

# **Global Paleomagnetic Data Analysis:**

## **Improved Methods of Reconstructing Plate Motions Using Paleomagnetic Data**



Chenjian Fu  
Department of Geology  
Kent State University

A thesis submitted for the degree of  
*Master of Science*

2021

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Background and Motivation . . . . .	2
1.1.1	Techniques Used in Relative and Absolute Plate Motion Studies . . . . .	2
1.1.2	Application of Paleomagnetism to Plate Tectonics . . . . .	4
1.1.3	Fact 1: Not All Regions on the Earth Surface Are Solid . . . . .	9
1.1.4	Fact 2: Not All Data Are Created Equal . . . . .	9
1.1.4.1	Age Uncertainty . . . . .	10
1.1.4.2	Position Uncertainty . . . . .	10
1.1.4.3	Data Consistency . . . . .	10
1.1.4.4	Data Density . . . . .	12
1.1.4.5	Publication Year . . . . .	12
1.2	Objectives . . . . .	13
1.2.1	Motivation and General Approach . . . . .	13
1.2.2	Research Questions or Hypotheses . . . . .	14
1.2.2.1	Question 1 . . . . .	14
1.2.2.2	Question 2 . . . . .	15
1.2.2.3	Question 3 . . . . .	15
1.2.2.4	Question 4 . . . . .	15
1.2.2.5	Question 5 . . . . .	15
1.2.2.6	Question 6 . . . . .	15
<b>2</b>	<b>Methodologies</b>	<b>16</b>
2.1	Introduction . . . . .	17
2.2	Methods . . . . .	19
2.2.1	Comparing Apparent Polar Wander Paths (APWPs) . . . . .	19
2.2.2	APWP Pairs Used in This Study . . . . .	22
2.2.3	Significant Spatial Difference . . . . .	23
2.2.4	Shape Difference . . . . .	23

2.2.4.1	Mean Length Difference . . . . .	23
2.2.4.2	Mean Angular Difference . . . . .	25
2.2.4.3	Significance Testing of Shape Difference . . . . .	25
2.2.5	Composite Path Difference ( $\mathcal{CPD}$ ) . . . . .	27
2.2.6	Fit Quality (FQ) . . . . .	27
2.3	Results and Discussion . . . . .	28
2.3.1	Discrimination of Difference Metrics . . . . .	28
2.3.1.1	$d_s$ . . . . .	29
2.3.1.2	$d_l$ . . . . .	30
2.3.1.3	$d_a$ . . . . .	31
2.3.1.4	$\mathcal{CPD}$ . . . . .	32
2.3.2	A Discussion on Weights . . . . .	32
2.3.3	Application to Real Paleomagnetic Data . . . . .	34
2.4	Conclusions . . . . .	38
<b>3</b>	<b>Methods for Producing a Reliable APWP</b>	<b>40</b>
3.1	Introduction . . . . .	41
3.1.1	How An APWP Is Generated and Key Factors . . . . .	42
3.1.1.1	Picking/Binning Step . . . . .	43
3.1.1.2	Filtering Step . . . . .	43
3.1.1.3	Calculation Step . . . . .	44
3.1.2	Existing Solutions and General Issues . . . . .	45
3.2	Methods . . . . .	46
3.2.1	General Approach . . . . .	46
3.2.2	Paleomagnetic Data . . . . .	46
3.2.2.1	GPMDB Field Codes . . . . .	46
3.2.2.2	Paleomagnetic Data of Three Representative Continents	46
3.2.3	APWP Generation . . . . .	49
3.2.3.1	Picking/Binning . . . . .	52
3.2.3.2	Filtering . . . . .	54
3.2.3.3	Weighting . . . . .	57
3.2.3.4	Scaling of Weights . . . . .	60
3.2.4	Reference Paths . . . . .	60
3.2.5	Comparison Algorithm . . . . .	65
3.3	Results . . . . .	65

3.3.1	Baseline Results: 10/5 Myr Window/Step, Fixed Hotspot Reference . . . . .	65
3.3.2	Effects of Windowing Method . . . . .	66
3.3.3	Effects of Filtering and Weighting . . . . .	69
3.3.3.1	Filter Aggression . . . . .	72
3.3.3.2	Filter Performance . . . . .	73
3.3.3.3	Weighting Performance . . . . .	75
3.3.4	Effects of Window Size . . . . .	77
3.3.4.1	Overall Change . . . . .	80
3.3.4.2	Relative Performance of Methods . . . . .	80
3.3.4.3	Other Window Sizes and Steps . . . . .	84
3.3.5	Moving versus Fixed Hotspot Reference . . . . .	87
3.3.5.1	Overall Change . . . . .	87
3.3.5.2	Relative Performance of Methods . . . . .	87
3.4	Discussion . . . . .	95
3.4.1	Question: Why APP Methods Generally Produce Better Similarities than AMP Methods? . . . . .	95
3.4.1.1	Perspective of Conceptual Difference between AMP and APP . . . . .	95
3.4.1.2	Perspective of Stability Comparison between AMP and APP . . . . .	97
3.4.1.3	Summary . . . . .	98
3.4.2	Question: Why AMP Methods Sometimes Unexceptionally Produce Better Similarities than APP Methods? . . . . .	98
3.4.3	Question: Why Weighting Is Not Affecting? . . . . .	101
3.4.4	Question: Why Best and Worst Methods Sometimes Are Not Consistent? . . . . .	107
3.4.5	Question: Are There Particular Parts of the Path That Are More Variable? Do Different Methods Affect Different Parts of the Path Differently? . . . . .	107
3.5	Final Conclusions . . . . .	108
3.5.1	Universal Rules of Ways of Processing Paleomagnetic Data . .	108
3.5.2	Conditional Rules of Ways of Processing Paleomagnetic Data .	109
3.5.3	Summary . . . . .	109

<b>Appendices</b>	<b>120</b>
<b>Appendix A Methods for Manipulating Data on A Globe in Chapter 2</b>	<b>121</b>
A.1 Test if A Coeval Pole Pair Is Distinguishable with the Bootstrap . . . . .	121
A.2 Bivariate Sampling . . . . .	122
A.3 Synchronization . . . . .	123
A.3.1 Equally Treated Random Weights . . . . .	124
<b>Appendix B Methods for Constraining Paleopoles for Tectonic Plates in Chapter 3</b>	<b>127</b>
B.1 120–0 Ma North America . . . . .	127
B.2 120–0 Ma India . . . . .	129
B.3 120–0 Ma Australia . . . . .	129

# List of Figures

1.1	Geocentric axial dipole model . . . . .	5
1.2	The hemispheric ambiguity and absolute paleolongitude indeterminacy with a single paleomagnetic pole (paleopole) . . . . .	7
1.3	All published paleomagnetic data from North America . . . . .	8
1.4	Example of AMP moving averaging effects . . . . .	11
2.1	Parts of APWPs of supercontinent fragments share the same geometry	18
2.2	Examples showing great circle distance is a bad index of similarity .	19
2.3	Eight examples of APWP comparisons . . . . .	21
2.4	Geometric difference definition between two APWPs . . . . .	24
2.5	Testing on geometry . . . . .	26
2.7	Mean spatial, length, angular differences . . . . .	30
2.8	Mean spatial, length, angular differences (without interpolations) .	31
2.9	Criteria of pair comparisons . . . . .	35
2.10	Reproducing Torsvik et al. 2012 Earth Sci Rev Fig.13a . . . . .	36
3.1	Distribution of 120–0 Ma North American paleopoles . . . . .	48
3.2	Distribution of 120–0 Ma Indian paleopoles . . . . .	50
3.3	Distribution of 120–0 Ma Australian paleopoles . . . . .	51
3.4	How moving average methods work . . . . .	53
3.5	Visualization of equation that relates $\alpha_{95}$ , and R and N . . . . .	61
3.6	120–0 Ma MHM vs FHM predicted APWP of North America . . .	62
3.7	120–0 Ma MHM vs FHM predicted APWP of India . . . . .	63
3.8	120–0 Ma MHM vs FHM predicted APWP of Australia . . . . .	64
3.9	$\mathcal{CPD}$ of each plate’s paleomagnetic APWPs vs its FHM predicted APWP	67
3.10	Best and worst $\mathcal{CPDs}$ (10/5 Myr window/step) . . . . .	68
3.11	$\mathcal{CPD}$ of each plate’s paleomagnetic APWPs vs its FHM predicted APWP (AMP vs APP) . . . . .	70

3.13	Differences of $\mathcal{CPD}$ of each plate's paleomagnetic APWPs vs its FHM predicted APWP . . . . .	74
3.14	$d_s$ of each pair of poles for North American 10/5 Myr APWPs . . . . .	75
3.15	$d_l$ of each pair of segments for North American 10/5 Myr APWPs . . . . .	76
3.16	$d_a$ of each pair of segment-orientation-changes for North American 10/5 Myr APWPs . . . . .	77
3.17	$\mathcal{CPD}$ of each plate's paleomagnetic APWPs vs its FHM predicted APWP (20/10 Myr window/step) . . . . .	78
3.18	Best and worst $\mathcal{CPDs}$ (20/10 Myr window/step) . . . . .	79
3.20	Sliding window and step sizes vs $\mathcal{CPD}$ . . . . .	86
3.21	$\mathcal{CPD}$ of each plate's paleomagnetic APWPs vs its MHM predicted APWP . . . . .	88
3.22	Best and worst $\mathcal{CPDs}$ (10/5 Myr window/step; MHM) . . . . .	89
3.23	$\mathcal{CPD}$ of each plate's paleomagnetic APWPs vs its MHM predicted APWP (20/10 Myr window/step) . . . . .	90
3.24	Best and worst $\mathcal{CPDs}$ (20/10 Myr window/step; MHM) . . . . .	91
3.25	Differences between results from FHM and MHM . . . . .	92
3.26	APP spatially better than AMP (arrow) . . . . .	95
3.27	APP spatially better than AMP (dot) . . . . .	96
3.28	Paleomagnetic APWP's FQ score vs significant difference score (i.e. $\mathcal{CPD}$ ) . . . . .	103
3.29	Paleomagnetic APWP's FQ score vs raw difference score (i.e., no statistical testing) . . . . .	104
3.30	Proportional changes of Wt 1–5 to 0: Paleomagnetic APWP's FQ score vs $\mathcal{CPD}$ . . . . .	105
3.31	Proportional changes of Wt 1–5 to 0: Paleomagnetic APWP's FQ score vs raw difference score . . . . .	106
3.32	Flowchart of making a reliable paleomagnetic APWP . . . . .	110
A.1	How $N \leq 1$ uncertainty ellipse is simulated . . . . .	124
A.2	Comparisons of Pairs a-g with random weights involved . . . . .	125
B.1	Final filtered datasets for analysis on 120–0 Ma North America . . . . .	128
B.2	Final filtered datasets for analysis on 120–0 Ma India . . . . .	130
B.3	Final filtered datasets for analysis on 120–0 Ma Australia . . . . .	131

# List of Tables

3.1	List of the used fields and field codes of the GPMDB. . . . .	47
3.2	List of all picking/binning algorithms developed here. . . . .	54
3.3	List of all weighting algorithms developed here. . . . .	57
3.4	Consistency check on comparisons of picking methods' performance between 20/10 and 10/5 Myr window/step. Notes: E means expected; UE means unexpected. . . . .	83
3.5	Equal-weight 120–0 Ma $\mathcal{CPD}s$ for the three representative 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. . . . .	85
3.6	Performance statistics of all the picking and weighting methods for 10/5 and 20/10 Myr window/step path comparisons. . . . .	94
3.7	One example of the Type 1 rare cases where AMP gives better similarity result than APP does from North America (10/5 Myr window/step). Only statistically significant values are listed here. . . . .	100
3.8	One example of the Type 2 rare cases where AMP gives better similarity result than APP does from North America (20/10 Myr window/step). Only statistically significant values are listed here. . . . .	100
3.9	One example of the Type 3 rare cases where AMP gives better similarity result than APP does from India (10/5 Myr window/step). Only statistically significant values are listed here. Note that for the bold-number ages, there is no mean poles at all for the “Pk 22 (Wt 2)” case. . . . .	101
B.1	Rotation parameters for making three reference paths . . . . .	132
B.2	Calculated North American, Indian and Australian reference APWPs (5 Myr interval) in the hotspot references . . . . .	137

# **Chapter 1**

## **Introduction**

*The first chapter introduces paleomagnetism-based paleogeographic reconstruction technique and highlights the motivation of the research conducted in the thesis.*

## 1.1 Background and Motivation

Reconstructing past paleogeographies, especially the motion of plates and their interactions through time, is a key component of understanding the Earth’s geological history, including deciphering tectonics (e.g. supercontinent reconstruction), paleoclimate history, and the evolution of life. Since the advent of plate tectonics, it has been the background for nearly all geologic events. In addition, plate reconstructions form the basis of global or regional geodynamic models.

### 1.1.1 Techniques Used in Relative and Absolute Plate Motion Studies

The earliest quantitative effort to model plate kinematics was fitting conjugate passive margins of the Atlantic (Bullard et al., 1965; Wessel et al., 2007). They showed that the Atlantic could be closed using a single Euler pole (using Euler’s theorem on rotation). Then it became fitting conjugate isochrons based on best-fitting marine magnetic anomaly and fracture zone data (McKenzie and Sclater, 1971), which minimizes the misfit area between two isochrons. The *Hellinger* method (Hellinger, 1981) is a more advanced and generalised method which also fits conjugate isochrons based on best-fitting marine magnetic anomaly and fracture zone data, which however minimizes the sum of the misfits of conjugate data points that belong to a common isochron segment (Wessel et al., 2007) instead. These techniques mainly through fitting conjugate lines mentioned above are relatively accurate for quantitative analysis. However they give relative, not absolute, motions between plates, because plate motions can’t be tied into absolute location on Earth’s surface, since both plates are likely moving. In addition, they are limited to survey data from the seafloor, with a maximum age of no more than  $\sim$ 200 Ma (Müller et al., 2008).

Reference frames are a means of describing the motion of geologic features (e.g. tectonic plates) on the surface of the Earth, relative to a common point or “frame” (Shephard et al., 2012). An absolute reference frame is a frame that can be treated as fixed relative to the Earth’s geographic reference frame. In reality, it’s impossible to find a truly absolute reference frame, so we are actually looking for a frame that has limited (and hopefully known) motion, which approximates as “fixed” over geologically useful timescales and provides the most complete descriptions of plate motions. A commonly used absolute reference frame is the “Fixed Hotspot model” (Müller et al., 1993, 1999), covering ages from  $\sim$ 132 Ma to present-day, which assumes that the

linear volcanic chains found on most oceanic plates are artifacts of absolute plate motions over a upwelling plume from the deep mantle, which is assumed to be relatively fixed. The advantage of this “Fixed hotspot model” is that it is fairly straightforward if the assumption of fixed hotspots is correct. However, this model is limited to plates with well-dated volcanic hotspot chains (e.g. the Ninetyeast Ridge on the Indian Ocean floor and the Walvis Ridge in the southern Atlantic Ocean; see O’Neill et al., 2005) and dating can be difficult (e.g. diffuse volcanic centers possibly related to large diameter plume conduits could cause the existence of time reversals; see O’Neill et al., 2005). As for not well-dated hotspot tracks, for example, only about 5% of the seamounts (thought to be volcanic) in the Pacific are thought to be related to hotspot volcanism and radiometrically dated (39 per cent of these ages are less than 10 Ma; see Hillier, 2007). In addition, the fixed hotspot model is mostly confined to existing oceanic or thin continental crust because older oceanic lithosphere has been largely destroyed by subduction and old, thick continental crust mostly removed by erosion (Chu et al., 2013). Last, but not least, hotspots can be susceptible to drift that may be caused by changes in sub-lithospheric mantle flow (Tarduno et al., 2009). Generally, however, the drift rate is considered to be an order of magnitude less than the rate of plate motions, so only becomes significant over timescales of  $\sim$ 50 Myr or more (O’Neill et al., 2005; Tarduno, 2007). To overcome this source of error, the “Moving hotspot model” (O’Neill et al., 2005) uses mantle convection modeling to predict hotspot drift. Some are apparent success, e.g. by getting motions in the Indo-Atlantic and Pacific hotspot clusters to agree with each other, but it’s very dependent on the mantle model. Hybrid models attempt to overcome the shortcomings of each reference frame by combining them, e.g. combining a fixed hotspot frame from 100 Ma to 0 Ma (Müller et al., 1993) with a moving hotspot frame from  $\sim$ 132–100 Ma (O’Neill et al., 2005) (Hybrid hotspot model (Shephard et al., 2012)), combining a moving hotspot frame from 100–0 Ma (O’Neill et al., 2005) with a paleomagnetic model (reflect plate motion relative to the magnetic dipole axis but cannot provide paleolongitudes because of the axial symmetry of the Earth’s magnetic dipole field) (Torsvik et al., 2008) from 140–100 Ma (Hybrid paleomagnetic model (Shephard et al., 2012)), and combining a moving hotspot frame from 120–100 Ma (O’Neill et al., 2005) with a true polar wander (TPW) corrected paleomagnetic model (Steinberger and Torsvik, 2008) from 100–0 Ma (Hybrid TPW-corrected model (Shephard et al., 2012)).

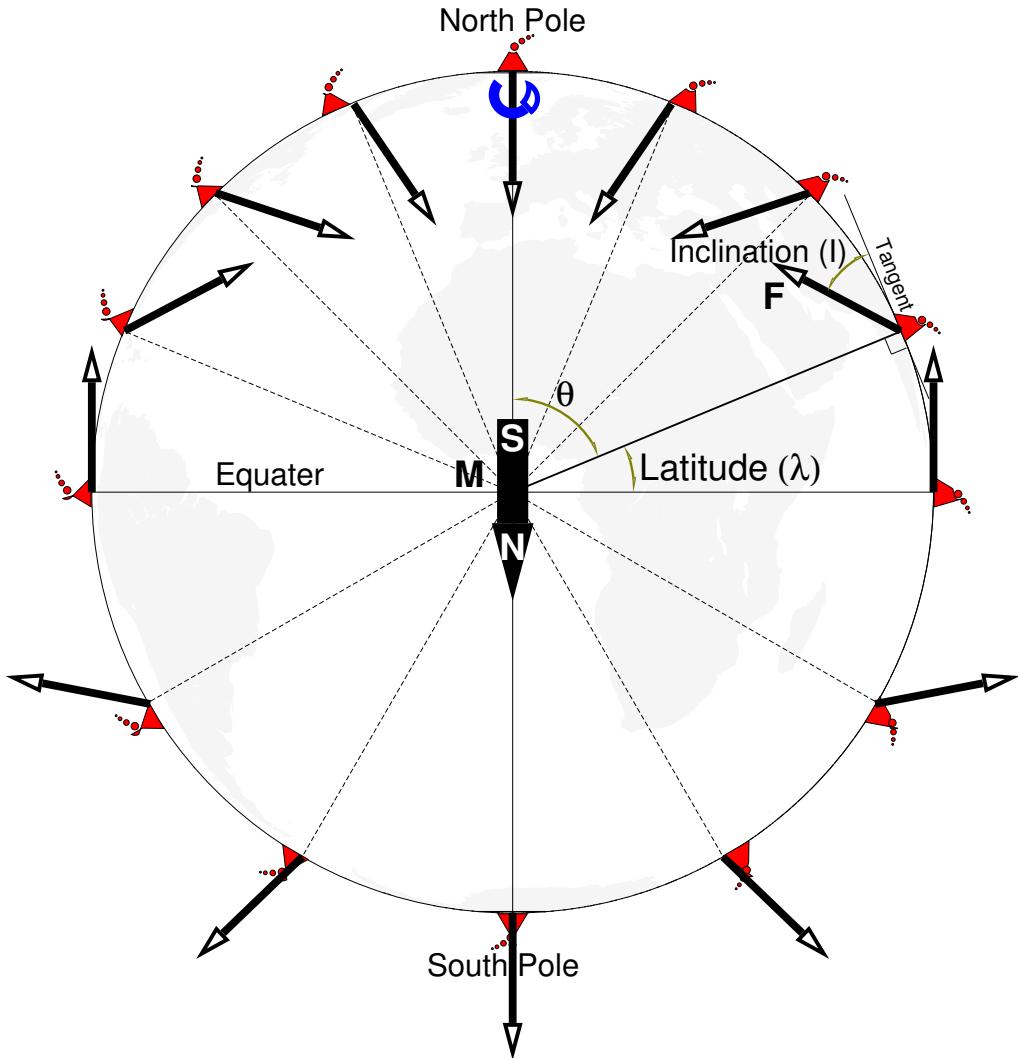
Recently another absolute reference frame “Subduction reference model” (van der Meer et al., 2010) tries to connect orogenies/sutures/subduction complexes’ on the

Earth surface with their corresponding subducted slabs in the mantle. Assuming that these remnants sank vertically through the mantle, the absolute location at which they were subducted can be reconstructed. In this way, this model mainly imposes a longitude correction on the above mentioned “Hybrid TPW-corrected model”, and can theoretically give past absolute locations of plates back to  $\sim$ 260 Ma based on the estimated age of the oldest slab remnants that can be reliably located in the mantle. While the “Subduction reference model” allows for reconstructions between  $\sim$ 260 Ma and 140 Ma, older than the other absolute models can predict, the model is strongly dependent on the vertical subduction assumption and resolution of seismic tomography models, so its uncertainty is high. Above all, importantly, if we can describe the absolute motion of one or a few key plates, the techniques for establishing the relative plate motions described in the second paragraph above can be used to construct plate circuits that allow a full kinematic description of plate tectonics to be developed.

As we can see, all of these above reconstruction methods are limited to recent geological history. For most of Earth history, concretely for times before  $\sim$ 170 Ma, the age of the oldest magnetic anomaly identification, paleomagnetism is the only accepted quantitative method for reconstructing plate motions and past paleogeographies.

### 1.1.2 Application of Paleomagnetism to Plate Tectonics

The geomagnetic field is generated by the convective flow of a liquid iron-nickel alloy in the outer core of the Earth. It is largely dipolar and can be represented by a dipole that points from the north magnetic pole to the south pole. However, the geomagnetic field varies in strength and direction over decadal–millennial timescales due to quadropole and octopole components of the field. The most spectacular variations in direction are occasional polarity reversals (normal polarity: the same as the present direction of the field; or the opposite, i.e. reverse polarity). Over a period of a few thousand years, the magnetic axis slowly rotates/precesses around the geographic axis and the Earth’s rotation axis (secular variation), but when averaged over 10,000 year timescales, higher order components of the field are thought to largely cancel out and the position of the magnetic poles aligns with the geographic poles. This is the geocentric axial dipole (GAD) hypothesis. In a GAD field, at the north magnetic pole the inclination (angle with respect to the local horizontal plane, see Fig. 1.1) of the field is  $+90^\circ$  (straight down), at the Equator the field inclination is  $0^\circ$  (horizontal) pointing north and at the south magnetic pole the inclination is  $-90^\circ$  (straight up)



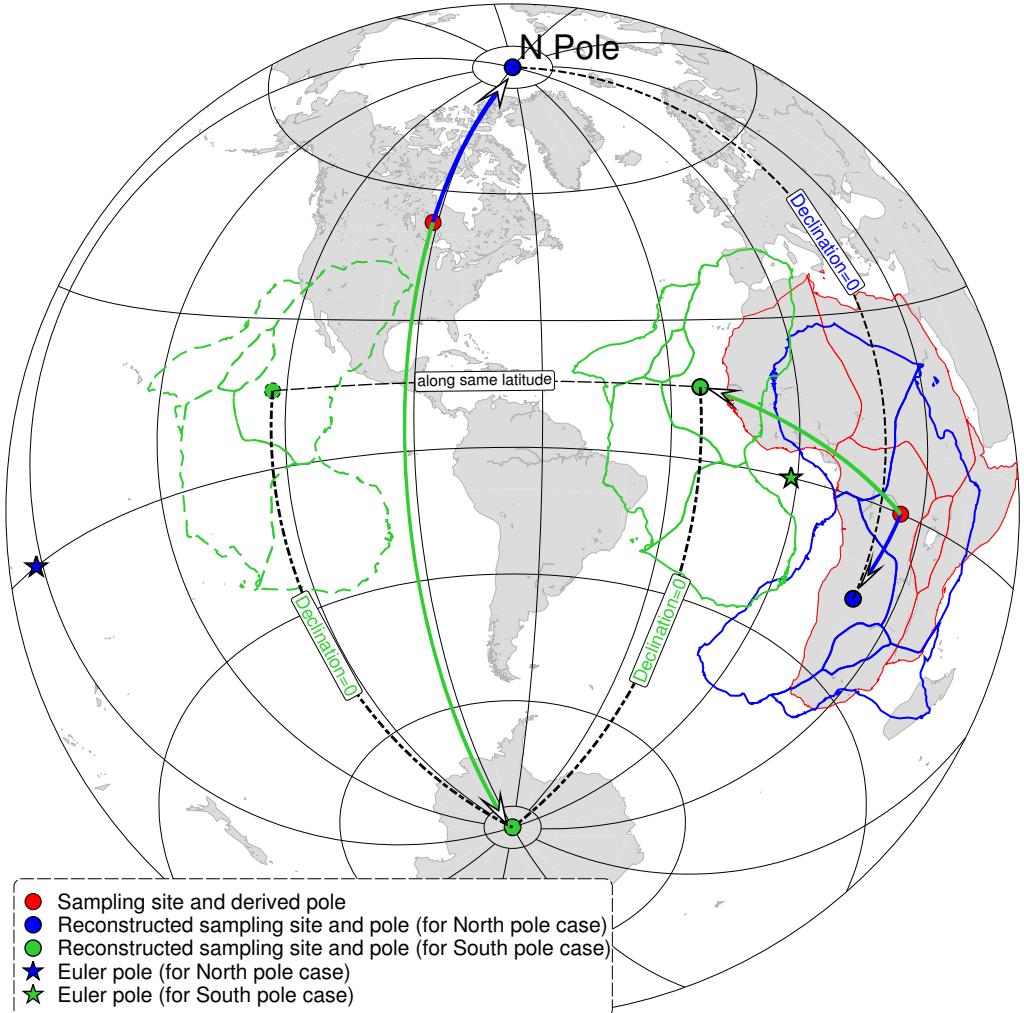
**Figure 1.1:** GAD model: Inclination ( $I = \tan^{-1}(2 \tan \lambda)$ ) of the Earth's magnetic field and how it varies with latitude, redrawn from Butler (1992); Tauxe et al. (2019); Torsvik et al. (2008). Magnetic dipole  $M$  is placed at the center of the Earth and aligned with the rotation axis;  $\lambda$  is the geographic latitude, and  $\theta$  is the colatitude.

(Fig. 1.1). Another direction parameter of the Earth's magnetic field is declination. It is the angle with respect to the geographic meridian, which is  $0^\circ$  everywhere in a time-averaged GAD field.

Magnetic remanence is the magnetization left behind in a ferromagnetic substance in the absence of an external magnetic field (Tauxe et al., 2019). The remanent magnetisation of rocks can preserve the direction and intensity of the geomagnetic field when the rock was formed, e.g. in the process of cooling, ferromagnetic materials in the lava flow are magnetized in the direction of the Earth's magnetic field, so the local direction of the field vector is locked in solidified lava. We are often interested

in whether the geomagnetic pole has changed, or whether a particular plate/terrane has rotated with respect to the geomagnetic pole (Tauxe et al., 2019). By measuring the direction of the remanent magnetisation, we can calculate a virtual geomagnetic pole (VGP) to represent the geomagnetic pole of an imaginary geocentric dipole which would give rise to the observed remanent declination and inclination. Collection of VGPs (or site-mean directions) allow calculating a “paleomagnetic pole”, also known as paleopole, at the formation level. Commonly a paleopole is a Fisherian mean (Fisher, 1953) (of VGPs) with a spatial uncertainty. A paleopole that plots away from the present geographic poles is assumed to be due to plate motions since the lava was solidified, which causes the paleopole to move with the plate (Torsvik et al., 2008). Based on measurements of the remanent inclination, the ancient latitude for a plate can be calculated when the rock formed from the dipole formula  $\tan(I) = 2 * \tan(latitude)$ . In addition, the remanent declination provides information about the rotation of a plate. Ideally, as a time average, a paleopole (which can be calculated from declination, inclination and the current geographic location of the sampling site) for a newly formed rock will correspond with the geographic north or south pole. To perform a reconstruction with paleopoles we therefore have to calculate the rotation (Euler) pole and angle which will bring the paleopole back to the geographic north or south pole, and then rotate the plate by the same amount of angle using the same Euler pole. This is how paleomagnetism can be used to reconstruct past positions of a plate. In our example (Fig. 1.2), a  $\sim$ 155 Ma paleopole (latitude=52.59°N, longitude=91.45°W) will be restored to the geographic pole by an Euler rotation of pole (0°, 178.55°E) with angle 37.41°, which rotates the sampling site from its present position of (0°, 25°E) to the Africa paleo-continent at (15.6998°S, 20.1121°E). So Africa must have drifted northwards since the Late Jurassic.

However, there are 2 problems with using paleopoles for constraining finite rotations (Tauxe et al., 2019). First, if only one paleopole is given alone without any geologic context, its polarity can be ambiguous, i.e. an upward inclination may be due to being located in the southern hemisphere during a normal polarity chron, or in the northern hemisphere during a reversed polarity chron (cf. the solid blue and solid green Africa in Fig. 1.2). In other words, we can't know if it's North pole or South pole, especially for paleomagnetic data with the Precambrian and early Paleozoic ages. Returning to the example above, if the  $\sim$ 155 Ma paleopole (52.59°N, 91.45°W) was formed during a period of reversed polarity, then it needs to be rotated to the South pole rather than the North pole. The necessary Euler rotation of pole (0°, 1.45°W) and angle 142.59° rotates the sampling site (0°, 25°E) on Africa to



**Figure 1.2:** Reconstruction of Africa with its  $\sim 155$  Ma paleopole. The red polygon is today's position of Africa, while the blue and green ones shows its reconstructed position at  $\sim 155$  Ma, if the pole was North and South pole, respectively. Dashed green polygon illustrates the ambiguity of paleolongitude from paleomagnetic data alone (sites at same latitude but different longitudes record the same Declination and Inclination in a GAD field).

( $15.6998^{\circ}\text{N}$ ,  $23.0121^{\circ}\text{W}$ ) indicating southward motion since the Late Jurassic. Second, because in a GAD field the declination equals zero everywhere (Fig. 1.2), paleomagnetic data doesn't register longitudinal motions of plates (the Euler pole for a plate moving purely to the east or west is at the geographic poles, so preserved paleopoles will experience zero rotation), which means we can position a plate at any longitude we wish subject to other geological constraints (cf. the solid and dashed green Africa in Fig. 1.2).

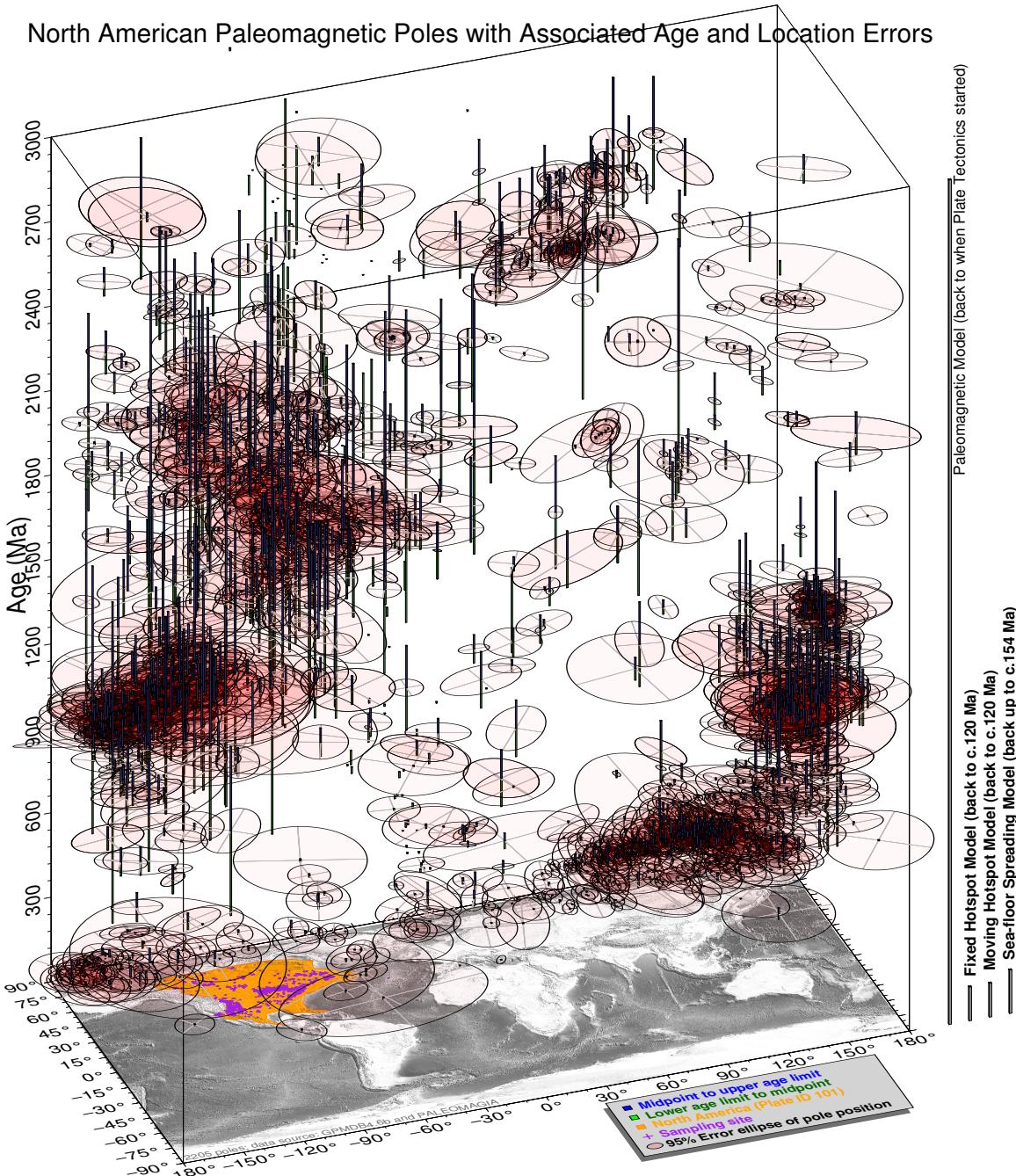
The data source used in this thesis is *Global Paleomagnetic Database* (GPMDB) Version 4.6b (McElhinny and Lock, 1996; Pisarevsky, 2005, updated in 2016 by the

Ivar Giaever Geomagnetic Laboratory team, in collaboration with Pisarevsky), which includes 9514 paleopoles for ages of 3,500 Ma to the present published from 1925 to 2016. GPMDB has been published in two ways: (1) IAGA GPMDB 4.6 online query: <http://www.ngu.no/geodynamics/gpmdb/>, which is now closed; (2) Microsoft Access system in .mdb format at NOAA's National Geophysical Data Center <https://www.ngdc.noaa.gov/geomag/paleo.shtml> (Pisarevsky and McElhinny, 2003) and CESRE's Paleomagnetism and Rock Magnetism project <https://wiki.csiro.au/display/cmfr/Palaeomagnetism+and+Rock+Magnetism>, which is later updated by Ivar Giaever Geomagnetic Laboratory <http://www.iggl.no/resources.html#data>.

An apparent polar wander path (APWP) is composed of poles of different ages from different sampling sites on the same stable (non-deforming) continent, chained together to form a record of motion relative to the fixed magnetic pole over geological time. It represents a convenient way of summarizing paleomagnetic data for a plate instead of producing paleogeographic maps at each geological period (Torsvik et al., 2008). As a preliminary study, the *North American Craton* (NAC) is chosen as a research object to develop techniques we want to think about. The NAC is one of best studied cratons in paleomagnetism with the GPMDB containing 2160 poles published since 1948 (Fig. 1.3). If we observe the latitudes, longitudes and age distribution of the NAC poles (Fig. 1.3), we actually can identify the general trend of its APWP. However, converting this data into a reliable, well-defined APWP can be challenging, due to the following issues:

### 1.1.3 Fact 1: Not All Regions on the Earth Surface Are Solid

If we consider the modern North America continent, the region west of the Rockies is actively deforming. Paleomagnetic data from such areas are likely to reflect local tectonic processes such as block rotation rather than rigid plate motions, and should be excluded. For example, the Rockies Mountain area was not included as my data selecting polygon (the transparent yellow area in Fig. 1.3). In order to investigate a specific craton or terrane or block's past paleogeographic motion, choosing an appropriate subregion without active tectonic activities, e.g. rotation or uplifting or rifting, to select data is often required. Such tectonics-free regions are usually called rigid. However, the difficulty of defining such tectonic boundaries makes appropriate spatial and temporal choices very difficult, particularly further in the geological past when cratonic configurations and active plate boundaries were very different to today. This leads to a question: What is the best way to constrain the data for a specific plate or block? The solution proposed in this thesis is described in Appendix B.



**Figure 1.3:** Much paleomagnetic data has been collected from the North American Craton. For younger geologic times, do we really need so much data to reconstruct accurately just like modern-day plate motions? The image shows distribution of all published paleopoles of the NAC over time, which are compiled from GPMDB 4.6b (Pisarevsky, 2005) and PALEOMAGIA (Veikkolainen et al., 2014).

#### 1.1.4 Fact 2: Not All Data Are Created Equal

APWPs are generated by combining paleopoles for a particular rigid block over the desired age range to produce a smoothed path. However, the NAC dataset illustrates

that uncertainties in the age and location of different paleopoles in the GPMDB can vary greatly.

#### 1.1.4.1 Age Uncertainty

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 NAC paleomagnetic data (Fig. 1.3) 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.

#### 1.1.4.2 Position Uncertainty

The uncertainties of pole latitudes and longitudes are plotted as 95% confidence ellipses (Fig. 1.3), 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.

#### 1.1.4.3 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. We want to look at the consistency of NAC and India’s data over smaller time periods, so the data is binned over a small time interval (e.g. 2 Myr) to see whether the paleopoles in each time interval overlap within their uncertainty ellipses, as they should. Sometimes, this is the case (Fig. 1.4a). Sometimes we have further separated poles with close ages (Fig. 1.4b).



**Figure 1.4:** Overlapping and further separated paleopoles of NAC and India. The oval ellipses are their 95% confidence uncertainties. The labels are their result number given in GPMDB 4.6b.

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

**Lithology** For this poor consistency of data (Fig. 1.4b), 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. For this case, all the three poles (Fig. 1.4b) are from sedimentary rocks. In addition, pole 1136 and 1137 (Result Number in GPMDB 4.6b)'s source rocks also contain redbeds (Opdyke et al., 1982), although the authors did not mention about the potential inclination shallowing. For pole 7095, although the source rocks do not contain redbeds, the authors did mention about possible inclination shallowing due to haematite grains (Gautam and Appel, 1994).

**Local Rotations** As discussed previously, local deformation between two paleomagnetic localities invalidates the rigid plate assumption and could lead to inconsistent paleopole directions. All the three poles (Fig. 1.4b) contain signals of local rotations (Opdyke et al., 1982; Gautam and Appel, 1994), e.g. pole 7095 has a signal which suggests the presence of a counter-clockwise local rotation of the Tinau Khola section (Gautam and Appel, 1994), and therefore do not reflect motions of the whole

rigid India plate in this case. So the discordance is likely due to local deformation (Fig. 1.4b), and we would ideally want to exclude or correct such poles from our APWP calculation.

**Other Factors** In Fig. 1.4, mean pole age (centre of age uncertainty) has just been binned. If any of the paleopoles have large age uncertainties, they could be different ages from each other and sample entirely different parts of the APWP. Conversely, if any of the paleopoles 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, because in order to average errors in orientation of the samples and scatter caused by secular variation, a “sufficient” number of individually oriented samples from “enough” sites must be satisfied (Tauxe et al., 2019; van der Voo, 1990; Besse and Courtillot, 2002). For example, pole 1136 (Fig. 1.4b) is from only 4 sampling sites, pole 1137 from only 3 sites and pole 7095’s site number not even given in the GPMDB 4.6b.

#### 1.1.4.4 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 uncertainties in both age and location (Fig. 1.3). The combination of lower data quality with lower data density means that a single ‘bad’ paleopole (with large uncertainties in age and/or location) can much more easily distort the reconstructed APWP, because there are few or no ‘good’ paleopoles to counteract its influence.

Data density also varies between different plates. E.g., we have a relatively high density of paleomagnetic data for NAC, but few paleopoles exist for Greenland and Arabia. Based on mean age (mean of lower and upper magnetic ages), for 120–0 Ma, GPMDB 4.6b has more than 130 poles for NAC, but only 17 for Greenland and 24 for Arabia.

#### 1.1.4.5 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? 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 (Besse and Courtillot, 2002). Briefly, the V90 criteria for paleomagnetic results includes seven criteria: (1) Well determined rock age and that magnetization age is the same is presumed; (2) At least 25 samples reported with Fisher precision  $\kappa$  (Fisher, 1953) greater than 10 and  $\alpha_{95}$  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). 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  $\alpha_{95}$  less than 10° in the Cenozoic and 15° in the Mesozoic. There are many potential ways to weight the data set which could obviously greatly influence the final result, and we want to test this. But there has been limited study of how effective these filtering/weighting 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.

## 1.2 Objectives

Our overarching aims are to develop rigorous, consistent and well-documented methods of reconstructing plate motions using paleomagnetic data, and to investigate the limits of paleomagnetic data on reconstructing individual plate motions, supercontinents, and global tectonic parameters like average rate of plate motion.

### 1.2.1 Motivation and General Approach

How has plate tectonics evolved over geologic history, in terms of average plate velocities, numbers of plates and so on? The only quantitative data we have prior to  $\sim 170$  Ma are paleomagnetic data. We know there are limitations, because we can’t constrain the longitudes of paleo-plates very well. When we look back through geologic history, how much good paleomagnetic data do we have, and how well does it reconstruct ‘true’ plate motions? We don’t know well the effects of data quality and density, which generally degrades further back in geologic history, on producing reliable APWPs. For the past  $\sim 130$ –200 Myr we have the highest density of

paleomagnetic data and also independent plate motion data from reconstructions of ocean spreading combined with hotspot reference frames. These independent data sources can help constrain plate motions in more accurate ways. This allows us to ask the question: How much paleomagnetic data do we need actually to reconstruct accurately known modern-day plate motions? If we can handle that, we can go back in time. For a certain density of paleomagnetic data that we have, how reliably can we talk about what's going on in the past given the much lower data distribution? It might turn out we don't need very much data to say something reasonably and reliably. We can test this by looking at the last 0–120 Ma where we can compare paleomagnetically derived plate motions with other methods of paleogeographic reconstruction. This does not only include the work of developing tools and algorithms to generate those paleomagnetically derived plate motions (to use paleomagnetic data to reconstruct APWP parameters that are known from other sources like ocean basins and hotspots), but also need us to know how good these tools are or which one is the best algorithm (to compare paleomagnetic APWPs with the known data sources predicted APWP). This can give insights into how well we can ‘know’ plate motions back in the past, and what data quality and density are necessary to reliably reconstruct a ‘true’ APWP.

As a preliminary analysis, some algorithms were made to separate/calculate out so-called good paleomagnetic data (at any particular time period for a particular craton, like here from 100 Ma to the present day for NAC). We are interested in what makes ‘good’ data, how we can identify it and filter it from the database, and how sometimes ‘bad’ data is only bad in the sense that it is poorly constrained in age or position or any other parameter, in which cases it might be possible to include it by e.g., weighting. A weighted mean pole can be calculated for a time interval with ‘better’ (more likely to be reliable) poles counting more than ‘worse’. For example, a pole with small  $\alpha_{95}$  and very well constrained age is more likely to reflect APWP position at the selected age point than a pole with large  $\alpha_{95}$  and very broad age range.

### 1.2.2 Research Questions or Hypotheses

Questions 1–4 focus on method development, whereas 5 and 6 start using them for plate tectonic research, especially in deep times.

### **1.2.2.1 Question 1**

What is the best way to turn a collection of individual poles, with different age constraints and uncertainties, into a smoothed APW path? This question, in fact, is about how to (1) choose a data-constraining polygon that represents a solid continent during a certain period; (2) pick (or bin) data within a certain window for Fisher statistical (Fisher, 1953) calculation; (3) do weighting for picked data according to different uncertainties or other kinds of standards of qualifications; (4) if the derived APWP is still not smoothed enough when compared with a reference path, is further smoothing necessary? Our goal here actually is to get a reliable result, i.e. a path generated to approximate the ‘real’ APWP with appropriate uncertainties.

### **1.2.2.2 Question 2**

Based on the consequences from the algorithms we developed, we can do research on why some algorithms are good, others bad for all plates? Why some algorithm performs well for a plate or two but not others?

### **1.2.2.3 Question 3**

How much paleomagnetic data do we need actually to accurately reconstruct known modern known plate motions? What insights does this give us into the reliability of reconstructions from earlier in geologic history?

### **1.2.2.4 Question 4**

Based on our analysis above, can we develop algorithms that look for matching segments of APWPs from different cratons, that might indicate they were part of the same continent or supercontinent?

### **1.2.2.5 Question 5**

What kind of dataset (in terms of data density and quality) is needed to accurately reconstruct a known APWP, or a shared APWP between two cratons? If we can establish some criteria for this, does it provide any insights into past reconstructions of plate motions (e.g., Rodinia)?

#### **1.2.2.6 Question 6**

Can we develop algorithms that use APWPs from multiple continents to estimate global average plate motion rates? Can we get a good sense of how much information is lost due to lack of data on longitudinal motions? Can we use this to draw any conclusions about long term trends (or lack thereof) in the style and vigour of global plate tectonics? (Possible further question: can data on relative continental motion acquired from matching APWP curves be incorporated to improve these estimates?)

In summary, this dissertation will not be able to help answer all the above questions. However, in the end the completion of this dissertation and solving the first three questions are hoped to be helpful solving the later questions in the future.

# Chapter 2

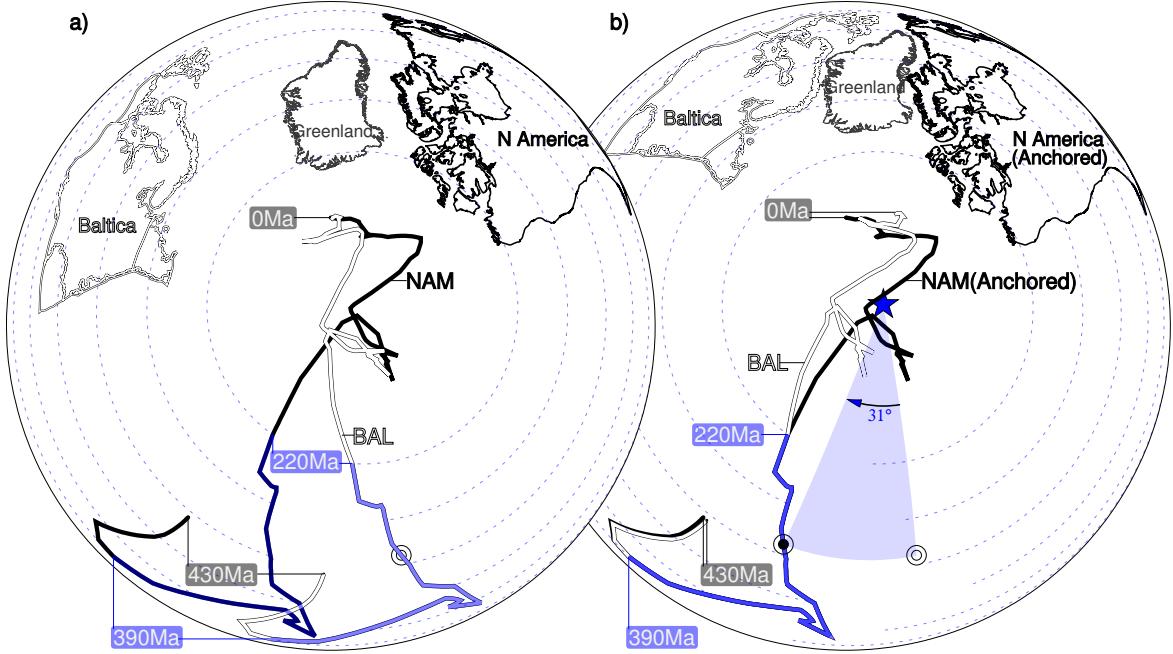
## Methodologies

*This chapter mainly describes the development of a new APW path similarity measuring tool used throughout the dissertation. Apparent polar wander paths (APWPs) based on paleomagnetic data are the principal means of describing plate motions through most of Earth history. Comparing the spatio-temporal patterns and trends of APWPs between different tectonic plates is important for testing proposed paleogeographic reconstructions of past supercontinents. However, thus far there is no clearly defined quantitative approach to determine the degree of similarity between APWPs. This paper proposes a new method of determining the degree of similarity between two APWPs that combines three separate difference metrics that assess both spatial separation of coeval poles, and similarities in the bearing and length of coeval segments using a weighted linear summation. Bootstrap tests are used to determine whether the differences between coeval poles and segments are significant for the given spatial uncertainties in pole positions. The Fit Quality is used to discriminate between low significance scores caused by comparing poorly constrained paths with large spatial uncertainties from those caused by a close fit between well-constrained paths. The individual and combined metrics are demonstrated using tests on synthetic pairs of APWPs with varying degrees of spatial and geometric difference. In a test on real paleomagnetic data, we show that these metrics can quantify the effects of correction for inclination shallowing in sedimentary rocks on Gondwana and Laurussia's 320–0 Ma APWPs. For an APWP pair, when one APWP's three individual metrics are all greater than or equal to, or less than or equal to the other one's, weighting is dispensable because the similarity ranking order becomes straightforward; otherwise assigning equal weights is recommended, although then decision makers are allowed to arbitrarily change weights according to their preferences.*

## 2.1 Introduction

Paleomagnetism is an important source of information on the past motions of the Earth's tectonic plates. The orientation of remanent magnetisations acquired by rocks during their formation record the past position of the Earth's magnetic poles. In older rocks, these virtual geomagnetic poles often appear to be increasingly offset from the modern day geographic poles. Because the Earth's geomagnetic field appears to have remained largely dipolar and centered on the spin axis for at least the last 2 billion years (Evans, 2006), this divergence is interpreted as recording the translation and rotation of a continent by the motion of tectonic plates in the time since the rock formed. An Apparent Polar Wander Path (APWP) is a time sequence of paleomagnetic poles (or, more commonly, mean poles that average all regional paleopoles of similar age) that traces the cumulative motion of a continental fragment relative to the Earth's spin axis.

