# Early Permian longitudinal position of the South China Block from brachiopod paleobiogeography

Robert J. Marks<sup>1</sup>, Nicolas Flament<sup>1</sup>, Sangmin Lee<sup>1</sup>, and Guang R. Shi<sup>1</sup>

<sup>1</sup>Environmental Futures, School of Science, University of Wollongong, New South Wales 2522, Australia.

Correspondence: Robert Marks (rm181@uowmail.edu.au) and Nicolas Flament (nflament@uow.edu)

Abstract. Knowledge of the past location of tectonic plates is essential to understanding the evolution of climate, ocean systems, and mantle flow. Tectonic reconstructions become increasingly uncertain back in geological timetimes. Paleomagnetic data constrain the past latitude of continental blocks, however, but not their past longitude unconstrained. For example, the longitude of the South China Block during the Early Permian Early Permian times is unknown. Paleobiogeographic data, which have long been used in tectonic reconstructions, make it possible to evaluate the faunal similarity between continental blocks. In this study, we use the Early Permian global brachiopod distribution for two Early Permian time periods (Asselian-Sakmarian times between ~298.9 Ma and ~290.1 Ma, and Artinskian-Kungurian times between ~290.1 Ma and ~274.4 Ma) from the Paleobiology Database to evaluate the correlation between faunal similarity and physical distance of continental blocks for three distinct tectonic reconstruction models. We use this approach to assess which of the three tectonic scenarios places the South China Block in a location that best accounts for the Early Permian brachiopod distribution data. Based on this analysis, the preferred tectonic reconstruction places the South China Block in a central position within the Paleo-Tethys Ocean instead of on its outskirts. The framework developed in this study is openly available and our approach could can be applied to other tectonic blocks, time periods, and faunal data.

#### 1 Introduction

Plate tectonic motions strongly influence Earth's internal and external systems (Valentine and Moores, 1970; Müller et al., 2008; Flament et al., 2017; Dutkiewicz et al., 2024), making it important to develop robust models of plate tectonic configuration to understand how these systems have evolved through deep timeover geological times. Constraining the latitude of a tectonic plate through over time can be achieved directly via paleomagnetic analysis (Torsvik and Van der Voo, 2002; Krivolutskaya et al., 2016), which determines the distance from the north or south North or South pole. Paleolongitude, however, cannot be constrained in absolute terms, and must be determined in relation to other features. This is commonly done by using either geophysical data from oceanic lithosphere (Scotese, 2004; Seton et al., 2012), which before recent times (< 60 Ma) is progressively lost to subduction through time is limited back in time due to subduction of oceanic lithosphere, or by relating orogens and stratigraphy across plate boundaries (Lehmann et al., 2010), which cannot provide information about isolated tectonic plates.

One method to constrain the relative locations of isolated tectonic blocks and for reconstructions of older times, at older times is paleobiogeographic similarity, which compares the similarity of ancient faunal assemblages between regions as a function of physical distance between tectonic blocks. Fossil data can be used to evaluate the past distance between tectonic plates (e.g. Piccoli et al., 1991; Lees et al., 2002) by assuming that the similarity between the faunal assemblages of two plates decreases as the physical distance between the plates increases. This assumption is based on the idea principle that continental configuration influences faunal distribution (Shi, 2001a, b; Zaffos et al., 2017). Tectonic plate configuration The configuration of tectonic plates controls climate (Valentine and Moores, 1972), ocean circulation (Valentine, 1971; Allison and Wells, 2006) (Valentine, 1971; Allison and Wells, 2006), and the locations of land and ocean barriers (Valentine and Moores, 1970), which are all factors that affect faunal distribution. Indeed, faunal similarity is expected to be negatively correlated with distance (Valentine, 1966), because the spread of species to geographically close areas is easier than to geographically distant areas. Statistical measures of the similarity between two faunal assemblages consider binary presence-absence data to assess either overlapping or unique fauna between two regions (Shi, 1993). Here, we present a new method-framework to apply these quantitative measures globally through a case study of the South China Block during the Early Permian.

25

30

35

Permian times (between ~298.9 Ma and ~274.4 Ma; ?), which causes a lack of geological evidence linking it to other blocks, thus limiting the data available for constraining available data to constrain the longitudinal position of the plate. In this contribution, we consider the Early Permian location of the SCB for three different plate tectonic reconstructions, with two placing it on the boundary of the Panthalassa Panthalassan Ocean and Paleo-Tethys Ocean (Wright et al., 2013; Matthews et al., 2016) (Wright et al., 2013; Matthews et al., 2016) and the other placing it much further west, locating it centrally within the Paleo-Tethys Ocean (Young et al., 2019). Paleomagnetic analysis of the Emeishan Large Igneous Province (Emeishan LIP) in the Mid Permian during mid-Permian times constrains the paleolatitude of the SCB to low latitudes around the equator (Krivolutskaya et al., 2016) at approximately ~260 Ma (Zhong et al., 2014). While this is not a direct latitudinal placement for the Early Permian times; consequently, the SCB can be assumed to be have been near this latitude during the Early Permian Early Permian times.

The differences in the proposed position of the SCB among the three reconstructions are largely due to differing methods to infer its Early Permian longitude. The eruption of the Emeishan LIP is attributed to a deep mantle plume, which has been suggested to originate from the edges of the African and Pacific Large Low Shear-Wave Velocity Provinces (LLSVPs) (Torsvik et al., 2008). The link between mantle plumes and LLSVPs has been used in some reconstructions to infer the longitude of plates on which Large Igneous Provinces LIPs are preserved, by placing such plates at the margin of one of the LLSVPs with the assumption that the LLSVPs have remained stationary since the Early Permian (Torsvik et al., 2010). In contrast, in some other reconstructions (e.g. Scotese, 2004; Young et al., 2019)(e.g. Scotese, 2004), the longitude is based on interpretations of tectonic activity preserved in the rock record. Here we These methods to determine paleolongitudes introduce uncertainties into the tectonic reconstructions, however, this is unavoidable since no direct evidence exists to constrain the past longitudinal

position of plates. By using the same faunal dataset to test each of the three tectonic reconstructions, the relationship between faunal similarity and physical distance is only influenced by the plate configurations for each tectonic reconstruction. This makes it possible to use statistical measures of faunal similarity to assess the compatibility of each of these SCB paleolongitude scenarios with global plate tectonic configuration, providing a test of paleolocation independent of the assumptions used to reconstruct the SCB and the uncertainties inherent to them.

#### 65 2 Data

#### 2.1 Plate Tectonic Reconstructions

We consider three tectonic reconstructions that are based on different approaches and present large variations in the location of the SCB during the Early Permian (Wright et al. (2013); Matthews et al. (2016); Young et al. (2019) Early Permian times (Wright et al. (2013); Matthews et al. (2016); Young et al. (2019) labelled here as W13Y19, M16 and Y19-W13, respectively). Reconstruction W13 primarily builds was built upon previous plate models from Scotese (2004), Golonka (2007) and Seton et al. (2012) by developing an updated location for the Australian plate based on paleoenvironments interpreted from faunal data. The placement of the SCB was an estimated transitional position based on information of tectonic activity in the rock record (Scotese, 2004). The relatively eastern placement of the SCB in-For reconstruction M16, the SCB was placed further east (133°E at 277 Ma) was-based on work by Domeier and Torsvik (2014) which attributes the Emeishan LIP to the western margin of the Pacific LLSVP under the assumption that the LLSVP has remained spatially fixed since the Emeishan LIP emplacement (Conrad et al., 2013). Reconstruction Y19 builds was built based on reconstruction M16 -using the same set of tectonic plate geometries, which makes a direct comparison between these two reconstructions possible. The global plate velocities required by the plate tectonic configuration in-used for reconstruction M16 were deemed to be unreasonably high, with the SCB moving during the Permian are large, requiring the SCB to move at speeds of up to 40 cm/yr between 260 Ma and 250 Ma. which is tectonically unreasonable (Young et al., 2019; Zahirovic et al., 2015) (Zahirovic et al., 2015; Young et al., 2019). Reconstruction Y19 uses was designed to have global plate configurations that give more tectonically reasonable plate velocities. This was done for the placement of the SCB by considering both records of tectonic activity to constrain the SCB longitude to a transitional position (similar to the approach of used for reconstruction W13), to define a position which gives more tectonically reasonable plate velocities (Young et al., 2019). All of the three reconstructions use paleomagnetic data as a latitudinal constraint which happens to coincide with the edge of a mobile large basal mantle structure produced by a mantle flow model (Young et al., 2019). While the edges of plume generating structures were considered for the placement of the SCB during the development of both reconstructions Y19 and M16, there remains key differences between the two. For reconstruction Y19, the SCB was placed over the eastern margin of the African basal mantle structure predicted by a mantle flow model, whereas the SCB was tied to the seismically imaged western margin of the Pacific LLSVP for reconstruction M16. Paleomagnetic data was used as the latitudinal constraint for the development of all three reconstructions, which is particularly effective because the north-south uncertainty that is typically a concern for paleomagnetic data is less pronounced for these SCB positions as it was situated near the equator. This study provides a chance to test the validity of the different reconstruction

models, particularly between-Here we investigate the compatibility of these different plate tectonic reconstruction models with faunal data, including reconstructions M16 and Y19, which tests that are designed using competing ideas on the mobility of LLSVPs.

