Design Rainfall Intensity–Frequency–Duration Surfaces for Australia

Methodology, Validation and Comparison with BoM 2016, LIMB 2021 and HARC 2024

Date: September 2026

Data: Bureau of Meteorology (BoM) daily rainfall and BoM Water Data Online (WDO) sub-daily rainfall, records to June 2026

Outputs: 25 durations (5 min – 7 days) × 13 AEPs (63.2% – 0.05%), 1 km grid (GDA94 / Australian Albers, EPSG:3577)

Author: Callan Schonrock

Contents

Executive summary

This report documents a new set of design rainfall Intensity–Frequency–Duration (IFD) surfaces for Australia, derived from Bureau of Meteorology (BoM) rain gauge records up to June 2026. The surfaces cover 25 durations from 5 minutes to 7 days and 13 annual exceedance probabilities (AEPs) from 63.2% to 1 in 2,000, on a 1 km grid. The work follows the same broad approach as the BoM 2016 IFDs that underpin Australian Rainfall and Runoff 2019 (ARR 2019): the Generalised Extreme Value (GEV) distribution, fitted by L-moments, with the L-moments regionalised across space. The main differences are in how the annual maxima are extracted and how the regionalisation is done.

What was done, in brief:

- Annual maxima were extracted with an IFD Calculator that treats missing data as missing (not as zero rain), uses data-quality codes, detects daily-read and accumulated totals inside pluviograph records and flags ("censors") the years in which they could hide a larger burst, and records the water year and date of every maximum. Water years with more than 25% of the record missing are not used.
- Gauges with at least 10 years between first and last record were processed. Short records inform the **index rainfall** (from 12 usable years), while L-CV and L-skewness, which need longer records, use gauges with at least 20 usable years. The mean of each short record is adjusted for the wetness of the years it happens to cover, using long-record neighbours.
- After quality control, 13,251 daily gauges (685,480 station-years) and 1,924 sub-daily gauges (54,207 station-years) were used.
- Restricted-to-unrestricted factors for daily gauges were derived from co-located daily and sub-daily gauges rather than assumed.
- Each gauge's annual maxima were summarised by three L-moments per duration: the mean (the **index rainfall**), L-CV (spread) and L-skewness (tail shape). Their sampling uncertainty was estimated by bootstrap, so that **short records carry less weight** in every subsequent step.
- Each L-moment was mapped with a **smooth covariate model plus a Gaussian-process (kriging) correction**, using covariates built for this study on a 1 km grid, including **direction-aware terrain descriptors**. The dense daily network sets the 1–7 day surfaces; sub-daily durations are modelled as **ratios to the 1-day value** at the pluviograph network and anchored back to the daily surface.
- Depths are smoothed across duration and forced to increase with duration and rarity at every cell.
- Accuracy was tested in two ways: **spatial block cross-validation** (whole 150 km blocks of gauges withheld), and a **hindcast** in which the surfaces were rebuilt from data up to 2012 only (the same period as BoM 2016) and both products were scored on the 2013–2026 annual maxima that neither had used.

Key findings. Cross-validation scores are measured against at-site estimates built from this study's own annual maxima and conventions, which favours this study. The hindcast (Section 8.5) is the fairest comparison with BoM 2016, and is given the most weight below. Because a single storm affects many gauges in the same year, the hindcast uncertainty was estimated by resampling whole years; differences are called significant only when the 95% interval excludes zero.

- **Daily durations (1–7 days): the products cannot be separated.** In the hindcast this study scores 1% better than BoM 2016 on a quantile score (95% interval −2.4% to +0.2%). BoM 2016's 1% AEP depths were exceeded 1.07 times as often as expected over 2013–2026 and this study's 1.36 times, but the intervals (0.59–1.65 and 0.79–2.03) overlap almost entirely and a paired test finds no difference in calibration. In cross-validation BoM 2016 has the smaller typical error (about 11% against 13% at 1% AEP), partly because many of these gauges were used to build it.
- **Short sub-daily durations (≤ 1 hour to 18 hours): this study is significantly closer to the pluviograph records.** In the hindcast it scores 5% (≤ 1 h; interval 2–8%) and 3% (1.5–18 h; interval 1–4%) better than BoM 2016, and it is significantly better calibrated: its 1% AEP depths were exceeded 2.1–2.3 times as often as expected against 2.8–3.2 times for BoM 2016. **At**

sub-daily gauges for 1–7 days BoM 2016 is better calibrated (2.0 against 2.4 times), with equal quantile scores.

- **Recent sub-daily extremes exceed both products.** Over 2013–2026, sub-daily annual maxima exceeded the 1% AEP depths of both products 2–3 times as often as expected, and 1.9–2.2 times as often even for this study's surfaces fitted to all data (which include those years). Recent pluviograph records are heavier-tailed than the long daily records at the same sites. This is consistent with recent intensification (HARC, 2024, adjusted its estimates to a 2010 climate for this reason) but may also reflect changes in instruments and logging.
- **Restricted-to-unrestricted factors** derived from about 500 co-located gauge pairs are 1.14 (1 day), 1.05 (2 days), 1.04 (3 days) and about 1.02 for 4–7 days. These are close to HARC's (2024) values for South East Queensland (1.15, 1.059) and smaller than those used by BoM 2016 for 2–6 days (1.11–1.03), which accounts for much of the 2–4% by which BoM 2016 exceeds this study at those durations.
- **Instrument change does not explain the recent sub-daily extremes.** At about 300 gauges that switched from chart pluviographs to tipping buckets, the chart-era short-duration maxima are 3–5% lower relative to their 1-day maxima, but gauges that were tipping buckets throughout show a similar drift over the same span, and the net step is not significant. No instrument adjustment was applied.
- **South East Queensland.** At the 91 BoM sub-daily gauges in the HARC extent, all products are below the recorded 1% AEP depths and this study is furthest below (−18.6%%, against −11.5%% for BoM 2016 and −12.7%% for HARC 2024). Admitting shorter records did not close this gap. **Within the HARC and LIMB extents the local studies should be preferred for sub-daily durations**; for daily durations all products agree within a few percent.
- **Terrain matters in a direction-dependent way.** Windward slopes facing easterly to north-easterly flow raise extreme daily rainfall along the eastern escarpments (up to +15–20% in the Wet Tropics) beyond what mean annual rainfall explains.

Status and recommended use. This is a research dataset. It is internally consistent (smooth, and monotone in duration and AEP), its methods are documented and reproducible, and it performs comparably with BoM 2016 overall and significantly better at sub-daily durations up to 18 hours. At 1–7 days the two cannot be separated at daily gauges, and BoM 2016 is better calibrated at sub-daily gauges; for AEPs rarer than 1% neither product can be tested against the available record. It is not a replacement for BoM 2016 or for the local studies within their extents, and it does not account for the recent change in sub-daily extremes, which affects all products.

1 Introduction

1.1 Purpose

Design rainfall IFDs are the starting point for most design flood estimates in Australia. The BoM 2016 IFDs (Green et al., 2015a; Bureau of Meteorology, 2016) remain the national standard, but they are based on data to 2012, use a national regionalisation scheme, and estimate sub-hourly rainfalls largely by scaling. Local studies in South East Queensland (WMAwater, 2021, "LIMB"; HARC, 2024) have shown that with more local data, and in particular more sub-daily data, IFDs can differ materially from BoM 2016.

This study produces an independent, Australia-wide set of IFD surfaces using BoM data to June 2026. It is intended to be transparent and reproducible: every step is scripted, every modelling choice is tested against withheld data, and the influence of each covariate is mapped so that engineers can see why the surface looks the way it does at their site.

1.2 Scope

Table 1-1 Scope of the IFD surfaces

Item	Value
Region	Mainland Australia and Tasmania (islands > 5 km from the grid excluded)
Durations (25)	5, 10, 15, 20, 25, 30, 45 min; 1, 1.5, 2, 3, 4.5, 6, 9, 12, 18 h; 1, 1.25, 1.5, 2, 3, 4, 5, 6, 7 days
AEPs (13)	63.21%, 50%, 39.35% (2-yr ARI), 20%, 18.13% (5-yr ARI), 10%, 5%, 2%, 1%, 0.5%, 0.2%, 0.1%, 0.05%
Grid	1 km cells, GDA94 Australian Albers (EPSG:3577), one GeoTIFF per duration with one band per AEP
Distribution	GEV, fitted by L-moments (Hosking & Wallis, 1997)
Rainfall type	Unrestricted (sliding-window) burst depths, point rainfall

1.3 Key terms

Table 1-2 Terminology

Term	Meaning in this report
AEP	Annual exceedance probability: the chance a depth is exceeded in any one year. 1% AEP ≈ 1-in-100-year.
Annual maximum series (AMS)	The largest burst depth of a given duration in each water year of record.
Index rainfall (L1)	The mean of a gauge's annual maxima for a duration. Sets the overall size of its IFD.
L-CV, L-skewness	Dimensionless L-moment ratios describing the spread and the tail shape of the annual maxima.
Growth factor	Depth at a given AEP divided by the index rainfall. Set by L-CV and L-skewness.
Covariate	A mapped physical quantity (e.g. mean annual rainfall, terrain) used to predict L-moments where there are no gauges.
Gaussian process (GP)	A statistical form of kriging: a smooth spatial surface fitted to what the covariates did not explain.
Spatial block cross-validation	Withholding all gauges in a 150 km block, predicting them from the rest, and scoring the error.
Hindcast	Building the surfaces from data up to 2012 only and scoring them on 2013–2026 data.
Restricted / unrestricted	Daily gauges report fixed 9 am–9 am totals (restricted); design rainfall refers to the worst 24 h window (unrestricted).

Term	Meaning in this report
Censored year	A year in which an accumulated total of unknown timing could contain a larger burst than the maximum found; not used.

2 Rainfall data

2.1 Sources

Daily-read gauges. BoM daily rainfall records (9 am to 9 am totals), 17,659 stations. After the minimum record span (10 years) and extraction, 14,702 stations had usable annual maxima. The median open period of the gauges used is 1929–2010.

Sub-daily gauges. BoM Water Data Online (WDO) continuous rainfall records (pluviographs, tipping-bucket and automatic weather stations), 3,121 stations, to June 2026. After the minimum record span (10 years) and extraction, 2,276 had usable annual maxima. The median record spans 1990–2026.

No council or third-party data were used. This is a deliberate choice for a national product, but it is the main data difference from the LIMB and HARC studies, which added several hundred council and ALERT gauges in South East Queensland.

2.2 Annual maxima extraction

Annual maxima were extracted with the project's IFD Calculator. Each input value is treated as the rainfall total since the previous timestamp, and each interval between timestamps is classified before any burst depths are computed:

Table 2-1 Classification of rainfall intervals in the extraction

Interval type	Definition	Treatment
Missing	Blank or negative value; WDO quality code 210 (not of release quality) or −1/255 (missing); the interval after a missing row if longer than 1 hour (gauge offline)	Excluded; counts towards the year's missing fraction
Dry	Zero rainfall	No rain
Resolved	Rainfall over an interval no longer than the grid step (1 min sub-daily, 1 day daily), or ≤ 0.5 mm after a long dry gap (first tip of an event)	Placed at the end of the interval
Accumulated	Rainfall over a longer interval: daily readings inside pluviograph records, multi-day accumulations in daily records, coarse logging	Spread evenly (a lower bound for short windows) and remembered

- Sliding-window burst depths were computed on a 1-minute (sub-daily) or 1-day (daily) grid for every target duration.
- Water years start at the middle of each gauge's driest 30-day period, so storms are not split between years.
- For each water year and duration the maximum depth, its end time, the missing fraction of the year and the largest unresolved accumulation are recorded. A year is **censored** for a duration when an accumulation longer than that duration holds more rain than the maximum found, because the true maximum could then be larger and cannot be recovered. A year is **incomplete** when more than 25% of it is missing (Section 3.1).
- Only years that are neither incomplete nor censored are used. Multi-day accumulations in the daily records are kept as single observations rather than spread across the days they cover.
- No restricted-to-unrestricted factors are applied during extraction; they are derived from co-located gauges and applied afterwards (Section 3.3).
- Gauges with less than 10 years between first and last record were not processed (Section 5.1 explains how short records are weighted). Gauges are not screened on how well their at-site fit matches their data, because such a screen preferentially removes gauges with unusual extremes.

2.3 Coverage

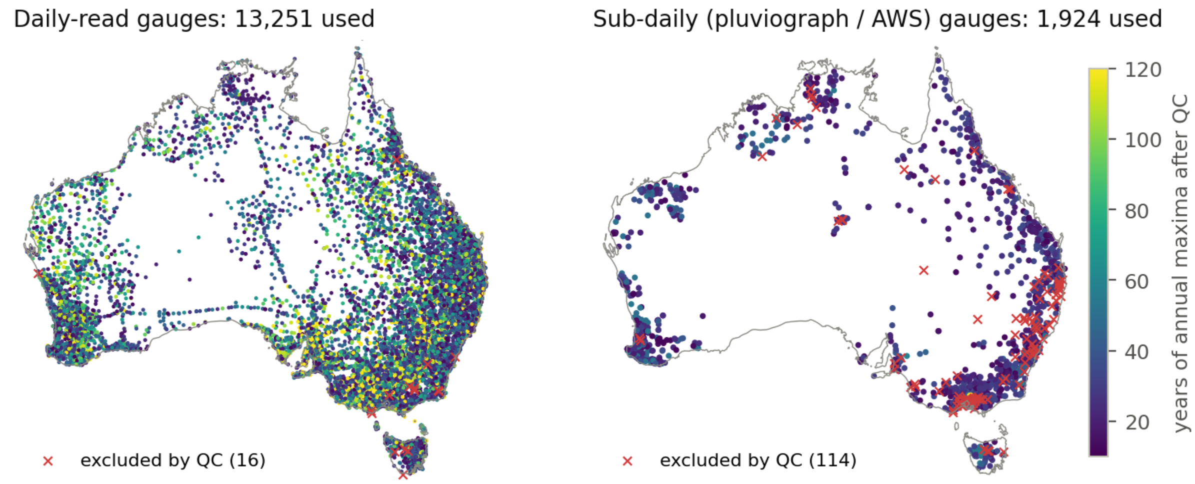

Figure 2-1 Gauges used after quality control, coloured by years of usable annual maxima. Red crosses mark gauges excluded by the quality checks in Section 3.

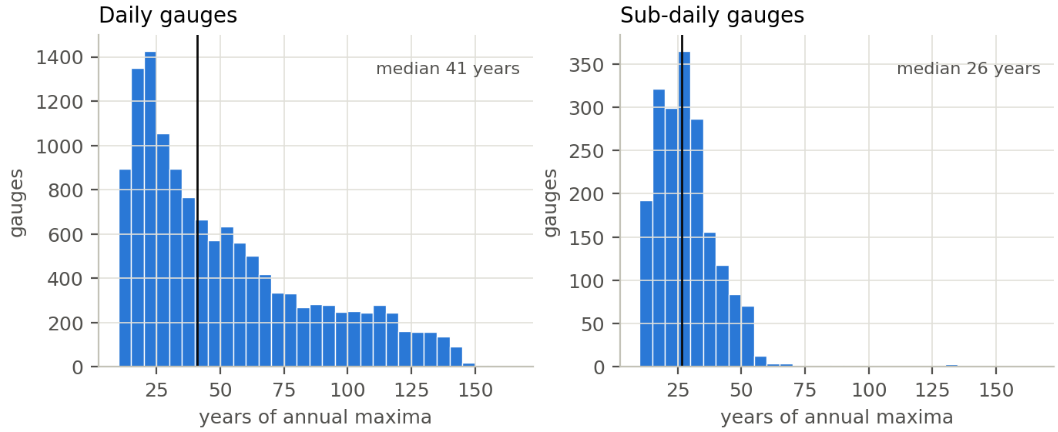

Figure 2-2 Distribution of record length (years of usable annual maxima).