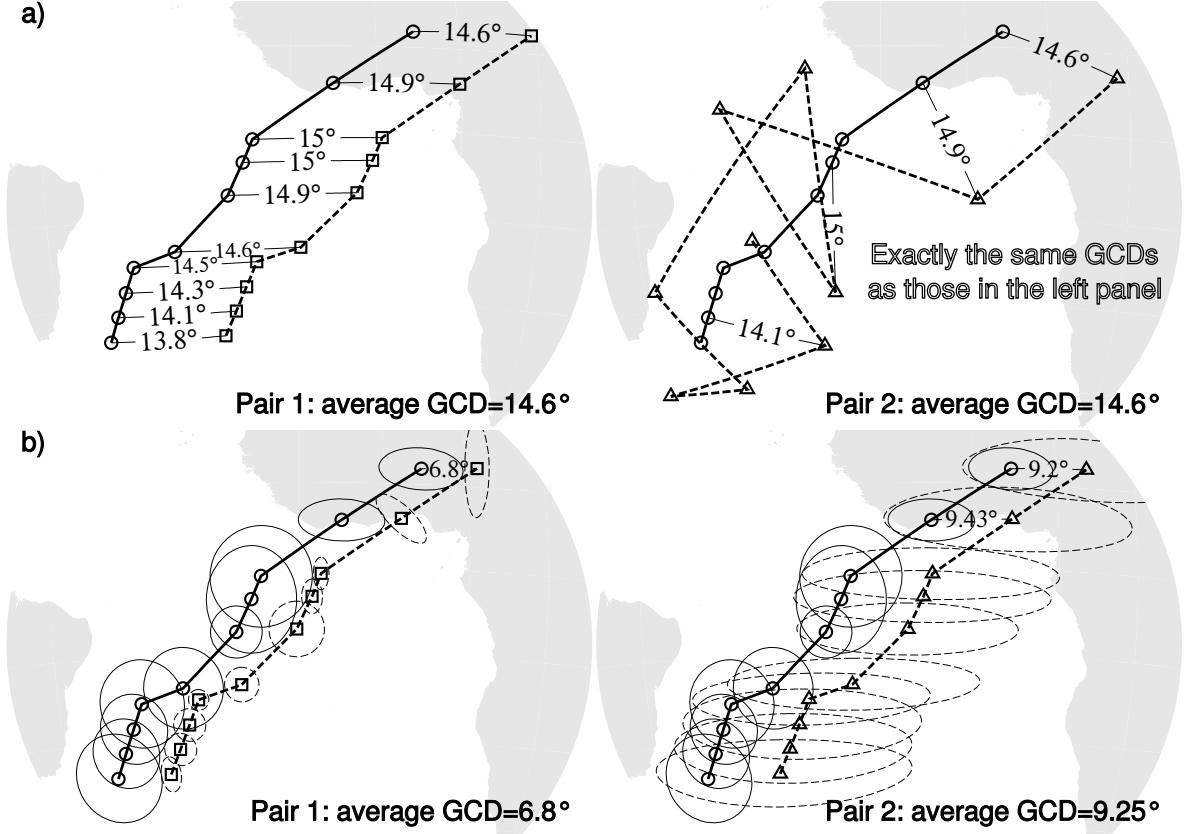
Investigations of the Earth's past tectonic evolution and paleogeography often involve comparing APWPs. For example, if two now separated continental fragments were once part of the same supercontinent, their APWPs should share the same geometry during the interval that this supercontinent existed. If the supercontinent has been correctly reconstructed, the APWPs should also overlap during this interval (Fig. 2.1). APWP comparisons can be used to assess plate motion models generated using different datasets and/or fitting techniques (Besse and Courtillot, 2002; Beaman et al., 2007; Sager, 2007; Torsvik et al., 2008; Domeier et al., 2011a, for example); significant deviations from the known APWP for a continent can also be used to identify local tectonic rotations (Geuna et al., 2010; Chandler et al., 2013, for example). Despite the clear importance of measuring APWP similarity, these comparisons remain largely qualitative in nature, involving visual comparisons of specific APWP segments and checking if they have overlapping 95% confidence limits (Besse and Courtillot, 2002; Beaman et al., 2007; Geuna et al., 2010; Domeier et al., 2011a, for example). Where quantitative measures are used, the mean great circle distance (GCD) between coeval poles on the APWP pair has been commonly used as a generalised difference metric for a pair of APWPs, with a lower score indicating that they are more similar (Sager, 2007; Torsvik et al., 2008, for example). However, because GCD is simply a measure of spatial separation and does not incorporate geometric information about the two paths being compared, it is possible for pairs with clearly different similarities to have similar mean GCD scores (Fig. 2.2a). Due to the inherent time variability of the geomagnetic field, uncertainties arising from the sampling



**Figure 2.1:** (a) The APWPs for North America (black) and Baltica (grey) are spatially distinct, but their Late Paleozoic–Early Mesozoic sections are geometrically similar due to them both being part of the supercontinent Pangaea. (b) Reversing the opening of the Atlantic Ocean by rotation around a reconstruction pole (blue star) results in the overlap of these two APWPs between 390 million years ago (Ma) and 220 Ma, validating the proposed paleogeography. The effects of this rotation on Baltica and its APWP (BAL) are illustrated by the motion of the circle marker (before: blank center; after: dark center), respectively. General Perspective projection. APWPs and rotation parameters from Torsvik and Cocks (2016).

and measurement of remanent magnetisations, and uncertainties in the magnetization age, the mean paleopoles that make up an APWP also have associated spatial uncertainties. The significance of a GCD score is therefore not immediately obvious. A score that indicates a relatively large difference between two paths may not be significant if the spatial uncertainties are large; a small difference could be significant if the spatial uncertainties are small (Fig. 2.2b).

We have developed an improved quantitative method of calculating the similarity between two APWPs, or coeval segments of APWPs, in the form of a composite difference score that compares both their spatial overlap and geometry. Our method incorporates statistical significance testing, allowing paths with associated spatial uncertainties to be rigorously compared to each other. The validity and effectiveness of this method, and its superior discrimination compared to a mean GCD score, are demonstrated by comparing the published APWP of the North America Plate to seven derivative paths with different degrees of spatial and geometric noise applied



**Figure 2.2:** (a) How the average GCD similarity metric ignores path geometry: *Pair 1* (circles and squares, left) is clearly more similar than *Pair 2* (circles and triangles, right), but for both pairs each GCD remains constant. (b) How GCD also ignores spatial uncertainties. The average GCD separation between coeval points is smaller for *Pair 1* (circles and squares, left) than *Pair 2* (circles and triangles, right). But if spatial uncertainties (plotted as 95% confidence ellipses) are considered, this ranking is not trustworthy: it is *Pair 2* that is statistically indistinguishable from the reference path. Azimuthal Orthographic projection.

(Fig. 2.3).

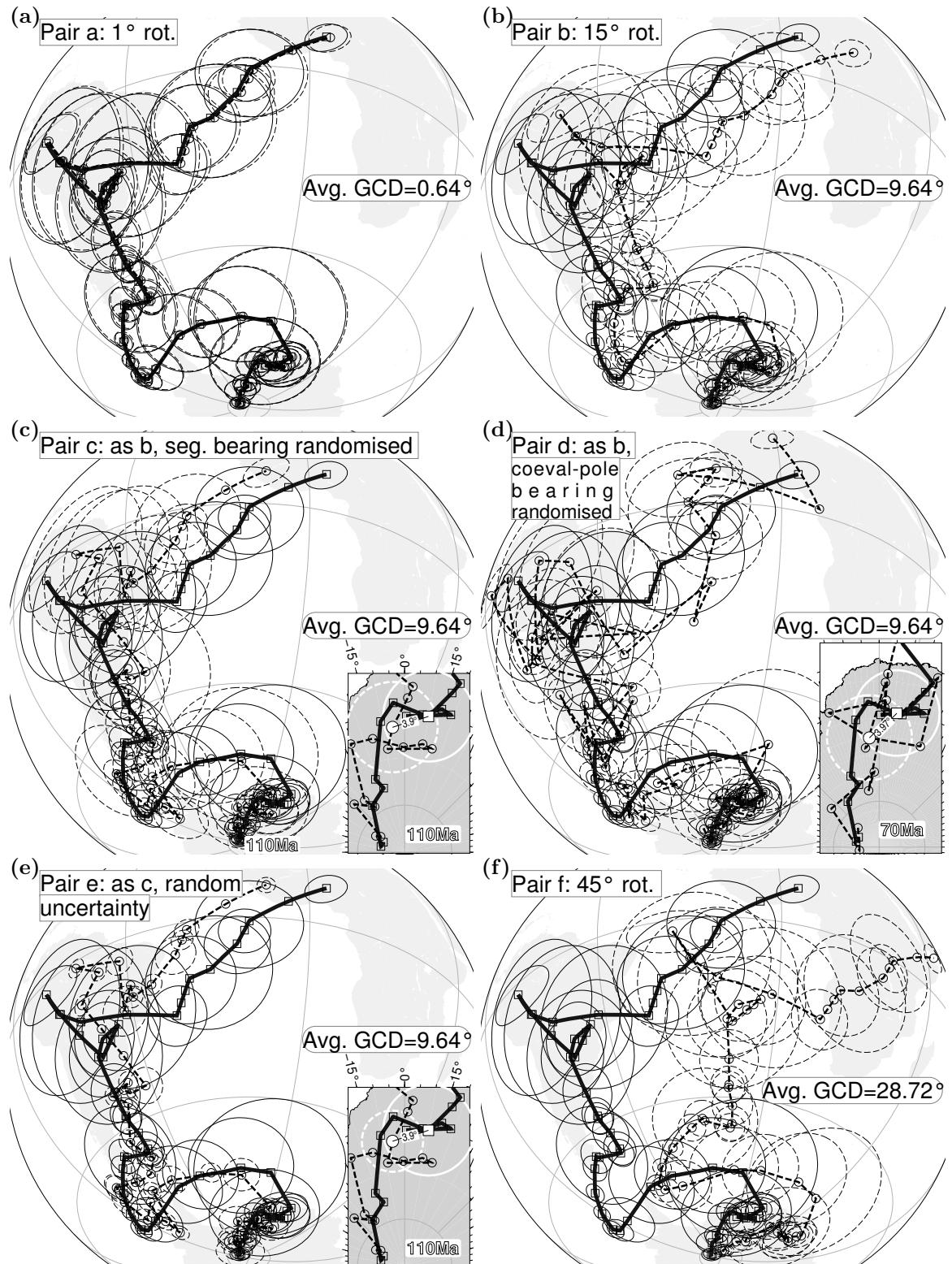
We also test our algorithm on real paleomagnetic data, demonstrating that this tool can be used to quantitatively assess the effects of different corrections (in this case, bulk corrections for inclination shallowing in sediments) on the similarity between APWPs from different continents.

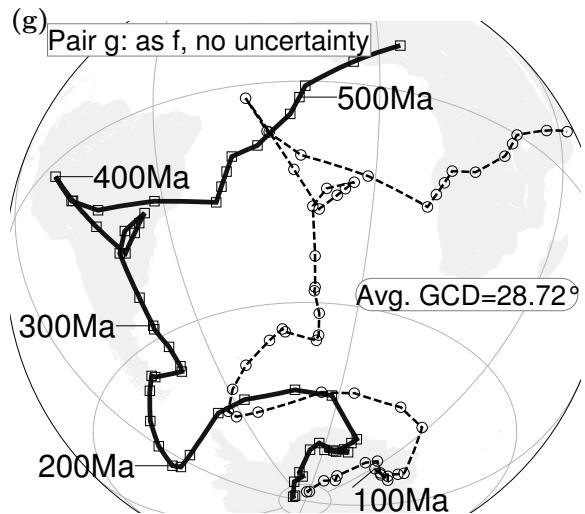
## 2.2 Methods

### 2.2.1 Comparing Apparent Polar Wander Paths (APWPs)

An APWP consists of a sequence of ( $n$ ) mean poles,  $P_1, P_2, \dots, P_n$ , which average the published paleopoles from a particular continent for a particular time interval. Each mean pole has associated longitude ( $\phi$ ), latitude ( $\lambda$ ), and age ( $t$ ). Spatial

uncertainty is represented by a 95% confidence ellipse described by semi-major axis **dm** with azimuth  $\beta$  (angle east of north) and perpendicular semi-minor axis **dp** (e.g. Fig. 2.3).





**Figure 2.3:** APWP pairs used to validate new path comparison method. In each case the Phanerozoic APWP for Laurentia/North America (squares, bold line) at 10 million-year (Myr) timesteps (Torsvik et al., 2012, “RM” column of its Table 3), is compared to a transformed copy (circles, dashed line): (a) 1° finite rotation applied to all the mean poles and their 95% uncertainty ellipses around an Euler pole at (125° E, 88.5° S); (b) as (a), but 15° rotation around same Euler pole; (c) after rotation as in (b), the orientation of each APWP segment is randomised whilst keeping their GCD length and the coeval poles’ GCD fixed; (d) after rotation as in (b), the bearing between coeval poles is randomised whilst keeping their GCD spacing fixed; (e) as (c), but with randomly varied and relatively smaller associated spatial uncertainty; (f) as (a), but 45° rotation around same Euler pole; (g) as (f), but with zero associated spatial uncertainty. Azimuthal Orthographic projection.

For two continents that were once part of a supercontinent, their APWPs for the period should perfectly overlap when rotated into a common reference frame (Fig. 2.1). However, due to (i) the spatial uncertainty associated with the mean poles that form an APWP, (ii) differences in the density and quality of data available to calculate paleopoles for different continents in coeval time periods, and (iii) uncertainties and possible errors in the rotations used to represent past plate motions, a perfect match will not be obtained. Instead, two ‘matching’ paths should share a generally similar geometry, and largely overlap with each other when rotated into a common reference frame. A quantitative measure of the spatial and geometric similarity between these two paths should ideally allow us to distinguish between non-identical paths that are similar within the associated uncertainties, and non-identical paths that are actually different, due to differential motion between the two continents or a poorly constrained reconstruction.

### 2.2.2 APWP Pairs Used in This Study

To assess the performance of the evaluation method developed here, we apply it to seven different path scenarios (Fig. 2.3) generated from transformations of the 530–0 Ma Phanerozoic APWP for Laurentia (Torsvik et al., 2012). Almost exactly identical paths generated by rotating one by a degree around an Euler pole at (-55°, 88.5°) (*Pair a*, Fig. 2.3a) represent an ideal case of matching paths in the same spatial reference frame. Matching paths that have been rotated out of the same reference frame by small (15°; *Pair b*, Fig. 2.3b) and large (45°, *Pair f*, Fig. 2.3f) amounts around the same rotation pole represent small and large reconstruction errors, respectively. Random noise added to the path (*Pair c*, Fig. 2.3c, *Pair d*, Fig. 2.3d) or the associated uncertainties (*Pair e*, Fig. 2.3e) represent differences in data source and/or quality. In the final pair (*Pair g*, Fig. 2.3g), spatial noise has been reduced by removing *Pair f*'s pole uncertainties.

These seven cases allow evaluation of the performance of any path comparison metric across a range of different spatial and geometric similarities. To be successful, such a metric must distinguish pairs with high spatial and geometric similarity (*Pair a*) from pairs with lower spatial (*Pair b*, *Pair f*, *Pair g*) or geometric (*Pair c*, *Pair d*) or both (*Pair e*) similarities.

To achieve more robust discrimination than the mean GCD, we propose combining a metric for spatial misfit (Mean Significant Spatial Difference) with metrics for geometric difference (Mean Significant Length and Angular Differences) using a weighted linear summation, as described in the following sections.

### 2.2.3 Significant Spatial Difference

As in previous quantitative comparisons (Sager, 2007; Torsvik et al., 2008, for example), the spatial separation of two APWPs is defined by the average GCD distance between their coeval poles, but we add a filter for spatial uncertainty based on the bootstrap approach (Tauxe et al., 1991). 1000 bootstrapped mean directions for each pole in a coeval pair were generated (the exact sampling method is dependent on the available information for the pole—see Supplementary Information for a full description) and their cumulative distributions in Cartesian coordinates were compared (Tauxe et al., 1991). Pairs that could not be distinguished at the 95% confidence interval had their GCD separation set to 0 prior to calculation of the mean GCD distance for all pairs. This distance is then normalised by dividing by 50°, which is referred to the empirical fact that a 95% confidence ellipse major semi-axis

of about  $25^\circ$  is considered unacceptably large by paleomagnetists (Butler, 1992), to obtain the significant spatial difference  $d_s$ . A  $d_s$  of zero indicates that the two paths are statistically indistinguishable from each other.

### 2.2.4 Shape Difference

The shape of an APWP is determined by the orientations and lengths of its geodesic segments, which are related to the location of the Euler stage pole that describes plate motions, and the rotation rate about that pole, respectively. The geometric similarity of two APWPs can therefore be assessed by comparing (i) the bearings, and (ii) the lengths of their coeval segments (Fig. 2.4), with the assumption that similar geometries are generated by a common set of stage rotations.

#### 2.2.4.1 Mean Length Difference

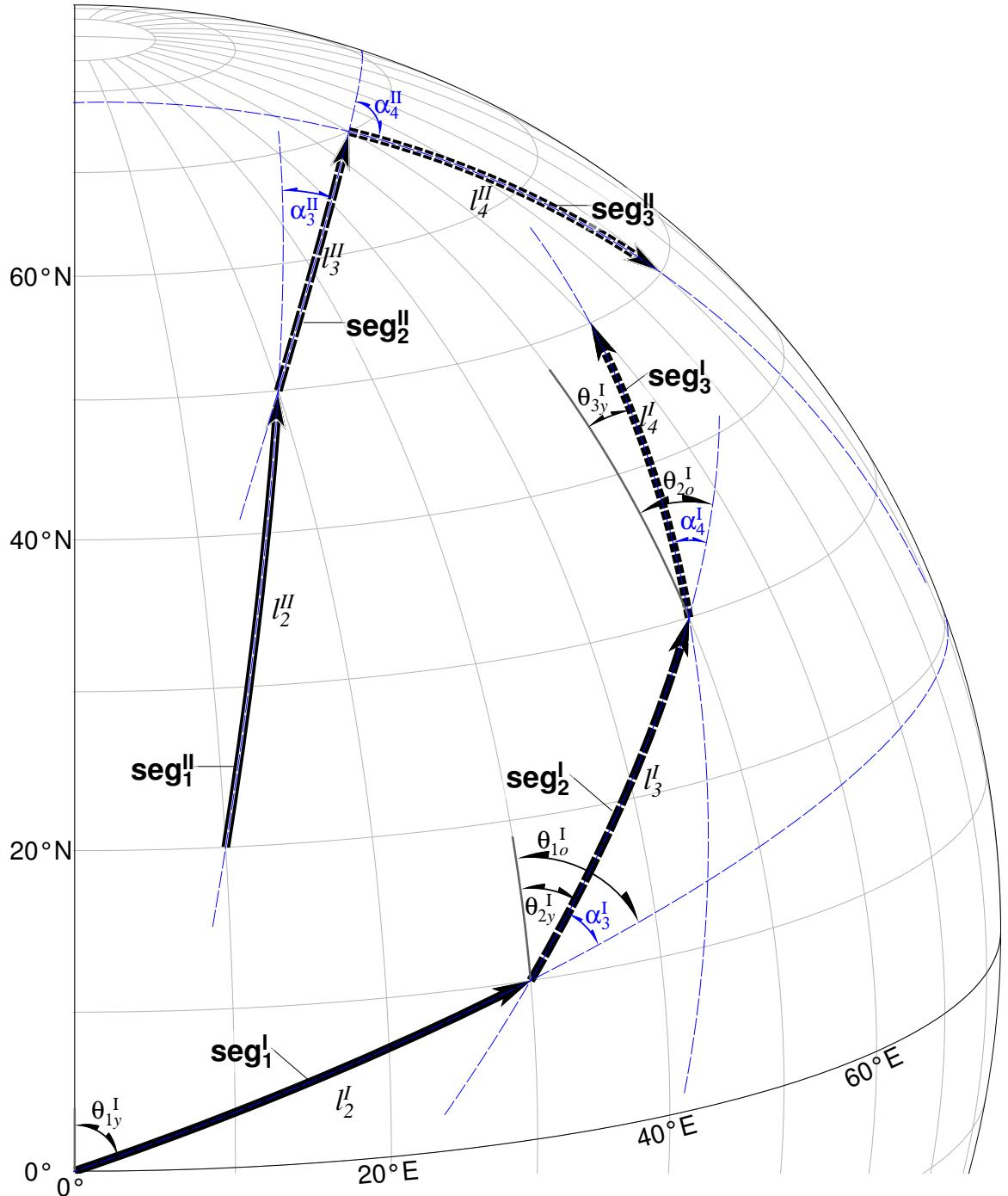
The mean length difference between the two APWPs  $traj^I$  and  $traj^{II}$  is the absolute sum of differences between the lengths of coeval path segments (e.g.  $l_2^I$  versus  $l_2^{II}$ ,  $l_3^I$  versus  $l_3^{II}$ ,  $l_4^I$  versus  $l_4^{II}$ , Fig. 2.4), normalised by dividing by the possible maximum distance the pole could wander during the whole period, such that:

$$d_l = \frac{\sum_{k=2}^n |l_k^I - l_k^{II}|}{D_{polar} * (t_n - t_1)}, \quad \forall k \in \{2, 3, \dots, n\},$$

where  $|l_k^I - l_k^{II}|$  is the length difference of one pair of coeval segments for an APWP pair ( $traj^I$  and  $traj^{II}$ ), e.g.  $|l_2^I - l_2^{II}|$  for the beginning coeval segment pair. The normalising parameter  $D_{polar}$  is  $2.7^\circ/\text{Myr}$ , derived from estimates of magnitude of maximum plate velocity (Swanson-Hysell et al., 2009; Kulakov et al., 2014, up to about 30 cm/year). A  $d_l$  approaching 1 would result from a comparison between a virtually stationary APWP and one associated with a rapidly moving plate.

#### 2.2.4.2 Mean Angular Difference

The mean angular difference describes the degree of consistency between the polar-wandering directions (defined as the bearing of the older pole in a segment with respect to the younger one) of two APWPs. In order to robustly compare two APWPs that have not necessarily been rotated into the same reference frame, it is more useful to define the APWP geometry relative to the path itself, rather than an external reference frame. Therefore the bearing of a segment is expressed as the change in geographic bearing with respect to the previous segment ( $\alpha_3$  and  $\alpha_4$ , Fig. 2.4). For



**Figure 2.4:** Geometric differences between coeval sections of two different APWPs ( $\text{seg}_1^I$ - $\text{seg}_2^I$ - $\text{seg}_3^I$  &  $\text{seg}_1^{II}$ - $\text{seg}_2^{II}$ - $\text{seg}_3^{II}$ ) can be described by comparing segment lengths (e.g.  $l_2^I$  versus  $l_2^{II}$ ) or changes in bearing of coeval segments relative to their previous segment (e.g.  $\alpha_3^I$  versus  $\alpha_3^{II}$ ). Segments are along great circles (blue dashed lines). Azimuthal Orthographic projection.

example,  $\alpha_3^I$  is the result of subtracting the geographic azimuth  $\theta_{2y}^I$  from  $\theta_{1o}^I$ , where “y” stands for young end of segment and “o” for old end of segment. The first segment

cannot record a relative bearing change: a path with  $n$  poles therefore consists of  $n-1$  segments which are described by  $n-2$  relative angles. The defined range of bearing values is set as  $-180^\circ$  to  $180^\circ$ , with clockwise (east) changes in direction defined as positive, e.g.  $\alpha_3^{II}$  and  $\alpha_4^{II}$ , and anticlockwise (west) changes defined as negative, e.g.  $\alpha_3^I$  and  $\alpha_4^I$ .

The mean angular difference  $d_a$  between two paths  $traj^I$  and  $traj^{II}$  can then be defined as

$$d_a = \frac{\sum_{k=3}^n \Delta\alpha_k}{180 * (n - 2)},$$

where

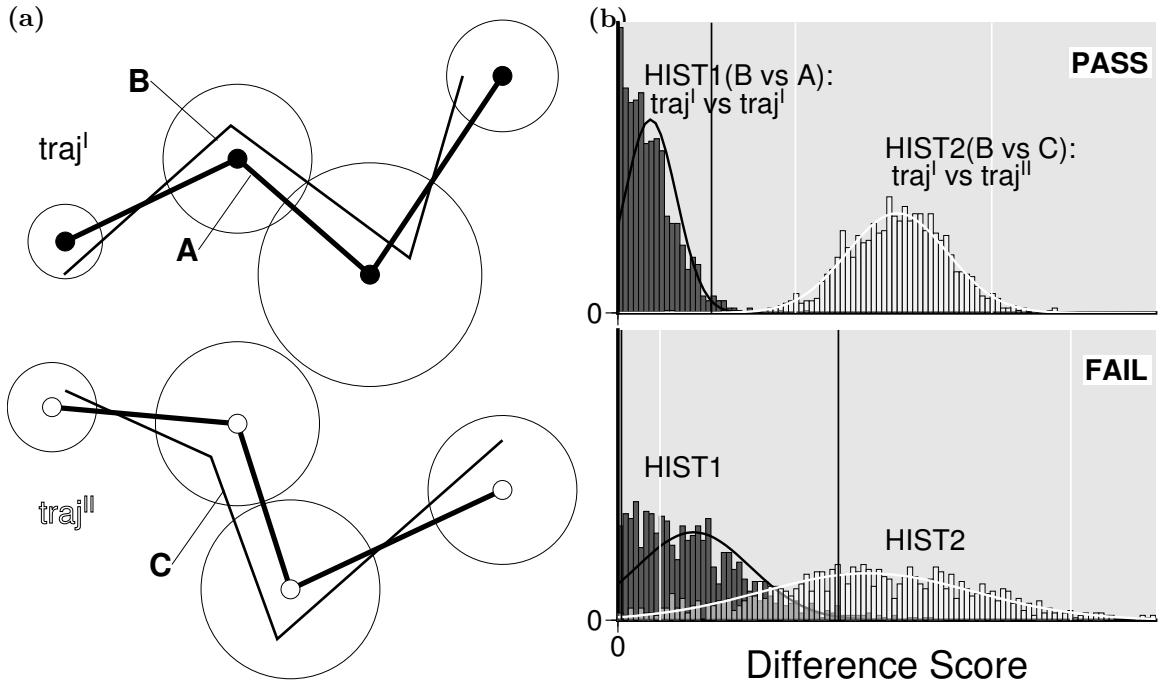
$$\Delta\alpha_k = \begin{cases} |\alpha_k^I - \alpha_k^{II}|, & \text{if } |\alpha_k^I - \alpha_k^{II}| \leq 180; \\ 360 - |\alpha_k^I - \alpha_k^{II}|, & \text{otherwise.} \end{cases} \quad \forall k \in \{3, 4, \dots, n\}.$$

$d_a$  is normalised by the maximum possible angular deviation of  $180^\circ$ . A score of 0 indicates exactly matching changes in the bearing of coeval segments along the length of the two paths, and a score of 1 indicates all segment bearings are antiparallel.

#### 2.2.4.3 Significance Testing of Shape Difference

Due to associated spatial uncertainty, the mean poles in an APWP trace out one possible path within a range of possible geometries (Fig. 2.5a). If the length and angular difference scores for one path fall within the range of possible scores for the other, two APWPs may not in fact be significantly different from each other. Significance testing for the shape difference scores is performed on each coeval segment pair as follows (Fig. 2.5b):

- A bootstrapped distribution of possible geometries for each segment in a path can be created by resampling the two mean poles that define the original segment, in the same manner as described in Section 2.2.3 and the Supplementary Information.
- A histogram of statistically indistinguishable length and/or angular difference scores (HIST1, Fig. 2.5b) is created by comparing the resampled paths with the original for each  $traj^I$  segment.
- This distribution is then compared to the histogram of difference scores created by resampling the coeval segments of  $traj^I$  and  $traj^{II}$  (HIST2, Fig. 2.5b).



**Figure 2.5:** Significance testing for the geometric metrics,  $d_l$  and  $d_a$ . (a) Illustration of how paths  $\text{traj}^I$  and  $\text{traj}^{II}$  can be re-sampled within their uncertainty ellipses, with B being a possible trajectory of  $\text{traj}^I$  and C being a possible trajectory of  $\text{traj}^{II}$ . (b) Upper: If the 95% confidence interval (black vertical lines are its upper and lower bounds) for the distribution of difference scores HIST1, generated by comparing multiple resamplings of  $\text{traj}^I$  with the original trajectory (A versus B) does not overlap with the 95% confidence interval (bounded by white vertical lines) for the distribution of scores HIST2, generated by comparing resamplings of  $\text{traj}^I$  and  $\text{traj}^{II}$  (B versus C), then the original difference score for  $\text{traj}^I$  and  $\text{traj}^{II}$  is statistically distinguishable; Lower: If the confidence intervals overlap, then the two paths are not distinguishable.

- If the two bootstrapped distributions HIST2 and HIST1 do not overlap at the given significance level (e.g. the upper and lower bounds of a 95% confidence intervals, Fig. 2.5b), then the difference score is interpreted to be significant. If not, then the bearings or lengths of the coeval segments are indistinguishable.

These tests allow a filter for spatial uncertainty to be added to the  $d_a$  and  $d_l$  metrics: prior to summation and normalisation, the difference score is set to zero for the coeval segments of  $\text{traj}^I$  and  $\text{traj}^{II}$  that are statistically indistinguishable.

### 2.2.5 Composite Path Difference ( $\mathcal{CPD}$ )

The three difference measures described above can be combined into a composite path difference ( $\mathcal{CPD}$ ) by means of a simple linear weighting rule,

$$\mathcal{CPD} = W_s \cdot d_s + W_a \cdot d_a + W_l \cdot d_l$$

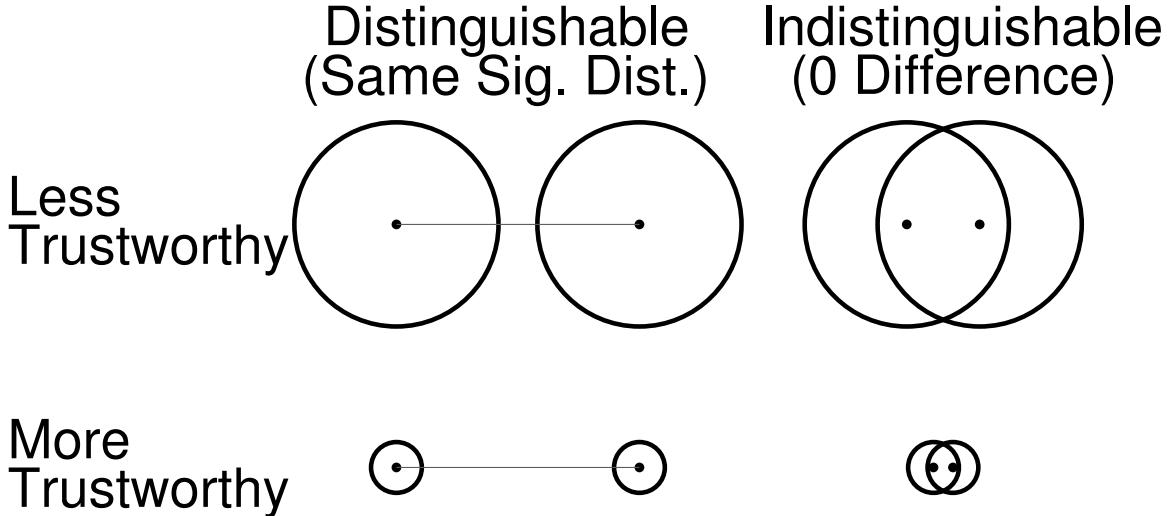
for  $0 < W_s, W_a, W_l < 1$ , where  $W_s$ ,  $W_a$  and  $W_l$  are weighting coefficients that sum to 1. Different weighting values allow the relative influences of spatial and geometric (length and angular) similarity to be varied (Section 2.3.2).

### 2.2.6 Fit Quality (FQ)

The three metrics are all tested to be significant based on the spatial uncertainties. However, the larger the uncertainties are, the less trustworthy the significant difference scores are (Fig. 2.6). Accordingly, we bring in a concept of “Fit Quality”, along the classification scheme of the reversal test (McFadden and McElhinny, 1990). For each mean pole of an APWP, it assigns a score based on the size of the spatial uncertainty (radius: A95, or  $(DM+DP)/2$ ): 1 if it is  $\leq 5^\circ$ , 2 if  $5^\circ < r \leq 10^\circ$ , 3 if  $10^\circ < r \leq 20^\circ$ , and 4 if it is  $> 20^\circ$ . These values are averaged for each APWP to give a “Fit Quality” score (from 1 to 4) for the difference score. This is then converted into an A/B/C/D letter grade, A if the average is  $< 1.5$ , B if  $1.5 \leq \text{avg} < 2.5$ , C if  $2.5 \leq \text{avg} < 3.5$ , and D if it is  $\geq 3.5$ , to indicate how easy it is to generate a low difference score. In other words, an A grade indicates that most poles are well-constrained and so it is fairly hard to have an indistinguishable path and a low difference score. A D grade indicates that most poles have large uncertainties so it is much easier to have a low difference score.

In addition, a short APWP segment tends to result in overlap of its two end poles’ spatial uncertainties. For example, if an APWP is generated at intervals of 10 Myr, the longest realistic segment-length would be about  $27^\circ$  (Swanson-Hysell et al., 2009; Kulakov et al., 2014, the maximum rate of plate movement is about 30 cm/yr). So the uncertainty size needs be less than  $\sim 13.5^\circ$  on average to make the segment length trustworthy. Therefore, to a certain extent, the “Fit Quality” also reflects the quality of the length metric if we give each path a grade for an APWP pair, e.g. A-A. The angular metric’s quality is related to both coeval mean poles and successive mean poles, so it has already been involved in the spatial and length quality. Moreover, given the fact that usually the mean significant length and angular differences are much lower than the mean significant spatial difference (e.g., Figs. 2.8b, 2.8d, 2.8f and Figs. 2.11b, 2.11d, 2.11f), the “Fit Quality” is capable to manifest the overall quality of all the three metrics and we should trust the difference scores if we get a B-B grade at least.

For example, *Pairs a, b, c, d* and *f*’s fit quality score is all 1.809–1.809, so their fit quality is B–B, which means that the mean poles in these APWP pairs have intermediate uncertainties on average so it is relatively hard to have a low difference score. *Pair e*’s fit quality is B–A (1.809–1.085). *Pair g*’s fit quality is A–A (0–0).



**Figure 2.6:** Difference scores from APWPs with large uncertainties are less trustworthy.

## 2.3 Results and Discussion

### 2.3.1 Discrimination of Difference Metrics

The performance of the individual metrics were tested by generating and ranking scores for each pair in Fig. 2.3. Scores for comparisons of the full 530 Myr paths, and sequential 100–130 Myr subsections were calculated for path pairs with (Fig. 2.7) and without (Fig. 2.8) poles with zero spatial uncertainties at 350 Ma, 360 Ma, 380 Ma, 390 Ma, 450 Ma, 460 Ma and 520 Ma calculated using linear interpolation by Torsvik et al. (2012).

#### 2.3.1.1 $d_s$

If ordered only in terms of spatial similarity, the desired order for the seven APWP pairs (Fig. 2.3), from most similar (lowest  $d_s$ ) to least similar (highest  $d_s$ ) is

$$\text{Pair a} < \text{Pair b} \approx \text{Pair c} \approx \text{Pair d} < \text{Pair e} < \text{Pair f} < \text{Pair g}. \quad (2.1)$$

This ordering is based largely on the mean GCD separations of each pair (Fig. 2.3), but also takes uncertainties into account: the relatively smaller uncertainty ellipses of *Pair e* and *Pair g* should lead to a higher  $d_s$  than *Pairs b-d* and *Pair f*, respectively. Without significance testing,  $d_s$  is directly proportional to mean GCD (Fig. 2.7a), which does not result in unique  $d_s$  for *Pair e* and *Pair g*. Significance testing reproduces the desired order (Fig. 2.7b).

Most  $d_s$  scores are also reduced with significance testing, with the largest reductions occurring where the path separations are low and 95% confidence ellipses for

coeval poles are more likely to overlap (e.g.,  $d_s$  for *Pair a* is reduced by 85%,  $d_s$  for *Pair f* is reduced by <2%).  $d_s$  that approach 0 for the 0–100 Ma and 100–200 Ma sub-paths, which are located close to the Euler pole used to separate the pairs and therefore remain in close proximity even after large rotations, also illustrate this effect. With significance testing, the 0–100 and 100–200 Ma sub-paths of *Pair f* and *Pair g* can still be distinguished (Fig. 2.8b and also Fig. 2.7b), due to no spatial uncertainty assigned to *Pair g*. In contrast, the older 300–400 Ma and 400–530 Ma sub-paths have a larger  $d_s$  than the whole path. This is because the 350 Ma, 360 Ma, 380 Ma, 390 Ma, 450 Ma, 460 Ma and 520 Ma pole coordinates are interpolated (Torsvik et al., 2012), and thus have no assigned spatial uncertainty on any of the test paths. Without the interpolated poles,  $d_s$  is always zero for *Pair a* and any of its sub-paths (Fig. 2.8b), which is expected.

Even with significance testing,  $d_s$  for *Pairs b-d* are the same (Fig. 2.7a and Fig. 2.7b) despite their different geometries (Fig. 2.3b, Fig. 2.3c and Fig. 2.3d), because GCDs between coeval poles and their uncertainties are the same. This result emphasises that a spatial difference metric alone cannot discriminate these pairs from each other. The comparison of *Pairs c* and *e* indicates that well-constrained mean poles with lower uncertainties make it relatively harder to have an indistinguishable APWP and a low difference score.

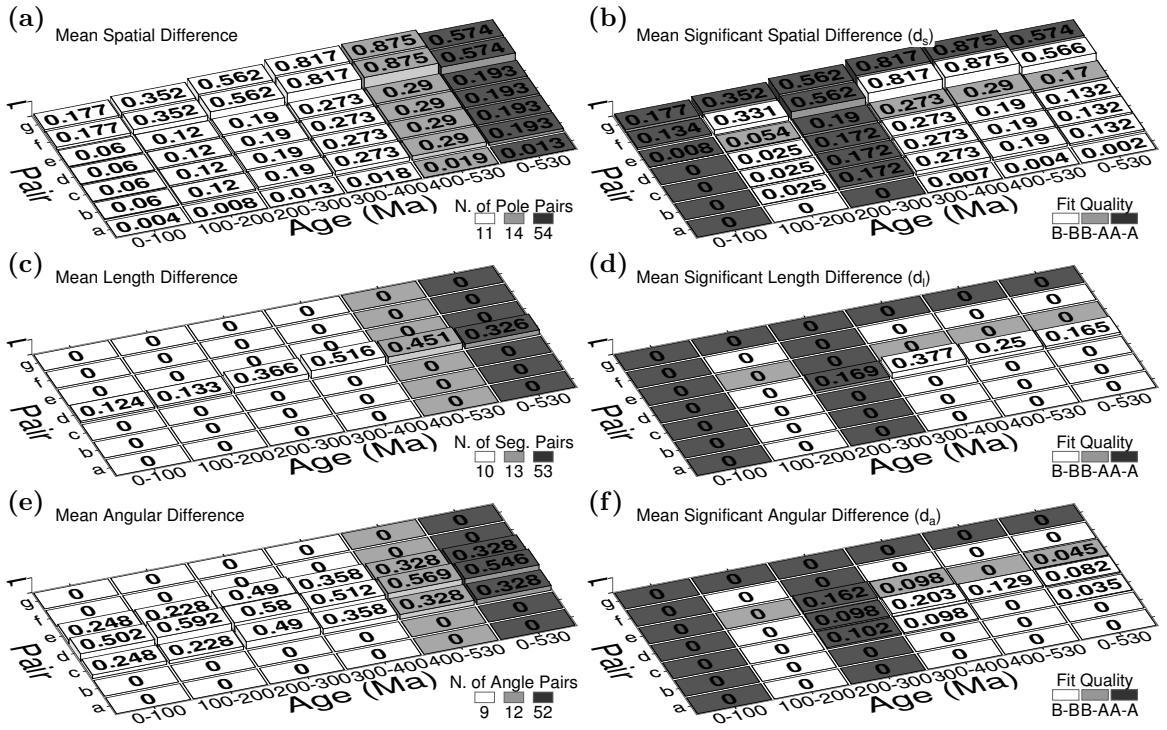
In summary, as Fig. 2.8b and also Fig. 2.7b illustrate,  $d_s$ , scaling with mean significant GCD, reproduces the expected order of spatial similarity (Order (2.1)) for the full path. It also compensates for the deficiency of the algorithm without statistical test (Fig. 2.8a and Fig. 2.7a) in differentiating *Pair f* and *Pair g*. Although our algorithm also works for APWPs with interpolations (e.g., Fig. 2.7; see how we do significance testing on the interpolated mean poles in the Supplementary Information), a meaningful and valid analysis should be based on the results with uninterpolated paleomagnetic APWPs (e.g., Fig. 2.8).

### 2.3.1.2 $d_l$

When ordered only according to the length similarity  $d_l$  the expected order is

$$0 = \text{Pair a} = \text{Pair b} = \text{Pair c} = \text{Pair e} = \text{Pair f} = \text{Pair g} < \text{Pair d} \quad (2.2)$$

Because only the path generated for *Pair d* allowed the length of coeval segments to vary, it is expected that other five pairs of APWPs have zero  $d_l$  for both the full-path and the five specified sub-paths even prior to significance testing (Fig. 2.8c and Fig. 2.8d), and this expected order is trivially reproduced. The effect of significance



**Figure 2.7:** Mean spatial, length and angular differences between two paths of the seven APWP pairs shown in Fig. 2.3. Left column: results without significance testing imposed in the metric; Right column: results with significance testing. Note that the spatial difference results of *Pairs b, c* and **d** are always the same for both the untested case (a) and the tested case (b). In addition, for those segments that do not begin from 0 Ma, their beginning segments are different from the 0–100 Ma sub-path’s and the full path’s. For example, for the 200–300 Ma sub-path, its beginning segment is the 200–210 Ma one.

testing (Fig. 2.8d) is to substantially reduce  $d_l$ . Many segment length differences do not pass the significance test because the angular uncertainties of the poles that define individual segments are large compared to the length of those segments.

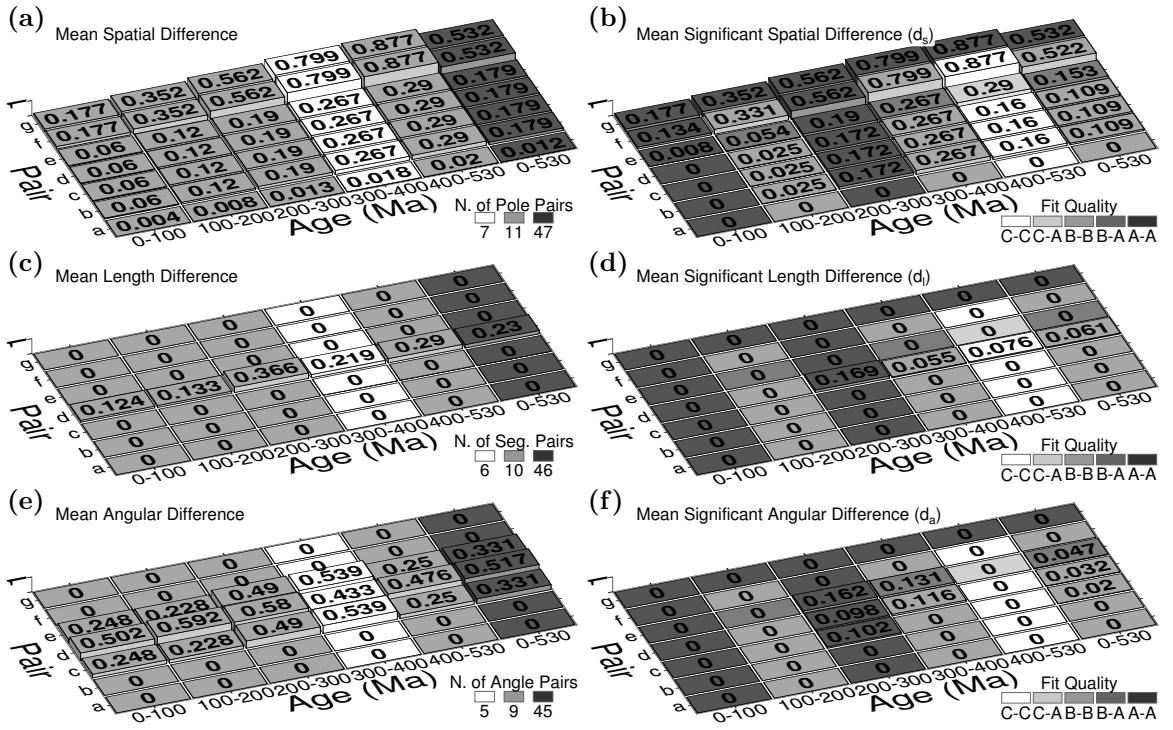
### 2.3.1.3 $d_a$

If ordered only according to angular similarity  $d_a$ , the expected order is

$$0 = \text{Pair a} = \text{Pair b} = \text{Pair f} = \text{Pair g} \leq \text{Pair c} \leq \text{Pair d} < \text{Pair e} \quad (2.3)$$

Because path geometries are not altered by a simple Euler rotation, only *Pairs c, d* and **e** are expected to have a non-zero  $d_a$ .  $d_a$  for *Pair d* and *Pair e* should be larger than *Pair c*’s due to more geometric variation and lower spatial uncertainties, respectively, although the expected ordering of *Pair d* and *Pair e* is less immediately obvious from visual inspection.

Without significance testing, non-zero  $d_a$  for *Pairs c, d* and **e** are consistently generated for both the full path and sub-paths (Fig. 2.8e).  $d_a$  for *Pair d* is usually



**Figure 2.8:** Mean spatial, length and angular differences between two paths of the seven APWP pairs with no interpolated poles shown in Fig. 2.3. Left column: results without significance testing imposed in the metric; Right column: results with significance testing. See explanation of Fig. 2.7.

higher (Fig. 2.8e), but there is no discrimination between *Pairs c* and *e*, which have the same score because geometrically they are identical. When significance testing is applied  $d_a$  is markedly reduced (Fig. 2.8f versus Fig. 2.8e), and is actually reduced to 0 for the two youngest and oldest sub-paths in all cases. This is somewhat expected because the segment lengths of the APWPs being tested are often of the same order as the angular uncertainty in their spatial position. As a result, a large range of different path geometries are possible within the specified uncertainty bounds, and the bearing of coeval segments has to be very large for the difference to be significant.

For the full paths and the 200–300 and 300–400 Ma sub-paths where  $d_a$  after significance testing is non-zero for *Pairs d*, *e* and (usually) *c*, *Pair e* can now be discriminated from *Pair c*, and consistently has the highest  $d_a$  of the 3 pairs.

In summary, our angular difference algorithm with statistical test (Fig. 2.8f) reproduces the expected order of angular similarity (Order (2.3)).

### 2.3.1.4 $\mathcal{CPD}$

When the seven different APWP pairs (Fig. 2.3) are rank-ordered in terms of the three criteria combined, their expected order is

$$\textit{Pair a} < \textit{Pair b} \leq \textit{Pair c} < \textit{Pair d} ? < \textit{Pair e} ? < \textit{Pair f} < \textit{Pair g} \quad (2.4)$$

In order to be useful, a path difference metric needs to reproduce this order. Note that a question mark is put on top of the “less than” symbols between *Pairs d* and *e*, and *Pairs e* and *f* because when comparing pairs with different spatial separation, geometric difference, and relative spatial uncertainty, it can be hard to objectively define which “should” have the highest similarity, and the ordering will depend on the relative weighting of  $d_s$ ,  $d_l$  and  $d_a$ . If the weightings are equal (i.e.,  $W_s = W_l = W_a = \frac{1}{3}$ ), significant CPD scores for paths without interpolated poles (i.e., using scores from Fig. 2.8) reproduce the expected order:

$$\mathbf{a}(0) < \mathbf{b}(0.036) < \mathbf{c}(0.043) < \mathbf{d}(0.067) \approx \mathbf{e}(0.067) < \mathbf{f}(0.174) < \mathbf{g}(0.177),$$

however *Pairs d* and *e* have almost identical scores and are not discriminated. However, their fit quality (B-B for *Pair d* and B-A for *Pair e*; Fig. 2.8b, Fig. 2.8d and Fig. 2.8f) indicates that *Pair e*’s  $\mathcal{CPD}$  is relatively more trustworthy. This order might also not be preserved with different applied weights. The impact of weighting will be discussed in the following section.

### 2.3.2 A Discussion on Weights

Although  $W_s$ ,  $W_a$  and  $W_l$  can be defined by user, this is subjective.

However, we do explicitly know that: when comparing two APWP pairs, our aim is to find the one whose similarity ranks higher. A simple subtraction between unsolved (because of unknown weights)  $\mathcal{CPDs}$  can help determine which pair’s similarity ranks higher. If a positive difference is obtained, no matter what  $W_s$ ,  $W_a$  and  $W_l$  values are assigned, the subtrahend pair’s similarity ranks higher; if the difference is always negative, the minuend pair’s similarity is always higher. In addition, the difference could be always zero. Interpretation is straightforward in these three scenarios. However, for some pairs, a positive, zero or negative CPD difference could result depending on the chosen weightings.

For example, for the full (i.e., 0–530 Ma) path with no interpolated poles, the mean significant spatial, length and angular differences  $d_s$ ,  $d_l$  and  $d_a$  are known (Fig. 2.8b,

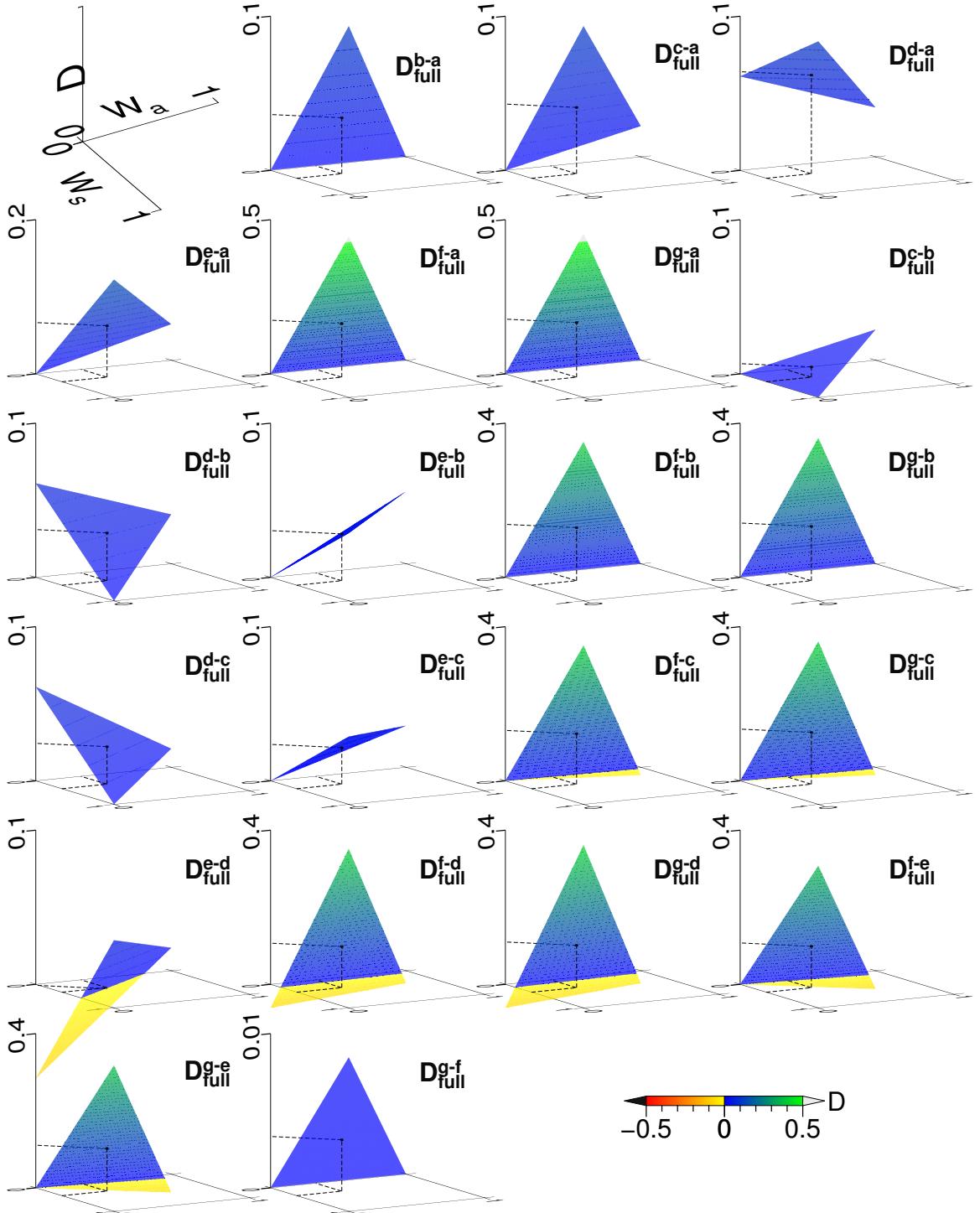
Fig. 2.8d and Fig. 2.8f). Also we know  $W_l = 1 - W_s - W_a$ . Then we do subtractions of  $\mathcal{CPD}s$  from each two APWP pairs:

$$\left\{ \begin{array}{l} D_{full}^{b-a} = \mathcal{CPD}_{full}^b - \mathcal{CPD}_{full}^a = 0.109W_s \\ D_{full}^{c-a} = \mathcal{CPD}_{full}^c - \mathcal{CPD}_{full}^a = 0.109W_s + 0.02W_a \\ D_{full}^{d-a} = \mathcal{CPD}_{full}^d - \mathcal{CPD}_{full}^a = 0.038W_s - 0.029W_a + 0.061 \\ D_{full}^{e-a} = \mathcal{CPD}_{full}^e - \mathcal{CPD}_{full}^a = 0.153W_s + 0.047W_a \\ D_{full}^{f-a} = \mathcal{CPD}_{full}^f - \mathcal{CPD}_{full}^a = 0.522W_s \\ D_{full}^{g-a} = \mathcal{CPD}_{full}^g - \mathcal{CPD}_{full}^a = 0.532W_s \\ D_{full}^{c-b} = \mathcal{CPD}_{full}^c - \mathcal{CPD}_{full}^b = 0.02W_a \\ D_{full}^{d-b} = \mathcal{CPD}_{full}^d - \mathcal{CPD}_{full}^b = -0.061W_s - 0.029W_a + 0.061 \\ D_{full}^{e-b} = \mathcal{CPD}_{full}^e - \mathcal{CPD}_{full}^b = 0.044W_s + 0.047W_a \\ D_{full}^{f-b} = \mathcal{CPD}_{full}^f - \mathcal{CPD}_{full}^b = 0.413W_s \\ D_{full}^{g-b} = \mathcal{CPD}_{full}^g - \mathcal{CPD}_{full}^b = 0.423W_s \\ D_{full}^{d-c} = \mathcal{CPD}_{full}^d - \mathcal{CPD}_{full}^c = -0.061W_s - 0.049W_a + 0.061 \\ D_{full}^{e-c} = \mathcal{CPD}_{full}^e - \mathcal{CPD}_{full}^c = 0.044W_s + 0.027W_a \\ D_{full}^{f-c} = \mathcal{CPD}_{full}^f - \mathcal{CPD}_{full}^c = 0.413W_s - 0.02W_a \\ D_{full}^{g-c} = \mathcal{CPD}_{full}^g - \mathcal{CPD}_{full}^c = 0.423W_s - 0.02W_a \\ D_{full}^{e-d} = \mathcal{CPD}_{full}^e - \mathcal{CPD}_{full}^d = 0.105W_s + 0.076W_a - 0.061 \\ D_{full}^{f-d} = \mathcal{CPD}_{full}^f - \mathcal{CPD}_{full}^d = 0.474W_s + 0.029W_a - 0.061 \\ D_{full}^{g-d} = \mathcal{CPD}_{full}^g - \mathcal{CPD}_{full}^d = 0.484W_s + 0.029W_a - 0.061 \\ D_{full}^{f-e} = \mathcal{CPD}_{full}^f - \mathcal{CPD}_{full}^e = 0.369W_s - 0.047W_a \\ D_{full}^{g-e} = \mathcal{CPD}_{full}^g - \mathcal{CPD}_{full}^e = 0.379W_s - 0.047W_a \\ D_{full}^{g-f} = \mathcal{CPD}_{full}^g - \mathcal{CPD}_{full}^f = 0.01W_s, \end{array} \right. \quad (2.1)$$

and we also have the following constraints of feasible regions

$$\left\{ \begin{array}{l} 0 < W_s < 1 \\ 0 < W_a < 1 \\ 0 < W_s + W_a < 1. \end{array} \right. \quad (2.2)$$

The linear equations (2.1) subject to (2.2) can be graphed in the three-variable ( $W_s$ - $W_a$ - $D$ ) coordinate system (Fig. 2.9). For all possible combinations of  $W_s$  and  $W_a$ , there is a consistent ordering of CPD scores such that  $\mathbf{a} < \mathbf{b} < \mathbf{c} < \mathbf{d}$ ,  $\mathbf{a} < \mathbf{b} < \mathbf{c} < \mathbf{e}$ , and  $\mathbf{a} < \mathbf{b} < \mathbf{f} < \mathbf{g}$  (Fig. 2.9). However, the ranking for *Pairs f, g* and *Pairs c, d, e*, or *Pair d* and *Pair e* has multiple possibilities, because their differences can be positive, negative or zero (Fig. 2.9). For this situation, assigning equal weights is recommended (giving centroid of all possible  $D$ s, Fig. 2.9; see also Supplementary Information for testing equally likely random weights) for deciding the rank order. With equal weights used, the order from most similar pair to least similar pair is  $\mathbf{a} < \mathbf{b} < \mathbf{c} < \mathbf{d} \approx \mathbf{e} < \mathbf{f} < \mathbf{g}$ . These conclusions do not contradict the expected Order (2.4).