100

105

110

120

We use used 277 Ma as the representative age of the Early Permian because W13 reconstructs the Early Permian plate tectonic configuration using paleoenvironments based on 277 Mato perform analysis because all three plate tectonic reconstructions have an explicitly defined position for the SCB at that age. This is due to-particularly important for reconstruction W13being based on the works of Golonka et al. (2006) and Golonka (2007), which considers the Early Permian to be 285 - 269 Ma, making, as positions for earlier Permian times are interpolated between the position at 277 Ma the midpoint, with the earliest Permian ages being grouped instead with the latest Carboniferous as many major continental collisions which formed Laurasia and Gondwana began during the Carboniferous and reached maturity during the earliest Permian (Golonka et al., 2006). This eauses many of the species present in earliest Permian brachiopod assemblages to be persistent from the Late Carboniferous (Shen et al., 2019), with and the next explicitly defined position during Late Carboniferous times at 306 Ma. Furthermore, profound climatic change occurred during Early Permian times (Fielding et al., 2008), with the peak of the Late Paleozoic Icehouse occurring near the Asselian-Sakmarian boundary (Fielding et al., 2008; Scotese et al., 2021), followed by the Artinskian Warming Event (Marchetti et al., 2022) which raised global average temperatures by 3-4°C (Scotese et al., 2021). This led to major changes in South China brachiopod taxonomic composition following sea-level rise during the middle of the Early Permian (Shen et al., 2019). Using this time allows for a better comparison between W13 and the other two reconstructions brachiopod faunal assemblages (Shen et al., 2013). Because of these climatic shifts and of the long duration of Early Permian times (27 Myr), over which tectonic plates may have moved considerably, we divided Early Permian times into two intervals: Asselian-Sakmarian times (between  $\sim$ 298.9 Ma and  $\sim$ 290.1 Ma; reconstructed at 295 Ma) and Artinskian-Kungurian times (between  $\sim$ 290.1 Ma and  $\sim$ 274.4 Ma; reconstructed at 277 Ma). We focused on Artinskian-Kungurian times because there are more fossil occurrences and all three tectonic reconstructions have an explicitly defined position at 277 Ma. Data were also analysed for Asselian-Sakmarian times to investigate the sensitivity of results to selected data and to showcase the applicability of our method to distinct periods. At 277 Ma, the SCB was placed further west in reconstruction Y19 by approximately 3,300 km compared to reconstruction M16 and 2,500 km compared to reconstruction W13; the centroid of the SCB is-was located at longitude 103°E in reconstruction Y19, 133°E in reconstruction M16 and 125°E in reconstruction W13. Because these reconstructed locations are different, we anticipate that their respective compatibility with brachiopod distribution can be established quantitatively.

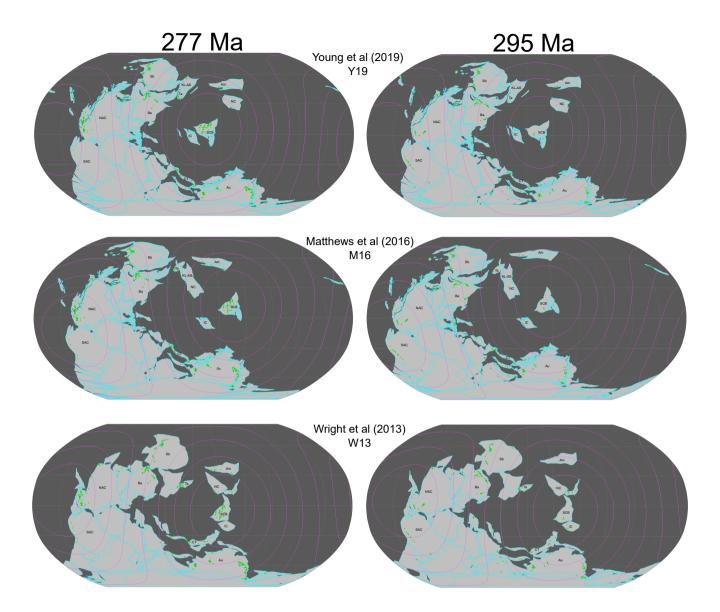


Figure 1. Reconstructions All three considered tectonic reconstructions at 295 Ma and 277 Ma with distance ranges as concentric circles at from 4,000 km 000-16,6,000 km from the SCB in 2,8,000 km , 10,000 km, and 12,000 km from the South China Block (SCB) increments. (a) Reconstruction of Young et al. (2019, Y19), (b) Reconstruction of Matthews et al. (2016, M16), and (c) Reconstruction of Wright et al. (2013, W13). Fossil occurrences are shown as green diskscircles, plate outlines as open cyan polygons, continents as filled light grey polygons, and oceans in dark grey. The name names of tectonic blocks are abbreviated as follows: Am: Amuria, Au: Australia, Ba: Baltica, CT: Cimmerian Terranes, IC: Indochina, KL-AS: Kunlun - Ala Shan, NACNA: North America Craton, NC: North China, SACSA: South America Craton, Sb: Siberia, SCB: South China Block, Ta: Tarim.

### 2.2 Brachiopod fossil data from the Paleobiology Database

135

140

145

150

The distribution of Early Permian brachiopods can be summarised by into three realms with distinct brachiopod faunal assemblages (Waterhouse and Bonham-Carter, 1975): Boreal (northern latitudes), Paleo-equatorial, and Gondwanan (southern latitudes) realms. These realms exhibit biogeographic patterns that match those expected from their continental configuration.

The Gondwanan and Boreal realms are dominated by large continental landmasses and cold-water environments which produce the expected that produce low diversity, while the Paleo-equatorial realm, which has warm waters and hosts characterised by warmer waters and many small island regions, has a correspondingly presents a high degree of biodiversity (Shen et al., 2013). The consistency between expected biogeographic patterns based on continental configuration and the biogeographic realms suggests that Early Permian brachiopod distribution should match the expected negative trend of decreasing faunal similarity with distance, making them an appropriate dataset for this study. Brachiopod fossils are also used commonly for understanding Palaeozoic biogeographic patterns, as they were the most dominant Palaeozoic benthic components in most marine conditions.

In order to conduct a reliable biogeographic approach at global scale, brachiopod biogeographic data were selected at the genus level; species-level data are also available, but genus-level data are generally less affected by taxonomic biases. The biogeographic data were obtained from the Paleobiology Database which provided an open, solid and updated dataset solely based on published taxonomies. The validity of the genera included in the dataset was checked before performing the biogeographic analyses, and reports of genera with uncertain identification were excluded. To further alleviate taxonomic uncertainty and demonstrate the flexibility of the developed framework, a sensitivity analysis was performed at the family level, with results presented for the strongest subset of results.

Early Permian brachiopod biogeographic data were downloaded from the Paleobiology Database (Paleobiology Database) on 2911/07/2021. Data on "occurrences" 2025 for both the Artinskian-kungurian interval and Asselian-Sakmarian interval. Data on 'occurrences' were retrieved for the taxa "brachiopoda" 'brachiopoda' with a taxonomic resolution of "'lump by genus"', excluding uncertain genera and using the following parameters: Interval the 'interval or Ma range= 299-272, age rule = major, outputs = coordinates, location, paleolocation' was set equal to 290.1-274.4 for Artinskian-Kungurian times and to 298.9-290.1 for Asselian-Sakmarian times, the 'age rule' was set to 'contains', 'outputs' were set to 'coordinates', and 'time binning' and 'classification' were enabled. The paleobiogeographic data set consisted of 15,670-9,988 fossil occurrences made up of 591 genera-502 unique genera belonging to 81 families for Artinskian-kungurian times, and 4, with columns used for analysis: "accepted\_name" 670 fossil occurrences made up of 359 unique genera belonging to 73 families for the Asselian-Sakmarian times. We analysed the following data frame columns: 'genus' (genus name), "lat" 'lat' (latitude), and "lng" 'lng' (longitude), and 'family' (family name).

The fossil record is known to be incomplete, although the degree to which this impacts studies using fossil data is contested (Benton et al., 2000). Preservation biases lead to differential preservation of biota between different environments so that regions in which large amounts of sedimentary rocks are deposited are more favourable for fossil preservation. Regions preserving Early Permian sedimentary rocks may preserve greater proportions of the regional biota because marine sedimentary

155 rock volume is positively correlated with biodiversity (Raup, 1976; Crampton et al., 2003). Observed variations in biodiversity have also been linked to the number of people working on sampling for a region (Sheehan, 1977), with a larger number of workers in a given region leading to more discovery of specimens. This causes larger biota sets to be known for developed regions is clear from Figure 1 in which the greatest point densities occur in present-day North America, Europe, China, and Australia. Such biases are likely to cause apparent differences in biodiversity between regions that are not representative of true faunal assemblages.

#### 3 Methods

# 3.1 Biogeographic indexesindices

Several binary similarity coefficients indices have been independently developed (Jaccard, 1907; Simpson, 1960; Lees et al., 2002) using varying conceptual bases to measure the similarity between two sets of fauna faunal data sets (Hohn, 2018). These coefficients indices are useful tools in paleobiogeography to quantify faunal similarity between biogeographic regions by comparing regional biota datasets based on presence-absence data (Schmachtenberg, 2008; Fallaw an (Fallaw and Dromgoole, 1980; Shi, 1993; Schmachtenberg, 2008). The underlying conceptual basis for each index introduces unique statistical biases for each measure, so multiple faunal similarity coefficients indices are commonly used in conjunction (Hohn, 2018; Simpson, 1960; Schmachtenberg, 2008) (Simpson, 1960; Schmachtenberg, 2008; Hohn, 2018).

The Jaccard Coefficient index (JC), Jaccard, 1907) measures the true similarity between any two sets of fauna, assuming that any differences in biodiversity are real. This means that the sampled distribution is assumed to be an accurate representation of the true distribution. It is calculated by dividing the intersection of the biota sets by the union of the biota sets:

$$JC = \frac{S1 \cap S2}{S1 \cup S2},\tag{1}$$

where S1 is the taxa set for the less diverse region (less genera) and S2 is the taxa set for the more diverse region (more genera).

The Simpson Coefficient index (SC, Simpson, 1960) takes into account that some differences in biodiversity between regions could be a result of sampling bias. This is done by not including the number of genera in the more diverse region in the calculation:

$$SC = \frac{S1 \cap S2}{n1},\tag{2}$$

where  $n_1$  is total number of genera in the less diverse region  $S_1$ .

The two-plate Mean Endemism index (ME, Lees et al., 2002) accounts for differences in area-biogeographic area and biogeographic effects between two regions by averaging the proportions of endemic fauna in each region:

$$ME = \frac{\left(\frac{S1 - S2}{n1}\right) + \left(\frac{S2 - S1}{n2}\right)}{2},\tag{3}$$

where  $n_2$  is the number of genera in the more diverse region  $S_2$ . JC and SC measure faunal similarity and are expected to be negatively correlated with distance, while ME measures faunal difference and is expected to be positively correlated with distance. For consistency, we use used cME = 1 - ME that is expected to be negatively correlated with distance. Values for each of the three indexes JC, SC and cME range between 0, indicating complete dissimilarity, and 1, indicating complete similarity.

We considered the natural logarithmic transformation of each index value  $(\ln(JC), \ln(SC), \text{ and } \ln(cME))$  to account for the potential of an exponential relationship between distance and faunal similarity as suggested by Piccoli et al. (1991). The logarithmic indexes have maximum value  $\ln(1) = 0$ , indicating total similarity, and get progressively more negative as the logarithm argument (linear index value) approaches 0 because  $\lim_{x\to 0^+} ln(x) = -\infty$ .

