Laser ablation mass spectrometry blast through detection in R

Alex Searle-Barnes[1], Andy Milton[1], Christopher Standish[1], Gavin Foster[1], and Thomas Ezard[1]

[1]University of Southampton

March 11, 2023

Abstract

Rational: Organisms that grow a hard carbonate shell or skeleton, such as foraminifera, corals or molluscs, incorporate trace elements into their shell during growth that absorbs the environmental change and biological activity they experienced. These geochemical signals locked within the carbonate are archives used in proxy reconstructions to study past environments and climates, to decipher taxonomy of cryptic species and to resolve evolutionary responses to climatic changes. Methods: Here we use a laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) as a time resolved acquisition to quantify the elemental composition of carbonate shells. We present the LABLASTER (Laser Ablation BLASt Through Endpoint in R) package, which imports a single time resolved LA-ICP-MS analysis, then detects when the laser has ablated through the carbonate as a function of change in signal over time, and outputs key summary statistics. We provide two worked examples within the package: a planktic foraminifera and a tropical coral. Results: We present the first R package that improves signal: noise ratios in data reduction workflows by automating the detection of when the laser has ablated through a sample using a smoothed time-series and subsequent removal of off-target data points. The functions are flexible and adjust dynamically to enhance the signal: noise ratio of the desired geochemical target. Visualisation tools for manual validation are also included. Conclusions: LABLASTER increases transparency and repeatability by algorithmically identifying when the laser has either ablated fully through a sample or across a mineral boundary and is thus no longer documenting a geochemical signal associated with the desired sample. LABLASTER's focus on better data targeting means more accurate extraction of biological and geochemical signals.

Introduction

Laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) is a powerful analytical tool to quantify the elemental composition of a wide variety of natural and anthropogenic materials. A laser beam is focussed to the surface of a target and then pulsed to ablate the sample. Particles from the ablated sample are subsequently transported into a mass spectrometer for detection based on the mass-to-charge ratio, which can be converted into a resolved isotopic or elemental profile. LA–ICP-MS has become increasingly popular with biogenic carbonates including foraminifer[1-3], coral skeletons[4] and molluscs[5], all of which act as archives of geochemical signals that can be used as proxy measurements to both reconstruct past environments and study the evolutionary response to long-term climate change.

Recent instrumentation advances enable LA-ICP-MS setups to collect comparable trace element to calcium ratio (TE/Ca) results to traditional solution based ICP-MS but with simpler sample preparation and higher throughput[6]. This solid sample and laser setup also allows for higher spatial resolution of the sample and avoids the heterogeneity averaging that occurs in solution based ICP-MS. Nevertheless, there is no package for R that combines high data throughput with the additional nuance that laser ablation data processing

requires to keep the maximum amount of relevant data. To fully leverage the gains of LA-ICP-MS, any software must be flexible enough to handle non-homogeneous samples.

Some ready-to-use free computer packages exist to process LA-ICP-MS data such as elementR[7] or the discontinued LAICPMS[8], both of which use the R environment, and LATools[9], which uses the Python environment. ElementR provides a point-and-click graphical user interface that slows data reduction throughput, while giving the user fine control over the data integration period. TERMITE[10] is not a package *per se* but is optimised for repeatable data reduction of homogenous samples, where the data integration period must be adjusted individually for each measurement and therefore requiring manual validation.

These three pieces of software provide a general end-to-end workflow to process experimental data into results rather than specialising on a particular data reduction step. In comparison, the LABLASTER package presented here contains a function that specialises in identifying when the laser is no longer recording the geochemical target of interest and is therefore designed for high-throughput processing that doesn't require user interaction once configured. A variety of integration time-range endpoint detection mechanisms are used in the literature, including k-means clustering[7], fixed time stamps[2], analyte signal below a given threshold[11], the mid-point between high and background signal counts[9] and even manual identification when the complexity of the samples is too great[9]. Here we fit a function over a first derivative to calculate the change in rate of signal change. As LA-ICP-MS increases in popularity and experiments become more complex, there is a need for repeatable algorithmic protocols that can deal with heterogeneous samples or where repeat measurements may have different integration times.

