New Zealand Paleoseismic Site Database: Data Dictionary

NJ Litchfield

GNS Science Report 2021/40
June 2022

GNS SCIENCE
TE PŪ AO

DISCLAIMER

The Institute of Geological and Nuclear Sciences Limited (GNS Science) and its funders give no warranties of any kind concerning the accuracy, completeness, timeliness or fitness for purpose of the contents of this report. GNS Science accepts no responsibility for any actions taken based on, or reliance placed on the contents of this report and GNS Science and its funders exclude to the full extent permitted by law liability for any loss, damage or expense, direct or indirect, and however caused, whether through negligence or otherwise, resulting from any person's or organisation's use of, or reliance on, the contents of this report.

BIBLIOGRAPHIC REFERENCE

Litchfield NJ. 2022. New Zealand Paleoseismic Site Database: data dictionary. Lower Hutt (NZ): GNS Science. 42 p. (GNS Science report; 2021/40). doi:10.21420/G5K4-ES33.

NJ Litchfield, GNS Science, PO Box 30368, Lower Hutt 5040, New Zealand

www.gns.cri.nz

ISSN 2350-3424 (online)
ISBN 978-1-99-101309-5 (online)
http://dx.doi.org/10.21420/G5K4-ES33

CONTENTS

TABLES

APPENDICES

ABSTRACT

The *New Zealand Paleoseismic Site Database* (new term, new database) contains paleoseismic data (grouped into slip rate, single-event displacement, earthquake timings and recurrence intervals) collected at specific sites along active faults throughout New Zealand. It was initially developed as part of the New Zealand National Seismic Hazard Model 2022 Revision Project (NSHM 2022, which began in earnest in 2020), and the purpose, design and contents of the database are described in a separate report. This report provides the definitions, formats and guidelines for compiling each attribute in the database, as well as templates for future compilations. A draft of this report was used during the initial compilation; subsequent tweaks are incorporated in this report and additional tweaks are anticipated to be required in the future. This report therefore not only describes how to fill out (compile or populate) attributes but also contains some information on the current state of the database and practical tips for compiling attributes based on experience that may be useful for other databases and future compilers. The database has two components, a stand-alone spreadsheet developed specifically for the NSHM 2022 project and a GIS feature class for uploading into the *AF.Points* (Active Fault Point features) layer of the New Zealand Active Faults Database. Attributes and fields are described for each of these components in a series of tables that include the attribute Definitions; Key Features, such as whether it is compulsory, the format, options and how uncertainties should be expressed; and Additional Information.

KEYWORDS

Paleoseismic, active fault, database, GIS, data dictionary

1.0 INTRODUCTION

The *New Zealand Paleoseismic Site Database* is a new database that contains paleoseismic data (grouped into slip rate, single-event displacement, earthquake timings and recurrence intervals) collected at specific sites along active faults throughout New Zealand – primarily on-land faults, but some information on offshore faults is also included. It was initially developed as part of the update of the New Zealand National Seismic Hazard Model 2022 Revision Project (NSHM 2022, which started in earnest in late 2020) but should be useful for a range of other purposes. The purpose, design and content of the database are described in a separate report (Litchfield et al. 2022).

This report provides the definitions, formats and guidelines for compiling each attribute in the database, as well as templates for future compilations (Appendix 1, Appendix 2). A draft of this report was used by compilers (listed in the acknowledgements) during the initial compilation to achieve a common understanding of the attributes and consistency of compilation. As expected, tweaks were made during the compilation to facilitate different types of data, which are incorporated in this version. It is likely that further tweaks will be required in the future as new types of datasets are added or the database is used for other purposes.

The report is divided into two sections, corresponding to the two components of the database. The first (Section 2) is the Excel spreadsheet that is the stand-alone database developed for the NSHM 2022 project. The second (Section 3) is the Geographic Information System (GIS) feature class that is designed to be uploaded into the *AF.Points* (Active Fault Point features) layer of the New Zealand Active Faults Database (NZAFD).

The attributes are described in a series of tables that include the attribute Definitions; Key Features, such as whether it is compulsory, the format, options and how uncertainties should be expressed; and Additional Information. Some of the additional information refers to the NSHM 2022 project but also includes practical tips for compiling attributes based on experience during the compilation that may be useful for other databases and/or future compilers.

The spreadsheet (Section 2) was originally adapted from the Uniform California Earthquake Rupture Forecast (UCERF) paleoseismic site databases (Field et al. 2013, particularly Appendix B for the slip-rate data), and so some of the attribute definitions are adapted from those. Most of the information for the GIS attributes (Section 3) is taken directly from the NZAFD Data Dictionary (Jongens and Dellow 2003).

2.0 PALEOSEISMIC SITE DATABASE SPREADSHEET

The master Paleoseismic Site Database (SiteDB) spreadsheet currently contains seven worksheets. Three contain the paleoseismic site data:

1. Slip Rate
2. Earthquake (EQ) Timings and Recurrence Interval (RI)
3. Single-Event Displacement.

These can be supplied as stand-alone files. The fourth worksheet,

4. Reference List

accompanies the first three worksheets and contains the references for all of the data in worksheets 1–3. The references are formatted according to the GNS Science Bibliographic Database, and the Bib ID corresponds to those contained within that database. The remaining worksheets are for calculations made for the purposes of paleoseismic site data compilation:

5. Net Calcs – for calculating net slip rates or single-event displacements from horizontal, vertical and/or dip-slip components.
6. SR Calcs – for calculating slip rates from offset measurements and ages.
7. Rake – for re-calculating rakes from the convention used in previous databases, such as the New Zealand Active Fault Model (Litchfield et al. 2014) to the Aki and Richards (1980) right-hand-rule convention.

Worksheets 1–3 are described in the following sections, and a template is contained in Appendix 1.

2.1 Slip Rate Worksheet

The Slip Rate worksheet is divided into two parts:

1. The left-hand part (columns A–I) contains Fault Data and is primarily data for specific faults or fault sections in the Community Fault Model (CFM; Van Dissen et al. 2021; Seebeck et al. 2022). The attributes are described in Table 2.1 and only need to be reported once for each fault. That is, the Fault Data are not repeated for every site.
2. The right-hand part (columns J–AU) contains Site Data; these are the data for an individual site or, in a few cases, data for two or more sites that have been combined together in publications. The attributes are described in Tables 2.2 and 2.3 and need to be compiled for each site.

The data in the Slip Rate worksheet are currently listed in alphabetical order of Fault name. The use of grey shades and borders in the worksheet is simply to assist with reading of the data; there are no specific rules for their use. Note that, because the Fault Data is not reported for each site, sorting the spreadsheet by Fault name will cause some sites to lose association with the fault.

Table 2.1 Slip Rate worksheet Fault Data attributes.

Attribute Name	Key Features	Definition	Additional Information
Fault	• Compulsory • Free text[1] • Omit 'fault' from name	Name of the fault to which the site data is applicable.	Where possible, the name should be from the NZAFD. Some off-fault sites (e.g. lakes, wetlands, marine terraces) have been assigned to the faults that have been inferred to have caused the deformation they record, but it is also possible to name them as 'Off-fault'.
CFM Name	• Compulsory • Free text • Spelled out in full	Name of the fault or fault section in the CFM to which the site data is applicable.	Most of the sites do not sit on a CFM fault in map view because of scale differences (CFM being regional scale), so a judgement call will need to be made as to which CFM fault the site is applicable. This may change as the CFM is updated, so will need to be updated accordingly.
CFM No.	• Compulsory • Number	Number (Fault_ID) of the fault or fault section in the CFM to which the site data is applicable.	This may change as the CFM is updated, so will need to be updated accordingly.
CFM Sense Dominant	• Compulsory • From a restricted list: - Dextral - Normal - Reverse - Sinistral	Dominant sense of movement on the fault or fault section in the CFM to which the site data is applicable.	This may change as the CFM is updated, so will need to be updated accordingly.
CFM Sense Secondary	• Optional • From a restricted list: - Dextral - Normal - Reverse - Sinistral	Secondary sense of movement on the fault or fault section in the CFM to which the site data is applicable.	This may change as the CFM is updated, so will need to be updated accordingly.