To calculate Early Permian faunal similarity indexes, the present-day fossil locations must first be-were first transformed to their Early Permian paleo-locations. The present-day latitude and longitude of each fossil occurrence were used to define a point on the Earth, which was then converted to a feature collection saved in a GPlates Markup Language ("gpml', gpml') file (Müller et al., 2018) using pyGPlates (Williams et al., 2017). A "'point-in-polygon" test was carried out with PyGPlates pyGPlates to assess which tectonic block each point belonged to, and to attribute the corresponding plate ID specific to each tectonic reconstruction. This made it possible to split the data into subsets of biota for each plate and to reconstruct the points back in time using PyGPlatespyGPlates. As a result of the reconstruction of to paleo-locations, one fossil occurrence point was located on a W13 plate boundary, and two points for reconstructions M16 and Y19 were located within a plate void. These points were discarded as they represented minor portions of the total data set dataset and including them would have introduced errors.

# 3.2 Calculating distance between plates

185

190

195

200

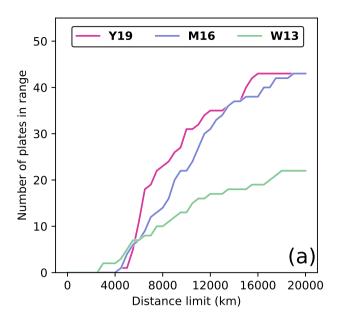
205

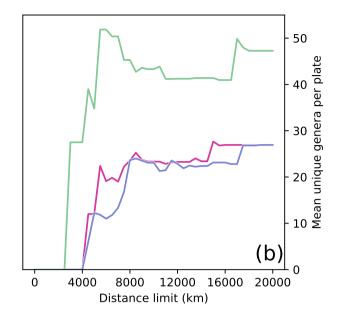
210

215

After each plate is—was assigned a faunal assemblagewe calculate, we calculated physical distances between each plate to establish the relationship between faunal similarity and distance. The physical distance between plates was calculated as great circle distances along the surface of a sphere with a constant radius of 6,371 km between plate centroids using pyGPlates. Some plate IDs were assigned to multiple polygons that collectively represented a larger plate. In these such cases, a new polygon was created from the multiple plate centroids and the centroid of that new polygon was used as reference for that plate.

A distance limit Distance limits  $d_l$  from the SCB was were imposed to obtain more meaningful results. The distance limit must include a large enough sample of plates to obtain a statistically meaningful relationship between distance and similarity (Fig. 2a) and to reduce the impact of noise from random sampling errors that are inherent to paleontological data sets. Larger distances, however, larger distances introduce continental land barriers between plates (Fig. 1), which complicates the relationship between biogeographic similarity and distance by making migratory distances much longer than considering great circle arc distances that are longer than reasonably expected for brachiopod migratory distances, because brachiopods are marine organisms. It is also important that plates within a distance limit have present, on average, a large enough faunal assemblage to make meaningful comparison with the SCB faunal assemblage (Fig. 2b). We For each reconstruction, we considered distance limits surrounding the SCB from 4,000 km to 12,000 km in 2,000 km intervals (beginning at the distance of the closest plate





**Figure 2.** Evaluating the effect of the distance limit on sample sizes during Artinskian-Kungurian times, considering only plates with at least one fossil occurrence. (a) Number of plates with that have centroids within the distance limit from the SCB centroid for each of the three considered tectonic reconstructions W13, M16 and Y19. (b) Average number of genera per plate for the set of plates within a given distance limit for each of the three considered tectonic reconstructions.

centroid, increasing by steps of 50 km in radius until a maximum radius of 20,100 km (just over half of Earth's maximum circumference) to test progressively larger tectonic plate data sets (see concentric circles in Fig. 1). A plate was considered within range if the minimum distance between its the plate centroid and the centroid of the SCB was smaller than the distance limit. If a plate centroid was within range, all fossil occurrences on the plate were considered.

# 3.3 Statistical Tests and Data Processing

220

225

230

All three biogeographic index values and their natural logarithms were calculated for each plate pair that consisted of the SCB and another plate. The Pearson Spearman rank correlation coefficient (*r*-value *r*<sub>s</sub>) was used to determine the strength monotonicity and direction of the linear relationship between faunal similarity as measured by the biogeographic indices and physical distance. A negative *r*-value *r*<sub>s</sub> is expected if faunal similarity decreases as physical distance increases. The Spearman correlation coefficient was selected as it has been argued that faunal similarity should decrease exponentially with physical distance (Piccoli et al., 1991), however, it has been found that using a natural logarithm transformation on the faunal similarity data does not always provide better results (Schmachtenberg, 2008), and often these results are very similar. As the Spearman rank correlation tests only for monotonic relationships, it has equal power in determining a relationship whether it is linear or exponential. The Spearman rank correlation is also more robust to outliers and tailed distributions than the linear Pearson

correlation coefficient (De Winter et al., 2016), which helps mitigating the uncertainty inherent to working with fossil datasets. The final benefit of the Spearman correlation is that, as it is a rank correlation which tests only for monotonicity, the order in which plates are encountered with increasing distance is more important than the distances themselves, making this test a more robust measure of global plate tectonic configuration than a purely linear or logarithmic relationship.

We carried out a one-tailed test to determine the statistical significance of the relationship between faunal similarity and physical distance for each distance limit in each tectonic reconstruction. We used formulated the null hypothesis that faunal distribution was random and unrelated to physical distance between plates, which was rejected at the 95% confidence interval if the pvalue was smaller-value was less than 0.05. A one-tailed test was chosen as because it allowed us to test for statistical significance only if the relationship was negative (decreasing faunal similarity as physical distance increases) as we consider. We considered a positive global relationship (faunal similarity increases with physical distance) is to be more indicative of an incorrect plate tectonic configuration rather than a true relationship. The strength of the correlation between faunal similarity and physical distance and the statistical significance of those relationships were then used to determine the most appropriate reconstruction and ideal distance limit within that reconstruction.

Data were prepared and analysed using Python Jupyter Notebooks (Kluyver et al., 2016). The pandas (McKinney, 2010), SciPy (Virtanen et al., 2020) and NumPy (Harris et al., 2020) libraries were used for statistical methods, the pyGPlates (Müller et al., 2018; Williams et al., 2017) library for reconstruction and geospatial operations, and the Matplotlib library (Hunter, 2007) to plot results.

#### 4 Results

235

240

245

250

260

Comparing genus-level faunal similarity indexes between the SCB and all other plates in reconstruction W13 reveals a clear inverse relationship between physical distance and faunal similarity Y19 revealed negative monotonic relationships for all three indices (Fig.3), with five of six indices having r < -0.2. The relationship is particularly strong for cME (Fig.3). 3The correlation was strongest for index SC ( $r_s = -0.1895$ , Fig. 3b), followed by index cME ( $r_s = -0.1396$ ; Fig. 3c), for which the relationship is statistically significant. These relationships and a weak correlation for index JC (Fig. 3a). None of these relationships were statistically significant, however, they include distant plates for which the migratory distance for brachiopods is much larger than the measured distance, meaning the results can be improved by imposing distance limits on the analysis. 255 All possible distance limits were examined for the strength and statistical significance of correlations for each of the three reconstructions (Figs. 4 and 5).

The correlations between faunal similarity and physical distance are expected to be negative for all distance limits, with more negative correlations showing a stronger inverse relationship (white section of Fig. 4), while positive correlations indicate the unexpected relationship of increasing faunal similarity as distance increases (grey section of Fig., which was the case for reconstructions Y19 and W13 where  $r_s$  was negative for > 90% of all  $d_l$  for the Artinskian-Kungurian. For reconstruction M16, however, positive correlations (grey section of Fig. 4) were obtained for 47% of all  $d_l$ . At smaller distance limits, all three reconstructions showed large instability in the relationship between faunal similarity and physical distance for each

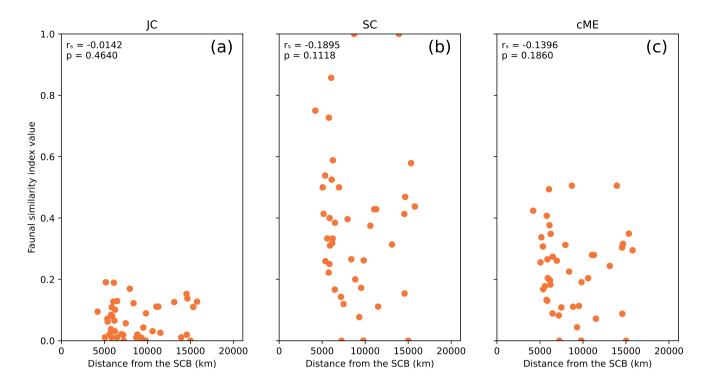


Figure 3. All considered biogeographic similarity indexes (and their logarithmic transformations) calculated between the SCB and all 31-43 other plates with genus-level brachiopod occurrences in reconstruction W13Y19 at 277 Ma, used to represent Artinskian-Kungurian times.

\*\*T\_E\*\* and pvalues\_values are presented for each index.

270

275

of the indices, with large fluctuations in  $r_8$  for small changes in  $d_l$  (Fig. 4). Indexes SC and cME and their logarithmic transformations consistently produce stronger correlations than indexes JC and In(JC). For indexes The relationship stabilised for all three reconstructions with all indices at larger distance limits (smaller fluctuations as new plates are added to the analysis), near  $d_l = 10,000$  km for reconstruction Y19, and  $d_l = 11,000$  km for reconstructions M16 and W13. Distance limits at or above those thresholds are most appropriate for further analysis. For indices SC and cME, reconstructions W13 and Y19 both produce stronger negative correlations than the and W13 peaked at  $r_8 \lesssim -0.45$  for  $d_l \gtrsim 10,000$  km (Fig. 4b and c), while peaks for reconstruction M16 reconstruction were at  $r_8 \leq 0.4$  for all indices. Index JC presented the overall weakest performance of all three faunal similarity indices, with none of the three considered reconstructions reaching  $r_8 < -0.4$  for any  $d_l > 10,000$  km. While strong correlations were obtained for reconstructions Y19 and W13, it is important to ensure the correlations are also statistically significant.