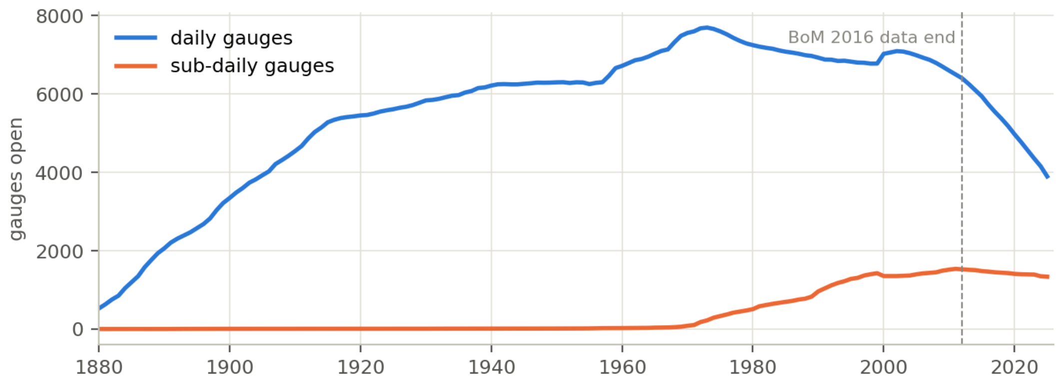

Figure 2-3 Number of gauges open in each year. The sub-daily network is much younger than the daily network: most pluviograph records start after 1980 and continue well beyond the end of the BoM 2016 data (2012).

Table 2-2 Data summary

	Daily gauges	Sub-daily gauges
Gauges with usable annual maxima	14,702	2,276
Used after QC (≥ 12 usable years; ≥ 20 for L-CV and L-skewness)	13,251	1,924

	Daily gauges	**Sub-daily gauges**
Station-years used (longest duration per gauge)	685,480	54,207
Median years of record	41	27
Durations	1–7 days	5 min – 7 days

3 Quality assurance

Rainfall records contain errors that are invisible in a single annual maximum but damaging to an IFD: missing periods, accumulated totals, relocated gauges and transcription errors. The extraction handles missing and accumulated data explicitly (Section 2.2); the checks below deal with what remains.

3.1 Completeness of water years

Across the sub-daily network, the 1-day annual maxima of years with 10–25% of the record missing are on average 0.99 times the median of complete years at the same gauge, i.e. essentially unbiased. Years with 25–50% missing average 0.92 times, and years with more than 50% missing 0.63 times. A maximum missing fraction of **25%** was therefore adopted. Treating missing data as zero rain, or keeping all years, would place many of these low maxima in the record.

As a residual check, the single Grubbs–Beck low-outlier test (IACWD, 1982) was applied in log space to each gauge and duration. At least one year was removed in 8% of daily and 9% of sub-daily gauge-durations; most low values that such a test would otherwise flag are removed earlier by the completeness rule.

3.2 Spatial consistency of daily gauges

Each daily gauge's 1-day index rainfall was compared with the median of its neighbours within 40 km. A gauge was excluded when the difference exceeded 4.5 standard deviations after allowing for its own sampling error (16 gauges).

3.3 Restricted-to-unrestricted factors

Daily gauges report fixed 9 am–9 am totals, which understate the worst 24-hour (or n-day) window. The conversion factors were derived from 548 pairs of daily and sub-daily gauges less than 3 km apart: for each pair and duration, the ratio of the sub-daily (unrestricted) mean annual maximum to the daily (restricted) mean over their common water years (at least 10), and the median over pairs. The resulting factors (Figure 3-1a) are 1.139, 1.053, 1.035, 1.024, 1.022, 1.017 and 1.017 for 1 to 7 days (forced non-increasing). They agree with HARC (2024) for South East Queensland (1.15 and 1.059 at 1 and 2 days) and are smaller than the BoM 2016 factors for 2–6 days (Section 9.2). The spread between pairs is large (interquartile range 1.07–1.21 at 1 day), so a single national factor per duration is an approximation.

Sub-daily gauges whose 1-day index rainfall differed from surrounding daily gauges by more than a factor of 1.45 were excluded (74 gauges); these are usually scaling or unit errors.

3.4 Physical plausibility of sub-daily records

Records with accumulated totals or spikes that survive the extraction produce 5-minute maxima almost as large as 1-hour maxima. A gauge was excluded when depth grew by less than 15% from 5 to 10 minutes, 12% from 15 to 30 minutes, or 10% from 30 to 60 minutes or 60 to 120 minutes (40 gauges, most of them short records admitted by the 10-year span; Figure 3-1b).

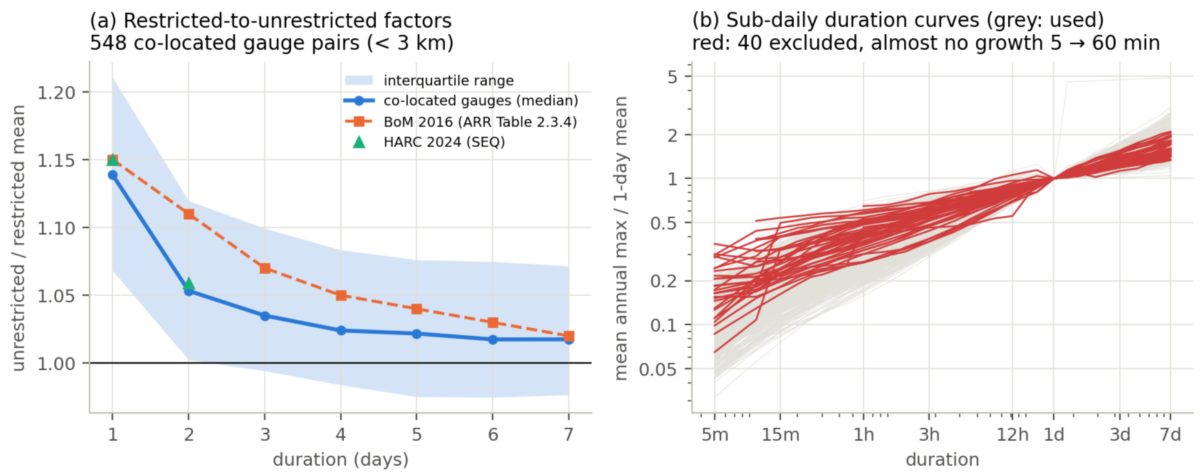

Figure 3-1 (a) Restricted-to-unrestricted factors from co-located gauges, with the BoM 2016 and HARC 2024 values. (b) Sub-daily duration curves relative to the 1-day mean; excluded gauges (red) show almost no growth between 5 and 60 minutes.

3.5 Gauges off the grid

52 gauges more than 5 km from the land grid (Antarctic, Macquarie, Norfolk, Lord Howe and other offshore stations) were excluded. Coastal gauges falling in an ocean cell were assigned the nearest land cell.

Table 3-1 Quality-control summary

Check	Excluded	Effect
Year completeness (> 25% missing)	(years, not gauges)	Removes years whose maxima are biased low
Censoring by unresolved accumulations	(years and durations)	Removes years whose short-duration maxima cannot be recovered
Grubbs–Beck low outliers	(years)	Residual screen
Daily spatial outliers	16 daily	Removes inconsistent gauges
Sub-daily vs daily mismatch (> ×1.45)	74 sub-daily	Removes scaling / unit errors
Sub-daily implausible duration growth	40 sub-daily	Removes accumulated / spike records
Off-grid stations	52 (both)	Island and Antarctic stations

3.6 Instrument era of sub-daily records

Older sub-daily records are digitised Dines pluviograph charts; most recent records come from tipping-bucket gauges and automatic weather stations. Chart digitisation can smooth short bursts, which would bias short-duration maxima low in the early part of a record. The instrument was identified for every station-month from the raw record: tipping buckets record increments that are multiples of 0.2 mm (or 0.5 mm), whereas digitised charts have 0.1 mm resolution.

At the 300–360 gauges with at least 8 usable years in each era, the change in the ratio of short-duration to 1-day annual maxima between the chart and tipping-bucket years was measured. Normalising by the 1-day maxima removes differences in climate and wetness between the two periods. Gauges that were tipping buckets throughout were split into early and late halves to measure how the same ratio drifts over time for reasons other than the instrument (for example climate), and this drift was subtracted.

Chart-era maxima at 10–45 minutes are 3–5% lower relative to their 1-day maxima than tipping-bucket maxima at the same gauges (Figure 3-2). About half of this is matched by the drift at gauges that never changed instrument (about 1% per decade), and the remaining step of 2–3% is not significant at any duration. **No instrument adjustment was therefore applied**, and instrument change cannot explain the 2–3 times excess of recent sub-daily extremes found in the hindcast (Section 8.5). The drift itself is consistent with a gradual increase of short-duration intensities relative to daily totals.

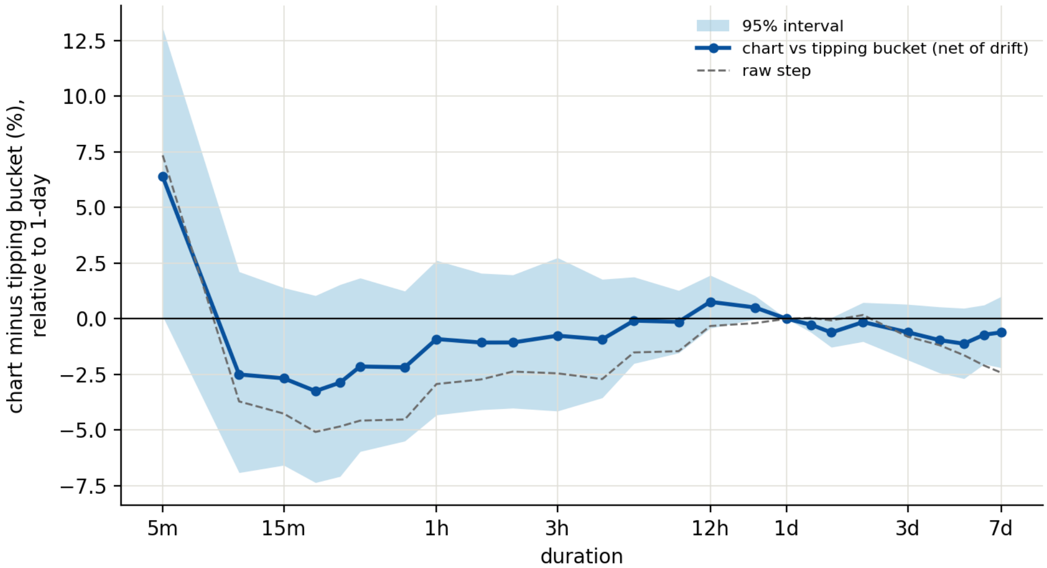

Figure 3-2 Difference in mean annual maximum (relative to 1-day) between chart-era and tipping-bucket years at the same gauges, before (dashed) and after (solid) subtracting the drift seen at gauges that never changed instrument, with 95% intervals.

4 Covariates

Covariates carry information from mapped physical quantities to places without gauges. All covariates were built for this study on the 1 km output grid from a 500 m DEM, the BoM mean annual rainfall grid and the national coastline. No directional weights are assumed; their strength is estimated from the data (Section 4.3).

4.1 Covariate definitions

Table 4-1 Covariates

Covariate	Definition	Why it matters
Mean annual rainfall (MAR)	BoM gridded MAR, log-transformed; raw and 25 km smoothed	Wetter climates have larger storms; captures much of the orographic signal already
Elevation	DEM averaged to 1 km, smoothed at 3, 10 and 30 km; local relief = 3 km minus 30 km	Large-scale uplift and local exposure
Distance to coast	Straight-line distance to the coastline (km), log-transformed	Moisture supply and coastal storm types
Latitude and longitude	Smooth spline terms	Large-scale storm regime (tropical, subtropical, temperate) not otherwise mapped
Windward slope (8 directions)	Rise in terrain (smoothed) over the last 5 km and 20 km upwind, for flow from N, NE, E, SE, S, SW, W, NW	Orographic lift where moist air meets rising ground
Upwind barrier (8 directions)	Highest smoothed terrain within 150 km upwind, above the site	Rain shadow behind ranges
Upwind ocean fetch (8 directions)	Distance to the ocean along the upwind direction, averaged over a fan of 7 rays (±30°)	How far moisture has to travel over land
Ocean-path barrier (8 directions)	Highest terrain crossed between the site and the ocean, fan-averaged	Whether a range stands between the site and its moisture source
Daily orographic signal (sub-daily models only)	The combined directional-terrain effect learned by the daily model, as one smooth layer	Lets the sparse pluviograph network borrow the daily network's terrain response

Directional descriptors were computed by marching rays across the grid. Single rays produced streaks where they grazed bays and inlets, so the fetch and ocean-path barrier layers use a fan of seven rays and are smoothed (Figure 4-1).

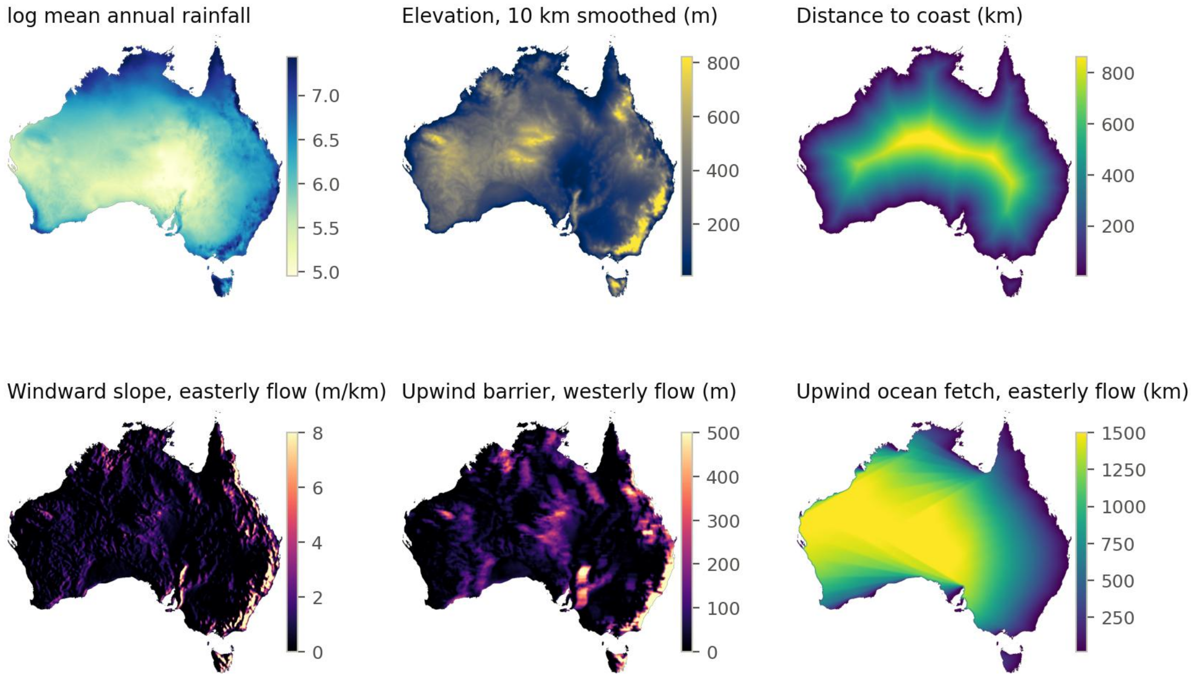

Figure 4-1 Selected covariates on the 1 km grid.