1 Free text in this report technically have no restrictions on length or characters (e.g. symbols or numbers) used. However, in practise, and particularly for attributes to be uploaded into the *AF.Points* layer of the NZAFD, the text should be kept as short as practicable, and only text and numbers used.

Attribute Name	Key Features	Definition	Additional Information
CFM Dip	• Compulsory • Number • Reported in degrees (0–90°) • Uncertainties should be included; in the format: - Preferred ± 2 sigma uncertainty, or - Preferred (minimum-maximum)	Dip assigned to the fault or fault section in the CFM to which the site data is applicable.	This may change as the CFM is updated, so will need to be updated accordingly. ± can be entered quickly using ALT-0177. The CFM also contains dip direction, but this is not needed for the paleoseismic site database and so is not included.
CFM Rake	• Compulsory • Number • Reported in degrees using the Aki and Richards (1980) right-hand-rule convention whereby: - 180/-180° = Dextral - -90° = Normal - 0° = Sinistral - 90° = Reverse • Uncertainties should be included; in the format: • Preferred ± 2 sigma uncertainty, or • Preferred (minimum-maximum)	Rake of the net sense of movement on the fault or fault section in the CFM to which the site data is applicable.	This may change as the CFM is updated, so will need to be updated accordingly. ± can be entered quickly using ALT-0177.
CFM Net SR	• Compulsory • Number • Reported in millimetres per year (mm/yr) • Uncertainties should be included; in the format: - Preferred (minimum-maximum)	Net slip rate for the fault or fault section in the CFM to which the site data is applicable.	This value should be derived from the site data in this database and, where applicable, updated in the CFM. There are no fixed rules for how this should be calculated, but the method should be noted in the CFM SR comments column. Ideally, it should be updated using CFM dip if this is significantly different from that used in the original publication. If the CFM fault dip changes, this value will need to be changed accordingly.
CFM Net SR Comments	• Compulsory • Free text	Comments regarding the net slip rate assigned to the fault or fault section in the CFM to which the site data is applicable.	The CFM slip rate was initially compiled from the New Zealand Active Fault Model (Litchfield et al. 2014), which is referred to as the '2014 AFM'. This attribute should therefore describe any updates to the net slip rate and, if so, the methods used to calculate it and/or the data sources.

Table 2.2 Slip Rate worksheet Site Data attributes.

Attribute Name	Key Features	Definition	Additional Information
Site DB ID	• Compulsory • Number	Site database identification number.	Entered in order of compilation and can have gaps but not duplicates. Must match the Site ID number in the GIS feature class (Section 3).
Site Name	• Compulsory • Free text or number	Name or number of the site.	From the original publication. Or, if there is none, one should be added that describes the location and feature (e.g. Saxton River T1).
Easting and Northing	• Compulsory • Number • Reported as a seven-digit figure (i.e. rounded to the nearest metre) • Coordinate System NZTM 2000	Site location XY coordinates.	This should be as accurate as possible, so could be from a reported grid reference or, if more appropriate, updated from the GIS feature class. For many sites, the locations were estimated using a published map and refined using orthophotos or LiDAR data, as noted in the Comments field in the GIS feature class. The location uncertainty is estimated in the Method_Accuracy field in the GIS feature class.
Local Strike	• Optional • Number • Reported in degrees (0–360°) • Reported as: - Preferred ± 2 sigma uncertainty, or - Preferred (minimum-maximum)	Strike of fault in the vicinity of the site.	Typically measured over an along-strike length of several hundred metres for small-scale (<20 m) displacements or between piercing points for larger offsets. This attribute was not prioritised during the initial compilation and to date has only been compiled for a few sites.
Reported SR	• Compulsory • Number • Reported in millimetres per year (mm/yr) • Format as per the original study • Reported as separate components (e.g. Right Lateral, Vertical), where available	The geologic slip rate as reported in the original study.	Because investigators report uncertainties in a variety of ways, this attribute is what was originally reported in the referenced study. This may include slip rate reported in individual components and/or a dip or net slip rate that may differ from the CFM SR if a different dip is used in the calculations. Where two components have been reported, they are entered one above the other. They can be separated using ALT-ENTER.

Attribute Name	Key Features	Definition	Additional Information
Reported Component (SR)	• Compulsory • From a restricted list: - DS = Dip Slip - LL = Left Lateral - NS = Net - RL = Right Lateral - V = Vertical	Component of slip rate reported in the Reported SR column.	Where two components have been reported, they are entered one above the other. They can be separated using ALT-ENTER.
Reported SR Min. and Max.	• Optional • Number • Reported in millimetres per year (mm/yr) • Reported as: - Preferred ± 2 sigma uncertainty, or - Preferred (minimum-maximum)	Reported slip rate expressed as minimum and maximum values.	If the minimum and maximum slip rates are reported in the referenced study, these values can be recorded here. ± can be entered quickly using ALT-0177.
SR Calculated Uncertainties	• Optional • Number • Reported in millimetres per year (mm/yr)	Calculated uncertainties for slip rates reported as a single value.	If no slip rate uncertainties are reported, then uncertainties may be calculated using the Reported Offset and Reported Age minimum and maximum values. The uncertainties are interpreted to be 95% confidence intervals (two sigma uncertainties).
CFM SR Net	• Compulsory • Number in millimetres per year (mm/yr) • Reported as: - Preferred (minimum-maximum)	Net slip rate calculated for the CFM, calculated or corrected for the CFM Dip (if significantly different than used in the original study).	Many slip rates are reported with only the horizontal or vertical component, so a net slip rate needs to be calculated. This can be done using the Net Calcs worksheet. The uncertainties can be calculated from the Reported Offset and Reported Age minimum and maximum values. If the offset and ages are only reported as ranges, then the preferred value is calculated from the mean value. If only one component is reported, and there is reason to believe there is, or could be, another component of motion, then the net SR should be prefixed with a > or ≥, respectively. If the CFM fault dip changes, this value will need to be changed accordingly.

Attribute Name	Key Features	Definition	Additional Information
QR1 Offset Feature, QR2 Dating and QR3 Overall	• Compulsory • A, B, C or D	Quality rankings of the offset feature measurements (QR1), dating method (QR2) and an overall ranking (QR3).	Defined in Table 2.3. This is largely taken from the UCERF3 geologic slip rate database (Appendix B of Field et al. 2013).
Reported Component (Offset)	• Compulsory • From a restricted list: - DS = Dip Slip - LL = Left Lateral - NS = Net - RL = Right Lateral - V = Vertical	Component of slip on the offset feature either reported from the study or inferred by the compiler.	-
Reported Offset	• Compulsory • Number • Reported in metres (m)	Reported offset used to calculate a slip rate.	Because investigators report uncertainties in a variety of ways, this attribute is what was originally reported in the referenced study. This may include slip rate reported in individual components and/or a dip or net slip rate that may differ from the CFM SR if a different dip is used in the calculations. Where two components have been reported, they are entered one above the other. They can be separated using ALT-ENTER. ± can be entered quickly using ALT-0177.
Offset Feature(s)	• Compulsory • Free text	Type of feature that is offset.	For example: terrace riser, channel, moraine, ridge.
No. of Measurements	• Optional • Number	Number of offset measurements obtained from a site for each component.	This refers to the number of individual offsets (e.g. channels or risers) measured rather than the number of measurements of a single offset or cumulative offset measurements across a multi-strand fault zone.
Reported Age	• Compulsory • Number in thousands of years (ka)	The reported age of the offset feature.	Because investigators report ages in a variety of ways (for example, as a range of values versus a preferred value with uncertainties), this column is what was originally reported in the referenced study to allow for double-checking of the Start and End ages. ± can be entered quickly using ALT-0177.