Faunal similarity indexes. The statistical relationship between faunal similarity and physical distance was strongest for reconstruction Y19. There was consistent statistical significance for indices SC and cME at 10, cME and their logarithmic transformations are significantly anti-correlated with distance from the SCB 000 km  $< d_1 < 14.000$  km (Fig. 4) in reconstruction 5b and c). Statistical significance was lowest for index JC, with significance only for Y19 when the distance limit is equal

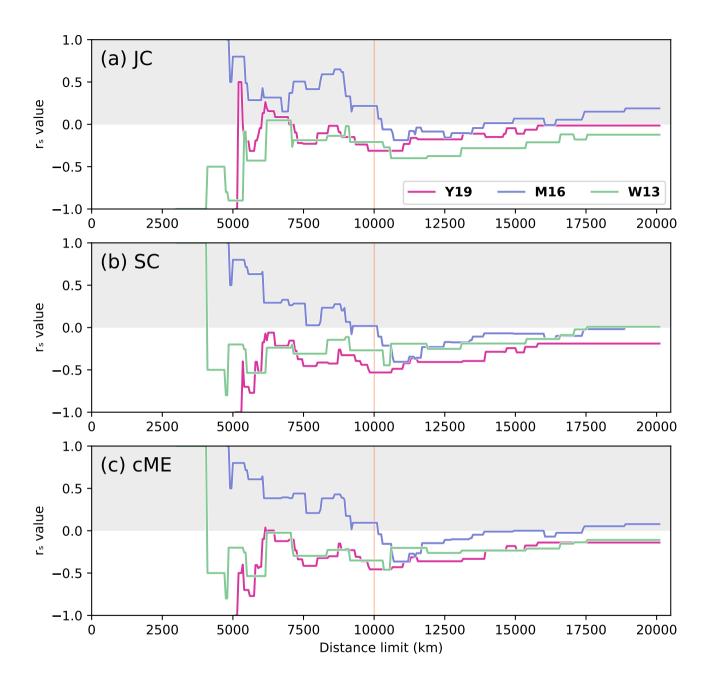


Figure 4. Pearson r value Spearman correlation coefficient  $r_s$  for all distance limits for each considered indexand its logarithmic transformation, measuring the anti-correlation of genus-level biogeographic similarity and distance when comparing the SCB to all other plates with centroids located within the distance limits on the #X-axis (only plates within the distance limit were considered in the correlation) for Artinskian-Kungurian times, represented by reconstructions at 277 Ma. Grey The grey region indicates positive correlation correlations, meaning faunal similarity increases with physical distance, which is an unexpected trend. The vertical orange line at  $d_l = 10.000$  km indicates the preferred distance limit.

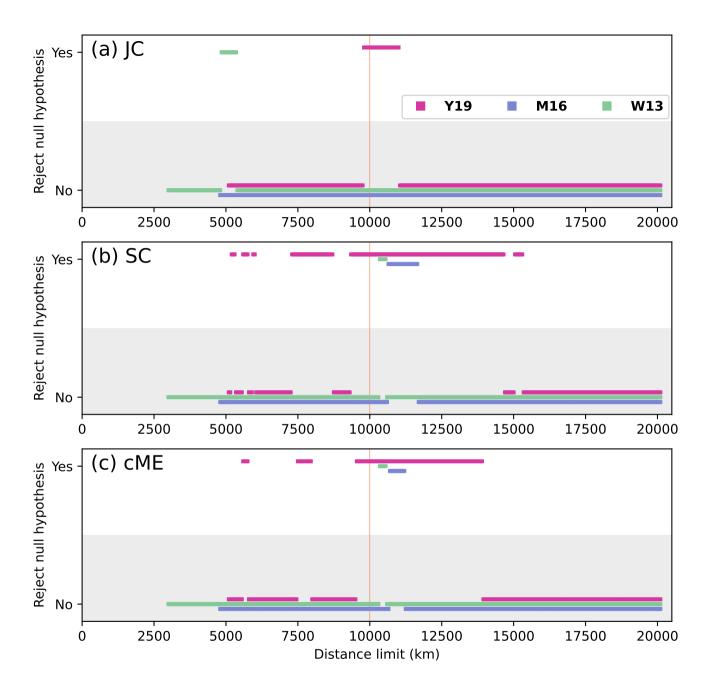


Figure 5. Statistical significance of the anti-correlation between each considered index and its logarithmic transformation and distance from the SCB for all three reconstructions values of  $r_s$  presented in Figure 4. A relationship is statistically significant ('Yes' in the graph) if the null hypothesis that faunal distribution is unrelated to physical distance between plates is rejected because the *p*-value is smaller less than 0.05. The vertical orange line at  $d_t = 10.000$  km indicates the preferred distance limit.

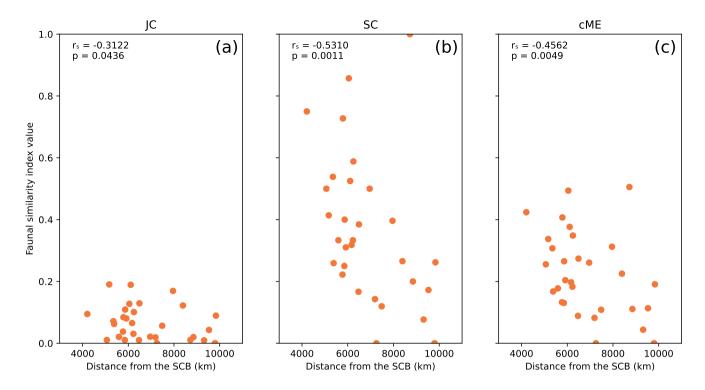


Figure 6. Relationships between biogeographical indexes genus-level faunal similarity and physical distance from the SCB for all 28-each index JC (a), SC (b) and cME (c), for 31 plates within  $d_t = 8d_t = 10,000$  km of the SCB in reconstruction Y19. Biogeographic indexes are shown in the top panels 277 Ma, used to represent Artinskian-Kungurian times. The Spearman correlation coefficient  $r_s$  and their logarithmic transformations p-value are shown in the bottom panels indicated on each subpot for each index.

to 8,000 kmor and 9,800 km  $< d_l < 10,000950$  km (Fig. 5). Other statistically significant anti-correlations are obtained for JC and cME with  $d_l = 12,000$  km in reconstruction Y19, and for SC with  $d_l = 6,000$  km a). Statistically-significant relationships occurred for some  $d_l$  for W13 and cME with M16, however, the ranges for  $d_l$  which were statistically significant for reconstructions W13 and M16 were far smaller than the statistically significant  $d_l$  ranges for reconstruction Y19. Statistical significance also occurred for Y19 for correlations with greater distance limits than for the other two reconstructions, up to  $d_l = 10,000$  km in reconstruction W13950 km (JC), 15,300 km (SC) and 13,900 km (cME). These results suggest that the Early Permian Artinksian-Kungurian longitude of the SCB in reconstruction Y19 is most consistent with the brachiopod biogeographic data among the three reconstructions, considered tectonic reconstructions. For reconstruction Y19, an optimal distance limit was determined at  $d_l = 10,000$  km based on  $r_s$  and p as a function of  $d_l$  (Figs. 4 and 5).

280

285

290

A strong anti-correlation between faunal similarity and physical distance from the SCB was obtained for reconstruction Y19 with  $d_l = 8d_l = 10,000$  km (Fig. 4)-6) during Artinskian-Kungurian times. In detail, the base indexes (Fig. 6a-c) tend to show slightly better correlations than their logarithmic transformations (Fig. 6d-f). all three indices presented strongly negative and statistically significant correlations. The strongest correlation are obtained for indexes was obtained for index SC and CME

(Fig. 6b-c), and both of these correlations are statistically significant (Fig. 5b-c). Strong (Fig. 6e-f( $r_s = -0.5310$ ), which also exhibited the largest range of data, with points spread between a maximum of 1 and minimum of 0. Strong ( $r_s < -0.3$ ) and statistically significant (Fig. 5e-f) anti-correlations correlations were also obtained for the logarithmic transformations of SC indices JC and cME. These results show that there is a robust, statistically significant decrease in faunal similarity as distance increases from the SCB as it is positioned in reconstruction Y19. In contrast, statistical significance cannot be established for index JC and its logarithmic transformation (Figs. 5a, d,during Artinskian-Kungurian times.

The relationship between faunal similarity and physical distance during Asselian-Sakmarian times did not favour one reconstruction as strongly as the Artinskian-Kungurian results (Fig. 7). Similar to Artinskian-Kungurian times, there were large fluctuations in  $r_s$  with small changes to  $d_l$  for  $d_l \leq 6a$ ) This is likely due to the SCB being a much larger faunal data set than most other plates, which overwhelms the JC index as it considers the union of the faunal data sets (Eq. 112,000 km for all reconstructions and indices. The most consistently statistically significant relationship occurred for reconstruction Y19 with index JC, which was statistically significant for 11,300 km  $< d_l < 14,550$  km with a minimum  $r_s = -0.3923$  and mean  $r_s = -0.3238$  (Fig. 7a). This suggests that accounting for biases in the number of fossil occurrences on each tectonic block is important to infer significant relationships between the faunal similarity and reconstructed distance between tectonic blocks because the SC and cME indices both were developed to account for this and have shown more success here, was the largest unbroken  $d_l$  range which is statistically significant during Asselian-Sakmarian times. For indices SC and cME, statistical significance was obtained most consistently for reconstruction W13 (Fig. 7b and c). This was primarily for lower  $d_l$ , with reconstruction W13 having no statistically significant correlations using  $d_l > 11.450$  km for any index. Statistical significance was obtained least consistently for reconstruction M16, although for index JC statistical significance is obtained for 12,800 km  $< d_l < 14,500$  km with  $r_s = -0.2840$  (Fig. 7a).

#### 5 Discussion

295

305

310

# 5.1 Position of South China Block in the during Early Permian times

Of the three considered reconstructions, the Early Permian

### **5.1.1** Artinskian-Kungurian times

The brachiopod fossil record is more consistent with the Artinskian-Kungurian location of the SCB in reconstruction Y19 is most supported by the brachiopod fossil records than in reconstructions W13 and M16. This is indicated by strong, statistically significant negative correlations between faunal similarity and distance from the SCB for this tectonic reconstruction tectonic reconstruction Y19 (Figs 4 and 5). The Early Permian Artinksian-Kungurian location of the SCB in reconstruction Y19 effectively splits the Paleo-Tethys Ocean into two arms (Fig. 1a), in contrast to its position in the eastern margin of the ocean in reconstructions W13 and M16 (Fig. 1b, e 1c and e).