**Figure 2.9:** Graphical depiction of  $\mathcal{CPD}$  differences ( $D$ ) between the seven APWP pairs for full-path (0–530 Ma) comparisons. If the planes derived from the equations intersect the  $D = 0$  plane at a point or in a straight line, that point or the infinite number of points (i.e., sets of  $W_s$ ,  $W_a$  values) on the line of intersection represent that the similarities of the minuend pair and the subtrahend pair are equal to each other. If  $D > 0$  or  $D < 0$  on the planes of the equations, the subtrahend pair or the minuend pair respectively owns higher similarities. The square dot locates the result when  $W_s = W_a = W_l = \frac{1}{3}$ .

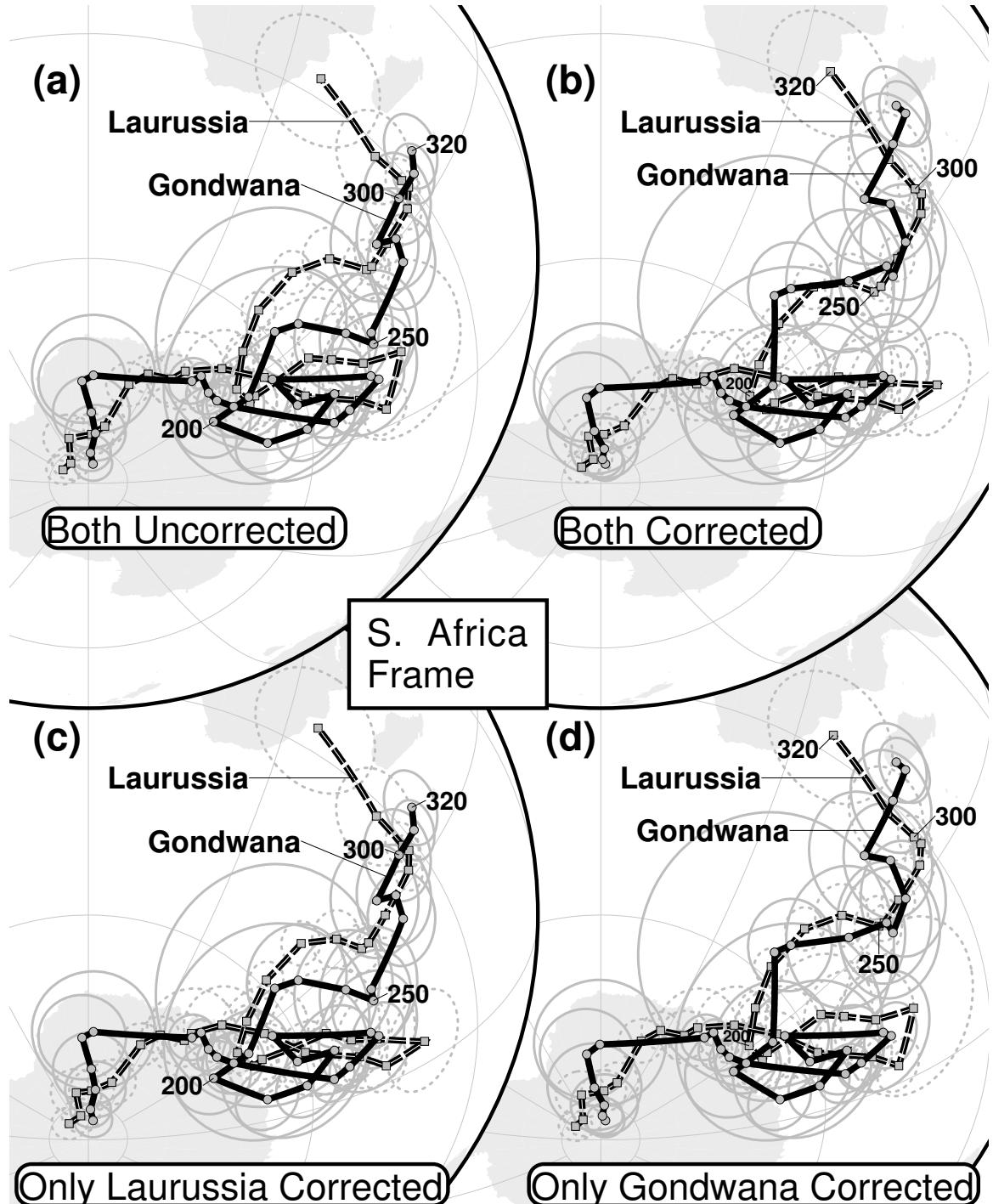
In summary, as Fig. 2.3 illustrates, mean GCD has trouble discriminating between *Pairs f* and *g*, and *Pairs c* and *e*, and also between intermediate similarities where the differences are mainly in path geometry (*Pairs b, c, d*). Our algorithm provides an improved solution for this problem. Obtaining similarity order can be straightforward, such as for *Pairs a, b, c, and d*, *Pairs a, b, c, and e*, or *Pairs a, b, f, and g*. In other words, when one APWP's three individual metrics are all greater than or equal to, or less than or equal to the other one's, weighting is irrelevant. However, when the ranking of individual metrics for a pair are not consistent (e.g., *Pair f* and *Pair c*; Fig. 2.9), obtaining similarity order is less straightforward. When this occurs, equally weighting is recommended for concluding the final rank order.

### 2.3.3 Application to Real Paleomagnetic Data

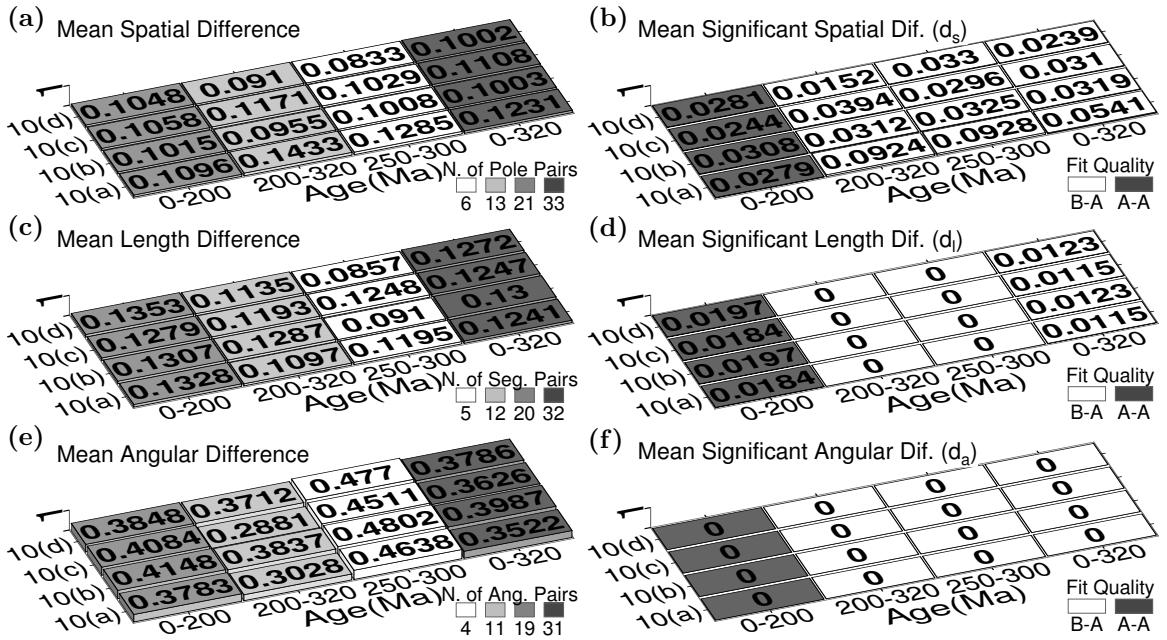
To illustrate how these metrics might be useful when applied to real paleomagnetic data, we compare 320–0 Ma APWPs for Gondwana and Laurussia calculated using a running mean method by (Torsvik et al., 2012), with paleopoles from sedimentary rocks both uncorrected (Fig. 2.10a) and bulk corrected for inclination shallowing ( $f=0.6$ ; Fig. 2.10b).

When comparing the 320–0 Ma paths (Torsvik et al., 2012, in their Fig. 13a) observed that a bulk correction for inclination shallowing applied to poles from sedimentary rocks reduced the mean GCD separation between poles, particularly in the Permian section of the path.

The full-path (320–0 Ma)  $d_s$  scores of *Pair 2.10a* and *Pair 2.10b* (Fig. 2.11) confirm that the corrected Gondwana and Laurussia APWPs (*Pair 2.10b*) are more similar than the uncorrected pair (*Pair 2.10a*). This difference is significant (Fig. 2.11b), and it is principally the result of an improved fit (lower  $d_s$ ) in the Permian (300–250 Ma) and Carboniferous to Triassic (320–200 Ma) sections. When geometry is considered,  $d_l$  and  $d_a$  scores without significance testing (Fig. 2.11c and Fig. 2.11e) are actually higher for corrected (*Pair 2.10b*) than uncorrected pair (*Pair 2.10a*), particularly in the same Permo-Triassic segment; however, none of these differences are statistically significant (Fig. 2.11d and Fig. 2.11f). Therefore, the equally weighted CPD score is actually worse for corrected pair (*Pair 2.10b*) (0.213 versus 0.1998 for *Pair 2.10a*), whereas with significance testing applied the CPD of the corrected pair is a clear improvement (0.0147 versus 0.0218). This emphasises the importance of significance testing for the geometric scores, particularly where the spatial uncertainties are relatively large compared to the step length, as is indicated by the fit quality score (B-A, Gondwana-Laurussia).



**Figure 2.10:** (a) 320–0 Ma APWPs (10 Myr step) for Gondwana and Laurussia (Torsvik et al., 2012, rotated to Southern Africa frame using the rotations from); (b) as (a), but both paths corrected for inclination shallowing; (c) and (d) as (a), but only Laurussia path and only Gondwana path respectively corrected for inclination shallowing. Note that all the paleomagnetic APWPs are reproduced using the same moving average method and same paleopoles for the APWPs in Figure 13(a) of Torsvik et al. (2012). Azimuthal Orthographic projection.



**Figure 2.11:** Mean spatial, length and angular differences between two paths of the four APWP pairs shown in Fig. 2.10. Left column: results without significance testing imposed in the metric; Right column: results with testing. Note that Pair 2.10b with both APWPs corrected for inclination shallowing is not the most similar pair according to both the untested (left column) and tested (right column) results.

Furthermore, if this analysis is extended to compare an Laurussia APWP corrected for inclination flattening with a Gondwana APWP that has not (*Pair 2.10c*), and vice versa (*Pair 2.10d*), removing the blanket correction from Gondwana poles has only a minor effect on the overall  $d_s$ , whilst removing it from the Laurussia poles actually improves  $d_s$  after significance testing (Fig. 2.11a and Fig. 2.11b). Comparison of the changes in sub-path scores for *Pairs 2.10b*, *2.10c* and *2.10d* suggests that the effect of the bulk flattening correction is sometimes positive and sometimes negative for different time periods, supporting arguments that flattening corrections need to be more judiciously applied (Bilardello and Kodama, 2010; Bilardello, 2016, for example). A more detailed study might further constrain the regions, continents and/or time periods where a correction is appropriate, and those where it is not. But for the purposes of this paper, this overview is sufficient to demonstrate the potential usefulness of our difference metrics when considering the effect of different techniques and corrections used to generate an APWP.

## 2.4 Conclusions

A new synthetic evaluation method is proposed in this paper to serve as a numerical tool for the purpose of quantitatively matching paleomagnetic APWPs. Multidimensional information tested by bootstrapping, such as overlap of coeval poles and shape of paths, are taken into account in the algorithm. This method can also be utilized to detect APWP subsections' degree of similarity by changing trajectory beginning and end poles. As an example of how this method can be applied, we confirm a previously published suggestion that applying a blanket correction for inclination shallowing in sedimentary rocks does significantly improve the fit between Carboniferous to Recent APWPs for Gondwana and Laurussia. However a more detailed analysis also indicates that such blanket inclination corrections are unlikely to produce the best possible fit.

## Computer Code Availability

**Name of Code** Spherical Path Comparison (spComparison)

**Developer** Chenjian Fu and Christopher J. Rowan

**Contact Address** 221 McGilvrey Hall, 325 S Lincoln St, Kent, OH 44242 USA

**Telephone Number** 4408479166

**E-mail** cfu3@kent.edu

**Year First Available** 2019

**Hardware Required** Intel (R) Core (TM) i7-6700 CPU @ 3.40GHz or higher; 8 GB DDR3 RAM or higher

**Software Required** GMT5 or higher; Python 3.6 or higher; Bash 4.4.23 or higher; Linux as the best platform, macOS also fine, for Windows further setup needed

**Program Language** Python 3 and Shell Scripting (Bash)

**Program Size** 75 KB (1500 lines)

**Details on How to Access the Source Code** The source code can be accessed from [https://github.com/f-i/Spherical\\_Path\\_Comparison](https://github.com/f-i/Spherical_Path_Comparison). Please use the provided Jupyter Notebook file “demo.ipynb” to reproduce some calculations shown in the paper.

# Chapter 3

## Methods for Producing a Reliable APWP

*This chapter mainly describes how to generate paleomagnetic APWPs using 168 different methods, and then the application of the new APW path similarity measuring tool described in the last chapter in finding the best APWP generating method(s). 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.*

### 3.1 Introduction

A paleomagnetic pole, also known as paleopole, has many attributes, including sampling site, number of samples, sample rock characteristics, age determination and its uncertainty, pole location and its uncertainty etc. And APWPs are generated by combining paleopoles into mean poles, for a particular rigid block over the desired age range to produce a smoothed path (see detailed definition of APWP in Section 1.1.2). APWP reflects cumulative motions of a continental fragment relative to the Earth's spin axis. So, in order to build an APWP, an object continent or continental fragment should be picked first. And then the general "boundary" of this continent or continental fragment should be determined tectonically. This closed tectonic "boundary" is then used to constrain the paleomagnetic datasets through picking only those with their sampling sites lying inside the closed "boundary". See Appendix B for examples about how the paleopole datasets are constrained for a particular tectonic plate during a specific time interval. Then the paleopoles that derive from those sampling sites are statistically combined together into mean poles when their ages are close. The final products, those mean poles, are then connected according to their estimated ages to make a time-series path which is called paleomagnetic APWP. So it is important to ask a question like what can be called a reliable paleomagnetic APWP or how to build a reliable APWP, because, for example (and also what we want to focus our work on here is), there are so many different perspectives to consider when we combine paleopoles into mean poles. For instance, how close a group of paleopoles' ages should be to be combined together into a mean pole, what statistical way should be used to do this combining, should those paleopoles be treated equally in this process of combining, if no what factor affects more than others, etc. These are the factors that might affect the end result.

In this chapter, paleopoles' attributes are used to weigh their influence on mean poles (see details in Section 3.2.3.3), or to determine if they should be omitted for producing a 'better' mean pole (see details in Section 3.2.3.2). Different key attributes, that can be quantified, or their combination are considered. In addition, when we are merging paleopoles to produce a smoothed mean path, we make choices not only about which data are included or excluded, but how data are combined. Moving averaging is used to combine data (see details in Section 3.2.3.1). With application of two moving average methods, where each moving window includes data with (1) only middle point of lower and higher age limits considered as paleopole's age and (2) whole dating uncertainty considered as paleopole's age range, different APWPs are

produced to see which attribute is more important and also which moving average method performs better. To answer these questions, we need a reference path (see details in Section 3.2.4) and a tool to measure similarity between the paleomagnetic APWPs and the reference path. Then the measuring tool developed and introduced in Chapter 2 is used to measure the similarity. We have two reference paths derived from other plate kinematics models. These reference models are thought to be more accurate and more precise, so that if the paleomagnetic APWPs are similar to the reference path, they are regarded as more ‘reliable’. The question is which model predicted reference path is a better reference or is there any difference between these two reference paths? In this chapter, we will also try to answer this question.

Finally no single attribute generally works better than others for all the continents or continental fragments. However, the moving average method, that considers the whole range of paleopoles’ temporal uncertainty when calculating mean pole’s age is needed, is proved that it is obviously a better way of generating reliable APWPs than the one that only considers the mid-point of temporal uncertainty as a paleopole’s age. And weighting is proved to be actually unnecessary. The reference paths derived from other models are not significantly different.

### 3.1.1 How An APWP Is Generated and Key Factors

As mentioned above, the first question to ask is that which tectonic plate the paleopole we are working on belongs to (see Section 1.1.3 and details about how it is solved in Appendix B). The fact is that paleomagnetic data is not just about poles but multiple attributes integrated and not all data are created equal. So once the host tectonic plate is determined, we encounter another problem: the attributes, for example, uncertainties of ages and also locations, can vary greatly for different paleopoles (see details in Sections 1.1.4.1 and 1.1.4.2). Paleopole’s attributes like age and location obviously can be quantified and then used to weigh paleopoles or to omit some of them through setting up a range of acceptable values. In addition, the age uncertainty can also be used to change the way how we do moving averaging when we calculate mean poles through considering the whole uncertainty range instead of only considering the mid-point of the uncertainty range in the traditional way (this is also how we get two different moving average methods). Besides age and position uncertainties of paleopoles, the data consistency is also needed to be investigated carefully (see details in Section 1.1.4.3), because some paleopoles’ polarity given in the GPMDB4.6b (McElhinny and Lock, 1996; Pisarevsky, 2005, updated in 2016 by the Ivar Giaever Geomagnetic Laboratory team, in collaboration with Pisarevsky)

could be wrong although most are correct. Data density is also vital (see details in Section 1.1.4.4) because paleomagnetism is basically built on statistics. In addition, when the paleopoles were published (publication year) is also a key factor that potentially indicates the quality of the paleopoles (see details in Section 1.1.4.5), simply because in old times, technology of magnetism measuring was not that advanced as today. In consideration of all these factors, we organise them in the general processing steps below in order to generate a paleomagnetic APWP.

### 3.1.1.1 Picking/Binning Step

The moving average method is used to combine paleopoles into mean poles. The moving time window/bin picks a group of paleopoles at each time step (see description in Section 3.2.3.1 and Fig. 3.4). So the key factors are:

**Width of time window** (trade off between number of paleopoles and amount of smoothing): Smoothness of paleomagnetic APWPs generated by moving averaging could change with different time window lengths and steps. If window length is too wide and step is too long, final path would be so smoothed that actual details would be missed. If window length is too narrow and step is too short, final path would be jerky piecewise so that too much noise would be introduced because there are too few datasets in each window. So a balance needs to be achieved.

**Age uncertainty in magnetisation**, particularly when that uncertainty is larger than the selection window. The mid-point is commonly used as a selection point (see an example in upper panel of Fig. 3.4), but when the minimum age is constrained by a field test (common when the age range is large; see Section 1.1.4.1), the assumption that this is the most likely age is questionable. As it is mentioned in Section 1.1.4.1, if field test shows magnetisation acquired prior to, for example, a folding event that is tens of millions of years after initial rock formation, a passed field test is not actually very useful. So age uncertainty is very large, the possibility of remagnetisation is certainly very high, which means the age is probably closer to the minimum age instead of the middle age.

### 3.1.1.2 Filtering Step

This step filters out bad data that is likely to be unreliable, based on known information about the paleopoles. These can be subdivided into two groups:

**Characteristics that indicate the paleopole is well-constrained (precise)**  
(e.g., has a small  $\alpha_{95}$  uncertainty ellipse, large  $\kappa$ )

**Characteristics that indicate the paleopole is reliable (accurate) :**

- Lithology, particularly the risk of inclination flattening in sediments (see Section 1.1.4.3)
- Risk of unidentified local rotations in deformed areas (see Section 1.1.4.3)
- Publication year — younger studies that use stepwise demagnetisation techniques more likely to remove overprints and isolate a primary remanence (see Section 1.1.4.5).
- Sampling strategy — need sufficient number of contributing samples and sites (N and B in Table 3.1) to have a good chance of averaging out secular variation (see Section 1.1.4.3).
- Field tests that constrain the magnetisation age (see Section 1.1.4.1).

### 3.1.1.3 Calculation Step

Normally all paleopoles are treated equally when calculating the Fisher mean (see how to calculate a Fisher mean using the formulas 6.12, 6.13, 6.14 and 6.15 in Butler, 1992, chap. 6; note that instead of direction declination and inclination expected in those formulas, pole longitude and latitude should be used), but as an alternative or supplement to filtering paleopoles can potentially be weighted according to the factors that potentially influence their precision and/or accuracy such as the A95 uncertainty or age uncertainty, and a weighted Fisher mean calculated, giving the ‘better’ paleopoles more influence on the result mean pole. So, for example, each paleopole has got a weight before calculating for a mean. In this dissertation these weights  $w_i$  are integrated into the  $\sum_{i=1}^N l_i$ ,  $\sum_{i=1}^N m_i$  and  $\sum_{i=1}^N n_i$  of the formulas 6.13 and 6.14 in Butler (1992, chap. 6), where  $w$  means weight and  $i$  is the same as in those two formulas. Then for a weighted Fisher mean, the two formulas 6.13 and 6.14 in Butler (1992, chap. 6) become

$$l = \frac{\sum_{i=1}^N l_i w_i}{R} \quad m = \frac{\sum_{i=1}^N m_i w_i}{R} \quad n = \frac{\sum_{i=1}^N n_i w_i}{R} \quad (3.1)$$

and

$$R = \left( \sum_{i=1}^N l_i w_i \right)^2 + \left( \sum_{i=1}^N m_i w_i \right)^2 + \left( \sum_{i=1}^N n_i w_i \right)^2 \quad (3.2)$$

Note that the trade-offs and difficulties are similar for filtering, e.g. giving poles with low A95 more weight presumes that they are closer to the actual mean, which is not necessarily the case. Therefore, weighing on different paleomagnetic factors is hoped to help us have some insight into these complex issues.

### 3.1.2 Existing Solutions and General Issues

As mentioned in the concluding paragraph of Section 1.1.4, if not all paleopoles are created equal, the question becomes: how should paleopoles of varying quality be combined? For a certain set of paleopoles, how can we produce an APWP that is both

1. well constrained: the mean pole for each time step has a low spatial uncertainty (small 95% confidence ellipse), and
2. accurate: the mean pole position is close to its actual position at each time step.

Previous work on this question have focussed on the filtering out so-called “bad” data before calculation of the mean pole. Commonly used schemes include:

1. van der Voo (1990) (V90). V90 includes seven criteria (see the details in the concluding paragraph of Section 1.1.4). The Q (quality) factor is the total number of criteria satisfied (0–7) (van der Voo, 1988). The Q factor is a very straightforward way to get a quantitative reliability score, and is widely used for filtering paleopoles prior to calculating APWPs (e.g. Torsvik et al., 2012; Ma et al., 2016; Franco, 2019). Further each paleopole can also be conveniently weighted in proportion to its Q factor in the 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.
2. Besse and Courtillot (2002) (BC02). BC02 data quality criteria use only paleopoles with  $\alpha_{95}$  less than  $10^\circ$  in the Cenozoic and  $15^\circ$  in the Mesozoic, derived from at least 36 samples from at least 6 sampling sites (see also the concluding paragraph of Section 1.1.4). While straightforward and convenient to apply, these stringent criteria mean some useful data may be filtered out and wasted, especially for a period where there are only limited number of paleopoles.

3. Schettino and Scotese (2005) (SS05). SS05 is similar to, but less stringent than, BC02. SS05 uses only paleopoles with  $\alpha_{95}$  of  $\leq 15^\circ$  and an age uncertainty of  $\leq 40$  Myr, derived from at least  $\geq 4$  sampling sites with  $\geq 4$  samples/site, and at least some sample demagnetisation (See also the last paragraph of Section 3.2.3.2).

Although they are often used, and effect the path and uncertainties of the resulting APWP, there has been limited study of how effective these filtering methods are at reconstructing a ‘true’ APWP compared to unfiltered data. Furthermore, the focus on the filtering stage ignores the possible impact of binning/windowing (i.e. combining paleopoles through moving-averaging) and weighting. This study more rigorously investigates the effects of the choices made at every stage on the resulting APWP. By focussing on paleomagnetic data from the last  $\sim 120$  Myr, where plate motions are independently constrained by reconstructions of seafloor spreading tied into a hotspot reference frame, we can also verify which choices actually lead to a better-constrained and accurate record of past plate motions.

## 3.2 Methods

### 3.2.1 General Approach

In this study, we use paleopoles extracted from the GPMDB4.6b (McElhinny and Lock, 1996; Pisarevsky, 2005) 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 (Table 3.2 and 3.3). 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.

### 3.2.2 Paleomagnetic Data

#### 3.2.2.1 GPMDB Field Codes

Data analysis includes manipulation of data fields/columns in the GPMDB. In the following content, the codes of the several specific fields used will be referred to. They are listed in Table 3.1 for easy reference.

**Table 3.1:** List of the used fields and field codes of the GPMDB.

Field Code	Meaning
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. $\alpha_{95}$
EP95	Radius of circle of 95% confidence about paleopole position
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
ROCKNAME	Name of sample rock
ROCKTYPE	Type of sample rock
STATUS	Indicates if results have been superseded
COMMENTS	General information including details of the origin of LOMAGAGE and HIMAGAGE. If the result is a combined pole this field contains information on the data included in the combined result

### 3.2.2.2 Paleomagnetic Data of Three Representative Continents

Collections of paleopoles with a minimum age (LOMAGAGE)  $\leq 135$  Ma for the North American, the Indian and Australian 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 and Samet, 2007) was used to find all poles within the geographic region that defines the rigid plate within the period of interest (see Appendix B for details):

**For North America,** the search region was defined by the North American craton (NAC), Avalon/Acadia and Piedmont blocks, as defined by the recently published plate model of Young et al. (2019). Following extraction, 58 paleopoles from southwestern North America that have been affected by regional rotations since 36 Ma (McQuarrie and Wernicke, 2006), were removed. The final dataset consists of 135 paleopoles (Fig. 3.1), with 76 of them ( $\sim 56\%$ ; average age uncertainty  $\sim 14$  Myr, average EP95  $\sim 9.3^\circ$ ) sampled from dominantly igneous sequences, 56 ( $\sim 42\%$ ; average age uncertainty  $\sim 27.5$  Myr, average EP95  $\sim 10.5^\circ$ ) sampled from mostly sedimentary

sequences, including 6 from redbeds, and 3 ( $\sim 2\%$ ) 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.

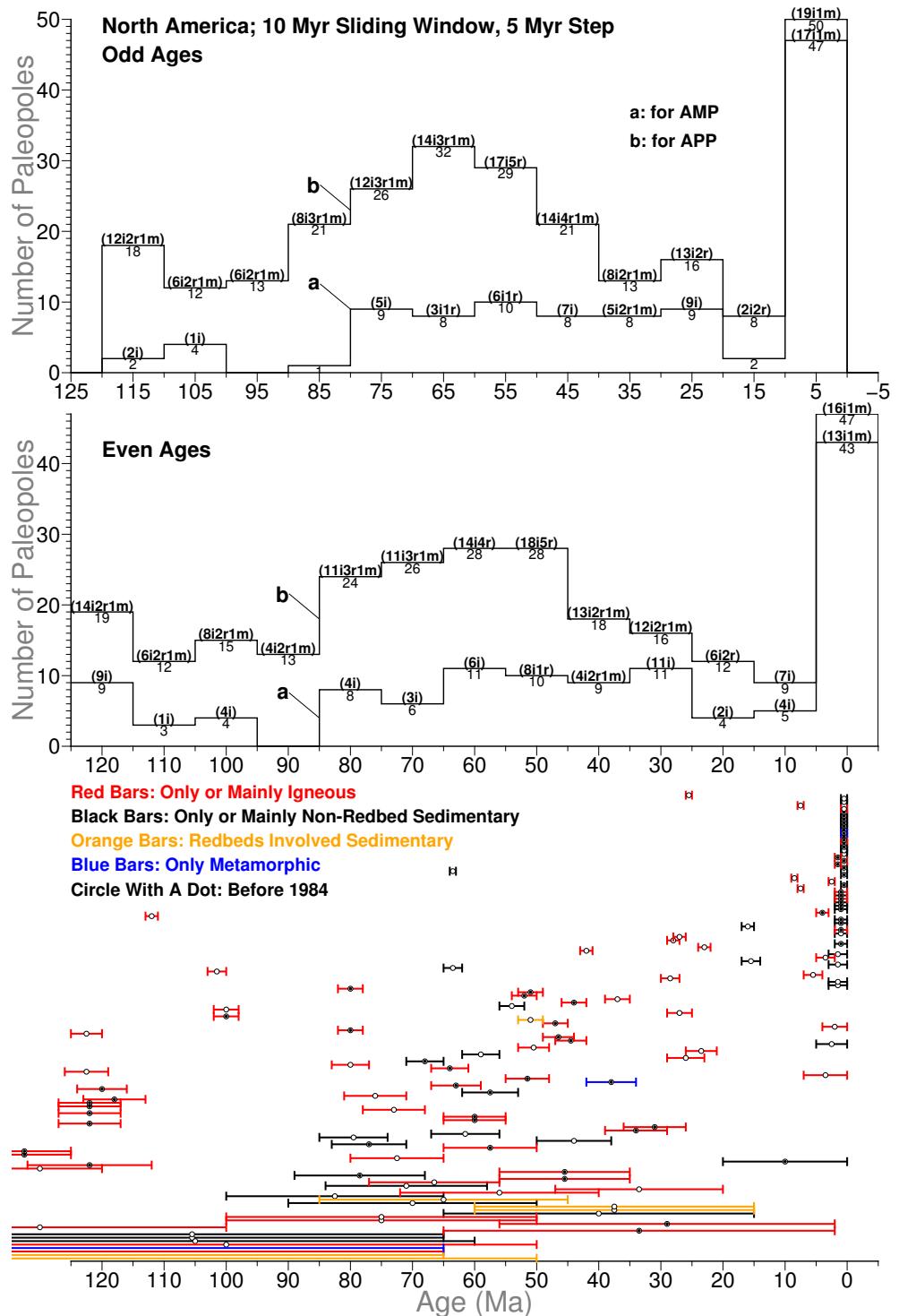
**For India,** the Indian block as defined by Young et al. (2019) was used, but following extraction 31 paleopoles associated with parts of the northern margin that have undergone regional rotations since the Jurassic (Gaina et al., 2015) were removed. The final dataset consists of 75 paleopoles (Fig. 3.2), with 39 of them (52%; average age uncertainty  $\sim 5$  Myr, average EP95  $\sim 7.7^\circ$ ) sampled from dominantly igneous sequences and 36 (48%; average age uncertainty  $\sim 14$  Myr, average EP95  $\sim 7^\circ$ ) sampled from mostly sedimentary sequences, including 3 from redbeds. There is a high concentration of poles from the latest Cretaceous–Early Cenozoic ( $\sim 70$ –60 Ma), many of which are igneous; in younger and older intervals, there are fewer, mostly sedimentary poles.

**For Australia,** the Australia, Sumba, and Timor blocks as defined by Young et al. (2019) were used, in combination with data from the Tasmania block younger than  $\sim 100$  Ma (with a maximum age (HIMAGAGE)  $\leq 100$  Ma), prior to which it was not fixed with respect to Australia (Young et al., 2019). The final dataset consists of 99 paleopoles (Fig. 3.3), with 61 of them ( $\sim 62\%$ ; average age uncertainty  $\sim 23.5$  Myr, average EP95  $\sim 10.9^\circ$ ) sampled from dominantly igneous sequences, and 38 ( $\sim 38\%$ ; average age uncertainty  $\sim 23$  Myr, average EP95  $\sim 9.4^\circ$ ) sampled from mostly sedimentary sequences, including 9 from redbeds. The temporal distribution of poles is relatively uniform.

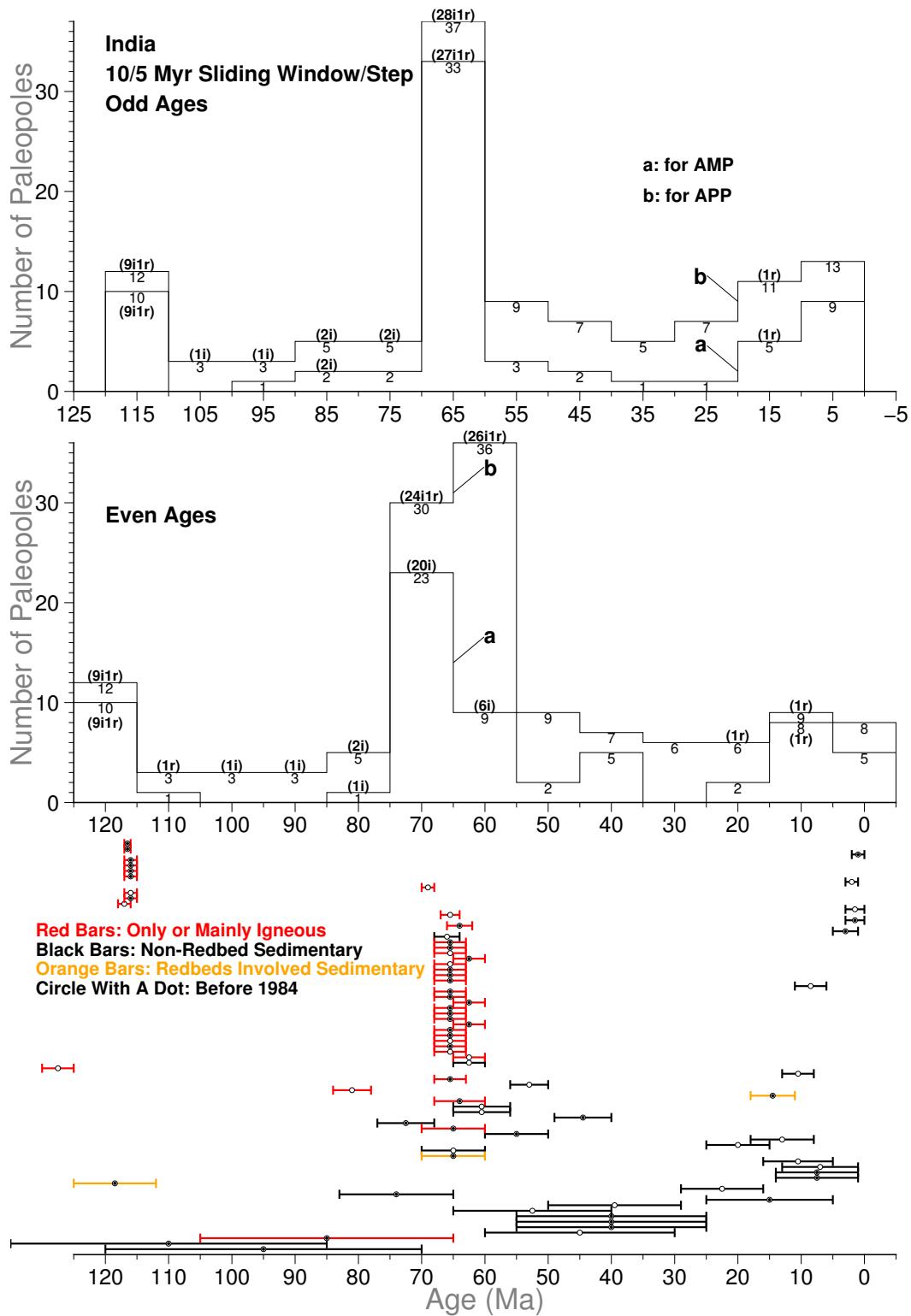
Compared with North American (Fig. 3.1) and Australian (Fig. 3.3) paleopoles, Indian paleopoles have a relatively lower density and a higher prevalence of sedimentary paleopoles, except during the period of  $\sim 70$ –60 Ma (Fig. 3.2). For North America and India, sedimentary paleopoles have significantly larger average age uncertainty, but about same for Australia, there is no igneous and sedimentary  $\alpha 95$  difference.

### 3.2.3 APWP Generation

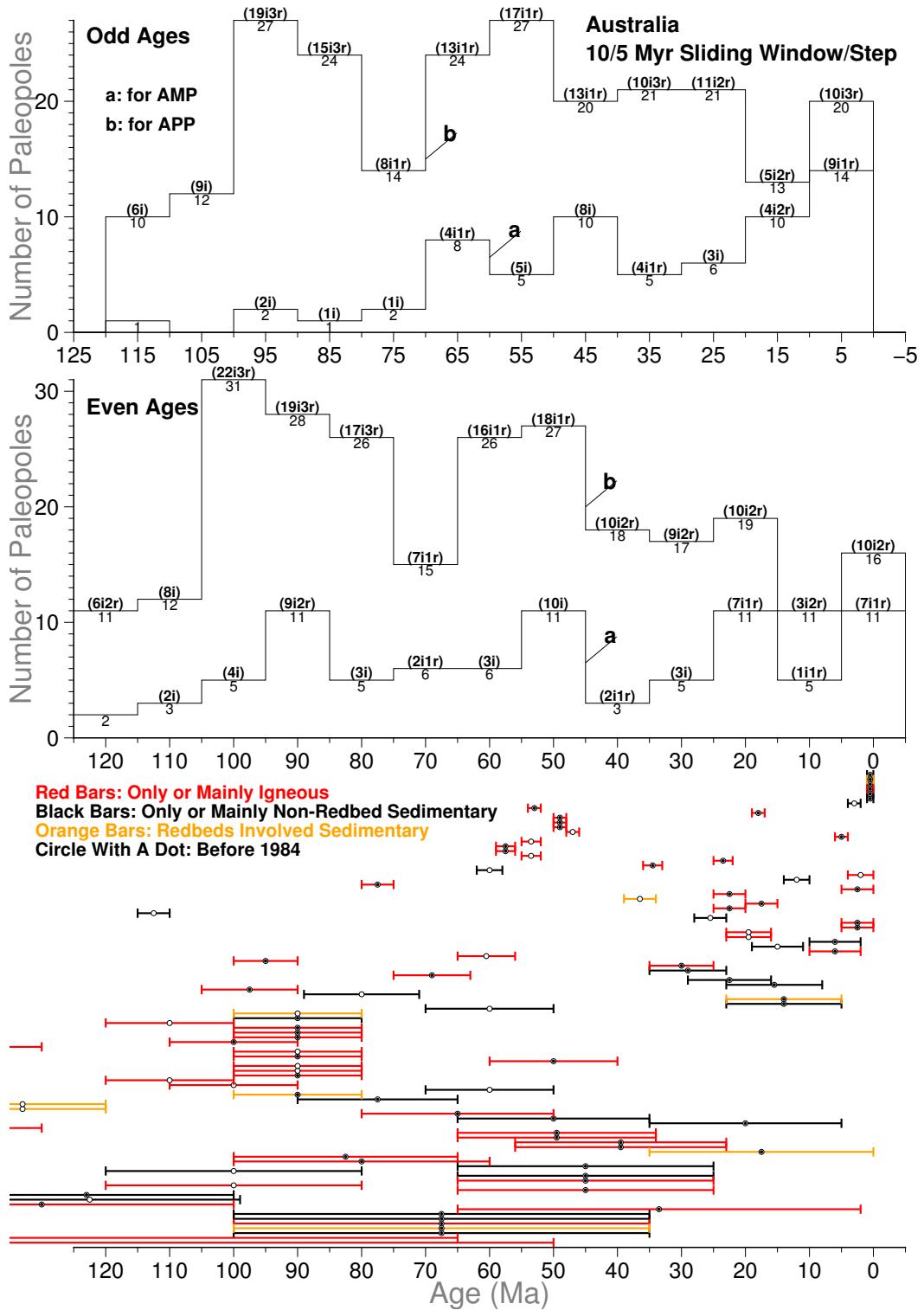
Multiple APWPs were generated using the selected poles for each of the three plates as follows:



**Figure 3.1:** Temporal distribution of 120–0 Ma NAC paleopoles in 10 Myr window length and 5 Myr step length. For distribution a, each bin only counts in the midpoints (circles) of pole uncertainty 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 3.2:** Temporal distribution of 120–0 Ma Indian paleopoles. For red bars, only one pole, 67–64 Ma (RESULTNO 8593), is from igneous and also sedimentary. See Fig. 3.1 for more information.



**Figure 3.3:** Temporal distribution of 120–0 Ma Australian 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. 3.1 for more information.

**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.

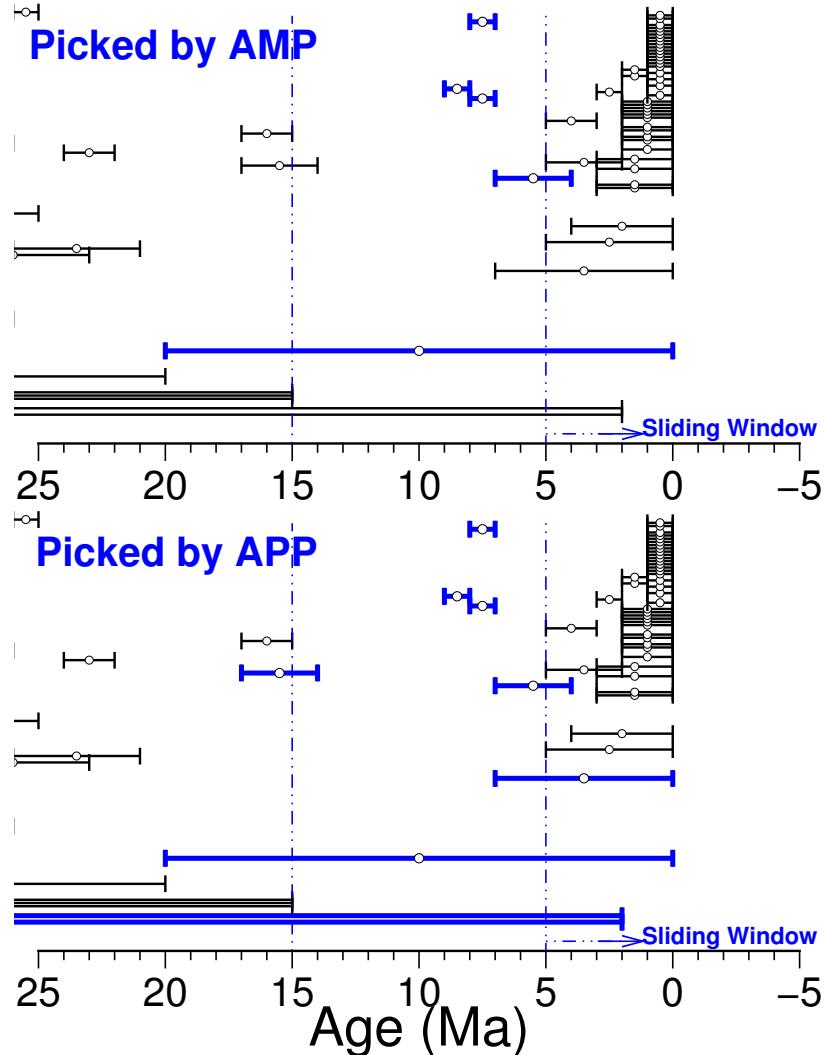
Twenty-eight different picking and filtering algorithms were tested (Table 3.2, referred to hereafter as Pk), in combination with 6 different weighting algorithms (Table 3.3, referred to hereafter as Wt), for the three plates. The effects of changing the time step length and width of the moving window, and the reference path, were also examined.

### 3.2.3.1 Picking/Binning

In studies where the moving window method is used to calculate an APWP (Torsvik and 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 3.2, Fig. 3.9 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 3.2, Fig. 3.9 and subsequent figures). The APP algorithm will pick more paleopoles in each moving window than the AMP algorithm (Figs. 3.1-3.3; Fig. 3.4).

If, for example, we have a paleopole with an acquisition age of 10–20 Ma, and we have a 2 Myr moving window with a 1 Myr age step, then it is included in just 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. Each original paleopole is therefore represented in the mean poles calculated over its entire possible acquisition age.



**Figure 3.4:** 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. 3.6) of only 5 paleopoles, while the APP method calculates the mean pole (dark circle in Fig. 3.6) of 9 paleopoles.

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.

### 3.2.3.2 Filtering

14 different filters or corrections (Table 3.2) 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 (Pk 0,1).**

**Removal of poles with large spatial and temporal uncertainties (Pk 2,3).** Paleopoles with both large  $\alpha_{95}$  ( $ED95 > 15^\circ$ , 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 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 (Pk 4,5, 6,7).** Pk 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 Pk 4,5 are derived from igneous-only rocks. Pk 6,7 select 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. Therefore Pk 6 also includes paleopoles from Pk 4, and Pk 7 also includes paleopoles from Pk 5.

**Correct sedimentary poles for inclination shallowing (Pk 8,9).** Rather than excluding paleopoles from sedimentary rocks, paleopoles coded as sedimentary or rebeds were instead corrected for inclination flattening using the flattening function

**Table 3.2:** List of all picking/binning algorithms developed here.

No.	Picking Algorithm
0	AMP: Age Mean Picking, see Section 3.2.3
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 rebeds)
13	APP (unflatten rebeds)
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 and Scotese (2005)

$\tan I_o = f \tan I_f$  (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 the rock type (ROCKTYPE field in the database) is not sedimentary dominated but contains sedimentary rocks. In these cases,  $f = 0.8$  is used instead, following the minimum anisotropy-of-thermal-remanence determined f-correction (Domeier et al., 2011a,b).

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

**Prefer poles from younger (Pk 14,15, 24,25) or older (Pk 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. Pk 14,15 removes any paleopoles published prior to 1983 (YEAR > 1983), the mean publication date for paleopoles in the GPMDB. Pk 24,25 takes a similar but less aggressive approach by excluding paleopoles that have been superseded (99 paleopoles) by later studies from the same sequence, which are presumed to represent a more accurate determination of the paleopole position. The “STATUS” field of the GPMDB4.6b indicates if a paleopole has been superseded. Conversely, removing paleopoles published after 1983 (Pk 16,17; YEAR ≤ 1983) should have a negative effect.

**Exclude suspected local rotations and secondary overprints (Pk 18,19), or correct for them where possible (Pk 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. Pk 18,19 removes paleopoles that were identified as such in the COMMENTS field: 66 paleopoles affected by local rotations are identified by carefully going through all the COMMENTS containing information about rotation; paleopoles affected by secondary overprints are extracted with the keyword “econd” identified in the COMMENTS. A subset (19) of the 66 paleopoles identified have a suggested correction associated with them; Pk 20,21 retains these paleopoles after applying the suggested correction.

**SS05 quality criteria (Pk 22,23).** As with Pk 2,3, SS05 (Schettino and Scotese, 2005) removes paleopoles with high spatial ( $\alpha_{95} > 15^\circ$ ) 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  $< 2$ ). Pk 26,27 also use these criteria, but further excludes superseded data.

Some of the picking (Table 3.2) and weighting (Table 3.3) methods developed here are also connected with the V90 Q factor (see Section 3.1.2). For example, Pk 2,3 and Wt 2, 4, 5 are related to the V90 criteria 1; Pk 2,3, 22,23, 26,27 and Wt 1, 3 are related to the V90 criteria 2; Pk 22,23 are related to the V90 criteria 3; The data constraining described in Appendix B is related to the V90 criteria 5; and Pk 18,19 are related to the V90 criteria 7.

