographic, biological, physical and chemical conditions (Zonneveld et al., 2013). To apply this knowledge to interpret downcore assemblages in terms of oceanography, testing and proxy-proxy intercomparison is required.

During their life cycle, dinos could produce cysts (Bravo & Figueroa, 2014). The resting cysts enable survival under adverse conditions, awaiting more favourable environmental circumstances, thereby protecting the organism (Matsuoka & Fukuyo, 2000). Dinocysts represent the only fossilizable remains of dinoflagellates (Taylor, 1980). Currently, approximately 80 extant dino species are known to produce cysts (Zonneveld et al., 2013), a small fraction of the more than 2000 known dino species (Matsuoka & Fukuyo, 2000). The morphology of dinocysts is related to both a range of environmental factors like temperature, salinity and nutrients, and the genetics of the motile dino from which it derives which means we can use dinocyst morphology as an expression of the diversity of dinos (keeping in mind that not all dinos produce a cyst) (Hoyle et al., 2019).

In order to understand and use dinocyst assemblages as a proxy for oceanographic conditions, the ecological affinities of the dinos that produce the cyst need to be known. This is studied by investigating sedimentary dinocyst compositions from modern ocean sediments throughout the world, compared to the oceanographic properties of the surface water overlying those sediment locations (Pospelova et al., 2008).

2.2 Literature thus far and research aims

The Southern Ocean dinocyst compilation of Marret et al. (2001) is the first study to assemble modern dinocysts assemblages, and the atlas of Marret and Zonneveld (2003) is a prime global collection regarding the modern distribution of dinocysts. Esper and Zonneveld (2007), Zonneveld et al. (2013), Prebble et al. (2013), Nooteboom et al. (2019), Marret et al. (2020) and Thöle et al. (2023) are the other modern dinocyst compilation studies.

Thöle et al. (2023) compiled a dataset of modern dinocyst distributions in the Southern Hemisphere, to better understand the relationship between dinocyst assemblages and the oceanographic conditions on the Southern Hemisphere. Using this dataset, a k-means clustering analysis was performed in RStudio, yielding nine clusters, each assigned surface water parameters based on surface models (Thöle et al., 2023). Thöle et al. (2023) demonstrates the potential of dinocyst assemblages as oceanographic proxy for ocean conditions (relative ocean warmth, relative changes in sea ice and nutrients). Hou, Stap, et al. (2023) interpreted broad paleotemperature from the overall dinocyst assemblages in the Miocene, following the cluster analysis of Thöle et al. (2023), but this effort lacked a solid statistical approach to assign clusters to downcore dinocyst assemblages.

This study build on that work by adapting the clustering approach to Python and extending it to include paleo data. By comparing downcore dinocyst assemblages with modern data and other proxy records for oceanographic conditions, such as TEX₈₆, it becomes possible to investigate whether dinocyst species maintain consistent oceanographic preferences over time, thereby assessing their reliability as a proxy for temperature. The decision to focus on temperature is based on the fact that temperature is an important environmental and climatological factor, and that there are more proxies available making proxy-proxy comparisons possible.

This research examines the temperature affinities of Southern Hemisphere dinocyst from the Neogene to Quaternary. The analyses were performed using a custom script developed in Spyder with Python. This is one of the few quantitative studies that have been performed for dinoflagellate cysts; an example is the study of De Schepper et al. (2011).

Two statistical methods were employed. The first involved a k-means clustering approach in which paleo samples were assigned to clusters defined by modern data, with each cluster representing distinct values for multiple oceanographic parameters. This enables a better understanding of environmental preferences of dinocyst assemblages. The clustering is method is chosen because it has already been used by Thöle et al. (2023) and Hou, Stap, et al. (2023) and this research can extend the knowledge concerning this method and also be linked to their data. The clustering method is also chosen because it means the spatial relationship between oceanographic conditions and dinocyst assemblages can be studied (Thöle et al., 2023).

The second analysis compared species individually, focusing on relative abundance and temperature. If a paleo sample was unexpectedly assigned to a particular cluster, the individual analysis could help identifying the species responsible. This approach enabled the identification of specific dinocyst species or assemblages that could serve as a reliable temperature proxy. This approach is chosen as an explanatory method to see if this method statistically works and if the results are complementary to the k-means clustering analysis.

<i>Dinocyst species:</i>	<i>Abbreviation:</i>
<i>Achomosphaera andalousiensis</i>	Aand
<i>Ataxiodinium choane</i>	Atax
<i>Barssidinium spp.</i>	Bars
<i>Brigantedinium spp.</i>	Bspp
<i>Bitectatodinium spongium</i>	Bspo
<i>Bitectatodinium tepikiense</i>	Btep
<i>Cerebrocysta spp.</i>	Cspp
<i>Corrudinium spp.</i>	Cosp
<i>Cryodinium meridianum</i>	Cmer
<i>Dalella chathamense</i>	Dcha
<i>Dapsilidinium spp.</i>	Dpas
<i>Dubridinium spp.</i>	Dubr
<i>Echinidinium aculeatum</i>	Eacu
<i>Echinidinium delicatum</i>	Edel
<i>Echinidinium granulatum</i>	Egra
<i>Echinidinium karaense</i>	Ekar
<i>Echinidinium spp.</i>	Espp
<i>Echinidinium transparantum</i>	Etra
<i>Echinidinium zonneveldlae</i>	Ezon
<i>Cyst of Gymnodinium spp.</i>	Gymn
<i>Impagidinium aculeatum</i>	Iacu
<i>Impagidinium pallidum</i>	Ipal
<i>Impagidinium paradoxum</i>	Ipar
<i>Impagidinium patulum</i>	Ipat
<i>Impagidinium plicatum</i>	Ipli
<i>Impagidinium sphaericum</i>	Isph
<i>Impagidinium strialatum</i>	Istr
<i>Impagidinium variaseptum</i>	Ivar
<i>Impagidinium velorum</i>	Ivel
<i>Islandinium brevispinosum</i>	Ibre
<i>Islandinium cezare</i>	Imic
<i>Islandinium minutum</i>	Imin
<i>Lejeunecysta oliva</i>	Loli
<i>Lejeunecysta spp.</i>	Lspp
<i>Lingulodinium machaerophorum</i>	Lmac
<i>Melitasphaeridium choanophorum</i>	Mcho
<i>Nematosphaeropsis labyrinthus</i>	Nlab
<i>Nucicla umbiliphora</i>	Numb
<i>Operculodinium centrocarpum sensu Wall and Dale 1966</i>	Ocen
<i>Operculodinium eirikianum</i>	Oeir
<i>Operculodinium israelianum</i>	Oisr
<i>Operculodinium janduchenei</i>	Ojan
<i>Operculodinium longispinigerum</i>	Olon
<i>Operculodinium spp.</i>	Ospp
<i>Cyst of Pentapharsodinium dalei</i>	Pdal
<i>Peridinoids</i>	Peri
<i>Cyst of Polarella glacialis</i>	Pgla

<i>Cyst of Polykrikos kofoidii</i>	Pkof
<i>Cyst of Polykrikos schwartzii</i>	Psch
<i>Polysphaeridium spp.</i>	Poly
<i>Polysphaeridium zoharyi</i>	Pzoh
<i>Cyst of Protoperidinium americanum</i>	Pame
<i>Cyst of Protoperidinium nudum</i>	Pnud
<i>Pyxidinopsis psilate</i>	Ppsi
<i>Pyxidinopsis reticulate</i>	Pret
<i>Pyxidinopsis spp.</i>	Pyxi
<i>Quinquecuspis concreta</i>	Qcon
<i>Selenopemphix antarctica</i>	Sant_all
<i>Selenopemphix dionaeacysta</i>	Sdio
<i>Selenopemphix nephroides</i>	Snep
<i>Selenopemphix quanta</i>	Squa
<i>Spiniferites belerius</i>	Sbel
<i>Spiniferites bulloideus</i>	Sbul
<i>Spiniferites delicatus</i>	Sdel
<i>Spiniferites dentacus</i>	Sden
<i>Spiniferites elongatus</i>	Selo
<i>Spiniferites lazus</i>	Slaz
<i>Spiniferites membranaceus</i>	Smem
<i>Spiniferites mirabilis (includes S. hyperacanthus)</i>	Smir
<i>Spiniferites pachydermus</i>	Spac
<i>Spiniferites pseudofurcatus</i>	Spse
<i>Spiniferites ramosus</i>	Sram
<i>Spiniferites ramosus and Spiniferites bulloideus</i>	Sram_bul
<i>Spiniferites rubinus</i>	Srub
<i>Spiniferites spp.</i>	Sspp
<i>Stelladinium reidii</i>	Srei
<i>Stelladinium robustum</i>	Srob
<i>Tectatodinium pellitum</i>	Tpel
<i>Trinovantedinium applanatum</i>	Tapp
<i>Tuberculodinium vancampoae</i>	Tval
<i>Votadinium calvum</i>	Vcal
<i>Votadinium spinosum</i>	Vspi
<i>Xandarodinium xanthum</i>	Xxan

Table 1: Species list

3. Materials

3.1 Oceanographic setting

The Southern Ocean is characterized by the Antarctic Circumpolar Current (ACC) is currently the strongest current in the world (Barker & Thomas, 2004), and influences the global climate significantly with sinking carbon and thus reducing atmospheric CO₂ (Böning et al., 2008). This current started approximately 33.5 Ma when the Drake Passage started to open up and began to fully thrive with the deepening of the passage (Kennett, 1977).

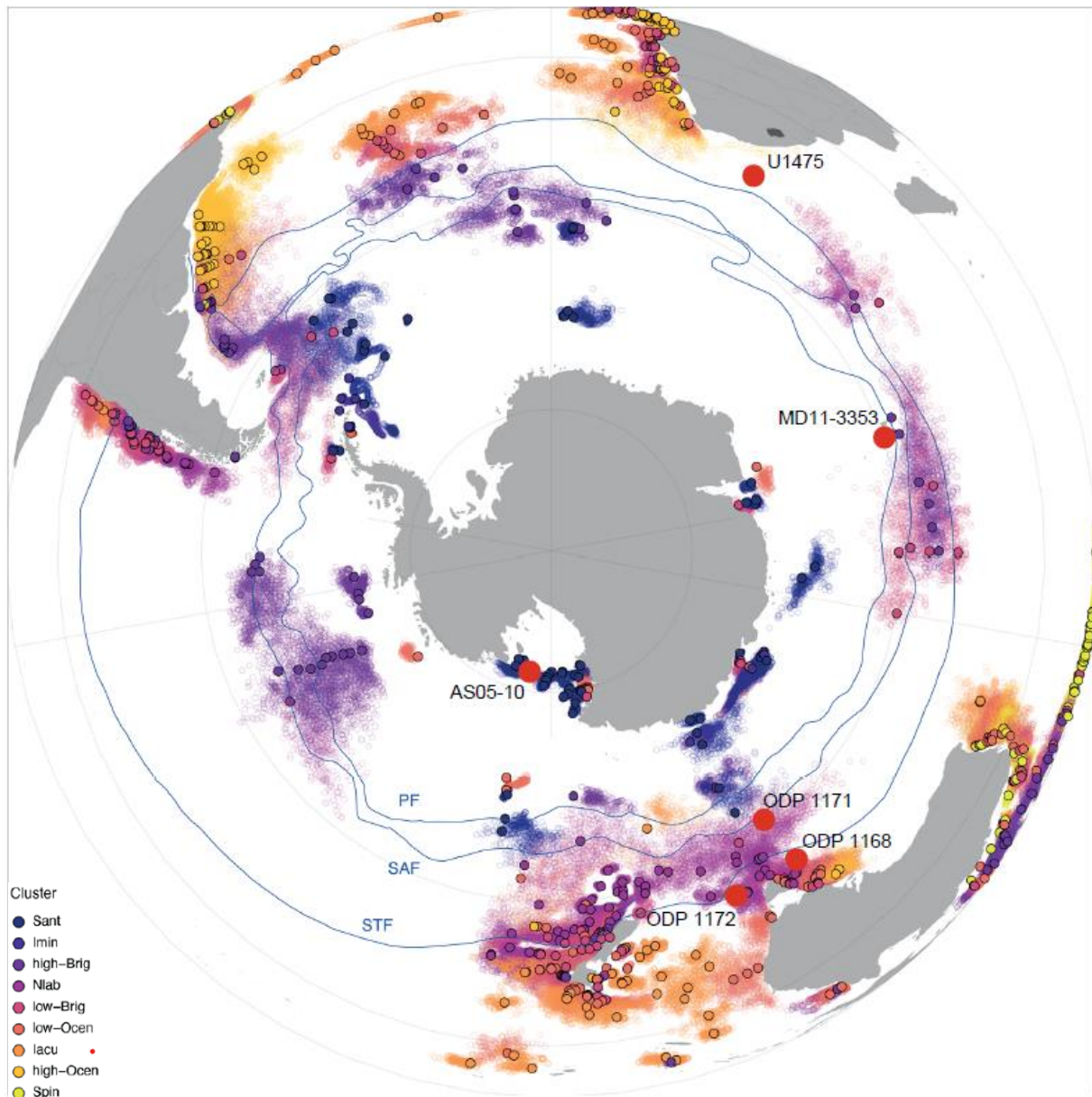


Figure 1: Map of the research area based on Thole et al. (2023). Coloured dots represent SH_655 datapoints. The colours represent the different clusters which the samples are assigned to. RedU14 dots represent drill cores used in this study. PF – Polar Front, SAF – Subantarctic Front, STF – Subtropical Front.

The ACC includes three major oceanographic fronts based on physical and chemical characteristics: the Subtropical Front (STF), the Subantarctic Front (SAF), and the Polar Front (PF) (Sokolov & Rintoul, 2002). These fronts drive latitudinal variations in oceanic conditions in the Southern Ocean. The STF, the warmest of three, has summer temperatures of approximately 16 °C (Sikes et al., 2009).

The SAF, the strongest front, serves as the main motor of the ACC (Sokolov & Rintoul, 2002). The PF, the southernmost front, features the coldest temperatures, reaching as low as -2 °C (Veth et al., 1997).

3.2 Modern data

Modern dinocyst data were sourced from Thöle et al. (2023). In their dataset, a total of 73 new surface sediment samples were examined, 66 of them located close to Antarctica. Most of the samples are located in close proximity to land, for instance the coasts of New Zealand, southern Africa and South America, with the Atlantic Ocean having the highest concentration of samples (Thöle et al., 2023).

These new Southern Hemisphere samples were added by Thöle et al. (2023) to the latest global compilation of dinocyst surface sediments constructed by Marret et al. (2020). This leads up to a total amount of 655 samples and the dataset is called SH_655 (Thöle et al., 2023), and PD in this research and its Python scripts.

<i>Cluster:</i>	<i>Name:</i>
<i>Cluster 1</i>	Low-Ocen Cluster
<i>Cluster 2</i>	Sant Cluster
<i>Cluster 3</i>	Sspp Cluster
<i>Cluster 4</i>	Low-Bspp Cluster
<i>Cluster 5</i>	Iacu Cluster
<i>Cluster 6</i>	Nlab Cluster
<i>Cluster 7</i>	High-Bspp Cluster
<i>Cluster 8</i>	Imin Cluster
<i>Cluster 9</i>	High-Ocen Cluster

Table 2: Cluster names

3.3 Drill core locations

This study uses multiple datasets to examine dinocysts and their assemblages over time: ODP 1168, ODP 1171, ODP 1172, MD11-3353, U1475 and AS0510 (Figure 1). These cores were used since they have a great spatial distribution, they all have a proxy temperature record and a palynology dataset.