4.2 Correlation with the rainfall statistics

Figure 4-2 shows rank correlations between the covariates and the 1-day L-moments at the daily gauges. The main points for engineers:

- **Mean annual rainfall is the strongest single predictor of the index rainfall** ($\rho = 0.73$), as found by HARC (2024). On its own it leaves a large residual pattern, however: the relationship differs between tropical and temperate Australia (Figure 4-4a).
- **Latitude and longitude carry the storm-regime signal** ($\rho = 0.50$ and 0.60 with the index rainfall; latitude is signed, so positive means larger to the north).
- **Many covariates are strongly inter-correlated** (e.g. MAR with easterly fetch, $\rho = -0.77$). The model therefore uses regularised (ridge) regression, and the influence of individual covariates must be read as a group rather than one at a time.
- **L-CV decreases in wetter climates** ($\rho = -0.28$ with MAR): extremes in dry regions are proportionally larger relative to a typical year. **L-skewness has only weak covariate relationships** and is mainly controlled by pooling many gauges.

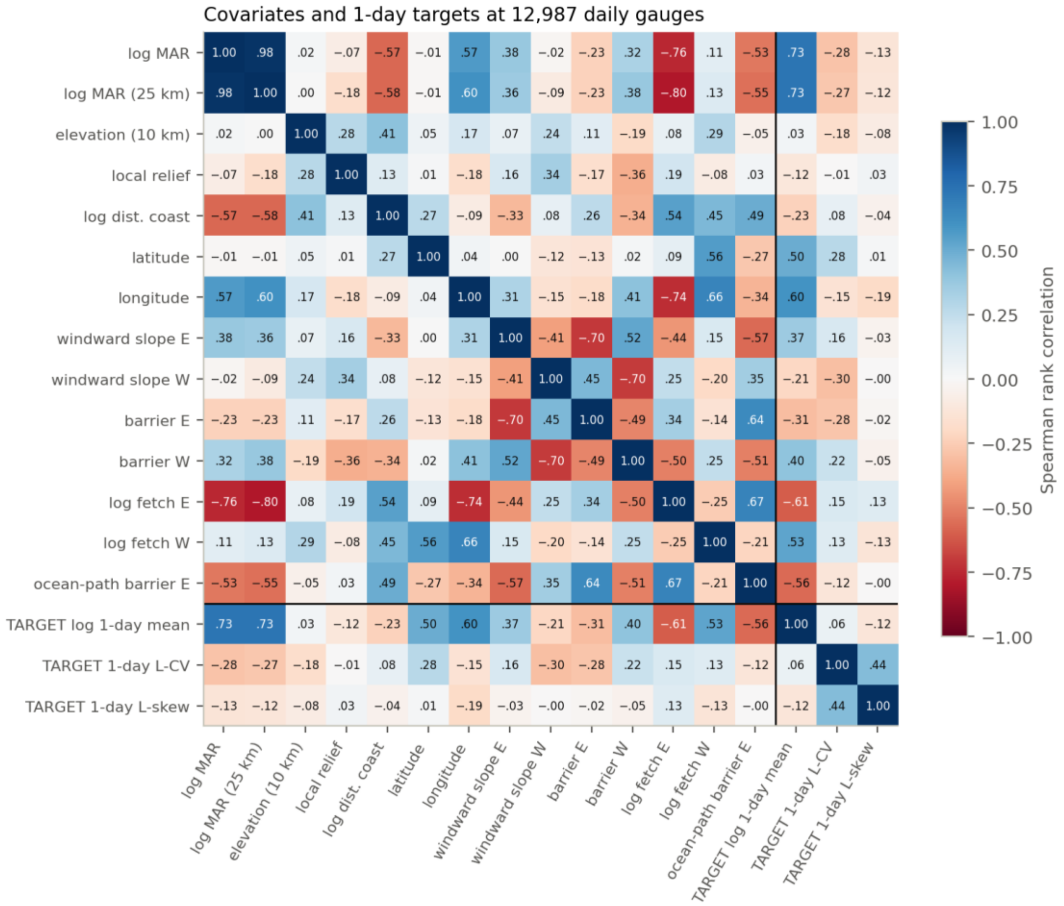

Figure 4-2 Spearman rank correlation between covariates (upper block) and the 1-day L-moments (lower rows) at daily gauges.

For sub-daily durations the relevant quantity is the ratio of the short-duration mean to the 1-day mean. Figure 4-3 shows that this ratio is lower in wet and orographic places (long-duration stratiform rain dominates their 1-day totals) and higher inland and to the north (convective bursts).

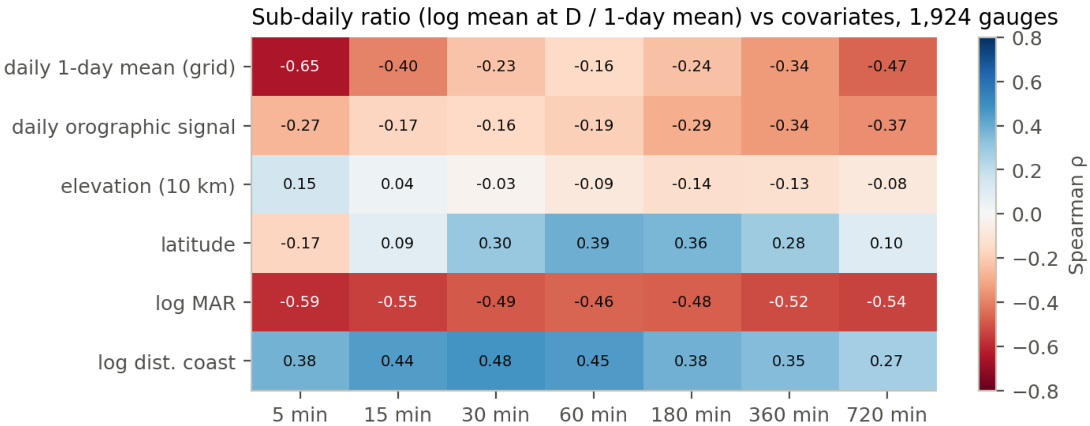

Figure 4-3 Rank correlation between the sub-daily ratio (log of the mean at duration D minus log of the 1-day mean) and covariates, by duration.

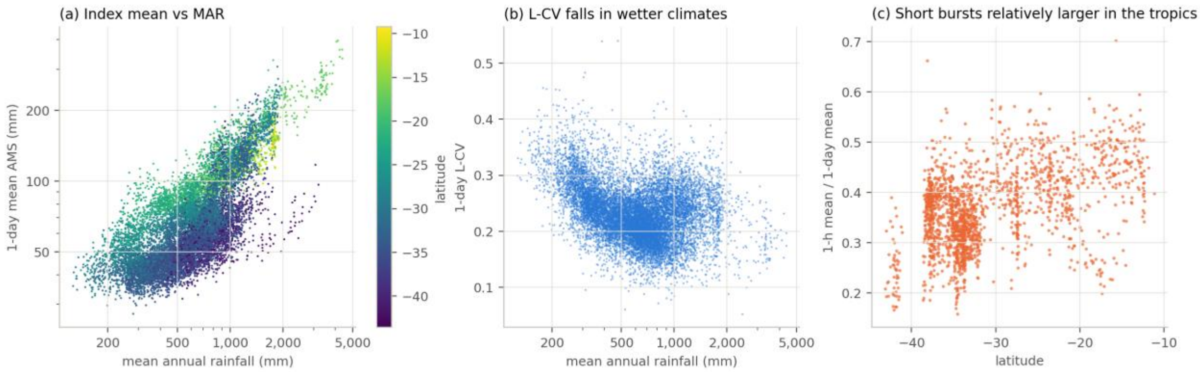

Figure 4-4 Key relationships: (a) index rainfall against MAR, coloured by latitude, showing separate tropical and temperate branches; (b) L-CV against MAR; (c) 1-hour to 1-day ratio against latitude.

4.3 Direction of the orographic effect

Rather than assuming storm directions, the residual of the 1-day index rainfall after the non-directional covariates was regressed, region by region and with non-negative weights and bootstrap averaging, on the eight windward-slope and barrier layers. The resulting weights (Figure 4-5) show that **extreme rainfall is enhanced on slopes facing easterly to north-easterly flow across eastern Australia** (East Coast Lows and tropical systems), whereas westerly-facing weights dominate only in south-west Western Australia. Mean annual rainfall is dominated by frequent westerly frontal rain, which is why directional weights tuned to average conditions are unsuitable for extremes.

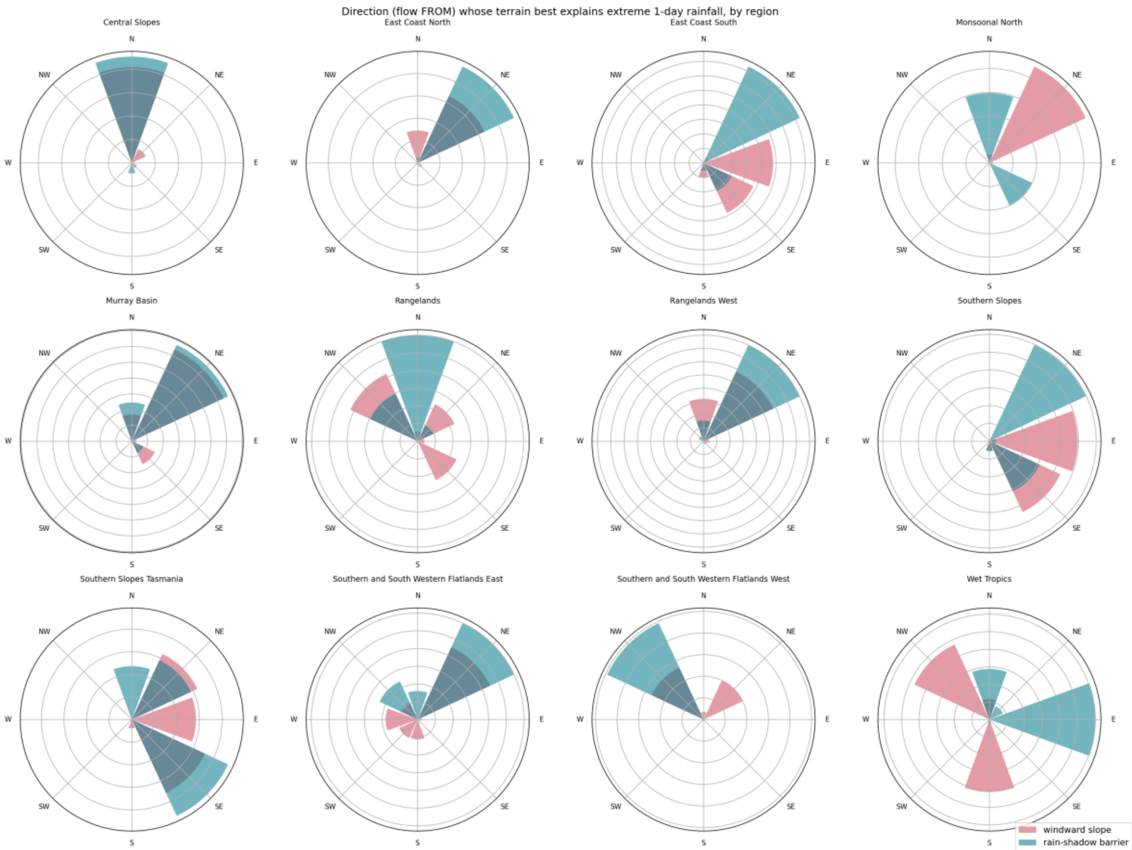

Figure 4-5 Data-derived directional weights per temporal-pattern region: windward-slope (red) and rain-shadow barrier (blue) weights for each direction of flow.

5 Statistical method

The method has four layers: at-site statistics, spatial models for each statistic, assembly of the full IFD at each grid cell, and consistency checks. Figure 5-1 summarises how the layers combine for the 1% AEP 24-hour depth.

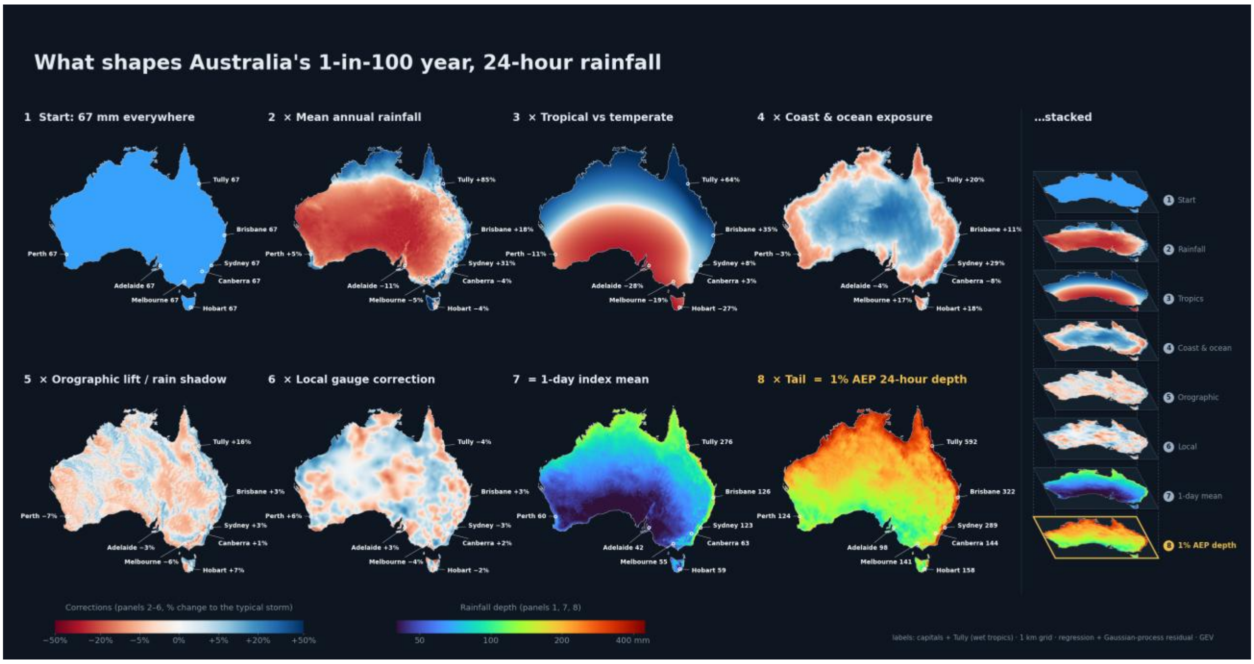

Figure 5-1 How the surface is built for the 1% AEP 24-hour depth: a typical storm (67 mm), multiplied by each covariate group's correction, gives the 1-day index rainfall, which the growth factor lifts to the design depth.

