DATA ARTICLE

A database for sea-level monitoring in French Polynesia

Jean-Pierre Barriot[1] | **Fangzhao Zhang[1,2]** | **Bernard Ducarme[3]** | **Guy Wöppelmann[4]** | **Gaël André[5]** | **Alban Gabillon[6]**

[1]Geodesy Observatory of Tahiti, Université de la Polynésie Française, Tahiti, French Polynesia

[2]College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, China

[3]Catholic University of Louvain, Georges Lemaître Centre for Earth and Climate Research (ELI), Louvain-La-Neuve, Belgium

[4]Littoral ENvironnement et Sociétés (LIENSs)-UMR-7266, Université de La Rochelle, La Rochelle, France

[5]Service Hydrographique et Océanographique de la Marine (SHOM), Brest, France

[6]Géopôle of the South Pacific, IT Team, Université de la Polynésie Française, Tahiti, French Polynesia

Correspondence
Jean-Pierre Barriot, Geodesy Observatory of Tahiti, Université de la Polynésie française, Tahiti, French Polynesia.
Email: jean-pierre.barriot@upf.pf

Funding information
Central French Government, Grant/Award Number: OBEMAP; Local Government of French Polynesia

Abstract

This article presents a curated database of the sea-level measurements acquired by the network of the five geodetic tide gauges managed over French Polynesia by the Geodesy Observatory of Tahiti from 13 June 2009 to 28 January 2021. A unique feature of this database, with respect to previous databases that host the same raw data, like the Intergovernmental Oceanographic Commission database (IOC, www.ioc-sealevelmonitoring.org) and the database of 'Réseaux de référence des observations marégraphiques' (REFMAR, http://refmar.shom.fr) is that all the time-tags of the raw measurements (1- or 2-min sampling) have been validated and, if necessary, corrected with a precision of 2 min (time shifts of up to 1 hr can be present in the raw data). Possible outliers have also been flagged, but not removed. In addition, smoothed hourly data are also provided, along with tidal analysis results and estimations of the sea-level trends for the five tide gauges, with respect to their local geodetic datum. The database, entitled 'SEA LEVEL collected from TIDE STATIONS in South Pacific Ocean from 2009-06-13 to 2021-01-28', can be accessed on the NOAA data servers as 'NCEI Accession 0244182' and contains two subsets: The first one is relative to the original sampling rate and the second one is relative to an hourly re-sampling with harmonic tide models for each tide gauge station.

1 | INTRODUCTION: SEA-LEVEL MONITORING IN FRENCH POLYNESIA

This study presents the results of the preprocessing, calibration, database encapsulation and preanalysis of the sea-level monitoring of the five stations installed and managed by the University of French Polynesia (UPF) in the French Polynesia part of the South Pacific area (Figures 1 and 2) and maintained by the hydrographic and oceanographic survey branch of the French Naval Hydrographic and Oceanographic Service (SHOM).

Spread over an area between 134°W–155°W longitude and 7°S–28°S latitude, French Polynesia also known as Tahiti and her islands, covers a vast expense (5,500,000 km^2) and remote (7,000 km from Los Angeles, 4,000 km from Australia) of ocean located in the middle of the South Pacific. This French overseas territory, with a large political autonomy, is made up of a swarm of 121 islands or islets, including high volcanic islands (35) and low coral islands or atolls (86), which together represent a surface area of 3,668 km^2 of emerged land (half of Corsica) and 12,800 km^2 of lagoons.

These islands, lying on the Pacific lithospheric intraplate, define five archipelagos dispersed along a general North-West, South-East direction:

- The Society Archipelago (1,590 km^2) composed of 14 islands (9 high volcanic islands and 5 atolls), itself divided into two groups: The Windward Islands with the capital island of Tahiti (itself a double island: Tahiti Nui and Tahiti Iti) and the Leeward Islands (including the very well-known touristic spot of Bora-Bora Island);
- The Tuamotu Archipelago including 79 atolls spread over an area of 850 km^2 (three main atolls: Rangiroa, Hao and Makemo);
- The Gambier Archipelago (31 km^2), a half-drowned volcano, made up of 9 volcanic islets (main islet: Mangareva) that are the remnants of the lips of the caldera, surrounded to the North and the East by a barrier reef;
- The Marquesas Archipelago (1,049 km^2), divided into the North Marquesas (main island Nuku Hiva), and some coral banks, and the South Marquesas (main island Hiva Oa), including several shoals;
- The Austral Archipelago covering an area of 148 km^2 and including 6 high volcanic islands (principal island: Tubuai), one atoll and one quasi-island (Neilson reef).