Ocean Drilling Program (ODP) 1168 is located at a depth of 2463 meters on Tasmania's western margin slope, near the STF (Exon et al., 2001). This entire core spans from the Late Eocene to the Holocene (Stickley et al., 2004), but in this study 5.3 to approximately 2.5 Ma is used. Going further down in time would increase the amount of extinct species. Furthermore, its palynology (Hou, Stap, et al., 2023) and U^{K}_{37} proxy data (Hou, Lamprou, et al., 2023) was used.

The ODP 1171 drilling site, situated on the southwestern slope of the South Tasman Rise at 215-meters depth (Exon et al., 2001), lies between the Subantarctic Front (SAF) and the Subtropical Front (STF). The entire core spans from the early Eocene to the Holocene (Exon et al., 2001), but for this study only to upper 450 kyr was used. Furthermore, the palynology constructed by Thöle et al. (2023) and the TEX_{86} data (Leutert, 2020) were used.

ODP 1172 is located at a depth of 2620 meters on the East Tasman Plateau (Brinkhuis et al., 2003), in cool subtropic water north of the STF (Exon et al., 2001). This core spans from the Late Cretaceous to the Holocene (Exon et al., 2001), for this study approximately 550 ka to 10 ka is used. Furthermore, the palynology (Thöle et al., 2023) and Mg/Ca proxy records were used (Nürnberg & Groeneveld, 2006); there is no TEX_{86} data for this core.

The MD11-3353 drilling core, drilled south of the French Southern and Antarctic Lands in the Polar Front Zone (Ai et al., 2020) at 1568 meters depth (Civel-Mazens et al., 2021). In this study the time range of approximately 170 ka to 70 ka is used. Furthermore, the palynology data and the TEX_{86} data (Thöle et al., 2019) were used.

U1475 is situated on the southwestern flank of the Agulhas Plateau, at a depth of 2669 meters (Hall et al., 2017). This area was influenced by multiple water masses, including the Lower Circumpolar Deep Water, the Indian Ocean Water and the North Atlantic Deep Water (Hall et al., 2017). The part of the

core used in this study spans from approximately 4.6 to 2.95 Ma, and its palynology and U^{K}_{37} proxy data were analysed both constructed by Hou (2024).

The AS05-10 site, is located in the Ross Sea, far south of the PF (Hartman et al., 2021), which is the closest of all sites to Antarctica. The time period used in this study spans approximately from 320 ka to 75 ka, the palynology and TEX_{86} data (Hartman et al., 2021) are used.

4. Methodology

4.1 Reproducing Thöle et al. (2023) in Python

The main structure at the beginning of was primarily derived from the R code developed by Thöle et al. (2023), which is now rewritten in Python and operates consistently. For the PD dataset, clustering was performed identically to the method described by Thöle et al. (2023), also with a similar amount of clusters ($k = 9$).

Next to the structure of the clustering method, the structure of the connection of samples to the surface variables was derived from Thöle et al. (2023). This connection leads to the determination of oceanographic surface variables values to samples, both past and present. Boxplots show the median and range of values of different oceanographic parameters across clusters. These will be used to construct graphs of the various oceanographic parameters per dataset through time.

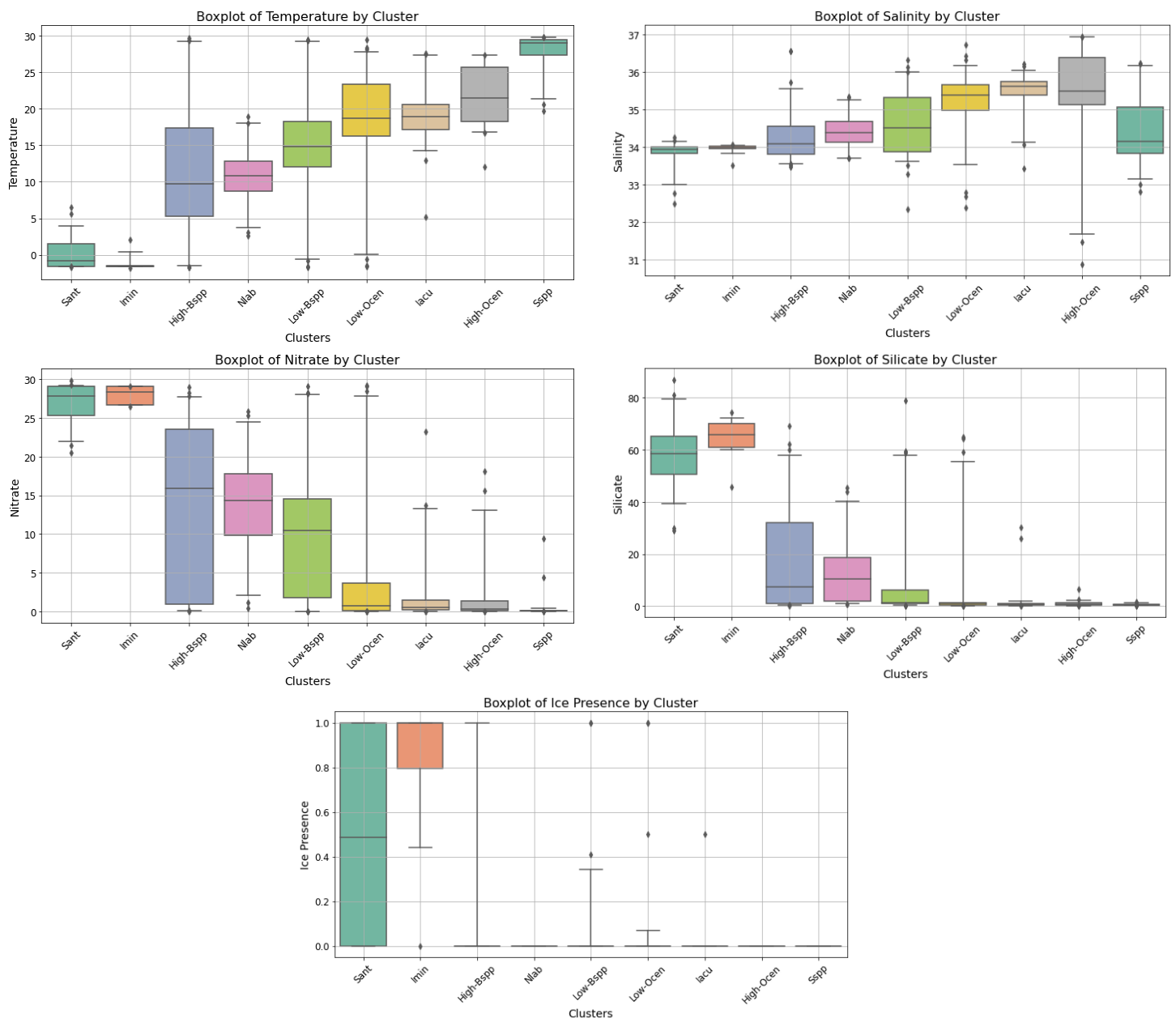


Figure 2: Surface water features per cluster based on the median for the : temperature, salinity, nitrate, silicate and ice presence

4.2 Prepping paleo data

The paleo data needs to be prepared correctly, before it can be used for the statistical analyses. Datasets of drill cores are uploaded and reshaped into a format suitable for the k-means clustering analysis. Not all dinocyst species in the drill core datasets are represented in the PD dataset. This discrepancy may arise due to several factors: (1) the dinocyst species is currently only found in the Northern Hemisphere, (2) the dinocyst species is extinct, (3) the dinocyst species exists in the Southern Hemispheres in very low abundances, and thus was excluded from PD or (4) the dinocyst species is categorized in the same genus under another species name.

Depending on the reason for the absence of a dinocyst species in PD, different approaches are required to address this issue. Expert knowledge is needed to ensure the correct decisions are made. When encountering a species absent from PD, firstly determine if it is extant or extinct. If extant, include the species to the PD dataset with a column entirely full with zeroes. When extinct the species can be grouped under a related dinocyst species within the same genus. If this is not possible the nearest species approach based on taxonomy is used.

Prior to clustering downcore data, it is essential to ensure that all datasets contained identical columns (dinocyst species) as PD, and that they are aligned in correct order (PD as the reference frame). With the Correspondence Analysis, it is examined if columns in the downcore datasets still need to be processed.

4.3 Adding paleo samples to cluster analysis

For the cluster analysis, the k-means algorithm is used. This is a widely applied clustering method that identifies cluster structure within datasets (Sinaga & Yang, 2020). This clustering method is also used by Thöle et al. (2023). The PD cluster boundaries determined by the k-means algorithm were fixed and remained unchanged during the analysis of downcore samples. For the downcore data, the focus is on determining in which clusters the samples belong without altering the cluster boundaries and centroids.

Proxy data for sea surface temperature (SST) was alongside the clustered temperature data for each drill core. For the U1475 core linear interpolation was necessary due to discrepancies between the age columns of the dinocyst and SST datasets. These discrepancies were minor and thus suitable for interpolation. For other datasets, interpolation was not required and the proxy SST data was used directly. U1475 was the only core that required interpolation in order to perform this study.

4.4 Individual dinocyst species analysis

In the individual analysis, dinocyst species were studied separately to investigate oceanographic affinities of individual species in the downcore samples, and compared it to their PD affinities. It is necessary to reevaluate the rearrangement of dinocyst species made for the cluster analysis, as alignment with PD was no longer necessary, thus requiring some transformation of the downcore datasets. For instance: extinct species no longer needed to be grouped with extant species for inclusion in the analysis. The categorization is based on taxonomic and morphological similarities or supra generic affinities. Since the objective is to evaluate species individually, it is preferable to minimize species grouping.

For the temperature analysis, the proxy data is used in the same manner as in the cluster analysis. For the U1475 core linear interpolation was necessary due to discrepancies between the age columns of the dinocyst and SST datasets. For the SST of the PD dataset, the data of Thöle et al. (2023) was linked to surface variables derived from the same study.

5. Results

5.1 Cluster analysis

5.1.1 Oceanographic conditions

The dominant cluster (most frequently abundant) for ODP 1168 is the *Low-Ocen* Cluster, corresponding to a temperature range of 16.3 to 23.6 °C (Figure 3.1). Several notable anomalies are observed, including two extended periods of fluctuation from 5100 – 4700 ka and from 3300 – 3100 ka (Table 3). Additional two distinct anomalies occur around 4200 ka and 3600 ka. Minor deviations are excluded from Table 3 due to their relatively minor divergence.

ODP 1168	Dominant cluster values	5100 - 4700 ka	4200 ka	3600 ka	3300 - 3100 ka
Temperature (°C)	16.2 - 23.4	8.7 - 29.4	8.7 - 12.8	8.7 - 12.8	8.7 - 25.7
Salinity (psu)	35.0 - 35.7	33.8 - 35.8	34.1 - 34.7	34.1 - 34.7	34.1 - 36.4
Nitrate (mmol/m ³)	0.1 - 3.7	0.1 - 17.8	9.9 - 17.8	9.9 - 17.8	0.1 - 17.8
Silicate (mmol/m ³)	0.5 - 1.3	0.3 - 18.6	1.8 - 18.6	1.8 - 18.6	0.4 - 18.6
Sea ice (%)	0	0	0	0	0
Clusters abundant	<i>Low-Ocen</i>	<i>Low-Ocen, Sspp, Iacu, Nlab</i>	<i>Nlab</i>	<i>Nlab</i>	<i>Low-Ocen, Iacu, Nlab, High-Ocen</i>

Table 3: Anomaly data ODP 1168

Drill core ODP 1171 exhibits a consistent trend with five irregularities observed across most oceanographic conditions (Figure 3.2). The most prevalent cluster for this dataset is the *Nlab* Cluster, associated with a temperature range of 8.7 to 12.8 °C, making it dominant. The anomalies are identified at approximately 450, 380, 300, 250 and 20 ka (Table 4).

ODP 1171	Dominant cluster values	450 ka	380 ka	300 ka	250 ka	20 ka
Temperature (°C)	8.7 - 12.8	16.2 - 23.4	-1.5 - 1.5	17.2 - 20.6	16.2 - 23.4	16.2 - 23.4
Salinity (psu)	34.1 - 34.7	35.0 - 35.7	33.8 - 34.0	35.4 - 35.8	35.0 - 35.7	35.0 - 35.7
Nitrate (mmol/m ³)	9.9 - 17.8	0.1 - 3.7	25.3 - 29.1	0.2 - 1.5	0.1 - 3.7	0.1 - 3.7
Silicate (mmol/m ³)	1.8 - 18.6	0.5 - 1.3	50.5 - 65.1	0.5 - 0.0	0.5 - 1.3	0.5 - 1.3
Sea ice (%)	0	0	0 - 100	0	0	0
Clusters abundant	<i>Nlab</i>	<i>Low-Ocen</i>	<i>Sant</i>	<i>Iacu</i>	<i>Low-Ocen</i>	<i>Low-Ocen</i>

Table 4: Anomaly data ODP 1171

Core ODP 1172 shows multiple anomalies at both the beginning and end of the represented time period, while the middle section is relatively stable (Figure 3.3). Phases of anomalies are observed, each characterized by multiple deviations in oceanographic conditions: 600 – 450, 440 – 410, 360 – 320, 140 and 90 – 5 ka (Table 5). The *Nlab* Cluster is dominant, with a temperature range of 8.7 to 12.8 °C.