Each discipline using LA-ICP-MS measure samples that have different matrixes and properties e.g., polished rock sections, powered pellets or carbonate shells. The worked examples presented here have been tailored to the field of ecology and evolution with a planktic foraminifer and the field of paleoclimate geochemistry with a tropical coral. The LABLASTER package will however work with any sample that the laser may ablate through and hit an undesired target. The foraminifera example here demonstrates how LABLASTER can be used with the specific needs of ecologists, whose data is often skewed and highly variable[7].

Here, we (1) improve current processing capabilities by dynamically identifying the end of the sample of carbonate subject and (2) implement this improved processing in the first freely available software to automate data extraction of a time resolved elemental depth profile. As demonstrated in the examples below, the end of the sample may be the maximum depth at a single spot location for a shell or a boundary between two minerals along a linear profile for a coral, but any non-homogeneous target sample is generally applicable.

An automated laser ablation setup often requires a constant firing time to be programmed into the controlling computer, with no regard for the heterogeneity or variation in thickness of the target. When samples are porous, have changes in mineralogy or variation in thicknesses within a single analytical session while using a consistent laser pulsing time, there is inevitably a chance that the laser will move across a mineral boundary or ablate through the entire depth, and thus the recorded data will not be restricted to only the area of interest. Any elemental measurement recorded after the laser has ablated through the sample is not of the target, it should be removed before subsequent statistical analysis. Because the time taken to ablate through a sample is not consistent, such corrections can be made manually on an ad hoc basic, but additional manual handling would be time-consuming, laborious and prone to subjective differences amongst operators. There are clear methodological benefits from the development of a repeatable workflow.

The LABLASTER package works alongside elementR or TERMITE application or can be run as a standalone process within private scripts, providing a flexible and versatile methodological improvement for heterogeneous samples that treats each sample individually to optimise signal: noise ratios. LABLASTER can batch process within a workflow, is customisable to the sensitivity for endpoint detection and does not require a point-and-click user interface. These features offer a higher throughput for data reduction compared to manual or alternative software methods and in retaining the maximum amount of on-target data for subsequent analysis.

Figure 1: LA-ICP-MS holes from each shot are visible.

Materials and Methods

The LABLASTER package contains one function named endPoint that calculates four items, and four example data sets to illustrate its use, including the foraminifer and coral examples presented here. The package requires R ([?] 3.5.0)[12] and has a number of dependencies as it calls functionality from the stats[12], smooth[13], ggplot2, magrittr[14] and scales[15] packages.

endPoint detection

The main function of the LABLASTER package is to detect when either the laser has ablated through a target (e.g., a carbonate shell or coral skeleton) or across a boundary in a transect. In the following sections, we illustrate this behaviour on a planktic foraminifer and a tropical coral as case studies.

Identifying the time range in a time resolved acquisition when the laser is ablating the target is essential for accessing and correlating the relevant data within the recorded time series. The LABLASTER package assumes the data frame supplied begins with the laser in focus of the desired target and the endPoint function determines the time stamp when the laser has ablated through the sample or across a boundary where the isotope signal changes rapidly. Keeping only the data between the start time and end time focusses subsequent analysis on only relevant target data.

Algorithm

The endPoint function requires a single data frame containing a minimum of two columns: a column containing a time index and a column containing isotope signal counts. This data frame is supplied to the function as the argument **detectDf** with the time index column specified as**timeCol** and the isotope signal column specified as**signalCol** . Any element that is abundant within the target but is scarce in the surrounding medium could be used to detect the ablate through endpoint time. In a case study below, we use "Ca44" containing the ^{44}Ca isotope counts per second as the signalCol as this is abundant within the foraminifera test but is not present in high quantities in the glass slide used for mounting.

