*Abbrev. Journal Name* (2018) **0**, 1–39

# Methods for Producing a Reliable APWP

Chenjian Fu[1⋆] Chris Rowan[2]
[1] *Department of Geology, Kent State University, 325 S Lincoln St, Kent, OH* 44242*, USA*
[2] *Department of Geology, Kent State University, 325 S Lincoln St, Kent, OH* 44242*, USA*

Received 2018 December 28; in original form 2018 November 22

**SUMMARY**
First, new picking/weighting methods developed here and previously published picking/weighting methods are compiled together to generate 168 different paleomagnetic APWPs. Then, the APWP similarity measuring tool is used to find which method$s$ is$(are)$ good or bad. The final results tell us that the "Age Position Picking (APP)" method is better than the "Age Mean Picking (AMP)" method for making a reliable paleomagnetic APWP and weighting is actually unnecessary.

**Key words:** Moving Average – Weighting – APWP – Paleomagnetism.

## 1 INTRODUCTION

APWPs are generated by combining paleomagnetic poles, also known as paleopoles, for a particular rigid block over the desired age range to produce a smoothed path. See the Appendix A for some examples how the paleopole datasets are constrained for a particular tectonic plate during a specific time interval.

### 1.1 Not All Data Are Created Equal

However, uncertainties in the age and location of paleopoles can vary greatly for different poles.

#### *1.1.1 Age Error*

Although remanent magnetizations are generally assumed to be primary, many events can cause remagnetisation (in which case the derived pole is 'younger' than the rock). If an event that has occurred since the rock's formation that should affect the magnetisation (e.g., folding, thermal overprinting due to intrusion) can be shown to have affected it, then it constrains the magnetisation to have been acquired before that event. Recognising or ruling out remagnetisations depends on these field tests, which are not always performed or possible. Even a passed field test may not be useful if field test shows magnetisation acquired prior to a folding event tens of millions of years after initial rock formation.

The most obvious characteristic we can observe from paleomagnetic data is that some poles have very large age ranges, e.g., more than 100 Myr. The magnetization age should be some time between the information of the rock and folding events. There are also others where we have similar position but the age constraint is much narrower, e.g. 10 Myr window or less. Obviously the latter kind of data is more valuable than the one with large age range.

⋆ Email: cfu3@kent.edu

#### *1.1.2 Position Error*

The errors of pole latitudes and longitudes are 95% confidence ellipses, which also vary greatly in magnitude. All paleopoles have some associated uncertainties due to measurement error and the nature of the geomagnetic field. More uncertainties can be added by too few samples, sampling spanning too short a time range to approximate a GAD field, failure to remove overprints during demagnetisation, etc.

#### *1.1.3 Data Consistency*

Paleopoles of a rigid plate or block should be continuous time series. For a rigid plate, two poles with similar ages shouldn't be dramatically different in location. Sometimes, this is the case. Sometimes we have further separated poles with close ages.

There are a number of possible causes for these outliers, including:

`Lithology`

For poor consistency of data, it is potentially because of different inclinations or declinations. The first thing we should consider about is their lithology. We want to check if the sample rock are igneous or sedimentary, because sediment compaction can result in anomalously shallow inclinations (Tauxe et al. 2019). In addition, we also can check if the rock are redbeds or non-redbeds. Although whether redbeds record a detrital signal or a later Chemical Remanent Magnetization (CRM) is still somewhat controversial, both sedimentary rocks and redbeds could lead to inconsistency in direction compared to igneous rocks.

`Local Rotations`

Local deformation between two paleomagnetic localities invalidates the rigid plate assumption and could lead to inconsistent VGP directions. So if discordance is due to local deformation, and we would ideally want to exclude or correct (if possible) such poles from our APWP calculation.

`Other Factors`

In most cases, mean pole age (centre of age error) has just been binned. If any of the poles have large age errors, they could be different ages from each other and sample entirely different parts of the APWP. Conversely, if any of the poles have too few samples, or were not sampled over enough time to average to a GAD field, a discordant pole may be due to unreduced secular variation.

#### *1.1.4 Data Density*

As we go back in time, we have lower quality and lower density (or quantity) of data, for example, the Precambrian or Early Paleozoic paleomagnetic data are relatively fewer than Middle-Late Phanerozoic ones, and most of them are not high-quality, e.g., larger errors in both age and location. The combination of lower data quality with lower data density means that a single 'bad' pole (with large errors in age and/or location) can much more easily distort the reconstructed APWP, because there are few or no 'good' poles to counteract its influence.

Data density also varies between different plates. E.g., we have a relatively high density of paleomagnetic data for North American Craton (NAC), but few poles exist for Greenland and Arabia. Based on mean age (mean of lower and upper magnetic ages), for 120–0 Ma, the **Global Paleomagnetic Database** (GPMDB) version 4.6b (McElhinny & Lock 1996; Pisarevsky 2005, updated in 2016 by the Ivar Giaever Geomagnetic Laboratory team, in collaboration with Pisarevsky) has more than 130 poles for NAC, but only 17 for Greenland and 24 for Arabia.

#### *1.1.5 Publication Year*

The time when the data was published should also be considered, because magnetism measuring methodology, technology and equipments have been improved since the early 20th century. For example, stepwise demagnetisation, which is the most reliable method of detecting and removing secondary overprints, has only been in common use since the mid 1980s.

In summary, not all paleopoles are created equal, which leads to an important question: how to best combine poles of varying quality into a coherent and accurate APWP?

### 1.2 Existing Solutions and General Issues

Paleomagnetists have proposed a variety of methods to filter so-called "bad" data, or give lower weights to those "bad" data before generating an APWP, e.g., two widely used methods: the V90 reliability criteria (van der Voo 1990) and the BC02 selection criteria provided by Besse & Courtillot (2002). Briefly, the V90 criteria for paleomagnetic results includes seven criteria: (1) Well determined age; (2) At least 25 samples with Fisher (Fisher 1953) precision $\kappa$ greater than 10 and $\alpha95$ less than 16°; (3) Detailed demagnetisation results reported; (4) Passed field tests; (5) Tectonic coherence with continent and good structural control; (6) Identified antipodal reversals; (7) Lack of similarity with younger poles (Torsvik et al. 1992). The total criteria satisfied (0–7) is then used as a measure of a paleomagnetic result's overall reliability, which is known as Q (quality) factor (Torsvik et al. 1992). Q factor is indeed a very straightforward way to get a quantitalized reliability score. Also it then can be conveniently used in the later calculations of APWPs (Torsvik et al. 1992). But at the same time this is a fairly basic filter that lumps together criteria that may not be equally important. Compared with V90, the BC02 criteria suggests stricter filtering, e.g., using only poles with at least 6 sampling sites and 36 samples, each site having $\alpha95$ less than 10° in the Cenozoic and 15° in the Mesozoic. B02 is also straightforward and convenient to use, but some useful data may be filtered out and wasted especially for a period where there are only limited number of data. In addition, there has been limited study of how effective these marking/filtering methods are at reconstructing a 'true' APWP, and for most studies after a basic filtering of 'low quality' poles, the remaining poles are, in fact, treated equally.

Above all, there haven't been any real attempts to study how APWP fits may be improved by filtering/weighting data. This paper is presented to address these issues.

**Table 1.** List of the used fields and field codes of the GPMDB.

| No. | Weighting Algorithm |
|---|---|
| LOMAGAGE | Lower best estimate of the magnetic age of the magnetisation component |
| HIMAGAGE | Upper best estimate of the magnetic age of the magnetisation component |
| B | Number of sites |
| N | Number of samples |
| ED95 | Radius of circle of 95% confidence about mean direction, i.e. $\alpha95$ |
| KD | Fisher precision parameter for mean direction |
| DP | Half-angle of confidence on the pole in the direction of paleomeridian |
| DM | Half-angle of confidence on the pole perpendicular to paleomeridian |
| K_NORM | Fisher precision parameter for Normal directions |

## 2 METHODS

### 2.1 General Approach

In this study, we use paleopoles extracted from the GPMDB to generate APWPs for the period 120–0 Ma. A range of possible APW paths for North America, India and Australia can be generated from the extracted sets of paleopoles using various binning, filtering and weighting methods (Tables. 2 and 3). These paths can then be compared to synthetic APWPs independently generated from an absolute plate motion model. The three plates chosen have different attributes, both in terms of the input data set and the nature of the reference APWP.

### 2.2 Paleomagnetic Data

#### *2.2.1 Used Specific Field Codes of GPMDB*

Data analysis includes a tremendous amount of manipulation of data fields/columns in the GPMDB. In the following content, several specific fields and field codes will be referred to. They are listed below for easy reference.

#### *2.2.2 Paleomagnetic Data of Three Representative Continents*

Collections of paleopoles with a minimum age (LOMAGAGE) $\leq$135 Ma for the North American (Plate ID 101), the Indian (ID 501) and Australian (ID 801) plates, were extracted from the GPMDB. In order to include valid paleopoles from blocks that moved independently prior to 120 Ma, which therefore have different assigned plate codes in the GPMDB, the spatial join technique (Jacox & Samet 2007) was used to find all poles within the geographic region that defines the rigid plate within the period of interest (see also Appendix A):

*for North America,* the search region was defined by the North American (ID 101), Avalon/Acadia (ID 108) and Piedmont (ID 109) blocks, as defined by the recently published plate model of (Young et al. 2018). Following extraction, 58 poles from southwestern North America that have been affected by regional rotations since 36 Ma (McQuarrie & Wernicke 2006), were removed. The final dataset consists of 135 paleopoles (Fig. 1), with 76 (about 56.3%) sampled from dominantly igneous sequences; 56 (about 41.5%) sampled from mostly sedimentary sequences, including 6 from redbeds); and 3 (about 2.2%) from metamorphic sequences. The principal features of the age distribution are a larger number of young (<5 Ma) poles, and relatively fewer poles in the Late Cretaceous and Miocene.

*for India,* the Indian block (ID 501) as defined by (Young et al. 2018) was used, but following extraction 31 paleopoles associated with parts of the northern margin that have undergone regional rotations since the Jurassic (Gaina et al. 2015) were removed. The final dataset consists of 75 paleopoles (Fig. 2), with 39 (52%) sampled from dominantly igneous sequences and 36 (48%) sampled from mostly sedimentary sequences, including 3 from redbeds). There is a high concentration of poles from the latest Cretaceous–Early Cenozoic (c.70–60 Ma), many of which are igneous; in younger and older intervals, there are fewer, mostly sedimentary poles.

*for Australia,* the Australia (ID 801), Sumba (ID 675), and Timor (ID 684) blocks as defined by (Young et al. 2018) were used, in combination with data from the Tasmania block (ID 805) younger than c. 100 Ma (with a maximum age (HIMAGAGE) $\leq$100 Ma), prior to which it was not fixed with respect to Australia (Young et al. 2018). The final dataset consists of 99 paleopoles (Fig. 3), with 61 (61.6%) sampled from dominantly igneous sequences and 38 (38.4%) sampled from mostly sedimentary sequences, including 9 from redbeds). The temporal distribution of poles is relatively uniform.

Compared with North American (Fig. 1) and Australian (Fig. 3) paleopoles, Indian paleopoles are in relatively lower density in general except during the period of about 70–60 Ma (Fig. 2).

### 2.3 APWP Generation

Multiple APWPs were generated using the selected poles for each of the three plates as follows:

*Picking/binning.* A moving average or moving window technique was used: paleopoles were selected for each APWP time step (initially 5 Myr step length from 0 to 120 Ma) if their age fell within a window centered on the current step age. In this study, the width of the moving window was always twice that of time step (i.e. initially 10 Myr), such that each window half-overlaps with its neighbours.

*Filtering.* Poles with characteristics thought to correspond to poor data quality, or lacking characteristics thought to correspond to good data quality, were discarded (or in some cases, corrected).

*Weighting.* Calculation of a weighted Fisher mean (Fisher 1953) of the remaining poles within each window, using weighting functions intended to increase the influence of higher quality poles relative to lower quality ones.

28 different picking and filtering algorithms were tested (Table. 2), in combination with 6 different weighting algorithms (Table. 3), for the three plates. The effects of changing the time step length and width of the moving window, and the reference path, were also examined.

#### 2.3.1 *Picking/Binning*

In studies where the moving window method is used to calculate an APWP (Torsvik & Smethurst 1999; Torsvik et al. 2008), a paleopole is generally considered to fall in the current window only if the mid-point of its age limits fall within that window. If paleopole has a large age uncertainty compared to the size of the moving window, it will not be included in the moving windows close to the beginning and end of the age range, which are arguably more likely magnetisation ages than the mid-point. To investigate this issue, we compare the performance of moving windows populated using the mid-point picking criterion, referred to hereafter as "Age Mean Picking" (AMP; even-numbered algorithms in Table. 2, Fig. 8 and subsequent figures), to a less restrictive picking criterion where a paleopole is included in the current moving window if any part of its age limits falls within that window, referred to hereafter as "Age Position Picking" (APP; odd-numbered algorithms in Table. 2, Fig. 8 and subsequent figures). The APP algorithm will pick more paleopoles in each moving window than the AMP algorithm (Figs. 1-3; Fig. 4).

If, for example, we have a paleopole which is constrained to within 10 and 20 Ma of age, and we have a 2 Myr moving window with a 1 Myr age step, then it is included just in the 14–16 Ma window (for the mid-point age of 15 Ma) for AMP. For APP, this paleopole falls in the 9–11, 10–12, 11–13, 12–14 . . . 17–19, 18–20 and 19–21 Ma windows. So the average poles are produced from each window, and each original paleopole is represented over its entire possible acquisition age.

A shorter step and narrower window will potentially increase the clustering of the selected paleopoles, but will reduce their number. Conversely, a longer step and wider window will increase the number of poles, but potentially decrease their clustering. To investigate these trade-offs, every picking/filtering and weighting method was also used to generate APWPs with a time step and window doubled to 10 Myr and 20 Myr, respectively. Paths generated using AMP and APP with no filtering, and every weighting method, with time steps from 1 Myr to 15 Myr in 1 Myr increments and window widths from 2 Myr to 30 Myr in 2 Myr increments, were also analysed. In all cases the oldest time step was 120 Ma.

#### 2.3.2 *Filtering*

14 different filters or corrections (Table. 2) were applied to both data picked using the AMP moving window method (even numbers) and data picked using the APP moving window method (odd numbers), resulting in a total of 28 unique picking algorithms. The filters or corrections can be characterised as follows:

*No modification (method 0/1).*

*Removal of poles with large spatial and temporal uncertainties (method 2/3).* Paleopoles with both large $\alpha95$ (ED95 $> 15^\circ$, following the BC02 threshold for the Mesozoic) and wide acquisition age limits (difference between HIMAGAGE and LOMAGAGE $>$ 20 Myr, following the V90 criteria about age within a

**Figure 1.** Temporal distribution of 120–0 Ma $NAC$ (101) paleopoles in 10 Myr window length and 5 Myr step length. For distribution a, each bin only counts in the midpoints (circles) of pole error bars (not including those right at bin edges); For distribution b, as long as the bar intersects with the bin (not including those intersecting only at one of bin edges), it is counted in. Inside the parentheses, i means igneous rocks derived (red bars; only two poles, 83–77 Ma and 80–65 Ma, from igneous and also sedimentary; only one pole, 72–40 Ma, from igneous and also metamorphic), r means sedimentary rocks with redbeds involved derived (orange bars), and m means metamorphic rocks derived (blue bars); the left are non-redbed sedimentary rocks derived (black bars; only two poles, 146–65 Ma [RESULTNO 6679] and 2–0 Ma [RESULTNO 1227], are from sedimentary and also metamorphic). The data published before 1984 are shown as circles with a dot.