Attribute Name	Key Features	Definition	Additional Information
Start Age (Pref, Min., Max.)	• Compulsory • Number in thousands of years (ka)	Age for the start of the interval over which the slip rate is calculated.	In most cases, the offset is assumed to have accumulated soon after the dated feature formed. However, an additional uncertainty is that the offset may start a considerable amount of time after the feature formed. This amount of time is usually not addressed in most studies and, for the purposes of this compilation, ignored unless specifically addressed in the study. This unknown amount of time biases the slip rate to be low.
End Age (Pref, Min., Max.)	• Optional • Number in thousands of years (ka)	Age for the end of the interval over which the slip rate is calculated.	Ideally in a slip rate study, both a start and end time would be used in the slip rate calculation to provide a true closed interval of time over which the offset accumulated. However, in most slip rate studies, the timing of the last earthquake is unknown, or not accounted for, and the open interval is included in the slip rate calculation. Although this potentially biases the slip rate to be low, it is considered to have a negligible effect on the slip rate if the interval of time is long. Reported end times are more common in trenching studies where there is event timing data and is especially important for rates calculated based on a limited number of offsets.
Dating Method	• Compulsory • Free text • Abbreviations: - OSL = Optically Stimulated Luminescence - C14 = Radiocarbon	Method used to constrain the age component of the slip rate estimate.	Sometimes this is not clear but should be filled in if possible and comments added to the Comments Regarding Dating column. Correlation is used when the age is assigned based on correlation to a geomorphic feature of known age or a climate period.
SR Timeframe Category	• Optional • Numbers in thousands of years (ka) • From a restricted list of categories: - <1 - 1–11 - 11–130 - 130–750 - 750–2600 - >2600	Generalised time intervals over which the slip rate applies.	Uses the USGS Quaternary Fault and Fold Database categories. Where the time frame spans more than one category, a judgement call should be used as to which category it falls into; for example, using a preferred value or the category that the majority falls into. This attribute was not prioritised during the initial compilation and so has only been compiled for a few sites. If it is compiled in future versions, it could be useful for documenting changes in slip rate with time.

Attribute Name	Key Features	Definition	Additional Information
No. of Events (Pref., Min., Max.)	• Optional • Number	Number of earthquake rupture events that contributed to the offset of the feature.	This is often unknown. This attribute was not prioritised during the initial compilation and so has only been compiled for a few sites.
Comments Regarding SR	• Compulsory • Free text	Comments specific to the geologic slip rate.	Notes on how the slip rate was calculated, including how the offset was measured and uncertainties calculated or assigned. May also include additional background information and any special issues that are noted by the investigators or compilers.
Comments Regarding Dating	• Compulsory • Free text	Comments specific to the offset dating.	Includes information such as uncertainties about the dating technique and any ages with 1 sigma uncertainties (typically OSL). For radiocarbon ages, the calibration curve used should be listed.
Data Source	• Compulsory • Free text	List of the publications from which the slip rate data has been obtained.	The full reference should be included in the Reference worksheet in the format used by the GNS Science Bibliographic Database (https://online.gns.cri.nz/online/login/bib).
Initially Compiled By	• Compulsory • Free text	The name of the person who first compiled the data.	Useful for tracking purposes if someone else subsequently updates the data.
Last Updated	• Compulsory • Date • DD/MM/YYYY	The date that the data was last updated.	Earlier versions should be managed with version control (e.g. saved in an archive folder).
Updated By	• Compulsory • Free text	The name of the person who last updated the data.	This currently does not include the organisation of each person, but this may be added.
Last QA'd	• Compulsory • Date • DD/MM/YYYY	The date that the data was last reviewed.	Earlier versions should be managed with version control (e.g. saved in an archive folder).
QA'd By	• Compulsory • Free text	The name of the person who reviewed the data.	-

Table 2.3 Slip rate quality rankings.

Quality Ranking	Category	Description
QR1 Offset feature	A – Well constrained	Identifiable piercing line or feature (typically at or near the surface) is well documented or can be independently verified from mapping or logs. If the offset feature is in the subsurface (e.g. submarine seismic lines), the data is presented and correlations appear reliable. The offset feature is described or documented. A range and (or) best-estimate value with uncertainties is provided or can be obtained from the data.
	B – Moderately constrained	For surficial or near-surficial features, only a best estimate or single value is given for the feature offset, and documentation does not allow for a range of values to be determined. For other determinations of offset (e.g. cross-sections, seismic lines), the offset value may be determined indirectly and may be somewhat model dependent. For cases where no range of offset values is reported, there is some confidence that there are relatively small uncertainties on the offset feature. An example of this would be an uplifted marine terrace that is correlated to a known sea-level high stand. Such features typically have small measurement uncertainties, as uplift is usually measured relative to current sea level.
	C – Poorly constrained	Major assumptions are involved in measuring the offset, or no information is given as to how the offset was measured. Correlation of the feature may by suspect, or other alternatives possible, but not described well enough to understand the range of possible values.
	D – Very poorly constrained	Reported offset is suspect, or so poorly constrained, that the slip rate calculated is not considered reliable. For the paleoseismic site database, sites with this quality ranking have generally not been compiled.
QR2 Dating	A – Well constrained	Radiometric dates (e.g. calibrated radiocarbon ages with the calibration curve used documented) or correlation to a well-dated datum (e.g. Last Glacial Maximum aggradation terrace or moraine). Uncertainties reported.
	B – Moderately constrained	Low-quality (few samples or material that only loosely date the offset feature, e.g. charcoal) radiometric dates or a general correlation to a known datum (e.g. tephra) or climatic event. If uncertainties are reported, they may not be formal uncertainties and only loosely constrained.
	C – Poorly constrained	Highly uncertain correlation or dating constraints poorly documented or not described, or only a minimum or maximum age constraint. Slip rates based on relative soil development or weathering rind dating are typically considered poorly constrained.
	D – Very poorly constrained	Reported age is suspect or so poorly constrained that the slip rate calculated is not considered reliable. For the paleoseismic site database, sites with this quality ranking have generally not been compiled.

Quality Ranking	Category	Description
QR3 Overall	A – Well constrained	Offset feature and dating are well constrained. Slip rate is believed to represent deformation across the entire width of the fault zone. Offset is also believed to have accumulated over enough earthquakes sufficient to provide a robust average rate.
	B – Moderately constrained	One or both components of the slip rate are less than well constrained. Offset feature may not span the full width of the fault zone, but investigators provide an assessment as to the degree of this.
	C – Poorly constrained	One or both components of the slip rate are poorly constrained, and the rate may not be reliable. Offset may not span the entire fault zone or may represent only a limited number of earthquakes, so the reported slip rate is unlikely to represent the fault slip rate or a long-term average over multiple earthquakes.
	D – Unreliable	The slip rate calculated is not considered reliable because the offset or dating constraints are unreliable. For the paleoseismic site database, sites with this quality ranking have generally not been compiled.

2.2 Earthquake Timings and Recurrence Interval Worksheet

The Earthquake (EQ) Timings and Recurrence Interval (RI) worksheet is divided into two parts:

1. The left-hand part (columns A–J) contains Fault Data and is primarily data for specific faults or fault sections in the CFM. It should be noted that the CFM does not explicitly contain earthquake timings or recurrence intervals, but they are compiled for use in other projects, such as the NSHM 2022. The attributes are described in Table 2.4 and only need to be reported once for each fault. That is, they are not repeated for every site.
2. The right-hand part (columns K–AJ) contains Site Data; these are data for individual earthquakes at a site or, in some cases, attributes for two or more sites that have been combined together in publications. This worksheet therefore differs from the other two, in that each row does not correspond to an individual site or combination of sites. The Site Data are described in Tables 2.5, 2.6 and 2.7 and need to be reported for each site.

The data are currently listed in the EQ Timings and RI worksheet in alphabetical order of the Fault name. The use of grey shades and borders in the worksheet is simply to assist with reading of the data; there are no specific rules for their use. Note that, because the Fault Data is not reported for each site, sorting the spreadsheet by Fault name will cause some sites to lose association with the fault.

Table 2.4 Earthquake Timings and Recurrence Interval worksheet Fault Data attributes.

Attribute Name	Key Features	Definition	Additional Information
Fault	• Compulsory • Free text • Omit 'fault' from name	Name of the fault to which the site data is applicable.	Where possible, the name should be from the NZAFD. Some off-fault sites (e.g. lakes, wetlands, marine terraces) have been assigned to the faults that have been inferred to have caused the deformation they record. Others have not been assigned where the causative fault is uncertain.
CFM Name	• Compulsory • Free text • Spelled out in full	Name of the fault or fault section in the CFM to which the site data is applicable.	Most of the sites do not sit on a CFM fault in map view because of scale differences (CFM being regional scale), so a judgement call will need to be made as to which CFM fault the site is applicable. This may change as the CFM is updated and so will need to be updated accordingly.
CFM No.	• Compulsory • Number	Number of the fault or fault section in the CFM to which the site data is applicable.	This may change as the CFM is updated, so will need to be updated accordingly.
CFM Sense Dominant	• Compulsory • From a restricted list: - Dextral - Normal - Reverse - Sinistral	Dominant sense of movement on the fault or fault section in the CFM to which the site data is applicable.	This may change as the CFM is updated, so will need to be updated accordingly.
CFM Sense Secondary	• Optional • From a restricted list: - Dextral - Normal - Reverse - Sinistral	Secondary sense of movement on the fault or fault section in the CFM to which the site data is applicable.	This may change as the CFM is updated, so will need to be updated accordingly.