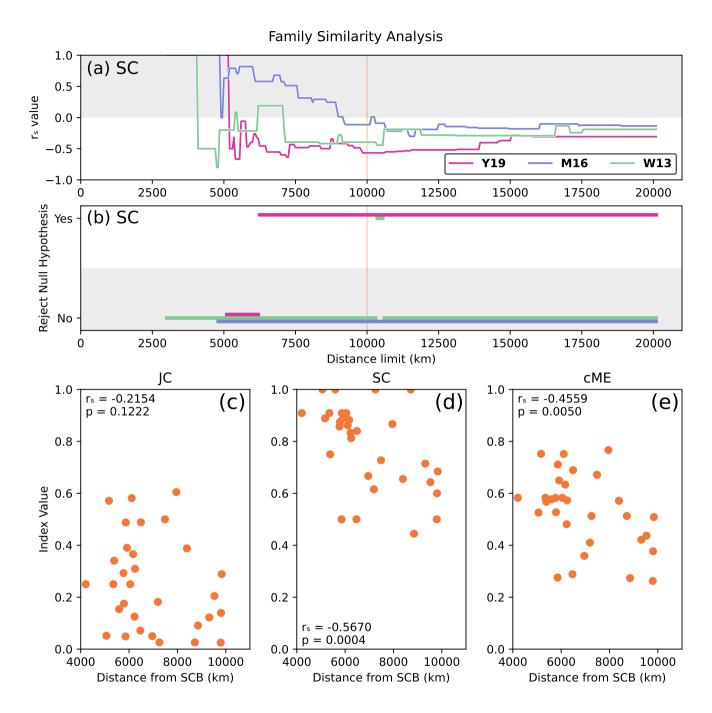


Figure 7. Faunal similarity and physical distance correlations for Asselian-Sakmarian times, represented using a reconstruction at 295 Ma. All three indices are presented for each reconstruction and all distance limits. Statistically significant  $r_8$  is shown as thick lines, while thinner lines represent statistically non-significant correlations.

Differences in the Early Permian Artinksian-Kungurian locations of tectonic blocks isolated within ocean basins (e.g. North China, Amuria) between the considered tectonic reconstructions introduce some uncertainty in the results. As explained in more details below, the main difference between reconstructions M16 and Y19 for the Early Permian Artinksian-Kungurian times is the location of the SCB (Fig. 1 1a, c, and e). Differences in the reconstructed location of other continental blocks affect the distances to the SCB and the relationships with biogeographic indexes. As such, the comparison between reconstruction W13 and reconstructions M16 and Y19 is less direct.

325

330

335

340

345

350

355

There are few differences in the geometry of Pangea between reconstructions M16 and Y19 (Fig. 1a, band c), with the Siberian and Baltica blocks slightly further south-east in reconstruction Y19 compared to their locations in reconstruction M16. There are also differences in the location of smaller, isolated blocks. For example, Indochina and Amuria are reconstructed at slightly different distances from the SCB in reconstructions M16 and Y19. Although the North China Block is at similar distances from the SCB in reconstructions M16 and Y19, it is to the east of the SCB in reconstruction M16 and to the west of the SCB in reconstruction Y19. In the absence of other significant differences, the major cause for variation in the relationship between biogeographic indices and distance for reconstructions M16 and Y19 is expected to be the difference in distance of the SCB from Pangean circum-Tethyan blocks between the two reconstructions. Indeed, the SCB is roughly 4,000 km further from the east Pangean coastline (comprised of plates such as Baltica, Armorica, the Pontides, Arabia, and Lutprimarily of Baltica and the Cimmerian Terranes), 2,000 km further from the Cimmerian terranes, and 1,000 km closer to the Australian continental block in reconstruction M16 than in reconstruction Y19 (Fig. 1a, band c).

The SCB is reconstructed slightly further to the south-west in reconstruction W13 than in reconstruction M16, and it is reconstructed significantly further to the south-east of its location in reconstruction Y19 (Fig. 1a, c, and e). Even though the locations of the SCB are closer between reconstructions W13 and M16 than between reconstructions W13 and Y19, the correlation results are more similar between reconstructions W13 and Y19 than between reconstructions W13 and M16 (Fig. 4). This result is likely due to other major differences in global paleogeography between reconstruction W13 and reconstructions M16 and Y19 (Fig. 1a, c, and e). The Cimmerian terranes are separated from Pangea and further to the north-east in reconstruction W13, which changes their distribution from the SCB from almost equidistant (the Cimmerian terranes are about 6,000 km away from the SCB in reconstruction Y19 and 8,000 km away from the SCB in reconstructions reconstruction M16 - (Fig. 1)-a and c), to mostly radially distributed from about 500 km to about 8,000 km away from the SCB in reconstruction W13—(Fig. 1e). The northern Pangean blocks (Siberia and Baltica) are further to the south-east and North China and Amuria are further to the south (such that North China borders the SCB) in reconstruction W13 than in reconstructions M16 and Y19. These differences may cause the distance to the SCB for many plates to be more similar between reconstructions W13 and Y19 than between reconstructions W13 and M16 (Fig. 4). The different relationships between biogeographic indexes and distance for reconstructions W13 and Y19 (Figs 4 and 5) are due to different distances between the SCB and other plates, including the Cimmerian terranes (see above), the North China Block that is about approximately 4,000 km away from the SCB in reconstruction Y19 but adjacent to it in reconstruction W13, Baltica that is 5,000 km away from the SCB in reconstruction Y19 and 7,000 km away in reconstruction W13, and the Australian block, that is 4,000 km from the SCB in reconstruction W13 and 6,000 km from the SCB in reconstruction Y19. It is important to consider that the SCB in reconstruction W13 differs from the reconstruction Y19 and M16 positions for the SCB not just longitudinally, but also latitudinally. This, along with the different set of tectonic blocks compared to Y19 and M16, means we cannot consider this as purely a comparison and constraint for SCB longitude. Nevertheless, this remains valuable as a comparison for the overall suitability of the SCB location within the global tectonic configuration of W13 compared to Y19 and M16.

The observed decrease in correlation for the

#### 5.1.2 Asselian-Sakmarian times

360

365

370

375

380

385

There is no clear result indicating which reconstructed SCB location is most favoured by brachiopod faunal distribution during Asselian-Sakmarian times (Fig. 7). Statistical significance was most consistently obtained for reconstruction W13, however, much of these statistically significant relationships were for short distance limits for which the relationship is inconsistent (Fig. 7). Statistically significant relationships were consistently obtained for reconstruction Y19 using index JC for large distance limits (Fig. 7a), however, this is not reproduced for indices SC and cME as the distance range  $d_I$  increases for, for which  $r_s$  was low without statistical significance for large distance limits. For Asselian-Sakmarian times, the weakest overall relationship between faunal similarity and physical distance from the SCB was obtained for reconstruction M16. For reconstructions Y19 and W13-M16 there are only minor changes in plate tectonic locations between 295 Ma and 277 Ma (Fig. -4)is expected. Indeed, correlation is expected to decrease as sample size increases (which is the case when more plates that earry more fossil occurrences are included). The absence of a trend between correlation and distance limit for reconstruction 1a, b, c, and d). There are small changes in location for the other isolated Tethyan plates and minor movement of the SCB between 295 Ma and 277 Ma in reconstruction M16could reflect the more complex relationship between sample size and distance limit for that reconstruction (Figs 1b and 2), but the Pangean coastline remains quite similar between the two time steps for both reconstructions. There is greater movement of tectonic blocks between the two time steps for reconstruction W13, in particular the Cimmerian Terranes separate from Pangea (Fig. 1e and f). Regardless of these changes in global plate configuration for any of the three reconstructions, plates are encountered in a very similar order at both 295 Ma and 277 Ma. This means the Spearman correlation should remain similar to what was obtained for Artinskian-Kungurian times. The lack of clear relationships for any of the reconstructions is likely a result of the smaller fossil dataset for Asselian-Sakmarian times (4,670 occurrences) than for Artinskian-Kungurian times (9,988 occurrences) and differences in the distribution of these fossils.

There are some noteworthy changes in fossil distribution for Tethyan tectonic plates between 277 Ma and 295 Ma. Indochina, the closest plate to the SCB, presents over 40 genera during Asselian-Sakmarian times, but not a single fossil occurrence attributed to the plate during Artinskian-Kungurian times (Fig. 1). There are also 65% more genera for North China and 30% more genera for the northern Cimmerian Terranes during Asselian-Sakmarian times compared to Artinskian-Kungurian times. Conversely, Amuria records over 40 genera during Artinskian-Kungurian times but only seven during Asselian-Sakmarian times, and Kunlun-Alashan records 20 genera during Artinskian-Kungurian times compared to 0 during the Asselian-Sakmarian times. Additionally, there are 65% more genera on the Tarim block and 25% more genera on the southern Cimmerian Terranes during Artinskian-Kungurian times (Fig. 1 point densities between the two columns).

These differences in fossil distribution between the two time periods are expected to have the greatest biogeographic connection with the SCB because they are for blocks in the same ocean basin as the SCB. The small sample sizes on these blocks during Asselian-Sakmarian times than during Artinskian-Kungurian times weaken the power of the statistical testing to determine a relationship (VanVoorhis et al., 2007). Results may be improved by expanding the dataset for Asselian-Sakmarian times by considering more marine phyla in the analysis (rather than just brachiopoda), which could be done with the framework introduced for this analysis.

# 5.2 Limitations and possible future improvements

# 5.2.1 Statistical significance

400

The fossil record is known to be incomplete, although the degree to which this impacts studies based on fossil data is debated (Benton et al., 2000). Preservation biases lead to differential preservation of biota between different environments so that regions in which large amounts of sedimentary rocks are deposited are more favourable for fossil preservation. Regions preserving Early Permian sedimentary rocks may preserve greater proportions of the regional biota because the volume of marine sedimentary rock is positively correlated with biodiversity (Crampton et al., 2003; Raup, 1976). Observed variations in biodiversity have also been linked to the number of people working on sampling for a region (Sheehan, 1977), with a larger number of workers in a given region leading to more discovery of specimens. This causes larger biota data sets to be known for developed regions as is clear from Figure 1, in which the greatest point densities occur in present-day North America, Europe, China, and Australia. Such biases are likely to cause apparent differences in biodiversity between regions that are not representative of true faunal assemblages. These sampling biases inherent to fossil data can make it difficult to effectively determine relationships, however, increasing the sample size makes the sampled distribution more closely resemble the true distribution, thus reducing the impact of sampling errors on the statistical test (?).