**Figure 2.** Temporal distribution of 120–0 Ma Indian (501) paleopoles. For red bars, only one pole, 67–64 Ma (RESULTNO 8593), is from igneous and also sedimentary. See Fig. 1 for more information.

**Figure 3.** Temporal distribution of 120–0 Ma Australian (801) paleopoles. For black bars, only four poles, 100–80 Ma (RESULTNO 1106), 10–2 Ma (RESULTNO 1208), 4–2 Ma (RESULTNO 140) and 1–0 Ma (RESULTNO 1963), are from sedimentary and also igneous. For red bars, only one pole, 65–25 Ma (RESULTNO 1872), is from igneous and also sedimentary rocks, and only one pole, 1–0 Ma (RESULTNO 1147), is from igneous and also metamorphic rocks. See Fig. 1 for more information.

**Figure 4.** An example of 10 Myr moving window and 5 Myr step in the two moving average methods, AMP and APP, based on poles of the $NAC$. White circles are the midpoints of low and high magnetic ages. The vertical axis has no specific meaning here. For example, for the window of 15 Ma to 5 Ma (the dashed-line bin), the AMP method calculates the Fisher mean pole (dark triangle in Fig. 5) of only 5 paleopoles, while the APP method calculates the mean pole (dark circle in Fig. 5) of 9 paleopoles.

half of a geological period; the average of the geological periods between 120 and 0 Ma [Quaternary, Neogene, Paleogene and Cretaceous] is about 20 Myr) which are less likely to provide a good estimate of the actual pole position within any specific age window, were excluded.

*Prefer poles from igneous rocks (methods 4/5, 6/7).* Method 4/5 removes paleopoles potentially affected by inclination flattening by selecting only paleopoles coded as igneous or mostly igneous (ROCKTYPE starting with "intrusive" or "extrusive"). In fact, most of the paleopoles picked by method 4/5 are derived from igneous-only rocks. Method 6/7 selects paleopoles coded as containing igneous (ROCKTYPE containing "intrusive" or "extrusive"); this is a less strict filter, because the dominant rock type could potentially be another lithology. So method 6 includes also poles from method 4, and method 7 includes poles from method 5.

*Correct sedimentary poles for inclination shallowing (method 8/9).* Rather than excluding paleopoles from sedimentary rocks, paleopoles coded as sedimentary or redbeds were instead corrected for inclination flattening using the flattening function $\tan I_o = f \tan I_f$ (King 1955), where $I_o$ is the observed inclination, $I_f$ is the unflattened inclination, and $f$ is the flattening factor (also known as shallowing coefficient; 1=no flattening, 0=completely flattened). Here $f = 0.6$ is used in all cases, following (Torsvik et al. 2012), unless when the rock type (ROCKTYPE field in the database) is not sedimentary dominated but contains sedimentary, $f = 0.8$ is used instead, following the minimum anisotropy-of-thermal-remanence determined f-correction (Domeier et al. 2011a; Domeier et al. 2011b).

**Table 2.** List of all Picking/Binning algorithms developed here.

| No. | Picking Algorithm |
|---|---|
| 0 | AMP: Age Mean Picking, see Section "APWP Generation" |
| 1 | APP: Age Position Picking |
| 2 | AMP ("$\alpha$95/Age range" no more than "15/20") |
| 3 | APP ("$\alpha$95/Age range" no more than "15/20") |
| 4 | AMP (mainly or only igneous) |
| 5 | APP (mainly or only igneous) |
| 6 | AMP (contain igneous and not necessarily mainly) |
| 7 | APP (contain igneous and not necessarily mainly) |
| 8 | AMP (unflatten sedimentary) |
| 9 | APP (unflatten sedimentary) |
| 10 | AMP (nonredbeds) |
| 11 | APP (nonredbeds) |
| 12 | AMP (unflatten redbeds) |
| 13 | APP (unflatten redbeds) |
| 14 | AMP (published after 1983) |
| 15 | APP (published after 1983) |
| 16 | AMP (published before 1983) |
| 17 | APP (published before 1983) |
| 18 | AMP (exclude commented local rot or secondary print) |
| 19 | APP (exclude commented local rot or secondary print) |
| 20 | AMP (exclude local rot or correct it if suggested) |
| 21 | APP (exclude local rot or correct it if suggested) |
| 22 | AMP (filtered using SS05 palaeomagnetic reliability criteria) |
| 23 | APP (filtered using SS05 palaeomagnetic reliability criteria) |
| 24 | AMP (exclude superseded data already included in other results) |
| 25 | APP (exclude superseded data already included in other results) |
| 26 | AMP (comb of 22 and 24) |
| 27 | APP (comb of 23 and 25) |

Notes: SS05, (Schettino & Scotese 2005)

*Remove redbeds (method 10/11) or correct them for inclination shallowing (method 12/13).* Bias toward shallow inclinations is also observed in paleomagnetic data derived from redbeds (Tauxe & Kent 2004; Krijgsman & Tauxe 2004; Tan et al. 2007; Bilardello & Kodama 2010, e.g., in central Asia, Mediterranean region, North America, etc.). This bias can be addressed by removing the source (method 10/11; ROCKTYPE containing redbeds), or correcting for inclination flattening, setting $f = 0.6$ as previously described (method 12/13). In the latter case, the assumption is being made that the redbeds are carrying a detrital paleomagnetic signal.

*Prefer poles from younger (methods 14/15, 24/25) or older (method 16/17) studies.* Advancements in equipment (e.g., cryogenic magnetometers) and analytical techniques (e.g., stepwise demagnetisation) mean that more recently published paleopoles are potentially more reliable than older ones. Method 14/15 removes any paleopoles published prior to 1983 (YEAR $>$ 1983)—the mean publication date for paleopoles in the GPMDB. Method 24/25 takes a similar but less aggressive approach by excluding paleopoles that have been superseded (99 datasets) by later studies from the same sequence, which are presumed to represent a more accurate determination of the paleopole position. Conversely, removing pale-

opoles published after 1983 (method 16/17; YEAR $\leq$ 1983) should have a negative effect.

*Exclude suspected local rotations and secondary overprints (method 18/19), or correct for them where possible (method 20/21).* Secondary remanence components and local tectonic deformation can both displace the measured pole position away from its "true" position. Such poles can be identified based on demagnetisation data, or comparison to the pre-existing APWP. Method 18/19 removes paleopoles that were identified as such in the COMMENTS field (all the paleopoles affected by local rotations are picked out by carefully going through all the datasets, including two groups with [19 datasets] and without [47 datasets] suggested corrections; the paleopoles affected by secondary overprints are extracted with the keyword "econd" contained in the COMMENTS). A subset (19 datasets) of these paleopoles have a suggested correction associated with them; method 20/21 retains these paleopoles after applying the suggested correction.

*SS05 quality criteria (methods 22/23).* As with method 2/3, SS05 (Schettino & Scotese 2005) removes paleopoles with high spatial ($\alpha95 > 15°$) and temporal (age range $>$ 40 Myr) uncertainty, but additionally remove paleopoles where samples had poor sampling coverage (sampling sites' quantity [B] of $< 4$, samples' quantity [N] of $< 4$ times of the sites [B]) and were not subjected to even a blanket demagnetisation treatment (laboratory cleaning procedure code DEMAGCODE $< 2$). Method 26/27 also uses these criteria, but further excludes superseded data.

### 2.3.3 Weighting

Following filtering, weights were assigned to each of the remaining paleopoles using one of the following six algorithms (Table 3), prior to calculation of a weighted Fisher mean:

*No weighting (Weighting 0).* Weighting factor=1 for all paleopoles.

*Weighting by sample and site number (Weighting 1).* Paleopoles derived from more individually oriented samples (observations; N) collected from more sampling levels/sites (B) are more likely to average out secular variation and accurately sample the GAD field (Tauxe et al. 2019; van der Voo 1990; Besse & Courtillot 2002), and are given a weighting closer to 1. Unfortunately, in the GPMDB, not all paleopoles' B or N are given. There are datasets with only number of sampling sites (B; at least greater than 1) given, but no number of samples (N) or only one sample given, so for this case, if B>1 and N $\leq$ 1, weight=$(1-\frac{1}{B})*0.5$ (for 120–0 Ma North America, there are 8 such datasets, India 4, Australia 1). If only N (at least greater than 1) is given, and B is missing or only one, i.e. B$\leq$ 1 and N>1, weight=$(1-\frac{1}{N})*0.5$ (for 120–0 Ma North America, there are 20 such datasets, India 26, Australia 22). If B $\leq$ 1 and N $\leq$ 1 (there are only 23 datasets for the whole GPMDB 4.6b, including 18 with both B and N informations missing; for 120–0 Ma North America, India and Australia, there is no such dataset), weight=0.2.

*Weighting by age uncertainty (Weighting 2)* Above a maximum age range that represents a well-constrained age, defined as half of each geological period in the Phanerozoic Eon (e.g., Quaternary, Neogene) (van der Voo 1990; Tauxe et al. 2019) or 15 Myr (the halves of the Paleogene, Cretaceous, and Jurassic periods are all at least 20 Myr, which is large for these relatively young geologic periods), whichever is smaller, paleopoles are given an increasingly small weight as the age uncertainty (the high magnetic age $-$ the low magnetic age) increases.

**Table 3.** List of all weighting algorithms developed here.

| No. | Weighting Algorithm |
|---|---|
| 0 | None (No weighting) |
| 1 | Larger numbers of sites (B) & observations (N), greater $weight$ ($w$): $w = \begin{cases} 0.2, \text{ if both B \& N are missing, or B} \leq 1 \text{ \& N} \leq 1; \\ (1-\frac{1}{B})*0.5, \text{ if only N is missing, or N} \leq 1 \text{ \& B>1;} \\ (1-\frac{1}{N})*0.5, \text{ if only B is missing, or B} \leq 1 \text{ \& N>1;} \\ (1-\frac{1}{B})*(1-\frac{1}{N}), \text{ if B>1 \& N>1.} \end{cases}$ |
| 2 | Lower age uncertainty, greater weight:<br>age_range=HIMAGAGE-LOMAGAGE<br>age_midpoint=(HIMAGAGE+LOMAGAGE)*0.5<br>if age_midpoint<2.58 (Ma; start of the Quaternary, according to GSA Geologic Time Scale),<br>$w = \begin{cases} 1, \text{ if age_range} \leq 1.29 \text{ (from } \frac{2.58-0}{2}) \\ \frac{1.29}{age_range}, \text{ if age_range>1.29;} \end{cases}$<br>if 2.58$\leq$age_midpoint<23.03 (Ma; Neogene),<br>$w = \begin{cases} 1, \text{ if age_range} \leq 10.225 \text{ (from } \frac{23.03-2.58}{2}) \\ \frac{10.225}{age_range}, \text{ if age_range>10.225;} \end{cases}$<br>if 23.03$\leq$age_midpoint<201.3 (Ma; Paleogene,Cretaceous,Jurassic),<br>$w = \begin{cases} 1, \text{ if age_range} \leq 15 \\ \frac{15}{age_range}, \text{ if age_range>15.} \end{cases}$ |
| 3 | Lower $\alpha95$, greater weight:<br>Positive half Normal distribution with a mean and standard deviation of 0 and 10, scaled with $10\sqrt{2\pi}$ (to make the peak reach 1)<br>$w = e^{-\frac{\alpha_{95}^2}{200}},$<br>where<br>$\alpha95 = \begin{cases} ED95 \\ DP, \text{ if ED95 is missing} \\ \frac{140}{\sqrt{KD*N}}, \text{ if ED95 \& DP are missing} \\ \frac{140}{\sqrt{K_NORM*N}}, \text{ if ED95, DP \& KD are missing} \\ \frac{140}{\sqrt{K_NORM*B}}, \text{ if ED95, DP, KD \& N are missing} \\ \frac{140}{\sqrt{1.7*B}}, \text{ if ED95, DP, KD \& K_NORM are missing,} \\ \quad \text{using the lowest KD in GPMDB, about 1.7} \end{cases}$<br>finally $w$=0 if this $\alpha95$ completely overlaps with another smaller $\alpha95$ whose paleopole is exactly derived from the same place and same rock. |
| 4 | Age error Position to bin (more overlap, greater weight):<br>wha, window high age; wla, window low age<br>$w = \begin{cases} \frac{wha-LOMAGAGE}{age_range}, \text{ if LOMAGAGE>wla \& HIMAGAGE>wha} \\ \frac{HIMAGAGE-wla}{age_range}, \text{ if LOMAGAGE<wla \& HIMAGAGE<wha} \\ \frac{wha-wla}{age_range}, \text{ if LOMAGAGE<wla \& HIMAGAGE>wha} \\ 1, \text{ if LOMAGAGE>wla \& HIMAGAGE<wha.} \end{cases}$ |
| 5 | Combining 3 and 4: average of the two weights from 3 and 4 |

*Weighting by spatial uncertainty (Weighting 3).* Paleopoles with a smaller $\alpha95$ confidence ellipse are given a higher weighting than those with a larger $\alpha95$, using a Gaussian/Normal distribution centered on 0 with standard deviation of 10. However, not all paleopoles' $\alpha95$ are given in the database. If $\alpha95$ is not given, DP (the semi-axis of the confidence ellipse along the great circle path from

site to pole) is assigned as $\alpha 95$. If DP is also not given, $\alpha 95$ was further approximated by $\frac{140}{KD*N}$, where KD is Fisher precision parameter for mean direction if this parameter is given, or Fisher precision parameter for Normal directions (K_NORM) if only K_NORM is given when KD is missing. If N is not given, B is used as N. If even K_NORM is also missing, the lowest KD value 1.7 in GPMDB 4.6b is used as KD. It is also worthwhile to mention that if samples, where two paleopoles are derived, are exactly from the same place and same rock, and one $\alpha 95$ is completely inside the other $\alpha 95$, a zero is assigned as the weight of the data with the larger $\alpha 95$. In fact, in the above described procedure A95 (circle of 95% confidence about mean pole) is a better alternative instead of $\alpha 95$, because A95 is directly reflecting the spatial uncertainty of the paleopoles. However, most paleopoles' A95s are not given in GPMDB 4.6b, so $\alpha 95$ is used instead since $\alpha 95$ is also indirectly reflecting the quality of the dataset.

*Weighting by degree of overlap between moving window and pole age uncertainty (Weighting 4).* If a large fraction of the age range for an individual paleopole falls within the current window, it is given a higher weighting than a pole where the overlap is smaller, because it is more likely to be close to the true pole position in the window interval. In other words, if window intersects with part of age range, weight= (intersecting part) / (age range width).

*Weighting by both spatial and temporal uncertainty (Weighting 5).* This weighting method is a combination of No 3 (but here the standard deviation is 15 though) and No 4. It takes the average of sums of the weights generated by weighting methods 3 and 4.

Some of the picking (Table. 2) and weighting (Table. 3) methods developed here are also connected with the V90 Q factors mentioned above. For example, Pt 2, 3 and Wt 2, 4, 5 are related to the V90 criteria 1; Pt 2, 3, 22, 23, 26, 27 and Wt 1, 3 are related to the V90 criteria 2; Pt 22 and 23 are related to the V90 criteria 3; The data constraining described in Appendix A is related to the V90 criteria 5; and Pt 18 and 19 are related to the V90 criteria 7.