The algorithm of the endPoint function first calculates a simple moving average over the isotope counts, with the number of points in the moving average controlled by the **smoothing** argument. This smoothed isotope count signal is then scaled between $0 - 1$ to convert to relative changes that works across all time resolved acquisition data sources. A larger amount of smoothing reduces the variance between adjacent data points, averaging out any signal spikes and therefore flattening the signal against time curve. A flatter time resolved profile curve gives greater distinction between the elevated signal of desired target ablation and the lower signal of the undesired under- or adjacent- target surface.

Using smoothing to de-spike and flatten the signal during the time of elevated signal reduces the magnitude of variance over time and therefore reduces the likelihood of a false positive detection in a change of ablated material composition, which is used as part of the detection in later steps of the algorithm. Over-smoothing the signal reduces the distinction from a sudden signal drop to a shallower gradual drop between the higher on-target and the lower off-target signal intensities that the algorithm uses to detect that the laser is no longer ablating the desired target causing a delayed detection. Over-smoothing can be identified manually using the visualisation tools by comparing the black and blue lines.

Next the algorithm calculates the number of data points per time step in the data frame supplied. The temporal resolution of the data can affect how sensitive the algorithm is to the rate of scaled signal change. The algorithm uses a moving window to calculate the rate of change in signal against time. Higher temporal resolution data can result in a smoother decline in isotope signal across the default time window causing a delayed endpoint detection, therefore using a wider time window captures a larger magnitude signal drop.

The algorithm uses the largest magnitude of negative rate of isotope signal change to identify that the laser is no longer on the desired target. Without a rapid signal drop, for example if the laser did not fully ablate through target or no mineral boundary was crossed then the algorithm will return the final observation in the provided data frame and a warning message is displayed to encourage use of the manual validation tools to check the results. The number of data points used as the width of the moving window is controlled by the **dt** argument.

Once the algorithm has identified the largest magnitude of negative rate of signal change, the corresponding time stamp is identified. As the largest signal drop occurs shortly after the laser has ablated through the sample or crossed a boundary, it is necessary to also remove the data that occurs between the final data point when the laser was on target and the largest signal drop. This elapsed time is calculated by dividing the moving window width by the number of data points per time step and is subtracted from the signal change time stamp. This earliest time stamp is when the laser was last ablating on the desired target and is returned as a numeric value in the returned data frame.

library(lablaster)

endPoint(detectDf, timeCol = "Time", signalCol = "Ca44", smoothing = 5, dt = 10, profile = "FALSE", timeUnits = "seconds")

Visualisation of the blast through algorithm

The endPoint function additionally provides visualisation tools into the mechanism of the algorithm. The smoothed scaled signal, rate of change and inferred bounds on the geochemical target are useful for diagnosing the real-life efficacy. A ggplot2 object is generated when the **profile** argument is set to TRUE. If the profile argument is TRUE, then the time units of the analysis are also required and specified with the **timeUnits** argument.

Returned values

The endPoint function returns a single data frame containing the values calculated.

$startTime contains the earliest timestep in the supplied time resolved acquisition as a numerical value.

$endTime contains the last timestep before the laser ablated through the carbonate shell as described above.

$df contains the supplied data frame with the same structure but containing only the rows that occur between the startTime and endTime.

$profile contains a visualisation of the endpoint mechanism as described above. This is only available if a profile was generated using profile = "TRUE".

Case study

A foraminifera setup

Experimental setup

A single planktonic foraminifer case study is included in the package as a worked example. Planktonic foraminifera are unicellular zooplankton distributed throughout the world's oceans as a key resource in understanding Earth's climate system[1, 2] and the evolution of biodiversity[16]. The thickness of a fossilised foraminifera test can vary substantially even within an individual. Some have thick solid chamber walls while others have a highly porous structure as a result of species specificities, biological controls and environmental influences. This natural variability is the motivation for our development of automated processing methods

and thus better control of geochemical "vital effects" (Kearns et al, submitted to *Paleoceanography and Paleoclimatology*). Foraminifera grow connected chambers throughout their life with those in the final whorl accessible for LA-ICP-MS analysis[17] (Figure 1). The case study is from the antepenultimate chamber of *Menardella exilis* foraminifera 72, identified hereon as "Foram-72-shot-3".

