




OPEN

DATA DESCRIPTOR

Mediterranean marine sediment cores database: unlocking paleoclimatic signals for the last 20,000 years

Allyson Viganò^{1,2}  & Claudia Agnini²

The Mediterranean Sea is a crucial area for studying past climate dynamics due to its unique geography and hydrology. We present PaleoMED20, a paleoclimatic database that spans nearly six decades of research in the Mediterranean, covering the last 20,000 years. With over 1,500 marine sediment cores, it includes oceanographic cruise data and descriptions of the related cores recovered. For each core, we gathered all available information on past environmental variables and their associated climatic indicators (proxies). Metadata were collected from over 400 scientific articles, and a sub-set of 36 well-studied and/or most relevant cores was selected from the total of 1,524 to explore the quality and utility of the database for further studies. Here we present how PaleoMED20 can be used to identify the available key environmental variables and proxies over the last 20,000 years, while also highlighting existing knowledge gaps. This database has the potential to inform future oceanographic expeditions in the Mediterranean and contribute to the development of strategies and policies for climate change mitigation and adaptation.

Background & Summary

The Mediterranean Sea has been a focal point for paleoclimatic research since the early days of the field¹. Its significance grew notably from the 1950s onward becoming a key reference area for paleoclimatic studies due to the numerous national and international projects, collaborations among various oceanographic institutions, and international oceanographic expeditions dedicated to its exploration/study.

Recently, the International Ocean Discovery Program (IODP), the leading marine drilling program, conducted two expeditions in the Mediterranean Sea: Expedition 401 and 402. Expedition 401, held from December 10, 2023 and February 9, 2024, explored the Mediterranean–Atlantic gateway, focusing on its evolution from the late Miocene to its present configuration². Expedition 402, which took place from February 9 to April 8, 2024, aim to investigate the early stages of seafloor spreading and the formation of oceanic crust in the Tyrrhenian Sea³. Together with previous oceanographic cruises, these two expeditions reinforce the role of the Mediterranean in enhancing our understanding of past climate dynamics, providing data for paleoclimatic reconstructions and future projections.

The importance of the Mediterranean Sea for (paleo)climatic research relies on its unique geographical and environmental feature: 1) its strategic position as an interface between the Atlantic Ocean and the surrounding continental landmasses makes it a key area for studying interactions between oceanic and terrestrial climate systems⁴ 2) its small size allows for a rapid and amplified response to climatic changes, which is crucial for understanding how regional climate systems react to global climatic events⁵.

Studying past climate conditions is essential to understand the Earth's climatic system and predicting its future evolution, and this is one of the primary goals of the scientific community⁶. This endeavour has become increasingly important in light of the current challenges posed by severe human-induced impacts such as extreme rainfalls and associated flooding, deteriorating water quality, and accelerated carbon emissions. All these factors influence sea water temperature and nutrient cycling, ultimately leading to ocean acidification⁷.

¹Dipartimento di Ingegneria Civile, Edile e Ambientale – ICEA, Università di Padova, 35131, Padova, Italy. ²Dipartimento di Geoscienze, Università di Padova, 35131, Padova, Italy. ✉e-mail: allyson.vigano@unipd.it

When assessing the ongoing hazards induced by specific climate forcing, it is crucial to recognize that predictions, assessments and simulations of risk are inherently influenced by a certain degree of uncertainty that needs to be minimized as much as possible even though can never be completely eliminated⁸.

In both present and geological past, uncertainty frequently arises from insufficient quantitative measurements (input data) and/or difficulties in establishing cause-and-effect relationships within complex processes that involve physical, biological, and human systems⁹. As we move further back in time, the quality and the quantity of documentary and instrumental data inevitably decline. This makes geological archive and paleoclimatic proxies indispensable for reconstructing past climate. Among these geological archives, marine sediment cores, offer continuous and well-preserved records, which are essential for investigating longer time-series with resolution (centennial to millennial) that is sufficient to allow for the identification of rapid climatic oscillations⁹. Given the importance of geological archives for understanding past climates, producing this synthesis of data from marine sediment cores collected in the Mediterranean Sea serves as a valuable tool for scientists. It offers easy access to published data, allows for the integration of existing information, and supports the incorporation of new findings, thereby encouraging broader discussion and more comprehensive interpretation.

The rationale for focusing on the last 20,000 years is based on several factors. First, this period offers a highly detailed understanding of climatic evolution and various climatic regimes, including the Last Glacial Maximum (LGM), deglaciation, and interglacial periods^{10–12}, offering a solid foundation for exploring how these climatic phases can inform future scenarios. Second, the dataset has sufficient temporal resolution—enhanced also by the construction of composites—to capture variations on a sub-millennial scale and to identify extreme and rapid climatic events (Supplementary Information, Figure S1). Although it would be desirable to expand the record further back in time, it presents objective limitations. Both resolution and data completeness tend to decrease rapidly as we move toward older intervals, bringing with it considerable uncertainties. Climatic events are typically identified by changes in one or more environmental variables (e.g. temperature) which, however, cannot be directly measured in the geological past. Instead, they are usually assessed indirectly using specific indicators, known as proxies, such as marine faunal assemblages, oxygen and carbon stable isotopes and/or other sediment properties¹³.

While numerous reviews from the Mediterranean area cover periods ranging from the last 500 years¹⁴ to 6000 years¹⁵, there are relatively fewer reviews based on marine sediment core records extending back 20,000 years. One possible explanation is that, while the data come from various cruises, not all of these expeditions were designed with the goal of recovering long sediment sequences. Many were conducted for different research purposes, such as the collection of modern surface sediments, which limited the depth and temporal coverage of the sediment cores recovered (Supplementary Information, Figure S2).

Most of the published reviews focus primarily on specific proxies, such as sediment elemental data¹⁶. Therefore, there is a need for a centralized Mediterranean dataset that includes all currently available proxies and extends the timeline back to 20,000 years. This is particularly significant as it allows for the study of climate beyond the current interglacial period, specifically during a time when the Earth was characterized by larger polar ice caps, lower CO₂ levels, and cooler temperatures¹⁷.

We compiled data from over 1,500 marine sedimentary cores from peer-reviewed studies and public data repositories. We provide an exploratory analysis of the primary environmental variables/parameters and climatic proxies used to reconstruct the past 20,000 years in the Mediterranean Sea, categorized by geographic areas.

The PaleoMED20 database¹⁸, as part of the RETURN project¹⁹, lays the groundwork for future research aimed at addressing these goals: (1) investigating underrepresented variables, (2) proposing new proxies for estimating (paleo)environmental parameters, and (3) guiding the planning of future oceanographic expeditions in the Mediterranean Sea.

Methods

Marine sediment cores metadata collection. PaleoMED20¹⁸ is a comprehensive compilation of meta-data sourced from shipboard data and/or scientific publications. It systematically organizes and documents published data related to Mediterranean marine sediment cores, providing easy, user-friendly access to the information needed by researchers and stakeholders. Paleoclimate records from various marine sedimentary cores located in the Mediterranean Sea were compiled from a systematic review of the public literature for inclusion in the database (Fig. 1a). These records were collected from publicly archives including PANGAEA²⁰ (<https://www.pangaea.de/>), NOAA²¹ (<https://maps.ngdc.noaa.gov/viewers/imlgs/samples>), SeaDataNet²² (<https://www.seadatanet.org/>) and ISMAR²³ databases, and using web search engines for the scientific literature. These include: Google Scholar (<https://scholar.google.it/>), Web of Science (<https://mjl.clarivate.com/home>), Scopus (<https://www.scopus.com/pages/home#basic>), Scientific Ocean Drilling Bibliographic Database²⁴ (<http://iodp.americangeosciences.org/vufind/>), UNESCO Digital Library²⁵ (<https://unesdoc.unesco.org/home>).

For the cruise expeditions and core collection, we used the search filters available from all public archives, specifically restricting our search to the Mediterranean Sea through the search terms: “Mediterranean Sea”, “Mediterranean”, and/or using the area selection tool (between latitude 46° N and 30° N and longitude 9° W and 40° E). We also used the “advanced search” option available in SeaDataNet (<https://www.seadatanet.org/>; “search data”) by selecting the platform type “research vessel”, the discipline “marine geology” and the sea region (“Mediterranean Region”). Following a literature review based on shipboard data from the recovered cores, an initial screening was conducted to retain only those cores that covered the time range of interest based on the relative chronologies reported.

We collected climatic data beginning with one of the pioneering expeditions in the Mediterranean, the Swedish Deep Sea Albatross Expedition²⁶ (1947–1948), and extending up to 2015. We structured the PaleoMED20 database¹⁸ chronologically based on the year of the oceanographic expedition/cruise and the

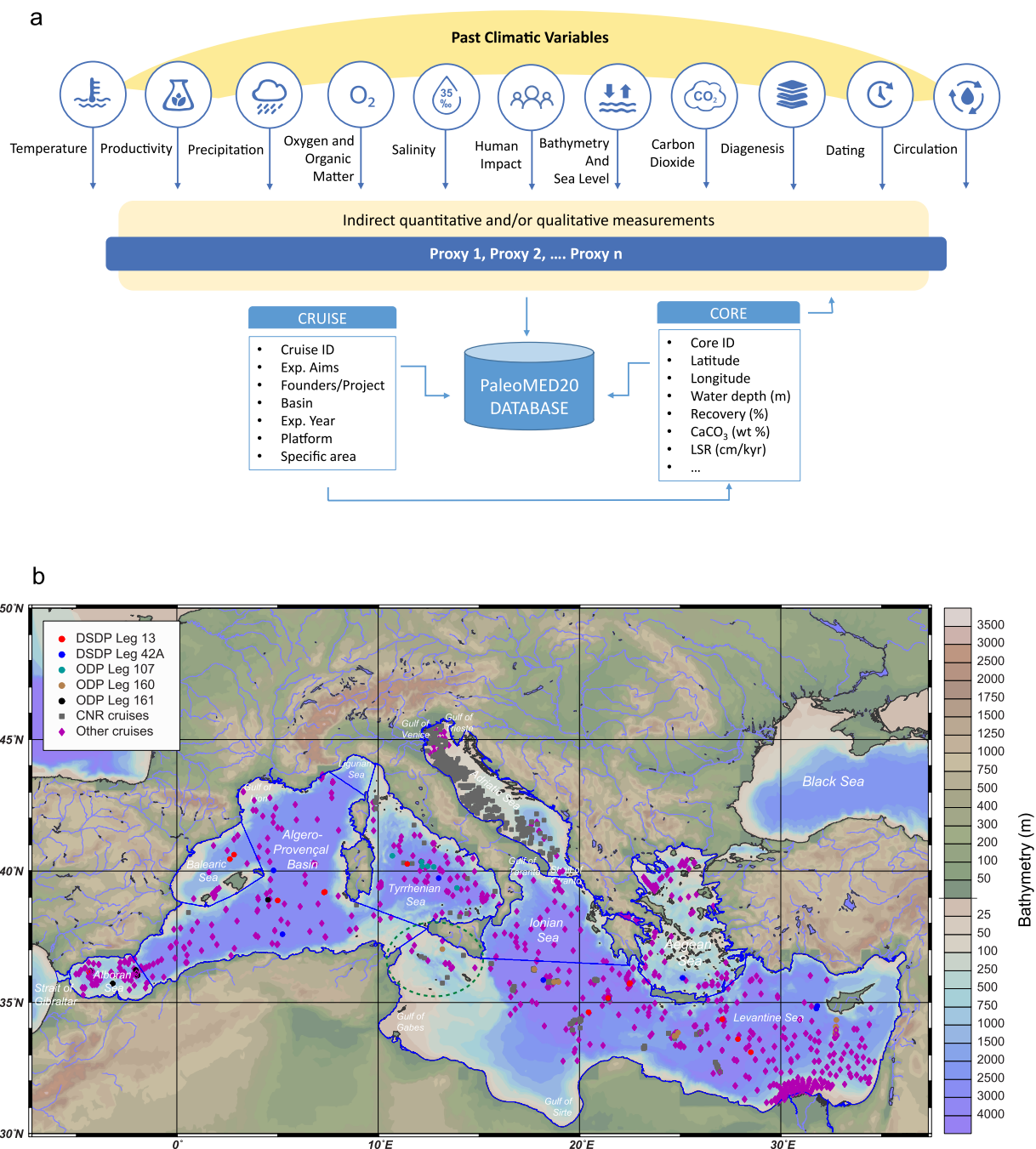


Fig. 1 (a) Structure, metadata, and key paleo climatic variables stored in the database. (b) Map of marine sedimentary cores drilled in the Mediterranean Sea. Blue lines are the limits of the Mediterranean sub-basins according to the IHO (International Hydrographic Organization; <https://www.marinerregions.org>). The dashed green line identifies the cores located in the Central MED area. Dots = cores recovered by DSDP/ODP expeditions; squares = cores recovered by CNR (National Research Council) Italian projects; diamonds = cores recovered by other European or national expeditions, research institutions and universities.