5.1 At-site statistics

For each gauge and duration the unbiased sample L-moments were computed from the screened annual maxima $x_1 \le \ldots \le x_n$:

$$b_0 = \mathrm{mean}(x),\ \ b_1 = \Sigma\,(i-1)/(n-1)\,x_i\,/\,n,\ \ b_2 = \Sigma\,(i-1)(i-2)/((n-1)(n-2))\,x_i\,/\,n$$

$$\lambda_1 = b_0,\ \ \lambda_2 = 2b_1 - b_0,\ \ \lambda_3 = 6b_2 - 6b_1 + b_0,\ \ \text{L-CV} = \lambda_2/\lambda_1,\ \ \text{L-skew} = \lambda_3/\lambda_2$$

The GEV depth for AEP p is the index rainfall times a growth factor that depends only on L-CV and L-skewness (Hosking & Wallis, 1997):

$$Q(p) = \lambda_1 \cdot g(\text{L-CV}, \text{L-skew}, p),\ \ g = \xi + (\alpha/k)[1 - (-\ln(1-p))^k]$$

where k is obtained from L-skewness by Hosking's approximation and α, ξ are scaled by L-CV. The **sampling variance** of each statistic was estimated by bootstrap resampling (300 resamples) of the gauge's own annual maxima. This is how record length enters the method: a gauge with 15 years has roughly twice the sampling variance of one with 30 years and receives correspondingly less weight. L-CV and L-skewness were then smoothed across duration at each gauge (penalised spline, weighted by sampling variance), because they are known to vary smoothly with duration and are noisy at any single duration.

Short records. Gauges need at least 12 usable years to inform the index rainfall and at least 20 to inform L-CV and L-skewness. The index rainfall is estimated reasonably well from a short record, while L-CV and especially L-skewness are too noisy from 12–19 years to add information. Including records of 12–30 years roughly doubles the sub-daily network compared with a 30-year minimum (for example from about 50 to about 100 BoM pluviographs in South East Queensland).

A short record samples only part of the climate: a gauge open from 2005 to 2026 includes the wet years of 2010–2012 and 2022 but none of the dry decades before. Its mean annual maximum was therefore adjusted by the ratio, at up to ten long-record daily gauges (40 or more usable years, within 75 km), of their mean annual maximum over the short record's years to their mean over their full record. Neighbours had to cover at least 70% of those years. The factor was shrunk towards 1 when few neighbours were available, and its uncertainty was added to the gauge's sampling variance. It was applied to all series with fewer than 40 usable years (median factor 0.995; 90% of factors between 0.92 and 1.07); sub-daily durations shorter than a day use the 1-day neighbours.

5.2 At-site IFDs from the IFD Calculator

The IFD Calculator also fits an at-site IFD to each gauge (the *_IFD.csv files). **The regional surfaces do not use these fits**; they are provided for engineers who need an at-site curve and are used only in the "vs calculator" columns of the QA tables. The default at-site fit is:

- GEV by L-moments (Hosking & Wallis, 1997) for each duration with at least 10 usable years;
- L-CV and L-skewness smoothed across log-duration with a record-length-weighted local-linear kernel (narrow for L-CV, wide for L-skewness), which reduces sampling noise without pretending to be spatial pooling;
- depth forced to increase with duration at each AEP (weighted isotonic regression) and with rarity at each duration.

An alternative composite fitter (a Gumbel or TCEV body blended into a GEV or log-Pearson III tail) is available as "-fit legacy". Both fitters also support a non-stationary option in which location and scale vary with a covariate (e.g. global mean temperature or year); it is not used for the regional surfaces. At-site IFDs from 30–60 years of data remain uncertain at rare AEPs (typically ±20–40% at 1% AEP); regional estimates are preferable where available.

5.3 Spatial model for each statistic

Each statistic (log index rainfall, L-CV, L-skewness) at each duration is modelled as

$$\mathrm{y(s) = f(covariates\ at\ s) + r(s) + \varepsilon(s)}$$

- **f** is a smooth trend: cubic spline terms for MAR, smoothed elevation, distance to coast, latitude and longitude, plus linear terms for the directional terrain covariates, estimated with a ridge penalty. Covariate values are clipped to the range seen at gauges, so the trend never extrapolates wildly.
- **r** is a Gaussian process (kriging) residual with a Matérn covariance. It corrects the trend where gauges consistently disagree with it and decays to zero away from gauges, so ungauged areas revert to the covariate trend rather than to a flat average.
- **ε** is noise with variance equal to a small nugget **plus each gauge's own bootstrap sampling variance**. This is the formal record-length weighting: short records are trusted less, and the GP will not bend the surface to fit a single short, noisy gauge.

GP parameters (range, sill, nugget) were estimated by maximum likelihood. For the 1-day index rainfall the fitted correlation range is about 75 km, the spatially structured residual has a standard deviation of about 6%, and the unstructured nugget about 4%, compared with a typical sampling error of 8% per gauge.

5.4 Combining daily and sub-daily information

The daily network has more than ten times the gauges and twice the record length of the sub-daily network. It is therefore used to set the 1-day to 7-day surfaces directly. For sub-daily durations the pluviograph network is used only for **ratios**:

$$\ln \lambda_1(\mathrm{D}) = \ln \lambda_1\mathrm{,daily(1\ day)} + \Delta(\mathrm{D}), \quad \Delta(\mathrm{D}) = \ln \lambda_1\mathrm{,sub(D)} - \ln \lambda_1\mathrm{,sub(1\ day)}$$

and similarly for L-CV and L-skewness. Δ(D) is mapped with its own covariate model, so short durations can respond differently to covariates from long durations (e.g. latitude and inland location raise short-duration ratios while orography lowers them). Because each gauge's ratio is taken relative to its own 1-day value, gauge-level biases in the pluviograph records cancel. The 30- and 36-hour durations, which daily data cannot resolve, are interpolated between the 1-day and 2-day daily surfaces using the shape of the sub-daily ratio curve.

The sub-daily ratio model uses the **daily orographic signal** and the gridded 1-day index rainfall as covariates, instead of the 40 raw directional layers. Both structures score equally well at gauges, but the compact form extrapolates far more safely into the interior, where there are few pluviographs (Section 6.1).

5.5 Assembly, smoothing and consistency

1. At each grid cell, the 25-duration curves of log index rainfall, L-CV and L-skewness are smoothed across log-duration (fixed penalised-spline smoothers), and L-CV and L-skewness are kept within physically reasonable limits.
2. GEV depths are computed for all 13 AEPs.
3. Depth must increase with duration, and intensity must not increase with duration (log-depth slope between 0 and 1 against log-duration). For rare AEPs the depth must also grow with duration at no less than 20% of the rate of the 50% AEP curve at that cell. This rule only binds at the 0.1% and 0.05% AEPs; at 99% of cells it changes the 0.05% AEP depth by less than 3%.
4. Depth must increase with rarity at every duration.
5. Constraints are enforced by isotonic (pool-adjacent-violators) regression, which makes the smallest change that satisfies them.

The trend is evaluated on the 1 km grid; the GP residual is evaluated on a 5 km grid and interpolated bilinearly, which is accurate because the residual varies over tens of kilometres.

6 Model selection and covariate influence

6.1 Choosing the model structure

Candidate structures were compared by spatial block cross-validation: gauges were grouped into 150 km squares, the squares were split into eight folds, and each fold was predicted from the other seven. The score is the out-of-sample error **beyond** the irreducible sampling error of each gauge, so that a model is not penalised for failing to reproduce noise.

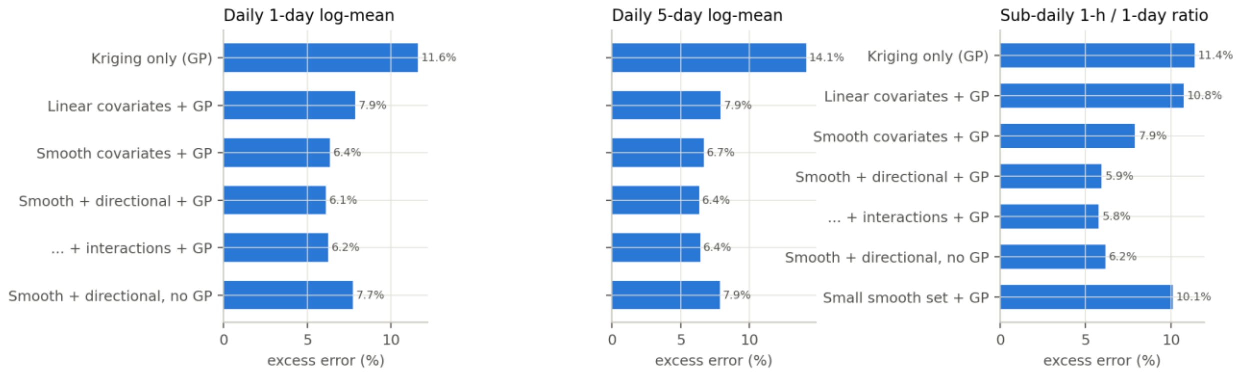

Figure 6-1 Out-of-sample error beyond sampling noise for candidate model structures. Kriging alone is clearly worst; smooth covariates with a GP residual are best.

Findings: (1) covariates roughly halve the error of kriging alone; (2) smooth (spline) responses beat linear ones; (3) the directional terrain terms give a further, consistent improvement for daily durations; (4) the GP residual is always worth keeping. Because the same cross-validation was used both to choose the structure and to report its skill, the reported skill is slightly optimistic. The structure comparison used gauges with at least 30-year record spans; the chosen structures were then validated on the full network by full-chain cross-validation and the hindcast. For L-CV and L-skewness, covariates add little beyond a smooth regional surface, so heavily pooled models are used. With 300 km blocks (mimicking gauge-poor regions) the compact sub-daily structure performs as well as the full directional one.

6.2 Strength of each covariate

Because the trend is additive in log space, each covariate group's contribution is exact and can be mapped as a percentage change relative to an average gauge. Effect maps for the 1-day index rainfall are shown in Figure 6-4 and are supplied as GeoTIFFs.

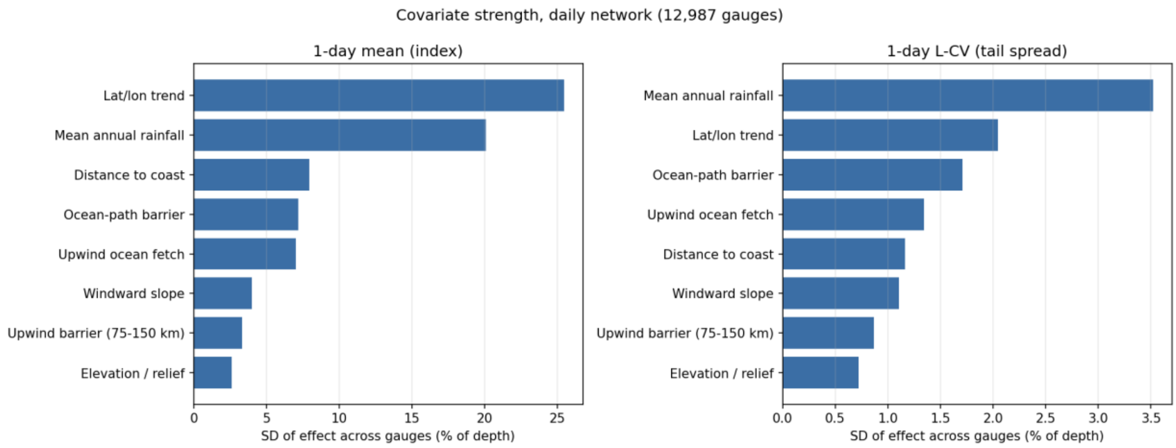

Figure 6-2 Covariate strength for the daily network: the spread (standard deviation across gauges) of each group's effect on the 1-day index rainfall and L-CV.

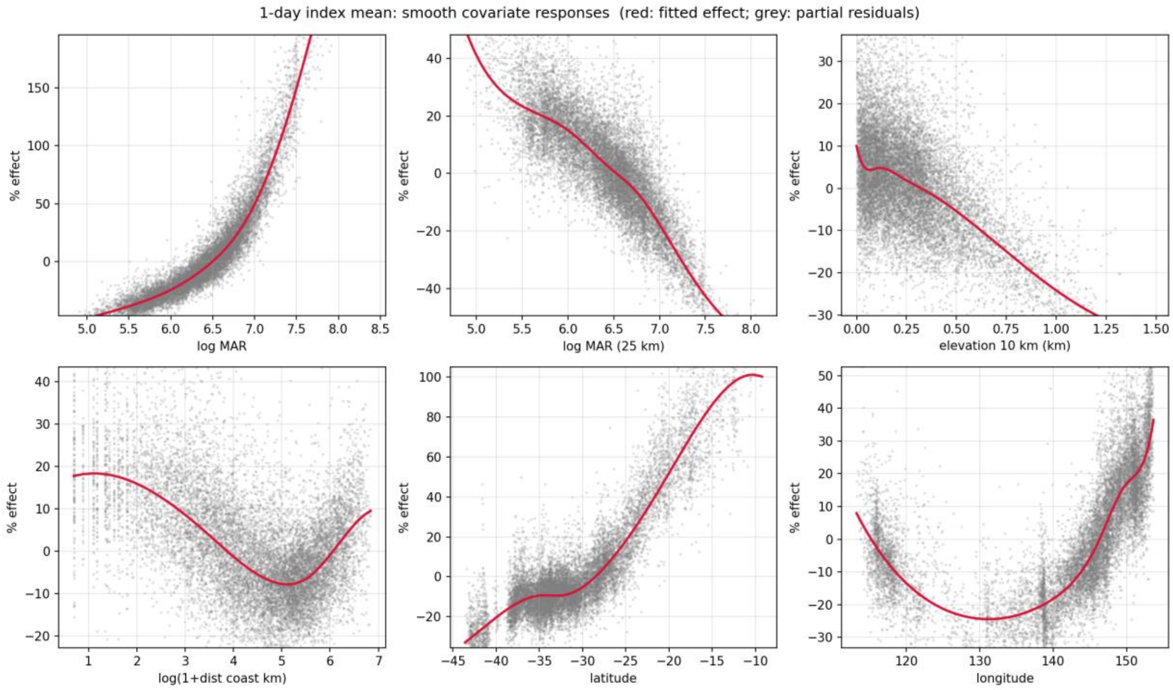

Figure 6-3 Fitted smooth responses of the 1-day index rainfall (red) with partial residuals (grey). Local and 25 km MAR act together: sites wetter than their surroundings gain, regional wetness gains less than proportionally.