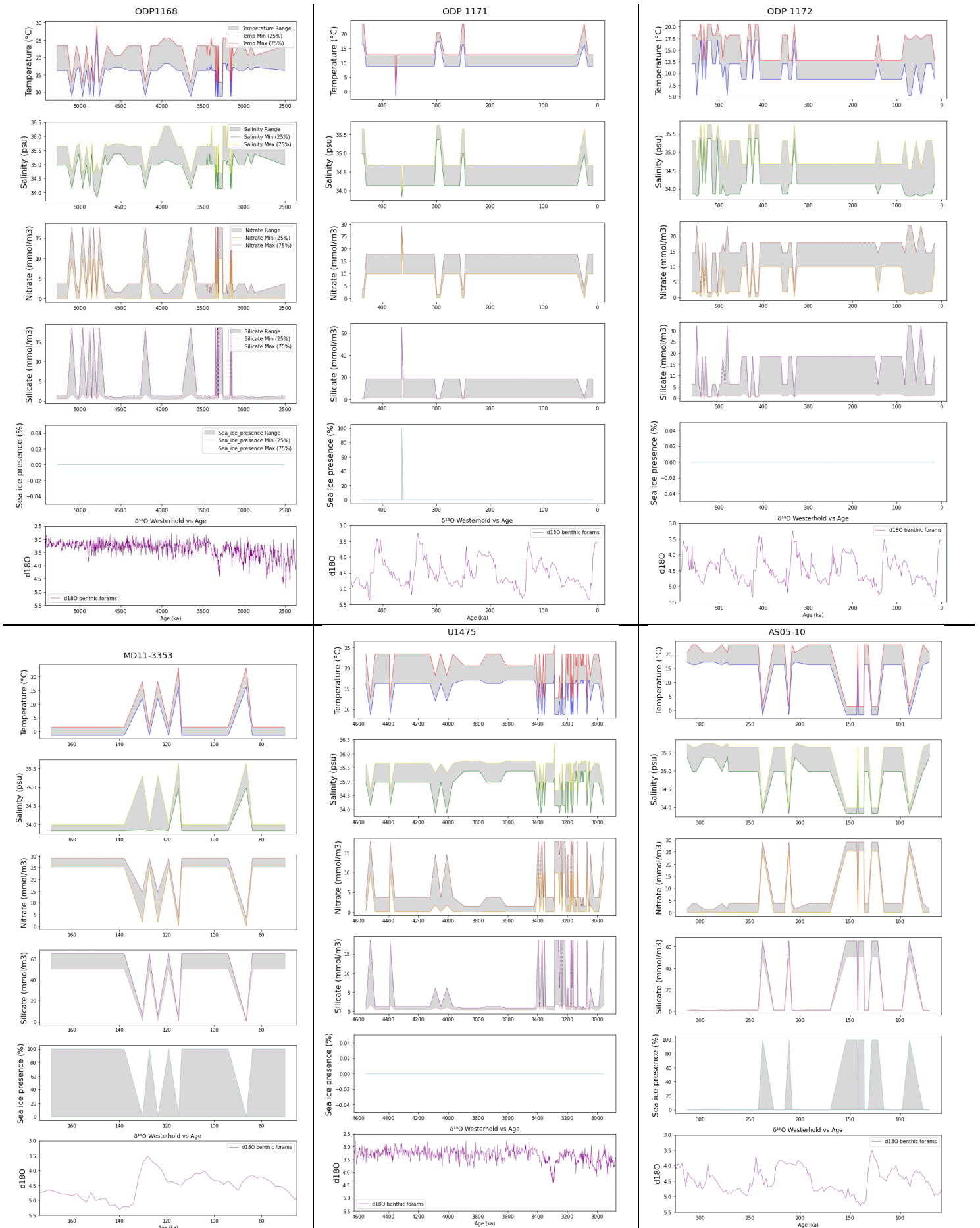


Figure 3: Oceanographic conditions per drill core through time based on k-means cluster analysis

ODP 1172	Dominant cluster values	600 - 450 ka	440 - 410 ka	360 - 320 ka	140 ka	90 - 5 ka
Temperature (°C)	8.7 - 12.8	5.2 - 20.6	8.7 - 20.6	8.7 - 20.6	12.1 - 18.2	5.2 - 18.2
Salinity (psu)	34.1 - 34.7	33.8 - 35.7	33.9 - 35.7	33.9 - 35.7	33.9 - 35.3	33.8 - 35.3
Nitrate (mmol/m ³)	9.9 - 17.8	0.2 - 23.5	0.2 - 17.8	0.2 - 17.8	1.8 - 14.5	1.0 - 23.5
Silicate (mmol/m ³)	1.8 - 18.6	0.5 - 32.1	0.5 - 18.6	0.5 - 18.6	1.1 - 6.2	0.9 - 32.1
Sea ice (%)	0	0	0	0	0	0
Clusters abundant	<i>Nlab</i>	<i>Low-Bspp, Iacu, Nlab, High-Bspp</i>	<i>Iacu, Nlab</i>	<i>Low-Bspp, Iacu, Nlab,</i>	<i>Low-Bspp</i>	<i>Low-Bspp, Nlab, High-Bspp</i>

Table 5: Anomaly data ODP 1172

For core MD11-3353 (Figure 3.4), the most dominant cluster is the *Sant* Cluster, which corresponds to a temperature range of -1.5 to 1.5 °C. Four anomalies are observed at 130, 125, 115 and 85 ka. The different oceanographic conditions deviate notably during these anomalies (Table 6).

MD11-3353	Dominant cluster values	130 ka	125 ka	115 ka	85 ka
Temperature (°C)	-1.5 - 1.5	12.1 - 18.2	12.1 - 18.2	16.2 - 23.4	16.2 - 23.4
Salinity (psu)	33.8 - 34.0	33.9 - 35.3	33.9 - 35.3	35.0 - 35.7	35.0 - 35.7
Nitrate (mmol/m ³)	25.3 - 29.1	1.8 - 14.5	1.8 - 14.5	0.1 - 3.7	0.1 - 3.7
Silicate (mmol/m ³)	50.5 - 65.1	1.1 - 6.2	1.1 - 6.2	0.5 - 1.3	0.5 - 1.3
Sea ice (%)	0 - 100.0	0	0	0	0
Clusters abundant	<i>Sant</i>	<i>Low-Bspp</i>	<i>Low-Bspp</i>	<i>Low-Ocen</i>	<i>Low-Ocen</i>

Table 6: Anomaly data MD11-3353

The dominant cluster in core U1475 is the *Low-Ocen* Cluster, associated with a temperature range of 16.2 to 23.4 °C (Figure 3.5). Four prominent deviations occur between 4450 and 4000 ka (Table 7). Between approximately 3400 and 2900 ka, anomalies are abundant which are categorized into two groups based on age.

U1475	Dominant cluster values	4450 ka	4400 ka	4100 - 4000 ka	3400 - 3350 ka	3300 - 2900 ka
Temperature (°C)	16.2 - 23.4	8.7 - 12.8	8.7 - 12.8	12.1 - 23.4	8.7 - 23.4	8.7 - 25.7
Salinity (psu)	35.0 - 35.7	34.1 - 34.7	34.1 - 34.7	33.9 - 35.7	34.1 - 35.7	33.9 - 36.4
Nitrate (mmol/m ³)	0.1 - 3.7	9.9 - 17.8	9.9 - 17.8	0.1 - 14.5	0.1 - 17.8	0.1 - 17.8
Silicate (mmol/m ³)	0.5 - 1.3	1.8 - 18.6	1.8 - 18.6	0.5 - 6.2	0.5 - 18.6	0.4 - 18.6
Sea ice (%)	0	0	0	0	0	0
Clusters abundant	<i>Low-Ocen</i>	<i>Nlab</i>	<i>Nlab</i>	<i>Low-Ocen, Low-Bspp</i>	<i>Low-Ocen, Low-Bspp, Nlab</i>	<i>Low-Ocen, Low-Bspp, Iacu, Nlab, High-Ocen</i>

Table 7: Anomaly data U1475

For core AS05-10, the dominant cluster is the *Low-Ocen* Cluster (Figure 3.6), corresponding to a temperature range of 16.2 to 23.4 °C. Several notable anomalies are present, at 240, 210, from 170 - 120 and at 90 ka (Table 8).

AS05-10	Dominant cluster values	240 ka	210 ka	170 - 120 ka	90 ka
Temperature (°C)	16.2 - 23.4	-1.5 - 1.5	-1.5 - 1.5	-1.5 - 23.4	-1.5 - 1.5
Salinity (psu)	35.0 - 35.7	33.8 - 34.0	33.8 - 34.0	33.8 - 35.7	33.8 - 34.0
Nitrate (mmol/m ³)	0.1 - 3.7	25.3 - 29.1	25.3 - 29.1	0.1 - 29.1	25.3 - 29.1
Silicate (mmol/m ³)	0.5 - 1.3	50.5 - 65.1	50.5 - 65.1	0.5 - 65.1	50.5 - 65.1
Sea ice (%)	0	0 - 100.0	0 - 100.0	0 - 100.0	0 - 100.0
Clusters abundant	<i>Low-Ocen</i>	<i>Sant</i>	<i>Sant</i>	<i>Low-Ocen, Sant</i>	<i>Iacu</i>

Table 8: Anomaly data AS05-10

5.1.2 Cluster temperature compared to proxy records

For ODP 1168, the $U^{K_{37}}$ provides a limited coverage of the older section of the core (Figure 4.1), limiting the comparison between the two temperature reconstructions. From approximately 4200 ka onwards, the $U^{K_{37}}$ reconstruction shows higher resolution, peaking around 3300 ka. An anomaly observed in the cluster temperature at 3600 ka most likely corresponds to a deviation in $U^{K_{37}}$ at approximately 3700 ka. From 3500 ka onward, both temperature reconstructions exhibit similar trends.

For ODP 1171, the resolution of the TEX_{86} records is sufficient for meaningful comparison between the two temperature reconstructions (Figure 4.2). The records generally display the same trend and remain within a similar temperature range. Occasionally, the TEX_{86} temperature deviates from the cluster range, but only for limited periods. The temperature anomaly at approximately 250 ka corresponds well with the peak in $d^{18}O$ record.

The proxy record of ODP 1172 is based on the Mg/Ca ratio, unlike TEX_{86} , which is predominantly used in this study. The proxy record spans the entire dinocyst record, although its resolution varies throughout (Figure 4.3). In the bottom section, the proxy data is slightly lower, by a few degrees Celsius, than the cluster temperature reconstruction. Higher up, both temperature reconstructions fall within the same temperature range.

The TEX_{86} proxy record for MD11-3353 covers the entire duration of the cluster temperature data (Figure 4.4). The anomaly occurring at approximately 130 ka is detected in both temperature reconstructions, but subsequent anomalies are only reflected in the cluster temperature. When the cluster temperature remains constant, the TEX_{86} temperature shows minimal variation, typically being slightly higher by a few degrees Celsius. The anomaly at 130 ka is consistent with a peak in the $d^{18}O$ record.

The $U^{K_{37}}$ record of the U1475 core does not cover the full length of the drill core, making comparisons for the bottom sections not possible (Figure 4.5). Further up in the core, the resolution of SST datapoints increases strongly. The proxy record does not reflect the large deviations observed in the cluster temperature but shows somewhat similar trends around 4050 ka.

The cluster temperature record of core AS05-10 exhibits substantial fluctuations in temperatures throughout the represented time period (Figure 4.6). The proxy record for this core has limited datapoints, resulting in an insufficient proxy temperature reconstruction. As a result, no meaningful comparison between the two temperature records can be made.

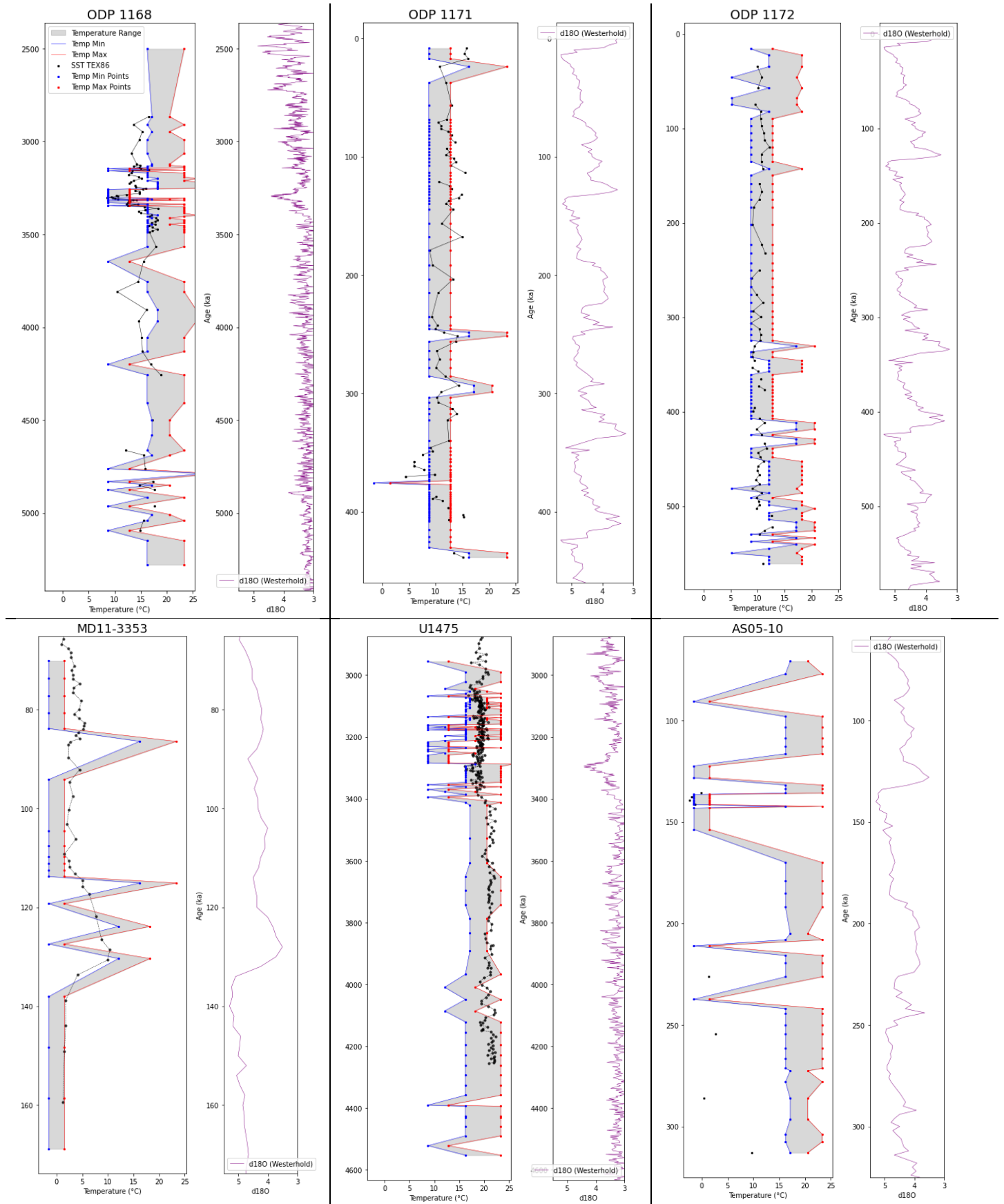


Figure 4: Temperature reconstructions based on *k*-means clustering method

5.2 Individual dinocyst species analysis

In this analysis, the temperature and relative abundance of dinocyst species from several datasets are compared to modern conditions (Figure 5). Species dominating the modern clusters are analysed to assess their past occurrences. The graph with all species abundant in the research can be found in Appendix J.

Nucicla umbiliphora (*Numb*) is a dinocyst species that is abundant around 0 °C in PD, with one outlier at 15 °C. This species is present in two drill cores: ODP 1172 and AS05-10. In AS05-10, temperatures align well, while in ODP 1172 the sole datapoint is an outlier at 9 °C.

Dinocyst species *Islandinium minutum* (*Imin*) is mostly abundant around 0 °C in the modern with relative abundances leading up to 1.0, but also has multiple occurrences at approximately 15 °C in lower abundance. In core ODP 1172, this species occurs around 10 °C with a maximum relative abundance of roughly 0.1. Core U1475 contains temperatures of 20 °C.

Spiniferites spp. (*Sspp*) is represented in more downcore datasets: ODP 1168, ODP 1171, ODP 1172 and U1475. In the modern this species is present from approximately 12 °C and higher (up to 30 °C), with the relative abundance increasing with temperature. The temperature of the downcore data roughly aligns with the modern, only the high end of the modern temperatures is not covered by the downcore data. The temperatures of ODP 1172 are somewhat on the lower end of the modern temperature range.

The *Sram_bul* species (*Spiniferites ramosus* and *Spiniferites bulloideus*) in the present is abundant for almost all temperatures, the relative abundance however does increase strongly with temperature. The species is present in four downcore datasets: ODP 1168, ODP 1171, ODP 1172 and U1475. These datasets align greatly with PD, only with slightly lower relative abundances. The paleo data however does not cover the warmer end of the modern record, from roughly 23 °C on.

Quinquecuspis concreta (*Ocen*) shows temperatures in the modern from 0 to 30 °C with the lowest abundances around 5 °C. Looking at the cores, it is abundant in ODP 1168, ODP 1171, ODP 1172, MD11-3353 and U1475. These cores are spread out across the entire modern temperature range, except above 23 °C, with corresponding relative abundances.