Statistical significance could generally be established for greater distance limits (Fig. 5) as increasing the distance limit increases the sample size, which increases the power of the statistical test to identify a relationship (VanVoorhis et al., 2007). This is clear from the calculation of the *t*-statistic that is used to determine the *p*-value:

$$t = r_s \sqrt{\frac{n-2}{1-r_s^2}},\tag{4}$$

where *n* is the sample size, meaning *t* is proportional to *n*. The probability density function for *t* is a bell-shaped curve centered about 0 (Kokoska and Zwillinger, 2000). The area under this curve from *t* → ∞ gives the *p*-value, thus as |*t*| increases, the *p*-value decreases (Kokoska and Zwillinger, 2000), meaning that when *r<sub>s</sub>* is kept constant the power to determine statistical significance increases with sample size *n* (Eq. 4). This explains why there are many cases in which, while *r<sub>s</sub>* values may be very similar for reconstructions W13 and Y19, the correlation for reconstruction Y19 is statistically significant while the correlation for reconstruction W13 is statistically non-significant. For all *d<sub>l</sub>* > 5,500 km, the sample size was greater for reconstruction Y19 than for reconstruction W13 (Fig. 2a), giving the statistical test greater power to identify a relationship.

### 5.3 Limitations and Possible Future Improvements

#### **5.2.1** Taxonomic level

425

430

435

440

450

In order to check any impact by possible biases in taxonomic identification of the fossil records, we analysed the same dataset at the family level. This effectively tested the sensitivity of the developed framework to changes in taxonomic rank of the input data. The relationships between family-level similarity and physical distance from the SCB for Artinskian-Kungurian times (reconstructed at 277 Ma) corroborate the SCB position for reconstruction Y19 as the most supported by faunal data (Fig. 8). The strong correlation between faunal similarity and physical distance for reconstruction Y19 maintains statistical significance across nearly 70% of distance limits and for all  $d_l > 6.250$  km (Fig. 8b). While index JC is not statistically significant at  $d_l = 10.000$  km (Fig. 8c), indices SC and cME both presented stronger relationships than were obtained at the genus-level ( $r_s = -0.5670$  and  $r_s = -0.4559$  respectively; Figs 8d and e). This result is consistent with the genus-level outcome, which indicates that there would be insignificant impact from taxonomic biases in our results. It also shows the framework is robust to changes in taxonomic ranks.

# 5.2.2 Faunal provinciality

Environmental variation across continental blocks can lead to multiple distinct biogeographic regions on a single block. The methods used in this study may lead to inaccuracies when biogeographic regions are selected that are far beyond the distance limit or when multiple regions that should be distinct become combined as one coherent region. This problem is exacerbated in reconstruction W13 with that presents a smaller number of plates and a larger area covered by each plate. Alternate methods to measure distance and define biogeographic regions may be implemented in future work to improve the method used in this study. Rather than selecting plates within a distance range, distances between each fossil occurrence and the SCB could be measured, with only occurrences within the distance limit being selected. This may separate distinct biogeographic regions on large plates, but it also risks splitting could also split blocks (e.g., the Cimmerian terranes in M16, Fig. 1bc). Multivariate statistical methods, such as ordination and cluster analysis (Shi, 1993), are other interesting possibilities to address this issue in future improvements of the framework introduced in this study. These methods may allow for greater resolution in defining biogeographic regions by identifying locations that should be distinct regions rather than defining each region as the entire tectonic block. It has also been found that a distance of 1400 suggested that an open ocean distance of 1400 km may be required between islands to develop a distinct faunal province (Shi, 1996). By taking these into account we could Taking such factors into account would potentially improve results for both large plates with multiple faunal provinces and clusters of islands comprised of individual tectonic blocks that may not be distinct faunal provinces.

The break-up of supercontinent Pangea and the demise of Late Paleozoic glaciation initiated in the Early Permian, generating dynamic climatic episodes during the time interval. The episodes have been interpreted to cause increases in biodiversity and changes in faunal assemblages (Zaffos et al., 2017; Angiolini et al., 2005). These biotic changes may be reflected in the data set considered in this study as it covers a long time period (> 27 Myr) which would lead to the grouping of multiple faunal assemblages that existed in the same faunal province but at distinct time periods with different climates. While all

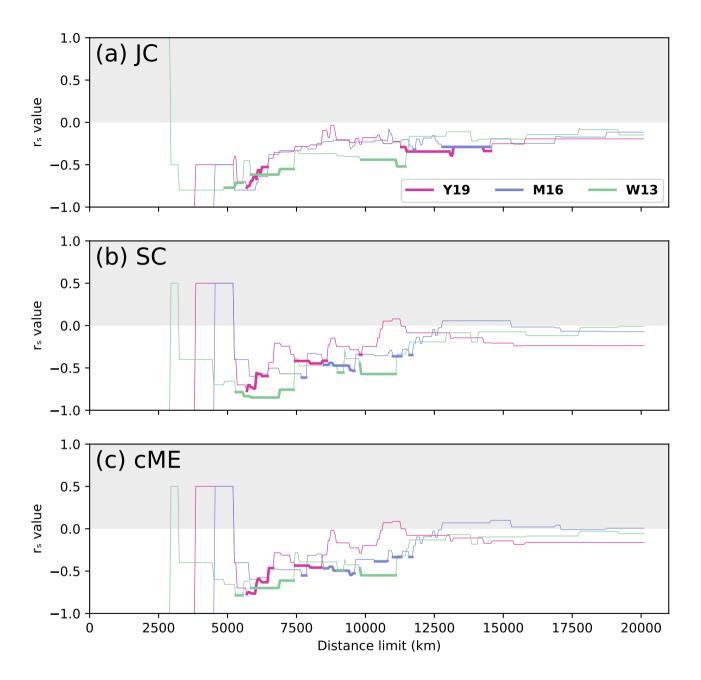


Figure 8. Relationship between faunal similarity and physical distance based on taxonomic data at the family level during Artinskian-Kungurian times (reconstructed at 277 Ma). Results selected based on success: a)  $r_s$  for index SC, all reconstructions and all distance limits, b) statistical significance for index SC, all reconstructions and all distance limits, c-e) relationships between faunal similarity and physical distance from the SCB for each index JC (c), SC (d) and cME (e), for 31 plates within  $d_l = 10,000$  km of the SCB in reconstruction Y19.

three reconstructions have differing All three reconstructions primarily differed in the longitude of continental configurations, most variation is longitudinal differences with corresponding with continental blocks having similar latitudes across each reconstruction (Fig. 1). This keeps continental blocks Continental blocks were in the same biogeographic realms across each reconstruction as they are defined latitudinally (Waterhouse and Bonham-Carter, 1975; Shen et al., 2013), meaning implying that the reconstructions should all be affected equally by this concern. This problem could also be limited in the future by selecting periods during which the climate was relatively stable although this could also reduce the size of the data set and make relationships more difficult to discernfaunal provinciality. Reconstruction W13 places the SCB in the southern hemisphere (Fig. 1e and f), whereas reconstructions M16 and Y19 place it in the northern hemisphere (Fig. 1a, b, c, and d). Despite this discrepancy, the SCB remains within warm water biogeographic realms, suggesting that latitudinal climate variations should not have a major effect on the results.

# 5.2.3 Limitations inherent to each of the three considered biogeographic indexes

455

460

480

485

Index JC measures the true similarity between two sets of fossil occurrences and is less reliable if the real biodiversity is unknown, which is likely the case with paleontological data. Indeed, the validity of index JC in paleobiogeographic studies is questionable (Simpson, 1960). The use of the union between two sets of fossil occurrences is problematic when one set is much larger than the other, and incidentally the SCB is one of the largest biota sets for Early Permian brachiopods, with nearly 150 genera present on the SCB, 3-7 times larger than the average number of genera per plate that the SCB is compared to (Fig. 2b). This effect is apparent in Figure 3 in which  $JC \approx 0$  and  $\ln(JC) \approx -5$  for many plates. In contrast, the absence of plates with  $\ln(SC) \approx -5$  or  $\ln(cME) \approx -5$  indicates that the lowest index values for SC and cME are equal to 0, rather than close to 0 is highlighted by the ideal index values (Fig. 3). Due to these limitations, 6), where the maximum values for index JC is expected to perform poorly relative to the other indexes (e.g. Fig. 5) are 0.19 compared to 1 (for index SC) and 0.51 (for index cME). This reduced range of values for index JC makes the determination of a linear relationship less effective (da Silva and Seixas, 2017), however, the use of the Spearman rank correlation alleviates this issue.

Index SC was developed to extrapolate the true similarity by assuming biodiversity is constant across regions (Simpson, 1960) and thus the calculation is limited by the smaller region. This risks overestimating the true similarity if there are real differences in biodiversity and is susceptible to large variations in value from small variations in set intersection when the smaller biota set is small. Index cME is limited by both of the issues that exist for indexes JC and SC, however, averaging each endemism ratio reduces the impact of each limitation. The Raup-Crick index (Raup and Crick, 1979), which considers how widespread each genus is across a given faunal assemblage and compares similarity between observed sets to similarity between randomly generated sets, could be used in future work as it has shown promise (Schmachtenberg, 2008).

Each of the three indexes indices considered in this study and the relationship to physical distance are subject to real-world complications to in the relationship they measure, typically due to factors such as land barriers, which lead migration distances to to migration distances that are much greater than measured physical distances. This is clearest for the connection between the SCB and North America (Fig. 1), where brachiopods are located on the eastern coast, implying that migratory pathways require travel either across the Panthalassan Ocean or north around Siberia. This leads to strong family-level faunal

similarity (Waterhouse and Bonham-Carter, 1975) because North America and the SCB occupy the same paleo-equatorial region, but limited genera-level similarity due to substantially limited brachiopod dispersal across the Panthalassan Ocean. This is in contrast to the physical distance, as measured by pyGPLATES, which is directly across the Paleo-Tethys Ocean and a much shorter distance than the actual dispersal pathway (Fig. 1). This is supported by the lack of statistical significance for genus-level similarity in reconstruction Y19 at the longest distance limits (Fig. 4). Nevertheless, statistically significant results were obtained at the largest distance limits when using family-level similarity.

#### 5.2.4 Possible interdependence of the considered datasets

490

505

520

While the Early Permian location of the SCB was made determined independently of faunal data in reconstructions W13, M16 and Y19, all three reconstructions used faunal data in the constraint of other tectonic blocks. Faunal data were used to constrain the location of the Australian block by interpreting paleoenvironments in reconstruction W13 (Wright et al., 2013), so that results may not be entirely independent for that plate. In contrast, faunal affinity was only used to determine connected tectonic plates during Devonian times in reconstruction M16 (Matthews et al., 2016), and to determine that Armorica and Perunica two plates were separate micro-continents at the Silurian-Devonian boundary in reconstruction Y19 (Young et al., 2019). We note that the dispersal of brachiopod larvae depends on ocean currents (Lam et al., 2018), which would be different based as well as larva types (?Carlson, 2016), and that ocean currents depend on the location of the SCB and other tectonic blocks in each of the considered tectonic reconstructions (Fig. 1).