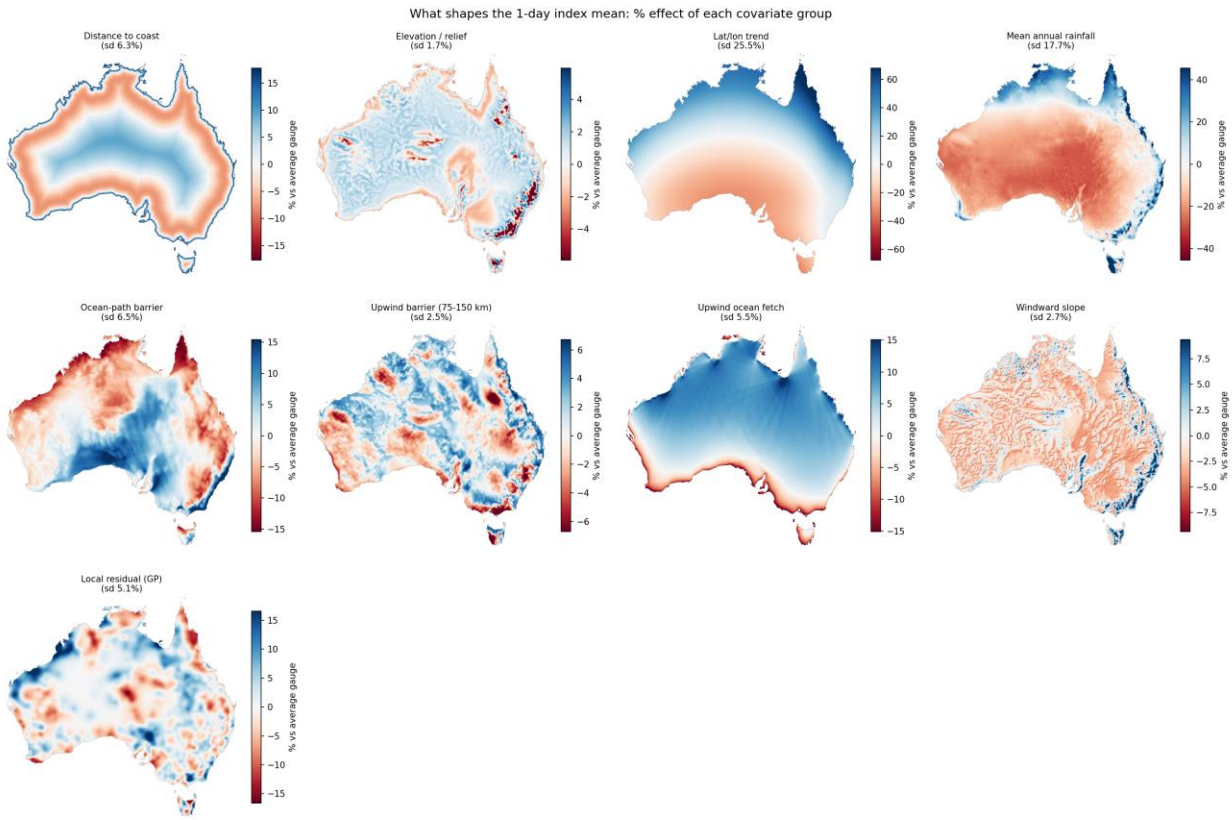

Figure 6-4 Effect of each covariate group on the 1-day index rainfall, as % relative to an average gauge. The local residual (GP) is small (standard deviation about 5%), showing that the covariates explain most of the pattern.

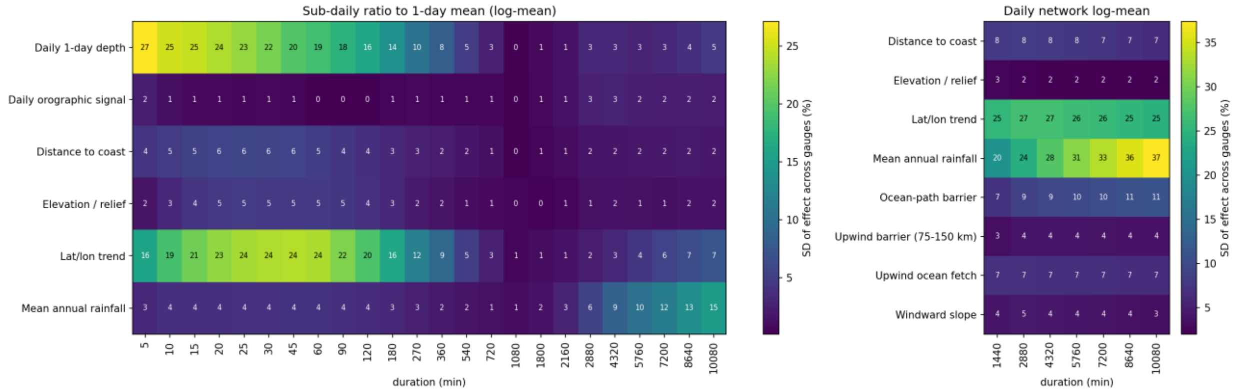

Figure 6-5 How covariate strength changes with duration. Short-duration ratios are driven by storm regime (latitude) and by the 1-day depth itself; mean annual rainfall regains influence at 5–7 days.

Interpretation for engineers. The terrain panel appears modest nationally (standard deviation about 3%) because most of the orographic signal is already inside mean annual rainfall. Locally it is not modest: windward escarpments gain 10–20% and lee slopes lose 5–10% beyond what MAR explains (e.g. +16% at Tully).

7 Results

7.1 Parameter surfaces

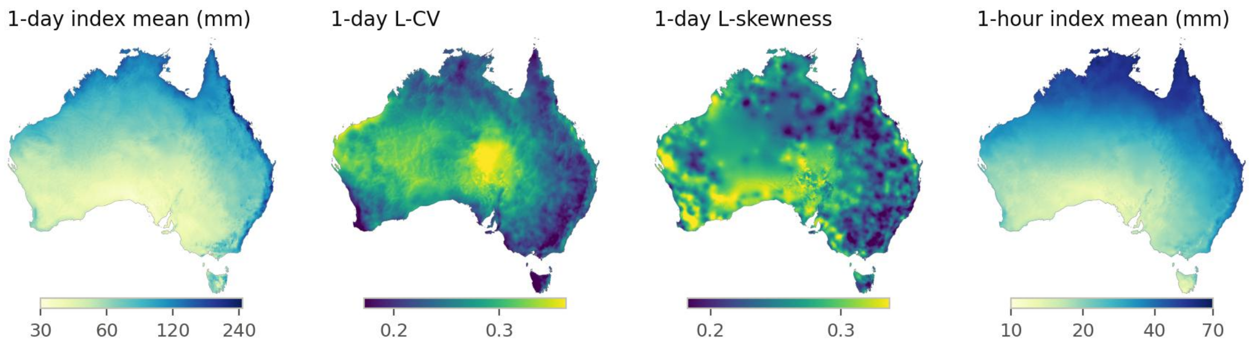

Figure 7-1 Regionalised parameters: 1-day index rainfall, 1-day L-CV and L-skewness, and 1-hour index rainfall.

7.2 Design rainfall surfaces

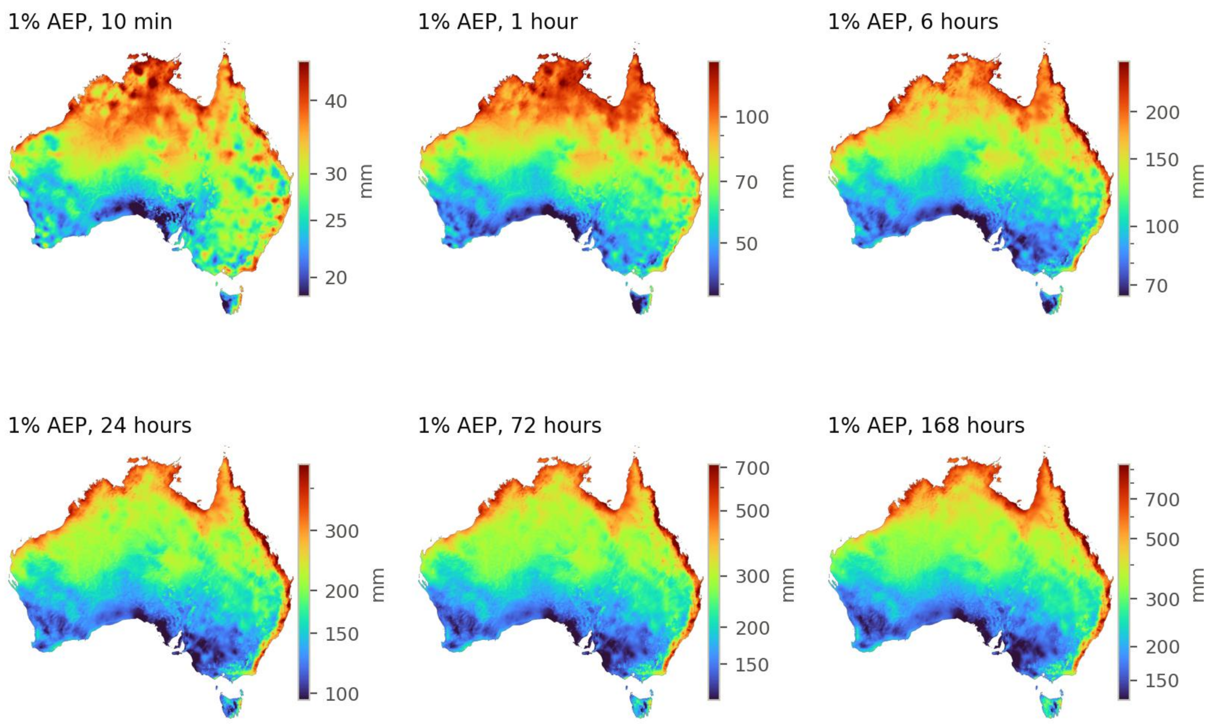

Figure 7-2 1% AEP design rainfall depths for six durations.

7.3 Grid consistency checks

Table 7-1 Consistency of the final grids

Check	Result
Cells checked (every 4th cell in each direction)	483,571
Depth decreasing with rarity (any duration)	0 cells
Depth decreasing with duration (any AEP)	0 cells
Intensity increasing with duration	0 cells
Log-depth vs log-duration slope, min / max	0.020 / 1.000
Neighbouring-cell change, 1% AEP 1 h: median / 99th / 99.99th percentile	0.13% / 1.5% / 3.2%
Cells needing a duration adjustment (all at rare AEPs)	0.7% of cell-AEP curves

The surfaces are smooth (no seams or artefacts; the largest cell-to-cell change is 3%), increase with duration and rarity everywhere, and never produce intensities that increase with duration.

8 Validation

8.1 Approach and reference

Two tests were used. **Spatial block cross-validation** withheld all gauges in a 150 km block, refitted every surface without them, and predicted full IFDs at the withheld gauges. All cross-validation numbers quoted for this study are these out-of-sample predictions, not the final grid. The **hindcast** (Section 8.5) rebuilt the surfaces from data up to 2012 only and scored them, alongside BoM 2016, on 2013–2026 data.

Cross-validated predictions were compared with an at-site GEV (L-moments) fitted to each gauge's usable annual maxima. **This reference is built from this study's own extraction and conventions** (the same annual maxima, conversion factors and data period to 2026). BoM 2016, LIMB and HARC used different data, periods, factors and extraction rules, so part of any difference between them and this reference reflects those choices rather than accuracy. The cross-validation therefore favours this study; the hindcast does not.

At-site GEV estimates are themselves uncertain (a 40-year record constrains the 1% AEP depth to roughly ±20%), so differences at single gauges are expected. Two measures are reported: the **median bias** across gauges and the **median absolute difference** (typical error) at a gauge.

8.2 Comparison with at-site estimates

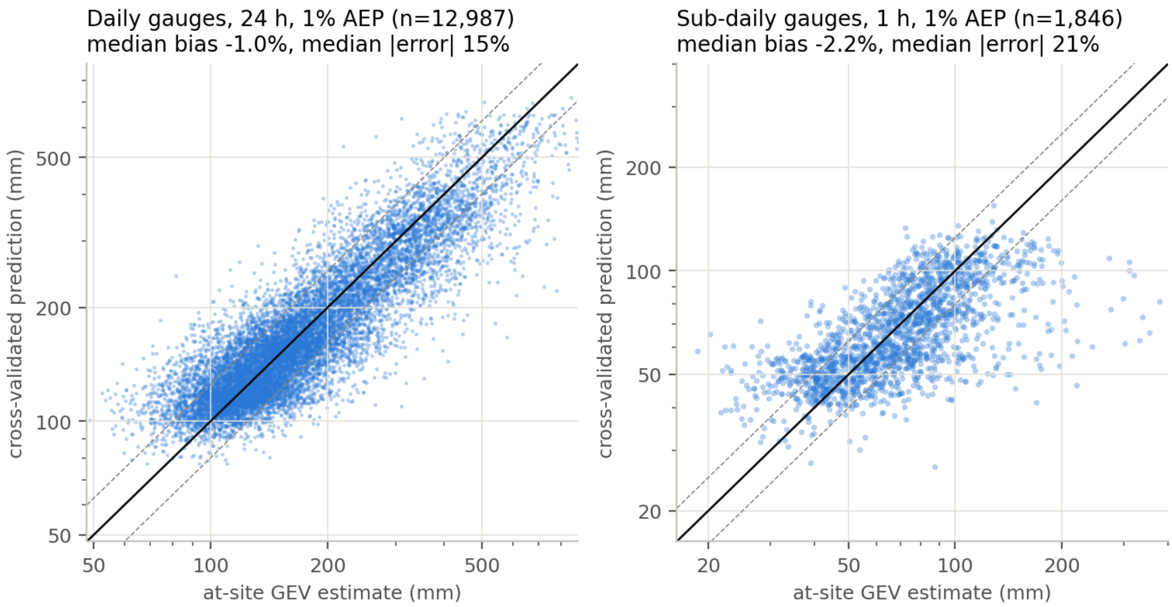

Figure 8-1 Cross-validated predictions of this study against at-site GEV estimates at 1% AEP. Dashed lines are ±25%.

Table 8-1 Median bias and median absolute difference against at-site GEV estimates (gauges with at least 20 usable years; all durations pooled)

AEP	Daily: this study (CV) bias / typical error	Daily: BoM 2016 bias / typical error	Sub-daily: this study (CV) bias / typical error	Sub-daily: BoM 2016 bias / typical error
50%	+0.3% / 5.7%	+2.3% / 4.9%	+1.3% / 7.9%	−2.1% / 7.7%
10%	−0.3% / 7.0%	+2.7% / 5.7%	−0.1% / 10.3%	−4.0% / 9.5%
1%	−1.6% / 13.2%	+1.9% / 11.0%	−3.8% / 20.1%	−8.3% / 18.4%
0.05%	−4.0% / 26.7%	−0.4% / 24.5%	−10.3% / 40.6%	−13.7% / 39.9%

Interpretation. At daily durations BoM 2016 has the smaller typical error at every AEP (but many of these gauges were used to build BoM 2016, whereas this study's values are for withheld gauges), and a small positive bias (2–3%) that comes mostly from 2–4 days, where its conversion factors are larger than those derived here. At sub-daily durations the typical errors are almost identical, and BoM 2016 is biased low, increasingly so for rarer AEPs. At rare AEPs a small negative bias is expected for any smooth regional estimate, because at-site GEV quantiles are inflated by sampling noise in L-skewness.