Nematosphaeropsis labyrinthus (*Nlab*) is showing a modern temperature range from 0 to 30 °C, with high relative abundances between roughly 3 and 19 °C. This species is abundant in almost all cores: ODP 1168, ODP 1171, ODP 1172, MD11-3353 and U1475. The relative abundances and temperatures align greatly between the past and the present data, only the warm end is not covered by the drilling cores. The species is mostly abundant in core ODP 1171.

The *Selenopemphix antarctica* (*Sant_all*) species is mostly abundant with colder temperatures, past and present, favourably -1 °C. The relative abundance decreases with increasing temperatures. The species is present in all downcore datasets and they align greatly with the modern data. Its abundance in U1475 is notable since this is on the warmer side of the temperature spectrum, with 20 °C. The relative abundances are however quite low, with a maximum of approximately 0.07.

The graph for *Brigantidium* spp. (*Bspp*) shows that in the modern this species is abundant in all temperatures with high abundances. At 0 and 30 °C it shows modern relative abundances all the way to 1.0. *Bspp* is abundant in all downcore dataset, some with higher relative abundances than the other, but definitely not as high as the modern data. The higher modern temperatures are not covered by the drilling cores, which show maximum temperatures up to approximately 23 °C.

Impagidinium aculeatum (*Iacu*) has a modern temperature range from -1 to 30 °C, the relative abundances are highest between 15 and 25 °C. The dinocyst species is abundant in all drill cores. The past and present data show a great fit, with the past having slightly lower relative abundances. AS0510

is showing some outliers when compared to the modern. The modern temperature is almost entirely covered by drill cores, except for 23 °C and up.

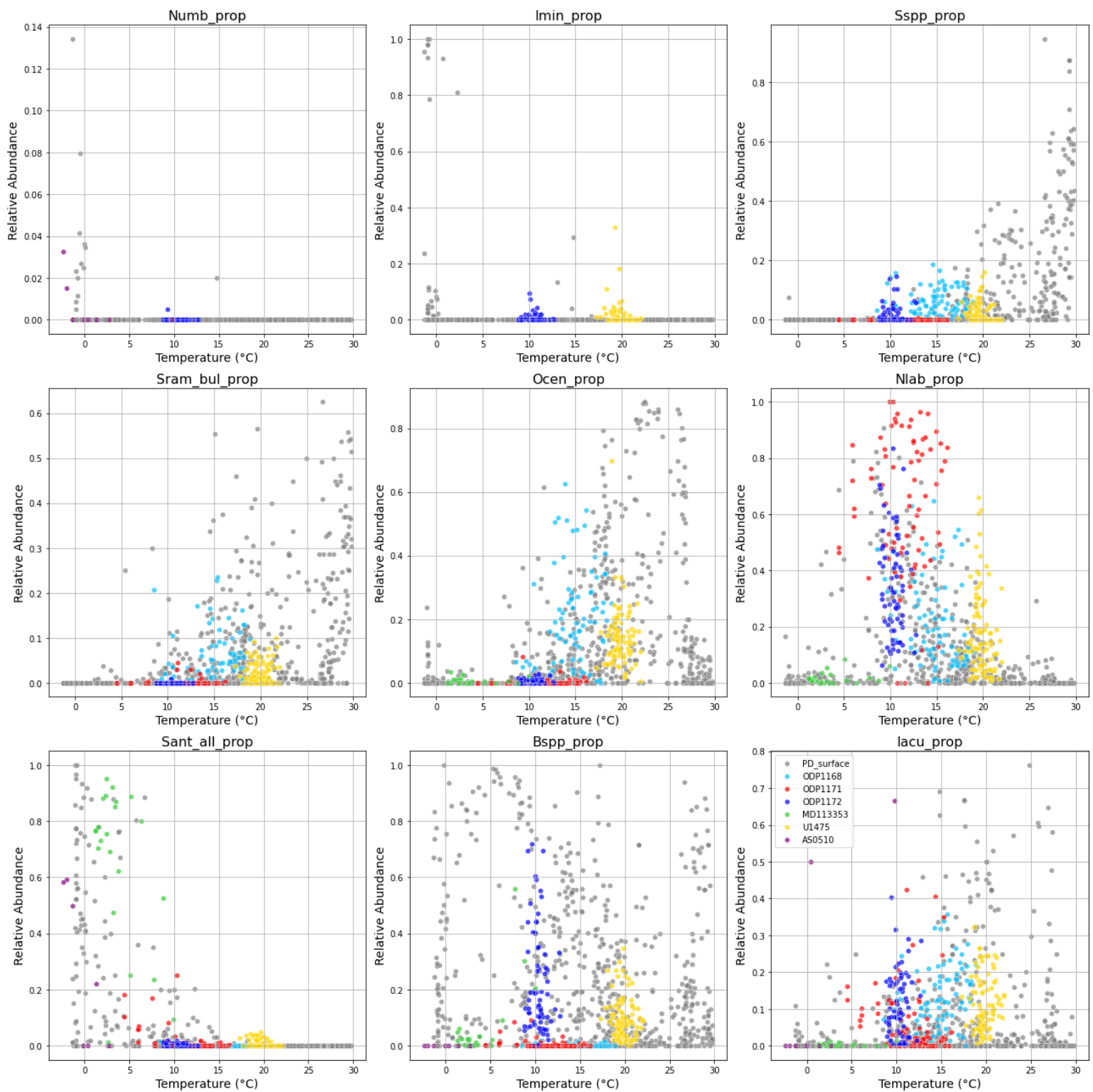


Figure 5: Relative abundance and temperature data of species dominant in the clustering analysis

6. Discussion

6.1 Temperature reconstruction comparison to proxy temperature

6.1.1 North of Subtropical front

Three cores are located north of the Subtropical Front (Figure 1): ODP 1168, ODP 1172 and U1475. Looking at the location of the cores and their dominant clusters (*Nlab*, Low-*Ocen* Clusters), the results align with the cluster distribution map of Thöle et al. (2023) (Figure 1). All three cores also show a fit between the reconstructed temperature and the proxy temperature.

For ODP 1168, the reconstructed temperature from dinocysts (16.2 - 23.4 °C dominantly) is slightly higher than modern STF temperatures around Tasmania (12 - 18 °C) (Ridgway & Ling, 2023), although the $U^{k_{37}}$ does align more closely with these temperatures. While the deviation between the reconstructed and modern temperature is not major, the Low-*Bspp* Cluster, with temperatures of 12.1 - 18.2 °C, would be more consistent with the modern data.

The low relative abundance of *Brigantedinium* spp. (*Bspp*) in this core (Figure 5) may explain why samples were not assigned to this cluster. The low relative abundance could be caused by preservation issues (Zonneveld et al., 2008) which leads to a lower abundance in the studied sediment compared to the actual in situ abundance.

U1475 has highly frequent fluctuations in the top part of the studied section that are driven by the decrease in relative abundance of *Quinquecuspidis concreta* (*Ocen*), a species indicative of warm, saline and nutrient poor waters (Santos et al., 2017). This decline allows other species, such as *Bspp* and *Impagidinium paradoxum* (*Ipar*), to become more prominent, altering the cluster assignments.

For *Spiniferites* spp. (*Sspp*), the reconstructed temperatures are somewhat lower than the modern temperature spectrum, with higher temperatures notably absent. This discrepancy is influenced by cores ODP 1171 and ODP 1172. An issue with *Sspp* could be that it functions as an umbrella species encompassing multiple *Spiniferites* species (Ellegaard, 2000). Some *Spiniferites* species may have been grouped differently in past and present datasets, potentially affecting the temperature affinity of *Sspp*. While *Spiniferites* species generally respond similarly to local environmental triggers (Pospelova et al., 2018), this minor deviation may attribute to inconsistencies.

6.1.2 Between Subtropical front and Subantarctic front

One core is located between Subtropical and Subantarctic front: ODP 1171. The temperature reconstruction of ODP 1171 aligns exceptionally well with TEX_{86} data in both absolute values and trends. Around 150 ka, TEX_{86} temperatures slightly exceed cluster-based temperatures, but the discrepancy is minor (a few degrees Celsius). This core exhibits the best fit among all cores (Figure 3), with temperatures consistent with modern Australian SSTs, ranging from 8 °C minimum in winter to an average of 16 °C in summer (Sikes et al., 2009). The TEX_{86} is showing more sample-to-sample variability than the reconstructed temperature, however this TEX_{86} variability is minor. This natural biological noise is not visible in the cluster reconstructed temperature.

6.1.3 South of Polar Front

Two cores are located south of the Polar Front: MD11-3353 and AS05-10 (Figure 1). Both cores show discrepancies with the modern temperatures recorded in this region. The MD11-3353 cluster temperature record is mostly stable, ranging between -1.5 and 1.5 °C, consistent with low PF temperatures (Barnes et al., 2006). However, significant anomalies occur due to cluster transitions, with temperatures spiking up to 23.4 °C. These fluctuations are extreme, but the TEX_{86} data does show a fluctuation to a lesser extent once for the researched time period which makes the reconstructed temperature more likely.

Sant maintains a high relative abundance throughout most of the core, except during anomalies. The sudden decrease in *Sant* results in the relative increase of other species such as *Impagidinium pallidum* (*Ipal*) and *Brigantedinium* spp. (*Bspp*), leading to different cluster assignments. This core contains only 12 species, strikingly lower than the modern data ($n = 64$) but not uncommon in the modern cold surface sediment samples, making it more vulnerable to misclassification in the clustering analysis when the population is not completely represented in the dinocyst counts. Sites close to Antarctica have low diversity. Important species that could dominate clusters are for instance not abundant, which could lead to assigning the wrong cluster to the sample.

Occasional spikes in relative abundance of *Impagidinium pallidum* (*Ipal*) (approximately 0.4 - 0.85) correspond to decreases in *Sant*. While *Ipal* is generally not a dominant species, its presence may have contributed to misclassification. However, *Ipal* might not be a reliable indicator of colder waters in older deposits since much warmer affinities are studied in paleo records (De Schepper et al., 2011).

MD11-3353 is close to a sector of the Southern ocean near the Polar Front, where the spatial coverage of surface sediment samples is extremely low. As a result, the cluster in Thöle et al. (2023) do not represent this ocean zone very well, and this might be the reason for the mismatch. What could possibly have a positive effect for the clustering analysis of this core is the increase in the amount of clusters (k), rather than the $k = 9$ of Thöle et al. (2023). This could cause a better clustering determination for this region. Misclassification affects the salinity, sea-ice presence, nitrate and silicate just like it does for temperature (Figure 3).

For core AS05-10, no meaningful comparison can be made between the two temperature reconstructions due to the sparse TEX_{86} data ($n = 10$). Notably, TEX_{86} data indicate temperatures below 5 °C, while dominant cluster temperatures range from 16.2 to 23.4 °C. This substantial discrepancy suggests that one of the temperature reconstructions is inaccurate for this core.

Given the core's location in the Ross Sea (Figure 1), the cluster temperature reconstruction is does not seem accurate. Smethie Jr and Jacobs (2005) indicate temperatures between -2.2 and 0 °C in the Ross Sea in the modern. Potential explanations for the cluster analysis temperature deviation include: (1) the remote location may be too specific for this cluster method, or (2) insufficient species diversity ($n = 9$) for robust clustering.

AS05-10 contains only 9 dinocyst species (Hartman et al., 2021), compared to 64 species in the PD dataset. High abundances of species like *Iacu* caused samples to be assigned to Cluster 5 (*Iacu* Cluster), which represents high temperatures. This cluster contains four species in this core. If more species were present, a better cluster fit could have been achieved, potentially leading to more accurate cluster assignments. However, this is not the case because of the restrictive character of the harsh circumstances and sea-ice in the area (Warny et al., 2006).

Examining individual species, it is remarkable that *Brigantedinium* spp. (*Bspp*) is present in only low relative abundances in AS05-10, even though it is highly abundant at similar temperatures in PD (Figure 5). If *Bspp* were more abundant, the samples might have been assigned to a different cluster with lower temperatures. In modern environments, *Bspp* exhibits a broad temperature range and high relative abundances throughout this temperature spectrum. However, in the colder cores, AS05-10 and MD11-3353, its relative abundance is remarkably lower than in the modern data. This may indicate that (1) the species was historically less abundant in colder areas, (2) another species dominated colder regions in the past, reducing the relative abundance of *Bspp*, or (3) other environmental conditions were less favourable. For example, *Bspp* is often associated with low salinity (Solignac et al., 2004). Additional cold-core samples could validate these observations and provide insights into this temperature range.

6.2 Other oceanographic features

A key consideration regarding the cluster analysis is that oceanographic conditions derived from this method are interdependent. Consequently, anomalies caused by changes in clusters are reflected across various oceanographic conditions. For example: if there is a temperature anomaly, there will automatically be a salinity anomaly as well. Since the clustered temperature reconstruction of AS05-10 is inconvenient, the values of other oceanographic features will be off as well. This also accounts for the anomalies of MD11-3353.

6.2.1 Salinity

For cores ODP 1169, ODP 1171, ODP 1172, U1475 and MD11-3353, the current sea surface salinity (Garcia-Eidell et al., 2019) align well with the dominant reconstructed salinity. The modern salinity for these locations is approximately 35.0 psu (MD11-3353 34.5 psu), while the reconstructed salinities are 34.1 - 34.7 psu for ODP 1171 and ODP 1172, 35.0 - 35.7 psu for cores ODP 1168 and U1475 and 33.8 - 34.0 psu for MD11-3353. Since there is no proxy for salinity, the anomalies cannot be compared.

The salinities however deviates tremendously for core AS05-10 with a modern salinity of approximately 34.0 psu (Garcia-Eidell et al., 2019) and the reconstructed salinity being 35.0 - 35.7 psu. This means there misclassification occurred on a large scale for this core.

6.2.2 Sea ice presence

The sea-ice reconstructions of ODP 1168, ODP 1172 and U1475 show no signs of sea-ice presence, and ODP 1171 only once and rather shortly, for the corresponding period which is consistent with other sea-ice reconstructions (McKay et al., 2012).

Looking at the location of MD11-3353 and AS05-10, southward of the Polar Front, it is logical that more sea-ice is present in the reconstruction. The MD11-3353 reconstruction is showing the possibility of sea-ice throughout most of the core, except for some small deviations, then no sea-ice would be present at all. The AS05-10 core is showing a few periods where sea ice could have been present, but dominantly it would have been a sea-ice free location, when looking at the reconstruction. The core is located at the edge of the Ross Ice Shelf, where sea ice most primarily abundant (Jacobs et al., 1970).

6.2.3 Nutrients

For the nutrients reconstructions, we focus on silicate and nitrate concentrations. ODP 1168, U1475 and AS05-10 have the same dominant reconstructed silicate (0.5 - 1.3 mmol/m³) and nitrate (0.1 - 3.7 mmol/m³) concentrations. These concentrations are logical for ODP 1168 and U1475 and align greatly with the modern concentrations for their locations (Freeman et al., 2018). The concentrations for AS05-10 are not logical for the Ross Sea at all, with modern concentrations of silicate (approximately 70 mmol/m³) and nitrate (approximately 25 mmol/m³) being extremely high. This reconstruction is thus incorrect.