Attribute Name	Key Features	Definition	Additional Information
RI	• Optional • Number • Reported in years • Uncertainties should be included; in the format: - Preferred ± 2 sigma uncertainty, or - Preferred (minimum-maximum)	Recurrence interval inferred for a fault or fault section in the CFM to which the site data is applicable.	The original purpose of this attribute was to compile a recurrence interval for a fault or fault section in the CFM for use in the NSHM 2022. However, recurrence interval is not included as a specific parameter in the CFM; therefore, compilation has not been prioritised to date. What is currently recorded here is a reported recurrence interval for a CFM fault or fault section calculated from earthquake timing data (generally from inter-event times). ± can be entered quickly using ALT-0177. This may be back-populated or updated in the future for comparison with the earthquake timings site data. This will require documentation of the calculation method used in recurrence interval comments.
No. of Earthquakes	• Optional • Number	The number of earthquakes used to calculate the recurrence interval reported in the RI column.	As noted above, recurrence interval is currently not included as a specific parameter in the CFM. However, if the recurrence interval attribute has been compiled from publications, then this attribute may also be compiled from the same publications. This may be back-populated or updated in the future for comparison with the earthquake timings site data.
QR RI	• Compulsory • A, B or C, whereby: - A (high) = from one or more sites with an Overall QR of 1 - B (medium) = from one or more sites with an Overall QR of 2 - C (low) = from one or more sites with an Overall QR of 3	Quality ranking of the recurrence interval.	As noted above, recurrence interval is currently not included as a specific parameter in the CFM. However, if the recurrence interval attribute has been compiled from publications, then this attribute must also be compiled. The quality ranking is derived from the Earthquake Timings Overall quality rankings in Table 2.7 but has been tested with only a few data, so may need revising.
RI Comments	• Compulsory • Free text	Comments regarding the recurrence interval determination.	Notes on how the recurrence interval was calculated and how uncertainties were calculated or assigned. May also include additional background information and any special issues that are noted by the investigators or compilers.

Table 2.5 Earthquake Timings and Recurrence Interval worksheet Site Data attributes.

Attribute	Key Features	Definition	Additional Information
Site DB ID	• Compulsory • Number	Site database identification number.	Entered in order of compilation and can have gaps but not duplicates. Should match the Site ID number in the GIS feature class (Section 3).
Site Name	• Compulsory • Free text or number	Name or number of the site.	From the original publication. Or, if there is none, one should be added, describing the location and preferably the feature (e.g. Hokuri Creek sections).
Easting and Northing	• Compulsory • Number • Reported as a full seven-digit figure • Coordinate System NZTM 2000	Site location XY coordinates.	From a reported grid reference or, if more appropriate, updated from the GIS feature class. For many sites, the locations were estimated using a published map and refined using orthophotos or LiDAR data, as noted in the Comments field in the GIS feature class. The location uncertainty is estimated in the Method_Accuracy field in the GIS feature class.
Reported RI	• Compulsory • Number • Reported in years	The recurrence interval as reported in the original study.	Because investigators report recurrence interval uncertainties in a variety of ways, this column is what was originally reported in the referenced study. ± can be entered quickly using ALT-0177.
No. of EQs	• Compulsory • Number	The number of earthquakes used to calculate the recurrence interval.	-
Reported RI Method	• Compulsory • Free text	Method used to calculate the recurrence interval in the original study.	The database currently deliberately prioritises recurrence intervals calculated from earthquake timings rather than from other methods, such as displacement divided by slip rate. These could be added in future if required but will require some different formatting if no earthquake timings are recorded at the site.

Attribute	Key Features	Definition	Additional Information
QR RI	• Compulsory • A (high), B (medium) or C (low), whereby: - A = ≥5 high or moderate-quality earthquakes (QR3 = 1 or 2) - B = ≥3 moderate-quality earthquakes (QR3 = 2) or calculated from three or more earthquakes during a well-dated time interval - C = ≥3 low-quality earthquakes (QR3 = 3) or calculated from three or more earthquakes during a poorly dated time interval	Quality ranking of the reported recurrence interval.	The individual quality rankings (QR1, QR2 and QR3) are defined in Table 2.6. This attribute has been tested with only a few data, so may need revising.
Earthquakes	• Optional • Free text • Following the convention: - EQ1 (LE) = last (most recent) event - EQ2 (PE) = penultimate event - EQ3 (APE) = antepenultimate event - EQ4, etc.	Earthquake number counting back in time from the last event.	There are a few sites that record earthquakes that cannot be explicitly separated out into specific events, based on poor dating controls. Some have been included, particularly if they have been used to calculate recurrence interval.
Reported EQ age	• Compulsory • Number • Include the reported units (e.g. AD, BC, ^{14}C yr BP, cal. yr BP, ka)	The ages of the specific earthquake as reported in the original study.	Because investigators report uncertainties in a variety of ways, this attribute is what was originally reported in the referenced study. Occasionally, the units are not stated and, if so, any assumptions made in converting these to the 'Old' and 'Young' attributes should be noted in the comments. ± can be entered quickly using ALT-0177.

Attribute	Key Features	Definition	Additional Information
Old (Pref., Min., Max.)	• Compulsory • Number • Reported in calibrated years BP.	The maximum age for the specific earthquake timing.	The use of 'Old' here is to signify that it is the older age constraint on the earthquake timing. If it is constrained by an individual age (e.g. a radiocarbon age), it may have preferred, minimum and maximum values; or, if it is a modelled age (e.g. using OxCal), it may have minimum and maximum values.
Young (Pref., Min., Max.)	• Compulsory • Number • Reported in calibrated years BP.	The minimum age for the specific earthquake timing.	The use of 'Young' here is to signify that it is the younger age constraint on the earthquake timing. If it is constrained by an individual age (e.g. a radiocarbon age), it may have preferred, minimum and maximum values; or, if it is a modelled age (e.g. using OxCal), it may have minimum and maximum values.
QR1 Earthquake Dating	• Compulsory • A, B or C	Quality ranking of the earthquake dating for the specified event.	Defined in Table 2.6.
QR2 Earthquake Evidence	• Compulsory • A, B or C	Quality ranking of the earthquake evidence.	Defined in Table 2.6.
QR3 Overall	• Compulsory • 1, 2 or 3	Overall earthquake quality (certainty) ranking.	Defined in Tables 2.6 and 2.7.
Comments	• Compulsory • Free text	Comments regarding the specific earthquakes.	Includes information such as earthquake evidence, dating technique and uncertainties and the calibration curve for radiocarbon dates. May also include additional background information and any special issues that are noted by the investigators or compilers.
Data Source	• Compulsory • Free text	List of the publications from which the earthquake timing data has been obtained.	The full reference should be included in the Reference worksheet in the format used by the GNS Science Bibliographic Database.
Initially Compiled By	• Compulsory • Free text	The name of the person who first compiled the data.	Useful for tracking purposes if someone else subsequently updates the data.

Attribute	Key Features	Definition	Additional Information
Last Updated	• Compulsory • Date • DD/MM/YYYY	The date that the data was last updated.	Earlier versions should be managed with version control (e.g. saved in an archive folder).
Updated By	• Compulsory • Free text	The name of the person who last updated the data.	This currently does not include the organisation of each person, but this may be added.
Last QA'd	• Compulsory • Date • DD/MM/YYYY	The date that the data was last reviewed.	Earlier versions should be managed with version control (e.g. saved in an archive folder).
QA'd By	• Compulsory • Free text	The name of the person who reviewed the data.	-

Table 2.6 Earthquake timings Site Data quality rankings.