### 3.2.3.3 Weighting

Following filtering, weights were assigned to each of the remaining paleopoles using one of the following six algorithms (Table 3.3), prior to calculation of a weighted Fisher mean:

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

**Table 3.3:** List of all weighting algorithms developed here.

No.	Weighting Algorithm
0	None (No weighting)
1	Larger numbers of sites (B) & observations (N), greater weight ( <i>w</i> ):  $w = \begin{cases} 0.2 & , \text{ if both } B \text{ & } N \text{ are missing, or } B \leq 1 \text{ & } N \leq 1 \\ (1 - \frac{1}{B}) \times 0.5 & , \text{ if } N \text{ is missing or } N \leq 1, \text{ & } B > 1 \\ (1 - \frac{1}{N}) \times 0.5 & , \text{ if } B \text{ is missing or } B \leq 1, \text{ & } N > 1 \\ (1 - \frac{1}{B}) \times (1 - \frac{1}{N}) & , \text{ if } B > 1 \text{ & } N > 1 \end{cases}$
2	Lower age uncertainty, greater weight:

continued on next page

**Table 3.3 — continued from previous page**

No.	Weighting Algorithm
	<p>age_range=HIMAGAGE-LOMAGAGE      age_midpoint = (HIMAGAGE+LOMAGAGE)×0.5      if age_midpoint&lt;2.58 (Ma; start of the Quaternary, according to GSA Geologic Time Scale),</p> $w = \begin{cases} 1 & , \text{if } \text{age\_range} \leq 1.29 \text{ (from } \frac{2.58-0}{2}) \\ \frac{1.29}{\text{age\_range}} & , \text{if } \text{age\_range} > 1.29 \end{cases}$ <p>if <math>2.58 \leq \text{age\_midpoint} &lt; 23.03</math> (Ma; Neogene),</p> $w = \begin{cases} 1 & , \text{if } \text{age\_range} \leq 10.225 \text{ (from } \frac{23.03-2.58}{2}) \\ \frac{10.225}{\text{age\_range}} & , \text{if } \text{age\_range} > 10.225 \end{cases}$ <p>if <math>23.03 \leq \text{age\_midpoint} &lt; 201.3</math> (Ma; Paleogene, Cretaceous, Jurassic),</p> $w = \begin{cases} 1 & , \text{if } \text{age\_range} \leq 15 \\ \frac{15}{\text{age\_range}} & , \text{if } \text{age\_range} > 15 \end{cases}$
3	<p>Lower <math>\alpha_{95}</math>, greater weight:      Positive half Normal distribution with a mean and standard deviation of 0 and 10, scaled with <math>10\sqrt{2\pi}</math> (to make the peak reach 1)</p> $w = e^{-\frac{\alpha_{95}^2}{200}},$ <p>where</p> $\alpha_{95} = \begin{cases} ED95 & \\ DP & , \text{if } ED95 \text{ is missing} \\ \frac{140}{\sqrt{KD \times N}} & , \text{if } ED95 \& DP \text{ are missing} \\ \frac{140}{\sqrt{K\_NORM \times N}} & , \text{if } ED95, DP \& KD \text{ are missing} \\ \frac{140}{\sqrt{K\_NORM \times B}} & , \text{if } ED95, DP, KD \& N \text{ are missing} \\ \frac{140}{\sqrt{1.7 \times B}} & , \text{if } ED95, DP, KD \& K\_NORM \text{ are missing,} \\ & \text{using the lowest KD value } \sim 1.7 \text{ in GPMDB4.6b,} \end{cases}$ <p>finally <math>w=0</math> if this <math>\alpha_{95}</math> completely overlaps with another smaller <math>\alpha_{95}</math> whose paleopole is exactly derived from the same place and same rock.</p>
4	<p>Age uncertainty Position to bin (more overlap, greater weight):      wha, window high age; wla, window low age</p> $w = \begin{cases} \frac{wha - LOMAGAGE}{age\_range} & , \text{if } LOMAGAGE > wla \& HIMAGAGE > wha \\ \frac{HIMAGAGE - wla}{age\_range} & , \text{if } LOMAGAGE < wla \& HIMAGAGE < wha \\ \frac{wha - wla}{age\_range} & , \text{if } LOMAGAGE < wla \& HIMAGAGE > wha \\ 1 & , \text{if } LOMAGAGE > wla \& HIMAGAGE < wha \end{cases}$
5	Combining 3 and 4: average of the two weights from 3 and 4

**Weighting by sample and site number (Wt 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; Besse and Courtillot, 2002; van der Voo, 1990), and are given a weighting closer to 1. Unfortunately, in the GPMDB, a paleopole's B or N is not always given. As shown in (Table 3.3 no. 1), in such cases the calculated weights were modified to give lower weights overall whilst still accounting for partial information, such as N being reported but not B.

For number of sampling sites B>1 and number of samples N≤1, there are 8 such paleopoles for 120–0 Ma North America, India 4, and Australia 1. For B≤1 and N>1, there are 20 such paleopoles for 120–0 Ma North America, India 26, and Australia 22. For B≤1 and N≤1, there are only 23 such paleopoles for the whole GPMDB 4.6b, including 18 with neither B value nor N value given; while for 120–0 Ma there is no such paleopole found in North America, India and Australia.

**Weighting by age uncertainty (Wt 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; here the geological period that the middle point of the paleopole's age range falls within is assigned) (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 (No. 2 in Table 3.3).

**Weighting by spatial uncertainty (Wt 3).** Paleopoles with a smaller  $\alpha_{95}$  confidence ellipse are given a higher weighting than those with a larger  $\alpha_{95}$ , using a Gaussian/Normal distribution centered on 0 with standard deviation of 10. However, a paleopole's  $\alpha_{95}$  is not always given in the database. If  $\alpha_{95}$  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 not given either,  $\alpha_{95}$  was further approximated by  $\frac{140}{\sqrt{KD \times N}}$ , 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 specimens, where two paleopoles are derived, are exactly from the same place and same rock (by checking if ROCKNAME, ROCKTYPE

and sampling site are the same), 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 paleopole.

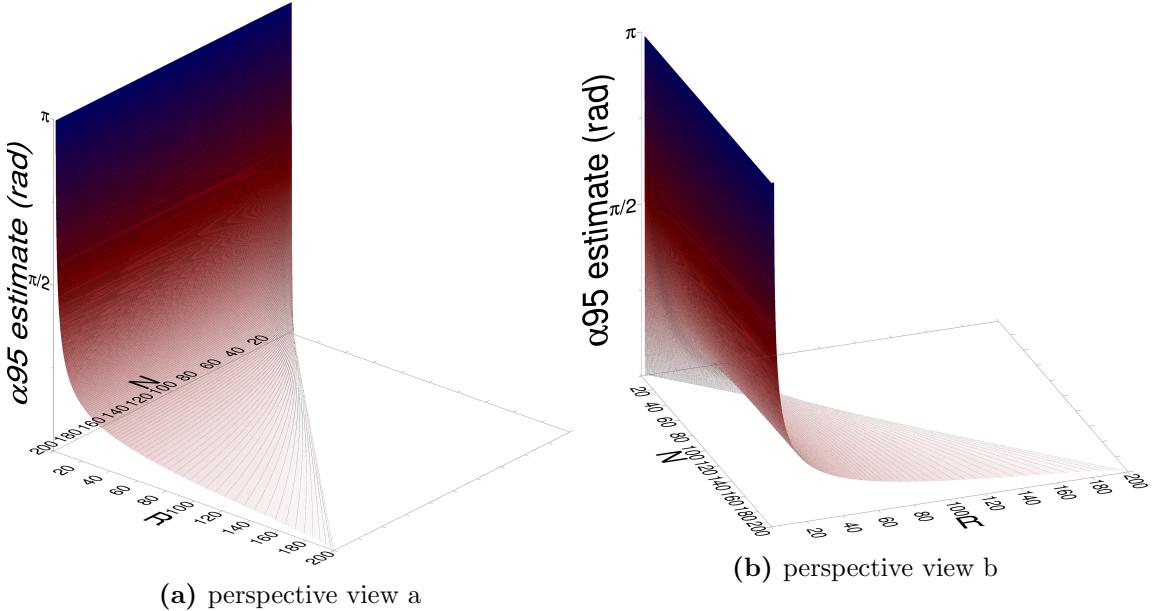
**Weighting by degree of overlap between moving window and pole age uncertainty (Wt 4).** If a large fraction of the age range for an individual paleopole falls within the current window, it is given a higher weighting than a pole where the overlap is smaller, because it is more likely to be close to 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 (Wt 5).** This weighting method is a combination of Wt 3 (but here the standard deviation of the weighting function is 15 though) and Wt 4. It takes the average of sums of the weights generated by Wt 3 and 4.

### 3.2.3.4 Scaling of Weights

The weights obtained from the six different weighting functions (Table 3.3) are then integrated into Fisher mean function (Fisher, 1953) to calculate a weighted Fisher mean. First, weight values are used to scale Cartesian x, y and z components of each individual paleopole's geographic coordinate. Then these individual coordinates are combined through Fisher resultant vector  $R$  function (see Tauxe et al., 2019, chap. 11). Therefore the mean pole location, its spatial uncertainty A95, and precision parameter are all weighted along with  $R$ .

Traditionally, weights are directly integrated by being multiplied with the variable we would like to do weighting to. For example, here weights can be directly multiplied with the Cartesian x, y and z components of each paleopole. However, this direct multiplying causes the decreasing of  $R$ , which further sensitively and extremely increases  $\alpha_{95}$  (Fig. 3.5), especially because  $R$  is always less than  $N$  and  $N$  is usually not that high (more than  $\sim 50$  is rather rare, around 10 averagely). The consequence would be that all the  $\alpha_{95}$ s of mean poles are extremely large in size and difficult to be spatially differentiated.



**Figure 3.5:** Visualization of Equation 11.9 of Essentials of Paleomagnetism: Fifth Web Edition, illustrating the relationship between the radius of the circle of 95% confidence ( $p=0.05$ ) about the mean,  $\alpha_{95}$ , resultant vector  $R$  and number of directions (or paleopoles)  $N$ .

Therefore, here weights are scaled before being multiplied with the Cartesian x, y and z components by

$$ScaledWeight = \frac{Weight_i * N}{\sum_i^N Weight_i},$$

where  $N$  is number of paleopoles for making a mean pole. So the scaled weight could be greater than 1 because it is actually scaled through being divided by the mean of the weights. This scaling does not only keep the effect of weighting but also avoid dramatically changing  $R$  and indirectly and extremely changing  $\alpha_{95}$ .

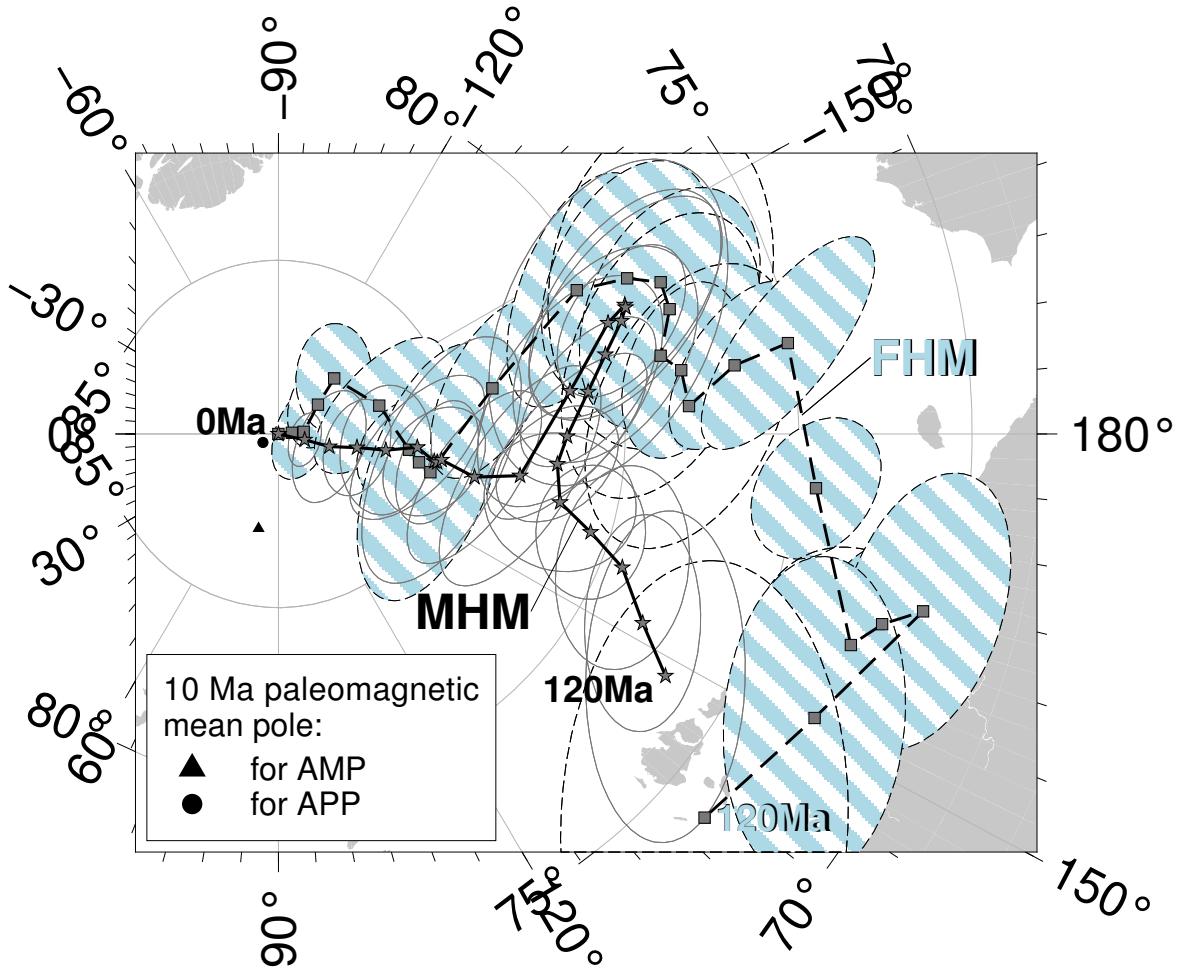
### 3.2.4 Reference Paths

A prediction of the expected APWP for any plate can be generated using a plate kinematic model (e.g. the last  $\sim 180$ –200 Myr of 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 Nubian Plate relative to the Indo-Atlantic hotspots back to 120 Ma. North America is linked to this absolute frame of reference across the Mid-Atlantic ridge, using North America-Nubia rotations from chron C1n (0–0.78 Ma) from DeMets et al. (2010), to chron C2An (2.7 Ma) from Shephard et al. (2012), to C5n.1ny (9.74 Ma) from Müller et al. (1999), to C5n.2o (10.949 Ma) from Gaina et al. (2013), to C6ny (19.05 Ma) from Müller et al.

(1999), to C6no (20.131 Ma) from Gaina et al. (2013), to C34ny (83.5 Ma) from Müller et al. (1999), from C34ny to  $\sim$ 118.1 Ma from Seton et al. (2012), and to closure at C34no (120.6 Ma) from Gaina et al. (2013). India is linked via the East African Rift Valley (Somalia-Nubia rotations from chron C1n (0–0.78 Ma) from DeMets et al. (2017), to C2A.2no (3.22 Ma) from Horner-Johnson et al. (2005), and to closure at C7.2m (25.01 Ma) and chron C34 (85 Ma) from Rowan and Rowley (2016), 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)). Australia is linked via the East African Rift Valley, then the SW Indian Ridge (E Antarctica-Somalia rotations from chron C1n (0–0.78 Ma) from DeMets et al. (2017), to C2A.2no (3.22 Ma) from Horner-Johnson et al. (2005), to C5n.2no (10.95 Ma) from Lemaux et al. (2002), to C13ny (33.06 Ma) from Patriat et al. (2008), to C29no (64.75 Ma) from Cande et al. (2010), to C34y (83 Ma) from Rowan and Rowley (2016), to 96 Ma from Marks and Tikku (2001), and to closure at chron M0 (120.6 Ma) from Müller et al. (2008)), and SE Indian Ridge (Australia-East Antarctica rotations from chron C1n (0–0.78 Ma) from DeMets et al. (2017), to C6no (20.13 Ma) from Cande and Stock (2004), to C8o (26 Ma) from Granot and Dymant (2018), to C17n.3no (38.11 Ma) from Cande and Stock (2004), to C34ny (83.5 Ma) from Whittaker et al. (2013), to the Quiet Zone Boundary (96 Ma) from Wessel et al. (2007), to full closure at 136 Ma from Whittaker et al. (2013)). The geomagnetic polarity timescales of Cande and Kent (1995) for Late Cretaceous and Cenozoic, and of Gradstein et al. (1994) for Early to Late Cretaceous time are used to convert from chron boundaries to absolute ages. A long table listing these rotation parameters with covariance uncertainties (Table B.1) is included in Appendix B, and the calculated rotations for the North American, Indian and Australian reference APWPs in the hotspot reference (Figs. 3.6-3.8) are also listed in Table B.2.

Where possible, poles which were published uncertainty estimates were used. Where no uncertainty estimates were available, values of the covariance matrix were set to an arbitrarily small value (1E–15). Where this occurs, the spatial uncertainties for the reference APWP are likely underestimated. However, but because the uncertainties for the Nubia-hotspot rotations are substantially larger than for rotations derived from fitting magnetic isochrons, the effect is small.

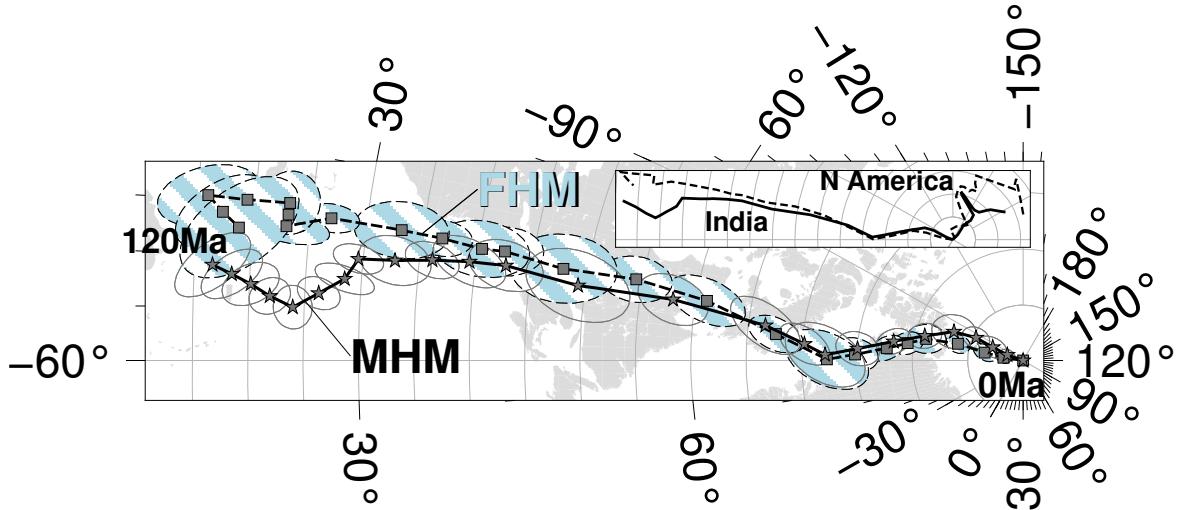
To reconstruct a reference APWP at the required time steps for comparison with the paleomagnetic APWPs, rotations and their associated uncertainties were interpolated between constraining finite rotation poles according to the method of Doubrovine and Tarduno (2008), assuming constant rates.



**Figure 3.6:** 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. 3.4.

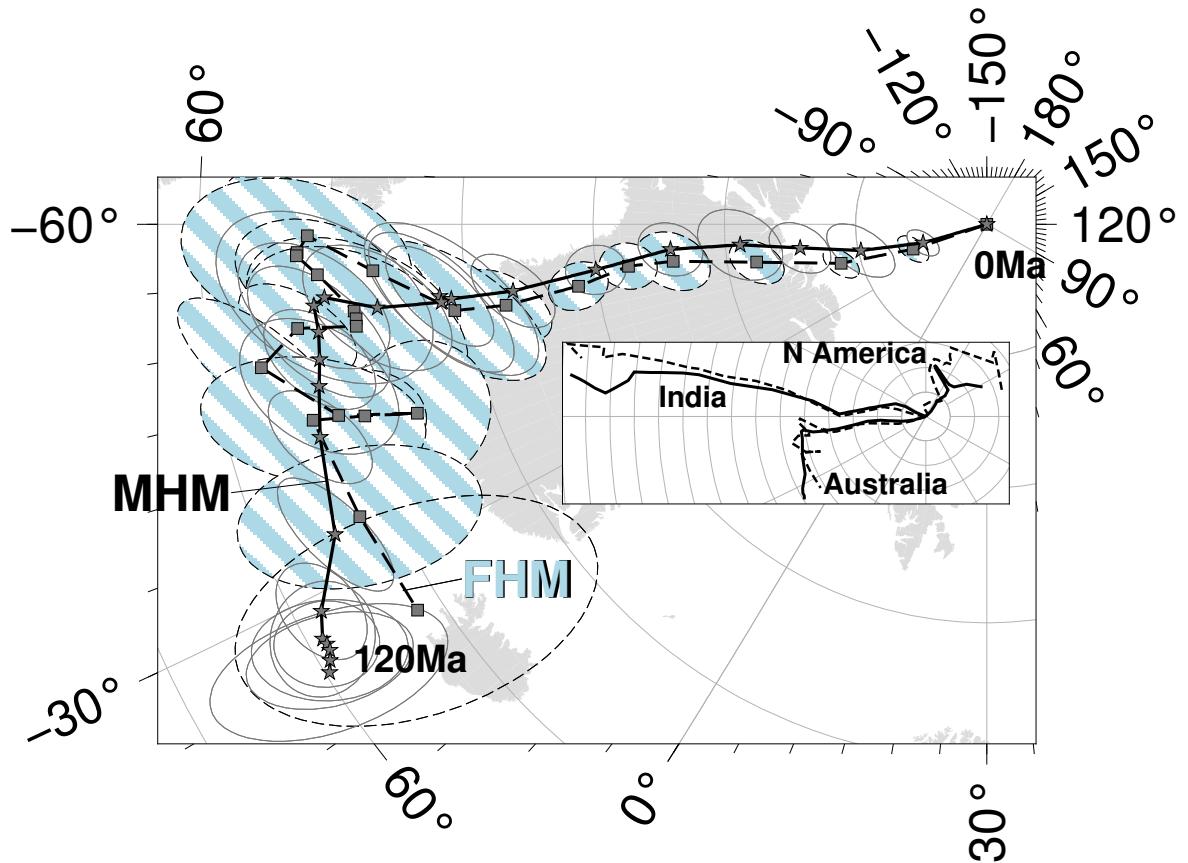
Neither the hotspot reference frame nor the paleomagnetic reference frame are truly fixed with respect to the solid Earth. In the former case, hotspots are not truly stationary in the mantle (Steinberger and O’Connell, 1998); in the latter, true polar wander (TPW) may also lead to differential movements 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  $\sim$ 120 Myr (Cottrell and 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 uncertain-

ties, 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. 3.6), India (Fig. 3.7) and Australia (Fig. 3.8). 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.



**Figure 3.7:** 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. 3.6.

When comparing the synthetic APW paths for the three plates (inset, Fig. 3.8), there are clear differences. The predicted mean north pole for North America at 120 Myr is still at  $\sim 75^\circ\text{N}$  (Fig. 3.6), 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  $\sim 50$ – $55$  Ma (Najman et al., 2010), is reflected by the 120 Ma predicted mean north pole being located at  $\sim 20^\circ\text{N}$  (Fig. 3.7). Australia represents an intermediate case, with north westerly plate motion from  $\sim 120$ – $60$  Myr changing to more rapid northward motion from  $\sim 60$ – $55$  Ma to the present (Wessel et al., 2007). When comparing the FHM and MHM tracks, significant differences in the oldest parts (before  $\sim 80$  Ma) are apparent for India and North America.



**Figure 3.8:** 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. 3.7, much slower moving North America shown in Fig. 3.6, and also relatively intermediate moving Australia.

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.

### 3.2.5 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.,  $W_s = W_l = W_a = \frac{1}{3}$ ). A lower  $\mathcal{CPD}$  score generally indicates a ‘better’ fit, although a lower score can also potentially result from comparison to a more poorly constrained path with large uncertainties, which are less likely to be significantly different. An additional ‘Fit Quality’ (FQ)

metric can help to distinguish such cases, by assigning the two paths being compared a letter score based on the average size of their uncertainty ellipses (Chapter 2.2.6). Here, the first letter refers to the generated APWP, and the second letter refers to the reference path. In this study, the reference path FQ score is fixed for each plate; because the spatial uncertainties for paths generated from plate motion models are small compared to those typical for paleomagnetic data, this reference path FQ is always rated ‘A’.

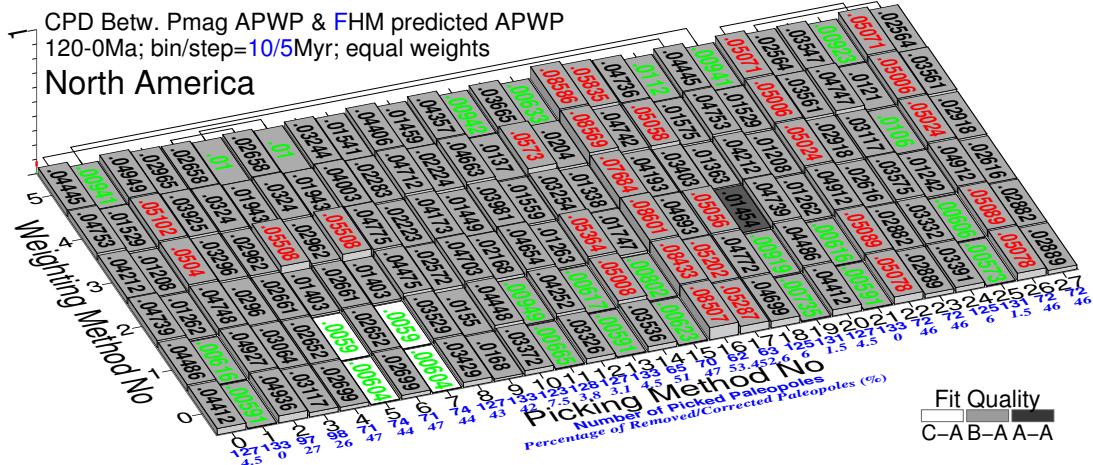
This does not only help find the most similar paleomagnetic APWP (from the best algorithm) to the reference APWP, but also help further test and demonstrate the validity of the similarity measuring tool in practise.

## 3.3 Results

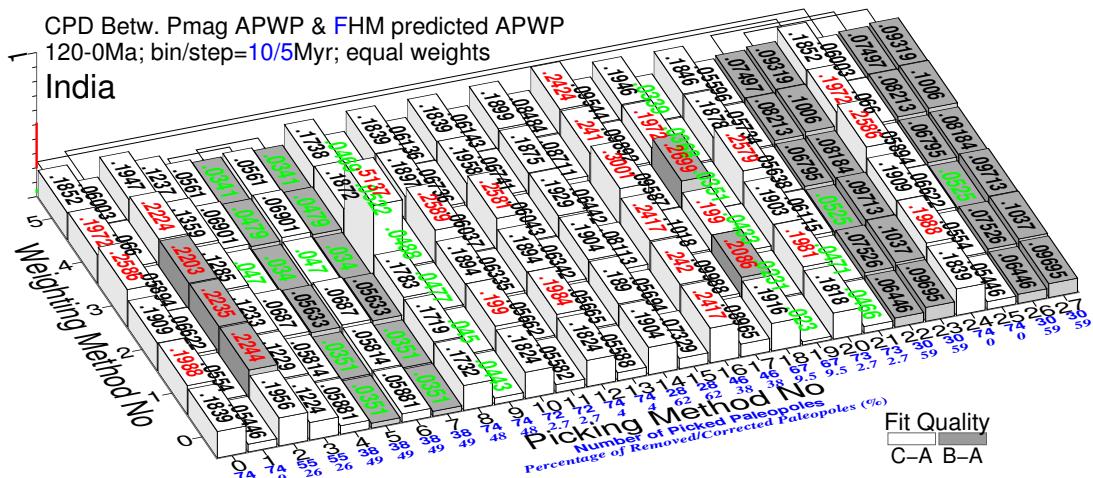
### 3.3.1 Baseline Results: 10/5 Myr Window/Step, Fixed Hotspot Reference

Fig. 3.9 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 (squares and dashed line in Figs. 3.6-3.8) for North America (Fig. 3.9a), India (Fig. 3.9b) and Australia (Fig. 3.9c). 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 0.0023 to 0.5137. The fits for paths with low difference scores are clearly much better than for those with high ones (Fig. 3.10). From Fig. 3.9, it is clear that:

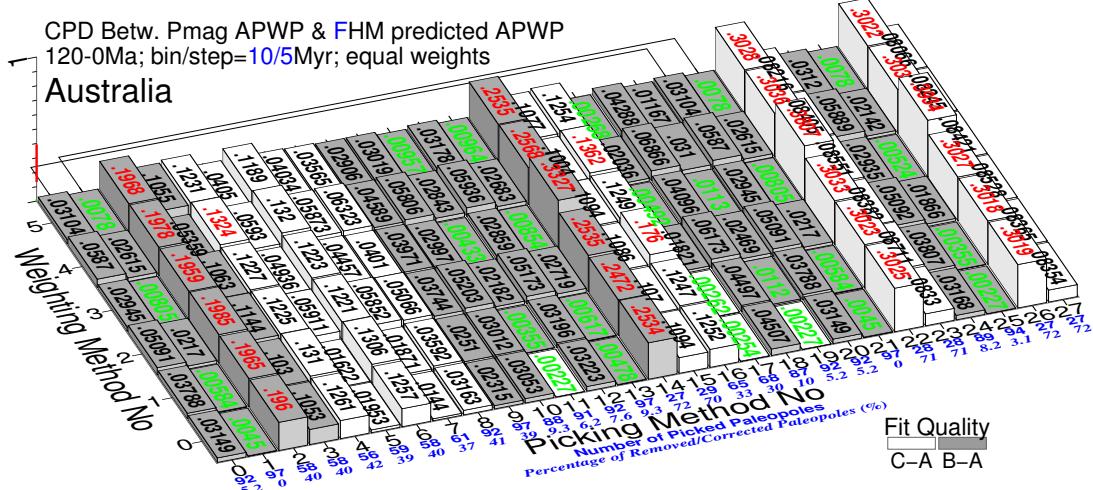
1. There is much more variation in scores along the horizontal axes than the vertical axes (see self-explanatory topography of bands in Fig. 3.9), suggesting that the choice of windowing and filtering method (Table 3.2) has a much greater impact than weighting (Table 3.3).
2. Many of the highest scores (worst fits) are associated with even-numbered picking and filtering methods, i.e., those which use the AMP windowing algorithm. Even so, Pk 4 and 6 are among the best methods for India.
3. 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



(a) North America: minimum 0.00573 (25(0)), maximum 0.08601 (16(2)), mean 0.032403, median 0.032395

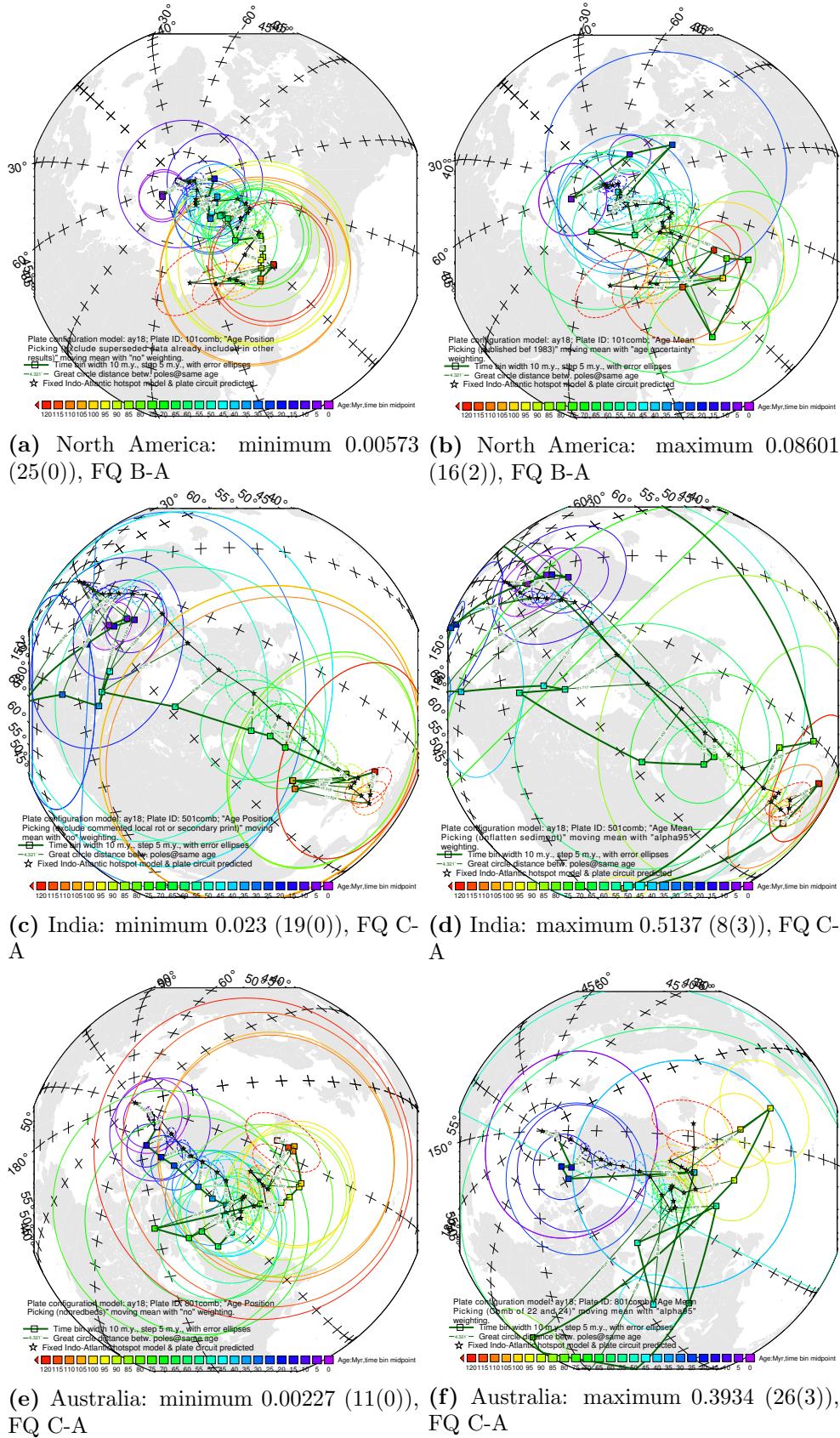


(b) India: minimum 0.023 (19(0)), maximum 0.5137 (8(3)), mean 0.1182, median 0.0835



(c) Australia: minimum 0.00227 (11(0)), maximum 0.3934 (26(3)), mean 0.08373, median 0.05

**Figure 3.9:** 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 Pk 1, corrected relative to each corresponding picking method (Pk 8,9, 12,13; 1 paleopole removed and 1 corrected by Pk 20,21 for India). Fit quality (FQ) for each score is color coded.



**Figure 3.10:** Path comparisons with best and worst difference values shown in Fig. 3.9. The parenthetical remarks are Pk No and Wt No.

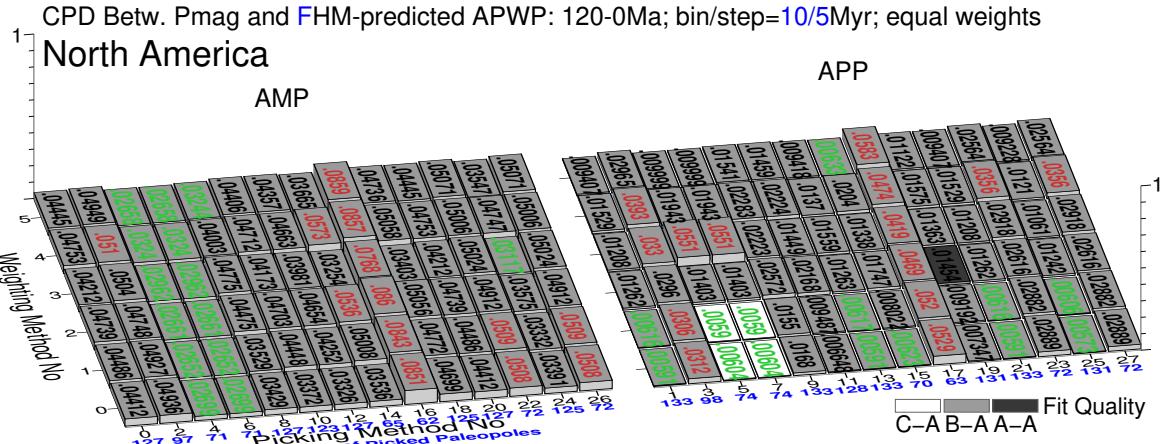
(Fig. 3.9a), and the Indian plate having the highest (Fig. 3.9b). The scores for the Australian plate are generally closer to the equivalent scores for the North American plate than to the scores for the Indian plate (Fig. 3.13).

4. Although there is some overlap (e.g., Pk 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, 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.
5. Relating to FQ ratings: North America's lower  $\mathcal{CPD}$  scores are also associated with consistently good FQ B-A grades; India's higher  $\mathcal{CPD}$  scores are also associated with lower FQ (mostly C-A, including the no-filter methods, so there are some improvements in FQ with some filters), and Australia is a mix of B-A (including the no-filter methods) and C-A (which represent a reduction in FQ with some filters). In general, lower  $\mathcal{CPD}$  scores also appear to be associated with better FQ (e.g. B-A rather than C-A).

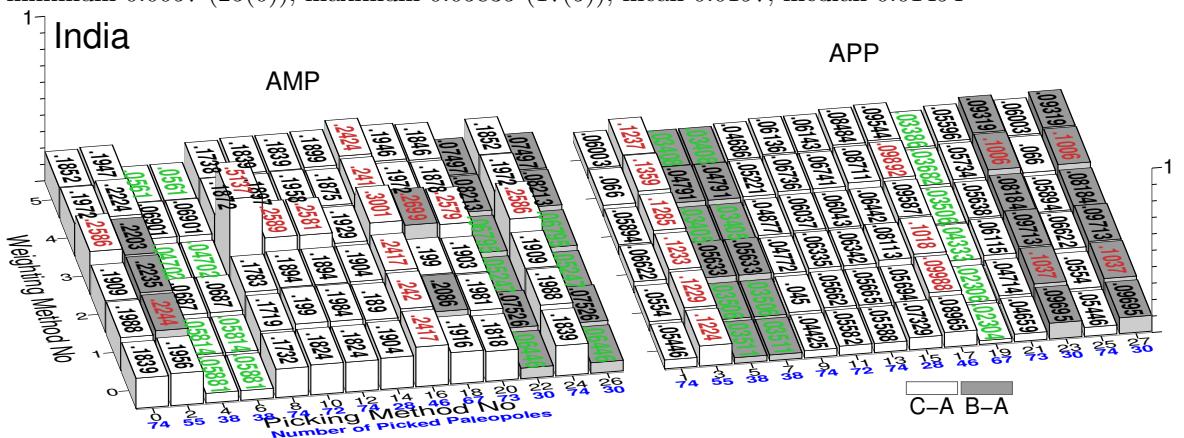
### 3.3.2 Effects of Windowing Method

Dividing  $\mathcal{CPD}$  scores according to whether the AMP or APP windowing method was used (Fig. 3.11) 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. 3.12). 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. 3.11b, 3.11c, 3.12b, 3.12c), and more than 1 standard deviation greater than the APP mean for North American APWPs (Figs. 3.11a, 3.12a).

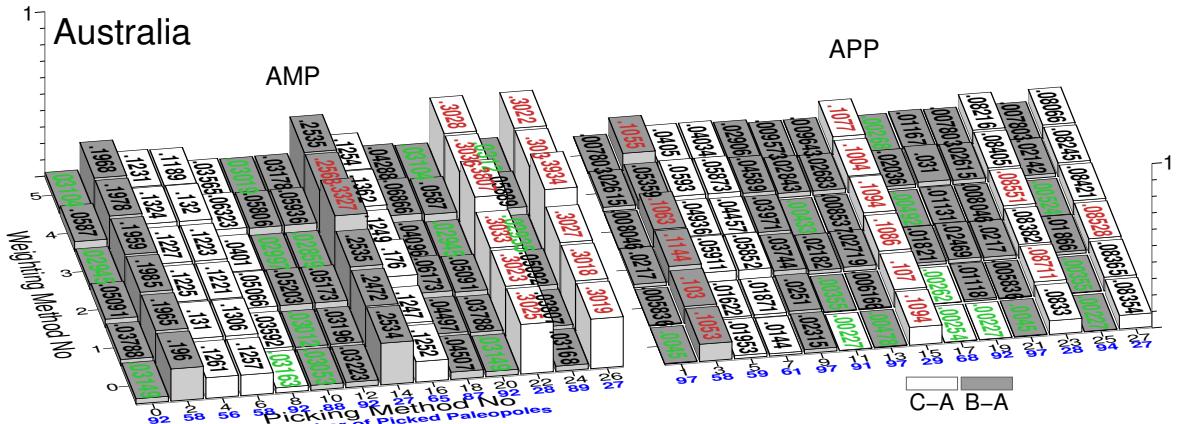
For each of the 84 possible combinations of filter method and weighting, the AMP-derived score is typically 3–5 times higher than the equivalent APP-derived score (Fig. 3.12, insets). APP-generated paths yield a lower  $\mathcal{CPD}$  score than the equivalent AMP-generated path for 82 (97.6%) of the 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 (Pk 4,5 and 6,7) or removes poles with large various uncertainties (Pk 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. 3.11a, 3.11b). In the former case, the scores are comparable and relatively low; in the latter case,



(a) AMP: minimum 0.0265 (4(1)), maximum 0.086 (16(2)), mean 0.0451, median 0.0457; APP: minimum 0.0057 (25(0)), maximum 0.05835 (17(5)), mean 0.0197, median 0.01494



(b) AMP: minimum 0.047 (4(3)), maximum 0.5137 (8(3)), mean 0.1688, median 0.1892; APP: minimum 0.023 (19(0)), maximum 0.1359 (3(4)), mean 0.0676, median 0.0602



(c) AMP: minimum 0.0286 (3(0)), maximum 0.3934 (26(3)), mean 0.12675, median 0.0938; APP: minimum 0.00227 (11(0)), maximum 0.11445 (3(2)), mean 0.0407, median 0.0256

**Figure 3.11:** Separated results from AMP and APP in Fig. 3.9. 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.

they are comparable but relatively high. For the Australian plate, only correcting for sedimentary inclination shallowing (Pk 8,9) produces comparable but moderate scores (Fig. 3.11c).

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 Wt 0 (no weighting) and 1 (by sample and site number) have been applied (Fig. 3.11a). This pattern is also seen for the Australian plate, but removal of poles published after 1983 (Pk 16,17) also results in significantly better performance of the APP method (Fig. 3.11c). For the Indian plate, the largest difference occurs when poles with sedimentary inclination shallowing (Pk 8,9) are corrected, or suspected overprints or local rotations are removed (Pk 18,19, Fig. 3.11b).

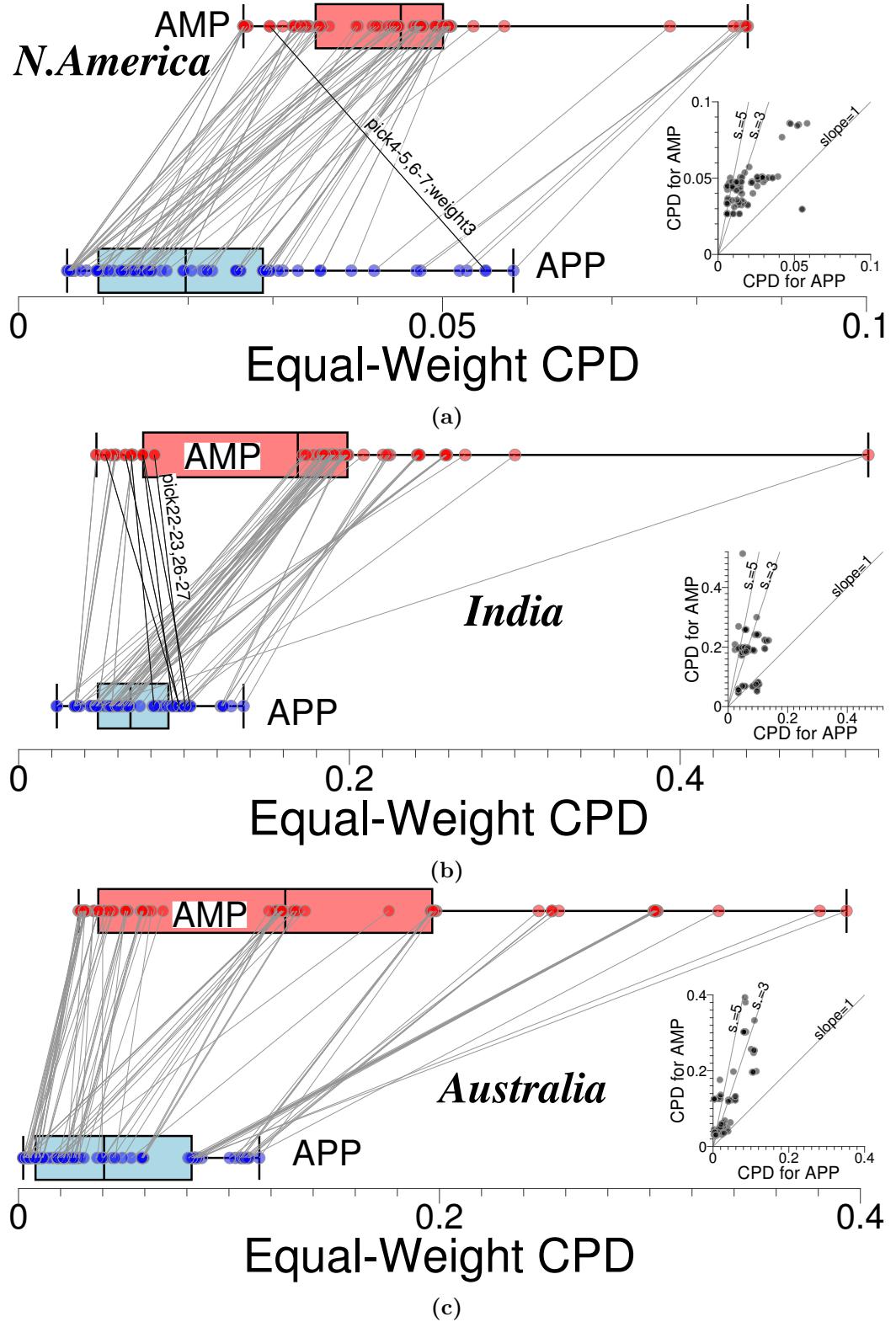
### 3.3.3 Effects of Filtering and Weighting

When  $\mathcal{CPD}$  scores are separated by windowing method (Fig. 3.11), 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 (Pk 0,1) and no weighting (Wt 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.

#### 3.3.3.1 Filter Aggression

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

1. removal of sedimentary paleopoles (Pk 4,5 and 6,7), which removes ~40–50% of the datasets on all 3 plates, with the highest proportion being removed on the Indian plate. Although Pk 4,5 are more strict, it does not remove many more paleopoles than Pk 6,7. For North America and India, the numbers of filtered paleopoles for Pk 4,5 and 6,7 are actually the same.



**Figure 3.12:** Box-and-whisker and cross (inset) plots of Fig. 3.11. 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  $\mathcal{CPD}$ .

2. correction of sedimentary paleopoles for inclination flattening (filter in Pk 8,9), which affects 38–48% of the dataset, with the highest affected proportion on the Indian plate.
3. removal of paleopoles with large temporal and spatial uncertainty (Pk 2,3, 22,23), particularly for the Australian plate, where the SS05 filtering criteria removes ~70% of the paleopoles. Filter in Pk 26,27 combines Pk 22,23 and 24,25, but no or very few (in the only case of Australia only 1 additional paleopole) are actually removed.
4. filtering based on publication date (Pk 14,15 and 16,17), with the ratio of pre/post 1983 poles varying from about 50/50 on the North American plate to about 70/30 on the Australian plate.

Conversely, filtering or correction for redbeds (Pk 10,11 and 12,13), local rotations and overprints (Pk 18,19 and 20,21), or superseded data (Pk 24,25) affected 4% or less of the paleopoles on any plate.

Note that there is an important trade-off that precision in a mean pole is determined by the number of contributing paleopoles ( $A_{95}$  is proportional to  $1/\sqrt{n}$ ), so an overly aggressive filter might improve accuracy at the expense of precision. This is particularly an issue where data density is low (see Section 1.1.4.4).

### 3.3.3.2 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:

1. For all 3 plates, higher  $\mathcal{CPD}$  scores are commonly associated with filtering based on  $\alpha_{95}$  and age range (Pk 2,3, 22,23, 26,27), with the exception of AMP-derived paths for India, where Pk 22 and 26 produce some of the lowest scores, and also improve FQ.
2. For North America and India, low scores are commonly associated with removal of non-igneous poles (Pk 4,5 and 6,7), particularly for AMP-derived paths. On the Australian plate, these filters are less effective, and also reduce FQ.
3. For North America and Australia, correction for inclination flattening generates  $\mathcal{CPD}$  scores very similar to scores with no filtering for AMP-derived paths (Pk 8 versus Pk 0; for Australia FQ reduced), and increases scores for APP-derived

paths (Pk 9 versus Pk 1). In contrast, for India generally there is a small decrease in  $\mathcal{CPD}$  scores compared to no filtering for both AMP- and APP-derived paths.

4. Many of the highest difference scores for North America and India occur when paleopoles published after 1983 are removed (Pk 16,17), whilst removing paleopoles published before 1983 (Pk 14,15) generates  $\mathcal{CPD}$  scores comparable to scores with no filtering (relatively much lower). In contrast, for Australia Pk 14,15 produces some of the highest  $\mathcal{CPD}$  scores, and Pk 16,17 have little effect on the  $\mathcal{CPD}$  score, although FQ is commonly reduced. It is noteworthy that the number of older studies (65/68; Fig. 3.9c) is almost 2.5 times of the number of newer studies (27/29) for the Australian plate.

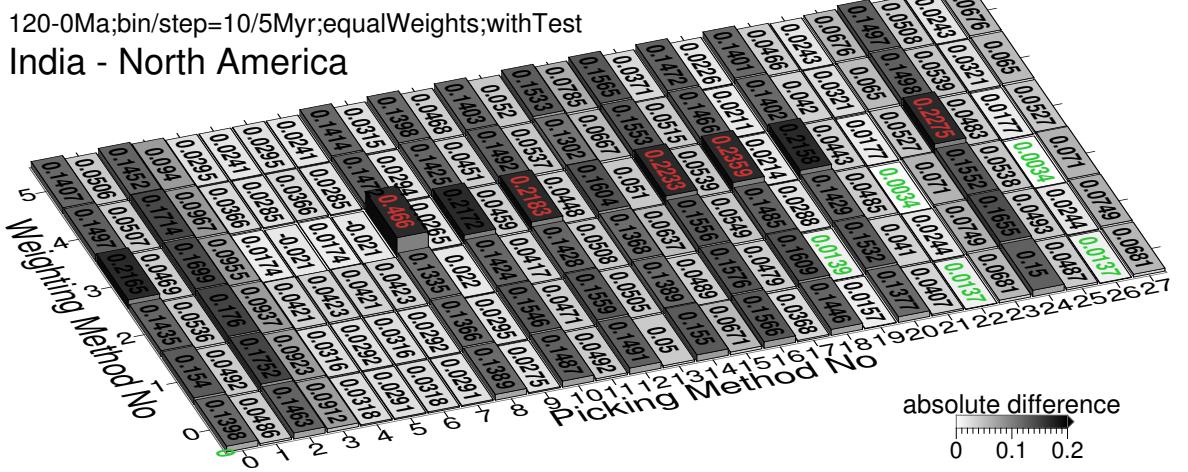
Whilst it is generally true that methods with low-aggression filters do not generate scores that differ much from the no-filtering scores, there are some exceptions:

1. For North America and Australia, removing superseded paleopoles (Pk 24,25) produces lower  $\mathcal{CPD}$  scores.
2. Removing paleopoles suspected to be affected by overprints or local rotations (Pk 18,19) consistently produces lower  $\mathcal{CPD}$  scores for Indian APP-derived paths.