The experimental setup is described fully in Kearns et al. (in review). Briefly major and trace elements in the foraminifera test were analysed using a New Wave UP193 laser ablation system (ArF source, 30 μm spot diameter, a fluence of 0.73 J/cm³ and 5 Hz pulse rate) coupled to an Agilent 8900 triple quadrupole inductively coupled plasma (ICP-QQQ) mass spectrometer in single quadrupole mode using a He and Ar gas mixture (900 ml/min) at the University of Southampton. Each laser spot pattern was sequenced with a 30 s warmup, 50 s laser pulse and 30 s washout. The default values within this endPoint function are based on outputs from this setup.

Initial processing of the time resolved acquisition data was performed within the R environment (version 4.2.2). The first 30 seconds were used to calculate the background signal and then removed. The next five seconds were removed as the isotopic signal begins to rise for the laser has started pulsing but the ablated material is still travelling through the system piping. As our washout time was enough to purge the system after each analysis, the duration of this signal rise was regular and reproducible.

endPoint detection function

The data frame containing the background corrected remainder of the acquisition (Figure 2b) was supplied to the endPoint function. With the first 35 seconds removed, we ensured that the laser was pulsing on the foraminifera's outer test wall and the ablated material was reaching the mass detector for the first row in the data frame supplied. In this example, we use the ^{44}Ca isotope measurements due to the high signal: noise ratio and the abundance of ^{44}Ca within the foraminifera test, which provides greater distinction between higher signal counts for the duration when the laser was on target and lower signal counts when the laser was ablating the glass slide mount. The endPoint function was used to dynamically identify the time step when the laser had fully ablated through the foraminifera test and to keep only the observations while the laser was ablating the target (Figure 2c).

endPoint(detectDf = foram72shot3, dt = 10, smoothing = 5, timeCol = "Time", signalCol = "Ca44", profile = "TRUE", timeUnits = "seconds")
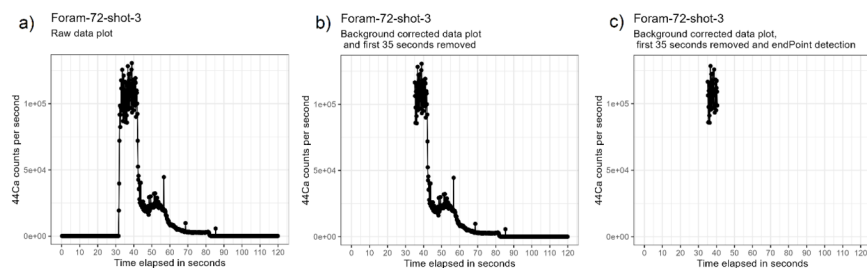

Figure 2: The processes implemented in the endPoint function in a time resolved acquisition of Foram-72-shot-3 as a visualisation of the returned data frame $df. Panel (a) visualises the entire raw ^{44}Ca data collected; panel (b) visualises the background corrected^{44}Ca with the first 35 seconds removed as this was before the laser was turned on and the ablated material is still travelling through the system and this is the data frame passed into the endPoint function; plot (c) visualises the target data retained after running the endPoint function, with both the first 35 seconds and post endPoint detection ^{44}Ca data removed.