Quality Ranking	Category	Description	Quality Ranking
QR1 Earthquake Dating	A – Good	Multiple (≥3) high-quality radiometric dates (e.g. AMS radiocarbon of a small fraction) or robust identification of a well-dated datum (e.g. microprobe identification of a tephra). Ages tightly bracket the event timing (e.g. obtained from immediately above or below an event horizon). Well-quantified uncertainties.	QR1 Earthquake Dating
	B – Average	Few and/or medium- to low-quality radiometric dates (e.g. pre-AMS radiocarbon, charcoal, OSL) or moderately constrained (e.g. field identification) correlation to a datum (e.g. tephra). Uncertainties reported.	
	C – Poor	No radiometric dates, or minimum or maximum dates only. Inferred correlation with a known datum (e.g. tephra), climatic event (e.g. end of LGM) or assumption (e.g. pre-1840 AD). Poor or no uncertainties.	
QR2 Earthquake Evidence	A – Strong	Well constrained from cross-cutting relationships in a trench (e.g. an event horizon) or a historical (documented by written records) earthquake.	QR2 Earthquake Evidence
	B – Moderate	Interpreted earthquake-related units in a trench (e.g. colluvial wedge), near-fault record (e.g. ponded sediments) or off-fault record with rigorous consideration of event origin (e.g. lakes).	
	C – Weak	Inferred earthquakes within unclear trench stratigraphy or from off-fault data with little consideration of event origin (e.g. landslide).	
Q3 Overall – see Table 2.7	1 – High	Both components – QR1 and QR2 – are high (A+A).	Q3 Overall – see Table 2.7
	2 – Moderate	One or both components – QR1 and QR2 – are moderate (A+B, B+A, B+B).	
	3 – Low	One or both components – QR1 and QR2 – are low (A+C, B+C, C+C, C+B, C+A).	

Table 2.7 Overall earthquake quality (certainty) ranking.

Matrix – Overall Earthquake Certainty Ranking (QR3) **1 = high, 2 = moderate, 3 = low**		**Quality of Earthquake Dating (QR1)**		
		A – Good	**B – Average**	**C – Poor**
Evidence Event is an Earthquake (QR2)	A – Strong	1	2	3
	B – Moderate	2	2	3
	C – Weak	3	3	3

2.3 Single-Event Displacement Worksheet

The Single-Event Displacement worksheet is divided into two parts:

1. The left-hand part (columns A–I) contains Fault Data and is primarily data for specific faults or fault sections in the CFM. It should be noted that the CFM does not explicitly contain single-event displacement, but it is compiled for use in other projects, such as the NSHM 2022. Their attributes are described in Table 2.8 and only need to be reported once for each fault. That is, they do not need to be repeated for every site.
2. The right-hand part (columns J–AA) is the Site Data and are the values obtained from an individual site or, in some cases, attributes for two or more sites that have been combined together in publications. The attributes are described in Tables 2.9 and 2.10 and need to be reported for each site.

The data are currently listed in the Single-Event Displacement worksheet in alphabetical order of the Fault name. The use of grey shading and borders in the worksheet is simply to assist with reading of the data; there are no specific rules for their use. Note that, because the Fault Data is not reported for each site, sorting the spreadsheet by Fault name will cause some sites to lose association with the fault.

Table 2.8 Single-Event Displacement worksheet Fault Data attributes.

Attribute Name	Key Features	Definition	Additional Information
Fault	• Compulsory • Free text • Omit 'fault' from name	Name of the fault to which the site data is applicable.	Where possible, the name should be from the NZAFD. Some off-fault sites (e.g. uplifted marine terraces) have been included.
CFM Name	• Compulsory • Free text • Spelled out in full	Name of the fault or fault section in the CFM to which the site data is applicable.	Most of the sites do not sit on a CFM fault in map view because of scale differences (CFM being regional scale), so a judgement call will need to be made as to which CFM fault the site is applicable. This may change as the CFM is updated and will need to be updated accordingly.
CFM No.	• Compulsory • Number	Number of the fault or fault section in the CFM to which the site data is applicable.	This may change as the CFM is updated, so will need to be updated accordingly.
CFM Sense Dominant	• Compulsory • From a restricted list: - Dextral - Normal - Reverse - Sinistral	Dominant sense of movement on the fault or fault section in the CFM to which the site data is applicable.	This may change as the CFM is updated, so will need to be updated accordingly.
CFM Sense Secondary	• Optional • From a restricted list: - Dextral - Normal - Reverse - Sinistral	Secondary sense of movement on the fault or fault section in the CFM to which the site data is applicable.	This may change as the CFM is updated, so will need to be updated accordingly.
CFM Dip	• Compulsory • Number • Reported in degrees (0–90°) • Uncertainties should be included; in the format: - Preferred ± 2 sigma uncertainty, or - Preferred (minimum-maximum)	Dip assigned to the fault or fault section in the CFM to which the site data is applicable.	This may change as the CFM is updated, so will need to be updated accordingly. ± can be entered quickly using ALT-0177.

Attribute Name	Key Features	Definition	Additional Information
CFM Rake	• Compulsory • Number • Reported in degrees using the Aki and Richards (1980) right-hand-rule convention whereby: - 180/-180° = Dextral - -90° = Normal - 0° = Sinistral - 90° = Reverse • Uncertainties should be included; in the format: - Preferred ± 2 sigma uncertainty, or - Preferred (minimum-maximum)	Rake of the net sense of movement on the fault or fault section in the CFM to which the site data is applicable.	This may change as the CFM is updated, so will need to be updated accordingly. ± can be entered quickly using ALT-0177.
SED Net	• Compulsory • Number • Reported in metres (m) • Uncertainties should be included; in the format: - Preferred (minimum-maximum)	Net single-event displacement for the fault or fault section in the CFM to which the site data is applicable.	The original purpose of this attribute was to compile a net single-event displacement for a fault or fault section in the CFM for use in the NSHM 2022. However, single-event displacement is not included as a specific parameter in the CFM; therefore, compilation has not been prioritised to date. What is currently recorded here is a reported single-event displacement for a CFM fault or fault section. This is either calculated from multiple single-event displacement measurements at a site (i.e. single-event displacement from multiple earthquakes) or single-event displacements measured at multiple sites along a fault. This information should be captured in the Net SED comments attribute.
Net SED Comments	• Compulsory • Free text	Comments regarding the single-event displacement for the fault or fault section in the CFM to which the site data is applicable.	Notes on how the single-event displacement was calculated and the data source. May also include additional background information and any special issues that are noted by the investigators or compilers. For example, this may include if the SED is considered to be small and contemporaneous with a larger SED on a nearby fault (i.e. secondary deformation).

Table 2.9 Single-Event Displacement worksheet Site Data attributes.

Attribute Name	Key Features	Definition	Additional Information
Site DB ID	• Compulsory • Number	Site database identification number.	Entered in order of compilation and can have gaps but not duplicates. Should match the Site ID number in the GIS feature class (Section 3).
Site Name	• Compulsory • Free text or number	Name or number of the site.	From the original publication. Or, if there is none, one should be added, describing the location and preferably the feature (e.g. Hokuri Creek channels 1).
Easting and Northing	• Compulsory • Number • Reported as a full seven-digit figure • Coordinate System NZTM 2000	Site location XY coordinates.	From a reported grid reference or, if more appropriate, updated from the GIS feature class. For many sites, the locations were estimated using a published map and refined using orthophotos or LiDAR data, as noted in the Comments field in the GIS feature class. The location uncertainty is estimated in the Method_Accuracy field in the GIS feature class.
Reported SED	• Compulsory • Number • Reported in metres (m)	The single-event displacement as reported in the original study.	Because investigators report single-event displacement uncertainties in a variety of ways, this column is what was originally reported in the referenced study. Where two components have been reported, they are entered one above the other. They can be separated using ALT-ENTER.
Reported Component (SED)	• Compulsory • From a restricted list: - DS = Dip Slip - LL = Left Lateral - NS = Net - RL = Right Lateral - V = Vertical	Component of single-event displacement reported in the Reported SED column.	Where two components have been reported, they are entered one above the other. They can be separated using ALT-ENTER.
No. of Measurements	• Compulsory • Number	The number of measurements used to measure or calculate the single-event displacement for a specific earthquake rupture.	In recent studies, displacements are calculated multiple times to obtain uncertainties (e.g. measured three times to calculate the preferred, minimum and maximum values).