### 2.4 Reference Paths

A prediction of the expected APWP for any plate can be generated using a plate kinematic model (e.g. the last c. 180–200 Myr plate motions reconstructed from spreading ridges in ocean basins) that is tied in to an absolute reference frame. Here, we use the rotations of (O'Neill et al. 2005), which describe motion of Nubia (plate ID 701) relative to the Indo-Atlantic hotspots. Such an absolute frame of reference based on hot spots has been extended back to 120 Ma (O'Neill et al. 2005). North America is linked to this reference frame across the Mid-Atlantic ridge, using North America-Nubia rotations from (DeMets et al. 2010) to C-Sequence chron C1no (0.78 Ma), from (Shephard et al. 2012) to chron C2An.2n (2.7 Ma), from (Müller et al. 1999) to chron C5n.1ny (9.74 Ma), from (Gaina et al. 2013) to chron C5n.2o (10.949 Ma), from (Müller et al. 1999) to chron C6ny (19.05 Ma), from (Gaina et al. 2013) to chron C6no (20.131 Ma), from (Müller et al. 1999) to chron C34ny (83.5 Ma), from C34ny to about 118.1 Ma (Shephard et al. 2012), and to closure at C34no (120.6 Ma) (Gaina et al. 2013). India is linked via the East African Rift Valley (Somalia-Nubia rotations to chron C1no [0.78 Ma] from (DeMets et al. 2017), to chron C2A.2no [3.22 Ma] from (Horner-Johnson et al. 2005), and to closure at C7.2m [25.01 Ma] from (Rowan & Rowley 2016) and C34 [85 Ma; Rowley, pers. comm.], and finally extended to 120 Ma because there was no known relative motion between Somalia and Nubia from 120 Ma to 85 Ma according to the rotations from (Müller et al. 2016)); and Australia via the East African Rift Valley, SW Indian Ridge (E Antarctica-Somalia rotations to chron C1no [0.78 Ma] from (DeMets et al. 2017), to chron C2A.2no [3.22 Ma] from (Horner-Johnson et al. 2005), to chron C5n.2no [10.95 Ma] from (Lemaux et al. 2002), to chron C13ny [33.06 Ma] from (Patriat et al. 2008), to chron C29no [64.75 Ma] from (Cande et al. 2010), to chron C34y [83 Ma; Rowley, pers. comm.], to 96 Ma from (Marks & Tikku 2001), and to closure at M0 [120.6 Ma] from (Müller et al. 2008)), and SE Indian Ridge (Australia-East Antarctica rotations to chron C1no [0.78 Ma] from (DeMets et al. 2017), to chron C6no [20.13 Ma] from (Cande & Stock 2004), to chron C8o [26 Ma] from (Granot & Dyment 2018), to chron C17n.3no [38.11 Ma] from (Cande & Stock 2004), to C34ny [83.5 Ma] from (Whittaker et al. 2013), to the Quiet Zone Boundary [96 Ma] from (Whittaker et al. 2007), to closure at 136 Ma from (Whittaker et al. 2013)). Above all, these kinematic models (see the collected rotation data for the above-mentioned plate circuits in the supplementary material) tied in to a hot spot absolute reference frame can guarantee the oldest predicted poles for the three continents back to 120 Ma.

To reconstruct a reference APWP at the required time steps for comparison with the paleomagnetic APWPs, rotations and their associated uncertainties were interpolated between constraining finite rotation poles according to the method of (Doubrovine & Tarduno 2008), assuming constant rates.

Neither the hotspot reference frame nor the paleomagnetic reference frame are truly fixed with respect to the solid Earth. In the former case, hotspots are not truly stationary in the mantle (Steinberger & O'Connell 1998); in the latter, true polar wander (TPW) may also lead to differential movements of the solid earth with respect to the spin axis (Evans 2003). In reality, it is difficult to untangle these effects. Whilst there is little clear evidence for significant TPW in the past about 120 Myr (Cottrell & Tarduno 2000; Riisager et al. 2004), modeling suggests that the effects of hotspot drift can start to become significant over 80–100 Myr timescales (O'Neill et al. 2005). Because paleomagnetic APWPs have large associated spatial errors, a synthetic APWP calculated using a fixed hotspot reference frame is unlikely to deviate significantly from the 'true' APWP, and most comparison experiments use a fixed hotspot model (FHM) reference path for North America (Fig. 5), India (Fig. 6) and Australia (Fig. 7). However, the full set of comparisons for the 28 picking methods and 6 weighting methods was also run for reference paths generated using the moving hotspot model (MHM) rotations of (O'Neill et al. 2005), which incorporate motions of the Indo-Atlantic hotspots relative to the mantle derived from mantle convection modeling.

When comparing the synthetic APW paths for the three plates (inset, Fig. 7), there are clear differences. The predicted mean north pole for North America at 120 Myr is still at about 75°N (Fig. 5), indicating rather slow drift with respect to the spin axis; this is due to a large component of the North American plate's absolute motion in the past 120 Myr being to the east. In contrast, the rapid northward motion of the Indian plate in the same period, particularly prior to its collision with Asia at about 50–55 Ma (Najman et al. 2010), is reflected by the 120 Ma predicted mean north pole being located at about 20°N (Fig. 6). Australia represents an intermediate case, with north westerly plate motion from about 120–60 Myr changing to more rapid northward motion from about 60–55 Ma to the present (Whittaker et al. 2007). When comparing the FHM and

MHM tracks, differences in the oldest parts (before about 80 Ma) are apparent for India and North America.

These differences in the reference path due to different plate kinematics is another variable that may affect the performance of the different weighting algorithm for different plates, in addition to the distribution and type of the contributing mean poles used to generate the paleomagnetic APWPs.

### 2.5 Comparison Algorithm

Comparisons between APWPs generated using different picking and weighting algorithms and the synthetic reference APWPs were performed using the composite path difference ($\mathcal{CPD}$) algorithm described in Chapter 2, with equal weighting given to the spatial, length and angular differences (i.e., $W_s = W_l = W_a = \frac{1}{3}$).

This does not only help find the most similar paleomagnetic APWP (from the best algorithm) to the reference APWP, but also help further test and demonstrate the validness of the similarity measuring tool in practise.

## 3 RESULTS

### 3.1 Baseline results: 10 Myr window, 5 Myr step, fixed hotspot reference

Fig. 8 shows the $\mathcal{CPD}$ scores for the APWPs generated with all 28 picking methods (AMP and APP with one of 14 separate filters applied) and one of 6 weighted mean calculations then applied, compared to the FHM reference path for North America (Fig. 8a), India (Fig. 8b) and Australia (Fig. 8c). The 27 lowest and highest of the 168 scores for each plate (values greater than 1 standard deviation from the mean) are marked in green and red, respectively. Different combinations of windowing method, filtering and weighting clearly affect the difference score, with $\mathcal{CPD}$ values ranging from 0.0037 to 0.5137. The fits for paths with low difference scores are clearly much better than for those with high ones (Fig. 9). From Fig. 8, it is clear that:

(i) There is much more variation in scores along the horizontal axes than the vertical axes (see self-explanatory topography of bands in Fig. 8), suggesting that the choice of windowing and filtering method (Table. 2) has a much greater impact than weighting (Table. 3).

(ii) Many of the highest scores (worst fits) are associated with even-numbered picking and filtering methods, i.e., those which use the AMP windowing algorithm. Even so, methods 4 and 6 are among the best methods for India.

(iii) The magnitude and range of $\mathcal{CPD}$ scores for each of the three plates is different, with the North American plate having the lowest magnitudes and range (Fig. 8a), and the Indian plate having the highest (Fig. 8b). In general, the scores of North America and Australia are relatively closer (Fig. 12).

(iv) Although there is some overlap (e.g., picking methods 19, 21 [best], and 2, 16 [worst] for all the three plates or for both India and Australia; 1, 11, 13, 19, 21, 25 [best], and 2, 14, 16, 22, 26 [worst] for both North America and Australia; 5, 7, 19, 21 [best], 2, 8, 16, 18 [worst] for both North America and India), the best- and worst-performing picking/filtering and weighting algorithms are not exactly the same for each plate.

### 3.2 Effects of windowing method

Dividing $\mathcal{CPD}$ scores according to whether the AMP or APP windowing method was used (Fig. 10) confirms that whilst the lowest $\mathcal{CPD}$ scores for paths generated by the AMP windowing algorithm are close to the lowest scores generated using the APP method, the highest scores are much higher (Fig. 11). The mean of the $\mathcal{CPD}$ scores for APWPs generated using AMP is greater than the maximum APP-derived score for the Indian and Australian plates (Figs. 10b, 10c, 11c, 11c), and more than 1 standard deviation greater than the APP mean for North American APWPs (Figs. 10a, 11a).

For each of the 84 possible combinations of filter method and weighting, the AMP-derived score is typically 3-5 times higher than the equivalent APP-derived score (Fig. 11, insets). APP-generated paths yield a lower $\mathcal{CPD}$ score than the equivalent AMP-generated path for 82 (97.6%) of the North America scores, 72 (85.7%) of the India scores, and 84 (100%) of the Australia scores. For the North American and Indian plates, filtering that prefers igneous poles (picking methods 4/5 and 6/7) or removes poles with large age uncertainties (methods 2/3, 22/23 and 26/27) is most likely to produce AMP scores close to (less than 1.5 times) or less than the APP scores (Figs. 10a, 10b). In the former case, the scores are comparable and relatively low; in the latter case, they are comparable but relatively high. For the Australian plate, only strict spatial and temporal uncertainty filtering (method 2/3) and correcting for inclination shallowing (method 8/9) produces comparable scores (Fig. 10c), producing comparable but moderate scores in both cases.

On the North American plate, the APP-derived $\mathcal{CPD}$ score is most likely to be significantly better (defined as more than 5 times higher) when weighting methods 0 (no weighting) and 1 (by sample and site number) have been applied (Fig. 10a). This pattern is also seen for the Australian plate, but removal of poles published after 1983 (picking method 16/17) also results in significantly better performance of the APP method (Fig. 10c). For the Indian plate, the largest difference occurs when poles with suspected overprints or local rotations are removed (picking method 18/19, Fig. 10b).

### 3.3 Effects of filtering and weighting

When $\mathcal{CPD}$ scores are separated by windowing method (Fig. 10), the effects of particular filtering and weighting methods become easier to discern. In general, different filters (rows) produce larger variations in scores than different weighting methods (columns). With the exception of AMP-derived paths for India, the $\mathcal{CPD}$ score for paths with no filtering (picking method 0/1) and no weighting (weighting method 0) is lower than the mean scores for that plate and windowing method. Therefore filtering and weighting at best slightly improves, and at worst significantly degrades, the APWP fit to the reference path.

#### *3.3.1 Filter aggression*

When considering the effects of filtering, it is important to consider how many poles within the data set have been removed (related to picking no 2–7, 10, 11, and 14–27) or corrected (8, 9, 12 and 13): if there is very little alteration of the data set, little change from no filtering (methods 0/1) would be expected. In terms of the numbers of poles affected, the most consequential filters are:

(i) removal of sedimentary poles (methods 4/5 and 6/7), which

**Figure 5.** MHM predicted 120–0 Ma APWP (solid line) for $NAC$ through the North America–Nubia–Mantle plate circuit. The FHM predicted path (dashed line with shaded uncertainties) is also shown for comparison. The age step is 5 Myr. Compared with the 10 Ma paleomagnetic mean pole calculated by the AMP method (dark triangle), the coeval mean pole derived from the APP method is closer to both FHM and MHM predicted 10 Ma poles, which indicates more data diluting the effect of outliers. See also the paleopoles that the two mean poles are composed of in Fig. 4.

removes 40-50% of the dataset on all 3 plates, with the highest proportion being removed on the Indian plate. Although method 4/5 is more strict, it does not remove many more poles than method 6/7. Especially for North America and India, the numbers of the filtered paleopoles by method 4/5 and 6/7 are actually the same.

(ii) correction of sedimentary poles for inclination flattening (filter 8/9), which affects 38-48% of the dataset, with the highest affected proportion on the Indian plate.

(iii) removal of poles with large temporal and spatial uncertainty (methods 2/3, 22/23), particularly for the Australian plate, where the SS05 filtering criteria removes about 70% of the poles. Filter 26/27 combines methods 22/23 and 24/25, but no or very few (in the case of Australia only 1 additional pole) are actually removed.

(iv) filtering based on publication date (methods 14/15 and 16/17), with the ratio of pre/post 1983 poles varying from about 50/50 on the North American plate to 70/30 on the Australian plate.

Conversely, filtering or correction for redbeds (methods 10/11 and 12/13), local rotations and overprints (methods 18/19 and 20/21 [one paleopole influenced by local rotation removed, and one corrected, for only India; about 2.7 per cent, Fig. 8b]), or superseded data (method 24/25) affected 4% or less of the poles on any plate.

### 3.3.2 *Filter performance*

Focussing on the filtering and weighting methods with aggressive filtering, some commonalities in the best- and worst-performing methods can be observed, although there are usually exceptions for particular plates and/or windowing methods:

(i) For all 3 plates, higher $\mathcal{CPD}$ scores are commonly associated with filtering based on $\alpha 95$ and age range (picking methods 2/3, 22/23, 26/27), with the exception of AMP-derived paths for India, where picking methods 22 and 26 produce some of the lowest scores.

**Figure 6.** MHM predicted 120–0 Ma APWP (solid line) for India through the India–Somalia–Nubia–Mantle plate circuit. Its age step is 5 Myr. The dashed line is the FHM predicted path shown for comparison. The inset shows paths for fast moving India and also much slower moving North America shown in Fig. 5.

(ii) For North America and India, low scores are commonly associated with removal of non-igneous poles (methods 4/5 and 6/7), particularly for AMP-derived paths. On the Australian plate, these filters are less effective.

(iii) For North America and Australia, correction for inclination flattening generates $\mathcal{CPD}$ scores very similar to scores with no filtering for AMP-derived paths (method 8), and increases scores for APP-derived paths (method 9). In contrast, for India there is a small decrease in $\mathcal{CPD}$ scores compared to no filtering for both AMP- and APP-derived paths.

(iv) Many of the highest difference scores for North America and India occur when paleopoles published after 1983 are removed (method 16/17), whilst removing paleopoles published before 1983 (method 14/15) generates $\mathcal{CPD}$ scores comparable to scores with no filtering (relatively much lower). In contrast, for Australia methods 14/15 produces some of the highest $\mathcal{CPD}$ scores, and methods 16/17 have little effect. The number of older studies (65/68; Fig. 8c) is almost 2.5 times of the number of newer studies (27/29) though.

Whilst it is generally true that methods with low-aggression filters do not generate scores that differ much from the no-filtering scores, there are some exceptions:

(i) For North America and Australia, removing superseded paleopoles (picking method 24/25) produces lower $\mathcal{CPD}$ scores.

(ii) Removing paleopoles suspected to be affected by overprints or local rotations (picking method 18/19) consistently produces lower $\mathcal{CPD}$ scores for Indian APP-derived paths.

#### *3.3.3 Weighting performance*

Compared to the variations resulting from different windowing methods and filters, Figs. 10, 13, 14 and 15 indicate that the effect of weighting the data prior to calculating a Fisher mean is generally small. Where an effect can be seen, it is negative, generating larger CPD scores.

(i) For APP-derived paths from the North American and Australian plates, weighting method 0 (no weighting) and 1 (weighting by sample and site number) appear to produce slightly better $\mathcal{CPD}$ scores than other weighting methods.

(ii) Weighting method 3 (weighting by spatial uncertainty) is most likely to generate much higher $\mathcal{CPD}$ scores, particular for AMP-derived paths, and particular for the Indian plate.

(iii) Weighting method 0, 1 and 5 are generally producing better similarity than 2, 3 and 4.

### 3.4 Effects of window size

Fig. 16 shows the $\mathcal{CPD}$ scores for the APWPs generated with all 28 picking and 6 weighting methods, compared to the FHM reference paths, with the picking time window width increased from 10 to 20 Myr, and the window step increased from 5 to 10 Myr.