8.3 Exceedance rates

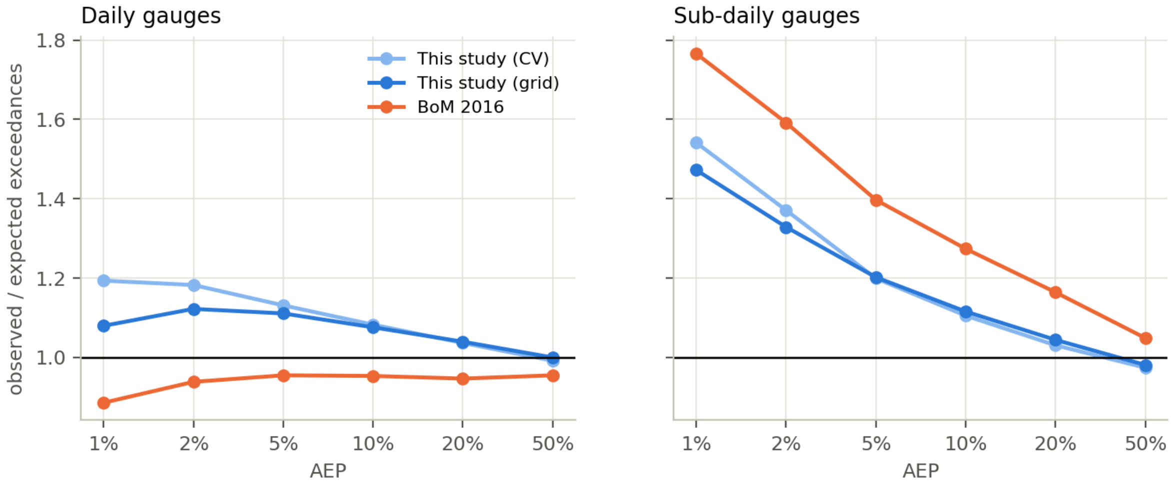

Figure 8-2 Observed divided by expected number of annual maxima exceeding the design depth (all years). 1.0 is the expected value; above 1 means the depth is exceeded more often than its AEP implies. “Grid” is the final surface, which is fitted to these gauges.

Over all years of record, this study's cross-validated 1% AEP depths are exceeded 1.19 times as often as expected at daily gauges and 1.54 times at sub-daily gauges; BoM 2016's 0.89 and 1.77 times. Random error in any prediction inflates these ratios, because the chance of exceeding a rare threshold rises steeply as the threshold falls. As a diagnostic of this study's growth curves: when each gauge's own index rainfall is combined with the regionalised growth curve, the 1% AEP ratio is 1.02 (daily) and 1.17 (sub-daily), so most of this study's excess comes from the index rainfall at withheld sites rather than from the tail shape.

8.4 Bias by duration and region

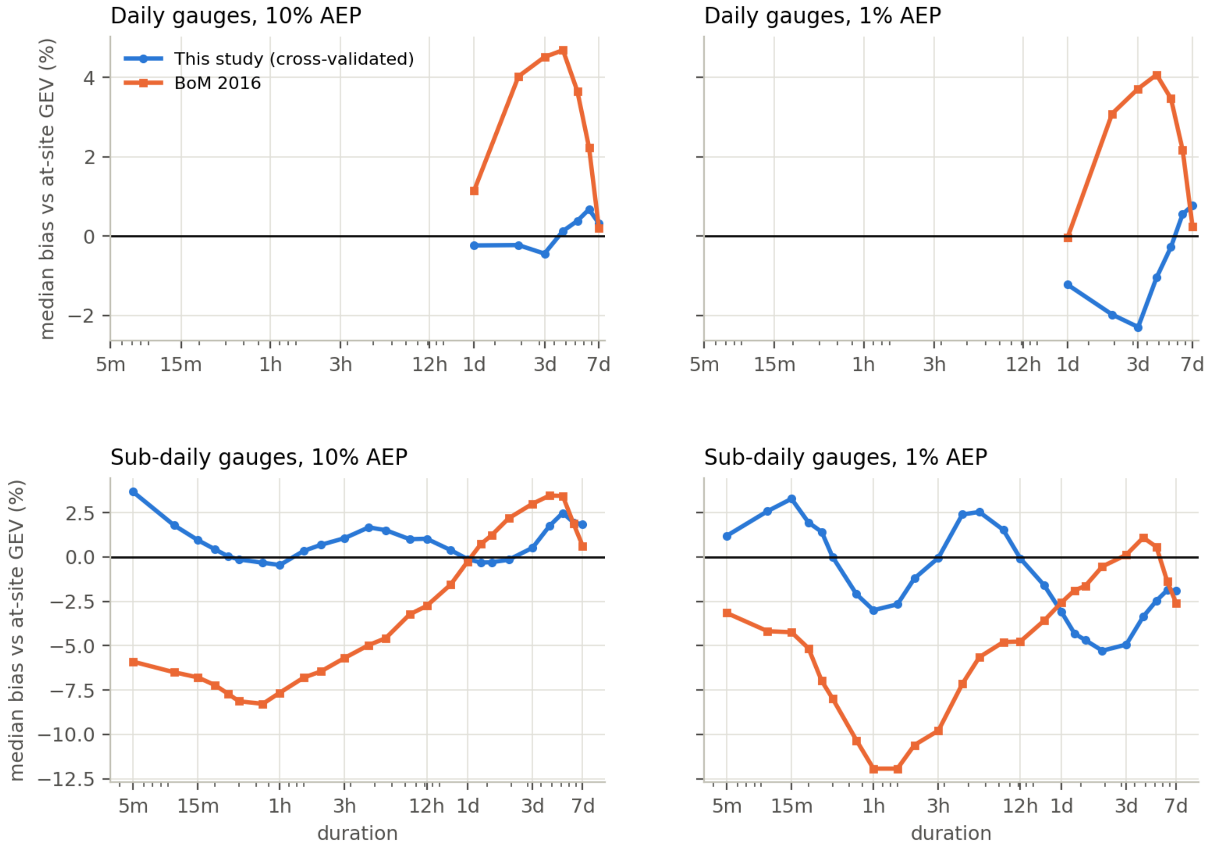

Figure 8-3 Median bias by duration for this study (cross-validated) and BoM 2016, at daily gauges (top) and sub-daily gauges (bottom).

At daily gauges, BoM 2016's bias at 1% AEP is small at 1 day, rises to a few percent at 2–4 days and falls again towards 7 days, tracking the ratio between its conversion factors and those derived here. At sub-daily gauges BoM 2016 is lowest at 30 minutes to 3 hours (about −10% to −15% at 1% AEP). This study is within a few percent at sub-daily durations up to 18 hours, and a few percent below the sub-daily records at 1% AEP at 1–7 days.

Table 8-2 Median bias / median absolute difference (%) at 1% AEP by temporal-pattern region, all durations pooled. SSW = Southern and South Western.

Region	Daily: BoM 2016	Daily: this study (CV)	Sub-daily: BoM 2016	Sub-daily: this study (CV)
Central Slopes	+3.3 / 10.3	−1.9 / 12.1	−0.6 / 20.9	−5.4 / 20.3
East Coast North	−0.4 / 11.9	−3.9 / 14.1	−6.8 / 19.7	−14.7 / 23.7
East Coast South	+0.1 / 11.5	+1.3 / 15.1	−5.4 / 19.1	+3.5 / 22.0
Monsoonal North	+1.1 / 12.4	−2.4 / 15.3	−5.7 / 19.1	−2.0 / 19.6
Murray Basin	+2.7 / 10.0	−0.5 / 11.8	−6.7 / 17.3	−2.0 / 19.0
Rangelands	+0.1 / 11.1	−2.3 / 13.9	−1.5 / 18.8	−3.4 / 21.0
Rangelands West	+1.9 / 10.1	−2.5 / 13.4	−1.1 / 13.2	−7.5 / 16.6
Southern Slopes	+3.3 / 11.0	−2.2 / 12.7	−12.0 / 22.9	−3.0 / 20.0
Southern Slopes Tasmania	+4.3 / 14.3	+8.0 / 17.6	+0.9 / 18.4	+8.2 / 22.2
SSW Flatlands East	+4.2 / 9.8	+3.1 / 12.7	−1.6 / 24.7	+1.4 / 23.3
SSW Flatlands West	+4.3 / 10.0	−0.4 / 11.7	−1.0 / 14.2	+5.1 / 17.6
Wet Tropics	+6.1 / 14.4	+0.4 / 15.4	+8.1 / 23.9	+8.5 / 20.7

BoM 2016 has the smaller typical error at daily gauges in every region. At sub-daily gauges this study has the smaller typical error in the Central Slopes, Southern Slopes, SSW Flatlands East and Wet Tropics, and BoM 2016 in the others, although with large biases in several (e.g. −12% in the Southern Slopes). Neither product is consistently better at sub-daily gauges on this measure; the hindcast (Section 8.5) is the more reliable comparison.

8.5 Hindcast: out-of-sample test on 2013–2026

BoM 2016 was built from data to 2012. To compare the two products on equal terms, the whole pipeline was re-run using only water years up to 2012, and both the resulting surfaces and BoM 2016 were scored against the annual maxima of water years 2013–2026 at the same gauges. Neither product used these data. Because a single storm year affects many gauges together, 95% intervals were obtained by resampling the 14 test years with replacement (2,000 resamples), keeping all gauges of a year together; differences against BoM 2016 were resampled in pairs. Two scores are reported: the **exceedance ratio** (observed over expected number of annual maxima above each AEP depth; 1.0 is ideal) and a **quantile score** (the average pinball loss over the 50%–1% AEP depths, a proper scoring rule for quantile predictions; lower is better), expressed relative to BoM 2016.

Table 8-3 Hindcast scores on water years 2013–2026, with 95% intervals from resampling whole years. Negative quantile-score differences mean this study's hindcast is better than BoM 2016.

Duration group	Gauges	Quantile score vs BoM 2016 (95% interval)	1% AEP exceed. ratio: BoM 2016	1% AEP exceed. ratio: this study (to 2012)	10% AEP exceed. ratio: BoM 2016	10% AEP exceed. ratio: this study (to 2012)
Daily 1–7 d	5,574	−1.1% (−2.4 to +0.2)	1.07 (0.59–1.65)	1.36 (0.79–2.03)	0.92	1.08
Sub-daily ≤ 1 h	1,520	−5.1% (−7.7 to −2.4)	3.20 (2.61–3.91)	2.26 (1.92–2.63)	1.72	1.23
Sub-daily 1.5–18 h	1,521	−2.7% (−4.1 to −1.1)	2.78 (2.16–3.46)	2.06 (1.63–2.52)	1.63	1.40
Sub-daily 1–7 d	1,571	+0.2% (−1.1 to +1.5)	2.00 (1.46–2.69)	2.36 (1.74–3.16)	1.18	1.36

Daily durations. The quantile scores are within about 1% and the difference is not significant. BoM 2016's 1% AEP depths were exceeded 1.07 times as often as expected and this study's 1.36 times, but 14 years of data cannot separate them: the intervals overlap almost entirely, and a paired comparison of how far each ratio is from 1 gives an interval from −0.33 to +0.28 (in log units), spanning zero. The median growth curves of the two products (1% AEP depth divided by 50% AEP depth) are almost identical at 1 day, so the difference is one of level rather than tail shape; at 1% AEP each 1% of depth changes the exceedance count by about 7%. At frequent AEPs BoM 2016's depths were exceeded less often than expected (0.92 at 10% and 50% AEP), i.e. they are slightly high in the body of the distribution.

Sub-daily durations up to 18 hours. This study's hindcast scores 3–5% better than BoM 2016, and both the score difference and the better calibration at 10% and 1% AEP are significant. This is consistent with the cross-validation result that BoM 2016 is low at these durations. **At sub-daily 1–7 days** the quantile scores are equal and BoM 2016 is significantly better calibrated.

Both products under-predict recent sub-daily extremes. Over 2013–2026, pluviograph annual maxima exceeded both products' 1% AEP depths 2–3.4 times as often as expected. Even this study's surfaces fitted to all data (which include these years) are exceeded 1.9–2.2 times as often at sub-daily gauges, against 1.06 at daily gauges. Part of this is the random-error inflation described in Section 8.3, but the size and consistency of the effect indicate that recent sub-daily records are heavier-tailed than the long-term record the surfaces represent (Section 9.5).

9 Comparison with BoM 2016, LIMB 2021 and HARC 2024

9.1 Methods

All four products fit the GEV distribution to annual maxima using L-moments and regionalise the L-moments. They differ in data, in how the index rainfall is mapped, in how L-CV and L-skewness are pooled, and in how sub-daily durations are handled. Table 9-1 extends HARC's (2024) comparison table to include this study.

Table 9-1 Comparison of methods (BoM 2016 and LIMB conversion factors as tabulated by HARC, 2024)

Step	BoM 2016 (ARR 2019)	LIMB 2021 (WMAwater)	HARC 2024 (QRA)	This study
Area	Australia	Brisbane, Ipswich, Lockyer Valley, Moreton Bay	South East Queensland (11 LGAs + buffer)	Australia
Data	BoM, to 2012	BoM + council + ALERT, to 2019	BoM + councils + Seqwater, to 2022–23; 810 sub-daily + 567 daily in SEQ	BoM only, to June 2026; 13,251 daily + 1924 sub-daily
Index rainfall	ANUSPLIN splines with elevation (24 h); polynomial across durations	Kriging with elevation and elevation variability	Kriging with MAR; adjusted to 2010 climate	Smooth covariate model (MAR, terrain, directional orography, coast, location) + GP; record-length-weighted
Sub-daily	BGLSR regression from daily data; scaling below 1 h	Local sub-daily data; 48 h taken as point of truth	Local sub-daily data	Sub-daily ratios to the daily 1-day value, with their own covariate model
L-CV, L-skew	Region of influence, 500 station-years	Region of influence, 100 station-years (distance in lat, lon, elevation)	Regional average ≤ 2 h; region of influence (500 yrs) ≥ 48 h; interpolated between	Smooth regional surface + GP with heavy nugget (equivalent to pooling), weighted by sampling variance
Conversion factors (1–7 d)	1.15, 1.11, 1.07, 1.05, 1.04, 1.03, 1.02	1.099, 1.020, 1.008 … 1.000	1.15, 1.059 (1–2 d)	1.14, 1.05, 1.04, 1.02, 1.02, 1.02, 1.02 (co-located gauges)
Consistency	Polynomial smoothing across durations	Polynomial smoothing; L-moment interpolation 6–48 h	L-moment smoothing; minor manual fixes	Penalised smoothing across durations + constrained isotonic enforcement
Validation	At-site comparison	At-site comparison	At-site comparison	Spatial block cross-validation, hindcast (2013–2026), at-site comparison