Attribute Name	Key Features	Definition	Additional Information
SED Net	• Compulsory • Number • Reported in metres (m) • Uncertainties should be included; in the format: - Preferred (minimum-maximum)	Net single-event displacement calculated or corrected for the CFM Dip (if significantly different than used in the original study).	Many single-event displacements are reported with only the horizontal or vertical component, so a net single-event displacement needs to be calculated. This can be done using the Net Calcs worksheet. If only one component is reported, and there is reason to believe there is, or could be, another component of motion, then the net single-event displacement should be prefixed with a > or ≥, respectively. If the CFM fault dip changes, or if there is reason to believe that the CFM dip is not representative of the dip at the ground surface, this value will need to be changed accordingly.
Net SED comments	• Compulsory • Free text	Comments regarding the net single-event displacement.	Notes on how the net single-event was derived, including any assumptions about the net single-event displacement being a minimum or maximum value (e.g. if one or more H or V components are not reported). May also include additional background information and any special issues that are noted by the investigators or compilers.
QR SED	• Compulsory • A, B or C	Quality ranking of the single-event displacement.	Defined in Table 2.10.
Offset Feature	• Compulsory • Free text	Type of feature that is offset.	For example: terrace riser, channel, moraine, ridge.
Comments regarding SED	• Compulsory • Free text	Comments specific to the single-event displacement.	Includes information such as how the offset was measured and calculation of uncertainties. For sites of multiple earthquake displacements, this includes which earthquake the single-event displacement applies to and the evidence for the number of earthquakes inferred. If from a historical earthquake, the name and year of the earthquake is recorded here.
Data Source	• Compulsory • Free text	List of the publications from which the slip rate data has been obtained.	The full reference should be included in the Reference worksheet in the format used by the GNS Science Bibliographic Database.
Initially Compiled By	• Compulsory • Free text	The name of the person who first compiled the data.	Useful for tracking purposes if someone else subsequently updates the data.

Attribute Name	Key Features	Definition	Additional Information
Last Updated	• Compulsory • Date • DD/MM/YYYY	The date that the data was last updated.	Earlier versions should be managed with version control (e.g. saved in an archive folder).
Updated By	• Compulsory • Free text	The name of the person who last updated the data.	This currently does not include the organisation of each person, but this may be added.
Last QA'd	• Compulsory • Date • DD/MM/YYYY	The date that the data was reviewed (quality assured).	Earlier versions should be managed with version control (e.g. saved in an archive folder).
QA'd By	• Compulsory • Free text	The name of the person who reviewed the data.	-

Table 2.10 Single-event displacement quality rankings.

Quality Ranking Category	Description
A – High	Multiple (≥3) measurements and/or measurements of single-event displacement for a specific earthquake rupture using high-resolution data (e.g. RTK GNSS, LiDAR) and clear fault-normal piercing lines that intersect a simple, well-defined fault zone. Well-quantified uncertainties.
B – Medium	Few (1–2) measurements and/or multiple measurements of single-event displacement for a specific earthquake rupture using features that are oblique to a broad, poorly defined fault zone. Uncertainties reported.
C – Low	Few (1–2) measurements of single-event displacement for a specific earthquake rupture or estimates from field data. No uncertainties reported, or a minimum or maximum only, or measurement method not stated.

3.0 PALEOSEISMIC SITE DATABASE GIS FEATURE CLASS

The GIS feature class contains data on paleoseismic sites for all three datasets (Slip Rate, EQ Timings and RI and Single-Event Displacement) in a single file.[2] The fields are a subset of those in the *AF.Points* layer in the NZAFD (Jongens and Dellow 2003), with five additional fields:

1. SiteDB_ID
2. Point_X
3. Point_Y
4. Parameter
5. Compiler.

There have also been some additions to the domains (restricted lists) within some fields to accommodate inclusion of new types of data (e.g. historical earthquake fault surface ruptures). The domains have also been re-ordered from the AF.Points Data Dictionary (Jongens and Dellow 2003) into alphabetical order to facilitate searching. The GIS feature class fields are described in Table 3.1, and a template is included as Appendix 2.

2 Several sites, such as trench sites, have three records in the GIS file – one for slip rate, one for earthquake timings and recurrence intervals and one for single-event displacement. They all have the same grid reference but a unique Site DB ID.

Table 3.1 Site Database feature class fields.

Attribute Name	Key Features	Definition	Additional Information
FIELD_NUMBER	• Optional • Free text	The reported site number in published studies.	This is a compulsory field in the NZAFD but is often not reported in publications.
FEATURE_NAME	• Compulsory • Free text	Name of the feature.	For example: ‘Saxton River T1/T2 riser’, ‘Haast Trench’.
METHOD	• Compulsory • From a restricted list: - Air Photo 1:500 - Air Photo 1:1000 - Air Photo 1:5000 - Air Photo 1:16,000 - Air Photo 1:25,000 - DSM - GPS – Differential - GPS – Field - GPS – RTK - Grid Reference - LiDAR - Location name - Map – <=1:20,000 scale - Map – 1:25,000 scale - Map – 1:50,000 scale - Map – 1:100,000 scale - Map – >=1:250,000 scale - Map – scale not specified - Orthophoto - Surveyed - Verbal description	The method used for locating the site.	This is the method used to locate the site geographically in this database. It is often initially compiled from a grid reference or a rough location eyeballed from a published map and then refined using LiDAR maps or orthophotos. This should be noted in the OTHER_INFO field. Verbal description is used in this compilation where the site location is from a description in the text (e.g. where there is no map or the map is not useful for a precise location). Generally, only one method should be given, but, if it is necessary to include two, they should be separated by a semicolon (**;**). The options in the restricted list should be entered in exactly the same format, i.e. the same use of capitals and spacing.

Attribute Name	Key Features	Definition	Additional Information
METHOD_ACC	• Compulsory • Number • Reported in metres (m)	The accuracy of the location method.	Generally, this should be lower for sites located using high-resolution data, such as LiDAR or GPS – RTK. Default values of 10 m and 5 m have often been assigned to these, respectively. However, in many cases, a precise site location is not possible or needed because it is approximating the centre-point of an offset feature (e.g. an offset terrace riser) comprising two separate elements – one element of an offset feature on either side of the fault.
FAULT_FEATURE	• Compulsory • From a restricted list (*can be prefixed with the word 'possible'*): - Base of scarp - Crush zone - Degraded scarp - Fault-associated - Fault-controlled - Fault in exposure - Fault in trench - Fold - Graben - Historical rupture - Modified scarp - Off-fault site - Scarp - Top of scarp	The fault feature that is captured at this site.	This is the feature that is interpreted to mark the fault at this location. It differs from the landscape feature that the fault deforms, which is the GEO_FEATURE field. The options in the restricted list must be entered exactly in the same format, i.e. the same use of capitals and spacing. **Restricted list definitions:** • Base of scarp – base of fault scarp where scarp has significant width. • Crush zone – zone of crushed rocks associated with fault movement. • Degraded scarp – eroded fault scarp. • Fault-associated – GEO_FEATURE is associated with fault, e.g. ridge rent, landslide, underfit stream. • Fault-controlled – GEO_FEATURE is controlled by fault , e.g. saddle, line of springs, guided stream. • Fault in trench – fault located in a trench. • Fault in exposure – fault located in an exposure. • Fold – technically not a fault feature, but some faults are expressed as a fold at the surface. • Graben – depressed block between fault scarps. • Historical rupture – ground-surface fault rupture in a post-1840 AD earthquake. The year and name should be noted in the OTHER_INFO field. • Modified scarp – scarp modified by anthropogenic activities. • Off-fault site – technically not a fault feature, but a general encompassing term for sites such as cores or marine terraces that are situated away from faults. • Scarp – step in the ground surface inferred to have been caused by displacement on a fault. • Top of scarp – top of fault scarp where scarp has significant width.