corresponding cores recovered. It includes a large variety of metadata related to oceanographic cruises and the cores retrieved (Table 1).

For the oceanographic cruises, we categorized the metadata under the following main labels: cruise ID, founders/project, expedition objectives, basin, expedition year, platform, and specific area. The specific area distinguishes among minor sub-basins (e.g. Alboran Sea, Tyrrhenian Sea, etc.) within the primary basin. Similarly, by narrowing the scale from cruise to core level, for each core, we provided the following metadata: core ID, latitude, longitude, water depth (m), core length (m), recovery (%), time represented (relative and/or absolute chronology), CaCO_3 (%), LSR (cm/kyr), lithology, reference (data source), publication year, DOI link, presence/absence of paleoclimatic data. Geodetic data are referenced to the WGS84 (World Geodetic System 1984) and

COLUMN NAME	UNITS	DESCRIPTOR	TYPE
Cruise ID		The name of the oceanographic cruise/expedition during which the core(s) were recovered	Character
Exp. Aims		Main objectives of the expedition	Character
Founders/Project		Names of the expedition's founders and/or associated projects	Character
Basin		Specific region of the Mediterranean Sea where a core was collected, categorized as WMED, EMED, central MED, or Adriatic Sea	Character
Exp. Year		The year of the expedition	Integer
Platform		Name of the research vessel	Character
Specific area		Name of the specific sea within the main basin	Character
Core ID		The name of the core recovered	Character
Latitude	Decimal degrees	Geographic coordinate of core location	Numeric
Longitude	Decimal degrees	Geographic coordinate of core location	Numeric
Water depth	Meters	The water depth at which the core was recovered	Numeric
Core length	Meters	Refers to the total length of a core	Numeric
Recovery	Percentage (%)	Refers to the proportion of the total length of a core that is successfully retrieved compared to the length of the core	Numeric
Relative_chronology		Epoch range represented/studied	Character
Absolute_chronology	kyr	Chronological range represented/studied	Numeric
CaCO ₃	Percentage (%)	Mean percentage of calcium carbonate documented within a core	Numeric
LSR	cm/kyr	The Linear Sedimentation Rate (LSR) refers to the rate of sediment accumulation that has occurred during the time interval represented by the core	Numeric
Archive		Marine sediment with additional information on the presence of sapropels and/or tephra layers	Character
Lithology		Description of the lithology and/or lithological units of the core	Character
Reference		Reference to the original publication (source) used to retrieve the data	Character
Publication_year		Year in which the publication was released	Numeric
DOI		Digital Object identifier of the original publication	Character
LINK		Link to the publication or to the compilation source	Character
Paleoclimatic data		Refers to the presence or absence of paleo climatic data (yes or no)	Character
Authors		The authors who studied the core	Character
Proxies		Quantitative or qualitative indicator used to infer indirect information about a variable, as analyzed by the authors in the core	Character
Variables investigated		Variable examined by the authors through the use of proxies	Character
Time start	kyr	Refers to the minimum age of the core as analyzed by the authors	Numeric
Time end	kyr	Refers to the maximum age of the core as analyzed by the authors	Numeric

Table 1. Description of the column names (labels) contained in the dataset.

are expressed in decimal degrees. The term “basin” specifies the “region” of the Mediterranean Sea from which a core was retrieved, categorized as Western Mediterranean (WMED), Eastern Mediterranean (EMED), Central Mediterranean (central MED) and Adriatic Sea. In addition, for each core, we provide bibliographic information, the time range studied by the authors and the paleoclimatic data available, including variables and proxies. This information is organized as follows: core ID, authors, other cores studied by the authors, relative and/or absolute chronology, time_start (kyr), time_end (kyr), LSR (cm/kyr), proxies, variables and comments.

Dataset-based classification of paleoclimatic variables. In this section, we have organized the climatic variables categories, as they are presented in the database. Marine sediment cores from the Mediterranean Sea offer high-resolution metadata, ranging from a millennial to centennial scale, or even decadal scale in regions with high-sedimentation rates¹⁶, allowing for detailed investigations of past environmental conditions. To facilitate navigation and ensure data reuse by researchers and various stakeholders, the paleoclimatic data included in PaleoMED20¹⁸ has been organized into 11 categories, each corresponding to a distinct environmental variable (Tables 2, 3):

1. **Temperature.** This category comprises two sub-variables: sea surface temperature (SST) and bottom water temperature (BWT). Of these, sea surface temperature is particularly significant for paleoclimate reconstructions and current climate policies²⁷, as the upper ocean is more directly linked to atmospheric circulation and stores more energy than the deep ocean^{28,29}.
2. **Chronology.** This category includes various chronological methods used in the literature to establish precise age constraints, which are crucial for analyzing time-series data³⁰. These methods encompass both relative dating—often based on shipboard observations—and absolute dating, each carrying uncertainties. The datasets have been acquired over several decades, which implies that data precision may vary according to the period and methodologies used for the measurements. In marine sediments, absolute dating methods primarily include radiocarbon (¹⁴C) dating (covering up to ~50,000 years)^{31,32} and ²¹⁰Pb dating for modern sediments deposited within the past 100–150 years ref. ³³. For the last 20,000 years, as reflected in

Main variable	Sub-variables	Specific parameters	Proxies
Temperature	Sea surface temperature (SST)		Fossils groups: calcareous nannofossils, diatoms, dinoflagellates, foraminifera, other marine fossil groups (e.g. pteropods, radiolarians...); faunal transfer functions; $\delta^{18}\text{O}$ vs VPDB (carbonates); $\delta^{44}\text{Ca}$ (foraminifera); coral bands (order Scleractinia); trace metals: Ba/Ca, Sr/Ca, Mg/Ca (e.g. from foraminifera and corals). Molecular biomarkers (marine): alkenones – U^{K}_{37} index; tetraunsaturated C37 (C37:4); TEX86 (Crenarchaeota - Picoplankton).
	Bottom water temperature (BWT)		Benthic foraminifera; $\delta^{18}\text{O}$ (benthic foraminifera); Mg/Ca (benthic foraminifera)
	Atmospheric temperature		δD (e.g. from ice cores); pollens; molecular biomarkers (terrestrial): n-nonacosane, n-hexacosanol (lipid compounds).
Dating			Cosmogenic isotopes: ^{10}Be , ^{14}C , ^{210}Pb ; tephra (Ar-Ar, K-Ar, ^{14}C radiometric/beta counting, AMS ^{14}C , (U-Th)/He, U-Pb, and $^{238}\text{U}/^{230}\text{Th}$ zircon dating, etc.)
Productivity and nutrients	Productivity	Biological productivity (export production)	Fossil groups: calcareous nannofossils, diatoms, dinoflagellates, foraminifera, other fossil groups (e.g. pteropods, radiolarians...); sapropels; $\delta^{15}\text{N}_{\text{org}}$; marine barite (BaSO_4); Ba excess (Babio); Ba_{org} ; Ba/Al; TOC (%) or C_{org} (%); $\delta^{13}\text{C}_{\text{org}}$ (of organic matter); $\delta^{13}\text{C}$ (carbonates); carbonate or opal content; accumulation rates: Opal AR = Opal Accumulation Rate; BFAR = Benthic Foraminiferal Accumulation Rate; NAR = Nannofossil Accumulation Rate; MAR = Mass Accumulation Rate (carbonates or opal); Marine Barite Accumulation Rates
		Seasonal upwelling	Cd/Ca (e.g. in corals)
		Nutricline position	<i>Florisphaera profunda</i> and N ratio (from calcareous nannofossils)
	Nutrients	Phosphorus and phosphate content	P/Ca and Cd/Ca
		Nitrate utilisation	$\delta^{15}\text{N}$ (organic matter); $\delta^{30}\text{Si}$ (diatoms)
Oxygen and organic matter	Anoxia		C_{org} ; sapropels; V/(V + Ni); authigenic U; other proxies: laminated sediments, black sediments, fossils traces (e.g. chondrites), taxonomic uniformism (cf benthic foraminifera), opportunistic taxa, only plankton and nekton. Elemental ratio: U/Th, Mn/Al, Ni/Co, V/Cr; Redox-Sensitive Trace Elements (RSTE): Mo, U, V, Cr, Re, Fe, Co, Cd, Ti, Sb, As, Mn, Ni, Cu, Zn; isotopes: $\delta^{57}\text{Fe}$ ($^{57}\text{Fe}/^{54}\text{Fe}$ ratios); $\delta^{98/95}\text{Mo}$
			Biomarkers: isorenieratene = green sulfur bacterial (GSB) pigment
	Sulfidic conditions		$\delta^{34}\text{S}$ (of barite or pyrite); pyrite (FeS_2); S species (e.g. S total, S pyrite, organic S and elemental S); S pyrite/Al
	Oxic conditions		V/Sc; benthic foraminifera (potentially infaunal taxa) for seasonal (monthly) oxygen fluctuations
	Organic matter	Origin of organic matter (terrestrial vs marine)	
Chemical quality			HI (hydrogen index)

Table 2. Inventory of significant variables and proxies in marine sedimentary core archives, in descending order from the most extensively studied to the least. The list includes main variables, sub-variables, specific parameters and proxies.

our dataset, conventional ^{14}C dating is the most widely applied method for dating carbon-bearing materials preserved in sediments, including tephra layers. These layers can also be dated using other radiometric techniques, such as K/Ar, Ar/Ar, (U-Th)/He, U-Pb, and $^{238}\text{U}/^{230}\text{Th}$ zircon dating³⁴ (PaleoMED20, “Variable and Proxies” section).

Since 1977, ^{14}C dating has become the most widely used chronological framework in paleoclimatology and related disciplines, owing to the advent of accelerator mass spectrometry (AMS), which significantly reduced the required carbon sample size and improved analytical precision, replacing traditional radiometric approaches^{35,36}.