#### *3.4.1 Overall change*

Along with the increasing sizes of moving window and step, mean, median and range values of $\mathcal{CPD}$ scores shrink, except for North America the mean value of $\mathcal{CPD}$ scores keeping unchanged (Fig. 8 and Fig. 16). For North America, most scores actually increase (i.e. 10-5 is better) except for methods 2, 4, 6, 16/17, 22/23, 26/27 where 20-10 better ('quality filtering' is less 'bad'?). For India, most scores are lower (20-10 better) except methods 22 and 26(all), 19(0,1,3,5), 21(0) and 27(1,3,4). Biggest decreases are associated with weighting 3 (i.e., effect of weighting is being diluted). For Australia, changes are largely negative, which means 20-10 is better. Also negative change is large for 3, 4, 6, 23, 27. For method 14, 10-5 is much better, whereas for 15, 20-10 is much better.

#### *3.4.2 Relative performance of methods*

(i) Method 19 is still among the best, and 16 still one of the worst, for all 3 plates.

**Figure 7.** MHM predicted 120–0 Ma APWP (solid line) for Australia through the Australia–East Antarctica–Somalia–Nubia–Mantle plate circuit. Its age step is 5 Myr. The dashed line is the FHM predicted path shown for comparison. The inset shows paths for fast moving India shown in Fig. 6, much slower moving North America shown in Fig. 5, and also relatively intermediate moving Australia.

(ii) Many of the biggest changes (decreases) for North America and India occur when paleopoles published after 1983 are removed and calculation is based on AMP (method 16). In contrast, for Australia the biggest changes (decreases) are associated with method 15, and methods 16/17 have little effect (Fig. 18).

(iii) Only for North America, the 10/5 Myr bin/step methods generally and unexceptionally produce better similarities than the 20/10 Myr methods do (Fig. 18).

(iv) All the APP methods with kinds of corrections, 1, 9, 13, 19, 21 and 25, show less changes for all the three plates. Methods 23 and 27 show less changes for North America and India, but more changes for Australia.

(v) APP still outperforms AMP. For North America and India, the percentage of factor of greater than 3 (about 28.6% and 3.8% respectively) is less than 10-5 (about 40.5% and 39.3% respectively), while for Australia the percentage (about 71.4%) is more than 10-5 (50%).

(vi) Both scores and AMP/APP difference are low for methods 4/5, 6/7.

(vii) Both scores and AMP/APP difference are high for methods 16/17.

(viii) Scores are high but AMP/APP difference are low for methods 2/3, 22/23 on North America and India.

(ix) Scores are low but AMP/APP difference are high for 0/1 on Australia.

(x) Comparable AMP/APP (20-10 vs 10-5) appears for 8/9 on Australian plate.

(xi) Method 1 is still a good performer.

(xii) Methods 4/5, 6/7 are still 'good' aggressive filters.

(xiii) Methods 2/3, 22/23 (at least 2 and 22) are still 'bad' aggressive filters. And the scores derived from these methods have control on overall range of scores.

(xiv) Methods 16/17 are still 'bad' for North America and India.

### *3.4.3 How time window/step size affects results*

Fitting curves by moving averaging change with different time window lengths and time increment lengths (i.e. steps) (e.g., the similarity of the pair in Fig. 17f is improved a bit compared to Fig. 9f). A balance needs to be made between having windows that are too wide and steps that are too long which will smooth the data so much we miss actual details in the APWP (e.g. those 20 Myr window 10 Myr step paleomagnetic paths in Fig. 17 and even 30/15 Myr win-

(a) Plate ID 101 with children: minimum 0.0053 (11(0)), maximum 0.0921 (16(3)), mean 0.0335, median 0.0342

(b) Plate ID 501: minimum 0.023 (19(0)), maximum 0.5137 (8(3)), mean 0.1182, median 0.0818

(c) Plate ID 801 with children: minimum 0.0037 (17(0)), maximum 0.3954 (22(3)), mean 0.08432, median 0.05037

**Figure 8.** Equal-weight composite path difference ($\mathcal{CPD}$) values between each continent's paleomagnetic APWPs and its predicted APWP from FHM and related plate circuits. The paths are in 10 Myr bin and 5 Myr step. The difference values less than one-standard-deviation interval of the whole 168 values (lower 15.866 per cent) are colored in green, more than one-standard-deviation interval (upper 15.866 per cent) colored in red. Exactly the same columns are connected. The percentages of removed paleopoles are derived relative to picking 1, corrected relative to each corresponding picking method (8, 9, 12, 13; 1 removed and 1 corrected by 20, 21 for India). Fit quality (FQ) for each score is color coded.

(a) North America: minimum 0.0053 (11(0)), FQ B-B

(b) North America: maximum 0.0921 (16(3)), FQ B-B

(c) India: minimum 0.023 (19(0)), FQ C-A

(d) India: maximum 0.5137 (8(3)), FQ C-A

(e) Australia: minimum 0.0037 (17(0)), FQ C-A

(f) Australia: maximum 0.3954 (22(3)), FQ C-A

**Figure 9.** Path comparisons with best and worst difference values shown in Fig. 8. The parenthetical remarks are Picking No and Weighting No.

(a) AMP: minimum 0.0292 (4(5)), maximum 0.0921 (16(3)), mean 0.0465, median 0.0464; APP: minimum 0.0053 (11(0)), maximum 0.05686 (5(3)), mean 0.0205, median 0.01594

(b) AMP: minimum 0.047 (4(3)), maximum 0.5137 (8(3)), mean 0.1684, median 0.1862; APP: minimum 0.023 (19(0)), maximum 0.1359 (3(4)), mean 0.06788, median 0.06023

(c) AMP: minimum 0.0305 (0(3)), maximum 0.3954 (22(3)), mean 0.1264, median 0.0951; APP: minimum 0.0037 (17(0)), maximum 0.11035 (15(3)), mean 0.04225, median 0.0282

**Figure 10.** Separated results from AMP and APP in Fig. 8. For each grid block (left: AMP; right: APP), the difference values less than one-standard-deviation interval of the whole 84 values are labeled in green, more than one-standard-deviation interval labeled in red.

(a)

(b)

(c)

**Figure 11.** Box-and-whisker and cross (inset) plots of Fig. 10. The $\mathcal{CPD}$s from same filter and weighting method (red and blue dots plotted with box-and-whisker) are connected; some special cases where $\mathcal{CPD}$ from AMP lower than from APP are highlighted using darker connecting lines. Dot symbols are semi-transparent so a darker color indicates a greater number of data at a given $\mathcal{CPD}$.

(a) Differences between Fig. 8b (India) and Fig. 8a (North America)

(b) Differences between Fig. 8c (Australia) and Fig. 8a (North America)

(c) Differences between Fig. 8b (India) and Fig. 8c (Australia)

**Figure 12.** Differences between grids in Fig. 8. The absolute difference values less than 1.96-standard-deviation interval of the whole 168 values are labeled in green, more than 1.96-standard-deviation interval labeled in red.

(a) For Weighting No. 0

(b) For Weighting No. 1

(c) For Weighting No. 2

(d) For Weighting No. 3

(e) For Weighting No. 4

(f) For Weighting No. 5

**Figure 13.** Tested spatial difference ($ds$) values (color shaded) between North American paleomagnetic APWPs and its predicted APWP from the FHM and related plate circuits. The paths are in 10 Myr bin and 5 Myr step. The labeled numbers on the grids are the numbers of site mean poles that are contributing to each mean path pole.

dow and step; Table. 5 and Fig. 19) and windows that are too narrow and steps that are too short which introduces noise by having too few poles in each window (e.g. 2 Myr window 1 Myr step; Table. 5 and Fig. 19). There is a dependence here on data density: higher density allows smaller windows/steps (this is one of the things we want to test with selective data removal in Chapter 4). A variety of ways of binning the data (here 30–2 Myr window size and half of the size as step) are being tested to see which one produces the better and more appropriately smoothed fit.

Note that there are 135, 75 and 99 paleopoles that compose of 120–0 Ma APWPs of North America, India and Australia respectively. Does the reason of 10/5 Myr generally better than 20/10 could be the relatively larger number of paleopoles for North America? Since theoretically for each sliding window, the more "bad" paleopoles it contains, the worse similarity we should obtain. In the contrary, the less paleopoles the window contains, the weaker the effect of averaging out "bad" poles' influence would be. So is there a threshold number of paleopoles for making an paleomagnetic APWP? For example, for making a 120–0 Ma APWP, do the results indicate the best number of paleopoles we need should be some value between 99 and 135? Here a test is implemented as follows. With the results from the 10/5 and 20/10 bin/step together, 2/1, 4/2, 6/3, 8/4, 12/6, 14/7, 16/8, 20/10, 24/12 and 30/15 Myr bin/step are also used to generate paleomagnetic APWPs for North America, India and Australia to see which bin/step size would make paleomagnetic APWP closest to reference path. Will the similarities they generate be generally worse than those the 10/5 Myr bin/step generates? Or will they be better first and then worse than those the 10/5 Myr bin/step generates when the bin/step sizes increase up to 20/10 Myr? For the best results (Table. 5), as expected, AMP needs wider sliding window and step to get closer to the reference path while APP does not (Fig. 19). Even the best sizes of sliding

(a) For Weighting No. 0

(b) For Weighting No. 1

(c) For Weighting No. 2

(d) For Weighting No. 3

(e) For Weighting No. 4

(f) For Weighting No. 5

**Figure 14.** Tested length difference (*dl*) values (color shaded) between North American paleomagnetic APWPs and its predicted APWP from FHM and related plate circuits. The paths are in 10 Myr bin and 5 Myr step. The labeled numbers on the grids are the averaged numbers of site mean poles that are contributing to each segment's two mean path poles.

window and step are assigned for AMP, the results from APP are still much better than those from AMP. Picking methods (directly related to N) are still the key influence factor of choosing a better sliding window size and step size of moving averaging, although weighting methods are also important.

**3.4.3.1 What to expect is** the difference values for larger window/step size should be generally lower than those for smaller window/step size, which further could result in more best methods and less worst methods.

**3.4.3.2 The results** are summarised in Table. 4, Table. 5 and Fig. 19.

**3.4.3.3 Conclusions:**

*If AMP has to be used, better results can be obtained through using large sizes of sliding window and step, commonly more than 24/12 Myr. In addition, we should be cautious when weighting 3 is used with AMP.*

(a) For Weighting No. 0

(b) For Weighting No. 1

(c) For Weighting No. 2

(d) For Weighting No. 3

(e) For Weighting No. 4

(f) For Weighting No. 5

**Figure 15.** Tested angular difference ($da$) values (color shaded) between North American paleomagnetic APWPs and its predicted APWP from FHM and related plate circuits. The paths are in 10 Myr bin and 5 Myr step. The labeled numbers on the grids are the averaged numbers of site mean poles that are contributing to each segment-orientation-change's three mean path poles.

*APP is still recommended, not only because the temporal uncertainty is incorporated into the algorithm but also the results from APP are not as that sensitive as AMP to the changes of sliding window and step sizes. In fact, for APP the results from different window and step sizes are much more stable than those from AMP (Fig. 19). This means we actually do not need to worry about what sizes should be chosen for the sliding window and step when we use APP method.*

### 3.5 Different reference path

Fig. 20 and Fig. 22 show the $\mathcal{CPD}$ scores for the APWPs generated with all 28 picking and 6 weighting methods, compared to the MHM reference paths instead, still with the same picking time window width 10 and 20 Myr, and the same window step 5 and 10 Myr respectively.

(a) North America (101 with children): minimum 0.0115 (13(0)), maximum 0.0635 (16(2)), mean 0.0334, median 0.02786

(b) India (501): minimum 0.0143 (6(3)), maximum 0.1543 (16(3)), mean 0.0738, median 0.0754

(c) Australia (801 with children): minimum 0.0022 (11(0)), maximum 0.3265 (22(3)), mean 0.0693, median 0.0362

**Figure 16.** As Fig. 8, here the paths are generated in 20 Myr bin and 10 Myr step. The difference values less than one-standard-deviation interval of the whole 168 values are colored in green, more than one-standard-deviation interval colored in red. Compare the numbers of picked paleopoles with those in Fig. 8.

(a) North America: minimum 0.0115 (13(0)), FQ B-B

(b) North America: maximum 0.0635 (16(2)), FQ B-B

(c) India: minimum 0.0143 (6(3)), FQ C-B

(d) India: maximum 0.1543 (16(3)), FQ C-B

(e) Australia: minimum 0.00217 (11(0)), FQ B-B

(f) Australia: maximum 0.3265 (22(3)), FQ C-B

**Figure 17.** Path comparisons with best and worst difference values shown in Fig. 16. The parenthetical remarks are Picking No and Weighting No.

(a) North America (101): Generally Bin/Step 10/5 Myr is better ($\frac{107}{168} \approx 63.7\%$)

(b) India (501): Generally Bin/Step 20/10 Myr is better ($\frac{145}{168} \approx 86.31\%$)

(c) Australia (801): Generally Bin/Step 20/10 Myr is better ($\frac{13}{21} \approx 61.9\%$)

**Figure 18.** Differences between grids in Fig. 8 (10 Myr bin, 5 Myr step) and Fig. 16 (20 Myr bin, 10 Myr step). The absolute difference values less than 1.96-standard-deviation interval of the whole 168 values are labeled in green, more than 1.96-standard-deviation interval labeled in red.

**Table 4.** Consistency check on comparisons of 20/10 window/step and 10/5 window/step. Notes: E: expected; UE: unexpected.

| Comparisons | | Consistency of Best | | | Consistency of Worst | | | If Difference Values for 20/10 Myr Bin/Step Are Lower (Y/N) | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10/5 | 20/10 | Y/N | Special Case(s) | Notes | Y/N | Special Case(s) | Notes | Mean | Median | Maximum | Minimum | All | If No, Unexpected Case (s) | Notes |
| | | | | | | FHM | | | | | | | | |
| Fig. 8a | Fig. 16a | Y | – | Same: Picking no. 1, 5, 7, 11, 13, 15, 19, 21 and 25 | N | Picking no. 2, 5, 7, 22 and 26 for 10/5 (E); 0, 10, 12, 20 and 24 for 20/10 (UE) | Same: Picking no. 8, 14, 16 and 18 | N | Y | Y | N | N | (0,1,8,10,11,12,18,19,20,21, 24,25)(0-5) (3,14)(0-3,5) (5,7,13,15)(0-2,4,5) 9(0,2,4,5) 23(2,3) 27(0-2); account for 29/42 | |
| Fig. 8b | Fig. 16b | Y | 2 more best: Picking no. 9 and 21 only for 10/5 (UE) | Same: Picking no. 4, 5, 6, 7 and 19 | N (almost Y) | Picking no. 8 for 10/5 (E); 23 and 27 for 20/10 (UE) | Same: Picking no. 0, 2, 10, 12, 16, 18, 20 and 24 | Y | Y | Y | Y | N | 15(3) 19(0,1,3,5) 21(0) (22,26)(0-5) 23(1,3,4) 27(1,3); account for 23/168 | |
| Fig. 8c | Fig. 16c | Y | – | Same: Picking no. 1, 11, 13, 17, 19, 21 and 25 | Y | 1 more worst: Picking no. 4 only for 10/5 (E) | Same: Picking no. 2, 14, 16, 22 and 26 | Y | Y | Y | Y | N | 0(1,2,5) (1,18,25)(0,1,3,5) 7(0) (8,10,20)(2,5) 9(1) 11(3) (12,24)(1,2) 13(0,1,3) (14,22,26)(0-2,4,5) (16,17)(0-5) 19(5) 21(0,1,5); account for 33/84 | |
| | | | | | | MHM | | | | | | | | |
| Fig. 20a | Fig. 22a | N | Picking no. 1, 13, 21 and 25 for 10/5 (UE); 22 and 26 for 20/10 (E) | Same: Picking no. 4, 5, 7, 11, and 15 | Y | 2 more worst: Picking no. 17, 22 only for 10/5 (E) | Same: Picking no. 0, 8, 10, 12, 14, 16, 18, 20 and 24 | N | Y | N | N | N | (0,4,8,10,11,12,14,15,18, 20,24,25)(0-5) (1,19,21)(0,1,5) 3(5) (5,7)(0-2,4,5) 6(1) (9,13)(0,1,3,5) 23(1,3) 27(1,3,5); account for 53/84 | |
| Fig. 20b | Fig. 22b | Y | 2 more best: Picking no. 22 and 26 only for 10/5 (UE) | Same: Picking no. 4, 5, 6, 7 and 19 | N | Picking no. 8, 20 for 10/5 (E); 23 and 27 for 20/10 (UE) | Same: Picking no. 0, 2, 10, 12, 16, 18 and 24 | Y | Y | Y | Y | N | 1,5 (4,15),3 (19,23,27),(0-5) 21,(0,1,35) 22,(1,4) 25,(3,5) 26,4; account for 5/28 | |
| Fig. 20c | Fig. 22c | Y | 3 more best: Picking no. 15, 23 and 27 only for 20/10 (E) | Same: Picking no. 1, 11, 13, 17, 19, 21 and 25 | Y | – | Same: Picking no. 2, 14, 16, 22 and 26 | Y | Y | Y | N | N | (0,24)(1,2,5) (1,11,13,18,19,21,25)(0,1,3,5) 4(0,3,5) (5,7)3 (8,17)(0-3,5) 12(1-3,5) 20(1,2,4,5) 10(1,2) (14,22,26)(0-2,4,5) 16(0-5); account for 10/21 | |

### *3.5.1 Overall change*

Although the reference path is changed to MHM, mean, median and range values of $\mathcal{CPD}$ scores are almost unchanged and comparable (Fig. 8 vs. Fig. 20, and Fig. 16 vs. Fig. 22).