ODP 1171 and ODP 1172 have a rather large range for silicate (1.8 - 18.6 mmol/m³) and nitrate (9.9 - 17.8 mmol/m³) concentrations. These concentrations are somewhat on the higher side of the modern concentrations (Freeman et al., 2018), but the difference between the reconstructions and the modern concentrations are however minor. The reconstructed silicate (50.5 - 65.1 mmol/m³) concentrations of MD11-3353 are higher than the modern concentrations (approximately 30 mmol/m³) (Freeman et al., 2018), the nitrate concentrations however do show a great fit.

6.3 Broader implications

When analysing the downcore data, it becomes evident that cores containing a greater diversity of dinocyst species align more closely with the proxy data. MD11-3353 (n = 12) and AS05-10 (n = 9) both have a low number of species and they are also the datasets with the least amount of samples:

MD11-3353 contains of 21 samples, AS05-10 has 48 samples, compared to 88 (ODP 1168), 90 (ODP 1172), 101 (ODP 1171) and 113 (U1475) samples. It appears that datasets with fewer species and samples are more vulnerable to inaccuracies in this clustering method, leading to misclassification.

The drill cores with fewest dinocyst species are located south of the PF. Research in this region is challenging due to the harsh environmental conditions. However, it is essential to obtain more data, both modern and paleo, from this area to improve the understanding of dinocyst dynamics. Including Northern Hemisphere data could also enhance the comprehension, offering insights across the globe.

Another valuable addition to this research would be to revisit the number of clusters used in the analysis. Thöle et al. (2023) determined that 9 clusters were optimal for a stable clustering system, though their study did not include downcore data. If this clustering system remains stable with an increased number of clusters, the clusters could potentially become more specific to particular regions, leading to a more detailed representation of surface water characteristics.

7. Conclusion

In this study, we performed two complementary statistical analysis to determine the temperature affinities of dinocyst species through time. Based on these analyses, we draw the following conclusions:

The clustering analysis shows rather consistent results: all cores north of the Polar Front have a temperature reconstructions derived from the clustering method that align closely with those from other proxies, such as TEX₈₆. This demonstrates that the k-means clustering method is functioning effectively.

However, cores located south of the Polar Front, MD11-3353 and AS05-10, show a less great fit between the two temperature reconstruction. Both cores are located south of the Polar Front (PF) and contain a low number of dinocyst species and sedimentary samples. This limited species diversity makes the samples more vulnerable to inaccuracies in the clustering analysis, which could cause misclassification.

Enhancing the robustness of the clustering method requires the inclusion of additional data, both modern and paleo. While such data is needed across all latitudes, the PF represents a critical gap in the modern dat. Filling this gap would improve the reliability of the clustering process. For the individual species analysis, incorporating a drill core from a region close to the equator would ensure the complete coverage of the modern temperature range, which is currently lacking.

Regarding individual species, the low relative abundance of *Brigantedinium* spp. in downcore data is particularly striking, as it contrasts with its high relative abundance in the present. *Brigantedinium* spp. appears to be a key species in cluster determination, yet there is a large difference between its past and present relative abundance patterns. This could possibly be caused by preservation issues.

Overall, this study demonstrates that the clustering method is a viable approach interpreting first-order quantitative oceanographic conditions from dinocyst assemblages. However, the reliability of the method depends on the dataset containing a sufficiently large number of dinocyst species.

8. Bibliography

- Ai, X. E., Studer, A. S., Sigman, D. M., Martínez-García, A., Fripiat, F., Thöle, L. M., Michel, E., Gottschalk, J., Arnold, L., & Moretti, S. (2020). Southern Ocean upwelling, Earth's obliquity, and glacial-interglacial atmospheric CO₂ change. *Science*, 370(6522), 1348-1352.
- Barker, P., & Thomas, E. (2004). Origin, signature and palaeoclimatic influence of the Antarctic Circumpolar Current. *Earth-Science Reviews*, 66(1-2), 143-162.
- Barnes, D. K., Fuentes, V., Clarke, A., Schloss, I. R., & Wallace, M. I. (2006). Spatial and temporal variation in shallow seawater temperatures around Antarctica. *Deep Sea Research Part II: Topical Studies in Oceanography*, 53(8-10), 853-865.
- Böning, C. W., Dispert, A., Visbeck, M., Rintoul, S., & Schwarzkopf, F. U. (2008). The response of the Antarctic Circumpolar Current to recent climate change. *Nature Geoscience*, 1(12), 864-869.
- Bravo, I., & Figueroa, R. I. (2014). Towards an ecological understanding of dinoflagellate cyst functions. *Microorganisms*, 2(1), 11-32.
- Brinkhuis, H., Sengers, S., Sluijs, A., Warnaar, J., & Williams, G. L. (2003). Latest Cretaceous to earliest Oligocene, and Quaternary dinoflagellate cysts from ODP Site 1172, East Tasman Plateau. Proceedings of the ocean drilling program. Scientific results,
- Civel-Mazens, M., Crosta, X., Cortese, G., Michel, E., Mazaud, A., Ther, O., Ikehara, M., & Itaki, T. (2021). Impact of the Agulhas Return Current on the oceanography of the Kerguelen Plateau region, Southern Ocean, over the last 40 kyrs. *Quaternary Science Reviews*, 251, 106711.
- Crosta, X., Kohfeld, K. E., Bostock, H. C., Chadwick, M., Du Vivier, A., Esper, O., Etourneau, J., Jones, J., Leventer, A., & Müller, J. (2022). Antarctic sea ice over the past 130 000 years—part 1: a review of what proxy records tell us. *Climate of the Past*, 18(8), 1729-1756.
- De Schepper, S., Fischer, E. I., Groeneveld, J., Head, M. J., & Matthiessen, J. (2011). Deciphering the palaeoecology of Late Pliocene and Early Pleistocene dinoflagellate cysts. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 309(1-2), 17-32.
- de Vernal, A., Rochon, A., Fréchette, B., Henry, M., Radi, T., & Solignac, S. (2013). Reconstructing past sea ice cover of the Northern Hemisphere from dinocyst assemblages: status of the approach. *Quaternary Science Reviews*, 79, 122-134.
- Ellegaard, M. (2000). Variations in dinoflagellate cyst morphology under conditions of changing salinity during the last 2000 years in the Limfjord, Denmark. *Review of Palaeobotany and Palynology*, 109(1), 65-81.
- Esper, O., & Zonneveld, K. A. (2007). The potential of organic-walled dinoflagellate cysts for the reconstruction of past sea-surface conditions in the Southern Ocean. *Marine Micropaleontology*, 65(3-4), 185-212.
- Exon, N., Kennett, J., & Malone, M. (2001). 1. LEG 189 SUMMARY. Proceedings of the Ocean Drilling Program,
- Freeman, N. M., Lovenduski, N. S., Munro, D. R., Krumhardt, K. M., Lindsay, K., Long, M. C., & Maclennan, M. (2018). The variable and changing Southern Ocean silicate front: insights from the CESM Large Ensemble. *Global Biogeochemical Cycles*, 32(5), 752-768.
- Garcia-Eidell, C., Comiso, J. C., Dinnat, E., & Brucker, L. (2019). Sea surface salinity distribution in the Southern Ocean as observed from space. *Journal of Geophysical Research: Oceans*, 124(5), 3186-3205.
- Hall, I., Hemming, S., LeVay, L., Barker, S., Berke, M., Brentegani, L., Caley, T., Cartagena-Sierra, A., Charles, C., & Coenen, J. (2017). Site U1475. In *Proceedings of the International Ocean Discovery Program*, 361. International Ocean Discovery Program.
- Hansen, J., & Sato, M. (2020). Global warming acceleration. *Columbia University Reports*.
- Hartman, J. D., Sangiorgi, F., Barcena, M., Tateo, F., Giglio, F., Albertazzi, S., Trincardi, F., Bijl, P., Langone, L., & Asioli, A. (2021). Sea-ice, primary productivity and ocean temperatures at

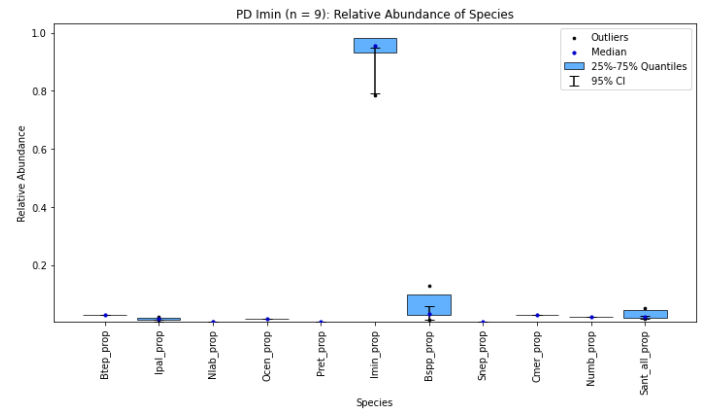
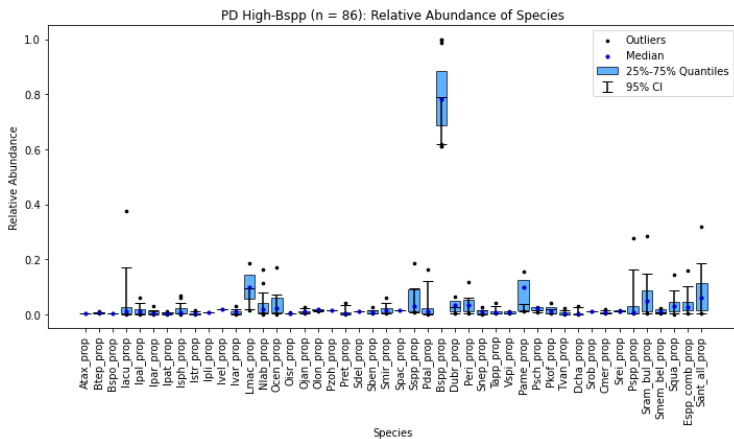
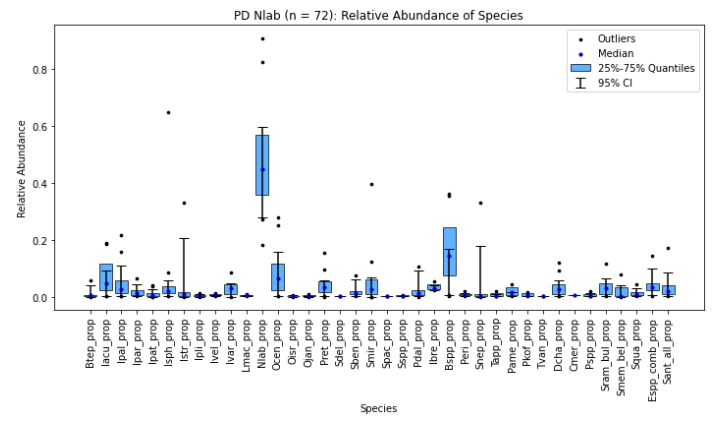
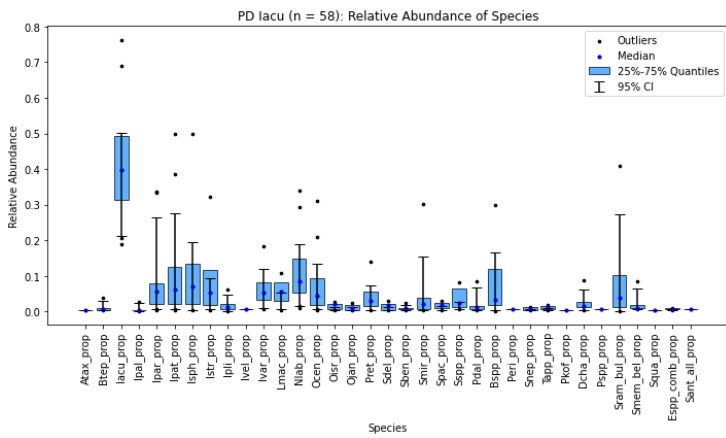
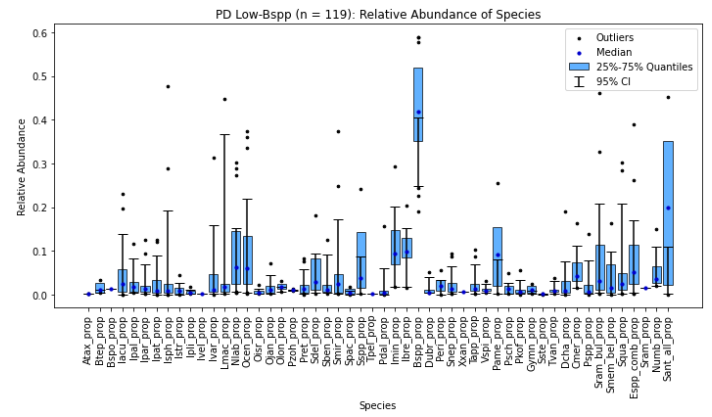
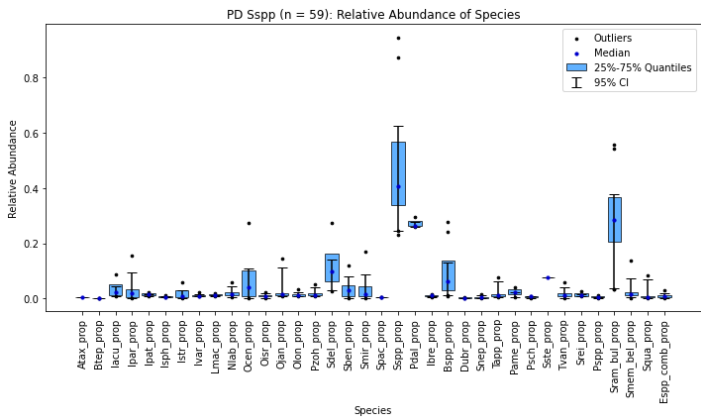
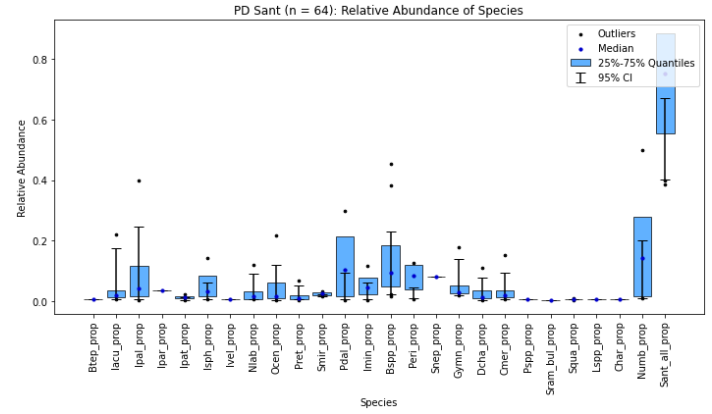
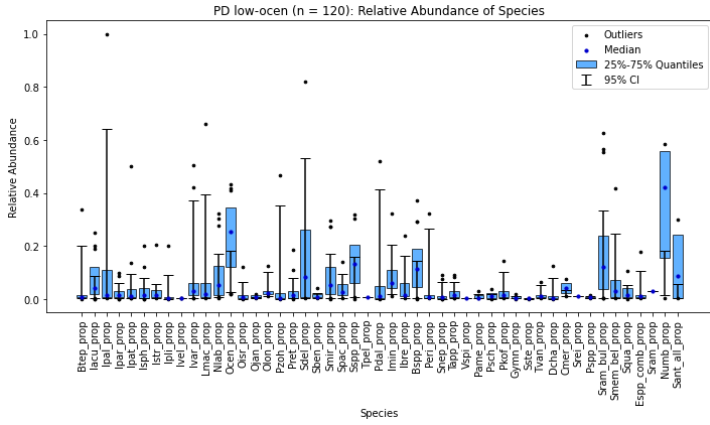
- the Antarctic marginal zone during late Pleistocene. *Quaternary Science Reviews*, 266, 107069.
- Hatté, C., & Schwartz, D. (2003). Reconstruction of paleoclimates by isotopic analysis: What can the fossil isotopic record tell us about the plant life of past environments? *Phytochemistry Reviews*, 2, 163-177.
- Hou, S. (2024). *Southern Ocean, Antarctic ice and climate interactions during Neogene cooling* [Utrecht University].
- Hou, S., Lamprou, F., Hoem, F. S., Hadju, M. R. N., Sangiorgi, F., Peterse, F., & Bijl, P. K. (2023). Lipid-biomarker-based sea surface temperature record offshore Tasmania over the last 23 million years. *Climate of the Past*, 19(4), 787-802.
- Hou, S., Stap, L. B., Paul, R., Nelissen, M., Hoem, F. S., Ziegler, M., Sluijs, A., Sangiorgi, F., & Bijl, P. K. (2023). Reconciling Southern Ocean fronts equatorward migration with minor Antarctic ice volume change during Miocene cooling. *Nature Communications*, 14(1), 7230.
- Hoyle, T. M., Sala-Pérez, M., & Sangiorgi, F. (2019). Where should we draw the lines between dinocyst “species”? Morphological continua in Black Sea dinocysts. *Journal of Micropalaeontology*, 38(1), 55-65.
- Jacobs, S. S., Amos, A. F., & Bruchhausen, P. M. (1970). Ross Sea oceanography and Antarctic bottom water formation. *Deep Sea Research and Oceanographic Abstracts*.
- Kennett, J. P. (1977). Cenozoic evolution of Antarctic glaciation, the circum-Antarctic Ocean, and their impact on global paleoceanography. *Journal of geophysical research*, 82(27), 3843-3860.
- Leutert, T. J. A., Alexandra; Martínez-García, Alfredo; Modestou, Sevasti; Meckler, Anna Nele (2020). *Temperatures from TEX86 measurements performed on middle Miocene sediment material from ODP Hole 189-1171C*.
<https://doi.org/https://doi.org/10.1594/PANGAEA.919351>
- Marret, F., Bradley, L., de Vernal, A., Hardy, W., Kim, S.-Y., Mudie, P., Penaud, A., Pospelova, V., Price, A. M., & Radi, T. (2020). From bi-polar to regional distribution of modern dinoflagellate cysts, an overview of their biogeography. *Marine Micropaleontology*, 159, 101753.
- Marret, F., de Vernal, A., Benderra, F., & Harland, R. (2001). Late Quaternary sea-surface conditions at DSDP Hole 594 in the southwest Pacific Ocean based on dinoflagellate cyst assemblages. *Journal of Quaternary Science: Published for the Quaternary Research Association*, 16(7), 739-751.
- Marret, F., & Zonneveld, K. A. (2003). Atlas of modern organic-walled dinoflagellate cyst distribution. *Review of Palaeobotany and Palynology*, 125(1-2), 1-200.
- Matsuoka, K., & Fukuyo, Y. (2000). Technical guide for modern dinoflagellate cyst study. *WESTPAC-HAB, Japan Society for the Promotion of Science, Tokyo, Japan*, 47.
- McElroy, M. B. (1983). Marine biological controls on atmospheric CO₂ and climate. *Nature*, 302(5906), 328-329.
- McKay, R., Naish, T., Carter, L., Riesselman, C., Dunbar, R., Sjunneskog, C., Winter, D., Sangiorgi, F., Warren, C., & Pagani, M. (2012). Antarctic and Southern Ocean influences on Late Pliocene global cooling. *Proceedings of the National Academy of Sciences*, 109(17), 6423-6428.
- Nooteboom, P. D., Bijl, P. K., van Sebille, E., Von Der Heydt, A. S., & Dijkstra, H. A. (2019). Transport bias by ocean currents in sedimentary microplankton assemblages: Implications for paleoceanographic reconstructions. *Paleoceanography and Paleoclimatology*, 34(7), 1178-1194.
- Nürnberg, D., & Groeneveld, J. (2006). Pleistocene variability of the subtropical convergence at East Tasman Plateau: evidence from planktonic foraminiferal Mg/Ca (ODP Site 1172A). *Geochemistry, Geophysics, Geosystems*, 7(4).