While different dating methods and age-depth models vary in their assumptions—such as reservoir age corrections, calibration curves, and linear interpolation—their reliability is thoroughly assessed in the original publications, all of which are cited in PaleoMED20¹⁸. To maintain the fidelity of the source material, PaleoMED20 preserves the original chronological frameworks provided by each study. Methodological details are documented where available, ensuring transparency and equipping users with the necessary information to evaluate or align age models across sediment cores according to their specific research goals.

- Productivity.** This category comprises information on biological productivity conditions (high/low), presence/absence of seasonal upwelling, nutricline depth, and type and amount of nutrients.
- Oxygen and organic matter.** Oxygen levels and organic matter content are grouped in the same category because both can serve as oxygen indicators. Low oxygen levels are typically associated with high organic matter content³⁷. While this merging represents a simplification, it was done to streamline the structure of the dataset and make it more accessible, not only to researchers but also to a wider audience.
- Rainfall.** This category includes several indicators for rainfalls such as vegetation types, amount of river runoff, aridity conditions and ice volumes. Precipitation, along with temperature, is a key driver of the hydrological cycle as it controls the movement and distribution of water through processes such as evaporation and condensation. Changes in precipitation impact freshwater availability, river flows, and groundwater levels, thereby influencing drought and flood patterns and affecting ecosystems and water resources²⁷.
- Salinity.** Salinity is a crucial environmental parameter that influences past and present oceanic currents and global heat transport which in turn regulates the (paleo) climate. Higher salinity increases water

Main variable	Sub-variables	Specific parameters	Proxies
Precipitation	Type of vegetation/ biomarkers		Pollens; n-alkanes produced by vegetation (TERR-alkanes); Average Chain Length (ACL) of n-alkanes; Glycerol Dialkyl Glycerol Tetraethers (GDGTs)
	River runoff	Fluvial input and source of sediments (terrigenous provenance)	Reworked coccoliths (RCs); sapropels; Mn/Ca; Ba/Ca; clay minerals: chlorite, illite, random mixed-layer clays, kaolinite, palygorskite and sepiolite, smectite. Specific elements: Ti; V and Ba (e.g. in sapropels); Co; Sr; Cr; Ni. Elemental ratios: Cr/V; Mg/Ca; Ba/Ca (e.g. foraminifera or corals); Mg/Al; K/Al; Rb/Al; Ca/Al; Sr/Ca. Mineral phases: dolomite; quartz; chlorite; mica(muscovite-illite); serpentine; feldspars. Isotopes: ϵ Nd; δ^7 Li; 87 Sr/ 86 Sr (in sapropels). Rare Earth Elements (REE): lanthanum (La), cerium (Ce), praseodymium (Pr), neodymium (Nd), promethium (Pm)
	Continental aridity	Dust (aeolian input)	Elemental ratios: Zr/Al, Ti/Al, Si/Al; elements: e.g. Silicon (Si); iron oxides: hematite; Eolian Sortable Silt (ESS)*; clay minerals: palygorskite and others (smectite, kaolinite, illite...); ions: K +; Rare Earth Elements (REE) ratios: La/Lu
	Ice volume		δ^{18} O in carbonates (e.g. bulk, foraminifera, corals,...), ice cores, tree rings,...; IRD = ice-raft detritous (coarse grained sediments)
Salinity			δ^{18} O (planktonic foraminifera); alkenones (SST) combined with δ^{18} O (planktonic foraminifera); transfer functions; Mg-calcite/calcite
Circulation and water masses	Circulation	Water circulation: water mass tracers/water exchanges variations	143 Nd/ 144 Nd (ϵ Nd from diagenetic Fe-Mn coatings on corals, foraminifera, from biogenic apatite of fish debris/teeth, from bulk sediment); 206 Pb/ 204 Pb; 176 Hf/ 177 Hf (expressed as ϵ Hf values); 187 Os/ 188 Os; 87 Sr/ 86 Sr; 10 Be/ 9 Be
		Deep-water formation	δ^{13} C (benthic and planktonic foraminifera)
		Deep-water flow/Bottom current velocity	231 Pa/ 230 Th (in sediments); Zr/Al (e.g. in contourites); radioactive 14 C (foraminifera); magnetic susceptibility; Sortable Silt (SS); UP10 index; glauconite mineral
		Atmospheric circulation	Pollens; dust; terrestrial biomarkers; grain-size of terrigenous fractions; major and minor elements
	Sea water gradients	Longitudinal and vertical gradients	$\Delta\delta^{18}$ O and $\Delta\delta^{13}$ C
	Water mass properties	Seawater pH	δ^{11} B from carbonates (e.g. corals order Scleractinia)
		Alkalinity	Ba/Ca (foraminifera)
Bottom waters conditions		Benthic foraminifera; δ^{13} C (benthic foraminifera); Cd/Ca (benthic foraminifera)	
Carbonate saturation state (Ω), CCD and lysocline		Dissolution patterns on microfossils	
Bathymetry and Sea level	Bathymetry		Benthic foraminifera; planktonic vs benthic foraminifera ratio (P/P + B)
	Sea level	Sea level fluctuations	Archeological remains; corals (order scleractinia) on coastlines; conversion from seawater δ^{18} O; foraminiferal δ^{18} O; gastropods (vermetid reefs); paleobeaches, marsh environments or lagoons; sedimentary facies; serpulid overgrowth on submerged speleothems; slope of river terraces; speleothems; system tracks of a depositional sequence; tidal notches
		Fluvial sediment transport (grain-size variations)	Si/Al; Zr/Rb; bulk density
Diagenesis and sediment alteration			Elemental ratio: Sr/Ca; magnetic parameters: ferrimagnetic minerals; bioturbation; low sedimentation rates; reworking of microfossils; microfossils etching/overgrowth
Human activities	Agriculture/Deforestation		Antropogenic pollens and benthic foraminifera
	Paleofires**		Charcoals (CHAR) (small carbonized particles)
CO ₂	Surface ocean		δ^{11} B (11 B/ 10 B); $U^{K_{37}}$; δ^{13} C (alkenone); δ^{13} C (phytol); sedimentary bulk $\delta^{13}C_{org}$
	Atmosphere		Stomata index (plant leaves); δ^{13} C (paleosols); CO ₂ (air from polar ice cores)

Table 3. Inventory of significant variables and proxies in marine sedimentary core archives, in descending order from the most extensively studied to the least, starting from the last parameter reported in Table 2. The list includes main variables, sub-variables, specific parameters and proxies. *used also for grain size analysis; **natural or antropogenic.

density, promoting water mass sinking and driving deep ocean circulation (i.e., thermohaline circulation)³⁸. Recently, sea surface salinity has been included to the list of key indicators for assessing the impact on marine ecosystems²⁷.

- 7. Circulation.** This category includes a series of indicators that describe water masses: water currents (e.g. water exchanges, deep water flows, bottom current velocity), sea water gradients (vertical and longitudinal) and water mass properties (e.g. carbonate saturation state Ω , CCD and lysocline; pH; alkalinity).
- 8. Bathymetry and sea level.** Bathymetry and sea level are more easily available in marine coastal areas but can also be documented in shelf and offshore environments. The global rise in sea level is currently a major focus of scientific research due to the impact on low-lying Mediterranean coasts. This rise contributes to coastal flooding, erosion, and increased salinity in freshwater reserves, making coastal settlements and activities more vulnerable³⁹. In the northern Adriatic lagoons and coastal plains, the mean eustatic rise in sea level is further exacerbated by local subsidence⁴⁰.
- 9. Diagenesis.** Diagenetic processes influence the preservation and, in turn, the interpretation of geological records over time. These processes alter chemical and/or physical features affecting the reliability of climate

reconstructions. Understanding diagenetic processes is essential for distinguishing between original signals and post-depositional alterations⁴¹. In this work, the variable “diagenesis” includes the alteration of sediments and its components (e.g. microfossils).

10. **Human impact.** This category includes the impact of human activities on Mediterranean landscape, particularly over the last two millennia¹⁵. Our database covers factors such as agriculture and deforestation and, to some extent, paleo fires.
11. **Carbon dioxide.** This category refers to surface ocean CO₂ concentrations derived from marine archives. However, past changes in CO₂ concentration can also be determined from terrestrial archives and ice cores (Table 3).

All these variables are estimated indirectly through qualitative and/or quantitative assessments. For instance, temperature and CO₂ levels can be quantitatively reconstructed, providing precise numerical data (with associated errors) that illustrate their long-term fluctuations over millennia. In contrast, other variables are often assessed qualitatively. For example, pollen assemblages from marine sediment cores can indicate shifts in vegetation patterns influenced by moisture availability, indirectly reflecting past precipitation regimes. While these qualitative estimations may lack numerical values comparable to modern instrumental/numerical data, they remain crucial for understanding past climates.

Dataset-based classification of proxies. In this paragraph, we present an overview of some of the proxies included in the database (“Variable and Proxies” spreadsheet). Proxies are derived from quantitative or qualitative indirect descriptors⁴², and include the biological, chemical, and physical properties of sediments. The use of proxies is essential for validating climate models for periods before the availability of instrumental records^{30,43}, which for the Mediterranean region extend back only a few centuries¹⁵. Thus, proxies are indispensable for understanding climate variability in the past. However, producing patterns of environmental descriptors remains challenging due to the limited spatio-temporal resolution and/or low reliability of certain proxy data⁴⁴. Additionally, proxies are sometimes local indicators, sourced from sediments transported by rivers and deposited in coastal or shelf environments or within the basin with minimal or no movements. Consequently, data from a single proxy could hold regional rather than global insights.

In open sea areas, most proxies are derived from biogenic marine sediments. Mediterranean Pliocene and Quaternary sediments often contain volcanic ashes (tephra layers)^{45,46} and sapropels (organic rich layers)⁴⁷. Tephra layers are unconsolidated pyroclastic (fragmental), ejected during volcanic eruptions, and deposited over land and/or the seafloor, forming a widespread blanket of the same age^{34,46}. They are useful for dating sediments because they can be traced to specific volcanic events. In particular, tephra from the Italian and Aegean volcanic provinces⁴⁵ are important tools for accurately dating the organic materials encased within or directly bracketing the tephra layer itself using different radiometric methods, including ¹⁴C radioactive decay (radiometric/beta counting, AMS) (Table 2), which is one of the most common method to date tephras erupted within the past 60,000 years ref. ⁴⁶.

Sapropels, on the other hand, form under conditions of very low or no oxygen. These anoxic conditions arise from the high primary productivity in the water, where organisms consume large amounts of oxygen, and/or from water stratification, which reduces the exchange of oxygen between the surface water and the deeper layers. Such conditions lead to sediments rich in organic matter, linked to the astronomically timed intensification of the African monsoon and its resulting rainfall. The rainwater, transported by the Nile River to its delta in the eastern Mediterranean, increases nutrient discharge (productivity) and promotes water stratification^{12,48}, which in turn enhance the preservation of organic matter.