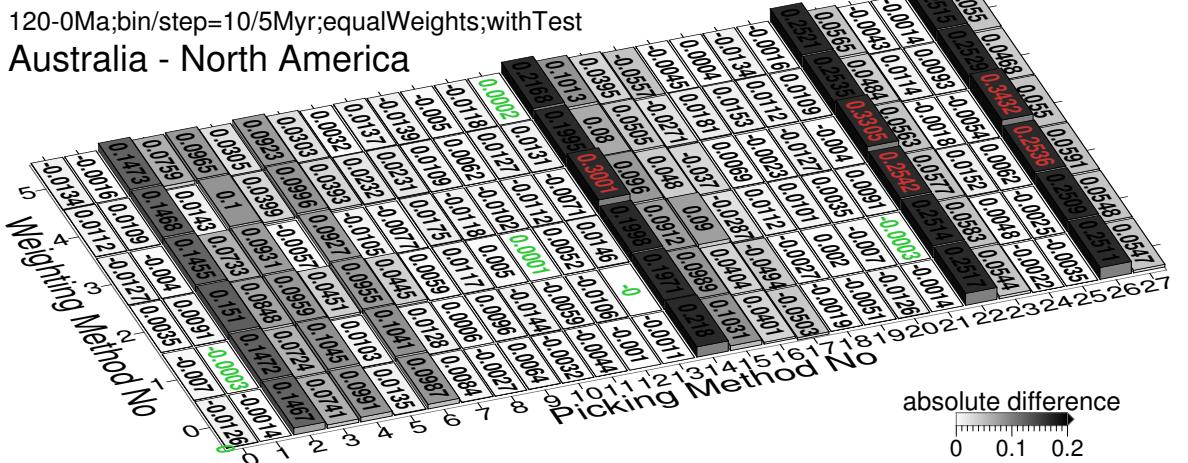
### 3.3.3.3 Weighting Performance

Compared to the variations resulting from different windowing methods and filters, Figs. 3.11, 3.14, 3.15 and 3.16 indicate that the effect of weighting the data prior to calculating a Fisher mean is generally small or diluted. Where an effect can be seen, it is negative, generating larger  $\mathcal{CPD}$  scores.

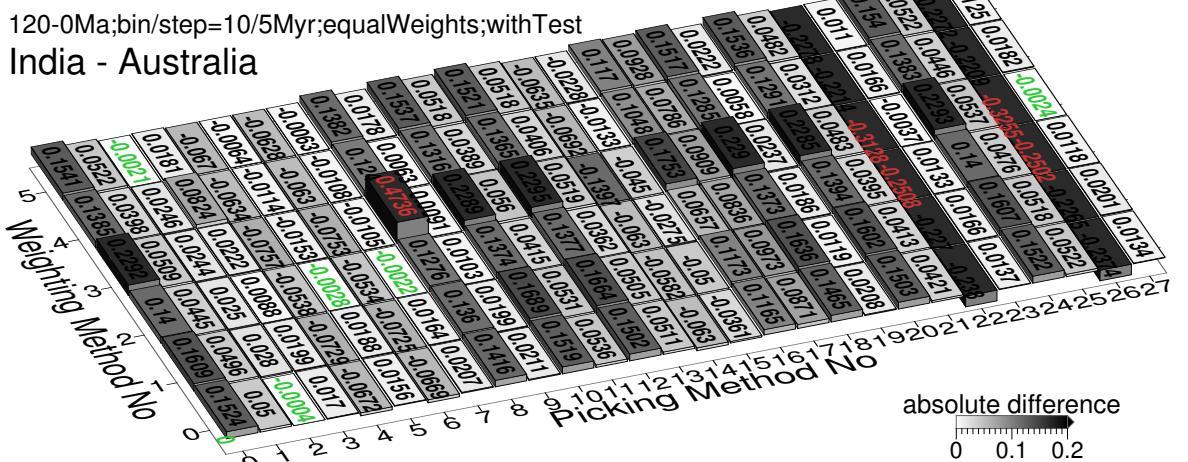
1. For APP-derived paths from the North American and Australian plates, Wt 0 (no weighting) and 1 (weighting by sample and site number) usually produce slightly better  $\mathcal{CPD}$  scores than other weighting methods.
2. Wt 3 (weighting by spatial uncertainty) seems most likely to generate much higher  $\mathcal{CPD}$  scores, particularly for AMP-derived paths, and particular for the Indian plate.
3. Wt 0, 1 and 5 generally produce lower similarity scores than Wt 2, 3 and 4.



(a) Differences between Fig. 3.9b (India) and Fig. 3.9a (North America)

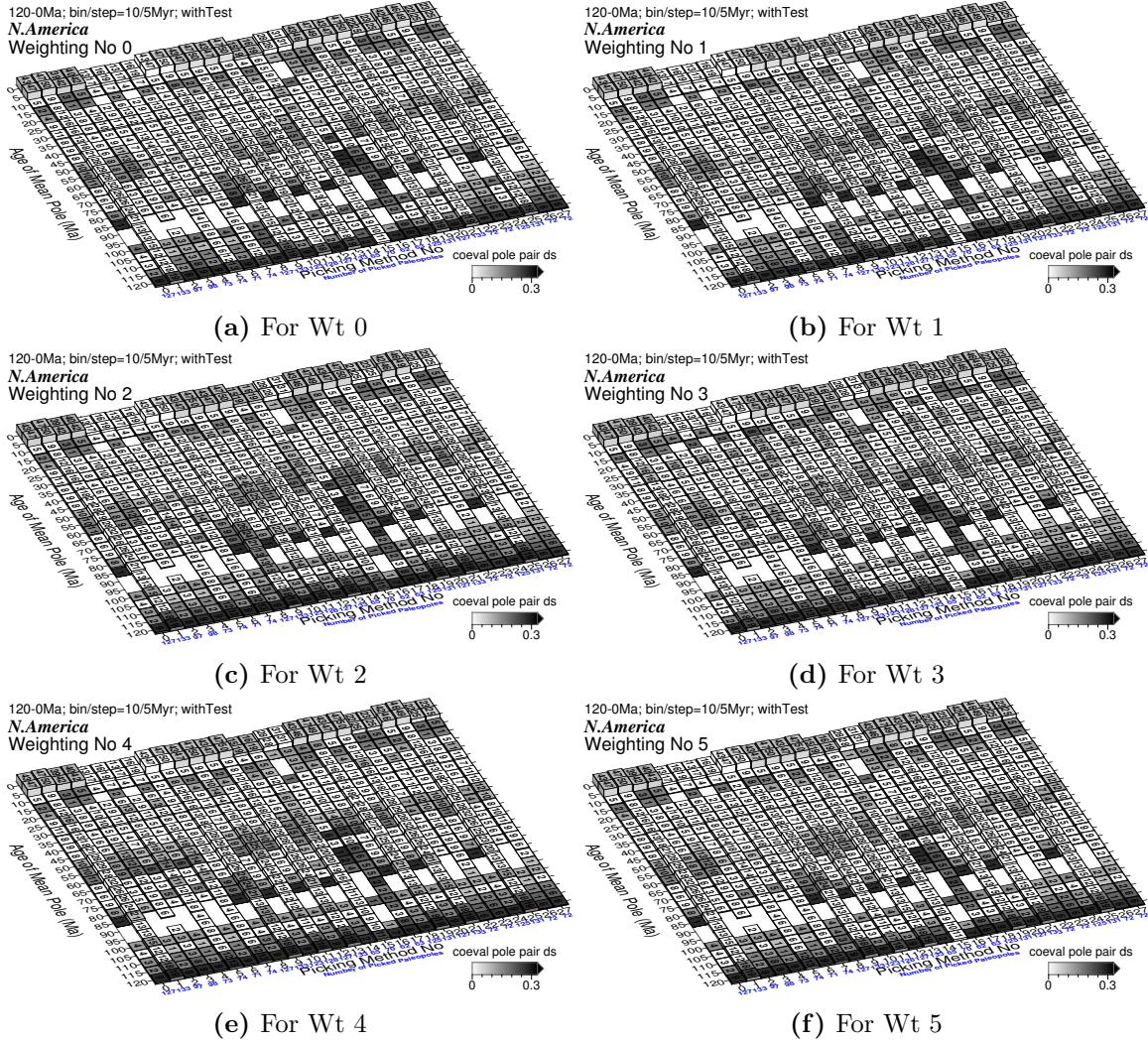


(b) Differences between Fig. 3.9c (Australia) and Fig. 3.9a (North America)



(c) Differences between Fig. 3.9b (India) and Fig. 3.9c (Australia)

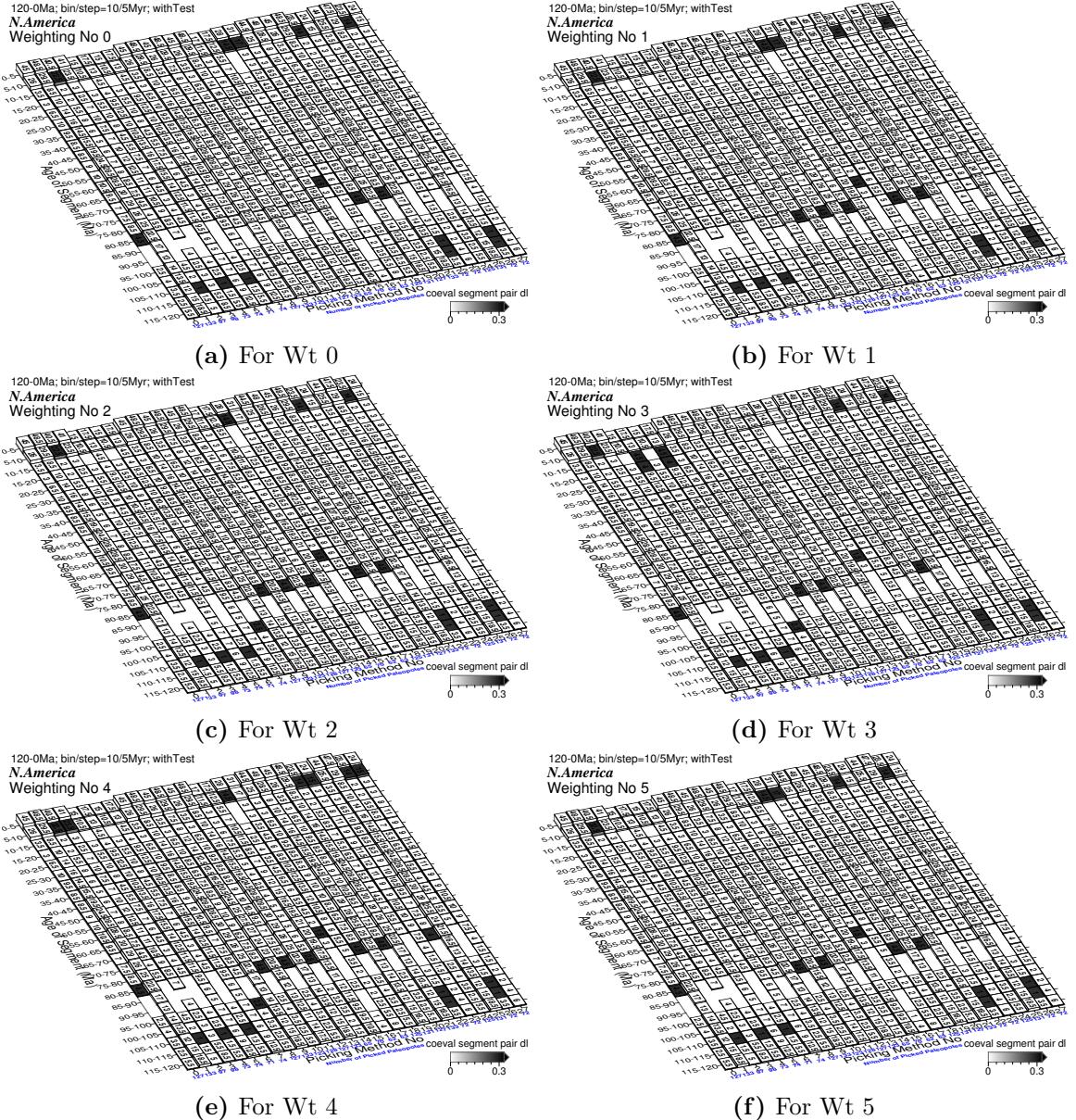
**Figure 3.13:** Differences between grids in Fig. 3.9. The absolute difference values less than 1.96-standard-deviation interval of the whole 168 values are labeled in green, more than 1.96-standard-deviation interval labeled in red.



**Figure 3.14:** Tested spatial difference ( $d_s$ ) 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 number labels on the grids (including grid heights) are the numbers of site mean poles that are contributing to make each mean path pole.

### 3.3.4 Effects of Window Size

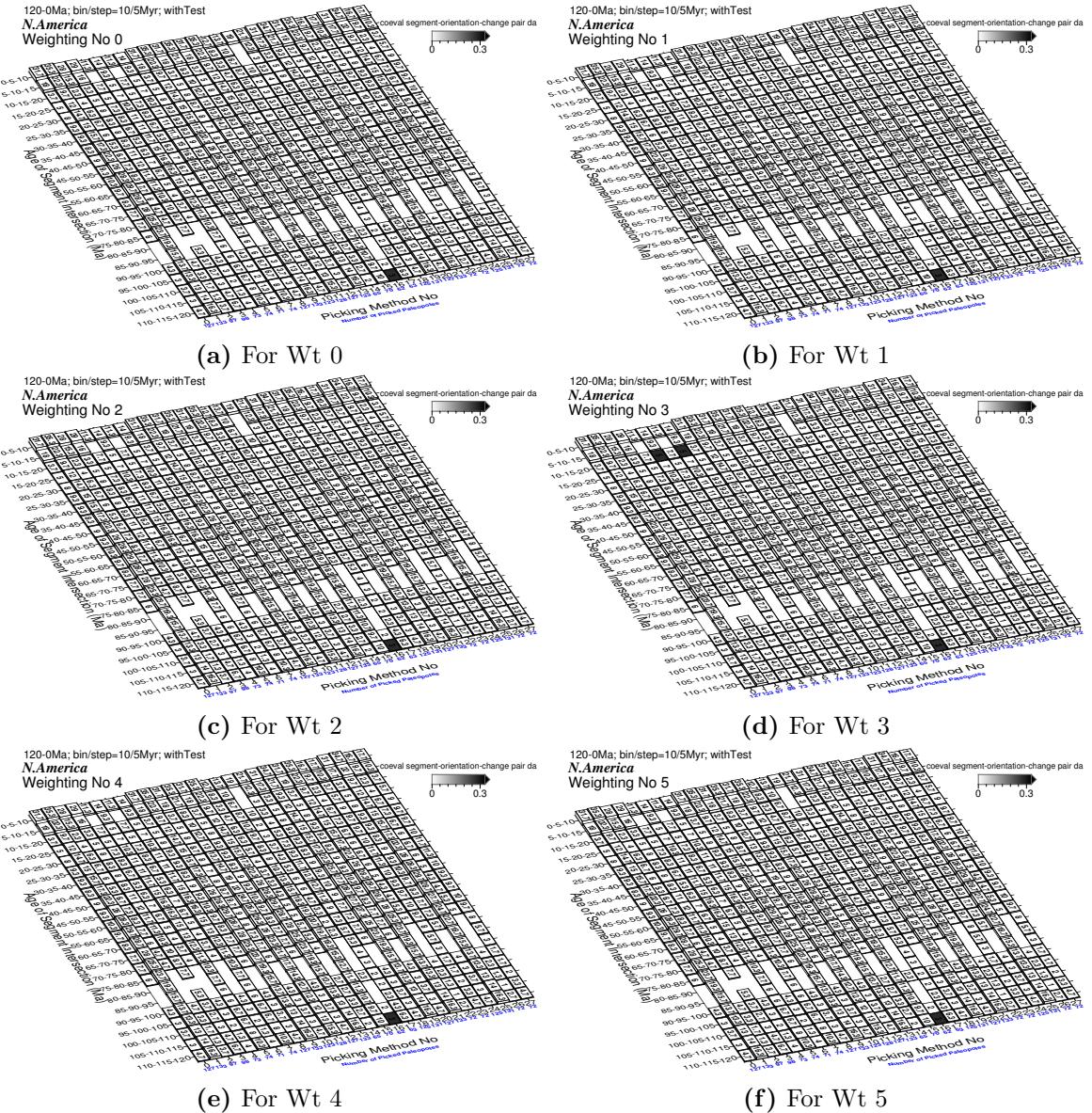
Fig. 3.17 shows the  $\mathcal{CPD}$  scores for the APWPs generated with the 28 picking and 6 weighting methods, compared to the FHM reference paths, with the picking time window width increased from 10 to 20 Myr, and the window step increased from 5 to 10 Myr.



**Figure 3.15:** Tested length difference ( $d_l$ ) 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.

### 3.3.4.1 Overall Change

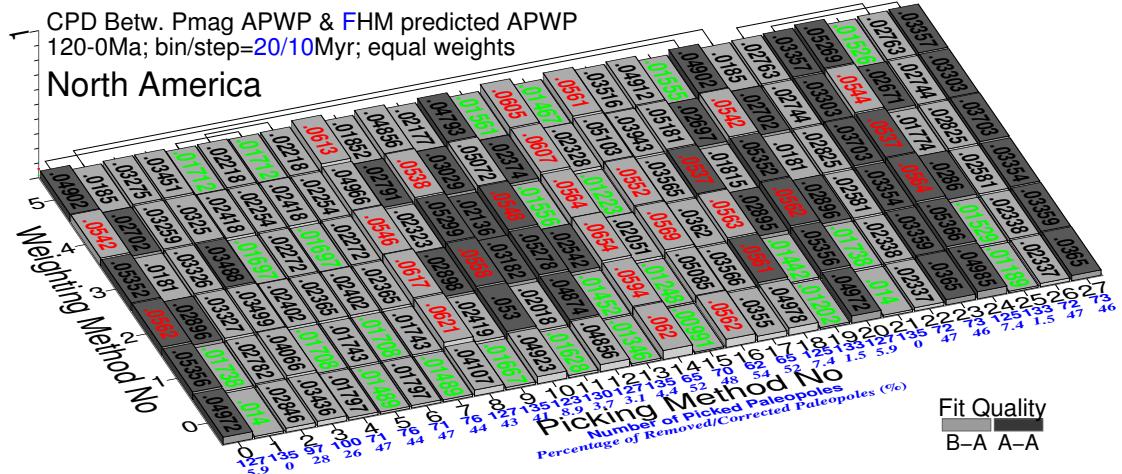
The overall effect of increased window and step size varies between plates (Fig. 3.19). 118/168 ( $\sim 70\%$ ) of equivalent  $\mathcal{CPD}$  scores for the North American plate increase, indicating reduced similarity with respect to the reference path (Fig. 3.19a). The increased occurrence of A-A rather than B-A FQ ratings (Fig. 3.17a versus Fig. 3.9a) indicate that in some cases this may be due to a better constrained path becoming more



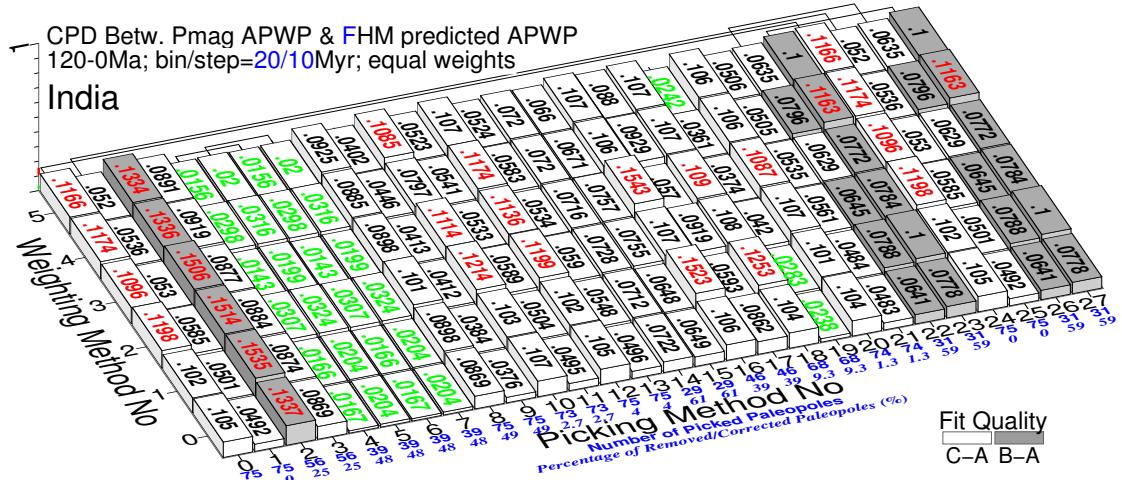
**Figure 3.16:** Tested angular difference ( $d_a$ ) values (color shaded) between North American paleomagnetic APWPs and its predicted APWP from FHM and related plate circuits. The paths are in 10 Myr bin and 5 Myr step. The labeled numbers on the grids are the averaged numbers of site mean poles that are contributing to each segment-orientation-change's three mean path poles.

distinguishable. Decreased scores are confined to particular, mostly AMP-derived picking methods: Pk 2, 22 and 26 (filtering based on  $\alpha_{95}$  and age uncertainties), 4 and 6 (removal of sedimentary paleopoles), and 16,17 (removal of post-1983 results). It is clearly due to the corresponding increase of N in each increased window for these AMP-derived quality filters.

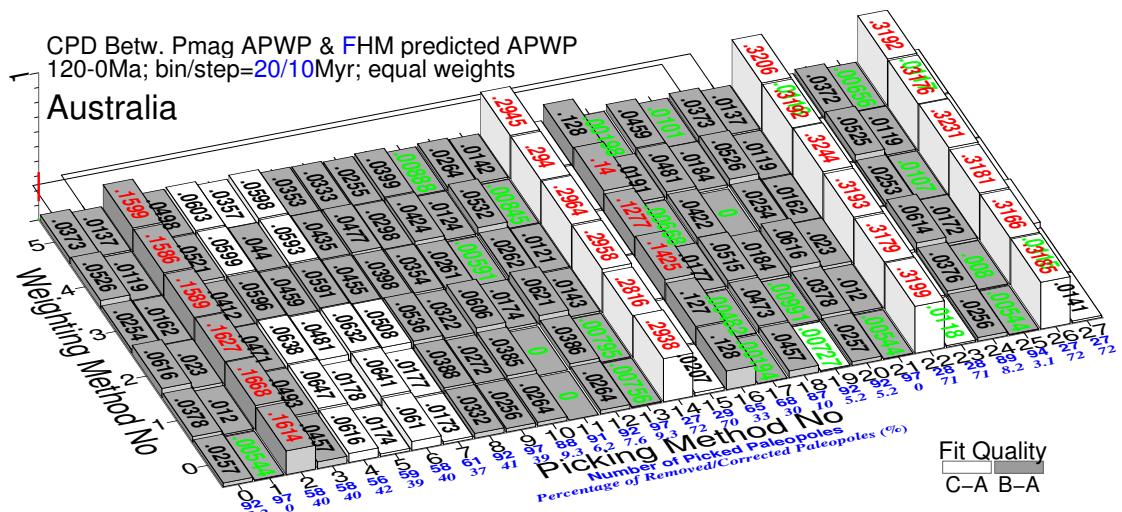
For the Indian and Australian plates, there is a trend towards decreased  $\mathcal{CPD}$



(a) North America: minimum 0.00991 (15(0)), maximum 0.0654 (14(2)), mean 0.0339, median 0.0296371

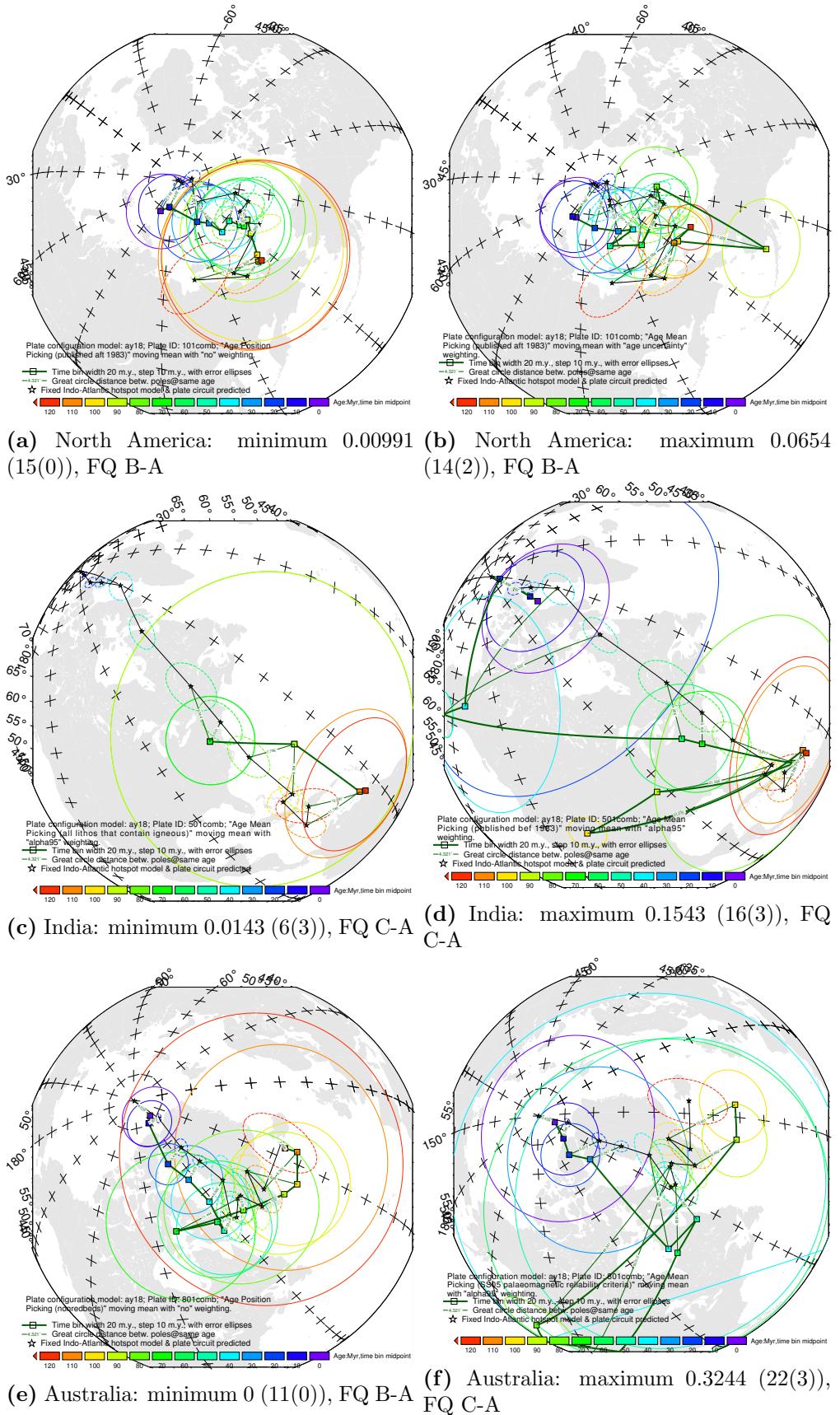


(b) India: minimum 0.0142822 (6(3)), maximum 0.154278 (16(3)), mean 0.07384, median 0.072088



(c) Australia: minimum 0 (11(0,1),19(3)), maximum 0.324438 (22(3)), mean 0.0682, median 0.036432

**Figure 3.17:** As Fig. 3.9, 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. 3.9.



**Figure 3.18:** Path comparisons with best and worst difference values shown in Fig. 3.17. The parenthetical remarks are Pk No and Wt No.

scores, with 153/168 (~91%) and 104/168 (~62%) of equivalent  $\mathcal{CPD}$  scores being lower than for the 10/5 Myr window/step, respectively (Fig. 3.19b, 3.19c). Where the score increases, the actual difference is close to 0, indicating very little change, with the exception of Pk 14 (AMP, removal of pre-1983 results) on the Australian plate. The largest score decreases are associated with Wt 3 (AMP,  $\alpha_{95}$  and age filtering) on the Indian plate, and Pk 15 (APP, removal of pre-1983 results) on the Australian plate. Most FQ ratings remain unchanged (Fig. 3.17b versus Fig. 3.9b, Fig. 3.17c versus Fig. 3.9c): most obviously, ratings for Pk 5 and 7 (APP, removal of sedimentary poles) on the Indian plate and Pk 14 on the Australian plate change from B-A to C-A.

### 3.3.4.2 Relative Performance of Methods

As for the baseline results (Section 3.3.1, Fig. 3.9), the effects of windowing and filtering, particularly the windowing method (APP versus AMP), are much more apparent than the effects of weighting. APP-derived scores still outperform AMP-derived ones, but the overall degree of difference is reduced for the North American and Indian plates, with percentage of APP-derived  $\mathcal{CPD}$  scores more than 3 times the equivalent AMP-derived score reducing to ~26% and ~4% (from ~44% and ~40%), respectively. For the Australian plate, this percentage increases from ~58% to ~70%.

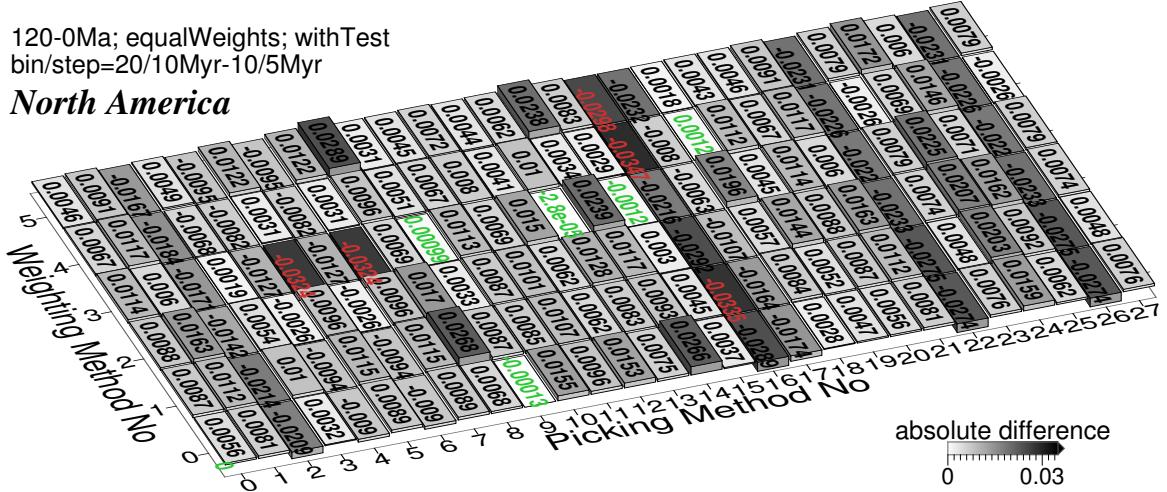
This is largely the result of larger changes in the AMP-derived  $\mathcal{CPD}$  scores compared to changes in the APP-derived ones with the wider time window. The scores for APP-derived methods Pk 1 (no filtering), other non-aggressive filters (Pk 13, 19, 21, 25), and Pk 9 (correction for inclination flattening), are particularly stable (Fig. 3.19) for all three plates. Filtering according to the SS05 criteria (Pk 23 and 27) also yields comparable scores for both time windows on the North American and Indian plates.

As for the 10/5 window, the lowest AMP/APP differences are for Pk 4,5 and 6,7 (igneous poles preferred – low scores) and Pk 2,3 and 22,23 (filtering for spatial and temporal uncertainty – high scores) for the North American and Indian plates; and 8,9 (correction for inclination shallowing – moderate scores) for the Australian plates. Likewise, the highest AMP/APP difference is still associated with Pk 0,1 (no filtering) on the North American and Australian plates, and Pk 16,17 on the Australian plate.

Overall, the relative performance of the different filtering methods when using the 20/10 window/step remains similar to that observed for the 10/5 window/step (Figs. 3.9, 3.17). Some of the lowest scores are still produced by Pk 1 (APP, no filtering). Of the aggressive filters, removing sedimentary paleopoles (Pk 4,5 and 6,7) still perform well, selecting only spatially and temporally well-constrained paleopoles (Pk

120-0Ma; equalWeights; withTest  
bin/step=20/10Myr-10/5Myr

### North America



(a) North America: Generally Bin/Step 10/5 Myr is better ( $\frac{59}{84} \approx 70.24\%$ )

120-0Ma; equalWeights; withTest  
bin/step=20/10Myr-10/5Myr

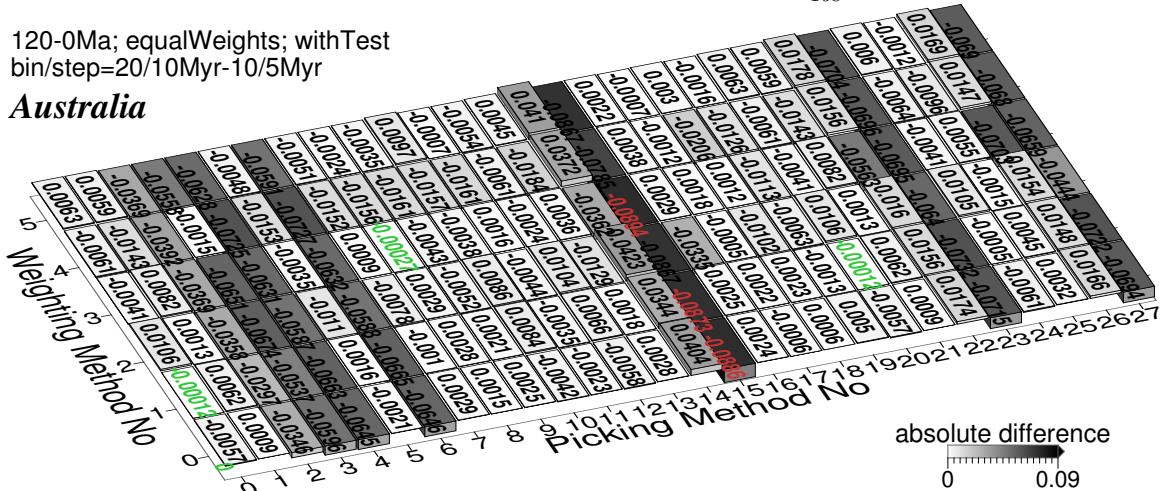
### India



(b) India: Generally Bin/Step 20/10 Myr is better ( $\frac{153}{168} \approx 91.07\%$ )

120-0Ma; equalWeights; withTest  
bin/step=20/10Myr-10/5Myr

### Australia



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

**Figure 3.19:** Differences between grids in Fig. 3.9 (10 Myr bin, 5 Myr step) and Fig. 3.17 (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.

2,3 and 22,23) still perform relatively poorly, and correction for inclination flattening (Pk 8,9) has little effect.

In addition, using only paleopoles published before 1983 (Pk 16,17) still produces relatively high *CPD* scores for the North American and Indian plates, although the scores for the AMP-derived Pk 16 are substantially reduced with the wider time window (Figs. 3.19a, 3.19b). Likewise, for Australia, using only paleopoles published after 1983 (Pk 14,15) still produces higher scores than Pk 16,17: scores for the wider time window are even higher for the AMP-derived Pk 14, but substantially lower for the APP-derived Pk 15 (Fig. 3.19c).

**Table 3.4:** Consistency check on comparisons of picking methods' performance between 20/10 and 10/5 Myr window/step. Notes: E means expected; UE means unexpected.

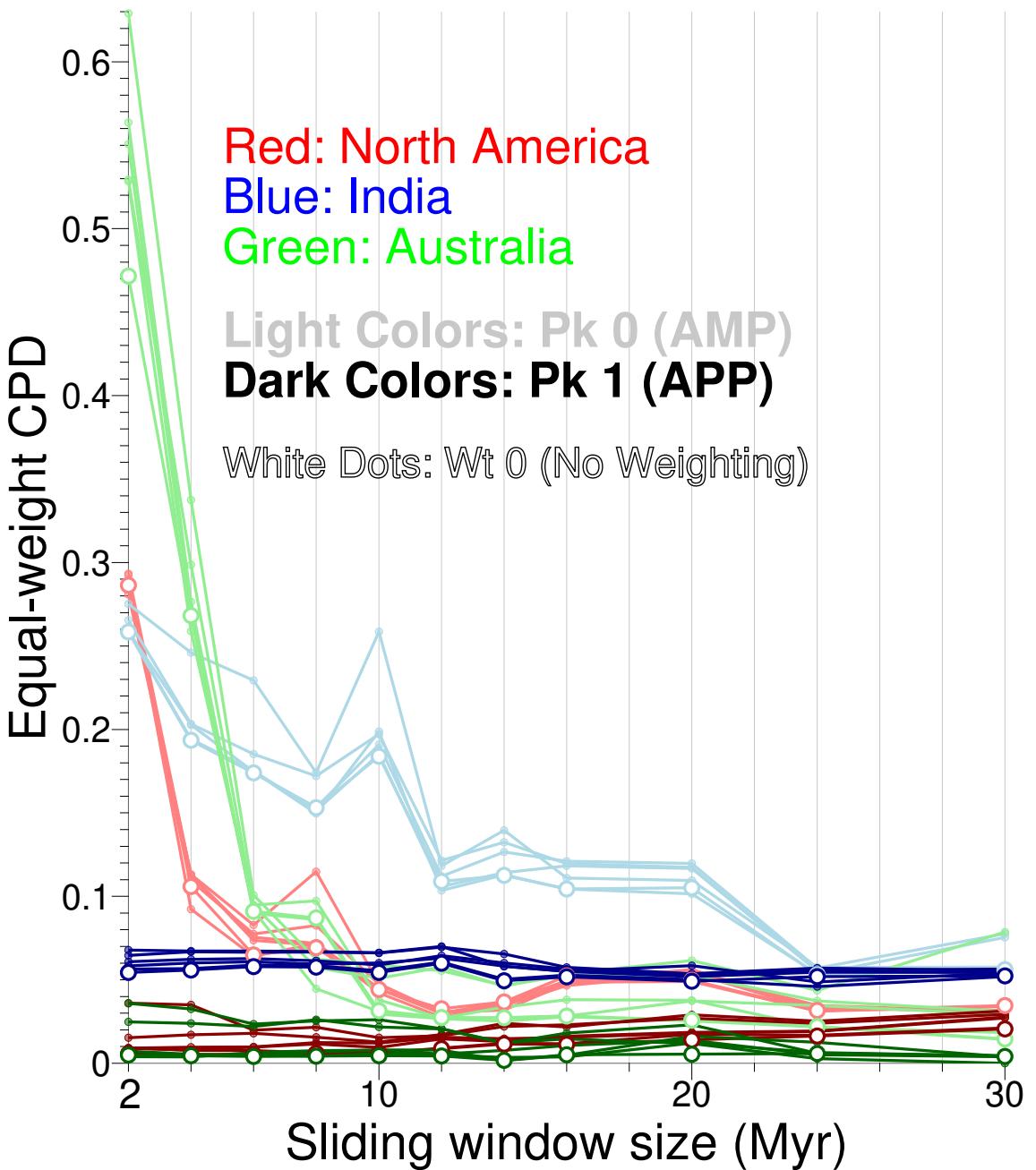
Comparisons				Consistency of best		Consistency of worst			If CPD values for 20/10 are lower (Y/N)				
10/5	20/10	Y/N	Special case	Same Pk	Y/N	Special case	Same Pk	Mean	Median	Max	Min	All	If N, UE case
FHM													
Fig. 3.9a	Fig. 3.17a	Y	only for 20/10 (E)	3 more best: Pk 4, 6, 9 11, 13, 15, 19, 21, 25	7, 11, 13, 15, 19, 21, 25	Pk 2, 5, 7, 26 for 10/5 (E); 0, 8, 10, 12, 20, 24 for 20/10 (UE)	22, 14, 16, 18	N	Y	Y	N	N	Positive values in Fig. 3.19a
Fig. 3.9b	Fig. 3.17b	Y	26 only for 10/5 (UE)	4 more best: Pk 9, 21, 22, 4–7, 19	11, 13, 17, 19,	N	10/5 (E); 23, 27 for 20/10 (UE)	22, 12, 18, 24	Y	Y	Y	N	Positive values in Fig. 3.19b
85	Fig. 3.9c	Fig. 3.17c	Y	2 more best: Pk 23, 27 only for 20/10 (E)	1, 11, 13, 17, 19, 21, 25	Y	Pk 4 only for 10/5 (E)	14, 22, 26	Y	Y	Y	N	Positive values in Fig. 3.19c
MHM													
Fig. 3.21a	Fig. 3.23a	N	(UE); 22, 26 for 20/10 (E)	Pk 1, 9, 11, 13, 19, 21, 25 for 10/5 (UE); 22, 26 for 20/10 (E)	5, 7, 15	N	Pk 16, 17 only for 10/5 (E); 8 only for 20/10 (UE)	10, 12, 14, 18, 20, 24	0, 10 12, 14, 18, 20, 24	N	Y	N	(0,8,10,11,12,14,18,20,24,25)(0–5) (1,11,13,19,21)(0,1,5) 3(0,3,5) (5,7)(0–2,4,5) 15(0–4) (9,23,27)(0,1,3,5) account for 60.71%
Fig. 3.21b	Fig. 3.23b	N	only for 10/5 (UE); 4, 6 only for 20/10 (E)	Pk 19, 21 only for 10/5 (UE); 4, 6 only for 20/10 (E)	5, 7, 22, 26	Y	3 more worst: Pk 8, 14, 20 for 10/5 (E)	0, 2, 10, 12, 16, 18, 24	0, 10, 12, 14, 16, 18, 24	Y	Y	Y	(1,25)(0,3,5) 15(3) (1,9,23,27)(0–5) 21(0,1,3–5) (22,26)(4); account for 19.05%
Fig. 3.21c	Fig. 3.23c	Y	27 only for 20/10 (E)	3 more best: Pk 15, 23, 13, 17, 19, 21, 25	1, 11, 13, 17, 19, 21, 25	Y	None	2, 14, 16, 22, 26	2, 14, 16, 22, 26	Y	Y	N	(0,20)(5) (5,7)(3) (1,11,13,16,18,19,21,25)(0,1,3,5) (8,17)(0–3,5) 9(0,1) 10(1,5) 12(2) 14(2,4,5) (22,26)(0–2,4,5) 24(1,2,5); account for 39.88%

### **3.3.4.3 Other Window Sizes and Steps**

**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.

**Table 3.5:** Equal-weight 120–0 Ma CPDs for the three representative continents' paleomagnetic APWPs compared with their FHM predicted APWPs. The best are in dark green and underlined, second best in light green.

Window/step size (Myr)	Pk 0					Pk 1				
	Wt 0	Wt 1	Wt 2	Wt 3	Wt 4	Wt 0	Wt 1	Wt 2	Wt 3	Wt 4
North America										
2/1	0.2864	0.2925	0.28685	0.280944	0.287	0.29324	<b>0.005759</b>	0.0091586	<b>0.01528</b>	0.03601
4/2	0.10584	0.09228	0.10938	0.11301	0.1131	0.11263	<b>0.00525</b>	0.007665	0.01701	<b>0.009612</b>
6/3	0.06486	0.064078	0.077426	0.073621	0.082867	0.075522	<u><b>0.005418</b></u>	<b>0.007522</b>	0.017936	<b>0.0097317</b>
8/4	0.06934	0.07168	0.082474	0.0707846	0.071686	0.071473	0.006	<b>0.006217</b>	<b>0.01556</b>	0.01242
10/5	0.04412	0.04486	0.04739	0.04212	0.04753	0.04445	0.00591	<b>0.00616</b>	<u><b>0.01262</b></u>	0.01208
12/6	<b>0.03271</b>	<b>0.02987</b>	<b>0.03024</b>	<b>0.0276143</b>	<b>0.03012</b>	<b>0.02988</b>	0.008736	<u><b>0.00902</b></u>	0.01715	<b>0.01642</b>
14/7(119-0)	0.0367	0.03695	<u><b>0.03498</b></u>	<u><b>0.0347</b></u>	<u><b>0.0319</b></u>	<u><b>0.0352</b></u>	0.0115	0.01198	0.023922	0.01477
16/8	0.052386	0.04923	0.04964	0.04653	0.047712	0.048523	0.01138	0.0117	0.02177	0.015409
20/10	0.04972	0.05356	0.0562	0.05352	0.0542	0.04902	0.014	0.01728	0.02896	0.0181
24/12	<b>0.031642</b>	<b>0.0342</b>	<b>0.03273</b>	<b>0.03469</b>	<b>0.0314765</b>	<b>0.031253</b>	0.0164	0.016737	0.02521	0.02397
30/15	<u><b>0.0345</b></u>	<u><b>0.0298</b></u>	<u><b>0.0317</b></u>	<u><b>0.0307</b></u>	<u><b>0.03402</b></u>	<u><b>0.0341</b></u>	0.0206	0.0211	0.0313	0.0272
India										
2/1	0.25838	0.258757	0.258701	0.275184	0.265442	0.258692	0.05446	0.05639	0.06456	0.06075
4/2	0.19345	0.20276	0.19344	0.24597	0.203264	0.19423	0.05588	0.056691	0.066803	0.062283
6/3	0.17396	0.17476	0.173975	0.229325	0.18514	0.175174	0.057887	0.05882	0.0664754	0.062734
8/4	0.15309	0.14966	0.15321	0.17412	0.172154	0.15124	0.05761	0.05848	0.06671	0.06139
10/5	0.1839	0.1845	0.1909	0.2586	0.1972	0.1852	0.0545	0.0554	0.0662	0.0589
12/6	0.108924	0.1035	<b>0.11205</b>	0.11831	0.121497	<b>0.10587</b>	0.059897	0.06068	0.06993	0.064499
14/7(119-0)	0.112537	0.112885	0.126554	0.139516	0.132359	0.114195	<b>0.04942</b>	<b>0.0502588</b>	0.057931	0.06018
16/8	<b>0.104461</b>	0.104463	0.121002	0.110942	0.119599	0.118336	<b>0.051735</b>	0.052813	<b>0.055188</b>	<b>0.056389</b>
20/10	0.1052	<b>0.1015</b>	0.1198	<b>0.1096</b>	<b>0.1174</b>	0.1166	<b>0.0492</b>	<b>0.0501</b>	0.0585	<b>0.0533</b>
24/12	<b>0.053143</b>	<b>0.05356</b>	<b>0.056986</b>	<b>0.05747</b>	<b>0.05558</b>	<b>0.0553047</b>	<b>0.051926</b>	<b>0.045995</b>	<b>0.04868</b>	<b>0.0557455</b>
30/15	<u><b>0.05617</b></u>	<u><b>0.0754578</b></u>	<u><b>0.0775947</b></u>	<u><b>0.0565421</b></u>	<u><b>0.056635</b></u>	<u><b>0.0565421</b></u>	0.0523614	<u><b>0.0519862</b></u>	<u><b>0.054158</b></u>	0.0563985
Australia										
2/1	0.471554	0.529417	0.52834	0.563622	0.628921	0.55071	0.00498465	0.0047458	0.0247455	0.0072776
4/2	0.26822	0.29862	0.25881	0.276464	0.33741	0.268965	0.004543	0.004578	0.023895	0.005417
6/3	0.090985	0.0947676	0.09448	0.08977	0.10083	0.091376	<b>0.0040955</b>	0.004847	0.02199	<b>0.0046074</b>
8/4	0.0870445	0.097201	0.057284	0.086078	0.06245	0.04463	<b>0.004265</b>	<b>0.004191</b>	0.026156	0.0067075
10/5	0.0315	0.039	<b>0.0509</b>	0.0305	0.0601	0.031	0.0045	0.0048	0.0199	0.0058
12/6	0.027348	<b>0.026942</b>	0.057592	0.026687	0.055426	<b>0.0270495</b>	0.004341	0.0049061	0.020529	0.007813
14/7(119-0)	0.02725	0.033687	<b>0.046242</b>	<b>0.0252737</b>	<b>0.04651</b>	<b>0.02722</b>	<b>0.0017226</b>	<b>0.00354029</b>	<b>0.0119085</b>	<b>0.00157378</b>
16/8	0.028261	0.037993	0.053856	0.0278993	0.05263	<b>0.0282745</b>	0.0050338	<b>0.0037223</b>	0.018136	0.0047548
20/10	<b>0.0257</b>	0.0378	0.0616	0.0254	0.0526	0.0373	0.00544	0.012	0.023	0.0162
24/12	<b>0.0219624</b>	<b>0.0214394</b>	<b>0.043667</b>	<b>0.0216832</b>	<b>0.0372685</b>	0.0345906	0.0058473	0.0063788	<b>0.0059937</b>	0.0125858
30/15	<b>0.014614</b>	<b>0.0183819</b>	<b>0.078627</b>	<b>0.014254</b>	<b>0.031029</b>	0.0293416	<b>0.0412276</b>	<b>0.0044694</b>	<b>0.00448627</b>	<b>0.00397289</b>



**Figure 3.20:** Plot of the equal-weight  $\mathcal{CPD}$  scores collected in Table 3.5. Note that here the step size is always half of the sliding window size and the reference path is the FHM derived.

**The results** are summarised in Table 3.4, Table 3.5 and Fig. 3.20. Based on these results, the effect of doubling the size of the time window and step is negligible.  $\mathcal{CPD}$  scores for APWPs generated with no filtering (Pk 0,1) over a wider range of time windows and steps, from 2/1 to 30/15 are shown in Fig. 3.20 and Table 3.5. Scores

for APP-derived paths (Pk 1) remain stable over the whole tested range. AMP-derived paths (Pk 0) have higher scores for windows narrower than 10 Myr, but are much more stable for windows wider than 10–12 Myr, whilst remaining generally higher than the equivalent APP scores.

### 3.3.5 Moving versus Fixed Hotspot Reference

Fig. 3.21 and Fig. 3.23 show the  $\mathcal{CPD}$  scores for the APWPs generated with all 28 picking and 6 weighting methods, compared to the MHM reference paths (stars and solid line in Figs. 3.6-3.8) with picking time windows and steps of 10/5 Myr and 20/10 Myr.

#### 3.3.5.1 Overall Change

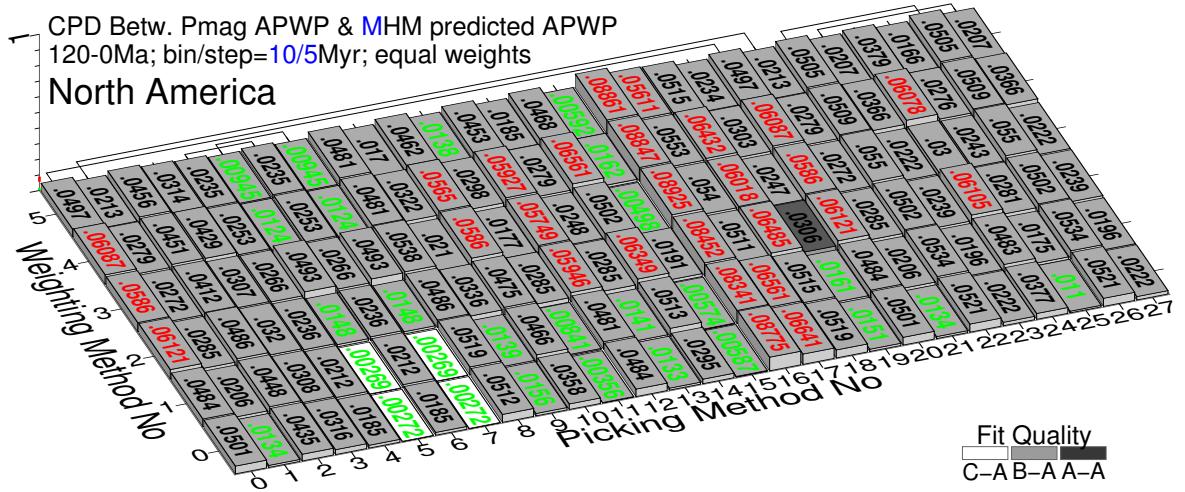
Although the reference path is changed to MHM, mean, median and range values of  $\mathcal{CPD}$  scores are almost unchanged and comparable (Fig. 3.9 versus Fig. 3.21, and Fig. 3.17 versus Fig. 3.23).

For 10/5 Myr window/step, the absolute differences (Fig. 3.25) are all lower than 0.066, and actually 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. 3.25), for North America (Fig. 3.25a), FHM derived path is a slightly better reference in general, while for both India (Fig. 3.25b) and Australia (Fig. 3.25c), generally MHM derived path is a slightly better choice.