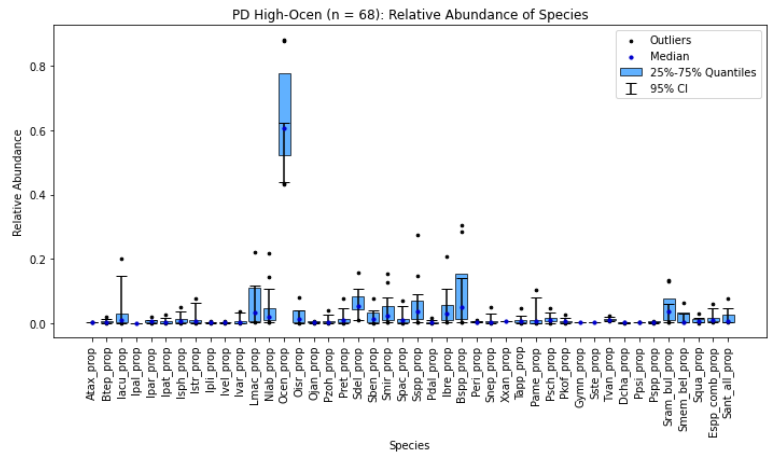
- Pospelova, V., de Vernal, A., & Pedersen, T. F. (2008). Distribution of dinoflagellate cysts in surface sediments from the northeastern Pacific Ocean (43–25 N) in relation to sea-surface temperature, salinity, productivity and coastal upwelling. *Marine Micropaleontology*, 68(1-2), 21-48.
- Pospelova, V., Zonneveld, K. A., Heikkilä, M., Bringué, M., Price, A. M., Esenkulova, S., & Matsuoka, K. (2018). Seasonal, annual, and inter-annual Spiniferites cyst production: a review of sediment trap studies. *Palynology*, 42(0), 162-181.
- Prebble, J., Crouch, E., Carter, L., Cortese, G., Bostock, H., & Neil, H. (2013). An expanded modern dinoflagellate cyst dataset for the Southwest Pacific and Southern Hemisphere with environmental associations. *Marine Micropaleontology*, 101, 33-48.
- Radi, T., & de Vernal, A. (2008). Dinocysts as proxy of primary productivity in mid–high latitudes of the Northern Hemisphere. *Marine Micropaleontology*, 68(1-2), 84-114.
- Ridgway, K., & Ling, S. (2023). Three decades of variability and warming of nearshore waters around Tasmania. *Progress in Oceanography*, 215, 103046.
- Rintoul, S. R. (2011). The southern ocean in the earth system. *Science diplomacy: Antarctica, science, and the governance of international spaces*.
- Santos, A., de Araujo Carvalho, M., de Oliveira, A. D., & Mendonça Filho, J. G. (2017). Paleoenvironmental changes and influence on Operculodinium centrocarpum during the Quaternary in the Campos Basin, southwestern Brazil. *Journal of South American Earth Sciences*, 80, 255-271.
- Sikes, E., Howard, W., Samson, C., Mahan, T., Robertson, L., & Volkman, J. (2009). Southern Ocean seasonal temperature and Subtropical Front movement on the South Tasman Rise in the late Quaternary. *Paleoceanography*, 24(2).
- Sinaga, K. P., & Yang, M.-S. (2020). Unsupervised K-means clustering algorithm. *IEEE access*, 8, 80716-80727.
- Smethie Jr, W. M., & Jacobs, S. S. (2005). Circulation and melting under the Ross Ice Shelf: estimates from evolving CFC, salinity and temperature fields in the Ross Sea. *Deep Sea Research Part I: Oceanographic Research Papers*, 52(6), 959-978.
- Sokolov, S., & Rintoul, S. R. (2002). Structure of Southern Ocean fronts at 140 E. *Journal of Marine Systems*, 37(1-3), 151-184.
- Solignac, S., de Vernal, A., & Hillaire-Marcel, C. (2004). Holocene sea-surface conditions in the North Atlantic—contrasted trends and regimes in the western and eastern sectors (Labrador Sea vs. Iceland Basin). *Quaternary Science Reviews*, 23(3-4), 319-334.
- Spector, D. L. (2012). *Dinoflagellates*. Academic Press.
- Stickley, C., Brinkhuis, H., McGonigal, K., Chaproniere, G., Fuller, M., Kelly, D., Nürnberg, D., Pfuhl, H., Schellenberg, S., & Schönfeld, J. (2004). Late Cretaceous–Quaternary biomagnetostratigraphy of ODP Sites 1168, 1170, 1171, and 1172, Tasmanian Gateway. *Proceedings of the Ocean Drilling Program, Scientific Results*,
- Taylor, F. (1980). On dinoflagellate evolution. *BioSystems*, 13(1-2), 65-108.
- Taylor, F., Hoppenrath, M., & Saldarriaga, J. F. (2008). Dinoflagellate diversity and distribution. *Biodiversity and conservation*, 17, 407-418.
- Thöle, L. M., Amsler, H. E., Moretti, S., Auderset, A., Gilgannon, J., Lippold, J., Vogel, H., Crosta, X., Mazaud, A., & Michel, E. (2019). Glacial-interglacial dust and export production records from the Southern Indian Ocean. *Earth and Planetary Science Letters*, 525, 115716.
- Thöle, L. M., Nooteboom, P. D., Hou, S., Wang, R., Nie, S., Michel, E., Sauermilch, I., Marret, F., Sangiorgi, F., & Bijl, P. K. (2023). An expanded database of Southern Hemisphere surface sediment dinoflagellate cyst assemblages and their oceanographic affinities. *Journal of Micropalaeontology*, 42(1), 35-56.
- Veth, C., Peeken, I., & Scharek, R. (1997). Physical anatomy of fronts and surface waters in the ACC near the 6 W meridian during austral spring 1992. *Deep Sea Research Part II: Topical Studies in Oceanography*, 44(1-2), 23-49.

- Wang, P., Yang, S., Li, Z., Song, Z., Li, X., & Hu, X. (2024). Role of the Antarctic Circumpolar Circulation in current asymmetric Arctic and Antarctic warming. *Geophysical Research Letters*, 51(13), e2024GL110265.
- Warny, S., Wrenn, J. H., Bart, P. J., & Askin, R. (2006). Palynology of the NBP03–01A transect in the Northern Basin, western Ross Sea, Antarctica: A late Pliocene record. *Palynology*, 30(1), 151-182.
- Wefer, G., Berger, W., Bijma, J., & Fischer, G. (1999). Clues to ocean history: a brief overview of proxies. *Use of proxies in paleoceanography: examples from the South Atlantic*, 1-68.
- Zonneveld, K. A., Marret, F., Versteegh, G. J., Bogus, K., Bonnet, S., Bouimetarhan, I., Crouch, E., de Vernal, A., Elshanawany, R., & Edwards, L. (2013). Atlas of modern dinoflagellate cyst distribution based on 2405 data points. *Review of Palaeobotany and Palynology*, 191, 1-197.
- Zonneveld, K. A., Versteegh, G., & Kodrans-Nsiah, M. (2008). Preservation and organic chemistry of Late Cenozoic organic-walled dinoflagellate cysts: a review. *Marine Micropaleontology*, 68(1-2), 179-197.

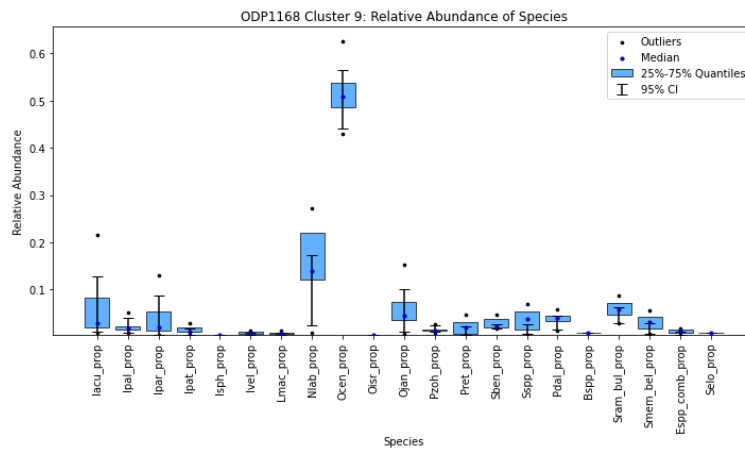
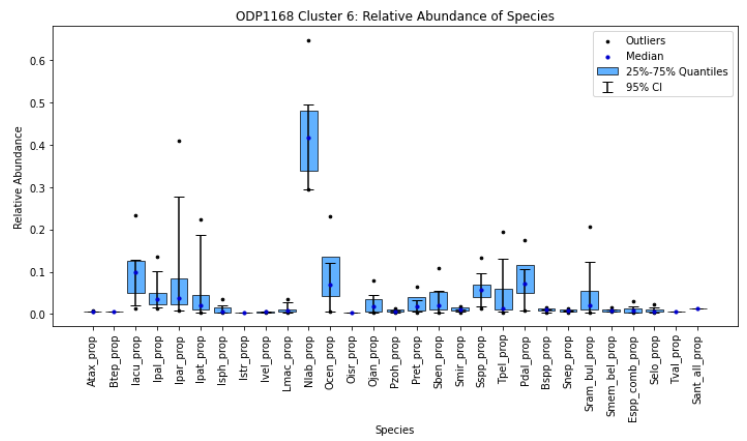
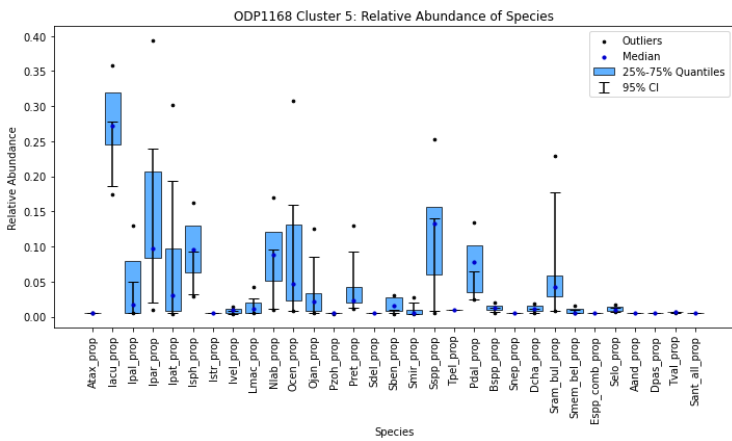
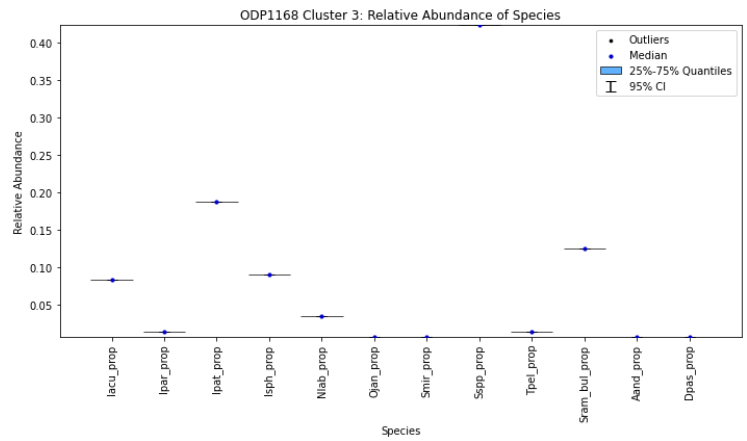
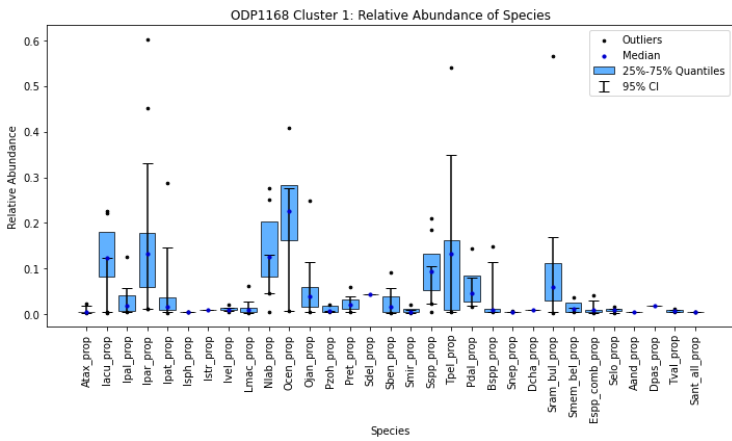
9. Appendix