In a sapropel, various properties serve as proxies, such as Total Organic Carbon (TOC) %, microfossil content and abundance, and trace elements. These proxies can be used to reconstruct different parameters, including productivity, oxygen levels, and precipitation regimes (Tables 2, 3). Traditional common proxies for productivity also include biogenic opal content or organic matter content, although the latter is susceptible to preservation biases⁴⁹. Alternative geochemical indicators for past productivity include: carbonates or opal accumulation rates, barium (Ba) and marine barite⁵⁰, as well as $\delta^{15}\text{N}$ in marine organic matter⁵¹. Traditionally, changes in the abundance of benthic foraminiferal assemblages have been used to infer past oxygen conditions⁵². Additionally, oxygen levels can also be reconstructed using redox-sensitive metal proxies and redox-sensitive element isotope systems⁵³. The methods used by different authors for the concentration measurements of various elements and their ratios (e.g., U, Th, Co, Sr, Ni, etc.) may vary, and the associated instrumental or technique-related biases can differ. Users should be aware that differences in instrumentation, calibration protocols, and analytical resolution may introduce variability in elemental concentration data. Therefore, a basic understanding of the underlying methods is essential when comparing datasets across studies. Nonetheless, these techniques are thoroughly discussed in the original publications, which are properly referenced in the database.

It is noteworthy that the same variable can be reconstructed using multiple indicators. This cross-checking of different proxies for a single variable increases the reliability of the reconstruction. On the other hand, the same proxy can be an indicator of different variables, complicating the interpretation of its observed changes over time.

For instance, precipitation can be inferred using multiple proxies: ⁸⁷Sr/⁸⁶Sr isotopic ratios and major elements⁵⁴ or pollens, but the latter could be controlled by other factors, such as temperature⁵⁵ and human activities⁵⁶.

Common proxies used in paleoclimate research are listed in Tables 2, 3, and are included in PaleoMED20¹⁸. Additional information, including the underlying assumptions for their use, methodological notes, and references, is available in the ‘Variables and Proxies’ section of the PaleoMED20 database¹⁸.

Data Records

The dataset is available at Figshare¹⁸, with this section being the primary source of information on the availability and content of the data being described. All the sources provided in this dataset are openly available.

Our database provides a comprehensive overview of 1,524 marine sediment cores collected from the WMED, EMED, Central MED, and the Adriatic Sea (Fig. 1b), with metadata further categorized by geographic areas. Although the Adriatic Sea is part of the EMED, it is treated separately due to its paleoclimatic relevance.

Ocean Data View software (version 5.6.5)⁵⁷ was used to map the distribution of the cores. The boundaries of the Mediterranean Sea and its sub-basins were defined using the Mediterranean Sea Area shapefile provided by the International Hydrographic Organization (IHO)⁵⁸.

This database also offers a well-organized collection of key environmental variables and climatic proxies used to reconstruct the last 20,000 years in the Mediterranean. To better understand the location and distribution of the cores, a summary of the Mediterranean's main features and basin divisions, along with additional information on record density over the last 20,000 years, is provided in the Supplementary Information. The data are archived in the Figshare repository¹⁸ under the name “PaleoMED20 database.csv”, a Microsoft Excel file containing the following spreadsheets:

1. **Core database:** contains the final core compilation; the data structure is described in the metadata file and in Table 1;
2. **Core distribution:** includes a description of the distribution of cores across the Mediterranean Sea, based on the information extracted from the database (equivalent to Fig. 2 and Table 4);
3. **Paleoclimatic data:** contains the paleoclimatic information of the cores, including proxies and climatic variables analyzed;
4. **Metadata:** contain the file metadata (equivalent to Table 1).
5. **Variables and Proxies:** contains a comprehensive list of all the variables and proxies compiled in this dataset.
6. **Record Density:** contains a filtered version of the ‘Paleoclimatic data’ spreadsheet focused on the following columns: Core ID, Authors, Time_start (in kyr), and Time_end (in kyr).

The Excel file is also organized with a proxy-centric filtering system, enabling the reader to efficiently access relevant information and facilitate cross-core comparisons. Below, we provide an overview of the data files stored in the database.

Core distribution. The core distribution file provides an overview of the number of cores recovered in the Mediterranean Sea by different oceanographic cruises and their distribution across the various basins.

In particular, since the 1950s, extensive oceanographic drilling has been primarily driven by European, national projects, research institutions, and/or universities, contributing to ~59% of the total marine sediment cores recovered, while CNR (Italian National Research Council) cruises have accounted for around 38% (Fig. 1b; Table 4). The Mediterranean Sea also hosts 46 DSDP (Deep Sea Drilling Project; 1968–1984)/ODP (Ocean drilling Project; 1985–2002)/IODP (Integrated Ocean Drilling Program; 2003–2013; International Ocean Discovery Program; 2013–2023 sites representing ~3% of the total (Table 4). The core distribution file also outlines an uneven distribution of cores, with higher densities in specific regions, particularly the Eastern Mediterranean and the Italian Adriatic coasts. Most cores were collected from the EMED (45%; Fig. 2), including the eastern basins and the Aegean Sea. The Adriatic Sea contributed approximately 30% of the total cores, while the WMED accounted for 24%, and the central Mediterranean only 1% (Fig. 2). These data are stored in the PaleoMED20 database.csv, “Core distribution” spreadsheet.

Paleoclimatic data. Paleoclimatic data are available for 465 cores out of 1,524 cores (31%) and are distributed as follows: 203 cores in the EMED, 169 in the WMED, 82 in the Adriatic Sea, and 11 in the central MED. The presence or absence of climatic data are reported in the PaleoMED20 database.csv, “Core database” spreadsheet, column “Paleoclimatic data”. For each core, we have documented the specific proxies used and the environmental variables reconstructed by the authors, based on the categories of variables outlined in the Methods section. These latter data are stored in the PaleoMED20 database.csv, “Paleoclimatic data” spreadsheet.

Variables and proxies. The table is organized into four main columns: 1) Variable: the climate variable being reconstructed; 2) Proxies: the abiotic and biotic indicators used to estimate the considered variable; 3) Assumptions: the underlying conditions under which a proxy can be used for variable reconstruction, along with its limitations; 4) Notes: additional comments.

Record density. The “Record Density” spreadsheet is a filtered version of the broader metadata and was used to assess the temporal coverage of the paleoclimate data across the last 20,000 years. The plot resulted (Figure S2, Supplementary Information) is constructed by assigning each record its temporal range, as defined by the Time_start and Time_end values reported by the original publications (“Authors”).

Technical Validation

No data were produced by the authors; the dataset only includes information from previously published and peer-reviewed studies (metadata) and publicly available online archives.

A subset of 36 cores was selected from the total of 1,524 cores recorded in the database to assess the technical quality and suitability for more in-depth studies. This selection was made to represent various Mediterranean

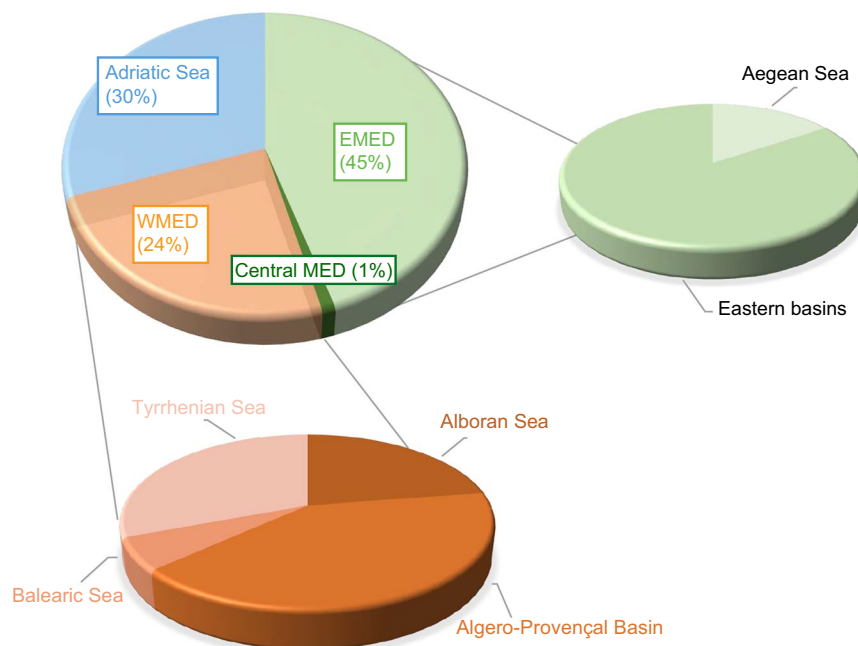


Fig. 2 Pie-chart of geographic distribution (%) of drilled cores recovered from the Western (WMED), Eastern (EMED), Central (Central MED) Mediterranean and from the Adriatic Sea. The EMED is further divided in: Eastern basins (including Ionian and Levantine Sea, and minor basins between them) and the Aegean Sea. The WMED is subdivided in: Tyrrhenian Sea, Alboran Sea, Algero-Provençal basin (which include the Gulf of Lion) and the Balearic Sea.

Cruises	Number of marine sediment cores	%
DSDP Leg 13	14	0.92
DSDP Leg 42 A	8	0.52
ODP Leg 107	7	0.46
ODP Leg 160	11	0.72
ODP Leg 161	6	0.39
CNR cruises	585	38.39
Other cruises	893	58.60
Total	1524	100.00

Table 4. Number and percentage of marine sediment cores collected from each oceanographic cruise.

sub-basins and water depths, including both well-studied cores and those that may be of interest for future research. (Fig. 3a,b). This spot-checking⁵⁹ focused on the following aspects:

- **Consistency and Completeness:** Ensuring the data was free of inconsistencies and missing values.
- **Utility:** Verifying that the data provided meaningful insights for both current and future research.
- **Reliability:** Cross-checking the data against established literature benchmarks for accuracy and reproducibility.

This approach offers an initial assessment of the PaleoMED20¹⁸ database's usefulness, providing an exploratory analysis of the key paleoclimatic variables used to reconstruct the last 20,000 years.

Core selection. The cores were chosen based on their geographic location (specific sub-basin), the availability of studies, reliable dating (e.g., multiple radiocarbon dates), continuity and high temporal resolution of the records, elevated sedimentation rates, variety of different proxies, balanced representation of various climate proxies, and the quality of the available data.