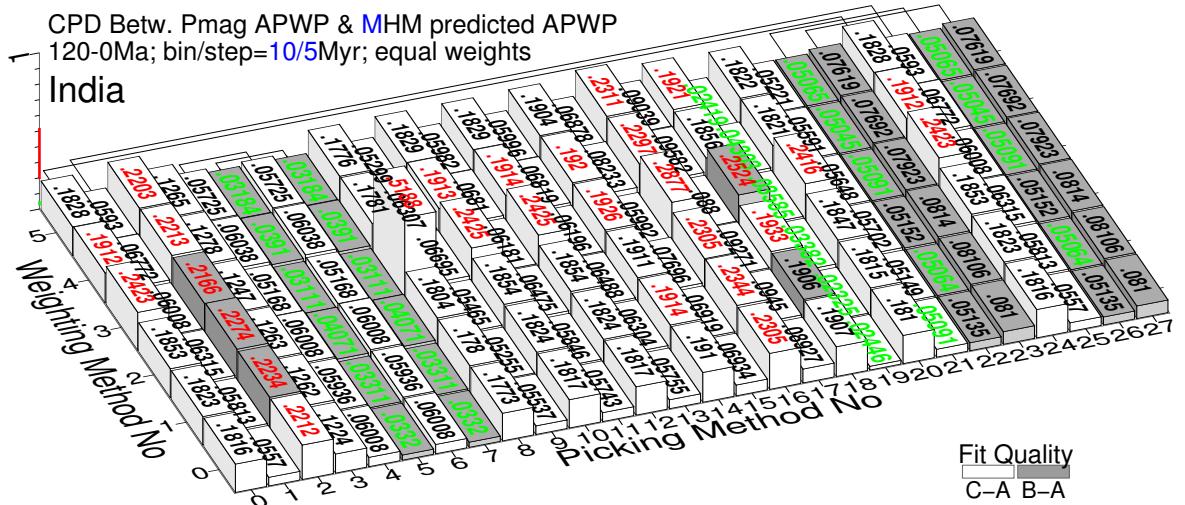
For North America, large changes seem favored by Wt 3, and small changes seem favored by Wt 5 or Pk 15 (Fig. 3.25). Pk 3 brings minor changes to both North America and India. Pk 22 and 26 show large changes and Pk 17, 19, 21 and 25 show minor changes for both India and Australia.

#### 3.3.5.2 Relative Performance of Methods

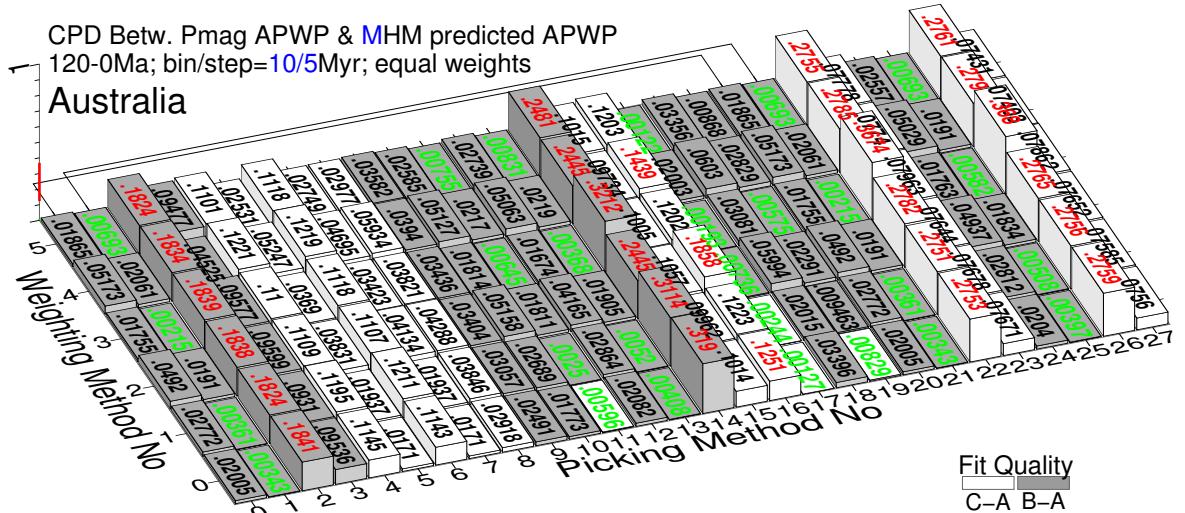
1. When 10/5 Myr bin/step is applied, Pk 19 is still among the best, and Pk 16 still one of the worst, for all 3 plates. Even when 20/10 Myr bin/step is applied, 19 is still relatively a “good” method, and 16 a “bad” one.
2. APP still outperforms AMP. For 10/5 Myr bin/step, the percentage of factor of 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%).



(a) North America: minimum 0.00268588 (5(1)), maximum 0.0892467 (16(3)), mean 0.03674, median 0.03177075

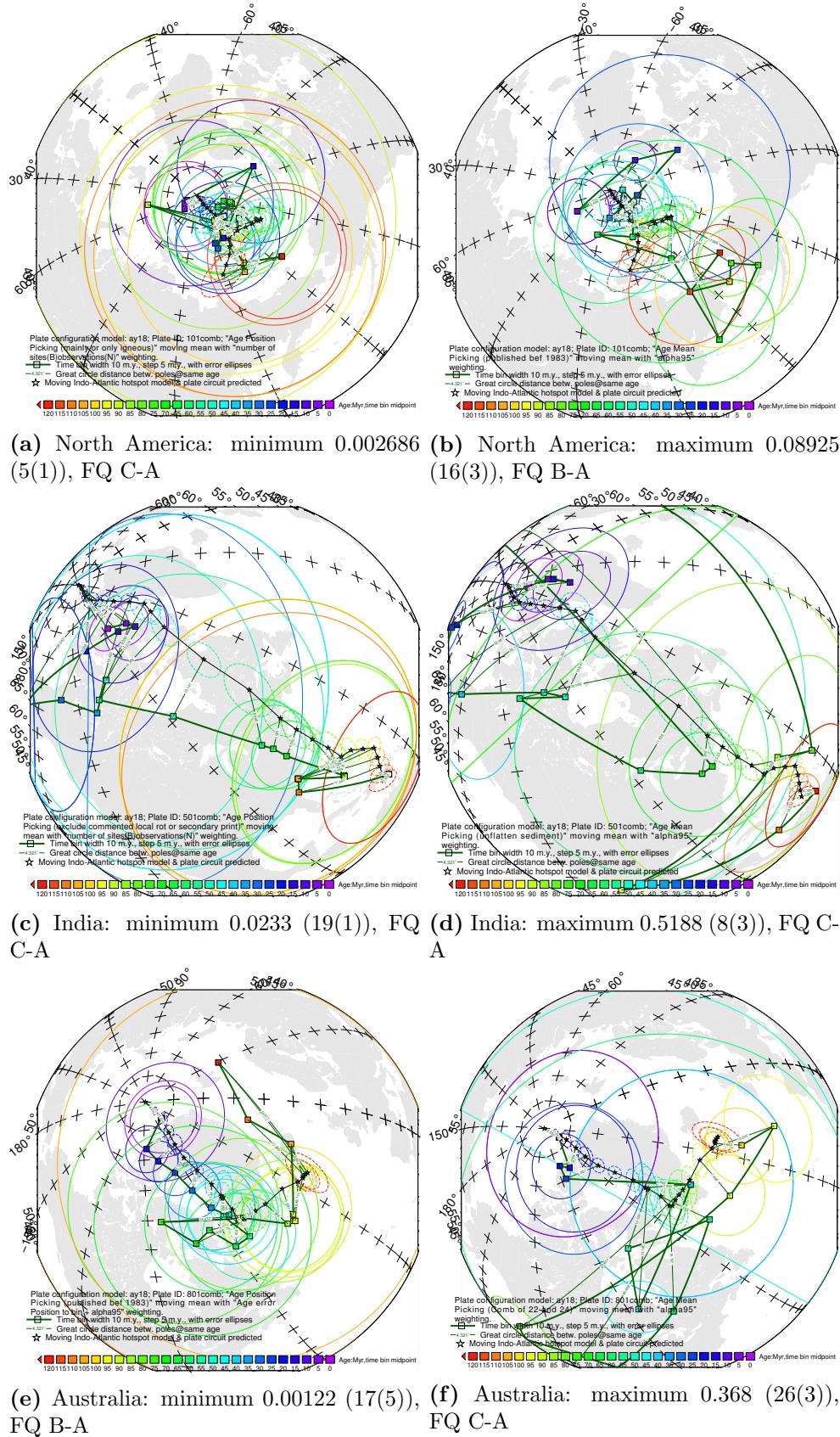


(b) India: minimum 0.0232517 (19(1)), maximum 0.51876 (8(3)), mean 0.11364, median 0.076556

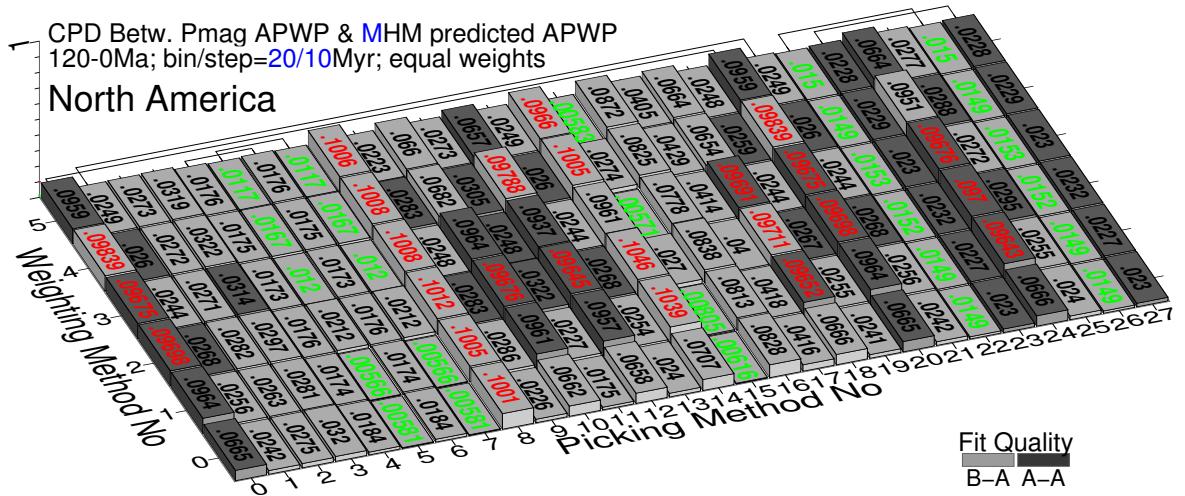


(c) Australia: minimum 0.00122074 (17(5)), maximum 0.367952 (26(3)), mean 0.077, median 0.03893

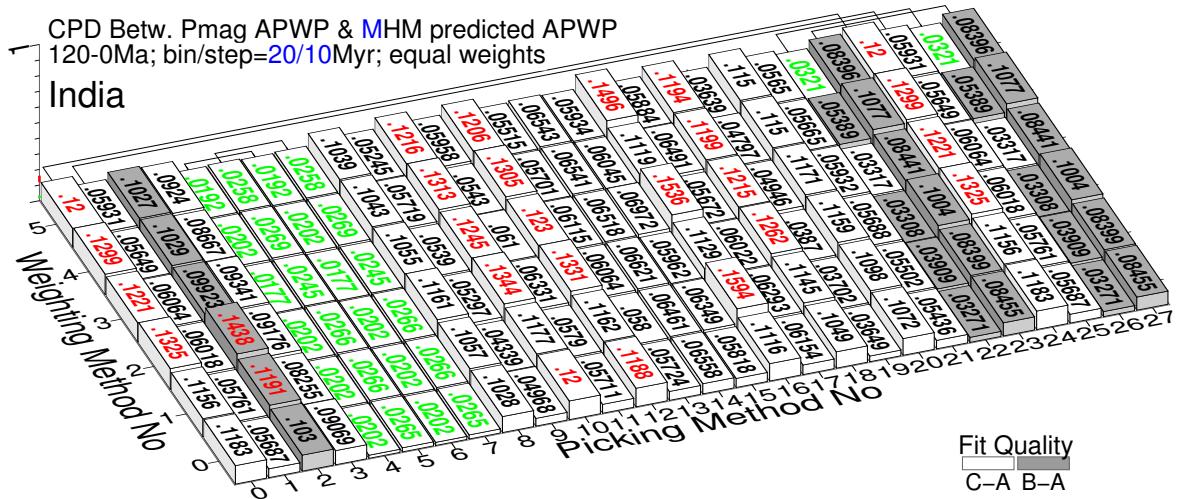
**Figure 3.21:** As Fig. 3.9, here the reference path is predicted from MHM. See the numbers of the picked paleopoles for methods in Fig. 3.9.



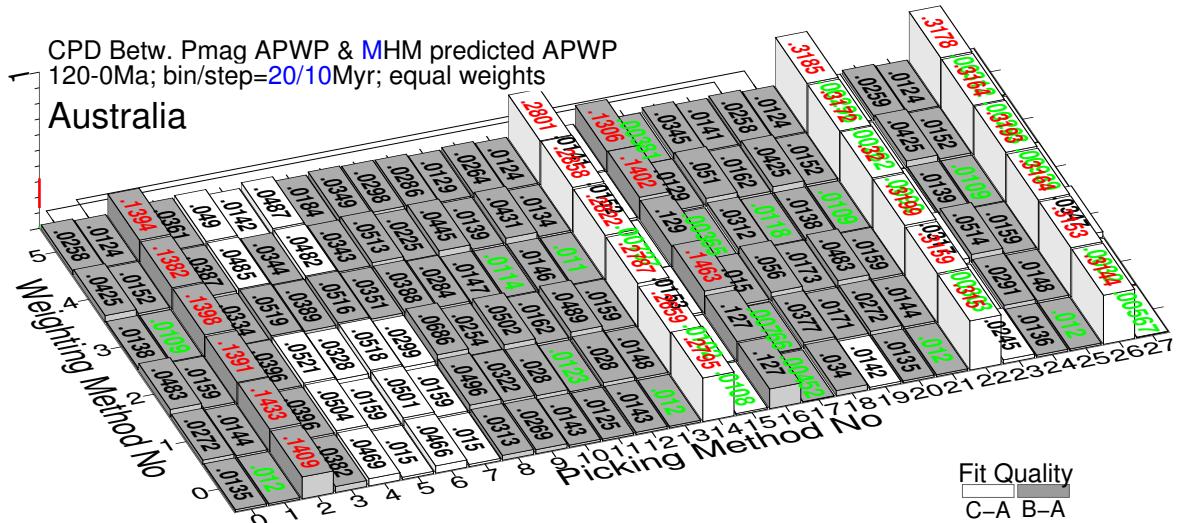
**Figure 3.22:** Path comparisons with best and worst difference values shown in Fig. 3.21. The parenthetical remarks are Pk No with Wt No.



(a) North America: minimum 0.00565784 (5(1)), maximum 0.104618 (14(2)), mean 0.043777, median 0.02679955

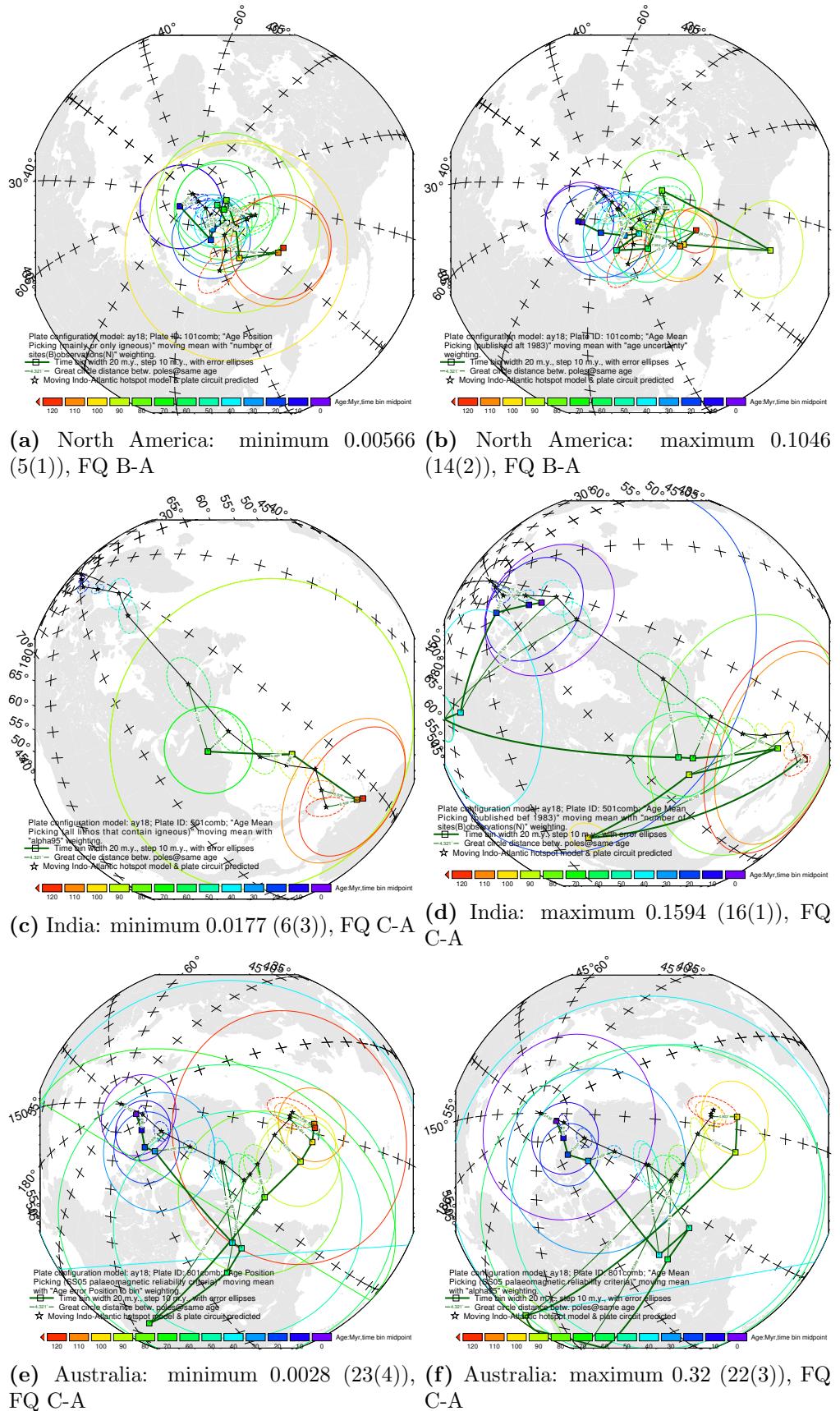


(b) India: minimum 0.0177 (6(3)), maximum 0.15937 (16(1)), mean 0.0745, median 0.061346

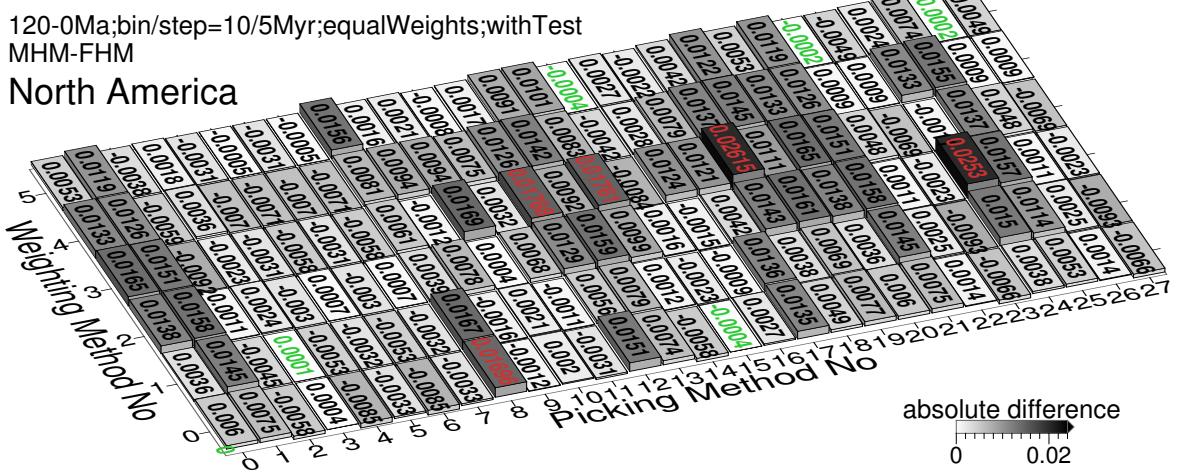


(c) Australia: minimum 0.00282 (23(4)), maximum 0.31998 (22(3)), mean 0.062766, median 0.02803

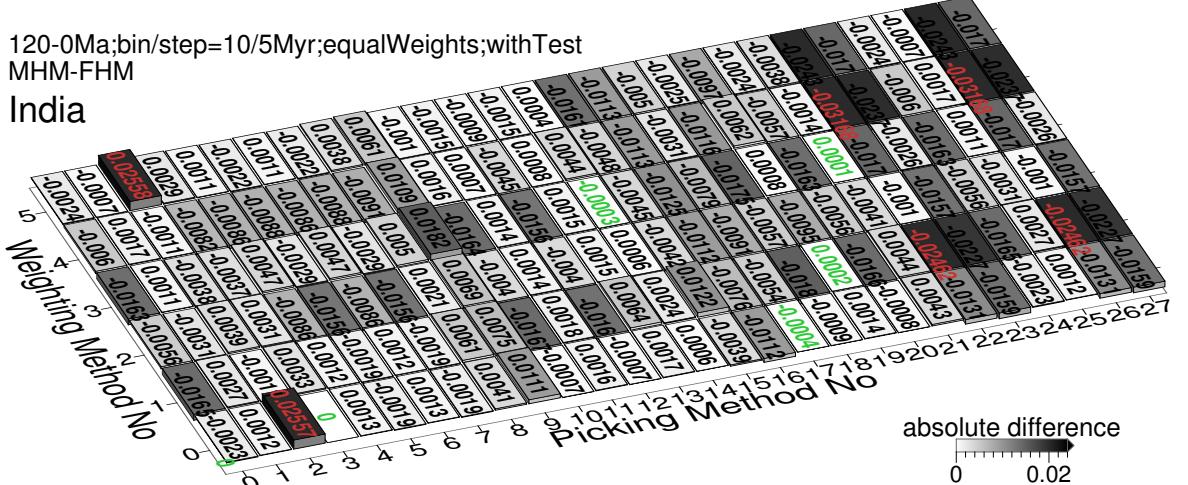
**Figure 3.23:** As Fig. 3.17, here the reference path is predicted from MHM. See the numbers of picked paleopoles in Fig. 3.9.



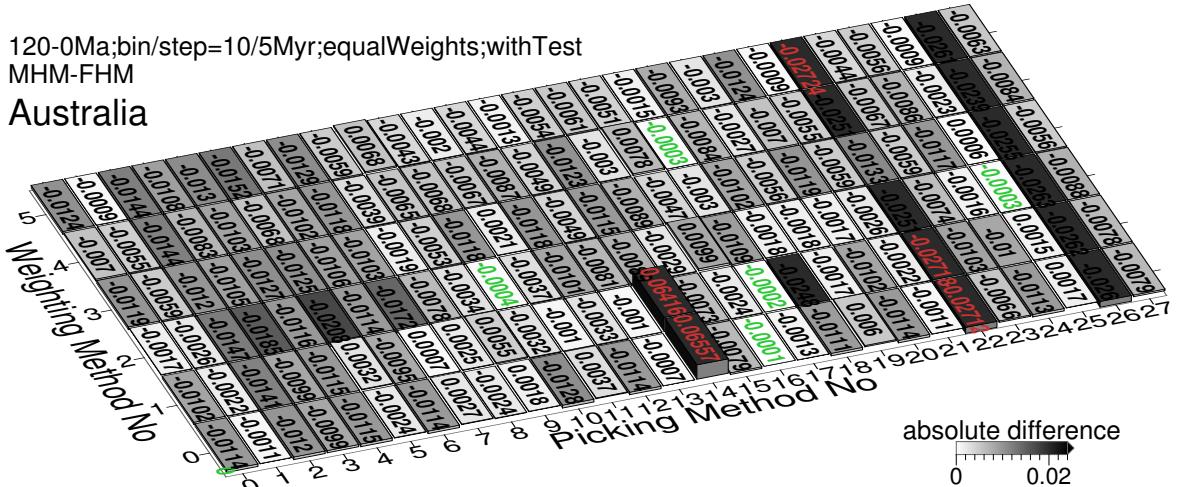
**Figure 3.24:** Path comparisons with best and worst difference values shown in Fig. 3.23. The parenthetical remarks are Pk No and Wt3 No.



(a) Fig. 3.21a-Fig. 3.9a; Percentage for positive values: ~66.67%



(b) Fig. 3.21b-Fig. 3.9b; Percentage for positive values: ~34.52%



(c) Fig. 3.21c-Fig. 3.9c; Percentage for positive values: ~19.62%

**Figure 3.25:** Differences between results from two different reference paths, FHM (Fig. 3.9) and MHM (Fig. 3.21) 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.

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 ( $\sim$ 44.05%) is more than FHM (28.6%).

3. Both scores and AMP/APP difference are still low for Pk 4,5, 6,7.
4. Both scores and AMP/APP difference are still high for Pk 16,17 (Pk 17 gives low scores for Australia).
5. Still scores are high but AMP/APP difference are low for Pk 2,3.
6. Still scores are low but AMP/APP difference are high for Pk 0,1 on Australia.
7. Comparable AMP/APP (20/10 versus 10/5, or FHM versus MHM) still appears for Pk 8,9 on Australian plate.
8. Pk 1 is still a good performer.
9. Pk 4,5, 6,7 are still ‘good’ aggressive filters.
10. Pk 2,3 (at least 2) are still ‘bad’ aggressive filters.
11. For 10/5 Myr bin/step, Pk 22 is still better than 23 for India. For 20/10 Myr bin/step, Pk 22 is still better than 23 for both North America and India.
12. Pk 16,17 are still ‘bad’ for North America and India.

**Table 3.6:** Performance statistics of all the picking and weighting methods for 10/5 and 20/10 Myr window/step path comparisons.

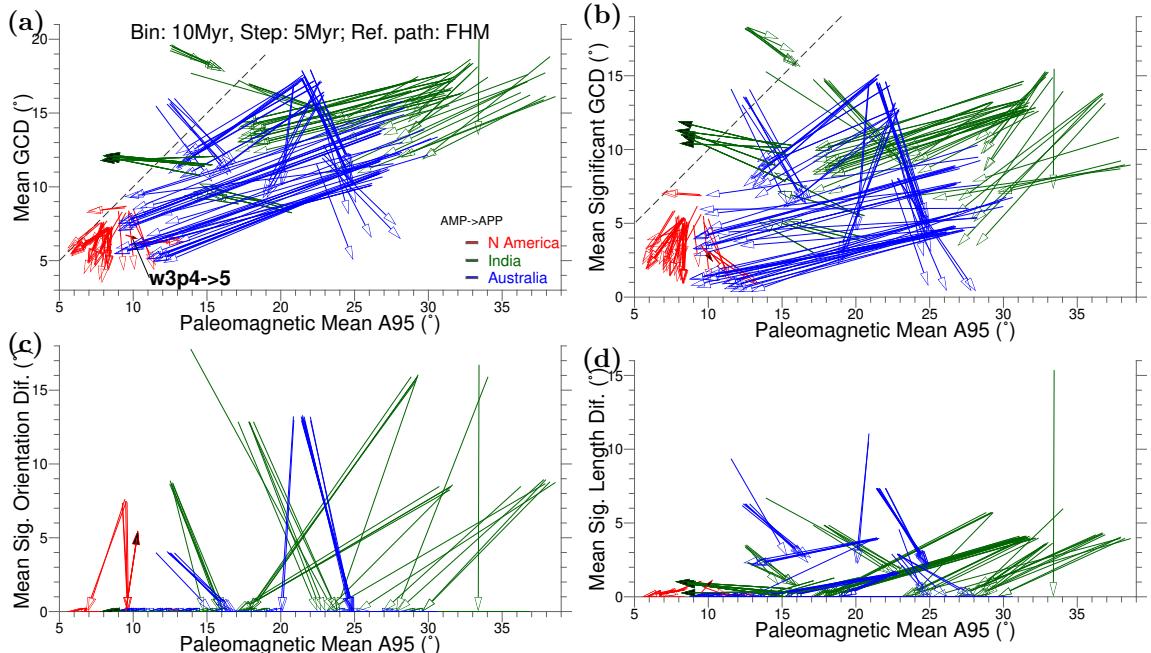
Hotspot model for ref path	CPD grid	Best		Worst		% of APP better than AMP	Count occurrences of each Wt being best for 28 Pks			Pk14,15 (New Studies) better than Pk16,17(Old)
		Pk	Wt	Pk	Wt		0	1	2	
Fig. 3.9a	1, 5, 7, 11, 13 15, <b>19</b> , 21, 25	0, 1, 3, 5	2, 5, 7, 14, <b>16</b> , 17, 18, 22, 26	97.6	<b>9</b>	4	4	7	0	5
Fig. 3.9b	4–7, 9, <b>19</b> <b>21</b> , 22, 26	0–5	0, 2, 8, 10, 12, <b>16</b> , 18, 20, 24	85.7	<b>15</b>	2	2	6	1	Y
FHM	Fig. 3.9c	1, 11, 13, 17, <b>19</b> , 21, 25	0, 1,	2, 4, 14, <b>16</b> , 22, 26	100	10	5	1	7	3
Fig. 3.17a	1, 4–7, 9, 11, 13 15, <b>19</b> , 21, 25	3, 5	0, 8, 10, 12, 14, <b>16</b> , 18, 20, 24	72.6	<b>14</b>	2	3	3	6	N
Fig. 3.17b	4–7, <b>19</b>	0–5	0, 2, 10, 12, <b>16</b> , 18, 20, 23, 24, 27	69	11	0	9	1	1	Y, 4y2n
Fig. 3.17c	1, 11, 13, 17, <b>19</b> , 21, 23, 25, 27	0, 3–5	2, 14, <b>16</b> , 22, 26	100	9	<b>10</b>	—	1	—	N
MHM	Fig. 3.21a	1, 5, 7, 9, 11, 13, 15, <b>19</b> , 21, 25	0, 0–5	0, 10, 12, 14, <b>16</b> –18, 20, 24	97.6	<b>10</b>	9	3	4	0
Fig. 3.21b	5, 7, <b>19</b> , <b>21</b> , 22, 26	0–5	0, 2, 8, 10, 12, <b>14</b> , <b>16</b> , 18, 20, 24	85.7	<b>12</b>	2	0	7	3	Y
Fig. 3.21c	1, 11, 13, 17, <b>19</b> , 21, 25	0–3, 5	2, 14, <b>16</b> , 22, 26	98.8	6	4	2	10	2	N
Fig. 3.23a	5, 7, 15, 22, 26	0, 0–5	0, 8, 10, 12, 14, 18, 20, 24	76.2	<b>10</b>	8	1	4	1	4y5n, Y
Fig. 3.23b	4–7, 22, 26	0–5	0, 2, 10, 12, <b>16</b> , 18, 24	70.2	<b>6</b>	<b>7</b>	6	3	6	Y, 3y3n
Fig. 3.23c	1, 11, 13, 15, 17, <b>19</b> , 21, 23, 25, 27	2, 22, 26	14, <b>16</b> , 22, 26	100	<b>12</b>	1	10	4	1	N

## 3.4 Discussion

### 3.4.1 Question: Why APP Methods Generally Produce Better Similarities than AMP Methods?

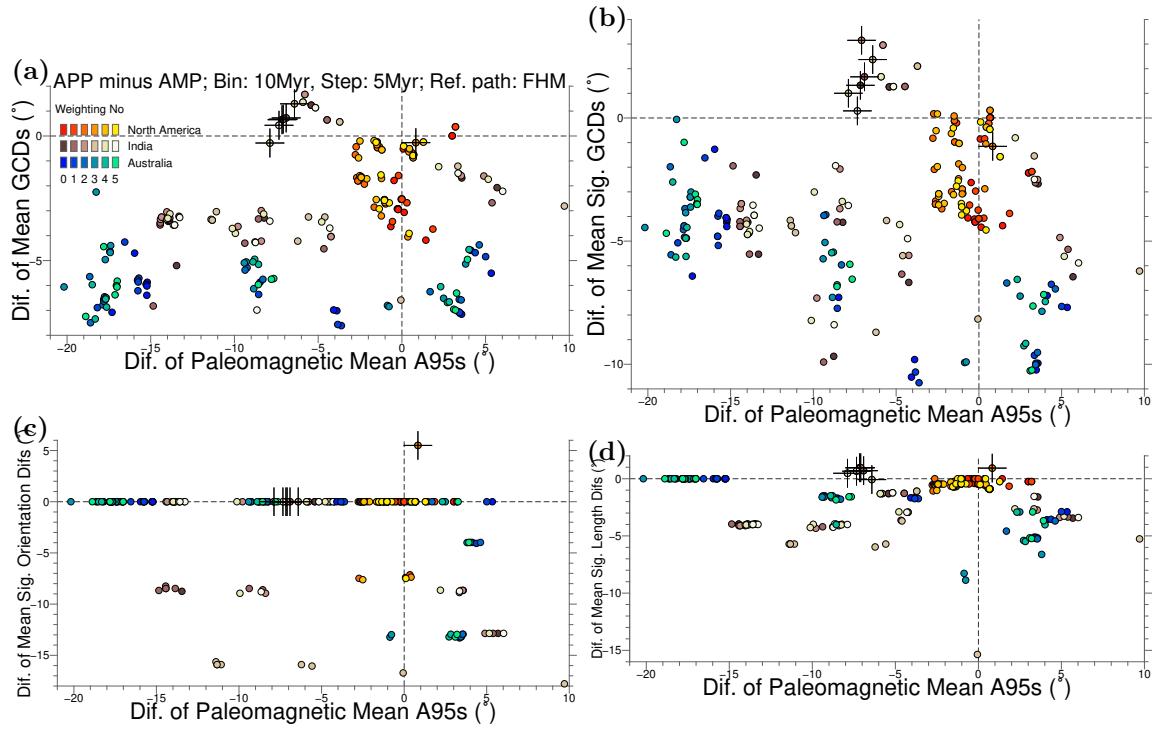
#### 3.4.1.1 Perspective of Conceptual Difference between AMP and APP

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. 3.26a and Fig. 3.26b). Compared with AMP, APP usually improves both and generates paths with higher accuracy and also higher precision (generally increasing number of contributing paleopoles).



**Figure 3.26:** Paleomagnetic APWP's 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. Arrowtails are the results from AMP, while arrowheads are from APP. Black color filled arrowheads are the small number of special cases of AMP derived equal-weight CPDs better than APP (see details in Fig. 3.9 and Table 3.6).

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. 3.26 and Fig. 3.27). However, if the added paleopoles were all or mostly “bad”,



**Figure 3.27:** Differences of APP and AMP coordinates shown in Fig. 3.26. Crosses locates the small minority cases of AMP derived equal-weight  $\mathcal{CPD}$ s better than APP (see details in Fig. 3.9 and Table 3.6).

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 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 3.6). 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.

### 3.4.1.2 Perspective of Stability Comparison between AMP and APP

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. 3.18e is improved a bit compared to Fig. 3.10e). 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. 3.18 and even 30/15 Myr window and step; Table 3.5 and Fig. 3.20) 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 3.5 and Fig. 3.20). 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 the future). 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? Or is there a threshold number of average paleopoles per window for making a reliable paleomagnetic APWP? 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 3.5), as expected, AMP needs wider sliding window and step to get closer to the reference path while APP does not (Fig. 3.20). Even the best sizes of sliding 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.

#### 3.4.1.3 Summary

**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 Wt 3 is used with AMP.

**APP is still recommended,** not only because the temporal uncertainty is incorporated into the algorithm but also the results from APP are not as 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. 3.20). 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.

#### 3.4.2 Question: Why AMP Methods Sometimes Unexceptionally Produce Better Similarities than APP Methods?

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. 3.9a, there are only two special (of 84 APP versus AMP comparisons) cases Pk (Wt) 4(3), 6(3) better than 5(3), 7(3) respectively. Compared with the Pk (Wt) 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 two 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 one more significant  $d_l$  segment pair (Table 3.7). 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) versus 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 Pk (Wt) 2(0) versus 3(0) for 20/10 Myr window/step North America, all their  $d_a$  and  $d_l$  are indistinguishable. Compared with the results from AMP, although the coeval pole GCDs are generally unchanged or decreased or even increased (but not too much) for APP, this spatial improvement is not able to offset the negative effects of also generally unchanged or decreased or even increased (but not too much) paleomagnetic A95s, which potentially brings more statistically distinguishable coeval poles (e.g. the 20 Ma and 110 Ma poles for Pk 3 and Wt 0; Table 3.8). This further causes greater distinguishable mean  $d_s$  from the APP methods. The similar phenomenon occurs to Fig. 3.17a Pk 2 versus 3 with Wt 2, 3 and 5, Pk 4 versus 5 with Wt 1, 3 and 5, and Pk 6 versus 7 with Wt 1, 3 and 5, and so on.

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 are paleopoles 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 3.9) 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. 3.9b Pk 22 versus 23 (actually exactly the same as Pk 26 versus 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 3.6).

Other Type 1 (e.g. Table 3.7) cases: Fig. 3.17a Pk (Wt) 2(1) versus 3(1). Fig. 3.21a 4(3) versus 5(3), 6(3) versus 7(3).

Other Type 2 (e.g. Table 3.8) cases: Fig. 3.17a Pk (Wt) 2(0, 2, 3, 5) versus 3(0, 2, 3, 5), 4(1, 3, 5) versus 5(1, 3, 5), 6(1, 3, 5) versus 7(1, 3, 5). Fig. 3.17b 4(0–5) versus 5(0–5), 6(0–5) versus 7(0–5), 14(2, 3) versus 15(2, 3), 22(0, 2, 3) versus

**Table 3.7:** One example of the Type 1 rare cases where AMP gives better similarity result than APP does from North America (10/5 Myr window/step). Only statistically significant values are listed here.

Age (Ma)	FHM predicted	Pk4(Wt3)					Pk5(Wt3)			
		ds		dl		ds		da		
DM/DP ()	Pmag A95 ()	Dist ()	Age (Ma)	Diff ()	Pmag A95 ()	Dist ()	Age (Ma)	Diff ()		
10	1.44607/0.793714	14.876819	<b>9.937</b>	105–110	5.91855				<b>10–15–20</b>	<b>126.59</b>
15	1.2875/0.816514					2.0857		<b>11.805</b>		
25	2.48031/1.10915					6.3358		<b>6.873</b>		dl
55	3.58782/2.14032	4.6347	<b>5.372</b>						Age (Ma)	Diff ()
60	4.85938/3.17602					6.5922		6.215	<b>10–15</b>	<b>13.52</b>
65	3.68984/2.30014					8.6632		<b>7.6</b>	<b>15–20</b>	<b>14.68</b>
75	2.6435/1.54052	9.0812	<b>8.836</b>							
100	2.8983/2.68346	8.892	<b>8.455</b>							
105	2.32328/1.74639	5.3	<b>5.03</b>							
110	4.13015/2.25964	3.8	<b>9.8064</b>							
115	4.63512/2.58006	19.6676	<b>9.3345</b>			8.5		<b>11.704</b>		
120	7.34408/4.06043	3.515	<b>17.35</b>			7.728		<b>15.258</b>		

**Table 3.8:** One example of the Type 2 rare cases where AMP gives better similarity result than APP does from North America (20/10 Myr window/step). Only statistically significant values are listed here.

Age (Ma)	FHM predicted	ds					
		Pk2(Wt0)			Pk3(Wt0)		
		DM/DP (°)	Pmag A95 (°)	Dist (°)	Pmag A95 (°)	Dist (°)	
0	0		3.97	<b>5.714</b>	3.97		<b>5.714</b>
10	1.44607/0.793714	3.879		<b>6.034</b>	3.879		<b>6.034</b>
20	1.58039/1.10047				6.771		<b>6.934</b>
50	3.57782/1.61328	3.8644		<b>7.304</b>	4.03		<b>8.6</b>
60	4.85938/3.17602	5.716		<b>8.457</b>	5.55		<b>7.367</b>
100	2.8983/2.68346	10.769		<b>7.308</b>	10.769		<b>7.308</b>
110	4.13015/2.25964				3.29		<b>8.311</b>
120	7.34408/4.06043	3.38		<b>16.41</b>	3.083		<b>16.728</b>

23(0, 2, 3), 26(0, 2, 3) versus 27(0, 2, 3). Fig. 3.17c 4(2) versus 5(2). Fig. 3.21c 8(5) versus 9(5). Fig. 3.23a 2(0–5) versus 3(0–5), 4(2) versus 5(2), 6(2) versus 7(2), 22(0–5) versus 23(0–5), 26(0–5) versus 27(0–5). Fig. 3.23b 4(0–5) versus 5(0–5), 6(0–5) versus 7(0–5), 14(3) versus 15(3).

Combined Type 1 and 2 cases: Fig. 3.17a Pk (Wt) 22(0–5) versus 23(0–5), 26(0–5) versus 27(0–5). Fig. 3.17b Pk (Wt) 22(1, 4, 5) versus 23(1, 4, 5), 26(1, 4, 5) versus 27(1, 4, 5). Fig. 3.23b 22(0–5) versus 23(0–5), 26(0–5) versus 27(0–5).

**Table 3.9:** One example of the Type 3 rare cases where AMP gives better similarity result than APP does from India (10/5 Myr window/step). Only statistically significant values are listed here. Note that for the bold-number ages, there is no mean poles at all for the “Pk 22 (Wt 2)” case.

FHM predicted	Pk22(Wt 2)					Pk23(Wt2)					
	ds		ds		dl						
Age (Ma)	DM/DP (°)	Pmag	DM/DP (°)	Dist (°)	N	Pmag	DM/DP (°)	Dist (°)	N	Age (Ma)	Diff (°)
0	0	<b>6.28</b>		<b>12.72</b>	<b>2</b>	<b>23.54</b>		<b>18.14</b>	<b>3</b>	<b>80–85</b>	6.286
10	1.12124/0.673225	5.4/3.1		29.9	1	5.4/3.1		29.9	1	<b>110–115</b>	16.684
15	1.1347/0.8127					5.4/3.1		28.28	1		
60	4.79687/3.07133					8.817		8.28	20		
70	4.26508/2.48783					3.26		4.464	20		
75	2.6777/1.57975					5		4.477	1		
80	4.20828/2.50294					5		3.358	1		
85	2.50744/1.24746	5		7.632	1	5		7.632	1		
90	3.88998/1.43423					5		10.884	1		
95	2.23389/1.6247					5		11.099	1		
100	2.8062/2.59819					5		11.4155	1		
105	2.32328/1.74639					5		14.908	1		
110	4.55519/2.49218					6.8/4.9		13.962	1		
115	4.63512/2.58006	10.73		10.508	5	10.73		10.508	5		
120	6.02639/3.3319	10.73		10.508	5	10.73		10.508	5		

Other Type 3 (e.g. Table 3.9) cases: Fig. 3.21a 4(2–5) versus 5(2–5). Fig. 3.21b 22(0–5) versus 23(0–5), 26(0–5) versus 27(0–5).

### 3.4.3 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 to each other (Fig. 3.9, Fig. 3.17, Fig. 3.21, Fig. 3.23). This closeness is also generally observed in the form of clusters in Fig. 3.28 and Fig. 3.29. In addition, from the general statistics of performance of the six weighting methods shown in Table 3.6, Wt 0 mostly performs the best or at least the second best, which means no weighting works better in general.

When the above-mentioned question, about why APP generally produces better fits than AMP, is tackled, we already find that both accuracy (how closely do the pairs/segments/angles match) and precision (how large are the uncertainties on the pairs/segments/angles, i.e. how difficult do they distinguish) can be the factors that finally determine the difference score. Although another factor, resolution (how many pairs/segments/angles are actually being compared) can also influence the difference score (e.g. Table 3.5), here this factor is not relevant to comparisons between different weightings for a certain picking method, because the numbers of picked paleopoles are the same for the six weighting methods.

Therefore, at the very basic level, for example, a lower score is the result of one of, or combination of:

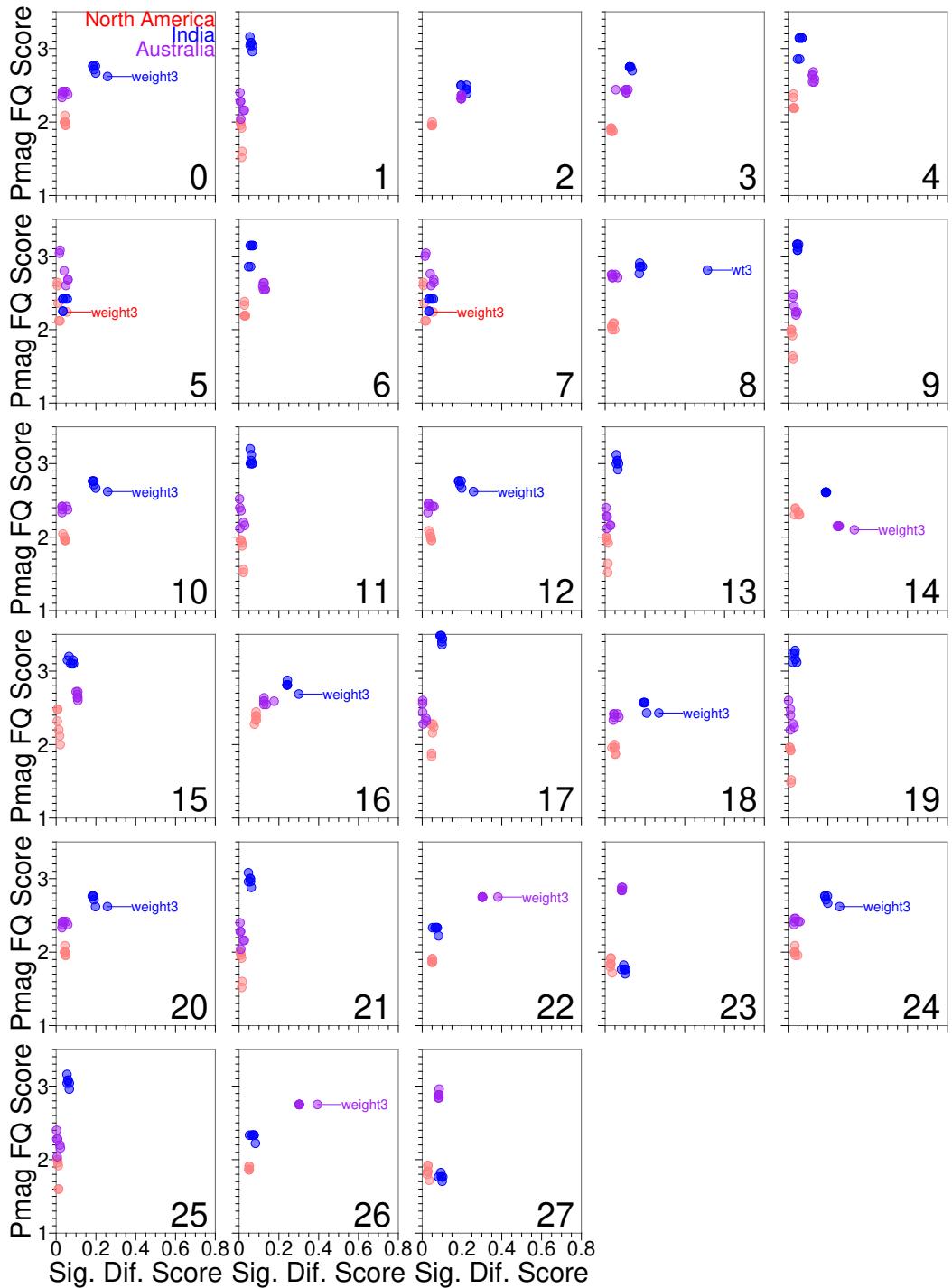
1. A reduction in the difference scores of significantly different pairs/segments, straightforwardly interpreted as a better fit (improved accuracy).
2. Previously significantly different pairs/segments becoming insignificant. This can occur either because the fit is better (improved accuracy), or because of an increase in the uncertainty of the mean poles (they become less distinguishable — decreased precision).

Fig. 3.30 and Fig. 3.31 show that the proportions of results from Wt 1–5 to result from 0 are generally in the second quadrant of the coordinate plane, where proportional change in difference score is positive whereas proportional change in paleomagnetic FQ score is negative. This means the five weightings (Wt 1–5) do make effects, not obviously in accuracy but mainly in improving precision, which could potentially expose more pairs of distinguishable poles/segments/angle-changes. Or even accuracy is improved in a small amount, improved precision intends to cancel out the effects from improved accuracy.

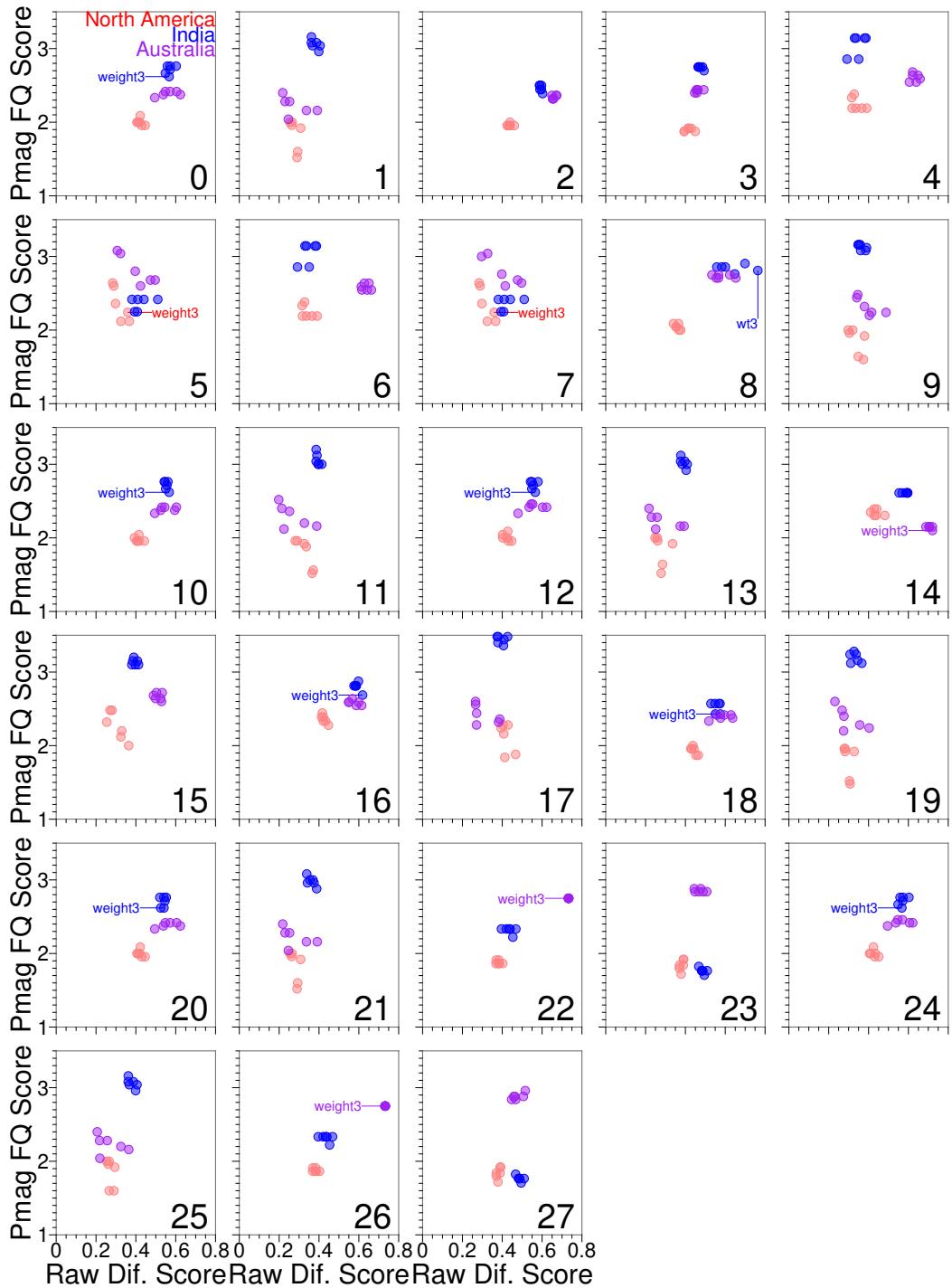
There are a few special cases that one or two of the six weighting methods gives a result with a dramatically worsened difference score, e.g. for weighting method Wt 3 (Fig. 3.9). From the labeled dots of Wt 3 in Fig. 3.28, Fig. 3.29, Fig. 3.30 and Fig. 3.31, we can get a general impression that Wt 3 is indeed improving precision but not accuracy (at least not enough to offset the effects from improved precision) so that this precision improvement is also the culprit that worsens the final score. Highly improved precision without corresponding improved accuracy would potentially bring more significant differences in shape metrics. Wt 3 is exactly performing in this form.