Appendix A: Boxplots PD



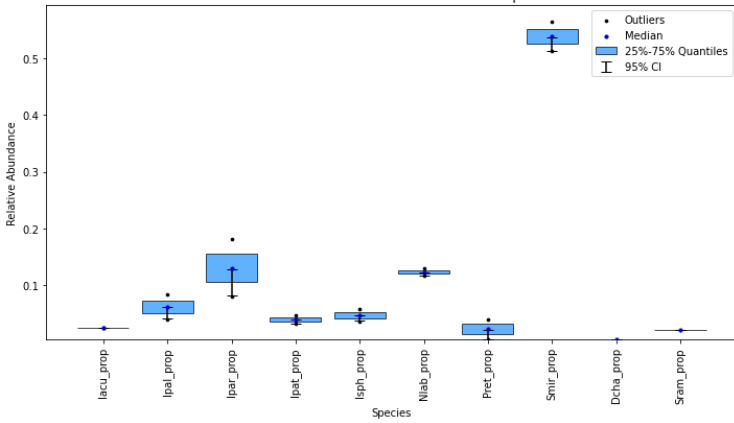


Appendix B: Boxplots ODP 1168

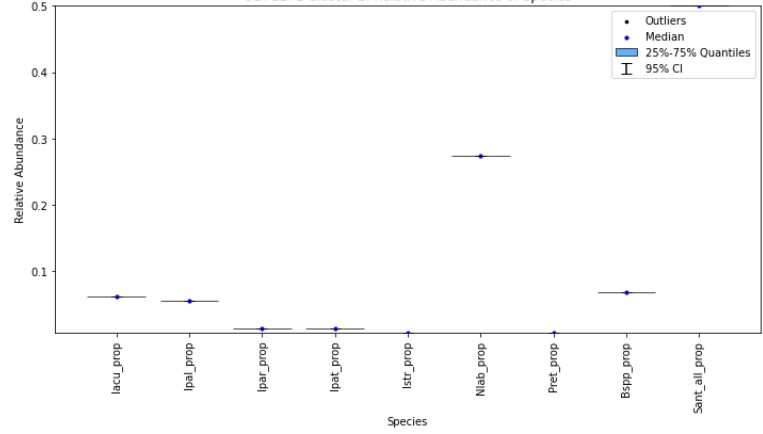


Appendix C: Boxplots ODP 1171

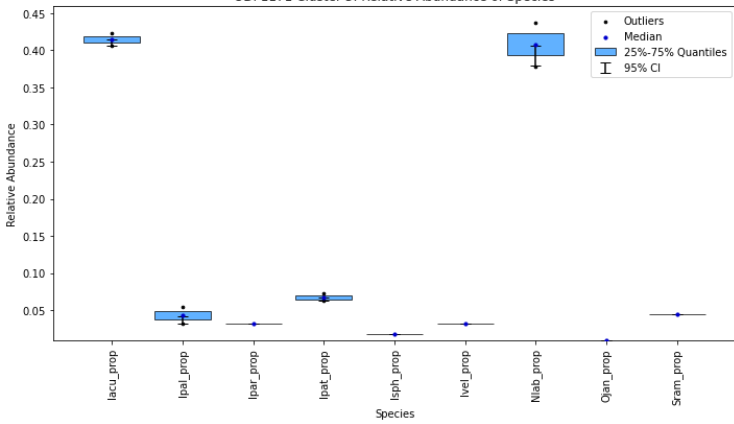
ODP1171 Cluster 1: Relative Abundance of Species



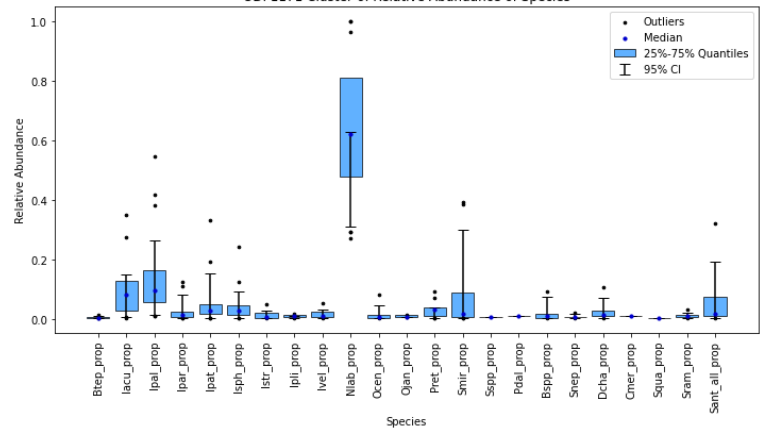
ODP1171 Cluster 2: Relative Abundance of Species



ODP1171 Cluster 5: Relative Abundance of Species

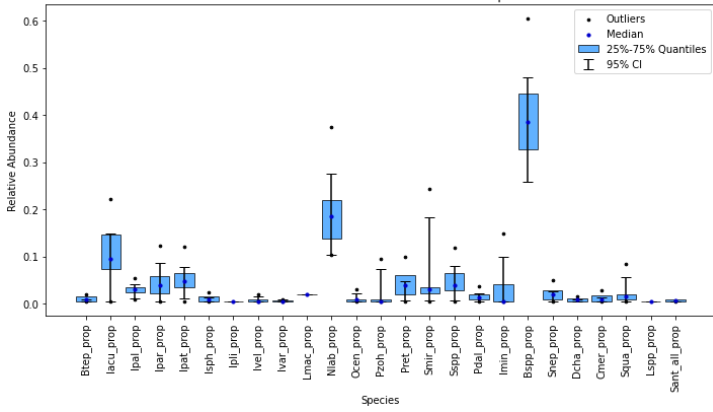


ODP1171 Cluster 6: Relative Abundance of Species

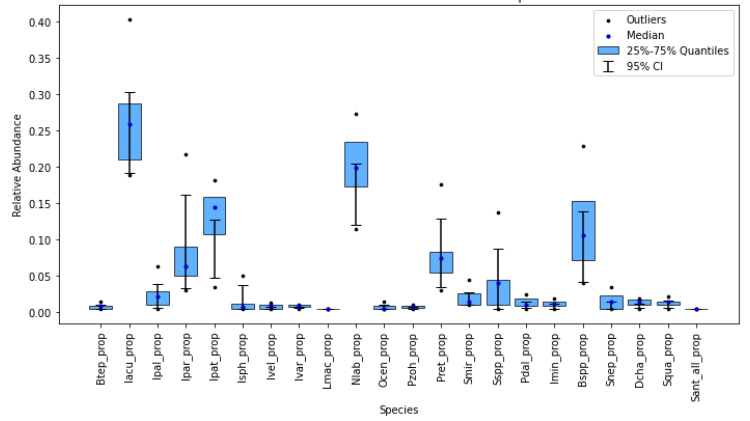


Appendix D: Boxplots ODP 1172

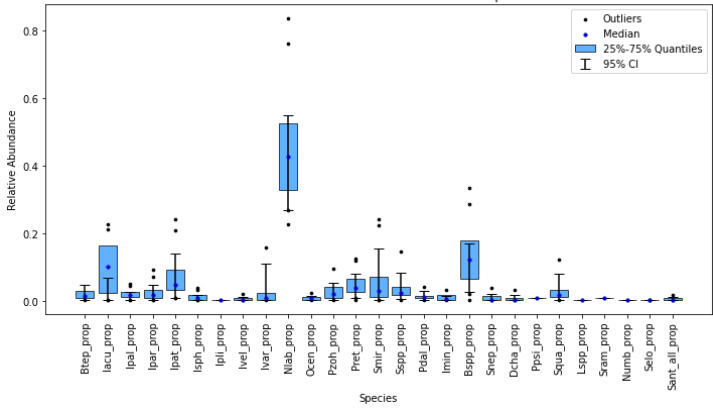
ODP1172 Cluster 4: Relative Abundance of Species



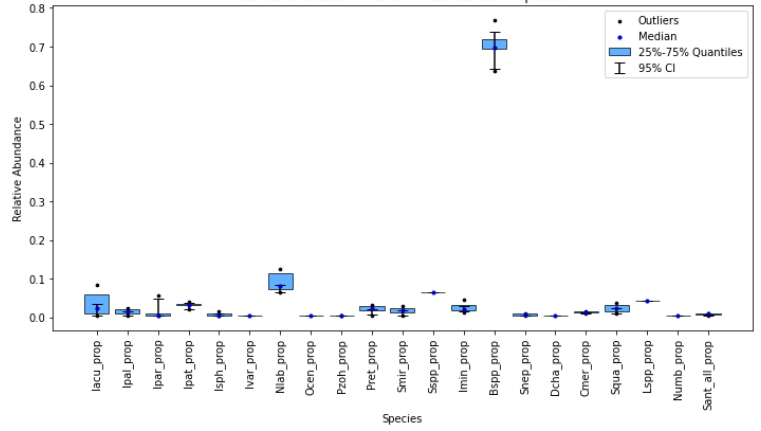
ODP1172 Cluster 5: Relative Abundance of Species



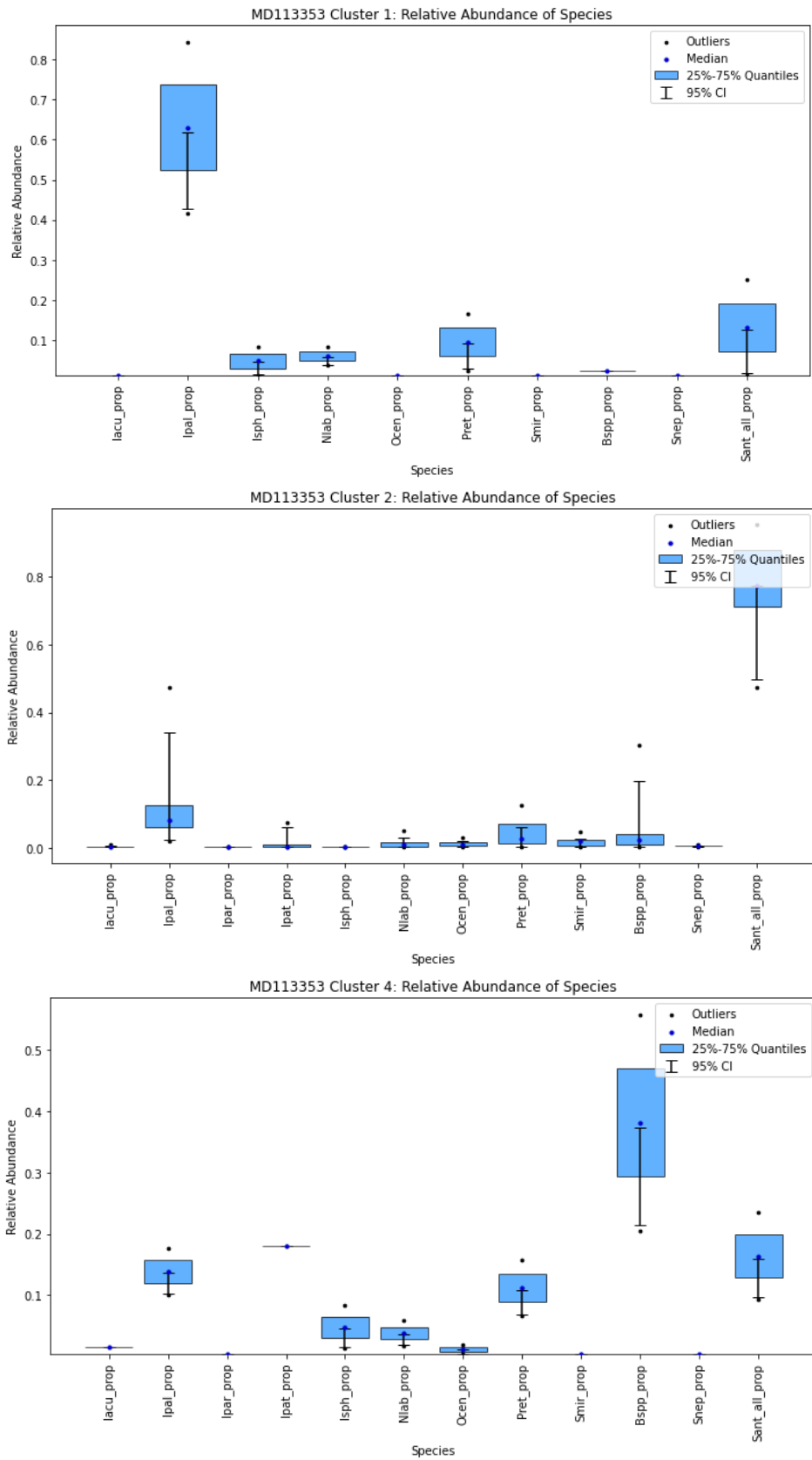
ODP1172 Cluster 6: Relative Abundance of Species