In the WMED, we selected a total of 10 cores, subdivided as follows: four in the Alboran Sea (ODP 976, KS8231, MD95-2043, TTR14-300G), three in the Balearic Sea/Algero-Provençal basin (ODP 975, MD99-2343, KESC9-14) and three in the Tyrrhenian Sea (ET91-18, KET8019, BS79-33). For the Central Mediterranean, we chose three cores (ODP 963, MD04-2797, CS70-5). The Adriatic Sea was considered separately, with 12 cores selected (Fig. 4): four from the Northern Adriatic (AD78-150, VE04-32, AD78-161), four from the Central

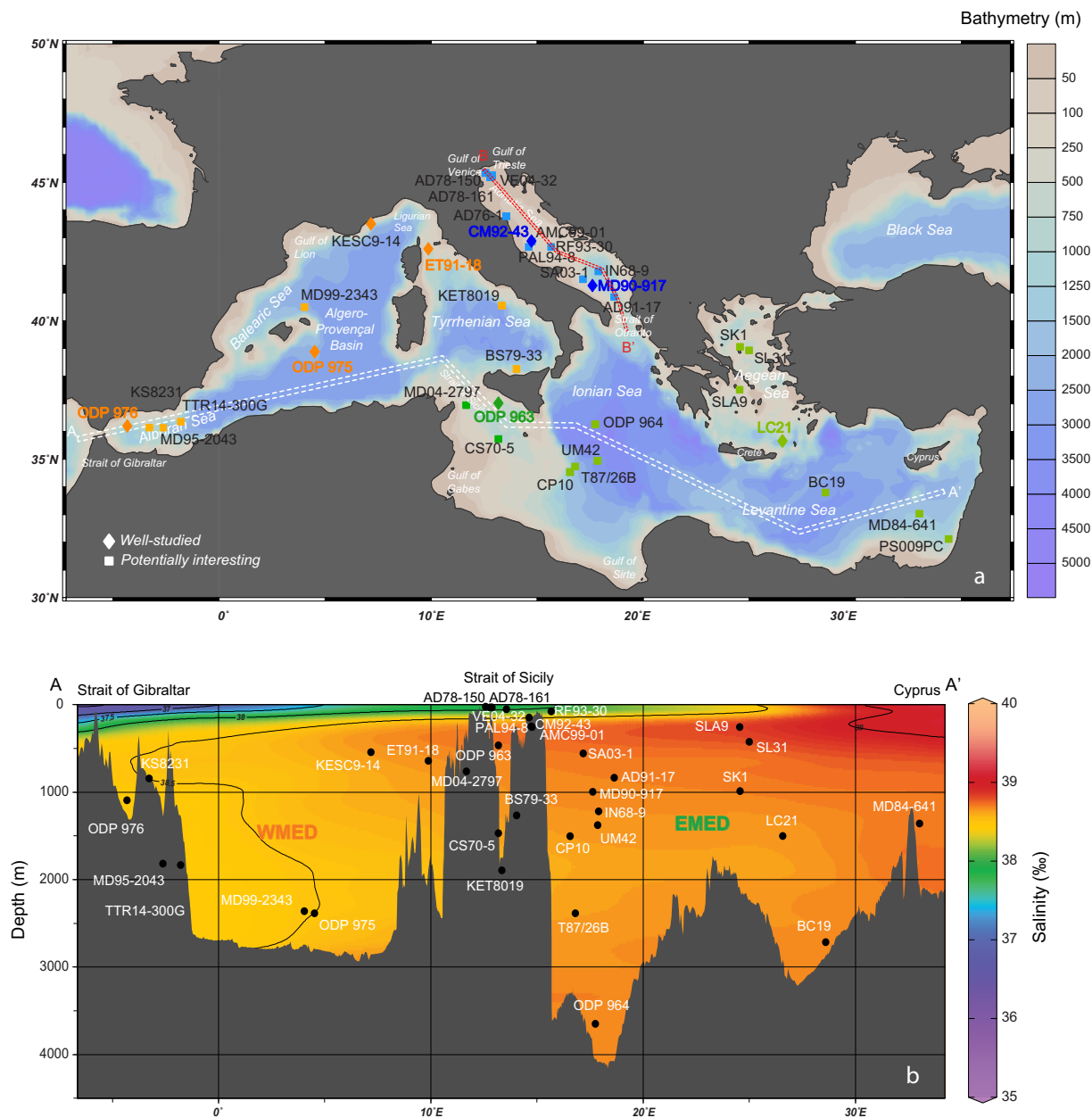


Fig. 3 (a) Bathymetric map of the Mediterranean Sea with locations of well-studied (diamonds) and/or potentially interesting cores (squares). Symbols colors: orange and light orange = WMED; dark green = Central MED; light green = EMED; blue and light blue = Adriatic Sea. (b) The W-E present-day salinity transect as reported in Fig. 3a (dashed red line). Figures were generated using the Ocean Data View software (version 5.6.5)⁵⁷. Vertical bathymetry profiles derived from the General Bathymetric Chart of the Ocean (GEBCO) 2023 ref.⁷⁵. Salinity data is from SeaDataNet (<https://www.seadatanet.org>) and was retrieved from SeaDataNet infrastructure at the end of July 2019.

Adriatic (CM92-43, AMC99-01, PAL94-8, RF93-30), and four from the Southern Adriatic (MD90-917, SA03-1, AD91-17, IN68-9). In the EMED, we selected 11 cores: seven from the eastern basins, including the Ionian Sea, Levantine Sea and smaller basins in-between (ODP 964, UM42, T87/26B, CP10, PS009PC, BC19, MD84-641) and four from the Aegean Sea (LC21, SLA9, SL31 and SK1).

Distribution and availability of (paleo)environmental data across Mediterranean regions. Paleoclimatic variables slightly vary among the Mediterranean Sea basins. In the Alboran Sea, most studies focus on temperature (38%), followed by precipitation (15%) and other environmental variables, with ODP Site 976 and MD95-2043 being the most extensively studied cores (Fig. 5a). The Algero-Balearic Sea shows a similar pattern, with temperature (22%) and AMS 14C dating (17%) being the primary focus, and ODP Site 975 standing out for its extensive documentation (Fig. 5b).

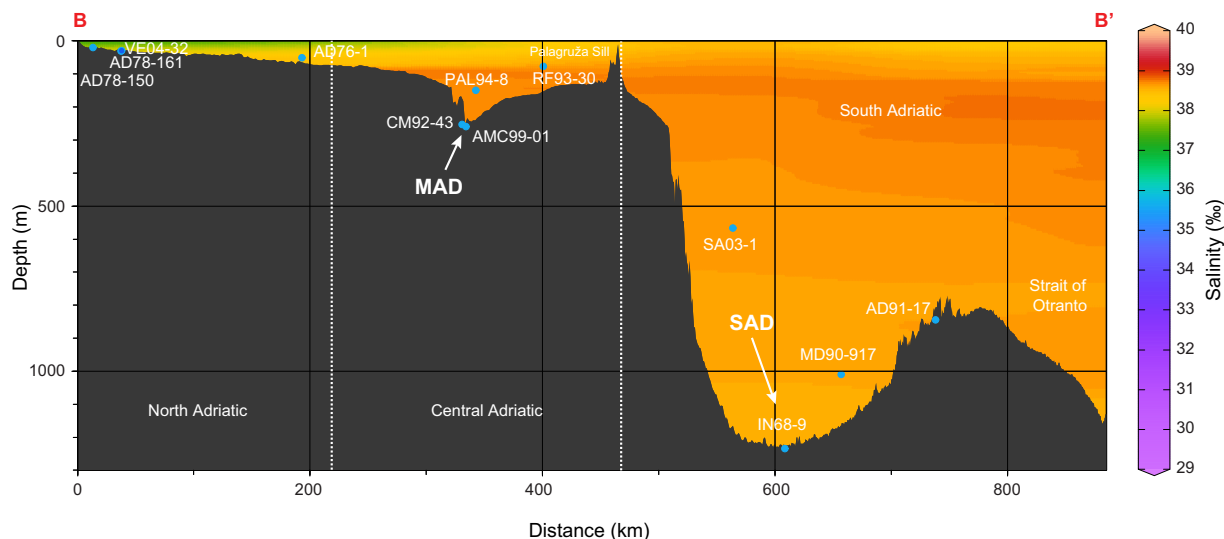


Fig. 4 BB' present day salinity transect running longitudinally across the Adriatic Sea as reported in Fig. 3a (dashed red line). According to bathymetry, the Adriatic Sea is divided into three sub-basins: North, Central and South⁷⁶. Vertical bathymetry profiles derived from the General Bathymetric Chart of the Ocean (GEBCO) 2023 ref.⁷⁵. Salinity data is from SeaDataNet (<https://www.seadatanet.org>) and was retrieved from SeaDataNet infrastructure at the end of July 2019. The water depth for the Adriatic cores is also reported. MAD = Middle Adriatic Depression, SAD = Southern Adriatic Depression.

In the Tyrrhenian Sea, temperature (38%) and salinity (19%) are the dominant variables, with core ET91-18 being the most studied (Fig. 5c). In the Central Mediterranean (Fig. 1b, dashed green line), temperature (29%) and productivity (23%) are the most studied, with ODP Site 963 being the best-monitored core (Fig. 5d).

The Adriatic Sea features 10 variables, with temperature (28%) and dating (20%) being the most documented, and cores like CM92-43 and MD90-917 being key representatives (Fig. 6). In the Northern Adriatic, the number of variables available is lower compared to the southern regions.

In the Eastern basins (Ionian and Levantine Sea), productivity (22%) is a major focus (Fig. 7a), strongly linked to the study of Early Holocene (Greenlandian) sapropel S1, one of the most investigated and debated events in the Eastern Mediterranean over the last 20,000 years. Data from the Aegean Sea include temperature (20%) and dating (24%), with core LC21 providing the highest number of variables among the Aegean cores (Fig. 7b).

Main climate parameters and their spatial distribution across the Mediterranean Sea. The distribution of environmental data across the Mediterranean Sea is generally uniform, but specific variables are more frequently documented in certain regions. For example, bathymetry and sea level data are well-represented in the Adriatic, Aegean, and the Strait of Sicily; carbon dioxide data are concentrated in the Ionian-Levantine Sea; and data on human activity are most prominent in the Adriatic (Fig. 8a).

Water temperature (including both SST and BWT) is the most studied variable in the Mediterranean Sea, as previously suggested⁴², accounting for 26% on the total variables investigated (Fig. 8b). The focus on this variable, both in the past and present, is largely due to its significant impact on marine ecosystems and the projected rise in surface water temperature for the Mediterranean Sea in the near future⁶⁰. This rise is expected to lead to an increase in water mass stratification, alter biogeochemical cycles, affect marine species, and nutrient cycles²⁷.

In addition to temperature, dating of marine sediment cores (17%) represents the fundamental constraint for any paleoclimatic study (Fig. 8a).

Productivity, which in our database includes primary production, export production, and nutrients, is the third most studied parameter in the marine sediments of the Mediterranean Sea (15%; Fig. 8b). In modern paleoclimatology, productivity is primarily reconstructed using fossil fauna and flora, the sedimentary accumulation of biogenic components (such as organic carbon, opal, and biogenic barium), and other geochemical indicators (es. Ba/Al, Ba_{tot}, etc.) (Table 2). Oxygen concentration and organic matter is also frequently reported (13%) followed by precipitation (10%), salinity (7%) and circulation (6%) (Fig. 8a). Bathymetry and sea level, though less frequently studied (3%), hold particular importance in regions with large continental shelves, such as the Aegean Sea⁶¹ and the Adriatic Sea^{40,62-64}. Carbon dioxide, human activity, and diagenesis are among the least studied parameters (Fig. 8b). Over the past 20,000 years, CO₂ levels (ppm) have fluctuated significantly due to natural processes and, more recently, human activities^{65,66}. Carbon dioxide data represents only 0.3%. Marine proxies used to reconstruct past pCO₂ are mainly reported from the EMED and typically include alkenone-based and phytol-based proxies⁶⁷ and sedimentary bulk organic matter⁶⁸ (Table 3). Human activity (1%) and diagenetic processes (2%) are also documented. Human impact data are primarily from the Adriatic Sea, due to its historical record of human-environment interactions⁶⁹, while diagenetic studies are mostly confined to the Ionian/Levantine basins (Fig. 7a). Historically, diagenesis has received less attention in paleoclimatological studies