Five geodetic tide gauges have been established since 2007 by the Geodesy Observatory of Tahiti over four of these archipelagos (Figure 1 and Table 1, except for the Marquesas), with exact locations mainly dictated by logistical issues (essentially presence of well-stabilized coastal facilities, frequency of airplane/boat links with Tahiti). These five stations mainly focus on the monitoring of the long-term evolution (over many years) of mean sea level (Cazenave et al., 2018). Other tide gauges exist, maintained by the University of Hawaii (UH) and the French Atomic Energy Commission (CEA), but they mainly focus on the monitoring of tsunamis (the typical duration of a tsunami event is a few tens of hours).
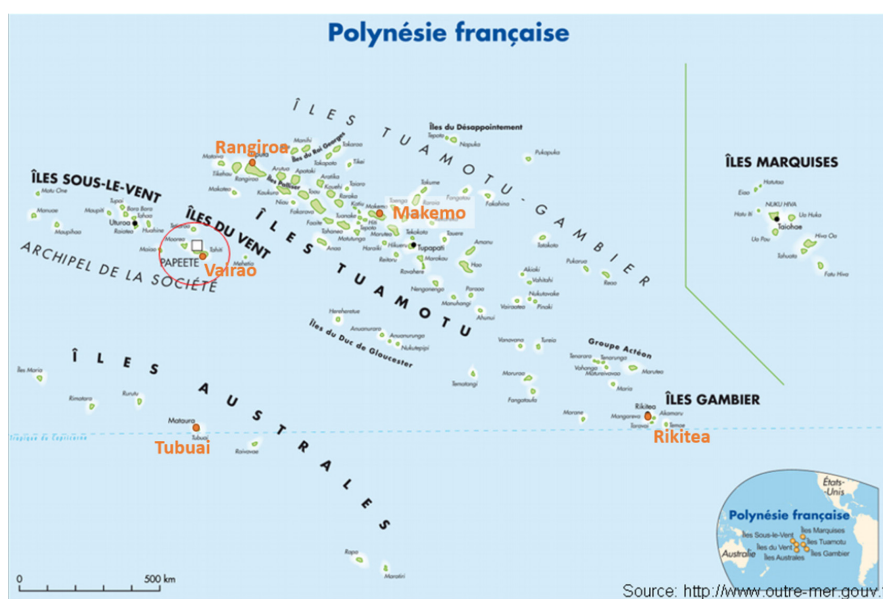

FIGURE 1 The network of geodetic tide gauges in French Polynesia (red points). The area covered in this map is larger than Europe.

FIGURE 2 The typical tide gauge geodetic station of Tubuai (austral archipelago, see Figure 1 for location). One can see the RADAR altimeter (vertical cornet), the GOES UHF Yagi antenna on the roof as well as the GPS radome and a solar panel.

2 | STRUCTURE OF A SEA-LEVEL MONITORING TIDE GAUGE IN FRENCH POLYNESIA

All the five tide gauges (see Figure 1) operating in French Polynesia obey the guidelines set up by GLOSS (IOC, Manual and Guides, 2020). We also refer to Hague et al. (2021) and Marcos et al. (2021). Each tide station is maintained by SHOM following this schedule:

- Typical 18-month maintenance: firmware and configuration upgrade, cable changes, troubleshooting, measurements and reference validation (Van de Casteele test over 12 hours, see Miguez et al., 2008) using a Vegapuls61 mobile RADAR and data retrieval.
- Three-year period maintenance: complete vertical levelling of the sea-level observation site, geodetic observation on benchmarks.

From the beginning of the network to 2020–2021, we can summarize the acquisition process as follows:

- Rangiroa, Vairao, Makemo: hourly satellite data transmission with samples of 60 s every 2 minutes of averaged sea-level records from the two sensors with no time-tag transmitted (60 s average of 4 instantaneous measurements done every 15 s). Only one over two of these 60 s data records is sent and stored locally on the tide gauge. Data are stored locally with a GPS time-tag. The Rangiroa tide gauge was dismantled in January 2021 because of the reconstruction of the dock on which it was located. Its identical reconstruction is scheduled for June 2022.
- Mangareva/Rikitea: satellite data transmission every 5-min of 60 s averaged sea-level records from the two sensors with no time-tag transmitted (60 s average of 4 instantaneous measurements done every 15 s). All these data records are stored locally on the tide gauge with a GPS time-tag.
- Tubuai: satellite data transmission every 15 min of 60 s averaged sea-level records with no time-tag transmitted (60 s average of 4 instantaneous measurements done every 15 s). All these data records are stored locally on the tide gauge with a GPS time-tag.