For 10/5 Myr bin/step, the absolute differences (Fig. 24) are all lower than 0.041, and most are less than 0.01. This indicates that for comparing with paleomagnetic APWPs choosing fixed or moving hotspot model for generating a reference path is not quite different. Therefore selecting fixed or moving model for having a reference path is not a priority. However, based on the signs of the differences between the scores from FHM and MHM (Fig. 24), for North America (Fig. 24a), FHM derived path is a slighly better reference in general, while for both India (Fig. 24b) and Australia (Fig. 24c), generally MHM derived path is a slightly better choice.

For North America, neigher more nor less change is favored by any particular picking method or weighting method. For example, methods 8(0,4), 12(3), (14,24)2 show more change and (9,22,26)1, (23,27)2, and 24(5) show less. However, methods 23 and 27 show more change for India, and 22 and 26 show more change for Australia. Methods 8, 17 and 21 show less change for India, and 1, 21 and 25 show less change for Australia.

### *3.5.2 Relative performance of methods*

(i) When 10/5 Myr bin/step is applied, method 19 is still among the best, and 16 still one of the worst, for all 3 plates. Even when 20/10 Myr bin/step is applied, 19 is still relatively a "good" method, and 16 a "bad" one.

(ii) APP still outperforms AMP. For 10/5 Myr bin/step, the percentage of factor of greater than 3 (about 19.05%, 38.1% and 48.81% for North America, India and Australia respectively) is less than FHM (about 40.5%, 39.3% and 50%). For 20/10 Myr bin/step, the percentage of factor of greater than 3 (about 3.6% and 46.43% for India and Australia) is still less than FHM (about 3.8% and 71.4%) whilst for North America the percentage (about 44.05%) is more than FHM (28.6%).

(iii) Both scores and AMP/APP difference are still low for methods 4/5, 6/7.

(iv) Both scores and AMP/APP difference are still high for methods 16/17 (method 17 gives low scores for Australia).

(v) Still scores are high but AMP/APP difference are low for methods 2/3.

(vi) Still scores are low but AMP/APP difference are high for 0/1 on Australia.

(vii) Comparable AMP/APP (20-10 vs. 10-5, or FHM vs. MHM) still appears for 8/9 on Australian plate.

(viii) Method 1 is still a good performer.

(ix) Methods 4/5, 6/7 are still 'good' aggressive filters.

(x) Methods 2/3 (at least 2) are still 'bad' aggressive filters.

(xi) For 10/5 Myr bin/step, method 22 is still better than 23 for India. For 20/10 Myr bin/step, method 22 is still better than 23 for both North America and India.

(xii) Methods 16/17 are still 'bad' for North America and India.

## 4 DISCUSSIONS

The following discussions will be in Q&A style.

### *4.0.1 Question: Why the APP methods generally produce better similarities than AMP methods do?*

Paleomagnetic (Mean) A95 represents precision (how well constrained calculated poles are), and (mean) coeval poles' GCD represents accuracy (how close calculated poles are to the reference path; Fig. 25a and Fig. 25b). Compared with AMP, APP usually

**Table 5.** Equal-weight 120–0 Ma $\mathcal{CPD}$s for the three continents' paleomagnetic APWPs compared with their FHM predicted APWPs. The best are in dark green and underlined, second best in green and third in light green.

| Window, Step (size in Myr) | N America Pk 0 | | | | | | N America Pk 1 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Wt 0 | Wt 1 | Wt 2 | Wt 3 | Wt 4 | Wt 5 | Wt 0 | Wt 1 | Wt 2 | Wt 3 | Wt 4 | Wt 5 |
| 2, 1 | 0.2864 | 0.26 | 0.2877 | 0.2801 | 0.2632 | 0.2619 | 0.00747 | 0.00776 | 0.01688 | 0.01283 | 0.0372 | 0.0097 |
| 4, 2 | 0.08205 | 0.08422 | 0.1034 | 0.09653 | 0.10295 | 0.09863 | 0.0064145 | 0.00711 | 0.0182 | 0.011423 | 0.03606 | 0.00909 |
| 6, 3 | 0.06657 | 0.06817 | 0.06788 | 0.06229 | 0.08557 | 0.07634 | 0.00627 | 0.007797 | 0.01754 | 0.01254 | 0.02113 | 0.008955 |
| 8, 4 | 0.0614 | 0.0772 | 0.0653 | 0.06214 | 0.0903 | 0.0646 | 0.006 | 0.01099 | 0.01576 | 0.01399 | 0.02027 | 0.01271 |
| 10, 5 | 0.0349 | 0.0458 | 0.0486 | 0.046 | 0.0488 | 0.0454 | 0.0059 | 0.0062 | 0.0126 | 0.0151 | 0.0153 | 0.0094 |
| 12, 6 | 0.0318 | 0.0316 | 0.0325 | 0.0298 | 0.0323 | 0.0299 | 0.0087 | 0.009 | 0.017 | 0.0126 | 0.0191 | 0.0145 |
| 14, 7 (119-0 Ma path) | 0.0367 | 0.0348 | 0.0353 | 0.0369 | 0.0319 | 0.0352 | 0.00996 | 0.01198 | 0.023922 | 0.013 | 0.0203 | 0.0113 |
| 16, 8 | 0.0493 | 0.0492 | 0.0496 | 0.0486 | 0.0477 | 0.0485 | 0.0114 | 0.0117 | 0.023902 | 0.0123 | 0.0207 | 0.0129 |
| 20, 10 | 0.0497 | 0.0536 | 0.0538 | 0.0557 | 0.0517 | 0.0525 | 0.014 | 0.0174 | 0.029 | 0.0181 | 0.0233 | 0.0156 |
| 24, 12 | 0.0274 | 0.0304 | 0.0327 | 0.0324 | 0.0315 | 0.0313 | 0.0138 | 0.0143 | 0.0221 | 0.0191 | 0.0203 | 0.0192 |
| 30, 15 | 0.0345 | 0.0298 | 0.0317 | 0.0307 | 0.03402 | 0.0307 | 0.0174 | 0.01797 | 0.0276 | 0.02402 | 0.0252 | 0.02414 |

| Window, Step (size in Myr) | India Pk 0 | | | | | | India Pk 1 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Wt 0 | Wt 1 | Wt 2 | Wt 3 | Wt 4 | Wt 5 | Wt 0 | Wt 1 | Wt 2 | Wt 3 | Wt 4 | Wt 5 |
| 2, 1 | 0.249664 | 0.249951 | 0.249985 | 0.257479 | 0.250294 | 0.249702 | 0.0550264 | 0.0558274 | 0.0653912 | 0.0616122 | 0.0725822 | 0.0583844 |
| 4, 2 | 0.193452 | 0.202757 | 0.193438 | 0.241803 | 0.210005 | 0.194226 | 0.0570142 | 0.059523 | 0.0670142 | 0.0622828 | 0.0672158 | 0.0601206 |
| 6, 3 | 0.173961 | 0.174758 | 0.173975 | 0.229325 | 0.185139 | 0.175174 | 0.0578869 | 0.0588201 | 0.0639928 | 0.0627534 | 0.0672496 | 0.0624135 |
| 8, 4 | 0.154308 | 0.149658 | 0.151995 | 0.174118 | 0.165387 | 0.150164 | 0.0576095 | 0.0584774 | 0.0667115 | 0.0613883 | 0.0686417 | 0.05941 |
| 10, 5 | 0.1839 | 0.1845 | 0.1909 | 0.2586 | 0.1972 | 0.1852 | 0.0545 | 0.0554 | 0.0662 | 0.0589 | 0.066 | 0.06 |
| 12, 6 | 0.108924 | 0.105626 | 0.114955 | 0.118308 | 0.121497 | 0.105872 | 0.0598967 | 0.0629866 | 0.0657908 | 0.0644987 | 0.0695085 | 0.0629451 |
| 14, 7 (119-0 Ma path) | 0.112537 | 0.112885 | 0.126554 | 0.116174 | 0.132359 | 0.120914 | 0.0547932 | 0.0502588 | 0.0579931 | 0.060018 | 0.0654112 | 0.0582519 |
| 16, 8 | 0.104461 | 0.1241 | 0.124436 | 0.110942 | 0.119599 | 0.118336 | 0.0517351 | 0.0528129 | 0.0551881 | 0.056389 | 0.0574883 | 0.0550421 |
| 20, 10 | 0.1052 | 0.1015 | 0.1198 | 0.1096 | 0.1174 | 0.1072 | 0.0492 | 0.0501 | 0.0526 | 0.053 | 0.0577 | 0.052 |
| 24, 12 | 0.0531434 | 0.053561 | 0.056986 | 0.0574692 | 0.0555799 | 0.0553047 | 0.0519257 | 0.0459949 | 0.048681 | 0.0557455 | 0.0569792 | 0.0545615 |
| 30, 15 | 0.0737324 | 0.0754578 | 0.0775947 | 0.0575459 | 0.0565421 | 0.056635 | 0.0523614 | 0.0519862 | 0.054158 | 0.0563985 | 0.0555998 | 0.0543501 |

| Window, Step (size in Myr) | Australia Pk 0 | | | | | | Australia Pk 1 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Wt 0 | Wt 1 | Wt 2 | Wt 3 | Wt 4 | Wt 5 | Wt 0 | Wt 1 | Wt 2 | Wt 3 | Wt 4 | Wt 5 |
| 2, 1 | 0.589222 | 0.522048 | 0.523031 | 0.558687 | 0.544474 | 0.54495 | 0.00612207 | 0.00611382 | 0.025182 | 0.00696989 | 0.0355568 | 0.00581442 |
| 4, 2 | 0.268779 | 0.247297 | 0.251993 | 0.271535 | 0.353676 | 0.272178 | 0.00504377 | 0.00625515 | 0.029305 | 0.00753837 | 0.0335213 | 0.00689346 |
| 6, 3 | 0.0918779 | 0.0956757 | 0.0944798 | 0.0852725 | 0.100829 | 0.0922681 | 0.00488333 | 0.00536227 | 0.0208704 | 0.00654597 | 0.0236738 | 0.00644947 |
| 8, 4 | 0.0448139 | 0.0972008 | 0.0572843 | 0.0859725 | 0.0624494 | 0.0455424 | 0.00426485 | 0.00419086 | 0.0272409 | 0.00670754 | 0.0303 | 0.00757145 |
| 10, 5 | 0.0315 | 0.039 | 0.0509 | 0.0305 | 0.0601 | 0.031 | 0.0045 | 0.0048 | 0.0199 | 0.0058 | 0.0288 | 0.0089 |
| 12, 6 | 0.0288423 | 0.0376182 | 0.0594042 | 0.0279421 | 0.0569207 | 0.0285109 | 0.00692362 | 0.0072455 | 0.0220767 | 0.0102027 | 0.0220214 | 0.0111216 |
| 14, 7 (119-0 Ma path) | 0.0287639 | 0.0286588 | 0.0480163 | 0.0279289 | 0.0480853 | 0.0287173 | 0.00343606 | 0.00354029 | 0.0137145 | 0.00306601 | 0.0137356 | 0.0105205 |
| 16, 8 | 0.0299621 | 0.0398455 | 0.0556817 | 0.0293245 | 0.0512076 | 0.0299436 | 0.0115828 | 0.0057367 | 0.0201674 | 0.00643355 | 0.0162713 | 0.0123646 |
| 20, 10 | 0.0278 | 0.04 | 0.0612 | 0.0271 | 0.0425 | 0.0393 | 0.0079 | 0.0076 | 0.0197 | 0.0129 | 0.014 | 0.0106 |
| 24, 12 | 0.024437 | 0.0241335 | 0.0465703 | 0.0237562 | 0.0397431 | 0.036984 | 0.00584733 | 0.00637881 | 0.00599371 | 0.0056679 | 0.00547471 | 0.00516432 |
| 30, 15 | 0.0200862 | 0.0204561 | 0.0844176 | 0.0173471 | 0.0365786 | 0.034705 | 0.00412276 | 0.00444694 | 0.00448627 | 0.00397289 | 0.00369699 | 0.00387337 |

improves both and generates paths with higher accuracy and also higher precision (generally increasing N).