ODP1172 Cluster 7: Relative Abundance of Species

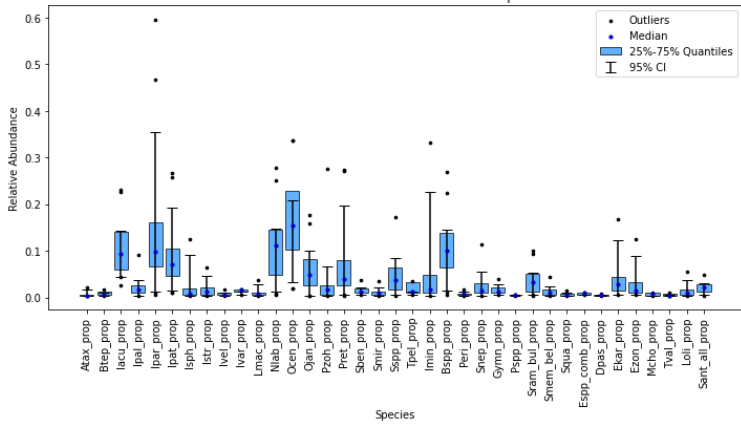


Appendix E: Boxplots MD11-3353

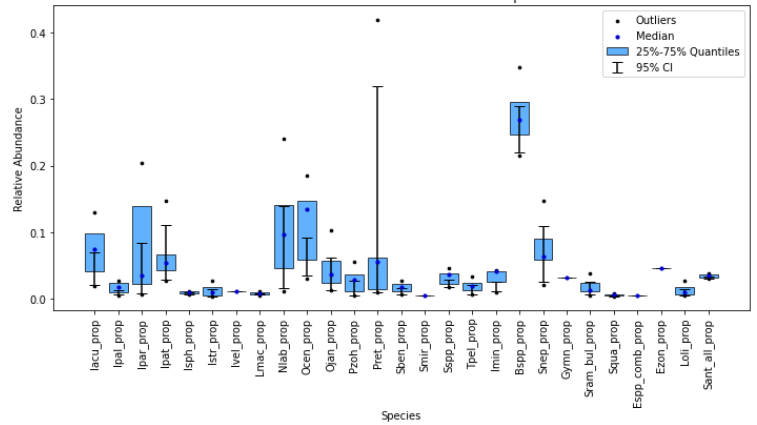


Appendix F: Boxplots U1475

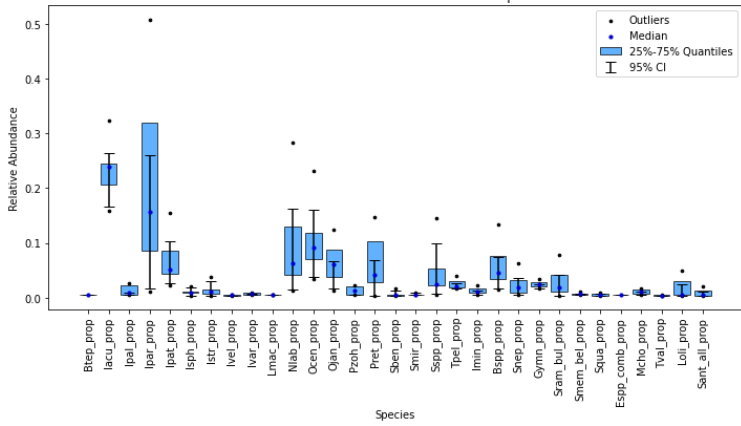
U1475 Cluster 1: Relative Abundance of Species



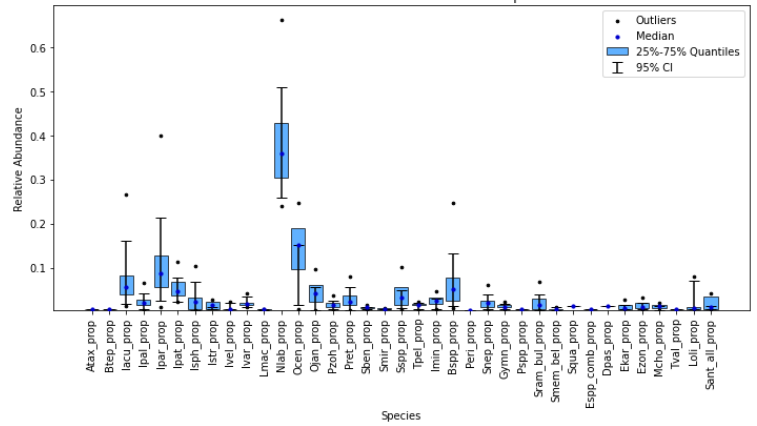
U1475 Cluster 4: Relative Abundance of Species



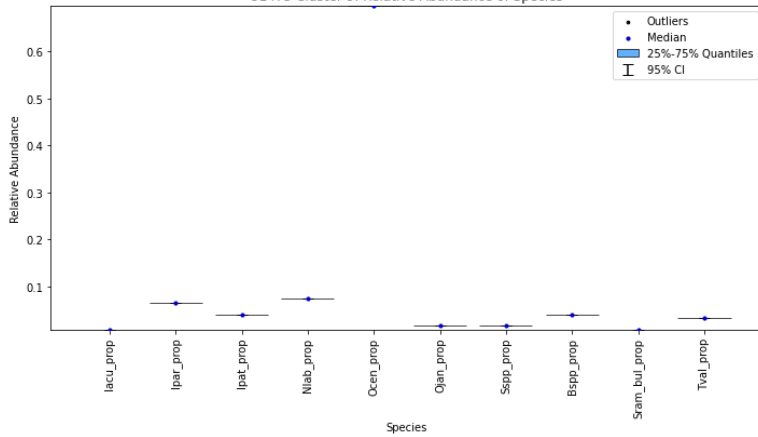
U1475 Cluster 5: Relative Abundance of Species



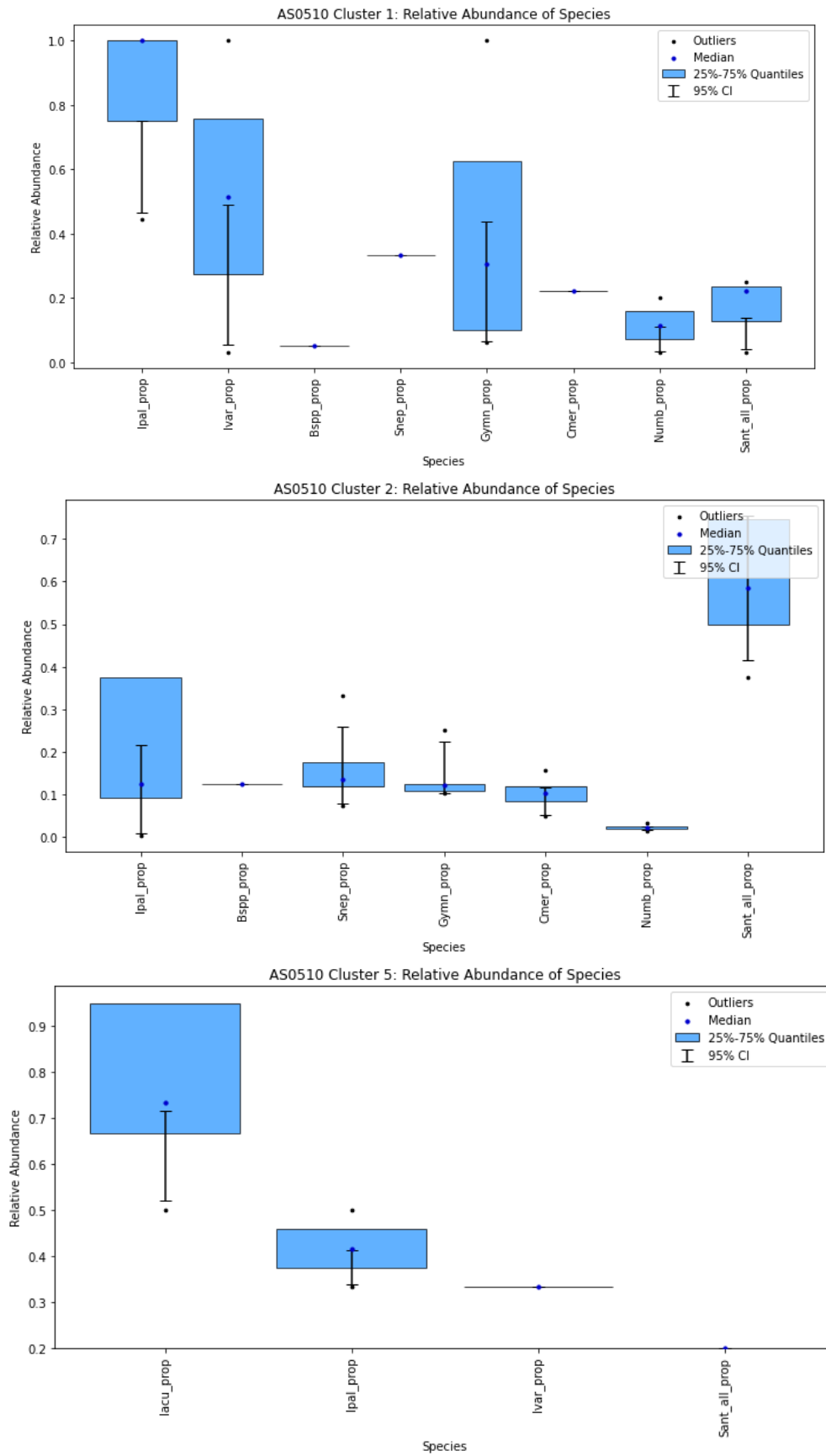
U1475 Cluster 6: Relative Abundance of Species



U1475 Cluster 9: Relative Abundance of Species

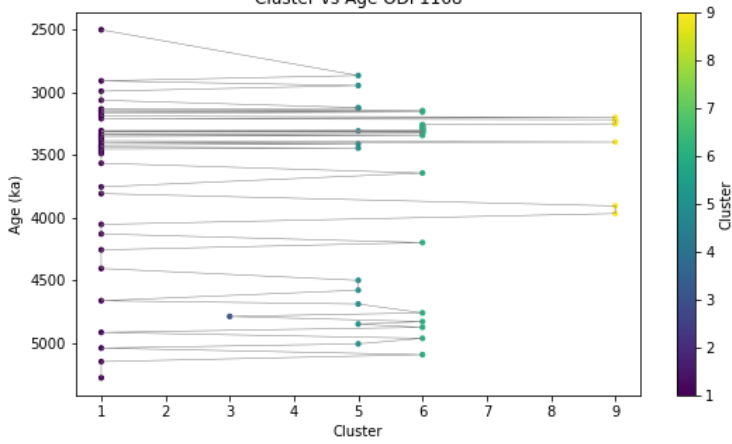


Appendix G: Boxplots AS05-10

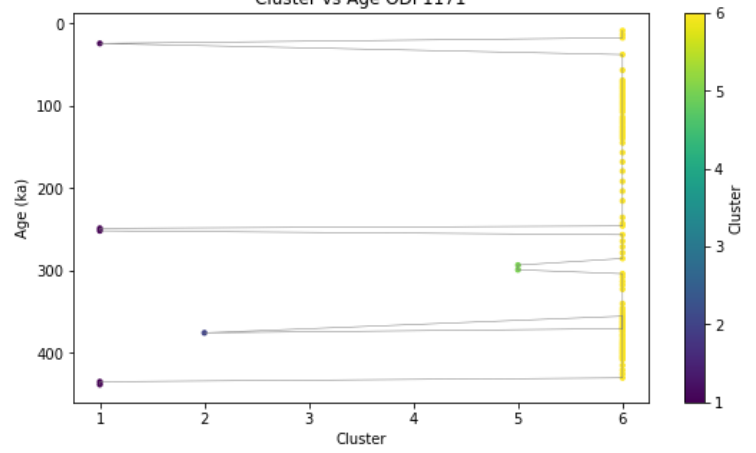


Appendix H: Cluster against Age

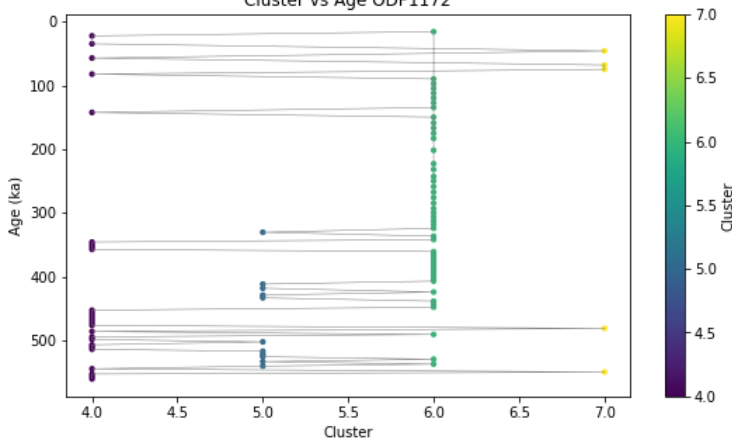
Cluster vs Age ODP1168



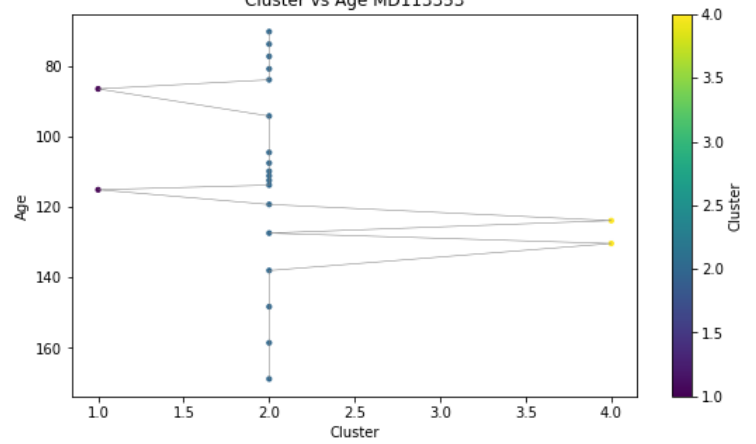
Cluster vs Age ODP1171



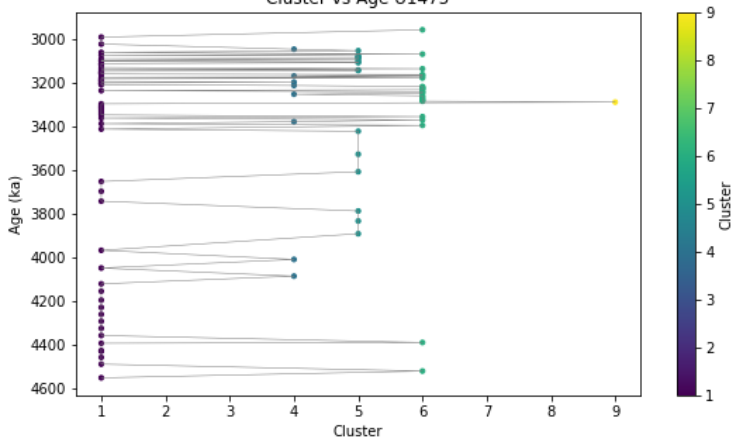
Cluster vs Age ODP1172



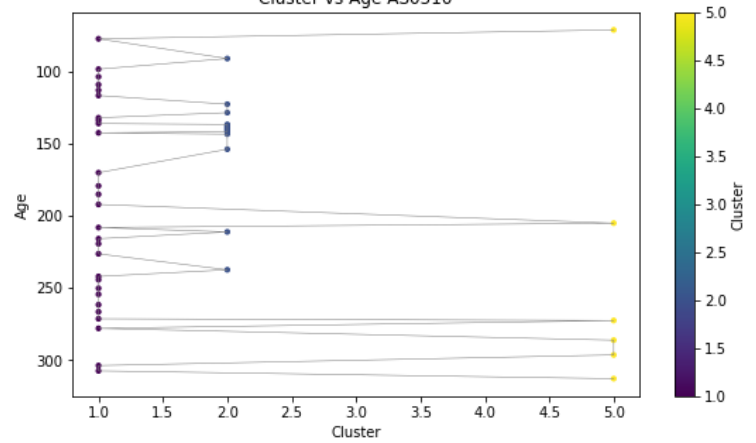
Cluster vs Age MD113353



Cluster vs Age U1475



Cluster vs Age AS0510



Appendix I: Individual dinocyst species analysis: relative abundance against temperature