In addition, for Wt 3, a small size of  $\alpha_{95}$  (high precision) could be caused by 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.

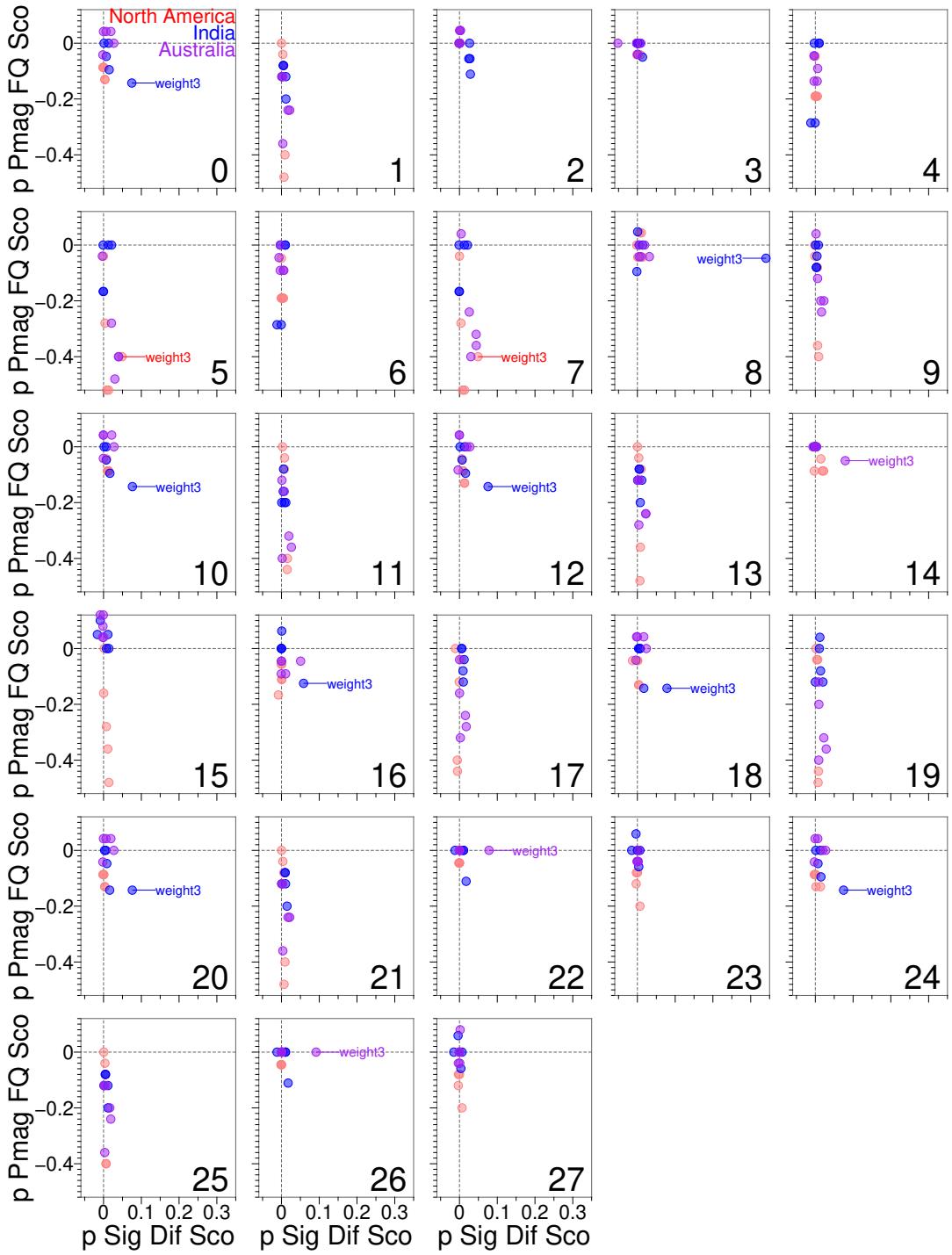
Generally, weighting is affecting because different weighting functions give obviously different results. However, interestingly weighting does not improve fit and generally no weighting (Wt 0) is giving the best fit, although in most cases weighting does improve precision. Wt 2 or 4 is not recommended, because they never have generated the best similarities (Table 3.6), 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



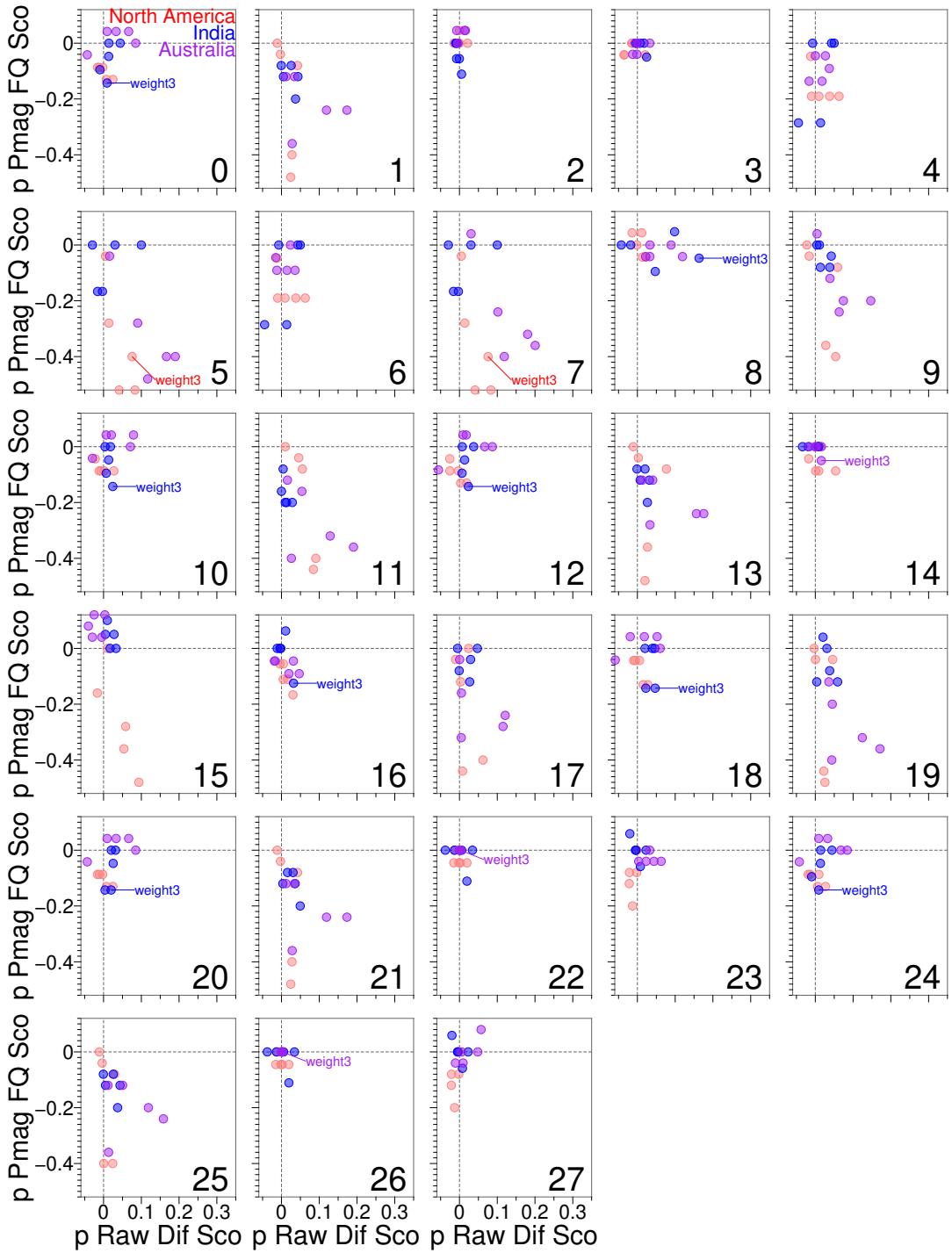
**Figure 3.28:** 10/5 Myr bin/step paleomagnetic APWP's FQ score (different from FQ, see the definitions of FQ and FQ score in Chapter 2) versus  $\mathcal{CPD}$  significant score (reference path: FHM predicted) for the 28 different picking methods. See the  $\mathcal{CPD}$  scores and Ppath-Rpath FQ in Fig. 3.9. Only those results dramatically worsened by Wt 3 are labeled.



**Figure 3.29:** 10/5 Myr bin/step paleomagnetic APWP's FQ score (different from FQ, see the definitions of FQ and FQ score in Chapter 2) versus raw difference score (reference path: FHM predicted) for the 28 different picking methods. See FQ versus significant difference score in Fig. 3.28. Only those results dramatically worsened by Wt 3 are labeled.



**Figure 3.30:** Proportion of Wt 1–5 to 0: Proportional change in 10/5 Myr bin/step paleomagnetic APWP's FQ score (different from FQ, see the definitions of FQ and FQ score in Chapter 2) versus proportional change in  $\mathcal{CPD}$  significant score (reference path: FHM predicted) for the 28 different picking methods. See the  $\mathcal{CPD}$  scores and Ppath-Rpath FQ in Fig. 3.9. Only those results dramatically worsened by Wt 3 are labeled.



**Figure 3.31:** Proportion of Wt 1–5 to 0: Proportional change in 10/5 Myr bin/step paleomagnetic APWP's FQ score (different from FQ, see the definitions of FQ and FQ score in Chapter 2) versus proportional change in raw difference score (reference path: FHM predicted) for the 28 different picking methods. See FQ versus significant difference score in Fig. 3.30. Only those results dramatically worsened by Wt 3 are labeled.

continent or some specific picking methods. For example, Wt 3 works generally fine with Australian data (Table 3.6). However, Wt 3 is not recommended for North America and India.

### **3.4.4 Question: Why Best and Worst Methods Sometimes Are Not Consistent?**

For all the three continents, North America, India and Australia, Pk 19 (APP with local rotation or secondary print excluded) and 21 (APP with local rotation and secondary print corrected) are consistently the best or at least the relatively better (for example, in Fig. 3.17b, Pk 21 results are not colored in green, but ranging from 0.0483–0.0561 that is still much less than the mean, 0.074 and also the median, 0.072); whereas Pk 16 (AMP with publications before 1983) and 18 (AMP with local rotation or secondary print excluded) are consistently the worst or at least the relatively worse.

Nevertheless, for each single continent, they have their own consistently best and worst picking methods that are not the best or worst for other continents. For example, Pk 15 (APP with publications after 1983) works well with North America. We know that 70 North American paleopoles (about 53%) have contributed to Pk 15 for 120–0 Ma. However, only 28 Indian paleopoles (about 38%) and 29 Australian paleopoles (about 30%) have contributed to Pk 15. In contrast, the Pk 17 (APP with publications before 1983) works well with only Australia. This also tells that number of paleopoles is a key factor that affects the fit score. The fact that Pk 22 (AMP with SS05 criteria) doesn't work well with only Australia also reflects the importance of this factor.

In summary, the reason why some best and worst methods are not consistent is generally that different continents have their unique data sets and get their own paleomagnetic studies in varying degrees.

### **3.4.5 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. 3.14, Fig. 3.15 and Fig. 3.16).

## 3.5 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.

### 3.5.1 Universal Rules of Ways of Processing Paleomagnetic Data

Although effects of filters (all the picking methods where number of paleopoles shrinks compared to Pk 0 and Pk 1; see Fig. 3.9) have a marginal change in reducing N (precision potentially going down), some filters do improve the similarity score, for example, Pk 25 (APP without superseded data) is always giving better scores than Pk 1 for all the three continents. However, Pk 0 and Pk 1 (no filtering and corrections applied) still generally perform well.

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 3.6). APP with most filters/corrections (Pk 3, 5, 7, 9, 11, ..., 27) are generally giving worse scores than APP without any filter/correction (Pk 1).

In most cases the APP methods produce better similarities than the AMP methods (Table 3.6).

Actually weighting is not improving the fit but improving precision generally. For quite many of the methods, no weighting is the best performer (Table 3.6). For example, score is likely worse for the combined methods of weight Pk 3 and AMP.

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.

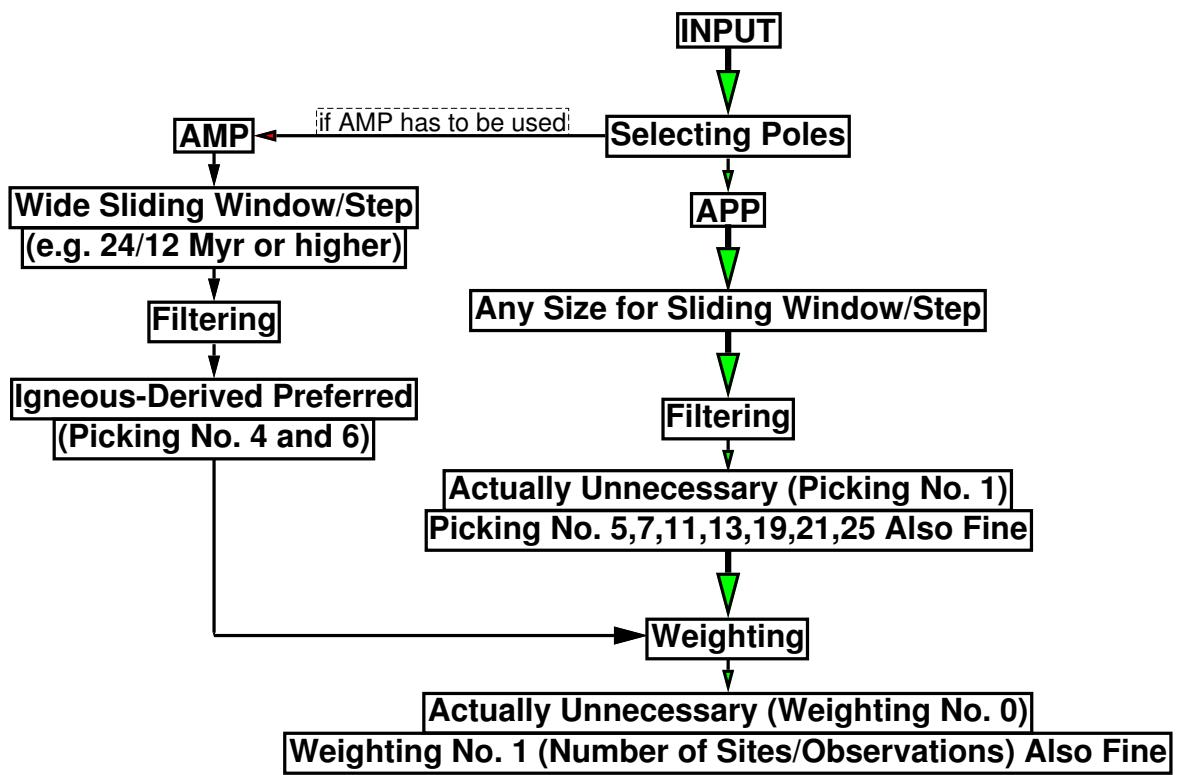
### **3.5.2 Conditional Rules of Ways of Processing Paleomagnetic Data**

Pk 16 (AMP with data from old studies) and 18 (AMP without data affected by rotations and secondary print) are not recommended for generating a paleomagnetic APWP.

### **3.5.3 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 3.5 and Fig. 3.20). 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. Pk 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, Wt 1 (related to the number of paleomagnetic sampling sites and observations) is generally better than the other four given weighting methods (Fig. 3.32).

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. Pk 4 and 6), which indicates that if we have fewer data, these data need to be better in quality (Fig. 3.32).



**Figure 3.32:** Flowchart for recommended procedure of processing paleomagnetic data.

# Bibliography

- M Beaman, W W Sager, G D Acton, L Lanci, and J Pares. Improved late Cretaceous and early Cenozoic paleomagnetic apparent polar wander path for the Pacific plate. *Earth Planet Sci Lett*, 262(1-2):1–20, 2007. doi: 10.1016/j.epsl.2007.05.036.
- J Besse and V Courtillot. Apparent and true polar wander and the geometry of the geomagnetic field over the last 200 Myr. *J Geophys Res*, 107(B11):2300, 2002. doi: 10.1029/2000JB000050.
- D Bilardello. The do’s and don’ts of inclination shallowing corrections. *Institute for Rock Magnetism Quarterly*, 26(3):1,9–11, 2016.
- D Bilardello and K P Kodama. Rock magnetic evidence for inclination shallowing in the early Carboniferous Deer Lake Group red beds of western Newfoundland. *Geophys J Int*, 181(1):275–289, 2010.
- E Bullard, F.R.S., J E Everett, and G Smith. The fit of the continents around the Atlantic. *Phil Trans Roy Soc Lond Math Phys Sci*, 258:41–51, 1965.
- R F Butler. *Paleomagnetism: Magnetic Domains to Geologic Terranes*. Blackwell Scientific Publications, Malden, Massachusetts, electronic edition, 1992.
- S. C. Cande and D. V. Kent. Revised calibration of the geomagnetic polarity timescale for the late cretaceous and cenozoic. *Journal of Geophysical Research: Solid Earth*, 100(B4):6093–6095, 1995. doi: <https://doi.org/10.1029/94JB03098>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/94JB03098>.
- S C Cande and J M Stock. Pacific—Antarctic—Australia motion and the formation of the Macquarie Plate. *Geophysical Journal International*, 157(1):399–414, 04 2004. ISSN 0956-540X. doi: 10.1111/j.1365-246X.2004.02224.x. URL <https://doi.org/10.1111/j.1365-246X.2004.02224.x>.

S C Cande, P Patriat, and J Dyment. Motion between the Indian, Antarctic and African plates in the early Cenozoic. *Geophysical Journal International*, 183(1): 127–149, 10 2010. ISSN 0956-540X. doi: 10.1111/j.1365-246X.2010.04737.x. URL <https://doi.org/10.1111/j.1365-246X.2010.04737.x>.

M T Chandler, P Wessel, and W W Sager. Analysis of Ontong Java Plateau palaeolatitudes: evidence for large-scale rotation since 123 Ma? *Geophys J Int*, 194(1): 18–29, 2013. doi: 10.1093/gji/ggt075.

G L Christeson, H J A Van Avendonk, I O Norton, J W Snedden, D R Eddy, G D Karner, and C A Johnson. Deep crustal structure in the eastern Gulf of Mexico. *J Geophys Res Solid Earth*, 119(9):6782–6801, 2014. doi: 10.1002/2014JB011045.

R Chu, W Leng, D V Helmberger, and M Gurnis. Hidden hotspot track beneath the eastern United States. *Nature Geosci*, 6(11):963–966, 2013.

R D Cottrell and J A Tarduno. Late Cretaceous True Polar Wander: Not so fast. *Science*, 288(5475):2281–2281, 2000. ISSN 0036-8075. doi: 10.1126/science.288.5475.2281p. URL <https://science.sciencemag.org/content/288/5475/2281.16>.

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

C DeMets, E Calais, and S Merkouriev. Reconciling geodetic and geological estimates of recent plate motion across the Southwest Indian Ridge. *Geophysical Journal International*, 208(1):118–133, 10 2017. ISSN 0956-540X. doi: 10.1093/gji/ggw386. URL <https://doi.org/10.1093/gji/ggw386>.

M Domeier, R van der Voo, and F B Denny. 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(1-2):38–52, 2011a. doi: 10.1016/j.tecto.2011.08.016.

M Domeier, R Van der Voo, R N Tomezzoli, E Tohver, B W H Hendriks, T H Torsvik, H Vizan, and A Dominguez. 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):1–26, 2011b.

doi: 10.1029/2011JB008495. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JB008495>.

P V Doubrovine and J A Tarduno. Linking the Late Cretaceous to Paleogene Pacific plate and the Atlantic bordering continents using plate circuits and paleomagnetic data. *Journal of Geophysical Research: Solid Earth*, 113(B7):B07104, 2008. doi: 10.1029/2008JB005584. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JB005584>.

D A D Evans. True polar wander and supercontinents. *Tectonophysics*, 362(1):303–320, 2003. ISSN 0040-1951. doi: 10.1016/S0040-1951(02)000642-X. URL <http://www.sciencedirect.com/science/article/pii/S004019510200642X>. Paleomagnetism Applied to Tectonics. A tribute to Rob Van der Voo.

D A D Evans. Proterozoic low orbital obliquity and axial-dipolar geomagnetic field from evaporite palaeolatitudes. *Nature*, 444(7115):51–55, 2006. ISSN 0028-0836. doi: 10.1038/nature05203.

R A Fisher. Dispersion on a sphere. *Proc Roy Soc London Ser A*, 217:295–305, 1953. doi: 10.1098/rspa.1953.0064.

D R Franco. Paleomagnetic evidence for inverse correspondence between the relative contribution of the axial dipole field and CMB heat flux for the past 270 Myr. *Scientific Reports*, 9:282, 2019. doi: 10.1038/s41598-018-36494-x.

C Gaina, T H Torsvik, D J J van Hinsbergen, S Medvedev, S C Werner, and C Labails. The African Plate: A history of oceanic crust accretion and subduction since the Jurassic. *Tectonophysics*, 604:4–25, 2013. ISSN 0040-1951. doi: 10.1016/j.tecto.2013.05.037. URL <http://www.sciencedirect.com/science/article/pii/S004019511300351X>. Progress in understanding the South Atlantic margins.

C Gaina, D J J van Hinsbergen, and W Spakman. Tectonic interactions between India and Arabia since the Jurassic reconstructed from marine geophysics, ophiolite geology, and seismic tomography. *Tectonics*, 34(5):875–906, 2015. ISSN 1944-9194. doi: 10.1002/2014TC003780.

P Gautam and E Appel. Magnetic-polarity stratigraphy of Siwalik Group sediments of Tinau Khola section in west central Nepal, revisited. *Geophysical Journal International*, 117(1):223–234, 04 1994. ISSN 0956-540X. doi: 10.1111/j.1365-246X.1994.tb03314.x. URL <https://doi.org/10.1111/j.1365-246X.1994.tb03314.x>.

S E Geuna, L D Escosteguy, and C O Limarino. Paleomagnetism of the Carboniferous-Permian Patquia Formation, Paganzo basin, Argentina: implications for the apparent polar wander path for South America and Gondwana during the Late Palaeozoic. *Geol Acta*, 8(4):373–397, 2010.

Felix M. Gradstein, Frits P. Agterberg, James G. Ogg, Jan Hardenbol, Paul van Veen, Jacques Thierry, and Zehui Huang. A mesozoic time scale. *Journal of Geophysical Research: Solid Earth*, 99(B12):24051–24074, 1994. doi: <https://doi.org/10.1029/94JB01889>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/94JB01889>.

R Granot and J Dymant. Late Cenozoic unification of East and West Antarctica. *Nature Communications*, 9(1):3189, 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-05270-w.

B Greiner. Euler rotations in plate-tectonic reconstructions. *Comput Geosci*, 25(3): 209–216, 1999.

S J Hellinger. The uncertainties of finite rotations in plate-tectonics. *J Geophys Res*, 86(NB10):9312–9318, 1981.

J K Hillier. Pacific seamount volcanism in space and time. *Geophysical Journal International*, 168(2):877–889, 02 2007. ISSN 0956-540X. doi: 10.1111/j.1365-246X.2006.03250.x. URL <https://doi.org/10.1111/j.1365-246X.2006.03250.x>.

B C Horner-Johnson, R G Gordon, S M Cowles, and D F Argus. The angular velocity of Nubia relative to Somalia and the location of the Nubia—Somalia—Antarctica triple junction. *Geophysical Journal International*, 162(1):221–238, 07 2005. doi: 10.1111/j.1365-246X.2005.02608.x.

E H Jacox and H Samet. Spatial join techniques. *ACM Trans. Database Syst.*, 32:7, 2007. doi: 10.1145/1206049.1206056.

J T Kent. The Fisher-Bingham distribution on the sphere. *J Roy Stat Soc B*, 44(1): 71–80, 1982.

R F King. The remanent magnetism of artificially deposited sediments. *Geophys Suppl Mon Not Roy Astron Soc Lett*, 7(3):115–134, 1955.

W Krijgsman and L Tauxe. Shallow bias in mediterranean paleomagnetic directions caused by inclination error. *Earth and Planetary Science Letters*, 222(2):685–695, 2004. ISSN 0012-821X. doi: 10.1016/j.epsl.2004.03.007.

E V Kulakov, A V Smirnov, and J F Diehl. Paleomagnetism of the ~1.1 Ga Coldwell Complex (Ontario, Canada): Implications for Proterozoic geomagnetic field morphology and plate velocities. *J Geophys Res Solid Earth*, 119(12):8633–8654, 2014. ISSN 2169-9356. doi: 10.1002/2014JB011463.

II Lemaux, J, R G Gordon, and J-Y Royer. Location of the Nubia-Somalia boundary along the Southwest Indian Ridge. *Geology*, 30(4):339–342, 04 2002. ISSN 0091-7613. doi: 10.1130/0091-7613(2002)030<0339:LOTNSB>2.0.CO;2. URL [https://doi.org/10.1130/0091-7613\(2002\)030<0339:LOTNSB>2.0.CO;2](https://doi.org/10.1130/0091-7613(2002)030<0339:LOTNSB>2.0.CO;2).

Y Ma, T Yang, W Bian, J Jin, S Zhang, H Wu, and H Li. Early Cretaceous paleomagnetic and geochronologic results from the Tethyan Himalaya: Insights into the Neotethyan paleogeography and the India–Asia collision. *Scientific Reports*, 6: 21605, 2016. doi: 10.1038/srep21605.

K.M Marks and A.A Tikku. Cretaceous reconstructions of East Antarctica, Africa and Madagascar. *Earth and Planetary Science Letters*, 186(3):479–495, 2001. ISSN 0012-821X. doi: 10.1016/S0012-821X(01)00262-X. URL <http://www.sciencedirect.com/science/article/pii/S0012821X0100262X>.

M W McElhinny and J Lock. Iaga paleomagnetic databases with access. *Surveys in Geophysics*, 17(5):575–591, 1996.

P L McFadden and M W McElhinny. Classification of the reversal test in palaeomagnetism. *Geophysical Journal International*, 103(3):725–729, 12 1990. ISSN 0956-540X. doi: 10.1111/j.1365-246X.1990.tb05683.x.

D McKenzie and J G Sclater. The evolution of the indian ocean since the late cretaceous. *Geophysical Journal of the Royal Astronomical Society*, 24(5):437–528, 1971. doi: 10.1111/j.1365-246X.1971.tb02190.x. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-246X.1971.tb02190.x>.

N McQuarrie and B P Wernicke. An animated tectonic reconstruction of southwestern North America since 36 Ma. *Geosphere*, 1(3):147–172, 2006. doi: 10.1130/GES00016.1.

R D Müller, J Y Royer, and L A Lawver. Revised plate motions relative to the hotspots from combined Atlantic and Indian-Ocean hotspot tracks. *Geology*, 21(3):275–278, 1993.

R D Müller, J Y Royer, S C Cande, W R Roest, and S Maschenkov. *New constraints on the Late Cretaceous/Tertiary plate tectonic evolution of the Caribbean*, volume 4, Chap. 2, pages 33–59. Elsevier, 1999.

R D Müller, M Sdrolias, C Gaina, and W R Roest. Age, spreading rates, and spreading asymmetry of the world’s ocean crust. *Geochem Geophys Geosyst*, 9(4):Q04006, 2008.

R D Müller, M Seton, S Zahirovic, S E Williams, K J Matthews, N M Wright, G E Shephard, K Maloney, N Barnett-Moore, M Hosseinpour, D J Bower, and J Cannon. Ocean basin evolution and global-scale plate reorganization events since Pangea breakup. *Annu Rev Earth Planet Sci*, 44(1):107–138, 2016.

Y Najman, E Appel, M Boudagher-Fadel, P Bown, A Carter, E Garzanti, L Godin, J Han, U Liebke, G Oliver, R Parrish, and G Vezzoli. Timing of India-Asia collision: Geological, biostratigraphic, and palaeomagnetic constraints. *Journal of Geophysical Research: Solid Earth*, 115(B12):B12416, 2010. doi: 10.1029/2010JB007673. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JB007673>.

C O’Neill, R D Müller, and B Steinberger. On the uncertainties in hot spot reconstructions and the significance of moving hot spot reference frames. *Geochem Geophys Geosyst*, 6(4):Q04003, 2005.

N D Opdyke, N M Johnson, G D Johnson, E H Lindsay, and R A K Tahirkheli. Paleomagnetism of the middle siwalik formations of northern pakistan and rotation of the salt range decollement. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 37(1):1–15, 1982. ISSN 0031-0182. doi: 10.1016/0031-0182(82)90055-4. URL <http://www.sciencedirect.com/science/article/pii/0031018282900554>. The geochronology and biochronology of the Siwalik Group, Pakistan.

E Pakyuz-Charrier, M Lindsay, V Ogarko, J Giraud, and M Jessell. Monte Carlo simulation for uncertainty estimation on structural data in implicit 3-D geological modeling, a guide for disturbance distribution selection and parameterization. *Solid Earth*, 9(2):385–402, 2018. doi: 10.5194/se-9-385-2018. URL <https://www.solid-earth.net/9/385/2018/>.

P Patriat, H Sloan, and D Sauter. From slow to ultraslow: A previously undetected event at the Southwest Indian Ridge at ca. 24 Ma. *Geology*, 36(3):207–210, 03 2008. ISSN 0091-7613. doi: 10.1130/G24270A.1. URL <https://doi.org/10.1130/G24270A.1>.

S A Pisarevsky. New edition of the Global Paleomagnetic Database. *Eos Trans AGU*, 86(17):170, 2005.

S A Pisarevsky and M W McElhinny. Global Paleomagnetic Data Base developed into its visual form. *Eos Trans AGU*, 84(20):192, 2003.

P Riisager, S Hall, M Antretter, and X Zhao. Early Cretaceous Pacific palaeomagnetic pole from Ontong Java Plateau basement rocks. *Geological Society, London, Special Publications*, 229(1):31–44, 2004. ISSN 0305-8719. doi: 10.1144/GSL.SP.2004.229.01.04. URL <https://sp.lyellcollection.org/content/229/1/31>.

C J Rowan and D B Rowley. Preserved history of global mean spreading rate: 83 Ma to present. *Geophysical Journal International*, 208(2):1173–1183, 07 2016. doi: 10.1093/gji/ggw277.

W W Sager. Divergence between paleomagnetic and hotspot-model–predicted polar wander for the Pacific plate with implications for hotspot fixity. *GSA Spec Paper*, 430:335–357, 2007. doi: 10.1130/2007.2430(17).

A Schettino and C R Scotese. 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(2):727–759, 2005.

M Seton, R D Müller, S Zahirovic, C Gaina, T Torsvik, G Shephard, A Talsma, M Gurnis, M Turner, S Maus, and M Chandler. Global continental and ocean basin reconstructions since 200 Ma. *Earth Sci Rev*, 113(3-4):212–270, 2012.

G E Shephard, H P Bunge, B S A Schuberth, R D Müller, A S Talsma, C Moder, and T C W Landgrebe. Testing absolute plate reference frames and the implications for the generation of geodynamic mantle heterogeneity structure. *Earth Planet Sci Lett*, 317-318(0):204–217, 2012.

B Steinberger and R J O’Connell. Advection of plumes in mantle flow: implications for hotspot motion, mantle viscosity and plume distribution. *Geophysical Journal International*, 132(2):412–434, 02 1998. ISSN 0956-540X. doi: 10.1046/j.1365-246x.1998.00447.x. URL <https://doi.org/10.1046/j.1365-246x.1998.00447.x>.

B Steinberger and T H Torsvik. Absolute plate motions and true polar wander in the absence of hotspot tracks. *Nature*, 452(7187):620–623, 2008.

N L Swanson-Hysell, A C Maloof, B P Weiss, and D A D Evans. No asymmetry in geomagnetic reversals recorded by 1.1-billion-year-old Keweenawan basalts. *Nature Geosci*, 2(10):713–717, 2009. doi: 10.1038/ngeo622.

J Tarduno, H-P Bunge, N Sleep, and U Hansen. The bent Hawaiian-Emperor hotspot track: Inheriting the mantle wind. *Science*, 324(5923):50–53, 2009.

J A Tarduno. On the motion of Hawaii and other mantle plumes. *Chem Geol*, 241 (3-4):234–247, 2007.

L Tauxe and D V Kent. *A Simplified Statistical Model for the Geomagnetic Field and the Detection of Shallow Bias in Paleomagnetic Inclinations: was the Ancient Magnetic Field Dipolar?*, pages 101–115. American Geophysical Union (AGU), 2004. ISBN 9781118665855. doi: 10.1029/145GM08. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/145GM08>.

L Tauxe, N Kylstra, and C Constable. Bootstrap statistics for paleomagnetic data. *J Geophys Res Solid Earth*, 96(B7):11723–11740, 1991. doi: 10.1029/91JB00572.

L Tauxe, S K Banerjee, R F Butler, and R van der Voo. Essentials of paleomagnetism: Fifth web edition. <http://earthref.org/MAGIC/books/Tauxe/Essentials/>, 2019.

T H Torsvik and L R M Cocks. *Earth History and Palaeogeography*. Cambridge University Press, Cambridge, 2016. doi: 10.1017/9781316225523.

T H Torsvik and M A Smethurst. Plate tectonic modelling: virtual reality with GMAP. *Comput Geosci*, 25(4):395–402, 1999.

T H Torsvik, M A Smethurst, R van der Voo, A Trench, N Abrahamsen, and E Halvorsen. Baltica. A synopsis of Vendian-Permian palaeomagnetic data and their palaeotectonic implications. *Earth Sci Rev*, 33(2):133–152, 1992.

T H Torsvik, R D Müller, R van der Voo, B Steinberger, and C Gaina. Global plate motion frames: Toward a unified model. *Rev Geophys*, 46(3):RG3004, 2008. doi: 10.1029/2007RG000227.

T H Torsvik, R van der Voo, U Preeden, C Mac Niocaill, B Steinberger, P V Doubrovine, D J J van Hinsbergen, M Domeier, C Gaina, E Tohver, J G Meert, P J A McCausland, and L R M Cocks. Phanerozoic polar wander, palaeogeography and dynamics. *Earth Sci Rev*, 114(3-4):325–368, 2012. doi: 10.1016/j.earscirev.2012.06.007.

D G van der Meer, W Spakman, D J J van Hinsbergen, M L Amaru, and T H Torsvik. Towards absolute plate motions constrained by lower-mantle slab remnants. *Nat Geosci*, 3(1):36–40, 2010.

R van der Voo. Paleozoic paleogeography of North America, Gondwana, and intervening displaced terranes: Comparisons of paleomagnetism with paleoclimatology and biogeographical patterns. *GSA Bulletin*, 100(3):311–324, 1988. ISSN 0016-7606. doi: 10.1130/0016-7606(1988)100<0311:PPONAG>2.3.CO;2. URL [https://doi.org/10.1130/0016-7606\(1988\)100<0311:PPONAG>2.3.CO;2](https://doi.org/10.1130/0016-7606(1988)100<0311:PPONAG>2.3.CO;2).

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

S van der Walt, S C Colbert, and G Varoquaux. The NumPy array: A structure for efficient numerical computation. *Comput Sci Eng*, 13(2):22–30, 2011.

T Veikkolainen, L Pesonen, and D D Evans. PALEOMAGIA: A PHP/MYSQL database of the Precambrian paleomagnetic data. *Studia Geophysica et Geodaetica*, pages 1–17, 2014.

P Wessel, R D Müller, and G Schubert. *Plate Tectonics*, book section 6.02, pages 49–98. Elsevier, Amsterdam, 2007.

J M Whittaker, S E Williams, and R D Müller. Revised tectonic evolution of the Eastern Indian Ocean. *Geochemistry, Geophysics, Geosystems*, 14(6):1891–1909, 2013. doi: 10.1002/ggge.20120. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/ggge.20120>.

A Young, N Flament, K Maloney, S Williams, K Matthews, S Zahirovic, and R D Müller. Global kinematics of tectonic plates and subduction zones since the late Paleozoic Era. *Geoscience Frontiers*, 10(3):989–1013, 2019. ISSN 1674-9871. doi: <https://doi.org/10.1016/j.gsf.2018.05.011>. URL <http://www.sciencedirect.com/science/article/pii/S1674987118301373>. Special Issue: Advances in Himalayan Tectonics.

# Appendices

# Appendix A

## Methods for Manipulating Data on A Globe in Chapter 2

### A.1 Test if A Coeval Pole Pair Is Distinguishable with the Bootstrap

Whether a pair of coeval mean poles are statistically distinguishable from each other is investigated, as it can be determined by checking if the confidence intervals of their bootstrap means (based on two poles' uncertainty attributes) overlap (Tauxe et al., 1991). The 95% confidence bounds of the Cartesian coordinates of the bootstrap means are determined and compared. If the poles are distinguishable, the confidence bounds along at least one coordinate axis do not overlap. Otherwise, if the confidence bounds along all the three coordinate axes overlap, the poles are indistinguishable (Tauxe et al., 1991). The actual test used is dependent on the number of paleopoles ( $N$ ) used to calculate the mean pole in the APWP:

- when  $N > 25$  a simple bootstrap (Tauxe et al., 1991) generates a pseudo-mean pole from  $N$  directions drawn randomly from the original set of paleopoles. 1000 such simple Bootstraps are implemented here.
- when  $1 < N \leq 25$  a parametric bootstrap (Tauxe et al., 1991) generates a pseudo-mean from  $N$  directions drawn from a Fisher distribution with the same  $K$  and  $N$  as the mean pole. 1000 such parametric Bootstraps are implemented here.
- when  $N = 1$ , a pseudo-mean is drawn from a bivariate normal distribution, defined by the properties of the associated A95 uncertainty circle or dm/dp ellipse (see the following Section A.2). Here 1000 samples are drawn from such a normal distribution.

- if  $\mathbf{N}$  is not given, because for example sometimes the pole could be an interpolated result, a negligible A95 like  $0.1^\circ$  or  $0^\circ$  is assigned and the same sampling way as used for the  $\mathbf{N} = 1$  case is applied here. This is for the situation when only one of the coeval poles is interpolated, and one would like to keep this pair of poles. Note that if the coeval poles are both interpolated, we suggest directly removing this pair of poles.

**Special cases** Sometimes, like in the cases in Fig. 3 and Fig. 10, we have complete access to the parameters of the mean poles, e.g.  $\mathbf{N}$  and precision parameter  $\mathbf{K}$ , and also the paleopoles. However this is not necessarily true. If, for instance, we only have access to the path with only its mean poles and spatial uncertainties, we can keep the way of doing bootstrap sampling consistent for all the mean poles, and just draw bootstrapped means from a bivariate normal distribution based on each spatial uncertainty’s geometry. This is implemented through arbitrarily setting  $\mathbf{N}=1$ . The consistency of bootstrap sampling makes it independent of the state of knowledge of the underlying dataset and even the underlying method used to calculate the uncertainty. This means the method can be more generalisable beyond APWPs, because the metrics and the significance testing procedure are more broadly applicable to comparison of other trajectories with associated spatial uncertainties, such as hurricane tracks and bird migration routes.

The final results for each coeval pole pair of all the seven APWP pairs (Fig. 3), are given in the sub-folder “0.result\_tables”, which is contained in the main “data” folder. The results for length and angular differences are listed starting from the rows for the second and third poles respectively, simply because one pole can not compose a APWP segment and at least three poles could constitute an APWP orientation change.

## A.2 Bivariate Sampling

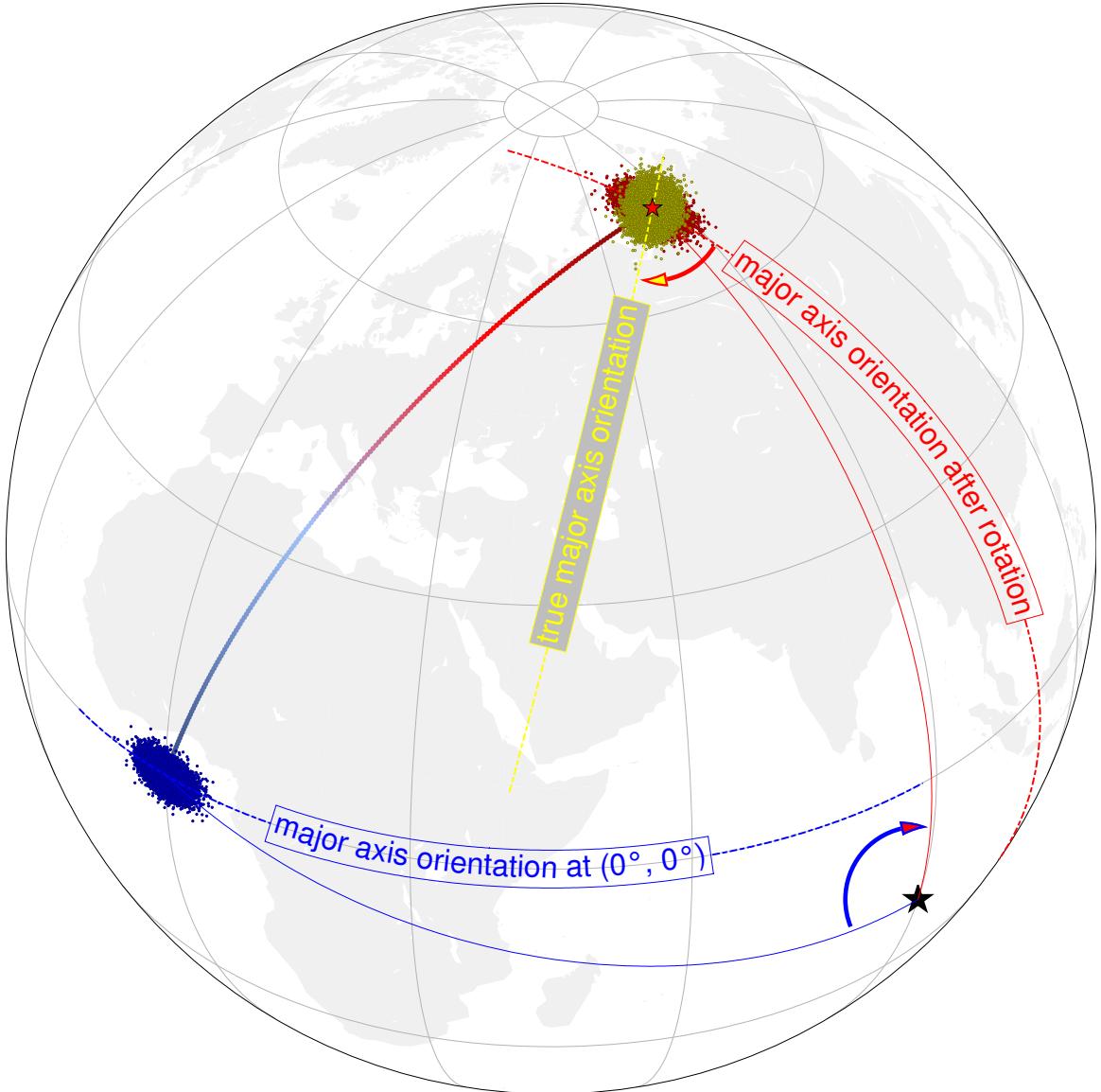
For some specific poles of the APWP, e.g., only one paleomagnetic pole makes up that “mean”, i.e.,  $N = 1$ , or even there is no paleomagnetic pole in that specific bin (i.e.,  $N = 0$ ) but an interpolated pole that might be given by authors at that specific age, the bivariate normal distribution is used to generate random samples based on its uncertainty ellipse’s semi axes and the major axis’ azimuth, then we use the cumulative distributions of Cartesian coordinates of those random samples to see if the confidence intervals overlap.

However, here the scenario is not a two dimensional (2D) domain, but rather a spherical surface. Directly simulating random points for an ellipse on a sphere is a complicated problem (Kent, 1982). An analogue approach is proposed here as follows. First, random points of a 2D bivariate normal distribution are generated with NumPy’s random sampling routine “`multivariate_normal`” (van der Walt et al., 2011). The lengths of the uncertainty ellipse’s semi-major and semi-minor axes are used as about 1.96 standard deviations of the bivariate normal distribution. The center of the ellipse is located at the intersection of the equator ( $0^\circ$  latitude) and the prime meridian ( $0^\circ$  longitude) with its major axis lying equator-ward (blue point cloud in Fig. A.1). Then according to the actual pole coordinates (red star in Fig. A.1), an Euler rotation (Greiner, 1999) (black star and blue angle arc) can be calculated along the great circle (progressing from blue to red) from the location ( $0^\circ, 0^\circ$ ) to the actual pole location. After those random points (blue points) are rotated using the same Euler rotation to the new locations (red point cloud in Fig. A.1), this elliptical cloud (red point cloud) then is adjusted to its actual azimuth (i.e., the major-axis azimuth of the pole’s uncertainty ellipse; the red dashed line rotated to the yellow dashed line using the red star as the Euler pole shown in Fig. A.1).

Note that directly using NumPy’s “`random.multivariate_normal`” or “`random.normal`” routine (2D calculations) and spherical trigonometry to draw random points for an elliptical uncertainty distorts the point cloud out of a bivariate normal distribution, especially at high-latitude areas (Pakyuz-Charrier et al., 2018, see the examples given by their Figure 7) and makes the simulation inaccurate. This analogue approach avoids producing declination and inclination vectors beforehand and directly generates random pole vectors, which saves the transformation from declination and inclination to pole and further helps keep us away from the distortion.

### A.3 Synchronization

This algorithm is developed for comparing time-synchronized APWPs. In other words, the compared APWPs should have the same timestamps. If the number of their timestamps are different, the unpaired *pole(s)* would be removed to make the timestamps the same before the comparison. APWPs with a pole interpolated for pairing an unpaired pole can be processed by our tool, as we noted earlier, but it is not recommended for a valid analysis. For example, for paleomagnetic APWPs, sometimes there are no paleopoles for a given time window ( $N = 0$ ); sometimes a mean pole is an interpolated result.

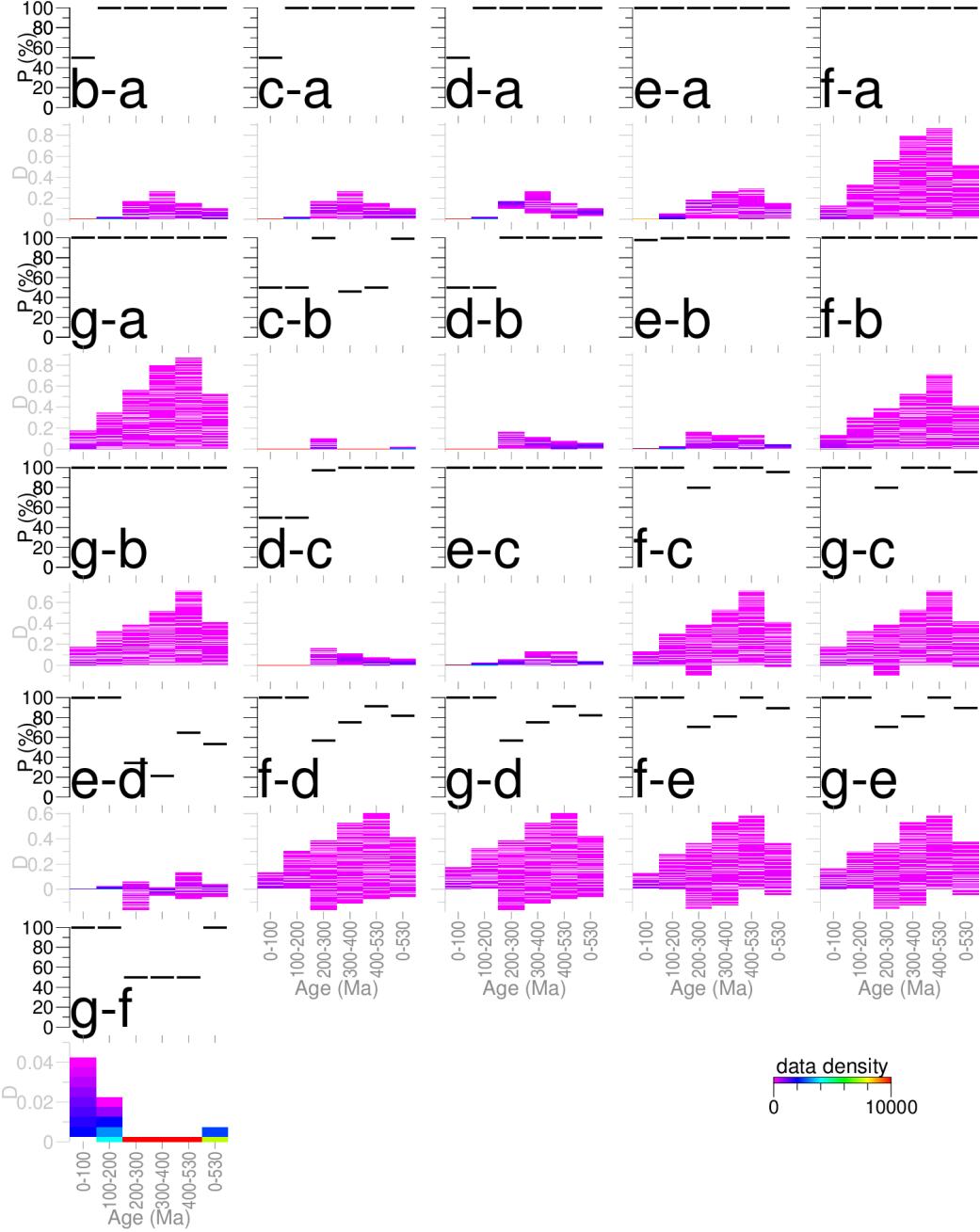


**Figure A.1:** Example of modeling random points for an ellipse uncertainty on the Earth’s surface. Sample points (blue) from a bivariate normal distribution centered at the intersection of the equator and the prime meridian are rotated to their new locations (red points) together with the uncertainty ellipse center (i.e. the  $0^\circ$  longitude  $0^\circ$  latitude point prior to the rotation) exactly rotated to its actual pole coordinate (red star), then adjusted to the true orientation (yellow dashed line).

### A.3.1 Equally Treated Random Weights

Assigning equally likely (not necessarily equal in value) random values to  $W_s, W_a, W_l$  is also tested. Three uniformly distributed random numbers with a given sum 1 are generated for, for example, 10 000 times here, and then are substituted into the CPD formula for deriving the seven APWP pairs’  $D_{full}, D_{0-100Ma}$  etc. to check the

possibility that “one pair is superior to the other pair” (Fig. A.2).



**Figure A.2:** Differences of CPDs between *Pair a*, *Pair b*, *Pair c*, *Pair d*, *Pair e*, *Pair f* and *Pair g*, when 10 000 sets of three uniformly random weights (with their sum 1) are applied. If the difference  $D$  is positive, the subtrahend pair ranks higher in similarity, and if it is negative, the minuend pair ranks higher. The  $y$  axis in each upper plot is for the percentage  $P$  that the subtrahend pair owns higher similarity.

The full-path results (Fig. A.2) again re-verify Order (4) and the results shown in Fig. 8. Although the possibility that *Pair d* is more similar than *Pair e* is

not significant (around 50%), the possibility that *Pairs f,g* are more similar than *Pairs c,d,e* is significant (more than 95%),

All the sub-path results (Fig. A.2) are explicable using the results shown in Fig. 8. For example, for 0–100 Ma, both *Pair a* and *Pair b* are assigned values of zero for all the three metrics  $d_s^{0-100Ma}$ ,  $d_l^{0-100Ma}$  and  $d_a^{0-100Ma}$  (Figures 8b, 8d and 8f), which means they are always undifferentiated.