The fact that APP increases the number of paleopoles (N) in each sliding window would potentially average out some "bad" (i.e. inaccurate) poles and improves the fit between the paleomagnetic APWPs and the model-predicted APWPs. The general effects that APP brings include the decreases in paleomagnetic A95s, or/and distances between compared coeval poles of paleomagnetic APWP and reference APWP (Fig. 25 and Fig. 26). However, if the added paleopoles were all or mostly "bad", the improvement of fit would not occur. So the improvement of fit is not only because of the increase in N, but also because the majority of the additional poles are "good". AMP only regards the time uncertainty of each pole as one mid-point. Then this mid-point is treated as the most likely age of that mean pole. This is actually incorrect. The age uncertainty of paleopole is not obtained from a probability density function derived from an observed frequency distribution. As defined, the time uncertainty's lower (older) limit is a stratigraphic age, and its upper (younger) limit could be also a stratigraphic age or be constrained by a tectonic event using the field tests (e.g. fold/tilt test and conglomerate test). So the true age of the pole could be any one that is not older than the lower limit and also not younger than the upper limit. In other words, the mid-point could be the true age of the pole, but it is not known as the most likely age of that pole. If the

**Table 6.** Performance statistics of all the picking and weighting methods.

| Grid | | Best No. | | Worst No. | | Proportion of APP Better Than AMP | For All 28 Picking Methods, Count of Occurrences of Each Weighting No. Being Best | | | | | | Picking 14/15 (Studies After 1983) Better Than 16/17 (Older) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Picking | Weighting | Picking | Weighting | | 0 | 1 | 2 | 3 | 4 | 5 | |
| FHM | Fig. 8a | 1, 5, 7, 11, 13, 15, **19**, **21**, 25 | 0, 1, 5 | 2, 5, 7, 8, 14, **16**, 17, 18, 22, 26 | 0, 1, 2, 3, 4, 5 | 27/28 | **15** | 4 | 2 | 3 | 1 | 4 | Y/Y |
| | Fig. 8b | 4, 5, 6, 7, 9, **19**, **21** | 0, 1, 2, 3, 4, 5 | 0, 2, 8, 10, 12, **16**, 18, 20, 24 | 0, 1, 2, 3, 4, 5 | 11/14 | **17** | 1 | 2 | 7 | 1 | 0 | Y/Y |
| | Fig. 8c | 1, 11, 13, 17, **19**, **21**, 25 | 0, 1, 3, 5 | 2, 6, 14, **16**, 22, 26 | 0, 1, 2, 3, 4, 5 | 27/28 | **11** | 4 | 1 | 8 | 3 | 1 | N/N |
| | Fig. 16a | 1, 5, 7, 11, 13, 15, **19**, **21**, 25 | 0, 1, 3, 5 | 2, 6, 14, 22, 26 | 0, 1, 2, 3, 4, 5 | 3/4 | **18** | 6 | 0 | 2 | 4 | 3 | N(5n1y)/Y |
| | Fig. 16b | 4, 5, 6, 7, **19** | 0, 1, 2, 3, 4, 5 | 0, 2, 10, 12, **16**, 18, 20, 23, 24, 27 | 0, 1, 2, 3, 4, 5 | 59/84 | **12** | 6 | 0 | 9 | 0 | 1 | Y/Y(4y2n) |
| | Fig. 16c | 1, 11, 13, 17, **19**, **21**, 25 | 0, 1, 3, 4, 5 | 2, 14, **16**, 22, 26 | 0, 1, 2, 3, 4, 5 | 41/42 | 7 | 9 | 0 | **11** | 1 | 2 | N/N(4n2y) |
| MHM | Fig. 20a | 1, 5, 7, 11, 13, 15, **21**, 25 | 0, 1, 2, 3, 4, 5 | 0, 8, 10, 12, 14, **16**, 17, 18, 20, 22, 24 | 0, 1, 2, 3, 4, 5 | 11/12 | **14** | 6 | 1 | 1 | 2 | 4 | Y/Y |
| | Fig. 20b | 4, 5, 6, 7, **19**, 22, 26 | 0, 1, 2, 3, 4, 5 | 0, 2, 8, 10, 12, **16**, 18, 20, 24 | 0, 1, 2, 3, 4, 5 | 11/14 | **12** | 3 | 0 | 9 | 1 | 3 | Y/Y |
| | Fig. 20c | 1, 11, 13, 17, **19**, **21**, 25 | 0, 1, 3, 5 | 2, 14, **16**, 22, 26 | 0, 1, 2, 3, 4, 5 | 20/21 | 8 | 5 | 1 | **9** | 4 | 1 | N/N |
| | Fig. 22a | 5, 7, 11, 15, 22, 26 | 0, 1, 2, 3, 4, 5 | 0, 8, 10, 12, 14, **16**, 18, 20, 24 | 0, 1, 2, 3, 4, 5 | 3/4 | **15** | 6 | 1 | 0 | 1 | 5 | N(5n1y)/Y |
| | Fig. 22b | 4, 5, 6, 7, **19** | 0, 1, 2, 3, 4, 5 | 0, 2, 10, 12, **16**, 18, 23, 24, 27 | 0, 1, 2, 3, 4, 5 | 29/42 | **8** | **8** | 2 | 5 | 3 | 2 | Y/N(5n1y) |
| | Fig. 22c | 1, 11, 13, 15, 17, **19**, **21**, 23, 25, 27 | 0, 1, 2, 3, 4, 5 | 2, 14, **16**, 22, 26 | 0, 1, 2, 3, 4, 5 | 41/42 | **9** | 5 | 0 | 8 | 5 | 1 | N/Y(4y2n) |

**Figure 19.** Plot of the data shown in Table. 5. Note that here the step size is always half of the sliding window size and the reference path is the FHM derived.

mid-point is the most likely age of a pole, AMP should generate a path that is closer to the reference. However, mostly APP generates better similarities (See the high proportions of APP better than AMP in Table. 6). Most reasonably, the mid-point should be regarded as one possibility of all uniformly (not necessarily normally bell shaped, or U shaped, or left or right skewed) distributed ages between the two time limits.

So APP remains the effect of a paleopole borne on the mean poles during all the period of its age uncertainty, and use the increased number of paleopoles (N) to average out the negative effect of those "bad" poles, including the paleopoles that should not be included at that age for mean pole.

#### 4.0.2 *Question: Why the AMP methods sometimes unexceptionally produce better similarities than APP methods do?*

Because of small number of paleopoles (not necessarily "bad") involved in each sliding window, the produced mean poles by AMP should be relatively far from its contemporary model-predicted pole. In other words, AMP intends to give fairly small change in accuracy. This also could potentially bring more distinguishable $d_s$ for AMP if the corresponding A95 is not large enough. For example, for Fig. 8a, there are only three special (of 84 APP vs. AMP comparisons) cases picking/weighting 4/3, 4/5, 6/3 better than 5/3, 5/5, 7/3 respectively. Compared with the picking/weighting 4/3 APWP, although most of the mean paleopoles are closer to the FHM predicted APWP and also the number of the significant pole pairs is one less for the APP derived path (i.e. 5/3), the A95s are smaller and most importantly there are one more significant $d_a$ orientation-change pair and two more significant $d_l$ segment pair (Table. 7). If we observe carefully, it is because of the much smaller 15 Ma A95 for 5/3. The similar phenomenon occurs to the case of 6/3 vs 7/3, a relatively much smaller paleomagnetic A95 causes more distinguishable $d_a$ and $d_l$ for the APP results, and they offset the improvement of spatial similarity $d_s$ APP brings.

For 4/5 vs 5/5, all $d_a$ and $d_l$ are indistinguishable. Compared with the results from AMP, although the coeval pole GCDs are all decreased for APP, this spatial improvement is not able to offset the negative effects of also generally decreased paleomagnetic A95s, which potentially brings more statistically distinguishable coeval poles (e.g. the 15 Ma and 30 Ma poles for picking 5 and weighting 5; Table. 8). This further causes greater distinguishable mean $d_s$ from the APP methods. The similar phenomenon occurs to Fig. 8b picking 4 vs 5 with all the six types of weightings, Fig. 8c picking/weighting 4/2–3 vs 5/2–3, 4/5 vs 5/5.

(a) North America (101 with children): minimum 0.002198 (11(0)), maximum 0.0926 (16(5)), mean 0.0369, median 0.0332

(b) India (501): minimum 0.0242 (19(5)), maximum 0.5188 (8(3)), mean 0.1141, median 0.0746

(c) Australia (801 with children): minimum 0.0024 (17(5)), maximum 0.3662 (22(3)), mean 0.0773, median 0.0411

**Figure 20.** As Fig. 8, here the reference path is predicted from MHM. See the numbers of picked paleopoles in Fig. 8.

(a) North America: minimum 0.002198 (11(0))

(b) North America: maximum 0.0926 (16(5))

(c) India: minimum 0.0242 (19(5))

(d) India: maximum 0.5188 (8(3))

(e) Australia: minimum 0.0024 (17(5))

(f) Australia: maximum 0.3662 (22(3))

**Figure 21.** Path comparisons with best and worst difference values shown in Fig. 20. The parenthetical remarks are Picking No and Weighting No.

(a) North America (101 with children): minimum 0.0058 (15(5)), maximum 0.1041 (8(4)), mean 0.0439, median 0.0267

(b) India (501): minimum 0.0177 (6(3)), maximum 0.162 (16(3)), mean 0.0759, median 0.0621

(c) Australia (801 with children): minimum 0.0047 (23(4)), maximum 0.3158 (26(3)), mean 0.0643, median 0.031

**Figure 22.** As Fig. 16, here the reference path is predicted from MHM. See the numbers of picked paleopoles in Fig. 8.

(a) North America: minimum 0.0058 (15(5))

(b) North America: maximum 0.1041 (8(4))

(c) India: minimum 0.0177 (6(3))

(d) India: maximum 0.162 (16(3))

(e) Australia: minimum 0.0047 (23(4))

(f) Australia: maximum 0.3158 (26(3))

**Figure 23.** Path comparisons with best and worst difference values shown in Fig. 22. The parenthetical remarks are Picking No and Weighting No.

(a) Fig. 20a-Fig. 8a; Percentage for positive values: about 60.12%

(b) Fig. 20b-Fig. 8b; Percentage for positive values: about 27.98%

(c) Fig. 20c-Fig. 8c; Percentage for positive values: about 19.62%

**Figure 24.** Differences between results from two different reference paths, FHM (Fig. 8) and MHM (Fig. 20) derived. The absolute difference values less than 1.96-standard-deviation interval of the whole 168 values are labeled in green, more than 1.96-standard-deviation interval labeled in red.

**Figure 25.** Paleomagnetic APWPs' mean A95 versus (a) "mean GCD", (b) "mean significant GCD", (c) "mean significant orientation difference", and (d) "mean significant length difference" between paleomagnetic APWP and its corresponding FHM-and-plate-circuit predicted APWP. Starting points of the arrows are results from AMP, while ending points are from APP. Black color filled arrow heads are the small number of special cases of AMP derived equal-weight $\mathcal{CPD}$s better than APP (see details in Fig. 8).

**Figure 26.** Differences of APP and AMP coordinates shown in Fig. 25. Crosses locates the small minority cases of AMP derived equal-weight $\mathcal{CPD}$s better than APP (see details in Fig. 8).

In addition, compared with AMP, APP potentially could generate more mean poles, because sometimes for some sliding window there is no paleopole involved at all for AMP while there is(are) paleopole(s) involved for APP. For APP, the mean poles at all ages should be composed of more paleopoles than it is for AMP, which should generally decrease both coeval pole distance and paleomagnetic A95. However, sometimes a rare case (e.g. the 0 Ma comparison shown in Table. 9) happens. It is sometimes that an additional very "bad" paleopole gets included by APP and this increases both coeval pole distance and paleomagnetic A95 even though N increases. Such cases include Fig. 8b picking 22 vs 23 (actually exactly the same as picking 26 vs 27) with all the six types of weightings.

So generally as we discussed in the last section APP decreases the distances between paleomagnetic APWPs and the hotspot and ocean-floor spreading model predicted APWP, and also the uncertainties of paleomagnetic APWPs. However, as we described in this section, special cases like decreased A95 potentially intends to make coeval poles differentiated if the coeval poles' distance is not decreased effectively or even increased, or very "bad" paleopoles got involved in some sliding windows, occurs. In summary, when the negative effect from these types of rare cases is beyond the positive effect the generally improved mean poles contribute, the composite difference score would increase. However, this phenomenon seldom occur (Table. 6).

Other Type 1 (e.g. Table. 7) cases: Fig. 16a picking/weighting 22/3 vs 23/3. Fig. 20a 6/3 vs 7/3.

Other Type 2 (e.g. Table. 8) cases: Fig. 16a picking/weighting 2/0–5 vs 3/0–5, 4/0–2 vs 5/0–2, 4/4 vs 5/4, 6/5 vs 7/5, 6/2–3 vs 7/2–3, 22/4–5 vs 23/4–5, 22/1 vs 23/1, 26/4 vs 27/4, 26/2 vs 27/2. Fig. 16b 14/3 vs 15/3. Fig. 16c 4/2 vs 5/2. Fig. 20a 4/0–1 vs 5/0–1. Fig. 20b 4/0–4 vs 5/0–4, 22/0–1 vs 23/0–1, 22/3–4 vs 23/3–4, Fig. 20c 4/3 vs 5/3, 4/0 vs 5/0, 8/5 vs 9/5, 8/3 vs 9/3. Fig. 22b 14/2–3 vs 15/2–3. Fig. 22c 4/4 vs 5/4.

Combination of both Type 1 and 2 cases: Fig. 16a picking/weighting 26/0–1 vs 27/0–1. Fig. 20a 4/3 vs 5/3. Fig. 20b 22/2 vs 23/2, 22/5 vs 23/5, 26/0–5 vs 27/0–5.

Other Type 3 (e.g. Table. 9) cases: Fig. 16b picking/weighting 6/0–5 vs 7/0–5, 22/0–5 vs 23/0–5, 26/0–5 vs 27/0–5. Fig. 20a 4/2–5 vs 5/2–5. Fig. 22a 2/0–5 vs 3/0–5, 4/2 vs 5/2, 4/4 vs 5/4, 6/2 vs 7/2, 22/0–5 vs 23/0–5, 26/0–5 vs 27/0–5. Fig. 22b 6/0–5 vs 7/0–5, 22/0–5 vs 23/0–5, 26/0–5 vs 27/0–5.

#### *4.0.3 Question: Why weighting is not affecting?*

Generally, weighting does not affect the similarities dramatically, because the six results from the six weighting methods are mostly very close. There are a few special cases that one of them dramatically changes similarity (e.g. for picking method (Pk) 5 and weighting method (Wt) 3, etc.; Fig. 8a). As follows, those normal cases with close results are discussed first. Those rare cases are examined later.

For Pk 0 in the North America (101) example (Fig. 8a), the reason of why Wt 1–4 do not produce better similarity scores than Wt 0 (i.e. no weighting) does is examined here. Take the comparison between Wt 0 and Wt 1 (Table. 10) as an example, although APP indeed decreases the values of most distances from the reference and most A95s, APP also brings one more significant differentiated segment-length difference (80–85 Ma coeval segments; Fig. 27) because of the slightly larger A95 and shorter coeval pole GCD for 80 Ma Pk 0 Wt 1. Although they are distinguishable, the overlapping of their distributions is quite small (Fig. 27b). Wt 2 and 4 cause one more pair of distinguishable mean poles ($d_s$) than other weighting methods. However, compared with Wt 0, Wt 5 (i.e. combination of $\alpha 95$ and age error position to bin) does affect and bring the best similarity, because Wt 5 decreases the GCDs of the 20, 80 and especially 100 Ma distinguishable pole pairs (by more than 2 degrees; Table. 10).

For Fig. 8a Pk 1, Wt 0 generates the best similarity. Wt 2–5 cause at least 3 more pairs of distinguishable mean poles ($d_s$) than other weighting methods. Wt 1 increases the GCDs of all the distinguishable pole pairs (0, 5 and 120 Ma).

For Fig. 8a Pk 2, Wt 0 generates the best similarity. Wt 1–5 cause at least 1 more pair of distinguishable mean poles ($d_s$) than Wt 0 (60 Ma). Wt 4 also increases another pair of distinguishable mean poles ($d_s$) (80 Ma).

As the above examples shown, although no weighting is often better than weighting, the results are actually rather close to each other and also APP still improves precision and accuracy of most poles for the cases with larger difference scores, although APP brings extra significant differences in shape metrics which is also the main cause of their larger difference scores (Table. 10). However, APP does not always improves precision or accuracy, for example, for the Pt 5 (mainly or only igneous derived) and Wt 3 (according to $\alpha 95$) results (Table. 11). For Pt 5, the result with no weighting is still the best, and the results from Wt 0–2, 4, 5 are close and APP generally improves precision and accuracy, just like the above mentioned general cases. Only for Wt 3, APP worsens precision and improves accuracy of most poles at the same time. This makes the difference score dramatically larger than others. In addition, Wt 3 brings more significant differences in shape metrics, which causes the difference score even larger (Table. 11). That Wt 3 sometimes gives dramatical worse similarity could be related to the way we do the weighting according to the size of $\alpha 95$, i.e. the smaller the $\alpha 95$ is, the larger the assigned weight is (see more details about Weighting No. 3 in the Weighting paragraph). However, small size of $\alpha 95$ (high accuracy) could be becausei of those sampled directions not covering enough long period (thought to be at least about $10^4$ years) to "average out" secular variation for giving a paleopole. That is to say, the smallest $\alpha 95$s could get the greatest weights that they should not deserve.