The method used to constrain the location of the SCB in M16 is also problematic, as it assumes that Large Low Shear Velocity Provinces (basal mantle structures) have remained stationary at least since the emplacement of the Emeishan LIP (Domeier and Torsvik, 2014), which is unnecessary since given that mantle flow models that predict mobile basal mantle structures are consistent with the record of large volcanic eruptions (Flament et al., 2022) (Flament et al., 2022; Bodur and Flament, 2023; Cucchiaro et al., 202

5.3 Comparison to previous work and possible future directions

Similar quantitative methods have been used to assess relationships between biogeographic indexes and distance (Lees et al., 2002; Belasky (Shi, 1993; Angiolini, 2001; Lees et al., 2002; Belasky et al., 2002; Shen et al., 2013), although on smaller spatial scales. Improved data availability and analytical tools have made these quantitative methods viable for application applicable at the global scale. Here, the SCB was compared to all other tectonic plates, rather than to a selection of geographically close plates.

Comparing Comparison to a small number of plates makes it possible to define a relative position within a region, whereas the comparison to all other plates makes it possible to elect a preferred location relative to global geography.

This work could be extended by testing more reconstructions of Early Permian tectonic plate configuration. Recent reconstructions of the configuration of Tethyan blocks during Early Permian times using faunal affinity suggest significantly different Tethyan arrangements than any of the three reconstructions tested in this study (?Xu et al., 2024). In particular, the three reconstructions tested here show minimal separation of the Cimmerian Terranes from Gondwana at 277 Ma, while faunal affinity suggests some Cimmerian Terranes may have separated from Gondwana during Artinskian times and been located much closer to the SCB

at 277 Ma (?Xu et al., 2024). In contrast, for Asselian-Sakmarian times, the Cimmerian Terranes are reconstructed to similar positions based on faunal affinity (?Xu et al., 2024) as they are in the reconstructions considered in this study (Fig. 1b, d, and f). Testing these differing configurations would be a valuable addition to this work, however, it is not currently possible as reconstructions based on faunal affinity (?Xu et al., 2024) only exist as series of static paleogeographic reconstruction maps, rather than as continuous plate tectonic reconstructions compatible with pyGPlates.

The framework developed here could be used to systematically evaluate the compatibility of tectonic reconstructions with faunal data, or applied to other fields of study such as paleo-oceanography. The framework is highly adaptable and can be applied to any reconstruction, geological period, and faunal data set that provides location data. For example, in this contribution we presented sensitivity analyses for distinct ages and taxonomic levels. Paleobiogeography has been used successfully as a qualitative method to track past ocean currents (Harper et al., 2005), with this framework providing and the framework introduced in this study opens the opportunity to extend these past works to larger scale, quantitative studiesquantitatively and to larger scales. Testing tectonic plate locations could also go beyond testing be applied beyond single tectonic plates, as was done here, by testing faunal affinity for all tectonic plates against all other tectonic plates across geological times. The framework could also be extended to any suitable fossil type. In principle, it should be possible to design a framework in which distances between all tectonic blocks would be a function of faunal similarity, which is one of the observations that led to the establishment of tectonic reconstructions (Jell, 1974). In such a framework, faunal constraints data would be considered alongside geological and paleomagnetic constraints as one method of constraint and validation for plate tectonic configuration alongside other, independent constraints such as paleomagnetic and stratigraphic data.

# 6 Conclusions

525

530

535

540 We assessed the Early Permian longitude of the SCB by using a new framework to evaluate the global relationship between faunal similarity and tectonic distance. We determined that Early Permian longitude the Artinskian-Kungurian position of the SCB in the reconstruction of Young et al. (2019) is more consistent with brachiopod distribution than the Early Permian longitude Artinskian-Kungurian position of the SCB in the reconstruction of Wright et al. (2013) and Matthews et al. (2016). It was found that the The strength of correlation between faunal similarity and physical distance between the SCB and circum-tethyan circum-Tethyan tectonic blocks was greatest-largest for Y19. This implies that the SCB could have been positioned centrally 545 within the Paleo-Tethys Ocean at 103°E, which is 20-30° further west than the other considered reconstructions which tend to in the reconstructions of Wright et al. (2013) and Matthews et al. (2016) that place the SCB on the margin of the ocean. This likely has important implications for Paleo-Tethys Ocean circulation and on for reconstructions of paleoclimate change throughout the Early Permian. The presented framework is also a valuable contribution Early Permian times. While the analysis of the position of the SCB during Asselian-Sakmarian times was inconclusive, it showcases the flexibility of the 550 framework introduced in this contribution. This highlights that the presented framework could be valuable to future studies on of paleogeography as it can be extended to refine global tectonic plate locations based on faunal data.

Code and data availability. All data sets and code used for analysis are openly available from Zenodo (Marks and Flament, 2025) as a collection of Jupyter notebooks that can be used to reproduce the analysis carried out in this contribution, or adapted to perform analysis for other plate tectonic reconstructions, times, or fossil types. Tectonic reconstruction data used within this manuscript are openly available for Y19, M16 and W13 following links provided in Young et al. (2019), Matthews et al. (2016) and Wright et al. (2013), respectively.

Author contributions. RM: Conceptualisation, Methodology, Software, Investigation, Formal analysis, Writing - Original Draft, Writing - Review and Editing, Visualisation; NF: Conceptualisation, Methodology, Software, Writing - Review and Editing, Visualisation, Supervision; SL: Conceptualisation, Methodology, Writing - Review and Editing, Supervision; GRS: Conceptualisation, Writing - Review and Editing

560 Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. N.F. was supported by Australian Research Council project grants LP170100863 and FT230100001, and G.R.S. was supported by Australian Research Council project grant DP230100323.

#### References

- Allison, P. A. and Wells, M. R.: Circulation in large ancient epicontinental seas: What was different and why?, Palaios, 21, 513-515, 2006.
- Angiolini, L.: Lower and Middle Permian brachiopods from Oman and Peri-Gondwanan palaeogeographical reconstructions, in: Brachiopods, pp. 366–376, CRC Press, 2001.
  - Angiolini, L., Brunton, H., and Gaetani, M.: Early Permian (Asselian) brachiopods from Karakorum (Pakistan) and their palaeobiogeographical significance, Palaeontology, 48, 69–86, 2005.
- Belasky, P., Stevens, C. H., and Hanger, R. A.: Early Permian location of western North American terranes based on brachiopod, fusulinid, and coral biogeography, Palaeogeography, Palaeoclimatology, Palaeoecology, 179, 245–266, 2002.
  - Benton, M. J., Wills, M., and Hitchin, R.: Quality of the fossil record through time, Nature, 403, 534–537, 2000.
  - Bodur, Ö. F. and Flament, N.: Kimberlite magmatism fed by upwelling above mobile basal mantle structures, Nature Geoscience, 16, 534–540, 2023.
  - Carlson, S. J.: The evolution of Brachiopoda, Annual Review of Earth and Planetary Sciences, 44, 409-438, 2016.
- 575 Conrad, C. P., Steinberger, B., and Torsvik, T. H.: Stability of active mantle upwelling revealed by net characteristics of plate tectonics, Nature, 498, 479–482, 2013.
  - Crampton, J. S., Beu, A. G., Cooper, R. A., Jones, C. M., Marshall, B., and Maxwell, P. A.: Estimating the rock volume bias in paleobiodiversity studies, Science, 301, 358–360, 2003.
- Cucchiaro, A., Flament, N., Arnould, M., and Cressie, N.: Large volcanic eruptions are mostly sourced above mobile basal mantle structures,

  Communications Earth & Environment, 6, 538, 2025.
  - da Silva, M. and Seixas, T.: The role of data range in linear regression, The Physics Teacher, 55, 371–372, 2017.
  - De Winter, J. C., Gosling, S. D., and Potter, J.: Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data., Psychological methods, 21, 273, 2016.
  - Domeier, M. and Torsvik, T. H.: Plate tectonics in the late Paleozoic, Geoscience Frontiers, 5, 303–350, 2014.
- Dutkiewicz, A., Merdith, A. S., Collins, A. S., Mather, B., Ilano, L., Zahirovic, S., and Müller, R. D.: Duration of Sturtian "Snowball Earth" glaciation linked to exceptionally low mid-ocean ridge outgassing, Geology, 52, 292–296, 2024.
  - Fallaw, W. and Dromgoole, E. L.: Faunal similarities across the South Atlantic among Mesozoic and Cenozoic invertebrates correlated with widening of the ocean basin, The Journal of Geology, 88, 723–727, 1980.
- Fielding, C. R., Frank, T. D., and Isbell, J. L.: Resolving the Late Paleozoic Ice Age in Time and Space, chap. The late Paleozoic ice age—A review of current understanding and synthesis of global climate patterns, Geological Society of America, 2008.
  - Flament, N., Williams, S., Müller, R., Gurnis, M., and Bower, D. J.: Origin and evolution of the deep thermochemical structure beneath Eurasia, Nature communications, 8, 1–9, 2017.
  - Flament, N., Bodur, Ö. F., Williams, S. E., and Merdith, A. S.: Assembly of the basal mantle structure beneath Africa, Nature, 603, 846–851, 2022.
- 595 Golonka, J.: Late Triassic and Early Jurassic palaeogeography of the world, Palaeogeography, Palaeoclimatology, Palaeoecology, 244, 297–307, 2007.
  - Golonka, J., Krobicki, M., Pajak, J., Giang, N., and Zuchiewicz, W.: Global plate tectonics and paleogeography of Southeast Asia, Faculty of Geology, Geophysics and Environmental Protection, AGH University of Science and Technology, Arkadia, Krakow, Poland, 2006.