9.2 Comparison with BoM 2016 across Australia

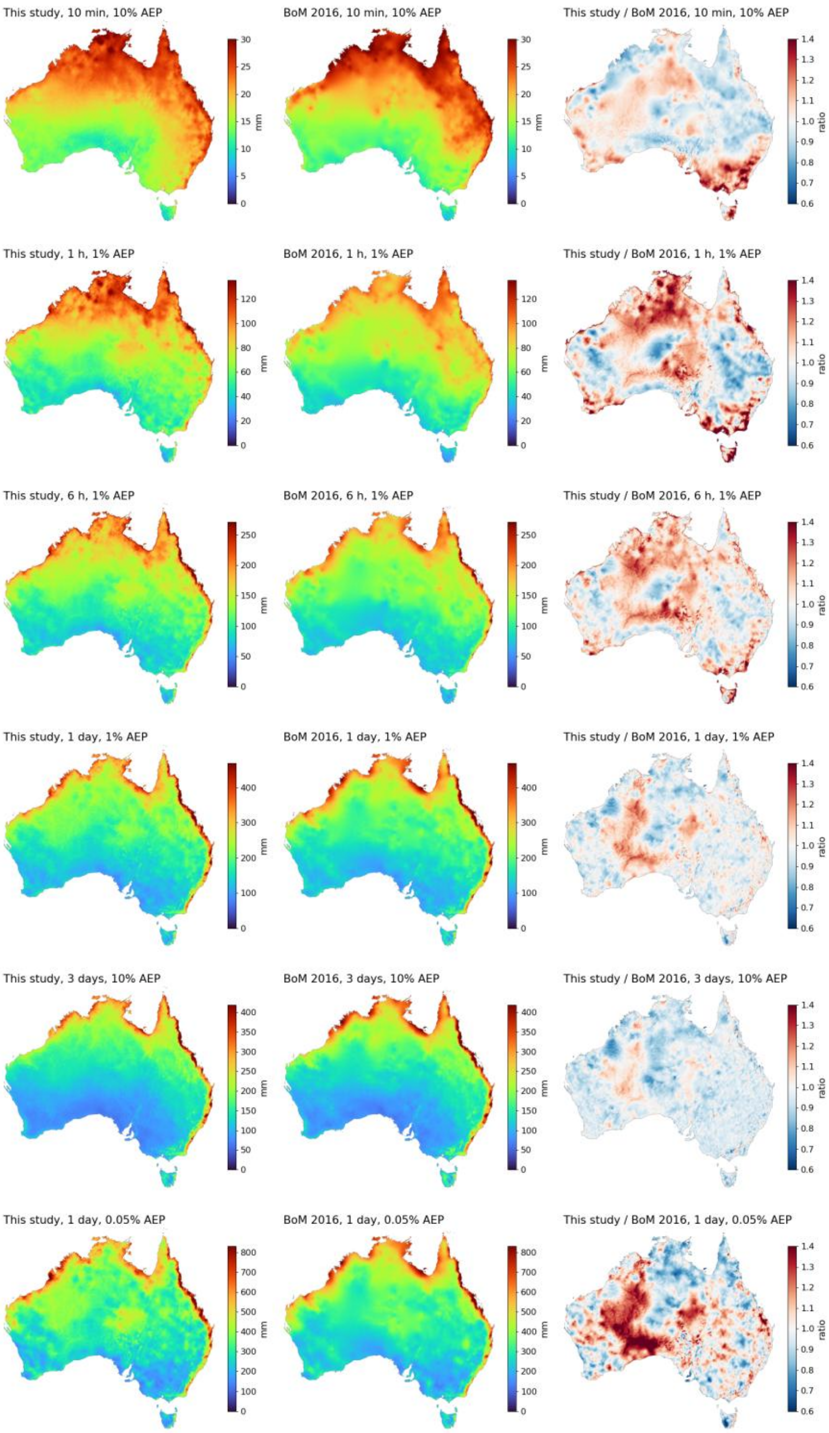

Figure 9-1 This study (left), BoM 2016 (centre) and their ratio (right) for selected durations and AEPs. Red: this study higher; blue: lower.

At 1 day the two products agree within about ±10% over most of the continent. At 2–4 days this study is generally a few percent lower, which is about the size of the difference in conversion factors (BoM 2016 uses 1.11, 1.07 and 1.05 at 2–4 days; the co-located gauges give 1.05, 1.04 and 1.02). At sub-daily durations this study is higher in the south-east, Tasmania and the tropical north (often +10% to +30% at 1 hour, 1% AEP) and lower over much of South East Queensland. In the arid interior, where there are few pluviographs, this study's sub-daily estimates are up to 40% above BoM 2016 and should be treated with caution.

9.3 South East Queensland: HARC 2024 and LIMB 2021

HARC (2024) and LIMB (2021) cover South East Queensland only. The products were sampled at the BoM gauges used in this study within the HARC extent (913 daily and 91 sub-daily gauges) and compared with the same at-site estimates. HARC and LIMB used many of these gauges plus several hundred council gauges, and the values for this study are from its final grid; all products are therefore effectively in-sample, and the comparison measures agreement with BoM records, not skill at ungauged sites.

Table 9-2 Agreement with at-site estimates at BoM gauges in South East Queensland (median bias, median absolute difference, and exceedance ratio at 1% AEP)

Product	Daily 10% bias	Daily 1% bias	Daily 1% |error|	Sub-daily 10% bias	Sub-daily 1% bias	Sub-daily 1% |error|	Sub-daily 1% exceed. ratio
This study	−0.5%	−1.7%	12%	−6.4%	−18.6%	25%	3.34
BoM 2016	+2.4%	−1.1%	12%	−4.7%	−11.5%	22%	2.16
HARC 2024	+0.7%	+2.3%	12%	−7.5%	−12.7%	19%	2.30

Table 9-3 As Table 9-2, for the LIMB extent (Brisbane, Ipswich, Lockyer Valley, Moreton Bay; 224 daily and 40 sub-daily gauges)

Product	Daily 10% bias	Daily 1% bias	Daily 1% |error|	Sub-daily 10% bias	Sub-daily 1% bias	Sub-daily 1% |error|	Sub-daily 1% exceed. ratio
This study	−0.9%	−2.7%	13%	−8.3%	−19.4%	25%	3.52
BoM 2016	+2.3%	−1.9%	14%	−6.1%	−10.7%	20%	2.04
HARC 2024	−1.3%	+0.8%	14%	−10.3%	−14.0%	19%	2.49
LIMB 2021	−3.3%	−6.2%	13%	−6.3%	−11.1%	19%	2.11

Daily durations. All products are within a few percent of the daily records at the BoM daily gauges and have similar typical errors (12–14% at 1% AEP). HARC 2024 sits 1–2% above the records at 1% AEP, which is expected: it deliberately adjusted its estimates to a 2010 climate, and a historical-period reference will sit below them. LIMB 2021 sits about 6% below the records in its extent.

Sub-daily durations. About 90 BoM pluviographs in the region passed quality control, about half of them with fewer than 20 usable years, so these figures are indicative. All products are below the recorded 1% AEP depths, and those depths were exceeded 2–3 times as often as expected, consistent with the heavy-tailed recent records found nationally (Section 8.5) and with the major events of 2011 and 2022. This study is furthest from the records (−18.6%%), because its tail shape comes from the long daily network; admitting short records, which roughly doubled the BoM pluviographs in the region, did not change this. HARC and LIMB, which used several hundred additional local pluviographs and (in HARC's case) a recent-climate adjustment, sit closer. **Within the HARC and LIMB extents, the local studies should be preferred for sub-daily durations.**

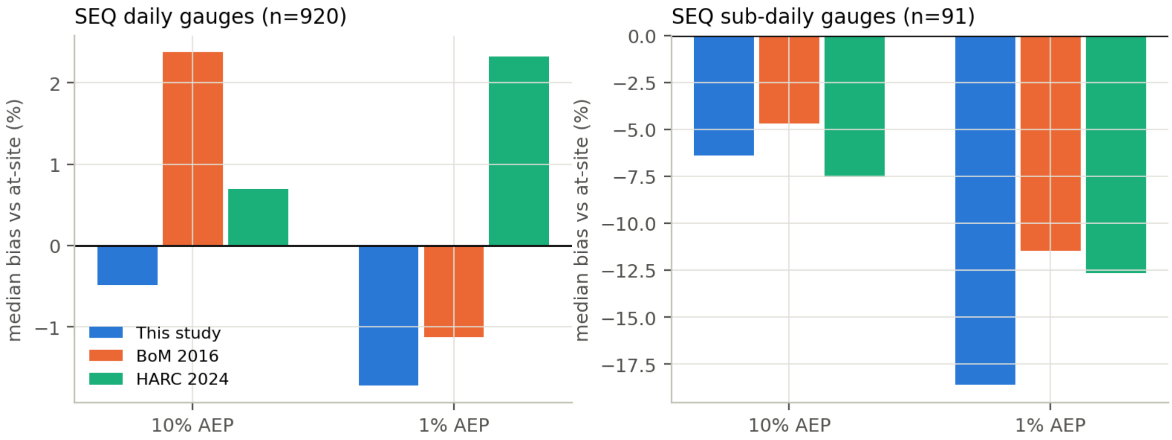

Figure 9-2 Median bias against at-site estimates at BoM gauges in South East Queensland.

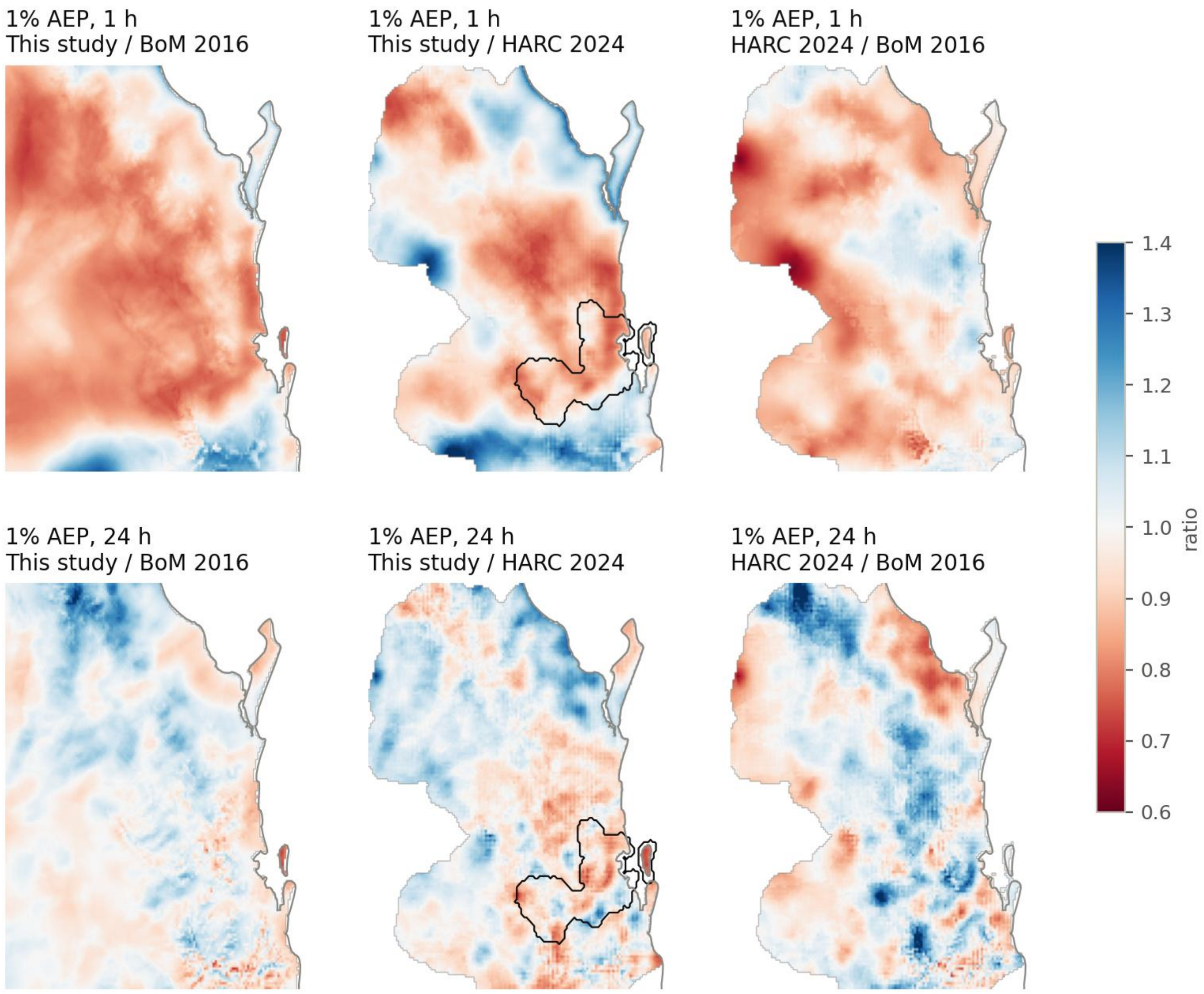

Figure 9-3 Ratios between products for 1% AEP at 1 hour (top) and 24 hours (bottom) across the HARC extent. The black outline in the centre column marks the LIMB extent.

At 1 hour, both this study and HARC 2024 are below BoM 2016 across most of South East Queensland, in line with HARC's finding that BoM 2016's short-duration estimates are high in the region. This study is lower still across central SEQ and higher in the Gold Coast hinterland and the far south. At 24 hours the three products agree within about ±15%, without a consistent pattern.

9.4 Example IFD curves

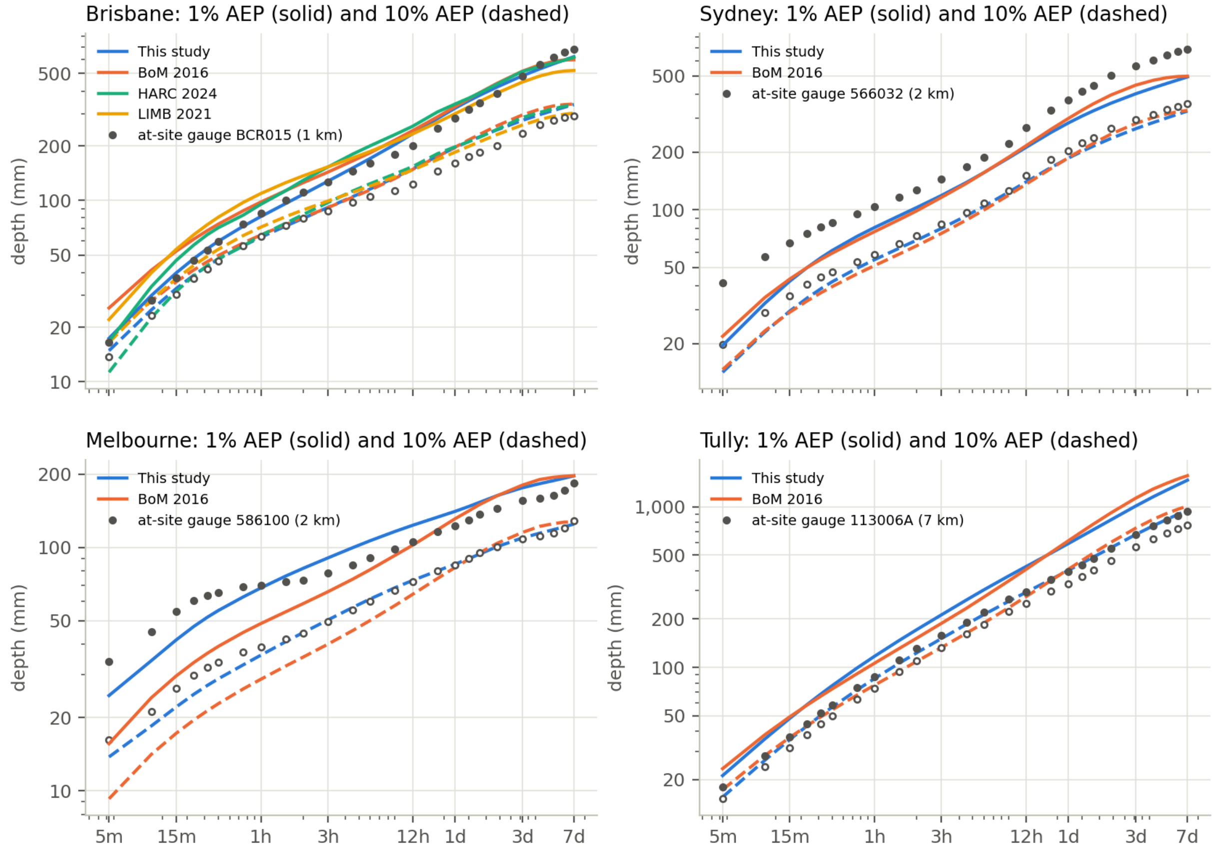

Figure 9-4 IFD curves at four locations for 1% AEP (solid) and 10% AEP (dashed), with at-site estimates from the nearest sub-daily gauge (dots) where one lies within 15 km.