Generally, weighting does not improve fit. In other words, Wt 0 is generaly the best. Wt 2 or 4 is not recommended, because they never have generated the best similarities (Table. 6), compared with other weighting methods. There is no general pattern about which weighting (of Wt 1–5) is better or worse. So weighting, for making a paleomagnetic APWP, is not absolutely necessary. However, there are some patterns about which weighting is better or worse for some specific continent. For example, Wt 3 prefers Australia (Table. 6). Wt 3 works fine with India. However, Wt 3 is not recommended for North America.

#### *4.0.4 Question: Why best and worst methods are not consistent?*

**4.0.4.1 Question: Why the picking method 21 is not among the best for 20 Myr binning and 10 Myr stepping with both space and shape tested?** As shown in Fig. 8, both the picking methods 19 and 21 are among the best. However, the picking method 21 is not one of the best any more for 20 Myr binning and 10 Myr stepping with both space and shape tested (Fig. 16). Further, in fact, in Fig. 16, we can see the picking method 21 is still one of the best for North America (101) and Australia (801), but just not for India (501). However, even for India (501), the difference values (ranging 0.0483–0.0535) produced by the picking method 21

**Table 7.** One example of the Type 1 rare cases where AMP gives better similarity result than APP does from North America (101). Only statistically significant values are listed here.

| FHM predicted | | picking 4 + weighting 3 | | | picking 5 + weighting 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | ds | | | ds | | | da | |
| Age (Ma) | A95 (°) | Age (Ma) | Pmag A95 (°) | Dist (°) | Age (Ma) | Pmag A95 (°) | Dist (°) | Age (Ma) | Diff (°) |
| 0 | 0 | 0 | 8.3 | **4.29** | 0 | 7.6 | **3.885** | ***10-15-20*** | **126.59** |
| 5 | 1.56039/0.87367 | 5 | 7 | **4.853** | 5 | 6.8 | **4.3453** | | |
| 10 | 2.89214/1.58743 | 10 | 12 | **9.91** | 10 | 8.6 | **5.79** | dl | |
| 15 | 2.575/1.63303 | 15 | 58/49 | **6.72** | 15 | 2.0857 | **11.8** | Age (Ma) | Diff (°) |
| 20 | 3.16077/2.20094 | ***20*** | 12.43 | **8.58** | | | | ***10-15*** | **13.52** |
| 25 | 4.96061/2.2183 | 25 | 6.76 | **6.96** | 25 | 6.3358 | **6.873** | ***15-20*** | **14.68** |
| 30 | 3.39692/2.57114 | 30 | 6.68 | **6.46** | 30 | 6.68 | **6.4583** | | |
| 115 | 9.27023/5.16012 | 115 | 8.6 | **12.2535** | 115 | 8.5 | **11.7** | | |
| 120 | 14.6882/8.12086 | 120 | 7.76 | **15.5744** | 120 | 7.728 | **15.258** | | |

**Table 8.** One example of the Type 2 rare cases where AMP gives better similarity result than APP does from North America (101). Only statistically significant values are listed here. Note that the number of the ages of the significant differentiated mean poles are the same for both paleomagnetic APWPs in this type of situation.

| Age | FHM predicted | ds | | | |
|---|---|---|---|---|---|
| | | picking 4 + weighting 5 | | picking 5 + weighting 5 | |
| (Ma) | A95 (°) | Pmag A95 (°) | Dist (°) | Pmag A95 (°) | Dist (°) |
| 0 | 0 | 8.3526 | **3.79** | 7.6664 | **3.391** |
| 5 | 1.56039/0.87367 | 7.5215 | **4.17** | 7.29 | **3.74** |
| 10 | 2.89214/1.58743 | 21.351 | **4.76** | 14.598 | **2.87** |
| 15 | 2.575/1.63303 | | | 10.153 | **10.71** |
| 30 | 3.39692/2.57114 | | | 6.9664 | **5.7245** |
| 115 | 9.27023/5.16012 | 9.635 | **11.935** | 9.452 | **11.276** |
| 120 | 14.6882/8.12086 | 8.0223 | **15.8183** | 7.943 | **15.511** |

**Table 9.** One example of the Type 3 rare cases where AMP gives better similarity result than APP does from India (501). Only statistically significant values are listed here. Note that for the bold-number ages, there is no mean poles at all for the "picking 22 (AMP) + weighting 0" case.

| FHM predicted | | picking 22 + weighting 0 | | | picking 23 + weighting 0 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | ds | | | ds | | | dl | |
| Age (Ma) | A95 (°) | Pmag A95 (°) | Dist (°) | N | Pmag A95 (°) | Dist (°) | N | Age (Ma) | Diff (°) |
| 0 | 0 | **6.85** | **12.973** | ***2*** | **23.6214** | **17.698** | ***3*** | **80-85** | 6.286 |
| 5 | 1.415/0.965 | 23.6214 | 15.941 | 3 | 23.6214 | 15.941 | 3 | **110-115** | 16.684 |
| 10 | 2.2425/1.34645 | 5.4/3.1 | 29.897 | 1 | 5.4/3.1 | 29.897 | 1 | | |
| **15** | 2.2694/1.62543 | | | | 5.4/3.1 | 28.28 | 1 | | |
| 70 | 8.53016/4.97567 | 4.0164 | 4.864 | 16 | 3.246 | 4.436 | 20 | | |
| **75** | 5.3554/3.1595 | | | | 5 | 4.477 | 1 | | |
| **80** | 8.41657/5.00588 | | | | 5 | 3.358 | 1 | | |
| 85 | 5.01489/2.49492 | 5 | 7.632 | 1 | 5 | 7.632 | 1 | | |
| **90** | 7.77997/2.86845 | | | | 5 | 10.884 | 1 | | |
| **95** | 4.46779/3.24941 | | | | 5 | 11.099 | 1 | | |
| **100** | 5.6124/5.19639 | | | | 5 | 11.4155 | 1 | | |
| **105** | 4.64657/3.49277 | | | | 5 | 14.908 | 1 | | |
| **110** | 9.11039/4.98436 | | | | 6.8/4.9 | 13.962 | 1 | | |
| 115 | 9.27023/5.16012 | 10.73 | 10.508 | 5 | 10.73 | 10.508 | 5 | | |

**Table 10.** My caption

| FHM predicted for 101 | | Pk 0 + Wt 0 | | Pk 0 + Wt 1 | | | | Pk 0 + Wt 5 | |
|---|---|---|---|---|---|---|---|---|---|
| | | ds | | ds | | dl | | ds | |
| Age (Ma) | A95 (°) | Pmag A95 (°) | Dist (°) | Pmag A95 (°) | Dist (°) | Age (Ma) | Diff (°) | Pmag A95 (°) | Dist (°) |
| 0 | 0 | 4.27286602 | 5.01 | 3.950661 | **5.05647** | **80-85** | **11.103** | 4.143 | 5.356 |
| 5 | 1.56039/0.87367 | 4.22350537 | 5.146 | 3.936601 | 5.1286 | | | 4.0534 | 5.407 |
| 10 | 2.89214/1.58743 | 20.9920176 | 3.076 | 19.826829 | **3.22624** | | | 19.868 | 3.36 |
| 15 | 2.575/1.63303 | 13.8698147 | 10.3 | 13.827757 | **10.34** | | | 13.85 | 10.2753 |
| 20 | 3.16077/2.20094 | 8.36501201 | 7.2162 | **8.463096** | 6.9973 | | | 8.413 | **6.7906** |
| 50 | 7.15565/3.22656 | 4.3991229 | 6.22 | **4.651517** | **6.273** | | | 4.3326 | 6.3563 |
| 55 | 7.17564/4.28065 | 5.6991191 | 8.647 | 5.670719 | **9.6724** | | | 5.52 | 8.53 |
| 60 | 9.71876/6.35204 | 7.71537555 | 9.498 | 6.889233 | 8.607 | | | 7.77 | 9.8237 |
| 80 | 8.76515/5.14459 | 6.29356332 | 9.26 | **6.452368** | 8.098 | | | 6.033 | **8.369** |
| 85 | 5.54221/2.65419 | 8.7/6.7 | 18.995 | 8.7/6.7 | 18.995 | | | 8.7/6.7 | 18.995 |
| 100 | 5.79659/5.36693 | 10.2720657 | 10.75658 | 9.286878 | 9.035 | | | 9.045 | **8.68** |
| 115 | 9.27023/5.16012 | 19.767437 | 9.074 | 18.483547 | **10.054** | | | 19.652 | 9.3547 |
| 120 | 14.6882/8.12086 | 3.56957955 | 17.3331 | 3.060561 | 17.062 | | | 3.606 | 17.47 |

are still closer to the left bound of the one-standard-deviation interval 0.0359–0.1072 and relatively farther from the mean 0.074, which means the picking method 21 is still a relatively better one.

#### *4.0.5 Question: Are there particular parts of the path that are more variable? Do different methods affect different parts of the path differently?*

The results may highlight the trade-off between more data diluting the effect of outliers, and fewer but 'better' data being more easily affected by a bad point that gets through the filters (Fig. 13, Fig. 14 and Fig. 15).

## 5 FINAL CONCLUSIONS

From the perspective of the general similarities between those paleomagnetic APWPs and the hot spot model and ocean-floor spreading model predicted APWPs, GAD hypothesis is proved valid for at least the last 120 Myr.

### 5.1 Universal Rules of Ways of Processing Paleomagnetic Data:

Although effects of filters (all the picking methods except Pk 1) have marginal change in reducing N (precision going down), some filters do improve the similarity score generally, for example, Pt 11 (APP without redbed-derived data) and 13 (APP with redbed-derived data corrected) are better than Pt 0 for all the three continents.

APP (adding data to a time window with overlapping age selection criterion) is better than AMP for making paleomagnetic APWPs, for both kinds of situations when there are lots of data (APP even better, e.g. for North America and Australia) and not much data (APP still a better option, e.g. for India; Table. 6). APP with filters (Pt 3, 5, 7, 9, 11, . . . , 27) is generally worse than APP without any filter (Pt 1).

Generally the APP methods produce better similarities than the AMP methods (Table. 6).

Actually weighting is not always affecting. In fact, for quite many of the methods, no weighting is the best performer (Table. 6). For example, score is likely worse for the combined methods of weight method 3 and AMP.

weighting is not always making similarities better

APP itself helps incorporate temporal uncertainty into the algorithm. With the bootstraps test helping incorporate spatial uncertainty into the algorithm together, both spatial and temporal uncertainties are successfully considered in APP methods.

### 5.2 Conditional Rules of Ways of Processing Paleomagnetic Data:

Picking method no. 16 (AMP with data from old studies) is not recommended, e.g. before 1983) for generating a paleomagnetic APWP.

### 5.3 Summary

According to the results we have from the three continents, North America, India and Australia, using the similarity measuring tool developed in Chapter 2, it is recommended that APP should be used to select the input paleopoles. According to all the paleomagnetic data we have from the three continents, the results from any size for sliding window and step are interestingly and extremely close to each other when the APP method is used, compared with the results from the AMP method (Table. 5 and Fig. 19). So any size for binning and stepping is ok when APP is used. Then filtering is actually not necessary. However, some filtering methods (e.g. Picking No. 5, 7 (igneous-derived), 11, 13 (nonredbeds or corrected redbeds derived), 19, 21 (non-local-rotation/reprinted or corrected-rotation derived) and 25 (non-superseded data derived)) are fine too and will not give worst or worse results than the other filtering methods (i.e. the left Picking methods). With APP used, weighting is actually not necessary either. If a weighting has to be used, Weighting No. 1 (related to the number of paleomagnetic sampling sites and observations) is generally better than the other four given weighting methods (Fig. 28).

If AMP has to be used, relatively wide sliding window and step are needed. According to our tests, more than 20/10 Myr is recommended. In addition, AMP works relatively better with igneous-derived data (i.e. Picking No. 4 and 6), which indicates that if we have fewer data, these data need to be better in quality (Fig. 28).

**Table 11.** Highest and lowest values for the same variable are highlighted in red and green respectively.

| FHM predicted for 101 | | Pk 5 + Wt 0 | | Pk 5 + Wt 1 | | Pk 5 + Wt 2 | | Pk 5 + Wt 3 | | | | Pk 5 + Wt 4 | | Pk 0 + Wt 5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ds | | ds | | ds | | ds | | da | | ds | | ds | |
| Age (Ma) | A95 (°) | Pmag A95 (°) | Dist (°) | Pmag A95 (°) | Dist (°) | Pmag A95 (°) | Dist (°) | Pmag A95 (°) | Dist (°) | Age (Ma) | Diff (°) | Pmag A95 (°) | Dist (°) | Pmag A95 (°) | Dist (°) |
| 0 | 0 | 7.458 | 2.058 | 7.4575 | 2.387 | 8.027 | 3.539 | 7.598 | **3.885** | 10-15-20 | **126.5907** | 7.7351 | 3.624 | 7.67 | 3.3909 |
| 5 | 1.56039/0.87367 | 7.3814 | 2.624 | 7.3814 | 2.995 | 7.887 | 3.876 | **6.8** | **4.3453** | | | 7.515 | 3.9475 | 7.29 | 3.74 |
| 10 | 2.89214/1.58743 | | | | | 15.208 | 3.402 | **8.602** | **5.79** | dl | | 16.783 | 4.2726 | 14.598 | 2.87 |
| 15 | 2.575/1.63303 | 12.421 | 9.2 | 12.4213 | 9.077 | 12.384 | 9.4 | **2.0857** | **11.805** | Age (Ma) | Diff (°) | 12.3843 | 9.4 | 10.153 | 10.71 |
| 20 | 3.16077/2.20094 | | | | | | | | | 10-15 | **13.52** | | | | |
| 25 | 4.96061/2.2183 | | | 6.463 | 6.2097 | | | **6.336** | **6.873** | 15-20 | **14.68** | 6.435 | 6.68 | | |
| 30 | 3.39692/2.57114 | | | | | | | **6.678** | **6.458** | | | | | 6.97 | 5.724 |
| 50 | 7.15565/3.22656 | | | | | | | | | | | 3.34 | 4.51 | | |
| 55 | 7.17564/4.28065 | | | | | | | | | | | 5.44 | 6.2034 | | |
| 65 | 7.37969/4.60029 | | | | | 7.6917 | 7.214 | | | | | | | | |
| 100 | 5.79659/5.36693 | | | | | | | | | | | 8.275 | 7.013 | | |
| 115 | 9.27023/5.16012 | | | | | 5.1 | 12.92 | 8.5 | 11.704 | | | 5.92 | 12.355 | 9.452 | 11.276 |
| 120 | 14.6882/8.12086 | 11.4266 | 13.435 | 11.4266 | 13.0664 | 4.7143 | 16.543 | 7.728 | 15.258 | | | 4.509 | 17.112 | 7.943 | 15.511 |

## ACKNOWLEDGMENTS

All images are produced using GMT (Wessel et al. 2013). Thanks to Ohio Supercomputer Center for their remote HPC resources.

## REFERENCES

Besse, J. & Courtillot, V., 2002. Apparent and true polar wander and the geometry of the geomagnetic field over the last 200 Myr, *J. Geophys. Res.*, **107**, 2300.

Bilardello, D. & Kodama, K. P., 2010. Palaeomagnetism and magnetic anisotropy of Carboniferous red beds from the Maritime Provinces of Canada: evidence for shallow palaeomagnetic inclinations and implications for North American apparent polar wander, *Geophys. J Int.*, **180**, 1013–1029.

Cottrell, R. D. & Tarduno, J. A., 2000. Late Cretaceous True Polar Wander: Not So Fast, *Science*, **288**, 2283.

Cande, S. C. & Stock, J. M., 2004. Pacific–Antarctic–Australia motion and the formation of the Macquarie Plate, *Geophys. J Int.*, **157**, 399–414.

Cande, S. C., Patriat, P. & Dyment, J., 2010. Motion between the Indian, Antarctic and African plates in the early Cenozoic, *Geophys. J Int.*, **183**, 127–149.