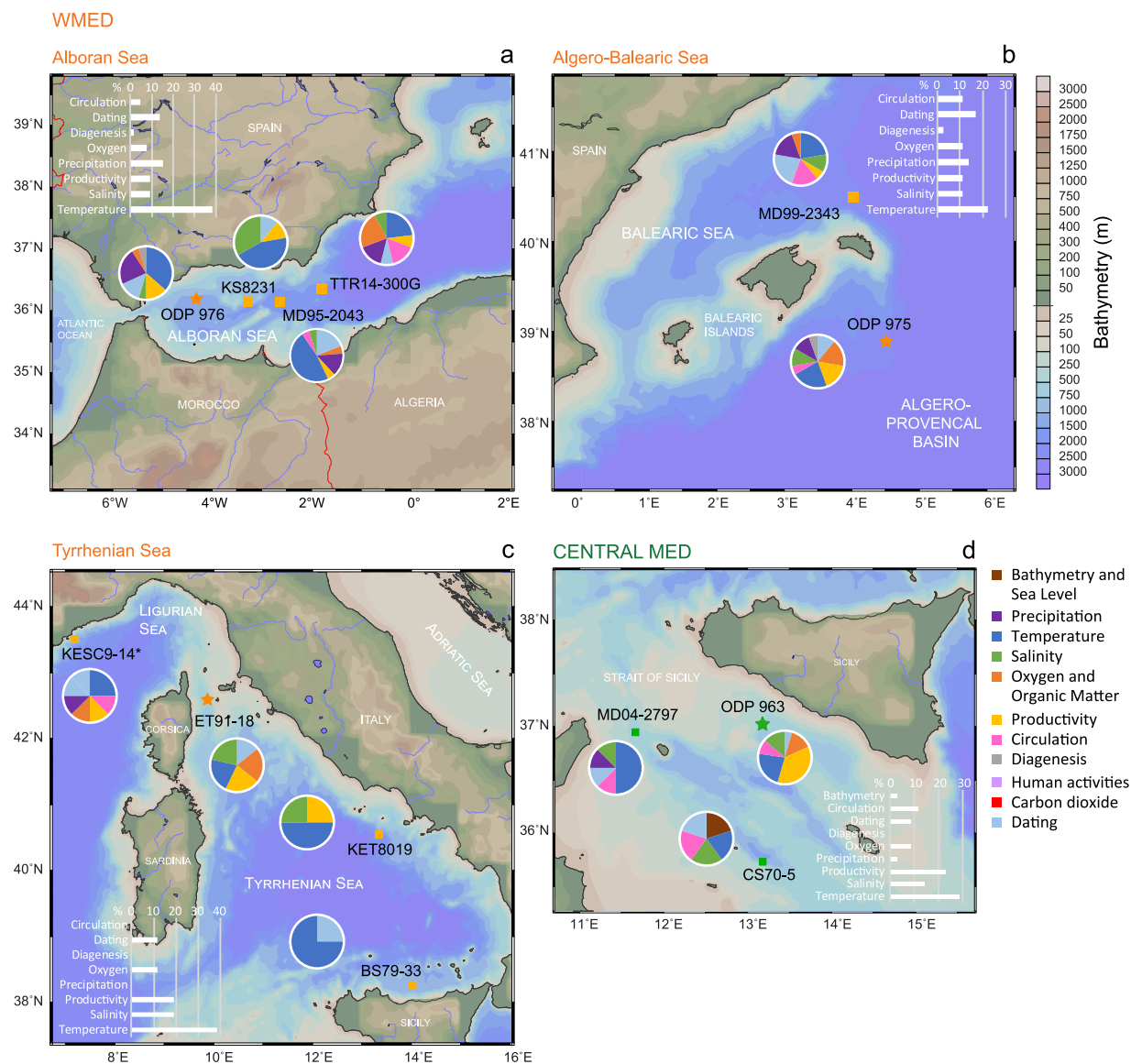


Fig. 5 Location of well-studied (stars) and potentially interesting cores (squares) for the late Pleistocene-Holocene, with focus on the last 20,000 years, for different sub-basins: (a) Alboran Sea, (b) the Algero/Balearic Sea, (c) the Tyrrhenian Sea, (d) the Central Mediterranean (Strait of Sicily). Temperature includes bottom and sea-surface water temperature (BWT and SST, respectively); productivity derives from nutrients and biological productivity; precipitation derives from river runoff, aridity, type of vegetation and ice volumes; circulation derives from water exchanges, deep water flow/bottom current velocity, deep water formation and water mass properties (pH, alkalinity, carbonate saturation state); human activities primarily include deforestation and agriculture. Dating comprises AMS ^{14}C , ^{210}Pb and tephra dating. For a comprehensive list of the variables examined in this study, refer to Table 2, as well as to the “Variables and Proxies” spreadsheet of the PaleoMED20 database¹⁸. MAD = Middle Adriatic Depression; SAD = South Adriatic Depression. *Core KESC9-14 in (c) belongs to the Algero-Balearic Sea even if located near the Ligurian Sea.

compared to other parameters. However, inadequate assessment of diagenetic overprints and/or alterations in the fossil record can introduce significant uncertainties and geochemical biases in climate reconstructions⁷⁰.

The Adriatic Sea has a distinct climatic data distribution compared to the rest of the Mediterranean. While data are uniformly distributed across the Western, Central, and Eastern Mediterranean, the Adriatic Sea exhibits regional variation. The northern Adriatic, with its wide continental shelf (Fig. 4), supports detailed bathymetry and sea-level studies, although fewer other variables are documented. In contrast, the deeper waters of the central and southern Adriatic allow for a broader range of climate proxies, enabling a more comprehensive documentation of climatic parameters.

Distribution of proxies across the mediterranean sea. This exploratory analysis reveals a distinction between abiotic (geochemical) and biotic (paleontological) proxies, providing a valuable starting point for

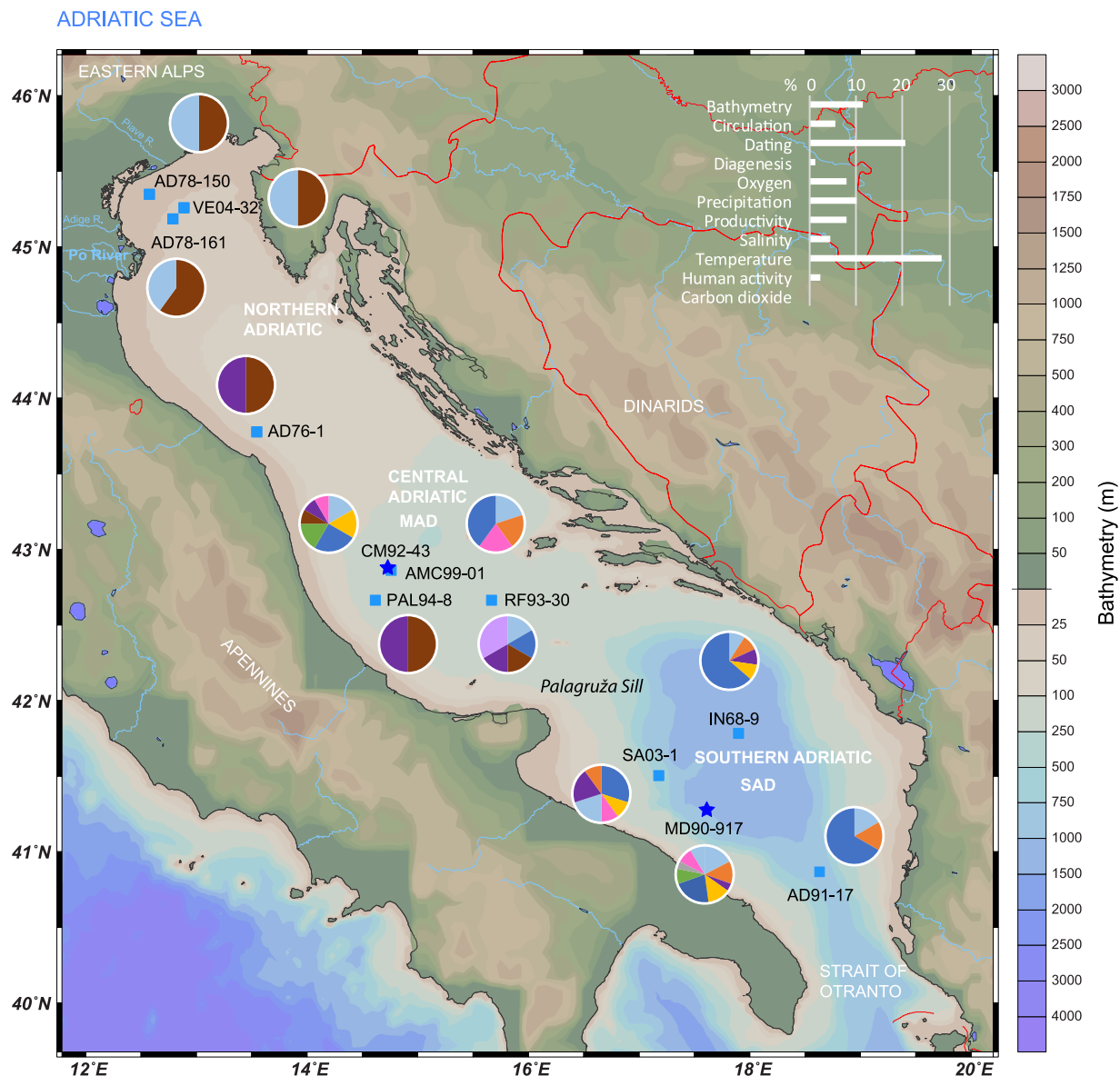


Fig. 6 Location of well-studied (stars) and potentially interesting cores (squares) for the late Pleistocene-Holocene with special focus on the last 20,000 years in the Adriatic Sea. Pie-charts report the main variables investigated for each core. The legend of the pie-charts is reported in Fig. 5.

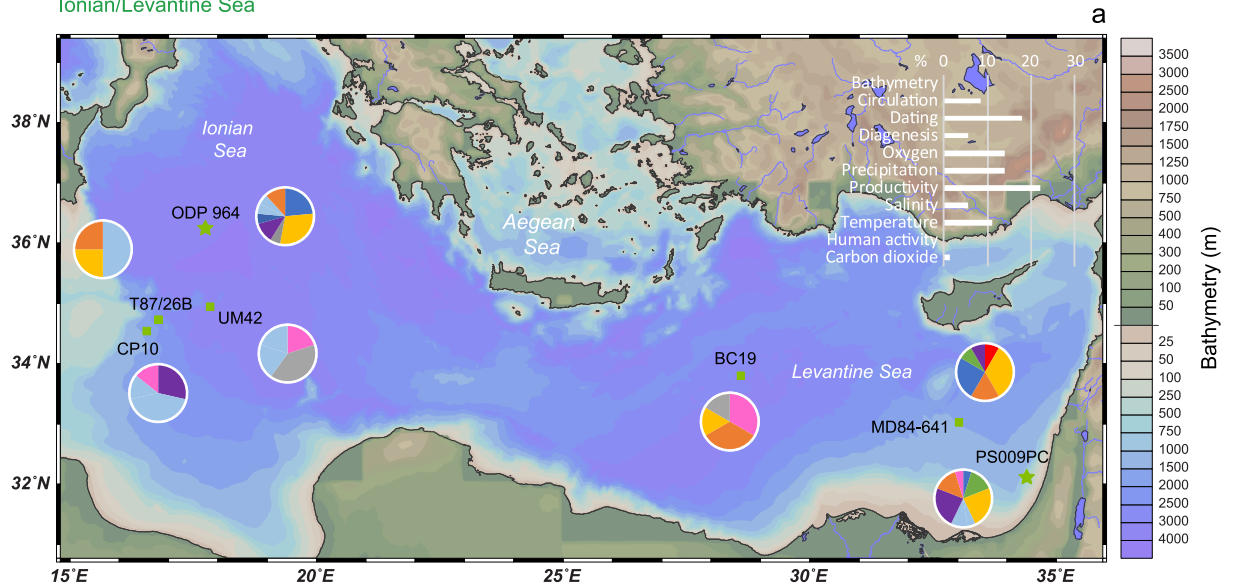
Mediterranean Quaternary paleoclimatic studies. Abiotic proxies are predominant, representing about 76% of the proxies available for the Mediterranean over the past 20,000 years (Fig. 8c). Among abiotic proxies AMS ^{14}C , stable carbon and oxygen isotopes, and elemental data are the most frequently investigated. The remaining 24% are biotic proxies including foraminifera, calcareous nannofossils, pollens, dinoflagellates, diatoms, and other minor faunal groups (pteropods, ostracods, echinoids, molluscs, etc.). Planktonic foraminifera, examined in ~70% of the total cores, is the most extensively studied biotic proxy. Following foraminifera, calcareous nannofossils and pollen are the next most investigated biotic proxies (Fig. 8c).