# Appendix B

## Methods for Constraining Paleopoles for Tectonic Plates in Chapter 3

(All the datasets used for Chapter 3 are openly accessible from [https://github.com/f-i/making\\_of\\_reliable\\_APWPs](https://github.com/f-i/making_of_reliable_APWPs).)

A polygon can be drawn around a set of paleomagnetic data, whose sampling sites we believe belong to a specific plate or rigid block. Each such plate tectonic polygon is assigned with an unique identification code, which is called Plate ID. Then the *Spatial Join* technique (Jacox and Samet, 2007) helps join attributes, especially Plate ID, 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.

### B.1 120–0 Ma North America

The data-constraining polygons are from the recently published plate model (Young et al., 2019) (Fig. B.1). North American craton (Plate ID 101) polygon in the plate model (Young et al., 2019), 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

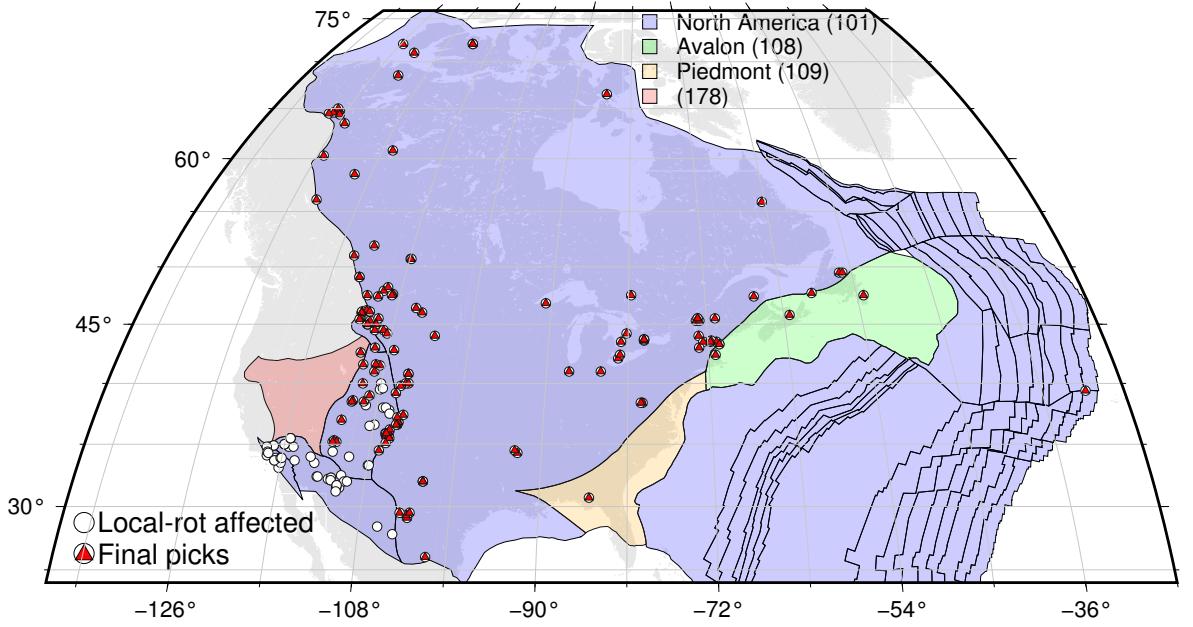
et al., 2019), 108 is fixed to 101 during the geologic period from Cretaceous to the present day. 109 is also fixed to 101 since  $\sim$ 300 Ma (Christeson et al., 2014). Then in order to be compared with the FHM (120–0 Ma) (Müller et al., 1993, 1999), the paleopoles with age ranging 120–0 Ma are further selected through constraining the lower magnetic age “LOMAGAGE  $\leq$  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 and 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. B.1).

Also based on this model of southwestern North America since 36 Ma (McQuarrie and Wernicke, 2006), part of the paleopoles constrained by the four small western terranes whose Plate IDs are also 101 (white circles in Fig. B.1) 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 and Wernicke, 2006)) should be included. So finally 135 of the 193 datasets remain (Fig. B.1). Spatially North American paleomagnetic data are mainly from the western and eastern margins of the plate.

## B.2 120–0 Ma India

Plate ID 501 polygons in the recently published Plate Model (Young et al., 2019) 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. B.2). The polygons are used to select the sampling sites of the paleopoles for India (Fig. B.2).

Based on the model of the tectonic interactions between India, Arabia and Asia since the Jurassic (Gaina et al., 2015) (Fig. B.2), 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. B.2). Spatially Indian paleomagnetic data are more evenly distributed on the India plate than North American and Australian poles.

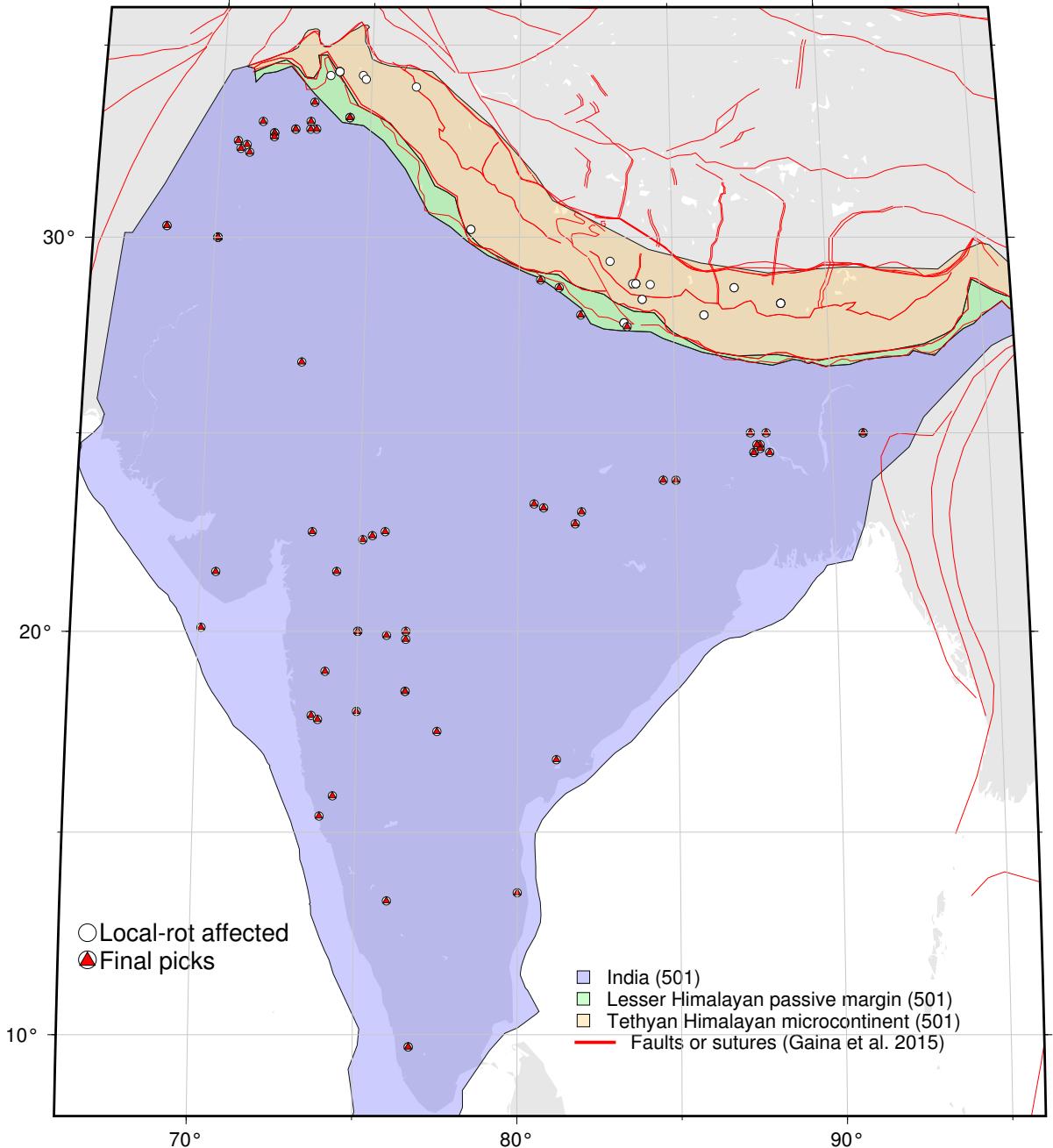


**Figure B.1:** The final filtered datasets (red triangle-inside-circles) for later analysis on 120–0 Ma North America. Those poles that had been influenced by local tectonic rotations are shown as white circles.

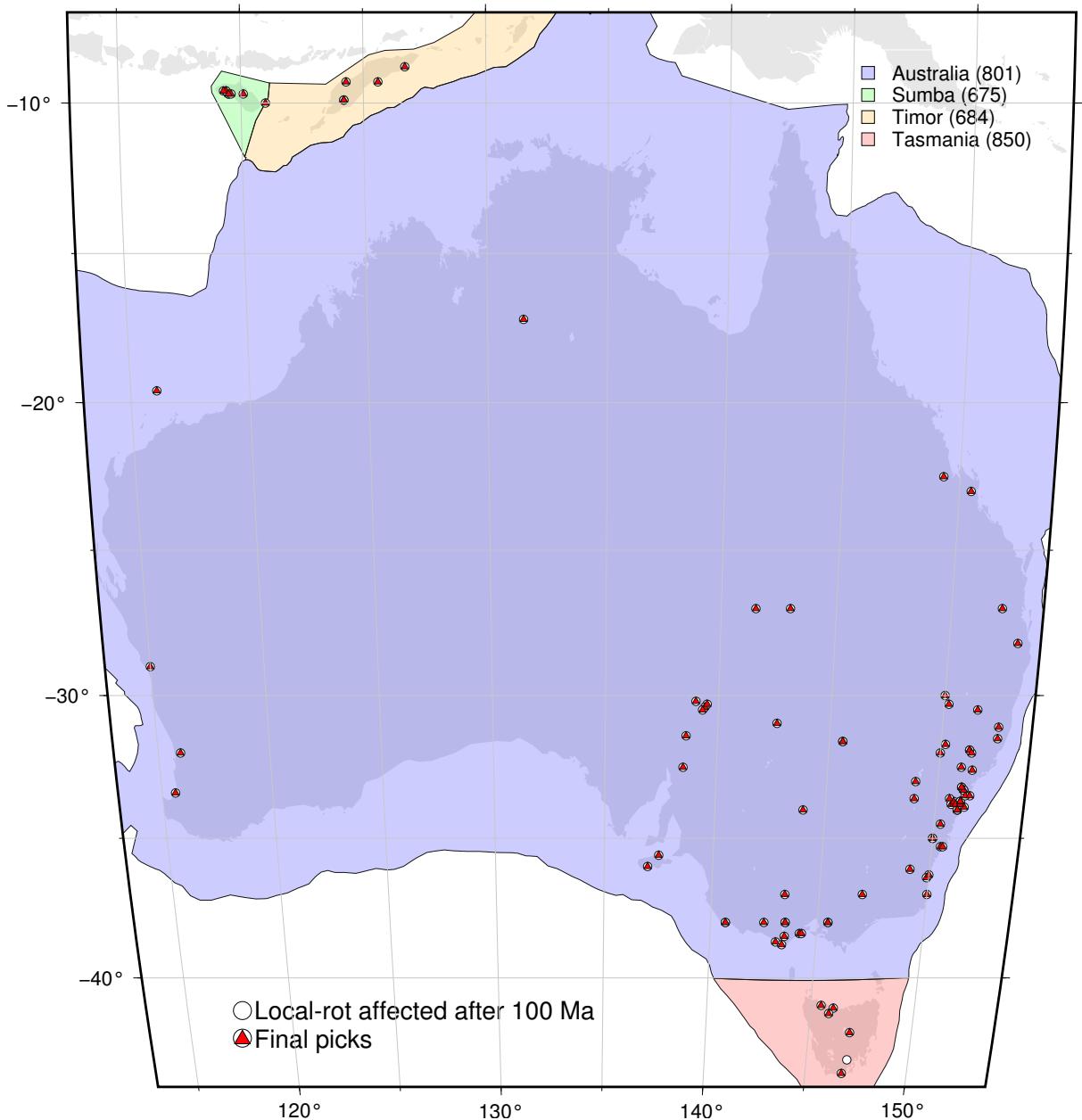
### B.3 120–0 Ma Australia

Plate ID 801 polygon in the recently published Plate Model (Young et al., 2019), including its children 675 (Sumba block) and 684 (Timor block) polygons for 120–0 Ma (Fig. B.3), is used to select the sampling sites of the paleopoles for Australia. According to the plate model rotation data (Young et al., 2019), 675 and 684 are fixed to 801 during the geologic period from  $\sim$ 145 Ma to the present.

On the southeast of the main Australia plate (the blue polygon in Fig. B.3), there is a triangle-shaped small polygon 850 (Tasmania block) which is fixed to 801 since  $\sim$ 100 Ma according to the Young et al. (2019) rotation data. With that attribute, 850 contributes more data younger than  $\sim$ 100 Ma for the later analysis. Ultimately the final 99 extracted datasets is shown in Fig. B.3.



**Figure B.2:** 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 B.3:** 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.

**Table B.1:** Rotation parameters for making three reference paths

Moving-Fixed	Chron	Age/Ma	Latitude	Longitude	Rotation	$\hat{\kappa}$	a	b	c	d	e	f	Source
NAm-Nubia	Ino	0.78	79.2	40.2	0.18	1	7.41E-9	-5.77E-9	4.29E-9	5.93E-9	-3.35E-9	5.15E-9	Demets et al. 2010
NAm-Nubia(NWA)	2An	2.7	78.8	38.3	0.65	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Cande et al. 1995 Shephard et al. EPSL 2012 Gurnis et al. CG2012
NAm-Nubia(NWA)	5n.1ny	9.74	80.98	22.82	2.478	2.46	4.72E-5	-4.05E-5	2.79E-5	4.35E-5	-2.9E-5	2.06E-5	Muller et al. 1999
NAm-Nubia(NWA)	5n.2o	10.949	81	22.9	2.85	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Gaina et al. 2013
NAm-Nubia(NWA)	6ny	19.05	80.89	23.28	5.244	1.96	6.13E-5	-4.97E-5	3.56E-5	5.2E-5	-3.59E-5	2.68E-5	Muller et al. 1999
NAm-Nubia(NWA)	6no	20.131	80.6	24.4	5.54	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Gaina et al. 2013
NAm-Nubia(NWA)	8n.1ny	25.82	79.34	28.56	7.042	2.53	1.71E-4	-1.71E-4	1.24E-4	1.96E-4	-1.41E-4	1.04E-4	Muller et al. 1999
NAm-Nubia(NWA)	13ny	33.058	75.99	5.98	9.767	1.19	8.32E-5	-8.47E-5	5.97E-5	1.03E-4	-7.14E-5	5.21E-5	Muller et al. 1999 Gaina et al. 2013
NAm-Nubia(NWA)	16ny	35.343	75.3	3.1	10.68	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Gaina et al. 2013
NAm-Nubia(NWA)	18n1ny?	38.43	74.54	0.19	11.918	1.65	1.98E-4	-2.08E-4	1.43E-4	2.41E-4	-1.64E-4	1.16E-4	Muller et al. 1999
NAm-Nubia(NWA)	18n.2no	40.13	74.5	-1.2	12.61	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Gaina et al. 2013
NAm-Nubia(NWA)	20ny	42.5	74.38	-2.8	13.56	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Muller et al. 1993 Shephard et al. 2012
NAm-Nubia(NWA)	20no	43.789	74.3	-3.6	14.09	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Gaina et al. 2013
NAm-Nubia(NWA)	21ny	46.26	74.23	-5.01	15.106	1.19	1.8E-4	-2.03E-4	1.36E-4	2.68E-4	-1.8E-4	1.26E-4	Muller et al. 1999
NAm-Nubia(NWA)	21no	47.906	74.9	-4.6	15.61	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Gaina et al. 2013
NAm-Nubia(NWA)	22ny	49.037	75.29	-4.26	15.95	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Muller et al. 1993 Shephard et al. 2012
NAm-Nubia(NWA)	22no	49.714	75.9	-3.5	16.15	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Gaina et al. 2013
NAm-Nubia(NWA)	24n.1ny	52.364	77.34	-1.61	16.963	3.08	2.65E-4	-3.13E-4	2.07E-4	4.05E-4	-2.67E-4	1.8E-4	Muller et al. 1999 Gaina et al. 2013
NAm-Nubia(NWA)	25ny	55.904	80.64	6.57	17.895	1.26	1.19E-4	-1.36E-4	8.64E-5	1.87E-4	-1.18E-4	7.87E-5	Muller et al. 1999 Gaina et al. 2013
NAm-Nubia(NWA)	30ny	65.58	82.74	2.93	20.84	1.07	9.12E-5	-1.07E-4	6.64E-5	1.58E-4	-9.83E-5	6.45E-5	Muller et al. 1999
NAm-Nubia(NWA)	31ny	67.735	82.3	-1.8	21.53	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Gaina et al. 2013
NAm-Nubia(NWA)	32n.2ny	71.59	81.35	-8.32	22.753	1.33	1.23E-4	-1.62E-4	9.7E-5	2.42E-4	-1.46E-4	9.1E-5	Muller et al. 1999
NAm-Nubia(NWA)	33ny	73.619	81.11	-10.64	23.74	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Muller et al. 1993Shephard et al. 2012
NAm-Nubia(NWA)	33no	79.075	78.64	-18.16	26.981	1.85	4.09E-5	-4.02E-5	2.34E-5	5.73E-5	-3.48E-5	2.38E-5	Muller et al. 1999
NAm-Nubia(NWA)	34ny	83.5	76.81	-20.59	29.506	1.82	6.2E-5	-4.83E-5	2.47E-5	6.96E-5	-4.29E-5	3.07E-5	Gaina et al. 2013
NAm-Nubia(NWA)	?	89.9	74.33	-22.65	33.86	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Muller,Roest 1992 Muller et al. 2008 Shephard et al. 2012
NAm-Nubia(NWA)	?	94.1	72.0	-24.39	36.49	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Muller et al. 2008 Shephard et al. 2012

Table B.1 continued from previous page

Moving-Fixed	Chron	Age/Ma	Latitude	Longitude	Rotation	$\hat{\kappa}$	a	b	c	d	e	f	Source
NAm-Nubia(NWA)	?	100	69.42	-23.52	40.46	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Muller_Roest1992 Muller et al. 2008 Shephard et al. 2012
NAm-Nubia(NWA)	?	106.9	68.08	-22.66	45.36	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Muller_Roest1992 Muller et al. 2008 Shephard et al. 2012
NAm-Nubia(NWA)	?	118.1	66.21	-21.0	53.19	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Muller_Roest1992 Muller et al. 2008 Shephard et al. 2012
NAm-Nubia(NWA)	34no interpol	120.6 0	65.9 57.728	-20.5 -51.2596	54.56 0	1	1E-15 1E-15	1E-15 1E-15	1E-15 1E-15	1E-15 1E-15	1E-15 1E-15	1E-15 1E-15	Gaina et al. 2013 O'Neill et al. 2005fixed-Hotspots
Nubia-Mantle	interpol	10	57.728	-51.2596	-1.3645	3.3043	9.495E-5	-6.072E-6	1.085E-5	2.8857E-5	-2.436E-5	6.987E-5	O'Neill et al. 2005fixed-Hotspots
Nubia-Mantle	interpol	20	40.412	-23.4948	-3.611	2.545	1.577E-4	4.682E-6	2.293E-5	8.242E-5	-6.609E-5	1.49E-4	O'Neill et al. 2005fixed-Hotspots
Nubia-Mantle	interpol	30	49.42	-61.8981	-5.795	3.172	8.149E-5	2.568E-5	-6.617E-6	8.819E-5	-5.922E-5	8.762E-5	O'Neill et al. 2005fixed-Hotspots
Nubia-Mantle	interpol	40	44.555	-55.735	-7.4188	0.7337	7.166E-4	-5.577E-5	-2.8646E-4	1.5166E-4	-5.034E-5	2.725E-4	O'Neill et al. 2005fixed-Hotspots
<sup>13</sup> <sub>35</sub> Nubia-Mantle	interpol	50	8.836	-34.845	-12.431	0.826	6.97E-4	2.808E-5	-3.607E-4	1.587E-4	-1.04E-4	3.44E-4	O'Neill et al. 2005fixed-Hotspots
Nubia-Mantle	interpol	60	14.244	-40.32	-13.725	2.081	1.496E-3	3.451E-4	-5.396E-4	7.096E-4	-2.723E-4	5.502E-4	O'Neill et al. 2005fixed-Hotspots
Nubia-Mantle	interpol	70	26.518	-47.4563	-14.604	0.65	8.564E-4	1.195E-4	-3.269E-4	2.863E-4	-1.276E-4	4.452E-4	O'Neill et al. 2005fixed-Hotspots
Nubia-Mantle	interpol	80	26.85	-46.9249	-17.705	0.473	7.972E-4	1.314E-4	-3.436E-4	2.874E-4	-1.396E-4	3.485E-4	O'Neill et al. 2005fixed-Hotspots
Nubia-Mantle	interpol	90	7.2356	-33.7327	-22.745	0.692	6.805E-4	2.6E-5	-1.336E-4	1.834E-4	-2.302E-4	6.163E-4	O'Neill et al. 2005fixed-Hotspots
Nubia-Mantle	interpol	100	10.826	-46.482	-26.59	1.308	3.705E-4	2.897E-6	5.225E-5	4.458E-4	3.313E-5	3.037E-4	O'Neill et al. 2005fixed-Hotspots
Nubia-Mantle	interpol	110	5.063	-34.8933	-31.787	0.475	6.694E-4	1.7826E-4	-1.243E-5	1.875E-4	6.704E-5	4.335E-4	O'Neill et al. 2005fixed-Hotspots
Nubia-Mantle	interpol	120	22.488	-37.2224	-30.632	66.181	1.436E-3	4.318E-5	1.78E-3	1E-3	1.134E-3	4.44E-3	O'Neill et al. 2005fixed-Hotspots
Nubia-Mantle	None	0	46.19	-87.86	0	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	O'Neill et al. 2005moving
Nubia-Mantle	None	10	46.19	-87.86	-1.92	0.92	9.94E-5	-3.8E-6	-1.18E-6	2.84E-5	-2.3E-5	8.68E-5	O'Neill et al. 2005moving
Nubia-Mantle	None	30	43.54	-69.67	-6.05	3.31	9.89E-5	2.14E-5	-3.56E-5	8.19E-5	-6.45E-5	1.18E-4	O'Neill et al. 2005moving
Nubia-Mantle	None	40	44.56	-54.31	-8.08	0.6	6.39E-4	-9.58E-5	-3.09E-4	1.54E-4	-2.27E-5	3.49E-4	O'Neill et al. 2005moving
Nubia-Mantle	None	50	36.97	-58.9	-10.26	1.24	7.6E-4	-1.29E-4	-4.44E-4	1.65E-4	-1.35E-5	4.53E-4	O'Neill et al. 2005moving
Nubia-Mantle	None	60	23.73	-42.14	-12.53	2.31	1.61E-3	8.69E-5	-6.22E-4	6.07E-4	-2.63E-4	7.51E-4	O'Neill et al. 2005moving

Table B.1 continued from previous page

Moving-Fixed	Chron	Age/Ma	Latitude	Longitude	Rotation	$\hat{\kappa}$	a	b	c	d	e	f	Source
Nubia-Mantle	None	90	14.6	-33.26	-16.24	2.47	5.47E-4	-9.38E-5	5.78E-6	1.18E-4	-7.96E-5	5.36E-4	O'Neill et al. 2005 moving
Nubia-Mantle	None	100	14.4	-29.63	-20.08	3.25	2.41E-4	-1.79E-5	1.48E-5	3.16E-4	-1.69E-5	1.92E-4	O'Neill et al. 2005 moving
Nubia-Mantle	None	120	17.03	-27	-29.72	9.06	7.81E-4	3.73E-4	4.38E-4	6.44E-4	4.17E-4	5.87E-4	O'Neill et al. 2005 moving
Ind-Somalia	Ino	0.781	19.57	27.97	-347	1	1.99E-6	3.41E-6	1.86E-9	6.82E-6	6.01E-7	4.67E-7	Bull et al. 2010
Ind-Somalia	2ny	1.778	21.59	30.83	-735	1	4.26E-7	6.51E-7	4.44E-8	1.32E-6	1.74E-7	2.11E-7	Bull et al. 2010
Ind-Somalia	2An.1ny	2.581	22.48	30.6	-1.074	1	2.94E-6	4.33E-6	-6.73E-7	7.77E-6	-7.56E-9	9.55E-7	Bull et al. 2010
Ind-Somalia	2An.3no	3.596	18.7	34.62	-1.642	1	4.51E-6	7.3E-6	-5.4E-7	1.42E-5	7.06E-7	1.35E-6	Bull et al. 2010
Ind-Somalia	3n.1ny	4.187	22.11	28.45	-1.74	1	4.31E-6	7.11E-6	-2.86E-7	1.34E-5	5.49E-7	8.8E-7	Bull et al. 2010
Ind-Somalia	3n.4no	5.235	22.05	31.64	-2.181	1	2.18E-6	3.53E-6	-3.32E-7	6.89E-6	1.84E-7	6.19E-7	Bull et al. 2010
Ind-Somalia	3An.1ny	5.89	22.89	28.11	-2.333	1	6.98E-6	1.15E-5	-1.71E-6	2.29E-5	-3.39E-7	2.37E-6	Bull et al. 2010
Ind-Somalia	3An.2no	6.57	21.32	30.92	-2.748	1	2.41E-6	4.68E-6	-5.41E-10	9.68E-6	4.53E-7	4.53E-7	Bull et al. 2010
Ind-Somalia	4n.1ny	7.43	22.61	30.57	-2.982	1	5.75E-6	1.17E-5	-2.81E-7	2.68E-5	1.7E-6	1.92E-6	Bull et al. 2010
Ind-Somalia	4n.2no	8.07	22.01	30.79	-3.281	1	4.62E-6	8.01E-6	-1.16E-6	1.58E-5	-6.77E-7	1.4E-6	Bull et al. 2010
Ind-Somalia	4Any	8.67	22.48	30.77	-3.489	1	4.46E-6	6.30E-6	-1.54E-6	1.36E-5	-3.14E-6	3.14E-6	Bull et al. 2010
Ind-Somalia	4Ano	9.03	22.59	30.78	-3.685	1	1.45E-6	1.82E-6	-6.92E-7	3.84E-6	3.34E-7	1.33E-6	Bull et al. 2010
Ind-Somalia	5n.1ny	9.74	23.59	30.49	-3.89	1	3.08E-6	4.97E-6	-2.38E-7	1.01E-5	1.36E-6	1.63E-6	Bull et al. 2010
Ind-Somalia	5n.2no	10.95	23.62	29.3	-4.311	1	4.43E-6	7.25E-6	-1.58E-7	1.58E-5	2.01E-6	1.56E-6	Bull et al. 2010
Ind-Somalia	5An.2no	12.4	23.8	29.26	-4.879	1	3.39E-6	5.42E-6	-5.85E-7	1.04E-5	5.02E-7	1.44E-6	Bull et al. 2010
Ind-Somalia	5ADno	14.607	24.6	29.14	-5.76	1	2.83E-6	6.13E-6	4.51E-7	1.4E-5	1.29E-6	3.29E-7	Bull et al. 2010
Ind-Somalia	5Cn.1ny	16.01	24.8	29.3	-6.399	1	1.27E-6	2.26E-6	-1.7E-7	5.7E-6	4.81E-7	5.08E-7	Bull et al. 2010
Ind-Somalia	5Dny	17.235	24.83	30.29	-7.149	1	1.76E-6	2.98E-6	-2.06E-7	6.53E-6	6.03E-7	7.15E-7	Bull et al. 2010
Ind-Somalia	5Eny	18.28	24.77	30.27	-7.637	1	2.47E-6	5.44E-6	3.58E-7	1.32E-5	1.44E-6	5.17E-7	Bull et al. 2010
Ind-Somalia	6no	20.13	25.41	30.6	-8.469	1	3E-5	3.6E-5	-1.63E-5	6.11E-5	-7.2E-6	1.81E-5	Bull et al. 2010
Ind-Somalia	13no	33.55	-19.35	-137.7	16.05	1	3.85E-5	5.31E-5	-2.14E-5	1.05E-4	-1.84E-5	2.19E-5	Calculated
Ind-Somalia	18n.2no	40.13	-19.48	-135.85	19.31	1	4.06E-5	6.05E-5	-2.33E-5	1.31E-4	-2.52E-5	2.37E-5	Calculated
Ind-Somalia	20ny	42.54	-19.94	-136.28	20.23	1	3.74E-5	4.99E-5	-2.03E-5	9.48E-5	-1.51E-5	2.08E-5	Calculated
Ind-Somalia	20no	43.79	-20.44	-136.77	20.63	1	3.73E-5	4.98E-5	-2.03E-5	9.5E-5	-1.52E-5	2.11E-5	Calculated
Ind-Somalia	21ny	46.26	-19.6	-137.27	21.82	1	3.73E-5	4.94E-5	-2.02E-5	9.33E-5	-1.46E-5	2.08E-5	Calculated
Ind-Somalia	21no	47.91	-19.49	-138.01	22.9	1	3.7E-5	4.87E-5	-1.99E-5	9.13E-5	-1.41E-5	2.05E-5	Calculated
Ind-Somalia	22no	49.71	-19.63	-140.64	24.1	1	3.72E-5	4.96E-5	-2.01E-5	9.42E-5	-1.48E-5	2.07E-5	Calculated
Ind-Somalia	23n.2no	51.74	-17.83	-142.26	26.82	1	3.8E-5	5.26E-5	-2.09E-5	1.06E-4	-1.75E-5	2.15E-5	Calculated
Ind-Somalia	24n.3no	53.35	-17.69	-144.8	28.76	1	3.86E-5	5.46E-5	-2.13E-5	1.14E-4	-1.9E-5	2.17E-5	Calculated
Ind-Somalia	25ny	55.9	-17.68	-147.99	31.52	1	4.13E-5	6.48E-5	-2.29E-5	1.52E-4	-2.53E-5	2.29E-5	Calculated
Ind-Somalia	26ny	57.55	-16.81	-148.61	33.86	1	3.91E-5	5.72E-5	-2.14E-5	1.25E-4	-2E-5	2.16E-5	Calculated
Ind-Somalia	27ny	60.92	-15.03	-149.25	38.15	1	3.97E-5	5.93E-5	-2.15E-5	1.34E-4	-2.02E-5	2.15E-5	Calculated
Ind-Somalia	28ny	62.5	-14.48	-149.55	40.23	1	4.93E-5	6.58E-5	-2.61E-5	2.74E-4	-3.79E-5	2.38E-5	Calculated
Ind-Somalia	29no	64.75	-15.39	-152.91	42.42	1	3.92E-5	5.79E-5	-2.08E-5	1.32E-4	-1.79E-5	2.1E-5	Calculated
Ind-Somalia	34ny	83	21.571	21.698	-52.702	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Rowan and Rowley 2016
Ind-Nubia(Madagas)	synthetic	83	22.16	19.22	-52.74	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Gibbons et al. 2013 Muller et al. 2017
Ind-Nubia(Madagas)	synthetic	100	-23.19	-152.79	55.76	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Gibbons et al. 2013 Muller et al. 2017

Table B.1 continued from previous page

Moving-Fixed	Chron	Age/Ma	Latitude	Longitude	Rotation	$\hat{\kappa}$	a	b	c	d	e	f	Source
Ind-Nubia(Madagas)	synthetic	106	-23.19	-152.79	55.76	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Gibbons et al. 2013 Muller et al. 2017
Ind-Nubia(Madagas)	synthetic	120.6	-24.58	-153.92	54.51	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Gibbons et al. 2013 Muller et al. 2017
Somalia-Nubia	1no	0.78	-33.83	34.4	0.04602	1	3.737E-9	3.426E-9	-8.39E-10	3.957E-9	-8.03E-10	6.66E-10	Demets et al. 2017 replace 2010 inverted DoF from 2010
Somalia-Nubia	2A.2no	3.22	-44.7	2.8	0.27048	1	4.29E-9	1.26E-9	-5.01E-9	1.62E-9	-1.44E-9	7.19E-9	Horner-Johnson et al. 2005
Somalia-Nubia	5n.2no	10.95	-27.4	43.28	0.4	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Rowan and Rowley 2016
Somalia-Nubia	6no	20.13	-27.4	43.28	0.8	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Rowan and Rowley 2016
Somalia-Nubia	C7.2m	25.01	-27.4	43.28	1	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Rowan and Rowley 2016
Somalia-Nubia	C34	85	-27.4	43.28	1	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Rowan and Rowley 2016
Somalia-Nubia	>85NoReMot	120	-27.4	43.28	1	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Inferred From GlobalEarthByte.230-0Ma.GK07_AREPS.rot
Australia-EAnt	1no	0.78	11.3	41.8	-.49374	1	2.605E-8	1.846E-8	9.725E-9	2.184E-8	1.785E-9	1.031E-8	Demets et al. 2010 inverted
Australia-EAnt	2An.1ny	2.58	-11.164	-139.7	1.655	5.33	2.81E-7	-3.35E-7	2.6E-7	5.15E-7	-4.87E-7	9.06E-7	Cande and Stock 2004 Krigsman et al. 1999 Age
Australia-EAnt	3An.1ny	5.89	-11.591	-139.23	3.83	2.12	2.94E-7	-3.9E-7	3.19E-7	6.27E-7	-5.83E-7	8.97E-7	From Mecke2005
Australia-EAnt	5n.2no	10.95	-11.896	-142.058	6.79	1.02	1.36E-7	-1.71E-7	5.56E-8	3.05E-7	-1.89E-7	4.47E-7	Cande and Stock 2004
Australia-EAnt	6no	20.13	-13.393	-145.63	12.051	1.06	1.61E-7	-1.85E-7	2.56E-8	3E-7	-1.59E-7	4.22E-7	Cande and Stock 2004
Australia-EAnt	8o	26	-13.79	-146.43	15.92	0.95	1.68E-7	-2.05E-7	1.05E-7	1.05E-7	-2.72E-7	5.04E-7	Granot and Dymet 2018
Australia-EAnt	8n.2no	26.55	-13.805	-146.444	15.919	1.01	1.95E-7	-2.24E-7	5.79E-8	3.71E-7	-2.39E-7	5.9E-7	Cande and Stock 2004
Australia-EAnt	10n.2no	28.75	-13.58	-146.016	17.319	1.17	2.32E-7	-2.56E-7	8.48E-8	3.94E-7	-2.77E-7	7.51E-7	Cande and Stock 2004
Australia-EAnt	12no	30.94	-13.396	-145.623	18.89	0.75	2.6E-7	-3.14E-7	5.53E-8	5.18E-7	-3.1E-7	9.82E-7	Cande and Stock 2004
Australia-EAnt	13no	33.55	-13.451	-145.623	20.495	1.07	2.44E-7	-3.26E-7	1.49E-7	5.56E-7	-3.43E-7	6.93E-7	Cande and Stock 2004
Australia-EAnt	17n.3no	38.11	-14.65	-146.525	22.882	0.49	6.88E-7	-8.39E-7	-2.51E-7	1.34E-6	-1.37E-7	1.31E-6	Whittaker et al. 2007-2013
Australia-EAnt	20no	43.79	14.92	32.5	-24.51	0.74	6.3E-7	-1E-7	1.68E-6	1.87E-6	-3.1E-6	6.76E-6	Whittaker et al. 2007-2013
Australia-EAnt	21ny	46.26	13.42	33.83	-24.62	5.32	1.328E-5	-2.024E-5	8.2E-7	3.203E-5	-2.28E-6	5.76E-6	Whittaker et al. 2013.replace2007
Australia-EAnt	24n.3no	53.35	10.48	35.17	-25.24	0.11	3.38E-7	-2.17E-7	4.29E-7	4.84E-7	-1.139E-6	4.281E-6	Whittaker et al. 2013.replace2007
Australia-EAnt	27ny	60.92	9.22	35.42	-25.43	0.17	3.09E-7	-1.93E-7	3.6E-7	5.07E-7	-1.206E-6	4.474E-6	Whittaker et al. 2013.replace2007
Australia-EAnt	31no	68.74	7.89	35.75	-25.61	0.19	3.11E-7	-1.85E-7	3.23E-7	4.95E-7	-1.158E-6	4.303E-6	Whittaker et al. 2013.replace2007
Australia-EAnt	32n.1ny	71.07	6.67	36.14	-25.75	0.14	2.5E-7	-1.82E-7	3.25E-7	4.82E-7	-1.087E-6	3.991E-6	Whittaker et al. 2013.replace2007
Australia-EAnt	33no	79.08	4.53	36.64	-26.13	0.13	2.52E-7	-1.79E-7	2.97E-7	4.75E-7	-1.031E-6	3.788E-6	Whittaker et al. 2013.replace2007

Table B.1 continued from previous page

Moving-Fixed	Chron	Age/Ma	Latitude	Longitude	Rotation	$\hat{\kappa}$	a	b	c	d	e	f	Source
Australia-EAnt	34my	83.5	1.02	37.28	-26.62	0.13	5.37E-7	-6E-9	-6.56E-7	5.73E-7	-8.19E-7	2.356E-6	Williams et al. 2011
Australia-EAnt	QZB Full-fit	96	-12.69	46.58	-29.06	1.03	3.53E-6	-4.06E-6	8.19E-6	6.09E-6	-1.35E-5	3.571E-5	Whittaker et al. 2007
Australia-EAnt	6no	-3.91	37.9	-30.86	1.04	2.85E-6	-2.63E-6	3.77E-6	4.01E-6	6.93E-6	1.38E-5	Whittaker et al. 2011 Whittaker et al. 2013 replaced 2007	
EAnt-Somalia	Ino	0.78	-20.58	115.29	-1.0062	1	3.661E-9	3.38E-9	-8.209E-10	4.124E-9	-1.068E-9	2.415E-9	Dennets et al. 2017 inverted.replace 2010
EAnt-Somalia	2A.2no	3.22	7.9	-44.1	0.42826	1	7.9E-10	8.8E-10	-8.2E-10	1.49E-9	-1.22E-9	1.15E-9	Horner-Johnson et al. 2005
EAnt-Somalia	5m.2no	10.95	14.6	-49.1	1.53	1	2.21E-7	2.36E-7	-9.2E-8	3.04E-7	-1.67E-7	2.45E-7	LeMaux et al. 2002
EAnt-Somalia	6no	20.131	10.8	-46	2.7	1	1.028E-7	8.83E-8	-2.34E-8	1.759E-7	-1.347E-7	2.237E-7	Patriat et al. 2008
EAnt-Somalia	8n.2no	26.554	14.3	-46.9	3.91	1	1.484E-7	1.489E-7	-9.66E-8	2.58E-7	-2.458E-7	3.702E-7	Patriat et al. 2008
EAnt-Somalia	13my	33.06	16.2	-44.7	5.66	1	8.56E-7	6.73E-7	-1.66E-7	6.9E-7	-3.74E-7	5.82E-7	Patriat et al. 2008
EAnt-Somalia	13no	33.55	12.69	-44.61	5.67	0.59	5.71E-7	5.12E-7	5.63E-7	2.35E-7	-2.53E-7	3.99E-7	Cande et al. 2010 basic
EAnt-Somalia	18n.2no	40.13	13.8	-43.75	7.05	0.63	6.6E-7	5.69E-7	5.56E-7	1E-9	-7.9E-8	6.8E-7	Cande et al. 2010 basic
EAnt-Somalia	20my	42.54	11.84	-42.32	7.53	0.93	2.36E-7	1.91E-7	2.31E-7	-2.28E-7	-2.63E-7	5.98E-7	Cande et al. 2010 all
EAnt-Somalia	20no	43.79	11.84	-42.16	7.87	0.78	1.64E-7	1.23E-7	1.79E-7	-1.95E-7	-2.42E-7	4.97E-7	Cande et al. 2010 all
EAnt-Somalia	21my	46.26	11.29	-41.54	8.49	0.55	1.89E-7	1.6E-7	2.24E-7	-2.76E-7	-3.25E-7	6.86E-7	Cande et al. 2010 all
EAnt-Somalia	21no	47.91	9.82	-40.7	8.83	1.86	2.26E-7	1.98E-7	2.67E-7	-3.69E-7	-4E-7	8.38E-7	Cande et al. 2010 all
EAnt-Somalia	22no	49.71	9.19	-40.63	9.21	1.4	2.73E-7	2.54E-7	3.7E-7	-4.7E-7	-5.4E-7	1.084E-6	Cande et al. 2010 all
EAnt-Somalia	23n.2no	51.74	9.31	-41.53	9.61	0.74	2.78E-7	2.77E-7	4.22E-7	-4.41E-7	-5.38E-7	9.45E-7	Cande et al. 2010 with-SWIR
EAnt-Somalia	24n.3no	53.35	10.16	-43.3	9.96	0.72	4.74E-7	4.34E-7	5.65E-7	-7.41E-7	-7.62E-7	1.398E-6	Cande et al. 2010 with-SWIR
EAnt-Somalia	25my	55.9	9.86	-45.24	10.49	0.52	1.149E-6	1.375E-6	1.932E-6	-2.016E-6	-2.601E-6	3.954E-6	Cande et al. 2010 with-SWIR
EAnt-Somalia	26my	57.55	10.64	-47.47	10.78	1.35	1.267E-6	1.409E-6	1.747E-6	-1.885E-6	-2.253E-6	3.33E-6	Cande et al. 2010 with-SWIR
EAnt-Somalia	27my	60.92	7.1	-45.8	11.08	1.15	1.623E-6	1.751E-6	2.193E-6	-2.459E-6	-2.843E-6	4.288E-6	Cande et al. 2010 with-SWIR
EAnt-Somalia	28my	62.5	4.75	-44.79	11.39	0.66	1.359E-6	1.549E-6	2.054E-6	-2.354E-6	-2.838E-6	4.486E-6	Cande et al. 2010 with-SWIR
EAnt-Somalia	29no	64.75	4.79	-45.56	11.7	0.53	9.16E-7	9.94E-7	1.27E-6	-1.651E-6	-1.913E-6	3.288E-6	Cande et al. 2010 with-SWIR
EAnt-Somalia	34y EAn-Afr	83	-1.8	-38.7	17.9	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Rowan and Rowley 2016
EAnt-Somalia	96	-3.06	-33.49	26.8	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Marks and Tilkku 2001EPSL	
EAnt-Somalia	M0	120.6	10.36	153.67	-41.56	1	1E-15	1E-15	1E-15	1E-15	1E-15	1E-15	Muller et al. 2008G3

Rotations for the plate pairs are given as motion of the ‘Moving’ plate relative to the ‘Fixed’ plate. The covariance matrix is given by the formula

$$\frac{1}{\hat{\kappa}} \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \text{ where the values of } a-f \text{ are given in radians}^2.$$

**Table B.2:** Calculated North American, Indian and Australian reference APWPs (5 Myr interval) in the hotspot references

Longitude	Latitude	Age/Ma	Azimuth	SemiMajor	SemiMinor
North America, FHM					
-179	90	0	0	0	0
-174.331	89.61	5	113.946	0.7802	0.4368
-175.024	89.266	10	117.593	1.446	0.7937
-143.465	88.579	15	145.472	1.287	0.8165
-135.153	87.731	20	137.589	1.58	1.1
-164.308	86.976	25	148.751	2.48	1.109
173.098	86.207	30	127.221	1.698	1.286
168.608	85.861	35	95.3017	1.95	1.066
165.88	85.483	40	95.9912	3.886	1.736
-167.954	83.693	45	130.354	2.862	1.303
-154.176	80.461	50	137.487	3.578	1.613
-155.85	79.001	55	138.82	3.588	2.14
-158.209	78.156	60	121.46	4.859	3.176
-162.178	78.169	65	132.308	3.69	2.3
-168.366	78.755	70	124.638	4.342	2.501
-170.935	78.245	75	119.072	2.643	1.541
-176.092	78.137	80	119.419	4.383	2.572
-171.38	76.706	85	132.099	2.771	1.327
-169.766	75.087	90	137.193	3.709	1.367
174.157	74.423	95	121.558	2.234	1.625
159.46	72.443	100	30.8058	2.898	2.683
162.224	71.777	105	85.1415	2.323	1.746
164.338	70.753	110	94.224	4.13	2.26
151.714	72.571	115	67.3948	4.635	2.58
137.618	73.573	120	50.9863	7.344	4.06
North America, MHM					
-179	90	0	0	0	0
167.448	89.222	5	93.9164	0.79	0.4434
166.095	88.489	10	97.1675	1.692	0.9353
170.16	87.697	15	98.0275	1.43	0.9188
171.448	86.871	20	88.655	1.876	1.31
174.498	85.977	25	126.491	2.586	1.186
169.985	85.437	30	119.932	1.758	1.246
169.891	85.352	35	102.791	1.88	1.024
170.954	85.237	40	107.055	3.863	1.725
167.564	84.214	45	112.528	2.921	1.197
170.182	82.928	50	112.937	3.648	1.478
-171.618	81.501	55	132.2	3.788	1.974
-161.291	79.974	60	133.697	4.997	2.988

**Table B.2 continued from previous page**

Longitude	Latitude	Age/Ma	Azimuth	SemiMajor	SemiMinor
-159.451	79.35	65	145.708	3.812	2.029
-159.851	79.359	70	148.362	3.806	1.859
-161.687	79.597	75	143.166	2.488	1.432
-166.266	80.301	80	138.098	3.482	2.166
-172.202	80.986	85	137.839	2.606	1.301
179.532	81.68	90	133.098	3.09	1.22
173.853	81.911	95	133.027	2.015	1.363
166.329	81.647	100	5.23552	2.217	1.972
162.461	80.565	105	75.9178	1.931	1.578
158.682	79.369	110	77.4283	2.952	1.897
152.419	78.189	115	65.0493	3.113	1.663
147.766	76.872	120	60.0476	4.742	2.28
India, FHM					
-179	90	0	0	0	0
-68.4652	88.2	5	16.7724	0.7074	0.4826
-71.7616	86.461	10	18.9733	1.121	0.6732
-74.4272	83.942	15	7.15217	1.135	0.8127
-72.6	81.043	20	170.291	1.79	1.302
-68.5899	79.462	25	156.305	1.443	0.9811
-65.2017	77.644	30	149.031	1.644	1.019
-62.0848	74.819	35	11.5526	2.123	1.469
-60.3729	72.289	40	25.2625	4.021	2.311
-66.2619	67.566	45	19.1795	2.617	1.427
-71.1592	61.003	50	14.6201	3.698	1.915
-72.6213	54.419	55	178.977	3.233	2.035
-72.3461	47.845	60	175.474	4.797	3.071
-73.3224	42.441	65	173.455	3.449	2.173
-73.2032	40.393	70	0.136912	4.265	2.488
-73.7248	36.782	75	178.704	2.678	1.58
-73.9922	33.116	80	178.443	4.208	2.503
-74.3076	26.836	85	2.40753	2.507	1.247
-73.0955	22.961	90	7.5674	3.89	1.434
-74.2136	23.036	95	5.96566	2.234	1.625
-75.3342	23.09	100	107.925	2.806	2.598
-75.2387	19.342	105	157.362	2.323	1.746
-75.3615	15.893	110	161.764	4.555	2.492
-73.9386	17.277	115	144.671	4.635	2.58
-72.6048	18.789	120	139.637	6.026	3.332
India, MHM					
-179	90	0	0	0	0
-79.4273	88.505	5	4.28395	0.723	0.4801
-83.6684	87.013	10	5.28613	1.444	0.8494

**Table B.2 continued from previous page**

Longitude	Latitude	Age/Ma	Azimuth	SemiMajor	SemiMinor
-85.8127	85.284	15	178.968	1.288	0.9161
-82.7702	83.243	20	169.091	2.016	1.52
-74.1526	80.865	25	158.372	1.528	1.127
-68.475	78.19	30	150.59	1.621	1.093
-64.1437	75.027	35	14.2827	2.02	1.489
-61.7224	72.127	40	30.0739	3.974	2.402
-64.3999	70.211	45	30.2175	2.618	1.439
-67.989	66.547	50	29.2119	3.662	1.833
-70.4059	57.978	55	15.7185	3.324	2.08
-70.4049	49.361	60	13.6689	4.86	3.013
-71.6567	42.727	65	11.0741	3.433	2.166
-71.5409	39.451	70	19.6466	3.506	1.922
-71.2003	36.142	75	14.9621	2.512	1.501
-70.7182	32.834	80	9.68229	3.3	2.132
-70.5195	29.586	85	11.9972	2.322	1.231
-68.3816	28.507	90	16.0164	2.807	1.109
-66.7679	26.262	95	20.8345	2.015	1.363
-65.2692	23.983	100	79.6564	1.972	1.753
-66.2778	21.878	105	151.058	1.931	1.578
-67.3405	20.117	110	152.924	2.704	1.738
-68.211	18.388	115	143.392	3.113	1.663
-68.9914	16.631	120	139.979	4.007	1.927
Australia, FHM					
-179	90	0	0	0	0
-41.4449	87.036	5	54.1118	0.689	0.3796
-45.0416	84.256	10	49.038	0.7835	0.4258
-50.6502	81.119	15	38.1494	1.117	0.7248
-53.2369	77.974	20	30.9554	1.476	1.028
-53.2848	76.256	25	16.07	1.106	0.869
-51.2849	74.267	30	167.746	1.167	0.8867
-50.2873	71.435	35	40.8833	1.936	1.056
-50.5986	69.481	40	42.2012	3.984	1.782
-55.5547	66.545	45	37.3027	2.536	1.191
-58.9951	64.115	50	33.5829	3.654	1.676
-57.5828	63.706	55	23.0261	3.095	1.867
-57.2977	63.671	60	16.0634	4.463	2.843
-55.5565	64.436	65	20.5328	3.214	1.989
-51.9574	65.69	70	29.6999	4.312	2.465
-51.2245	65.722	75	30.9868	2.626	1.529
-50.5474	65.686	80	32.5267	4.471	2.657
-51.107	63.469	85	39.4428	2.511	1.249
-48.3618	61.873	90	48.1966	3.89	1.435

**Table B.2 continued from previous page**

Longitude	Latitude	Age/Ma	Azimuth	SemiMajor	SemiMinor
-43.0424	64.33	95	53.4492	2.241	1.629
-41.1958	67.209	100	154.896	2.871	2.661
-42.376	65.281	105	21.3797	2.327	1.752
-43.2223	63.352	110	23.2216	4.278	2.343
-34.22	63.817	115	12.2265	4.637	2.583
-24.9452	64.093	120	13.5957	6.936	3.836
Australia, MHM					
-179	90	0	0	0	0
-43.8522	87.468	5	49.83	0.7027	0.3808
-48.2655	85.095	10	44.0171	1.459	0.7801
-52.7796	82.828	15	37.4668	1.275	0.8358
-55.2364	80.568	20	33.5627	1.731	1.21
-55.4555	77.907	25	25.7336	1.268	0.9666
-53.2917	75.002	30	1.25275	1.156	0.9391
-51.8626	71.779	35	45.6379	1.863	1.018
-51.898	69.415	40	46.9976	4.026	1.803
-52.3541	68.985	45	48.9107	2.568	1.132
-51.6966	69.015	50	51.1423	3.548	1.462
-51.9822	66.573	55	42.7177	3.232	1.823
-53.4926	64.622	60	36.3825	4.516	2.731
-52.8722	64.191	65	40.5529	3.247	1.883
-50.5454	64.244	70	49.5822	3.279	1.689
-48.1673	64.106	75	49.0042	2.465	1.426
-45.9043	63.879	80	46.2523	3.33	2.149
-41.6854	63.431	85	55.6345	2.326	1.233
-33.6915	62.706	90	67.5618	2.814	1.113
-28.7281	61.005	95	77.4292	2.028	1.373
-26.8608	60.554	100	135.582	2.143	1.914
-26.3797	60.593	105	26.4611	1.942	1.585
-25.8665	60.575	110	27.2651	2.877	1.85
-25.2025	60.386	115	17.2151	3.118	1.669
-24.4723	60.146	120	13.2747	4.514	2.173