Christeson, G. L., Van Avendonk, H. J. A., Norton, I. O., Snedden, J. W., Eddy, D. R., Karner, G. D. & Johnson, C. A., 2014. Deep crustal structure in the eastern Gulf of Mexico, *J. Geophys. Res. Solid Earth*, **401**, 183–195.

Doubrovine, P. V. & Tarduno, J. A., 2008. Linking the Late Cretaceous to Paleogene Pacific plate and the Atlantic bordering continents using plate circuits and paleomagnetic data, *J. Geophys. Res. Solid Earth*, **113**, B07104.

DeMets, C., Gordon, R. G. & Argus, D. F., 2010. Geologically current plate motions, *Geophys. J Int.*, **181**, 1–80.

Domeier, M., van der Voo, R. & Denny, F. B., 2011. Widespread inclination shallowing in Permian and Triassic paleomagnetic data from Laurentia: Support from new paleomagnetic data from Middle Permian shallow intrusions in southern Illinois (USA) and virtual geomagnetic pole distributions, *Tectonophysics*, **511**, 38–52.

Domeier, M., Van der Voo, R., Tomezzoli, R. N., Tohver, E., Hendriks, B. W. H., Torsvik, T. H., Vizan, H. & Dominguez, A., 2011. Support for an "A-type" Pangea reconstruction from high-fidelity Late Permian and Early to Middle Triassic paleomagnetic data from Argentina, *Journal of Geophysical Research: Solid Earth*, **116**, B12114.

DeMets, C., Calais, E. & Merkouriev, S., 2017. Reconciling geodetic and geological estimates of recent plate motion across the Southwest Indian Ridge, *Geophys. J Int.*, **208**, 118–133.

Evans, D. A. D., 2003. True polar wander and supercontinents, *Tectonophysics*, **362**, 303–320.

Fisher, R. A., 1953. Dispersion on a sphere, *Proc. Roy. Soc. London Ser. A.*, **217**, 295–305.

Gaina, C., Torsvik, T. H., van Hinsbergen D. J. J., Medvedev, S., Werner, S. C. & Labails, C., 2013. The African Plate: A history of oceanic crust accretion and subduction since the Jurassic, *Tectonophysics*, **604**, 4–25.

Gaina, C., van Hinsbergen D. J. J. & Spakman, W., 2015. Tectonic interactions between India and Arabia since the Jurassic reconstructed from marine geophysics, ophiolite geology, and seismic tomography, *Tectonics*, **34**, 875–906.

Granot, R. & Dyment, J., 2018. Late Cenozoic unification of East and West Antarctica, *Nature Communications*, **9**, 3189.

Horner-Johnson, B. C., Gordon, R. G., Cowles, S. M. & Argus, D. F., 2005. The angular velocity of Nubia relative to Somalia and the location of the Nubia—Somalia–Antarctica triple junction, *Geophysical Journal International*, **162**, 221–238.

Jacox, E. H. & Samet, H., 2007. Spatial Join Techniques, *ACM Trans. Database Syst.*. **32**, 7.

King, R. F., 1955. The remanent magnetism of artificially deposited sediments, *Geophys. Suppl. Mon. Not. Roy. Astron. Soc. Lett.*, **7**, 115–134.

Krijgsman, W & Tauxe, L., 2004. Shallow bias in Mediterranean paleomagnetic directions caused by inclination error, *Earth Planet. Sci. Lett.*, **222**, 685–695.

Lemaux, J., II, Gordon, R. G. & Royer, J-Y, 2002. Location of the Nubia-Somalia boundary along the Southwest Indian Ridge, *Geology*, **30**, 339–342.

Müller, R. D., Royer, J. Y. & Lawver, L. A., 1993. Revised plate motions relative to the hotspots from combined Atlantic and Indian-Ocean hotspot tracks, *Geology*, **21**, 275–278.

McElhinny, M. W. & Lock, J., 1996. IAGA Paleomagnetic Databases with access, *Surveys in Geophysics*, **17**, 575–591.

Müller, R. D., Roest, W. R., Royer, J. Y., Gahagan, L. M. & Sclater, J. G., 1997. Digital isochrons of the world's ocean floor, *J Geophys Res Solid Earth*, **102**, 3211–3214.

Müller, R. D., Royer, J. Y., Cande, S. C., Roest, W. R. & Maschenkov, S., 1999. New constraints on the Late Cretaceous/Tertiary plate tectonic evolution of the Caribbean, *Sedimentary Basins of the World*. **4**, 33–59.

Marks, K. M. & Tikku, A. A., 2001. Cretaceous reconstructions of East Antarctica, Africa and Madagascar, *Earth and Planetary Science Letters*, **186**, 479–495.

McQuarrie, N. & Wernicke, B. P., 2006. An animated tectonic reconstruction of southwestern North America since 36 Ma, *Geosphere*, **1**, 147–172.

Müller, R. D., Sdrolias, M., Gaina, C. & Roest, W. R., 2008. Age, spreading rates, and spreading asymmetry of the world's ocean crust, *Geochemistry, Geophysics, Geosystems*, **9**, Q04006.

Müller, R. D., Seton, M., Zahirovic, S., Williams, S. E., Matthews, K. J., Wright, N. M., Shephard, G. E., Maloney, K., Barnett-Moore, N., Hosseinpour, M., Bower, D. J. & Cannon, J., 2016. Ocean basin evolution and global-scale plate reorganization events since Pangea breakup, *Annu Rev Earth Planet Sci*, **44**, 107–138.

Najman, Y., Appel, E., Boudagher-Fadel, M., Bown, P., Carter, A., Garzanti, E., Godin, L., Han, J., Liebke, U., Oliver, G., Parrish, R., Vezzoli, G., 2010. Timing of India-Asia collision: Geological, biostratigraphic, and palaeomagnetic constraints, *J. Geophys. Res. Solid Earth*, **115**, B12416.

O'Neill, C., Müller, R. D. & Steinberger, B., 2005. On the uncertainties in hot spot reconstructions and the significance of moving hot spot reference frames, *Geochem. Geophys. Geosyst.*, **6**, Q04003.

Pisarevsky, S. A., 2005. New edition of the Global Paleomagnetic Database, *Eos. Trans. AGU*, **86**, 170.

Patriat, P., Sloan, H. & Sauter, D., 2008. From slow to ultraslow: A previously undetected event at the Southwest Indian Ridge at ca. 24 Ma, *Geology*, **36**, 207–210.

Riisager, P., Hall, S., Antretter, M. & Zhao, X., 2004. Early Cretaceous Pacific palaeomagnetic pole from Ontong Java Plateau basement rocks, in *Origin and Evolution of the Ongong Java Plateau*, **229**, pp. 31–44, eds Fitton, J. G., Mahoney, J. J., Wallace, P. J. & Saunders, A. D., Geological Society of London Special Publications.

Rowan, C. J. & Rowley, D. B., 2016. Preserved history of global mean spreading rate: 83 Ma to present, *Geophys. J Int.*, **208**, 1173–1183.

Steinberger, B. & O'Connell, R. J., 1998. Advection of plumes in mantle flow: implications for hotspot motion, mantle viscosity and plume distribution, *Geophys. J Int.*, **132**, 412–434.

Schettino, A. & Scotese, C. R., 2005. Apparent polar wander paths for the major continents (200 Ma to the present day): a palaeomagnetic reference frame for global plate tectonic reconstructions, *Geophys. J Int.*, **163**, 727–759.

Shephard, G. E., Bunge, H.-P., Schuberth, B. S. A., Müller, R. D., Talsma, A. S., Moder, C. & Landgrebe, T. C. W., 2012. Testing absolute plate reference frames and the implications for the generation of geodynamic mantle heterogeneity structure, *Earth Planet. Sci. Lett.*, **317–318**, 204–217.

Torsvik, T. H., Smethurst, M. A., van der Voo, R., Trench, A., Abrahamsen, N. & Halvorsen, E., 1992. Baltica. A synopsis of Vendian-Permian palaeomagnetic data and their palaeotectonic implications, *Earth Sci. Rev.*, **33**, 133–152.

Torsvik, T. H. & Smethurst, M. A., 1999. Plate tectonic modelling: virtual reality with GMAP, *Comput. Geosci.*, **25**, 395–402.

Tauxe, L. & Kent, D. V., 2004. A simplified statistical model for the geomagnetic field and the detection of shallow bias in paleomagnetic inclinations: Was the ancient magnetic field dipolar? *Geophys. Monogr. AGU*, **145**, 101–115.

Tan, X., Kodama, K. P., Gilder, S., & Courtillot, V., 2007. Rock magnetic evidence for inclination shallowing in the Passaic Formation red beds from the Newark basin and a systematic bias of the Late Triassic apparent polar wander path for North America, *Earth Planet. Sci. Lett.*, **254**, 345–357.

Torsvik, T. H., Müller, R. D., van der Voo, R., Steinberger, B., & Gaina, C., 2008. Global plate motion frames: Toward a unified model, *Rev. Geophys.*, **46**, RG3004.

Torsvik, T. H., van der Voo, R., Preeden, U., Mac Niocaill, C., Steinberger, B., Doubrovine, P. V., van Hinsbergen, D. J. J., Domeier, M., Gaina, C., Tohver, E., Meert, J. G., McCausland, P. J. A. & Cocks, L. R. M., 2012. Phanerozoic polar wander, palaeogeography and dynamics, *Earth Sci. Rev.*, **114**, 325–368.

Tauxe L., Banerjee S.K., Butler R.F. & van der Voo R., 2019. *Essentials of Paleomagnetism*, 5th web edn, Available on line.

van der Voo, R., 1990. The reliability of paleomagnetic data, *Tectonophysics*, **184**, 1–9.

Whittaker, J. M., Müller, R. D., Leitchenkov, G., Stagg, H., Sdrolias, C. G. & Goncharov, A., 2007. Major Australian-Antarctic Plate Reorganization at Hawaiian-Emperor Bend Time, *Science*, **318**, 83–86.

Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J. & Wobbe, F.,2013. Generic Mapping Tools: Improved version released, *Eos. Trans. AGU*, **94**, 409–410.

Whittaker, J. M., Williams, S. E. & Müller, R. D., 2013. Revised tectonic evolution of the Eastern Indian Ocean, *Geochem. Geophys. Geosyst.*, **14**, 1891–1909.

Young, A., Flament, N., Maloney, K., Williams, S., Matthews, K., Zahirovic, S. & Müller, D., 2019. Global kinematics of tectonic plates and subduction zones since the late Paleozoic Era, *Geosci. Front.*, **10**, 989–1013.

## APPENDIX A: CONSTRAIN PALEOPOLES FOR A CERTAIN TECTONIC PLATE

A polygon can be drawn around a set of paleomagnetic data, whose sampling sites we believe belong to a specific plate or rigid block. Then the *Spatial Join* technique (Jacox & Samet 2007) helps join attributes from the polygon to the paleomagnetic data based on the spatial relationship allowing data within this polygon to be extracted from the whole raw large dataset without splitting a subset just for a specific plate. That allows us to quickly select subsets of the database based on geographic constraints just as easily as for age. Of course, the boundary of this polygon must be reasonably along a tectonic boundary. Regions like those close to the plate boundaries are usually tectonically active (e.g. local rotations), so we should also be careful when we deal with the paleopoles derived from this type of locations.

### A1 120–0 Ma North America

The data-constraining polygons are from the recently published plate model (Young et al. 2018) (Fig. A1). Plate ID 101 polygon in the recently published Plate Model (Young et al. 2018), including its children 108 (Avalon/Acadia block) and 109 (Piedmont block) polygons for 120–0 Ma, is used to select the sampling sites of the paleopoles for North America. According to the plate model rotation data (Young et al. 2018), 108 is fixed to 101 during the geologic period from Cretaceous to the present day. 109 is also fixed to 101 since about 300 Ma (Christeson et al. 2014). Then in order to be compared with the FHM (120–0 Ma) (Müller et al. 1993; Müller et al. 1999), the paleopoles with age ranging 120–0 Ma are further selected through constraining the lower magnetic age "LOMAGAGE <= 135" (here it is not 120 but 135, because for the lower resolution case when the window length is 30 Myr, the Age Position Picking method will include those data with their lower magnetic age between 120 Ma and 135 Ma). In addition, the RESULTNO=6007 dataset should also be included according to a published plate kinematic model (McQuarrie & Wernicke 2006) with a relatively higher resolution of polygons and rotations, although the dataset is in the PlateID=178 polygon. In the end, 193 datasets in total are extracted (both white circles and red triangle-inside-circles in Fig. A1).

Also based on this model of southwestern North America since 36 Ma (McQuarrie & Wernicke 2006), part of the paleopoles constrained by the four small western terranes whose Plate IDs are also 101 (white circles in Fig. A1) in fact had gone through regional rotations and here are removed. However, the poles with age younger than 10 Ma located within the largest western 101 terrane (on the south of the smallest western 101 terrane; corresponding to the RANGE_ID=74 polygon in the model (McQuarrie & Wernicke 2006)) should be included. So finally 135 of the 193 datasets remain (Fig. A1). Spatially North American paleomagnetic data are mainly from the western and eastern margins of the plate.

### A2 120–0 Ma India

Plate ID 501 polygons in the recently published Plate Model (Young et al. 2018) also include the two small polygons of the northern "Lesser Himalayan passive margin of Greater Indian Basin" and "Tethyan Himalayan microcontinent of Greater India" (Fig. A2). The polygons are used to select the sampling sites of the paleopoles for India (Fig. A2).

Based on the model of the tectonic interactions between India, Arabia and Asia since the Jurassic (Gaina et al. 2015) (Fig. A2),

part of the paleopoles constrained by the north two small terranes whose Plate IDs are also 501 in fact had gone through regional rotations and here are removed. So finally 75 datasets are left (Fig. A2). Spatially Indian paleomagnetic data are more evenly distributed on the India plate than North American and Australian poles.

### A3 120–0 Ma Australia

Plate ID 801 polygon in the recently published Plate Model (Young et al. 2018), including its children 675 (Sumba block) and 684 (Timor block) polygons for 120–0 Ma (Fig. A3), is used to select the sampling sites of the paleopoles for Australia. According to the plate model rotation data (Young et al. 2018), 675 and 684 are fixed to 801 during the geologic period from c.145 Ma to the present.

On the southeast of the main Australia plate (the blue polygon in Fig. A3), there is a triangle-shaped small polygon 850 (Tasmania block) which is fixed to 801 since c.100 Ma according to the (Young et al. 2018) rotation data. With that attribute, 805 contributes more data younger than c.100Ma for the later analysis. Ultimately the final 99 extracted datasets is shown in Fig. A3.

**Figure 27.** Significance testing on the 80–85 segment length difference between North American paleomagnetic APWPs derived from picking 0 and weightings 0 and 1 and the FHM predicted APWP (Fig. 5; Table. 10). (a) The thin segment through stars is from FHM predicted path, and the bold solid and dashed are from paleomagnetic paths. (b) The results from Wt 1 are differentiated (Fig. 14a and Fig. 14b).

**Figure 28.** Flowchart for recommended procedure of processing paleomagnetic data.

**Figure A1.** The final filtered datasets (red triangle-inside-circles) for later analysis on 120–0 Ma North America. Those poles that had been influenced by local tectonic rotations are shown as white circles.

**Figure A2.** The final filtered datasets (red triangle-inside-circles) for later analysis on 120–0 Ma India. Those poles that had been influenced by local tectonic rotations are shown as white circles. The rifts, faults and detachments (red lines) around India are used to filter out those data that are influenced by local tectonic rotations.

**Figure A3.** The final filtered datasets (red triangle-inside-circles) for later analysis on 120–0 Ma Australia. Those poles that had been influenced by local tectonic rotations are shown as white circles. The Plate ID 850 helps increase the amount of qualified datasets for 100–0 Ma.