Attribute Name	Key Features	Definition	Additional Information
GEO_FEATURE	• Optional • From a restricted list, which can be added to: - Abandoned channel - Base of faceted spur - Beheaded stream - Break in slope - Captured stream - Damaged ridge - Displaced aggradational surface - Displaced degradational surface - Displaced sedimentary horizon - Displaced surface (unknown aggradational or degradational surface) - Displaced tephra horizon - Guided stream - Head of landslide - Head of spring (uppermost extent of spring) - Landslide - Line of springs - Offset channel (paleo stream) - Offset cultural feature - Offset fan - Offset gully - Offset lahar - Offset landslide - Offset marsh edge - Offset moraine - Offset ridge	The geological or geomorphological feature that is formed, displaced or associated/controlled by the active fault.	This is the landform feature that is deformed by the fault, not the fault itself. This generally matches the offset feature in the site database spreadsheet. Sometimes publications describing trenches or exposures do not specifically mention the landform, so the unit displaced (e.g. tephra or sedimentary horizon) is noted. Offset cultural feature was added in this compilation for anthropogenic features (e.g. fences, roads, tracks) offset in historical earthquakes. The options in the restricted list must be entered in exactly the same format, i.e. the same use of lower case and spacing.

Attribute Name	Key Features	Definition	Additional Information
	- Offset riser - Offset river - Offset slope - Offset spur - Offset stream - Offset terrace - Offset valley - Ponded drainage - Ridge rent (gravity failure feature that may be associated with fault movement) - Saddle - Seafloor - Spring - Treadwidth - Underfit stream - Uplifted terrace		
GEO_INFO	• Optional • Free text • Examples are: - Cultural feature details (fence, road, track, etc.) - Trench name* - Treadwidth in metres down- and upstream side - Fault scarp rejuvenated - Youngest surface preserved age - Youngest surface displaced age - Age of the tephra layer or sedimentary horizon - Dating of landslide	Geological or geomorphic information relevant to active fault movement history.	This has been broadened out to be more than just the geological or geomorphic information, as it can be useful to have other information here rather than in the OTHER_INFO field. ** Although the trench name is often the FEATURE_NAME, it is useful to re-state it here for searching and making trench site maps. The preferred format is: TRENCH. Name. (For example: TRENCH. Glenledi Road T16/01.)*

Attribute Name	Key Features	Definition	Additional Information
DOM_SENSE	• Optional • From a restricted list: - dextral - dipslip – if cannot tell whether normal or reverse - normal - reverse - sinistral - strikeslip – if cannot tell whether dextral or sinistral	Dominant or primary sense of movement on the fault.	Although this is listed as optional, this should be compiled for each site. The options in the restricted list must be entered in exactly the same format, i.e. lower case and single words. This differs from the spreadsheet, which uses capitals and abbreviations, in order to be compatible with the NZAFD.
SUB_SENSE	• Optional • From a restricted list: - dextral - dipslip – if cannot tell whether normal or reverse - normal - reverse - sinistral - strikeslip – if cannot tell whether dextral or sinistral	Subordinate or secondary sense of movement for oblique-slip faults.	The options in the restricted list must be entered in the same format, i.e. lower case and single words. This differs from the spreadsheet, which uses capitals and abbreviations, in order to be compatible with the NZAFD.
DOWN_QUAD	• Optional • From a restricted list: - N - NE - E - SE - S - SW - W - NW	Downthrown side of fault trace described in compass quadrants.	-

Attribute Name	Key Features	Definition	Additional Information
DIP	• Optional • Free text • Reported in degrees (0–90°) • Can be reported in the format (in order of decreasing preference): - 60 ± 5 - 60 (50–75) - 50–75 - c. 60 - c. 50–75 - ≤60 - ≥60 - <60 - >60	Mean dip angle of the fault plane.	The NZAFD states that the dip should be a generalised dip that represents the entire trace, though in practise it is often the dip measured in an exposure or trench. This has not been compiled for many sites to date.
DIP_DIR	• Optional • Free text • Reported in azimuth degrees (1–360°) • Can be reported as a single number or a range: - 315 - 290–345	Mean dip direction of the fault plane for the entire trace.	0° should be reported as 360°. For pre-existing quadrant reporting, e.g. NW, convert to a range of c. 290°–340° (25° either side of 315°). This has not been compiled for many sites to date.
STRIKE_SR	• Optional • Free text • Reported in millimetres per year (mm/yr = m/kyr) • Can be reported in the formats (in order of decreasing preference): - 1.5 ± 0.5 - 1.5 (1.0–2.5)	Strike-slip (horizontal) slip rate.	Although this is listed as optional, this should be compiled for any site for which this (right lateral, RL, or left lateral, LL) has been compiled in the slip rate worksheet.

Attribute Name	Key Features	Definition	Additional Information
	- 1.5 +1.0/-0.5 - 1.0–2.5 - c. 1.5 - c. 1.0–2.5 - ≤1.5 ± 0.5 - ≥1.5 ± 0.5 - <1.5 ± 0.5 - >1.5 ± 0.5 - ≤1.5 +1.0/-0.5 - ≥1.5 +1.0/-0.5 - <1.5 +1.0/-0.5 - >1.5 +1.0/-0.5 - ≤1.5 (1.0–2.5) - ≥1.5 (1.0–2.5) - <1.5 (1.0–2.5) - >1.5 (1.0–2.5) - ≤1.5 - ≥1.5 - <1.5 - >1.5		
STRIKE_SED	• Optional • Free text • Reported in metres (m) • Can be reported in same formats as STRIKE_SR	Strike-slip (horizontal) single-event displacement.	Although listed as optional, this should be compiled for any site for which this (RL or LL) has been compiled in the single-event displacement worksheet. If this is from a historical earthquake, the year and name of the earthquake should be recorded in the OTHER_INFO field.
VERT_SR	• Optional • Free text • Reported in millimetres per year (mm/yr = m/kyr) • Can be reported in same formats as STRIKE_SR	Vertical slip rate.	Although listed as optional, this should be compiled for any site for which this has been compiled (V) in the slip rate worksheet.

Attribute Name	Key Features	Definition	Additional Information
VERT_SED	• Optional • Free text • Reported in metres (m) • Can be reported in same formats as STRIKE_SR	Vertical single-event displacement.	Although listed as optional, this should be compiled for any site for which this has been compiled (V) in the single-event displacement worksheet. If this is from a historical earthquake, the year and name of the earthquake should be recorded in the OTHER_INFO field.
DIP_SR	• Optional • Free text • Reported in metres (m) • Can be reported in same formats as STRIKE_SR	Dip-slip slip rate.	Although listed as optional, this should be compiled for any site for which this has been compiled (DS) in the slip rate worksheet. This has generally only been compiled in the GIS feature class when it has been reported in a publication.
DIP_SED	• Optional • Free text • Reported in metres (m) • Can be reported in same formats as STRIKE_SR	Dip-slip single-event displacement.	Although listed as optional, this should be compiled for any site for which this has been compiled (DS) in the single-event displacement worksheet. This has generally only been compiled in the GIS feature class if it has been reported in a publication.
NET_SR	• Optional • Free text • Reported in millimetres per year (mm/yr = m/kyr) • Can be reported in same formats as STRIKE_SR	Net slip rate.	This has generally only been compiled in the GIS feature class if it has been reported in a publication.
NET_SED	• Optional • Free text • Reported in metres (m) • Can be reported in same formats as STRIKE_SR	Net single-event displacement.	This has generally only been compiled in the GIS feature class if it has been reported in a publication.
RI	• Optional • Free text • Reported in years • Can be reported in same formats as STRIKE_SR	Recurrence interval.	This has generally only been compiled in the GIS feature class if it has been reported in a publication.

Attribute Name	Key Features	Definition	Additional Information
LE	• Optional • Number • Reported in calibrated years before 1950 AD (cal. yr BP)	Last (most recent) event (earthquake).	The age is given in years prior to 1950 AD for compatibility with paleoseismicity ages derived from radiocarbon dates, e.g. an 1857 AD event = 93. Events after 1950 AD will be given a negative number, e.g. the 1987 Edgecumbe earthquake is given a value of -37. Although listed as optional, this should be compiled for any site for which this has been compiled in the EQ Timings and RI worksheet and for historical ground-surface ruptures in the Single-Event Displacement worksheet.
BIB_ID	• Optional • Free text	Bibliographic Database ID number.	To be obtained from the GNS Bibliographic Database for each published data source cited. If it is not included in the Bibliographic Database, it may be able to be added. Separate multiple entries with a comma.
TRACE_ID	• Optional • Number	The unique identification number of the parent trace.	This is compulsory in the NZAFD and should be obtained from *AF.Traces*. However, in practise, *AF.Traces* is in many places incomplete and so some sites do not sit on traces. A judgement call should be used as to whether to assign sites to the nearest trace or to not fill it in. If the trace changes, this needs to be updated.
SOURCE	• Compulsory • Free text • For published papers, reports and theses, examples of formats are: - Beanland (1989) - Amos et al. (2010) - Barnes and Pondard (2010) - Barrell et al. (2002, 2005) - Beanland (1995); Langridge and Ries (2014) - Langridge et al. (2005a, 2005b)	The name or shortened reference of the report, paper or published map that this data is, or will be, sourced from.	If the data is unpublished, then a description needs to be entered to allow users to find the data, e.g. author and project name. Separate multiple entries with a semicolon (;).