Figure 3 shows the manual validation tools to check that the endPoint function had successfully identified when the laser had burnt through the entire depth of the test wall. The black line is a scaled transformation of the smoothed signal specified in signalCol; the blue line is the scaled rate of change in signal over the

5

moving time window and the shaded red areas identify the data points removed from the returned data frame, $df, as these exceed the end point blast through detection time stamp.
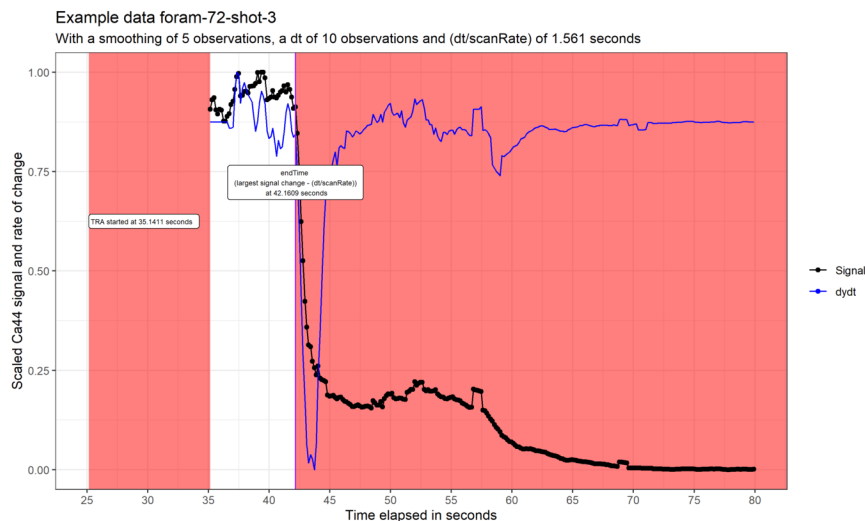

Example data foram-72-shot-3
With a smoothing of 5 observations, a dt of 10 observations and (dt/scanRate) of 1.561 seconds

Figure 3: A visualisation of the endPoint function for the Foram-72-shot-3 case study, showing the scaled smoothed [44]Ca signal in black, the scaled rate of signal change in blue and the shaded red areas identify the data point rows that are removed in the returned data frame.

Comparison without using endPoint detection

Foraminifera are frequently used as archives of geochemical signals that are used in proxy measurement reconstructions. Ecological evolutionary responses can be tracked and sea surface temperatures can be reconstructed with calibrated equations using a ratio of magnesium to calcium as an input, making this Mg/Ca (mmol/mol) a popular geochemical measurement[1, 6].

When processed without the endPoint function and the integration time was between 35 – 80 seconds elapsed, the median [24]Mg/[44]Ca ratio of Foram-72-shot-3 was 4.37±2.10 mmol/mol and when processed with the endPoint function with an integration time of 6 seconds was 2.46±1.35 mmol/mol. Use of the endPoint function thus increases accuracy (by focussing on the geochemical target) and also increases precision through higher quality sample data for statistical analysis.

A coral example

Experimental setup

A transect across a sample of tropical coral *Acropora polystoma* , cultured in artificial seawater in a closed coral mesocosm at the Coral Reef Laboratory, National Oceanography Centre, Southampton[18] is also included in the package as an example of an ablation line and is identified hereon as "coral-6". As polished coral sections can be irregular in shape, thickness and porosity, it is possible for the laser to ablate fully through at a thinner location or traverse a boundary between the sample and its mounting resin. In this example, the ablation line passes over a thinner section of coral and consequently the laser fully ablates through before the end of the analysis time.

endPoint detection function

This coral section has bands of higher and lower calcium concentration across the laser transect and consequently fluctuating higher and lower counts for the [43]Ca signal are observed. In this example, we select

the dt and smoothing parameters to de-spike these real signal fluctuations and using the included manual visualisation tool (Figure 4) to validate our selection was detecting the moment when the laser fully ablated through the coral.