Machine learning applications for data integration and age model harmonization. The increasing volume and complexity of paleoclimate data presents both challenges and opportunities for research. Artificial intelligence (AI) and machine learning (ML) offer powerful solutions to enhance databases. PaleoMED20¹⁸ is an ideal platform for developing advanced AI/ML applications because its metadata is systematically organized into well-defined categories - an essential feature for training robust machine learning models. Its flexible format (Excel/CSV) also facilitates seamless integration with AI tools and programming environments (e.g., Python, R, cloud-based platforms). Looking ahead, future advancements might include:

- 1) Automated age model harmonization across sediment cores using the alignment of multiple dating methods;

EMED

Ionian/Levantine Sea



Aegean Sea

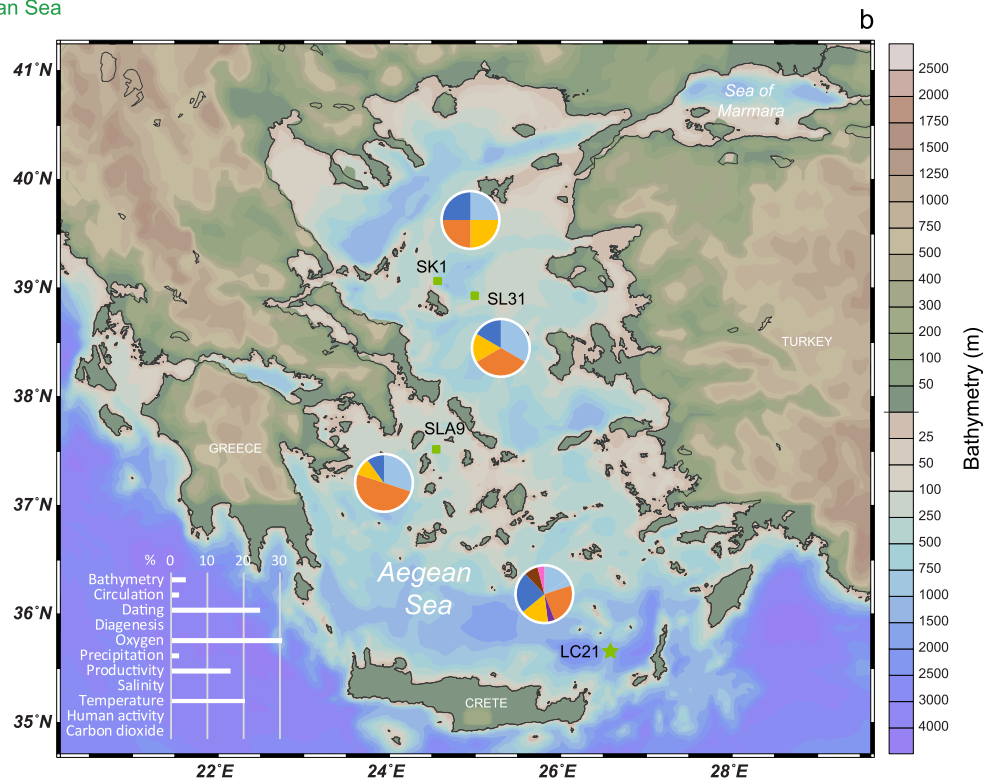


Fig. 7 Location of well-studied (stars) and potentially interesting cores (squares) for the late Pleistocene-Holocene with special focus on the last 20,000 years for the (a) Ionian/Levantine Sea and (b) the Aegean Sea. Pie-charts report the main variables investigated for each core. The legend of the pie-charts is reported in Fig. 5.

- 2) Proxy–environment relationship modeling using advanced ML techniques;
- 3) Automated data ingestion using Natural Language Processing (NLP); The management of such complex and heterogeneous datasets could therefore be supported, at least in part, by the application of ML and AI techniques. For example, in the case of age model harmonization, sequential algorithms such as Dynamic Time Warping or Bayesian modeling frameworks could be used to align stratigraphic signals across sediment cores^{71,72}. Similarly, for modeling proxy–environment relationships, algorithms like Random Forests or neural networks may help to capture nonlinear patterns between proxy data and climatic variables⁷³.

Finally, the ingestion of paleoclimate data from the vast and scattered scientific literature could be facilitated by Natural Language Processing (NLP) methods, such as named entity recognition or document classification⁷⁴.

Usage Notes

The PaleoMED20¹⁸ database offers valuable resources for guiding future research by enabling the exploration of under-documented variables, the introduction of new proxies, and the strategic planning of oceanographic expeditions in less-studied regions of the Mediterranean Sea. It supports researchers interested in investigating understudied proxies and allows for comparative climate analysis across basin scales, enriching the understanding of regional climate patterns. Access instructions and a detailed classification of proxies and variables are provided in the Methods section. The database is available in a user-friendly CSV format from the Figshare repository¹⁸, with no restrictions on re-use.

Code availability

No custom code has been used to produce this database.

Received: 31 January 2025; Accepted: 24 June 2025;

Published online: 01 July 2025

References

- de Rouville, P. Études sur les terrains tertiaires du département de l'Hérault. in *Mémoires de l'Académie des Sciences et Lettres de Montpellier. Section des Sciences, Tome II* 1–84 (1853).
- Flecker, R., Ducassou, E. & Williams, T. Expedition 401 Scientific Prospectus: Mediterranean–Atlantic Gateway Exchange. International Ocean Discovery Program. *Coll. Station. TX Int. Ocean Discov. Program, Coll. Station. TX, United States*. <https://doi.org/10.14379/iodp.sp.401.2023> (2023).
- Zitellini, N., Malinverno, A. & Estes, E. R. Expedition 402 Scientific Prospectus: Tyrrhenian Continent–Ocean Transition. International Ocean Discovery Program. *Int. Ocean Discov. Progr.* <https://doi.org/10.14379/iodp.sp.402.2023> (2023).
- Camuera, J. *et al.* Past 200 kyr hydroclimate variability in the western Mediterranean and its connection to the African Humid Periods. *Sci. Rep.* **12**, 1–13 (2022).
- Santinelli, C. DOC in the Mediterranean Sea. in *Biogeochemistry of Marine Dissolved Organic Matter: Second Edition* 579–608 <https://doi.org/10.1016/B978-0-12-405940-5.00013-3> (Academic Press, 2015).
- Intergovernmental Panel on Climate Change (IPCC). *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, Lee, H. & Romero, J. (eds.)]*. <https://doi.org/10.59327/IPCC/AR6-9789291691647> (2023).
- Bolle, H.-J. *Mediterranean climate: variability and trends*. (Springer Science & Business Media, 2003).
- Montanari, A. Uncertainty assessment for predictions of natural hazard events and impact. in *PNRR progetto RETURN, Dissemination workshop - Multi risk models from assessment to validation and implication for decision making* (2024).
- Moreno, A. *et al.* Links between marine and atmospheric processes oscillating on a millennial time-scale. A multi-proxy study of the last 50,000 yr from the Alboran Sea (Western Mediterranean Sea). *Quat. Sci. Rev.* **24**, 1623–1636 (2005).
- Olausson, E. Studies of deep-sea cores. *Reports Swedish Deep Sea Exped. 1947–1948* **8**, 353–391 (1961).
- Rohling, E. J. & De Rijk, S. Holocene climate optimum and last glacial maximum in the Mediterranean: The marine oxygen isotope record. *Mar. Geol.* **153**, 57–75 (1999).
- Rohling, E. J., Marino, G. & Grant, K. M. Mediterranean climate and oceanography, and the periodic development of anoxic events (sapropels). *Earth-Science Rev.* **143**, 62–97 (2015).
- Henderson, G. M. New oceanic proxies for paleoclimate. *Earth Planet. Sci. Lett.* **203**, 1–13 (2002).
- Camuffo, D. *et al.* 500-Year temperature reconstruction in the Mediterranean Basin by means of documentary data and instrumental observations. *Clim. Change* **101**, 169–199 (2010).
- Luterbacher, J. *et al.* A Review of 2000 Years of Paleoclimatic Evidence in the Mediterranean. *The Climate of the Mediterranean Region: From the Past to the Future* <https://doi.org/10.1016/B978-0-12-416042-2.00002-1> (Elsevier Inc., 2012).
- Martinez-Ruiz, F. *et al.* Paleoclimate and paleoceanography over the past 20,000 yr in the Mediterranean Sea Basins as indicated by sediment elemental proxies. *Quat. Sci. Rev.* **107**, 25–46 (2015).
- Clark, P. U. *et al.* The Last Glacial Maximum. *Science*. **325**, 710–714 (2009).
- Viganò, A. & Agnini, C. PaleoMED20: A Database of Mediterranean Marine Sediment Cores For the Last 20,000 Years. *Figshare* <https://doi.org/10.6084/m9.figshare.28303013> (2025).
- Fondazione Return. Spoke 8 – DS: Science underpinning climate services for risk mitigation and adaptation. Available at: <https://www.fondazionereturn.it/spoke8/>. (Accessed: 28th January 2025) (2025).
- Felden, J. *et al.* PANGAEA - Data Publisher for Earth & Environmental Science. *Sci. Data* **10**, 347 (2023).
- NOAA. Index to Marine and Lacustrine Geological Samples (IMLGS). *National Ocean Service website* Available at: <https://maps.ngdc.noaa.gov/viewers/imlgs/samples>. (Accessed: 2nd August 2024).
- SeaDataNet. Pan-European infrastructure for ocean & marine data management. *SeaDataNet website* Available at: <https://www.seadatanet.org/>. (Accessed: 10th July 2024).
- ISMAR. Spatial database - WebGIS. *ISMAR Spatial database website* Available at: <http://gismargrey.bo.ismar.cnr.it:8080/mokaApp/apps/ismarBoApp/index.html?null>. (Accessed: 7th December 2023).
- American Geosciences Institute. Scientific Ocean Drilling Database. Available at: <http://iodp.americangeosciences.org/vufind/>.
- UNESCO. UNESCO Digital Library. Available at: <https://unesdoc.unesco.org/home>.
- Kullenberg, B. On the salinity of the water contained in marine sediments. *Medd. Ocean. Inst. Göteborg*. **21**, 1–38 (1952).
- SNPA. Rapporto sugli indicatori di impatto dei cambiamenti climatici – Edizione 2021. *Report SNPA 21/2021* (2021).
- Rhein, M. *et al.* Observations: Ocean. in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (ed. Stocker, T. F. *et al.*) 255–315 (Cambridge University Press, 2013).
- Deser, C., Alexander, M. A., Xie, S. P. & Phillips, A. S. Sea surface temperature variability: Patterns and mechanisms. *Ann. Rev. Mar. Sci.* **2**, 115–143 (2010).
- Bradley, R. S. *Paleoclimatology: Reconstructing Climates of the Quaternary*. (Academic Press, 1999).
- Adolphi, F. *et al.* Radiocarbon calibration uncertainties during the last deglaciation: Insights from new floating tree-ring chronologies. *Quat. Sci. Rev.* **170**, 98–108 (2017).
- Hajdas, I. *et al.* Radiocarbon dating. *Nat. Rev. Methods Prim.* **1**, 1–26 (2021).