- Harper, D. A., Alsen, P., Owen, E. F., and Sandy, M. R.: Early Cretaceous brachiopods from North-East Greenland: biofacies and biogeography, Bulletin of the Geological Society of Denmark, 52, 213–225, 2005.
  - Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Fernández del Río, J., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E.: Array programming with NumPy, Nature, 585, 357–362, https://doi.org/10.1038/s41586-020-2649-2, 2020.
- Hohn, M. E.: Binary Coefficients Redux, in: Handbook of Mathematical Geosciences, pp. 143-160, Springer, Cham, 2018.
  - Hunter, J. D.: Matplotlib: A 2D graphics environment, Computing in Science & Engineering, 9, 90–95, https://doi.org/10.1109/MCSE.2007.55, 2007.
  - Jaccard, P.: La distribution de la flore dans la zone alpine, Revue générale des Sciences, pp. 961 967, 1907.
  - Jell, P.: Faunal provinces and possible planetary reconstruction of the Middle Cambrian, The Journal of Geology, 82, 319–350, 1974.
- Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., and Willing, C.: Jupyter Notebooks a publishing format for reproducible computational workflows, in: Positioning and Power in Academic Publishing: Players, Agents and Agendas, edited by Loizides, F. and Schmidt, B., pp. 87 90, IOS Press, 2016.
  Kokoska, S. and Zwillinger, D.: CRC standard probability and statistics tables and formulae, Crc Press, 2000.
- Krivolutskaya, N., Veselovskiy, R., Song, X., Lie-Meng, C., Yu, S.-Y., Gongalsky, B., Bychkova, Y. V., Svirskaya, N., and Smolkin, V.:
   Correlation of intrusive and volcanic rocks in Emeishan Large Igneous Province by geochemical and paleomagnetic data, in: Problems of Geocosmos, pp. 45–53, 2016.
  - Lam, A. R., Stigall, A. L., and Matzke, N. J.: Dispersal in the Ordovician: speciation patterns and paleobiogeographic analyses of brachiopods and trilobites, Palaeogeography, Palaeoclimatology, Palaeoecology, 489, 147–165, 2018.
- Lees, D. C., Fortey, R. A., and Cocks, L. R. M.: Quantifying paleogeography using biogeography: a test case for the Ordovician and Silurian of Avalonia based on brachiopods and trilobites, Paleobiology, 28, 343–363, 2002.
  - Lehmann, J., Schulmann, K., Lexa, O., Corsini, M., Kroener, A., Štípská, P., Tomurhuu, D., and Otgonbator, D.: Structural constraints on the evolution of the Central Asian Orogenic Belt in SW Mongolia, American Journal of Science, 310, 575–628, 2010.
  - Marchetti, L., Forte, G., Kustatscher, E., DiMichele, W. A., Lucas, S. G., Roghi, G., Juncal, M. A., Hartkopf-Fröder, C., Krainer, K., Morelli, C., et al.: The Artinskian Warming Event: an Euramerican change in climate and the terrestrial biota during the early Permian, Earth-Science Reviews, 226, 103 922, 2022.
  - Marks, R. and Flament, N.: Analysis scripts and datasets for Marks et al. "Early Permian longitudinal position of the South China Block from brachiopod paleobiogeography", Zenodo, https://doi.org/10.5281/zenodo.14961181, 2025.
  - Matthews, K. J., Maloney, K. T., Zahirovic, S., Williams, S. E., Seton, M., and Müller, R. D.: Global plate boundary evolution and kinematics since the late Paleozoic, Global and Planetary Change, 146, 226–250, https://doi.org/10.1016/j.gloplacha.2016.10.002, 2016.
- McKinney, W.: Data Structures for Statistical Computing in Python, in: Proceedings of the 9th Python in Science Conference, edited by Stéfan van der Walt and Jarrod Millman, pp. 56 61, https://doi.org/10.25080/Majora-92bf1922-00a, 2010.
  - Müller, R. D., Sdrolias, M., Gaina, C., Steinberger, B., and Heine, C.: Long-term sea-level fluctuations driven by ocean basin dynamics, science, 319, 1357–1362, 2008.
- Müller, R. D., Cannon, J., Qin, X., Watson, R. J., Gurnis, M., Williams, S., Pfaffelmoser, T., Seton, M., Russell, S. H., and Zahirovic, S.:

  GPlates: Building a virtual Earth through deep time, Geochemistry, Geophysics, Geosystems, 19, 2243–2261, 2018.
  - Paleobiology Database: https://paleobiodb.org/navigator/, accessed: 29-07-2021.

- Piccoli, G., Sartori, S., Franchino, A., Pedron, R., Claudio, L., and Natale, A. R.: Mathematical model of faunal spreading in benthic palaeobiogeography (applied to Cenozoic Tethyan molluscs), Palaeogeography, Palaeoclimatology, Palaeoecology, 86, 139–196, 1991.
- Raup, D. M.: Species diversity in the Phanerozoic: an interpretation, Paleobiology, 2, 289–297, 1976.
- Raup, D. M. and Crick, R. E.: Measurement of faunal similarity in paleontology, Journal of Paleontology, pp. 1213–1227, 1979.
  - Schmachtenberg, W. F.: Resolution and limitations of faunal similarity indices of biogeographic data for testing predicted paleogeographic reconstructions and estimating intercontinental distances: a test case of modern and Cretaceous bivalves, Palaeogeography, Palaeoclimatology, Palaeoecology, 265, 255–261, 2008.
  - Scotese, C. R.: A continental drift flipbook, The Journal of Geology, 112, 729–741, 2004.
- Scotese, C. R., Song, H., Mills, B. J., and van der Meer, D. G.: Phanerozoic paleotemperatures: The earth's changing climate during the last 540 million years, Earth-Science Reviews, 215, 103 503, 2021.
  - Seton, M., Müller, R. D., Zahirovic, S., Gaina, C., Torsvik, T., Shephard, G., Talsma, A., Gurnis, M., Turner, M., Maus, S., et al.: Global continental and ocean basin reconstructions since 200 Ma, Earth-Science Reviews, 113, 212–270, 2012.
  - Sheehan, P. M.: A reflection of labor by systematists?, Paleobiology, 3, 325–328, 1977.
- 650 Shen, S., Zhang, H., Zhang, Y., Yuan, D., Chen, B., He, W., Mu, L., Lin, W., Wang, W., Chen, J., et al.: Permian integrative stratigraphy and timescale of China, Science China Earth Sciences, 62, 154–188, 2019.
  - Shen, S.-z., Zhang, H., Shi, G. R., Li, W.-z., Xie, J.-f., Mu, L., and Fan, J.-x.: Early Permian (Cisuralian) global brachiopod palaeobiogeography, Gondwana Research, 24, 104–124, 2013.
- Shi, G. R.: Multivariate data analysis in palaeoecology and palaeobiogeography—a review, Palaeogeography, palaeoclimatology, palaeoecology, 105, 199–234, 1993.
  - Shi, G. R.: A model of quantitative estimate of marine biogeographic provinciality-a case study of the permian marine biogeographic provincialism, Acta Geologica Sinica, 1996.
  - Shi, G. R.: 4.2. 3 Palaeobiogeography of Marine Communities, Palaeobiology II, 2, 440, 2001a.

- Shi, G. R.: Terrane rafting enhanced by contemporaneous climatic amelioration as a mechanism of vicariance: Permian marine biogeography of the Shan-Thai terrane in Southeast Asia, Historical Biology, 15, 135–144, 2001b.
- Simpson, G. G.: Notes on the measurement of faunal resemblance, American Journal of Science, 258, 300–311, 1960.
- Torsvik, T. H. and Van der Voo, R.: Refining Gondwana and Pangea palaeogeography: estimates of Phanerozoic non-dipole (octupole) fields, Geophysical Journal International, 151, 771–794, 2002.
- Torsvik, T. H., Steinberger, B., Cocks, L. R. M., and Burke, K.: Longitude: linking Earth's ancient surface to its deep interior, Earth and Planetary Science Letters, 276, 273–282, 2008.
  - Torsvik, T. H., Burke, K., Steinberger, B., Webb, S. J., and Ashwal, L. D.: Diamonds sampled by plumes from the core–mantle boundary, Nature, 466, 352–355, 2010.
  - Valentine, J. W.: Numerical analysis of marine molluscan ranges on the extratropical northeastern Pacific Shelf 1, Limnology and oceanography, 11, 198–211, 1966.
- 670 Valentine, J. W.: Plate tetonics and shallow marine diversity and endemism, an actualistic model, Systematic Zoology, 20, 253–264, 1971.
  - Valentine, J. W. and Moores, E. M.: Plate-tectonic Regulation of Faunal Diversity and Sea Level: a Model, Nature, 228, 657–659, https://doi.org/10.1038/228657a0, 1970.
  - Valentine, J. W. and Moores, E. M.: Global tectonics and the fossil record, The Journal of Geology, 80, 167–184, 1972.

- VanVoorhis, C. W., Morgan, B. L., et al.: Understanding power and rules of thumb for determining sample sizes, Tutorials in quantitative methods for psychology, 3, 43–50, 2007.
  - Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.
  - Waterhouse, J. and Bonham-Carter, G.: Global distribution and character of Permian biomes based on brachiopod assemblages, Canadian Journal of Earth Sciences, 12, 1085–1146, 1975.

- Williams, S., Cannon, J., Qin, X., and Müller, D.: PyGPlates-a GPlates Python library for data analysis through space and deep geological time, in: EGU General Assembly Conference Abstracts, p. 8556, 2017.
- Wright, N., Zahirovic, S., Müller, R. D., and Seton, M.: Towards community-driven paleogeographic reconstructions: integrating open-access paleogeographic and paleobiology data with plate tectonics, Biogeosciences, 10, 1529–1541, https://doi.org/10.5194/bg-10-1529-2013, 2013.
  - Xu, H.-P., Zhang, Y.-C., Zhang, Y.-J., Qiao, F., and Shen, S.-Z.: Lower Cisuralian brachiopod faunas from the Lhasa and South Qiangtang blocks in Tibet and their biostratigraphical and palaeobiogeographical implications, Palaeoworld, 33, 679–705, 2024.
- 690 Young, A., Flament, N., Maloney, K., Williams, S., Matthews, K., Zahirovic, S., and Müller, R. D.: Global kinematics of tectonic plates and subduction zones since the late Paleozoic Era, Geoscience Frontiers, 10, 989–1013, https://doi.org/10.1016/j.gsf.2018.05.011, 2019.
  - Zaffos, A., Finnegan, S., and Peters, S. E.: Plate tectonic regulation of global marine animal diversity, Proceedings of the National Academy of Sciences, 114, 5653–5658, 2017.
- Zahirovic, S., Müller, R. D., Seton, M., and Flament, N.: Tectonic speed limits from plate kinematic reconstructions, Earth and Planetary Science Letters, 418, 40–52, 2015.
  - Zhong, Y.-T., He, B., Mundil, R., and Xu, Y.-G.: CA-TIMS zircon U–Pb dating of felsic ignimbrite from the Binchuan section: Implications for the termination age of Emeishan large igneous province, Lithos, 204, 14–19, 2014.