endPoint(detectDf = coral6, dt = 10, smoothing = 5, timeCol = "Time", signalCol = "Ca43", profile = "TRUE", timeUnits = "seconds")
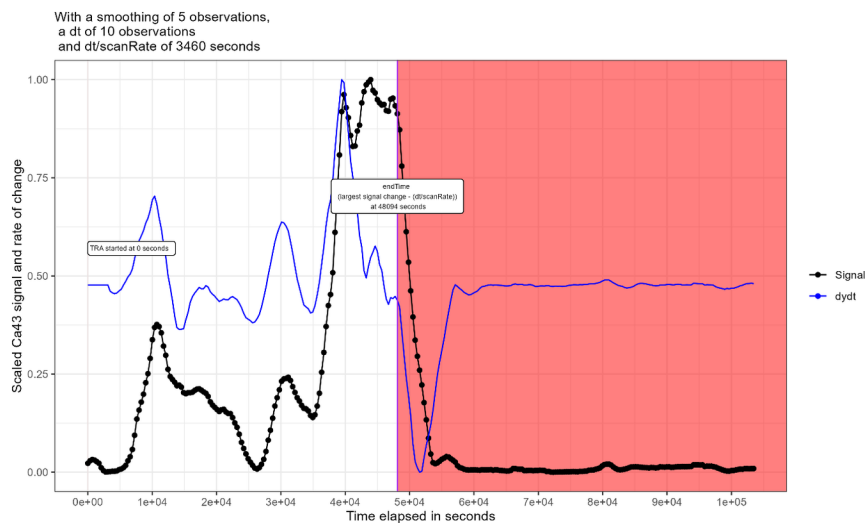

Figure 4: A visualisation of the endPoint function for the Coral-6 case study, showing the scaled smoothed ^{43}Ca signal in black, the scaled rate of signal change in blue and the shaded red areas identify the data point rows that are removed in the returned data frame.

Conclusion

While LA-ICP-MS has many benefits over traditional solution based ICP-MS methods for measuring major and trace elements in various carbonate objects, current analytical setups can be inflexible with regards to laser firing duration. Our freely available LABLASTER package implements methodological improvements to refine analytical workflows.

The endPoint function that implements these improvements is compatible for use within loops, aiding high-throughput and repeatable data cleaning of carbonate materials that are ubiquitously used in past climate reconstructions, geochemical ecology and evolution studies and cryptic taxonomic distinctions.

Acknowledgements

Conflict of interest statement

The authors do not declare a conflict of interest.

Author contributions

ASB designed the functions and wrote the code. JAM and GLF provided methodological support and laboratory analysis. CDS provided the coral data for additional algorithm testing and example data published within the LABLASTER package. THGE validated the functions and code. ASB drafted the manuscript and JAM, GLF, CDS and THGE provided comments on earlier drafts.

References

1. Eggins, S., De Deckker, P. & Marshall, J. (2003) Mg/Ca variation in planktonic foraminifera tests: implications for reconstructing palaeo-seawater temperature and habitat migration. *Earth and Planetary Science Letters,* **212,** 291-306, doi:10.1016/S0012-821X(03)00283-8.

2. Hathorne, E. C., Alard, O., James, R. H. & Rogers, N. W. (2003) Determination of intratest variability of trace elements in foraminifera by laser ablation inductively coupled plasma-mass spectrometry. *Geochemistry, Geophysics, Geosystems,* **4** , doi:10.1029/2003GC000539.

3. Fehrenbacher, J. S., Spero, H. J., Russell, A. D., Vetter, L. & Eggins, S. (2015) Optimizing LA-ICP-MS analytical procedures for elemental depth profiling of foraminifera shells. *Chemical Geology,* **407-408,** 2-9, doi:10.1016/j.chemgeo.2015.04.007.

4. Chalk, T. B., Standish, C. D., D'Angelo, C., Castillo, K. D., Milton, J. A. & Foster, G. L. (2021) Mapping coral calcification strategies from in situ boron isotope and trace element measurements of the tropical coral Siderastrea siderea. *Scientific Reports,* **11,** 472, doi:10.1038/s41598-020-78778-1.

5. Schöne, B. R., Zhang, Z., Radermacher, P., Thébault, J., Jacob, D. E., Nunn, E. V. & Maurer, A.-F. (2011) Sr/Ca and Mg/Ca ratios of ontogenetically old, long-lived bivalve shells (Arctica islandica) and their function as paleotemperature proxies. *Palaeogeography, Palaeoclimatology, Palaeoecology,* **302,** 52-64, doi:10.1016/j.palaeo.2010.03.016.