At 10% AEP all products follow the nearby gauges closely. At 1% AEP the Sydney pluviograph sits above both products at all durations and the Melbourne gauge above them at durations shorter than about 12 hours, whereas the Brisbane gauge lies within the spread of the four products and the Tully gauge sits below them. Single-gauge 1% AEP estimates carry ±20% or more sampling uncertainty; these four sites are illustrations, not a test.

9.5 Why sub-daily records sit above the daily-based surfaces

At several hundred co-located gauge pairs, continuous records have systematically higher L-CV and L-skewness than the neighbouring daily gauge for the same duration, and the difference grows with duration (Figure 9-5). At 7 days, restricted and unrestricted windows are almost identical, so the difference cannot come from the 9 am reading time. Because years more than 25% incomplete are excluded, missing data are unlikely to explain it. The hindcast (Section 8.5) shows that sub-daily maxima since 2013 exceed long-term estimates far more often than daily maxima do. The most likely explanation is the **record period** (sub-daily records median 1990–2026; daily records median 1929–2010), i.e. recent intensification of sub-daily extremes, possibly combined with changes in instruments and logging resolution, which the data supplied cannot separate.

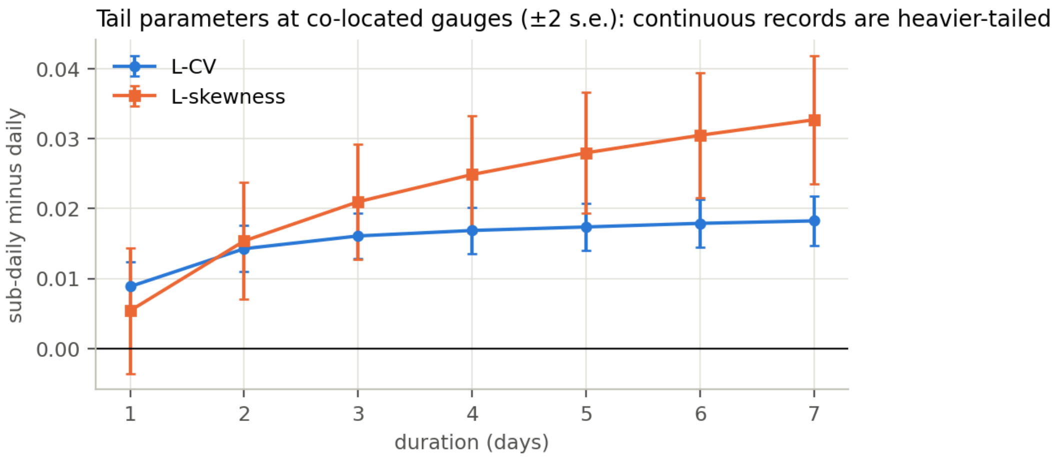

Figure 9-5 Difference in L-CV and L-skewness between sub-daily and co-located daily gauges, by duration (±2 standard errors).

This study keeps the daily network as the anchor for tail shape, because it is ten times larger and twice as long. The surfaces therefore represent the long-term historical climate. Users designing for current or future climate should apply climate-change adjustments as recommended in ARR, noting that this dataset is centred on an earlier period than HARC's 2010-adjusted estimates and that the hindcast indicates recent sub-daily extremes are under-represented by all the products compared.

10 Limitations and uncertainty

- **Non-stationarity.** The surfaces represent the historical record. The hindcast shows that recent sub-daily extremes are exceeded 2–3 times as often as expected, by this study and BoM 2016 alike; no climate adjustment has been applied.
- **Evaluation reference.** Cross-validation statistics use at-site estimates built with this study's own extraction and conventions and favour this study. The hindcast avoids this, but covers only 2013–2026 and BoM gauges.
- **Rare daily tail.** In the hindcast, this study's 1% AEP daily depths were exceeded 1.36 times as often as expected, against 1.07 for BoM 2016; the difference is not significant with 14 test years. AEPs rarer than 1% cannot be tested against the available record for any product.
- **Short records.** Records of 12–30 years inform only the index rainfall and are adjusted for the period they cover using long-record neighbours; the adjustment assumes the period anomaly at a gauge matches that of the daily network nearby.
- **Conversion factors.** Derived from about 500 co-located gauge pairs with a wide spread between pairs; one factor per duration is applied nationally.
- **Sub-daily coverage.** Pluviographs are concentrated in the south-east and on the coast. Sub-daily estimates in the interior and the north-west rely on the covariate model and carry larger uncertainty.
- **Gauge heterogeneity.** Sub-daily gauges differ by type and logging resolution, which adds 15–25% unexplained scatter at short durations.
- **Censoring and completeness rules.** Years are either used or not; partial information in incomplete or censored years is discarded rather than modelled.
- **Model selection.** The comparison of model structures (Section 6.1) used gauges with at least 30-year record spans; the chosen structures were validated on the full network by full-chain cross-validation and the hindcast.
- **Rare AEPs.** 0.5% to 0.05% AEP depths are extrapolations of the GEV fitted to 30–160 years of data and are not constrained by probable maximum precipitation estimates.
- **Point rainfall.** The grids are point estimates; areal reduction factors must be applied for catchments, as per ARR.
- **Covariate interpretation.** Several covariates are inter-correlated. Group effects are robust, but the split between individual covariates is not unique.

11 Recommendations

The surfaces are a research dataset for sensitivity testing and comparison. BoM 2016 remains the national standard. On the evidence of the hindcast: for sub-daily durations up to 18 hours outside South East Queensland these surfaces are closer to recent records than BoM 2016; for 1–7 days BoM 2016 is equally good at daily gauges and better calibrated at sub-daily gauges; and within their extents the local studies (HARC 2024, LIMB 2021) should be preferred for sub-daily durations.

References

Ball, J., Babister, M., Nathan, R., Weeks, W., Weinmann, E., Retallick, M. and Testoni, I. (Eds.) (2019). Australian Rainfall and Runoff: A Guide to Flood Estimation. Commonwealth of Australia (Geoscience Australia).

Bureau of Meteorology (2016). Design Rainfall Data System (2016). Commonwealth of Australia.

Green, J., Johnson, F., Beesley, C. and The, C. (2015a). Using regionalised L-moments to derive design rainfalls for Australia. Proceedings of the 36th Hydrology and Water Resources Symposium, Engineers Australia.

Green, J., Johnson, F., McKay, D., Podger, S., Sugiyanto, M. and Siriwardena, L. (2015b). Improvements in design rainfalls using new methods and data. Bureau of Meteorology.

HARC (2024). Review and update of design rainfall intensity-frequency-duration for eleven local government areas in South East Queensland. Draft report QRA00002 for the Queensland Reconstruction Authority. Hydrology and Risk Consulting.

Hosking, J.R.M. and Wallis, J.R. (1997). Regional Frequency Analysis: An Approach Based on L-Moments. Cambridge University Press.

IACWD (1982). Guidelines for Determining Flood Flow Frequency, Bulletin 17B. Interagency Advisory Committee on Water Data, U.S. Geological Survey.

Jakob, D., Taylor, B. and Xuereb, K. (2009). A pilot study to explore methods for deriving design rainfalls for Australia. Bureau of Meteorology.

Rasmussen, C.E. and Williams, C.K.I. (2006). Gaussian Processes for Machine Learning. MIT Press.

Wasko, C. et al. (2023). Climate change considerations for design rainfall in Australian Rainfall and Runoff. (as cited in HARC, 2024).

WMAwater (2021). Updated Local Design Rainfalls for Brisbane, Ipswich, Lockyer Valley and Moreton Bay (LIMB). Report 119057, final, June 2021.

Appendix A Sub-daily covariate effects

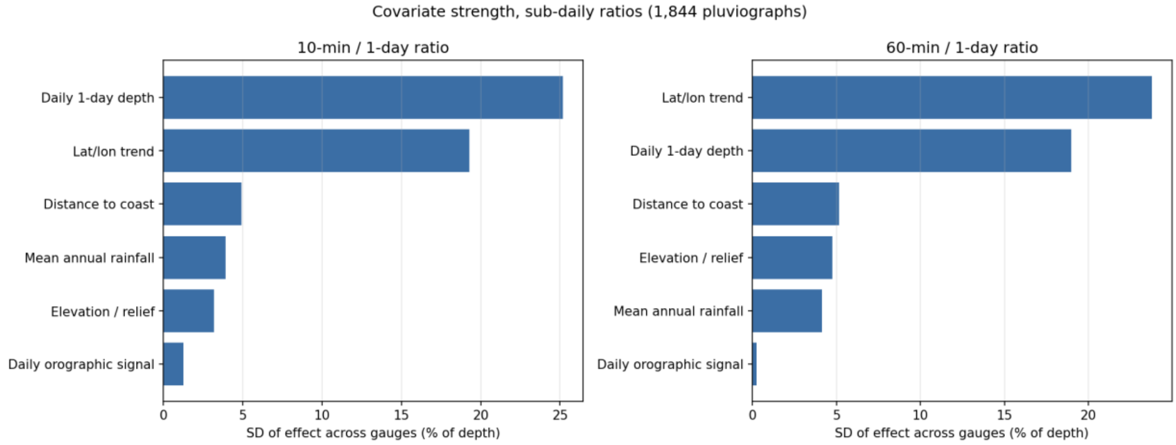

Figure A-1 Covariate strength for the sub-daily ratio models (10-minute and 60-minute).

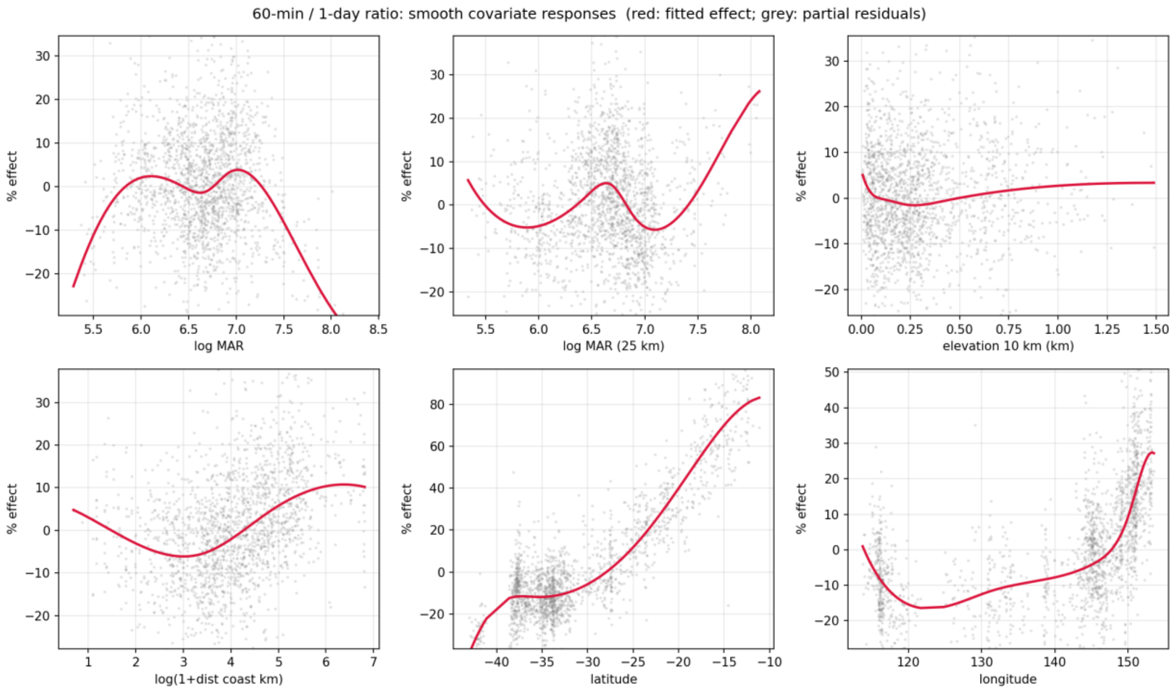

Figure A-2 Smooth covariate responses of the 60-minute to 1-day ratio.

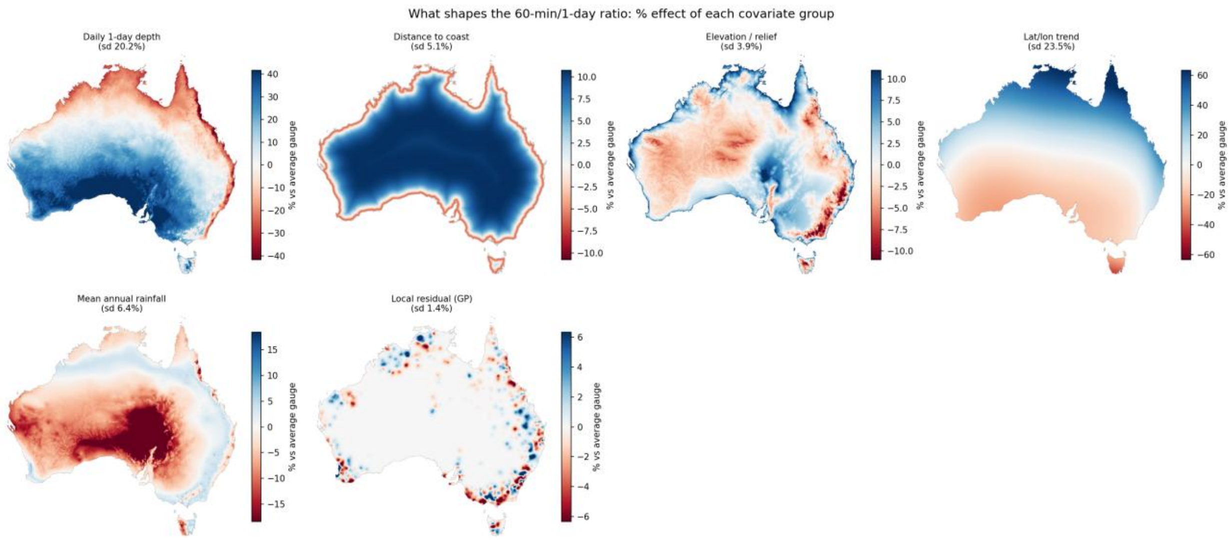

Figure A-3 Effect of each covariate group on the 60-minute to 1-day ratio (% relative to an average gauge). Opposing latitude and 1-day-depth effects are partly collinear and should be read together.