33. Nittrouer, C. A., Sternberg, R. W., Carpenter, R. & Bennett, J. T. The use of Pb-210 geochronology as a sedimentological tool: Application to the Washington continental shelf. *Mar. Geol.* **31**, 297–316 (1979).
34. Lowe, D. J. Tephrochronology and its application: A review. *Quat. Geochronol.* **6**, 107–153 (2011).
35. Bennett, C. L. *et al.* Radiocarbon Dating Using Electrostatic Accelerators: Negative Ions Provide the Key. *Science*. **198**, 508–510 (1977).
36. Törnqvist, T. E., Rosenheim, B. E., Hu, P. & Fernandez, A. B. Radiocarbon dating and calibration. in *Handbook of Sea-Level Research* (eds. Shennan, I., Long, A. J. & Horton, B. P.) 347–360 <https://doi.org/10.1002/9781118452547.ch23> (John Wiley & Sons, Ltd, 2015).
37. Engel, A., Kiko, R. & Dengler, M. Organic Matter Supply and Utilization in Oxygen Minimum Zones. *Ann. Rev. Mar. Sci.* **14**, 355–378 (2022).
38. Durack, P. J., Wijffels, S. E. & Boyer, T. P. Long-term Salinity Changes and Implications for the Global Water Cycle. *Int. Geophys.* **103**, 727–757 (2013).
39. EEA. Extreme sea levels and coastal flooding in Europe. Available at: <https://www.eea.europa.eu/en/analysis/indicators/extreme-sea-levels-and-coastal-flooding> (2024).
40. Trincardi, F., Argnani, A. & Correggiari, A. Note illustrative della Carta Geologica d'Italia alla scala 1:250,000 – Foglio NL33-7 “Venezia”. *ISPRA - Servizio Geologico d'Italia*. (2011).
41. Hesse, R. & Schacht, U. Early diagenesis of deep-sea sediments. *Dev. Sedimentol.* **63**, 557–713 (2011).
42. Wefer, G., Berger, W. H., Bijma, J. & Fischer, G. Clues to Ocean History: a Brief Overview of Proxies. in *Use of Proxies in Paleoclimatology: Examples from the South Atlantic* (eds. Fischer, G. & Wefer, G.) 1–68, https://doi.org/10.1007/978-3-642-58646-0_1 (Springer-Verlag Berlin Heidelberg, 1999).
43. Intergovernmental Panel on Climate Change. *Climate change and land: an IPCC Special Report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*. <https://doi.org/10.1017/9781009157988> (Cambridge University Press, 2022).
44. Masson-Delmotte, V. *et al.* *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. <https://doi.org/10.1017/9781009157896> (2021).
45. Narcisi, B. & Vezzoli, L. Quaternary stratigraphy of distal tephra layers in the Mediterranean—an overview. *Glob. Planet. Change* **21**, 31–50 (1999).
46. Zanchetta, G. *et al.* Tephrostratigraphy, chronology and climatic events of the Mediterranean basin during the Holocene: An overview. *Holocene* **21**, 33–52 (2011).
47. Murat, A. & Got, H. Middle and Late Quaternary depositional sequences and cycles in the eastern Mediterranean. *Sedimentology* **34**, 885–899 (1987).
48. Rohling, E. J. & Hilgen, F. J. The eastern Mediterranean climate at times of sapropel formation: a review. *Geol. en Mijnb.* **70**, 253–264 (1991).
49. Zonneveld, K. A. F. *et al.* Selective preservation of organic matter in marine environments; Processes and impact on the sedimentary record. *Biogeosciences* **7**, 483–511 (2010).
50. Paytan, A. & Griffith, E. M. Marine barite: Recorder of variations in ocean export productivity. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **54**, 687–705 (2007).
51. Sigman, O. M., Altabet, M. A., McCorkle, D. C., Francois, R. & Fischer, G. The $\delta^{15}\text{N}$ of nitrate in the southern ocean: Consumption of nitrate in surface waters. *Global Biogeochem. Cycles* **13**, 1149–1166 (1999).
52. Kaiho, K. Benthic foraminiferal dissolved-oxygen index and dissolved-oxygen levels in the modern ocean. *Geology* **22**, 719–722 (1994).
53. Hoogakker, B. A. A. *et al.* Reviews and syntheses: Review of proxies for low-oxygen paleoceanographic reconstructions. *Biogeosciences* **22**, 863–957 (2025).
54. Krom, M. D., Michard, A., Cliff, R. A. & Strohle, K. Sources of sediment to the Ionian Sea and western Levantine basin of the Eastern Mediterranean during S-1 sapropel times. *Mar. Geol.* **160**, 45–61 (1999).
55. Chevalier, M. *et al.* Pollen-based climate reconstruction techniques for late Quaternary studies. *Earth-Science Rev.* **210**, 103384 (2020).
56. Kaniewski, D. *et al.* Holocene palaeoecological archives of Eastern Mediterranean plant diversity: Past, present and future trends. *Anthropocene* **45**, 100430 (2024).
57. Schlitzer, R. *Ocean Data View*. (2022).
58. The International Hydrographic Organization (IHO). Mediterranean Sea - IHO Sea Area. *Limits of oceans and seas*. Available at: <http://marineregions.org/mrgid/1905> (1953).
59. Batini, C. & Scannapieco, M. *Data and information quality*. <https://doi.org/10.1007/978-3-319-24106-7> (Springer International Publishing, 2016).
60. EEA. *European sea surface temperature*. Available at: <https://www.eea.europa.eu/en/analysis/indicators/european-sea-surface-temperature?activeAccordion=309c5ef9-de09-4759-bc02-802370dfa366>. (Accessed: 25th June 2024) (2024).
61. Lambeck, K. Late Pleistocene and Holocene sea-level change in Greece and south-western Turkey: a separation of eustatic, isostatic and tectonic contributions. *Geophys. J. Int.* **122**, 1022–1044 (1995).
62. Ronchi, L., Fontana, A., Correggiari, A. & Asioli, A. Late Quaternary incised and infilled landforms in the shelf of the northern Adriatic Sea (Italy). *Mar. Geol.* **405**, 47–67 (2018).
63. Trincardi, F., Cattaneo, A., Asioli, A., Correggiari, A. & Langone, L. Stratigraphy of the late - Quaternary deposits in the central Adriatic basin and the record of short - term climatic events. *Memorie-Istituto Ital. di Idrobiol.* **55**, 39–70 (1996).
64. Storms, J. E. A., Weltje, G. J., Terra, G. J., Cattaneo, A. & Trincardi, F. Coastal dynamics under conditions of rapid sea-level rise: Late Pleistocene to Early Holocene evolution of barrier-lagoon systems on the northern Adriatic shelf (Italy). *Quat. Sci. Rev.* **27**, 1107–1123 (2008).
65. Marcott, S. A. *et al.* Centennial-scale changes in the global carbon cycle during the last deglaciation. *Nature* **514**, 616–619 (2014).
66. Monnin, E. *et al.* Atmospheric CO₂ concentrations over the Last Glacial Termination. *Science*. **291**, 112–114 (2001).
67. Witkowski, C. R., van der Meer, M. T. J., Blais, B., Sinnighe Damsté, J. S. & Schouten, S. Algal biomarkers as a proxy for pCO₂: Constraints from late quaternary sapropels in the eastern Mediterranean. *Org. Geochem.* **150**, 104123 (2020).
68. Fontugne, M. R. & Calvert, S. E. Late Pleistocene Variability of the Carbon Isotopic Composition of Organic Matter in the Eastern Mediterranean: Monitor of Changes in Carbon Sources and Atmospheric CO₂ Concentrations. *Paleoceanography* **7**, 1–20 (1992).
69. Ruiz-Redondo, A. *et al.* Mid and Late Upper Palaeolithic in the Adriatic Basin: Chronology, transitions and human adaptations to a changing landscape. *Quat. Sci. Rev.* **276**, 107319 (2022).
70. Viganò, A., Westerhold, T., Bown, P. R., Dunkley Jones, T. & Agnini, C. Calcareous nannofossils across the Eocene-Oligocene transition: Preservation signals and biostratigraphic remarks from ODP Site 1209 (NW Pacific, Shatsky Rise) and IODP Hole U1411B (NW Atlantic Ocean, Newfoundland Ridge). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **629**, 111778 (2023).
71. Lin, L., Khider, D., Lisiecki, L. E. & Lawrence, C. E. Probabilistic sequence alignment of stratigraphic records. *Paleoceanography* **29**, 976–989 (2014).
72. Birlo, S., Tylmann, W. & Zolitschka, B. Bayesian age-depth modelling applied to varve and radiometric dating to optimize the transfer of an existing high-resolution chronology to a new composite sediment profile from Holzmaar (West Eifel Volcanic Field, Germany). *Geochronology* **5**, 65–90 (2023).
73. Sobol, M. K. & Finkelstein, S. A. Predictive pollen-based biome modeling using machine learning. *PLoS One* **13**, e0202214 (2018).

74. Zhou, S., Peng, S. & Plank, B. CLIMATELI: Evaluating Entity Linking on Climate Change Data. In *Proceedings of the 1st Workshop on Natural Language Processing Meets Climate Change (ClimateNLP 2024)* (ed. Stambach, D. *et al.*) 215–222, <https://doi.org/10.48550/arXiv.2406.16732> (Association for Computational Linguistics, 2024).
75. GEBCO Compilation Group. *GEBCO 2023 Grid*. <https://doi.org/10.5285/f98b053b-0cbc-6c23-e053-6c86abc0af7b> (2023).
76. Vrdoljak, L., Režić, M. & Petričević, I. Bathymetric and geological properties of the Adriatic sea. *Rud. Zb.* **36**, 93–107 (2021).

Acknowledgements

This research was financially supported by University of Padova, PRIN (Prot. 2022T4XEBP) and a RETURN Extended Partnership, financed by the National Recovery and Resilience Plan - NRRP, Mission 4, Component 2, Investment 1.3–D.D. 1243 2/8/2022, PE0000005.

Author contributions

A.V. and C.A. designed the study, developed the interpretative framework of the work and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-025-05451-5>.

Correspondence and requests for materials should be addressed to A.V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025