6. Fehrenbacher, J., Marchitto, T. & Spero, H. J. (2020) Comparison of Laser Ablation and Solution-Based ICP-MS Results for Individual Foraminifer Mg/Ca and Sr/Ca Analyses. *Geochemistry, Geophysics, Geosystems,* **21,** e2020GC009254, doi:10.1029/2020GC009254.

7. Sirot, C., Ferraton, F., Panfili, J., Childs, A.-R., Guilhaumon, F. & Darnaude, A. M. (2017) elementr: An R package for reducing elemental data from LA-ICPMS analysis of biological calcified structures. *Methods in Ecology and Evolution,* **8,** 1659-1667, doi:10.1111/2041-210X.12822.

8. Rittner, M. & Müller, W. (2012) 2D mapping of LA-ICPMS trace element distributions using R. *Computers & Geosciences,* **42,** 152-161, doi:10.1016/j.cageo.2011.07.016.

9. Branson, O., Fehrenbacher, J. S., Vetter, L., Sadekov, A. Y., Eggins, S. M. & Spero, H. J. (2019) LAtools: A data analysis package for the reproducible reduction of LA-ICPMS data. *Chemical Geology,* **504,** 83-95, doi:10.1016/j.chemgeo.2018.10.029.

10. Mischel, S. A., Mertz-Kraus, R., Jochum, K. P. & Scholz, D. (2017) TERMITE: An R script for fast reduction of laser ablation inductively coupled plasma mass spectrometry data and its application to trace element measurements. *Rapid Communications in Mass Spectrometry,* **31,** 1079-1087, doi:10.1002/rcm.7895.

11. Fricker, M. B., Kutscher, D., Aeschlimann, B., Frommer, J., Dietiker, R., Bettmer, J. & Günther, D. (2011) High spatial resolution trace element analysis by LA-ICP-MS using a novel ablation cell for multiple or large samples. *International Journal of Mass Spectrometry,* **307,** 39-45, doi:10.1016/j.ijms.2011.01.008.

12. R Core Team. 2022. *R: A Language and Environment for Statistical Computing* . https://www.R-project.org: R Foundation for Statistical Computing.

13. Svetunkov, I. & Petropoulos, F. (2018) Old dog, new tricks: a modelling view of simple moving averages. *International Journal of Production Research,* **56,** 6034-6047, doi:10.1080/00207543.2017.1380326.

14. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., Grolemund, G., Hayes, A., Henry, L. & Hester, J. (2019) Welcome to the Tidyverse. *Journal of open source software,* **4,** 1686, doi:10.21105/joss.01686.

15. Wickham, H. & Seidel, D. 2022. *scales: Scale Functions for Visualization* . https://scales.r-lib.org.

16. Aze, T., Ezard, T. H. G., Purvis, A., Coxall, H. K., Stewart, D. R. M., Wade, B. S. & Pearson, P. N. (2011) A phylogeny of Cenozoic macroperforate planktonic foraminifera from fossil data.*Biological Reviews,* **86,** 900-927, doi:10.1111/j.1469-185X.2011.00178.x.

17. Lea, D. W., Mashiotta, T. A. & Spero, H. J. (1999) Controls on magnesium and strontium uptake in planktonic foraminifera determined by live culturing. *Geochimica et Cosmochimica Acta,* **63,**2369-2379, doi:10.1016/S0016-7037(99)00197-0.

18. D'angelo, C. & Wiedenmann, J. (2012) An experimental mesocosm for long-term studies of reef corals. *Journal of the Marine Biological Association of the United Kingdom,* **92,** 769-775, doi:10.1017/S0025315411001883.

With a smoothing of 5 observations,
a dt of 10 observations
and dt/scanRate of 3460 seconds