Attribute Name	Key Features	Definition	Additional Information
DATE_	• Optional • Free text • Reported as a date, e.g. 15/11/2010	Date of data collection/interpretation.	Often this is not known for legacy data, or only the year is known. This has not been compiled for many sites to date.
CONFID	• Compulsory if confidentiality clause exists between client and data owner • Reported as a date, e.g. 15/11/2010	Expiry date of confidentiality clause between the client and data owner.	Note that this does not actually hide the data from anyone who has access to the database. This has not been compiled for any sites to date.
ACCESS	• Compulsory • Text • From a restricted list: - Public - GNS only - GNS paleoseismologists only - Embargoed	Level of access to the data.	Note that this does not actually hide the data from anyone who has access to the database. Where ACCESS = Public, data can be published and passed on to external organisations. Acknowledgement must be given to the OWNER of the data. Where ACCESS = GNS only, data cannot be published or given to external organisations. Data can be used for research purposes with the permission of the relevant staff member. Where ACCESS = GNS paleoseismologists only, this usually means that a confidentiality clause exists between GNS and a client. In this case, data cannot be used for anything except for that stated in the confidentiality clause. Embargoed has been added for university data that is not to be published or passed on to external organisations.
AUTHOR	• Compulsory • Free text • Full or preferred names, if known, otherwise initials as in the associated publications	Author of the person who collected the data.	This is the person who collected or published the data, not the person who entered the data (which is entered in Compiler). For multiple authored studies, it is recommended to include the first author and any other key authors, including at least one GNS author, to assist with seeking of additional information in the future. Can include ‘et al.’. These should be separated by a comma.

Attribute Name	Key Features	Definition	Additional Information
OWNER	• Compulsory • Free text • Examples include: - Canterbury University - GNS Science - NIWA - Otago University - Victoria University of Wellington	Owner of the data.	In the NZAFD, this is a restricted list that can be added to, but there are multiple universities and organisations represented by the authors of the data now compiled. For multiple authored studies, it is recommended to include the organisation from the first author, GNS Science (if applicable) and then a judgement call on other organisations. These should be separated by a comma.
OTHER_INFO	• Optional • Free text	Additional information.	Recommend including any information about how the site was located and any uncertainties associated with that. Also, historical earthquake year and name.
SiteDB_ID	• Compulsory • Unique identifier	Site database identification number.	Must match the Site DB ID in the spreadsheet. Entered in order of compilation and can have gaps but not duplicates.
POINT_X and POINT_Y	• Compulsory • Number • Reported as a full seven-digit grid reference • Coordinate System NZTM 2000	Site location XY coordinates.	These have generally been calculated in the GIS after some refinements of the site locations using LiDAR data or orthophotos. These must match the Easting and Northing in the spreadsheet.
Parameter	• Compulsory • Text • From a restricted list: - SR (slip rate) - SED (single-event displacement) - EQ timings (earthquake timings and recurrence interval)	The site dataset compiled at this site.	Because the site database site IDs are unique numbers regardless of the type of data compiled, the parameter can be used to separate out the data compiled in the three worksheets in the site database.
Compiler	• Compulsory • Free text	The name of the person who compiled the data in the GIS feature class.	-

4.0 ACKNOWLEDGMENTS

The site database was compiled by Nicola Litchfield, Jade Humphrey (University of Canterbury), Robert Langridge, Genevieve Coffey and Regine Morgenstern (GNS Science), and they made helpful suggestions and changes to the attributes and the Data Dictionary during compilation. Russ Van Dissen and Regine Morgenstern (GNS Science) reviewed this report.

5.0 REFERENCES

Aki K, Richards PG. 1980. Quantitative seismology: theory and methods. Volume 1. San Francisco (CA): W.H. Freeman. 557 p.

Field EH, Biasi GP, Bird P, Dawson TE, Felzer KR, Jackson DD, Johnson KM, Jordan TH, Madden C, Michael AJ, et al. 2013. The Uniform California Earthquake Rupture Forecast, version 3 (UCERF3) – the time-independent model. Reston (VA): US Geological Survey. 97 p. + appendices. USGS Open-File Report 2013-1165.

Jongens R, Dellow GD. 2003. The active faults database of NZ: data dictionary. Lower Hutt (NZ): Institute of Geological & Nuclear Sciences. 27 p. (Institute of Geological & Nuclear Sciences science report; 2003/17).

Litchfield NJ, Humphrey J, Morgenstern R, Langridge RM, Coffey GL, Van Dissen RJ. 2022. New Zealand Paleoseismic Site Database: design and overview of version 1.0. Lower Hutt (NZ): GNS Science. 27 p. (GNS Science report; 2021/52).

Litchfield NJ, Van Dissen R, Sutherland R, Barnes PM, Cox SC, Norris R, Beavan RJ, Langridge R, Villamor P, Berryman K, et al. 2014. A model of active faulting in New Zealand. *New Zealand Journal of Geology and Geophysics*. 57(1):32–56. doi:10.1080/00288306.2013.854256.

Seebeck H, Van Dissen RJ, Litchfield NJ, Barnes PM, Nicol A, Langridge RM, Barrell DJA, Villamor P, Ellis SM, Rattenbury MS, et al. 2022. New Zealand Community Fault Model – version 1.0. Lower Hutt (NZ): GNS Science. 97 p. (GNS Science report; 2021/57).

Van Dissen RJ, Seebeck H, Litchfield NJ, Barnes P, Nicol A, Langridge RM, Barrell DJA, Villamor P, Ellis SM, Rattenbury MS, et al. 2021. Development of the New Zealand community fault model – version 1.0. In: *NZSEE 2021 Annual Technical Conference: turning challenges into positive legacies*; 2021 Apr 14–16; Christchurch, New Zealand. Wellington (NZ): New Zealand Society for Earthquake Engineering. Paper 92.

APPENDICES

This page left intentionally blank.

APPENDIX 1 SPREADSHEET TEMPLATE

The attached spreadsheet is a blank template that can be used for future compilation of the New Zealand Paleoseismic Site Database. It contains five worksheets:

1. Slip Rate
2. Earthquake (EQ) Timings and Recurrence Interval (RI)
3. Single-Event Displacement
4. Reference List
5. Restricted Lists.

Data can be entered into worksheets 1–4 using the definitions in this report. Some cells in worksheets 1–3 have domains (drop-down restricted lists). These are defined in worksheet 5, so worksheet 5 should not be edited.

APPENDIX 2 GIS FEATURE CLASS TEMPLATE

A blank GIS feature class is provided that may be useful for future compilation of the New Zealand Paleoseismic Site Database and for uploading into the *AF.Points* layer into the New Zealand Active Faults Database. The feature class is contained within a geodatabase and is able to be downloaded with the report from the GNS Science Online Shop.

www.gns.cri.nz

Principal Location

1 Fairway Drive, Avalon
Lower Hutt 5010
PO Box 30368
Lower Hutt 5040
New Zealand
T +64-4-570 1444
F +64-4-570 4600

Other Locations

Dunedin Research Centre
764 Cumberland Street
Private Bag 1930
Dunedin 9054
New Zealand
T +64-3-477 4050
F +64-3-477 5232

Wairakei Research Centre
114 Karetoto Road
Private Bag 2000
Taupo 3352
New Zealand
T +64-7-374 8211
F +64-7-374 8199

National Isotope Centre
30 Gracefield Road
PO Box 30368
Lower Hutt 5040
New Zealand
T +64-4-570 1444
F +64-4-570 4657