From 2020 to 2021:

- Vairao since 18 September 2020: hourly satellite data transmission with samples of 60 s every 2 min of averaged sea-level records from the two sensors with no time-tag transmitted (one instantaneous measurement *per* second). Only one over two of these 60 s averaged data records is transmitted, but all the 60 s averaged records are stored with GPS time-tags on the tide gauge.
- Makemo since 12/02/2020: same operating mode as Vairao.
- Mangareva/Rikitea since 24 April 2021: satellite data transmission every 5-min of 60 s averaged sea-level records from the two sensors with no time-tag transmitted (one instantaneous measurement *per* second). All these 60 s averaged data records are stored locally on the tide gauge with a GPS time-tag.
- Tubuai since 16 December 2020: satellite data transmission every 15 min of 60 s averaged sea-level records from the two sensors with no time-tag transmitted (one instantaneous measurement *per* second). All these 60 s averaged data records are stored locally on the tide gauge with a GPS time-tag.

Additional details about the acquisition and preprocessing are given in the following sections and additional

TABLE 1 Location of the geodetic tide gauges in French Polynesia managed by the Geodesy Observatory of Tahiti/UPF. We also mentioned the tide gauge of the University of Hawaii (Rikitea UH), which was used for comparison purposes (see Section 7).

Number	Station name	Starting date	Coordinates	Sensors	Management	Sampling	GOES DCP satellite transmission interval	Identification GLOSS	PSML
1	Makemo	Oct. 2013	16°37,63'S 143°34,15'W	RADAR, pressure, GPS	UPF	2 min	FAA04758 60 min	354	
2	Rangiroa	Feb. 2009	14°56,75'S 147°42,36'W	RADAR, pressure, GPS	UPF	2 min	B1F011F6 60 min	355	2,216
3	Mangareva/ Rikitea	Mar. 2012	23°07,07'S 134°58,13'W	RADAR, pressure, GPS	UPF	1 min	FAA0542E 5 min		
4	Tubuai	Dec. 2008	23°20,50'S 149°28,53'W	RADAR, pressure, GPS	UPF	1 min	FFF0073A 15 min	356	2,217
5	Vairao	Feb. 2011	17°48,35'S 149°17,7'W	RADAR, pressure, GPS	UPF	2 min	FAA022BE 60 min		2,242
6	Rikitea UH	June. 2008	23°07,33'S 134°58'W	Float sensor, RADAR, pressure	UH	1 min	3541F0D6 5 min	138	1,253

information about the stations can be found in Barriot et al. (2011).

3 │ TIDE GAUGE RAW DATA AND TIME-TAGGING

The sea-level data acquired by these five stations were (and still are) being transferred in real time by a UHF radio-link to the NOAA (National Oceanic and Atmospheric Administration of the USA) through the GOES satellite constellation (GOES - Geostationary Operational Environmental Satellite), but without neither GPS *in situ* time-tagging nor GPS positioning data, to save transmission bandwidth, and were *ex situ* time-tagged again upon arrival at the NOAA datacentres. A long-term data archive is provided by the Sea Level Monitoring Facility database of IOC (Intergovernmental Oceanographic Commission, www.ioc-sealevelmonitor ing.org).

This data transfer through the GOES satellite constellation was the only way to have real time transfer during most of this data acquisition period, as the only undersea optical cable was running between Hawaii and Tahiti and data transmission through commercial satellites was too expensive. All the raw data, including *in situ* GPS time-tagging and positioning, were also locally recorded on flash drives at the tide gauges and manually transferred to the UPF database when maintenance took place. Sometimes the local recording at the tide gauges was not operating, so the only source of data for this study was the IOC database, sometimes the GOES data link was not operating, so the only source of data was the local records on the tide gauges.

We started our validation process by doing a cross-comparison of the IOC datasets versus the raw UPF datasets. We found that the time-tagging on the IOC database was often off by several minutes with respect to the GPS time-tagging done at the tide gauges, and sometimes shifts up to 1 hr were found (Table 2). We therefore elected to retain the UPF dataset whenever possible. When the UPF and IOC time series have no overlap, it was nevertheless always possible to check possible time drifts by considering the coherence of the phase of the IOC recorded main tidal waves with a reconstructed phase (see thereafter).

The observation time corresponds to the middle of the time interval, that is, it is shifted of 30 s for a one-minute integration period. This time shift is indeed negligible for long-period phenomena which are the main goal of these sea-level measurements. For tidal analysis, a time shift of 30 s was sometimes during the analysis procedure.

TABLE 2 Number of time-tag errors (in minutes) *per* tide gauge identified in the IOC database records, with respect to the original or reconstructed time-tags

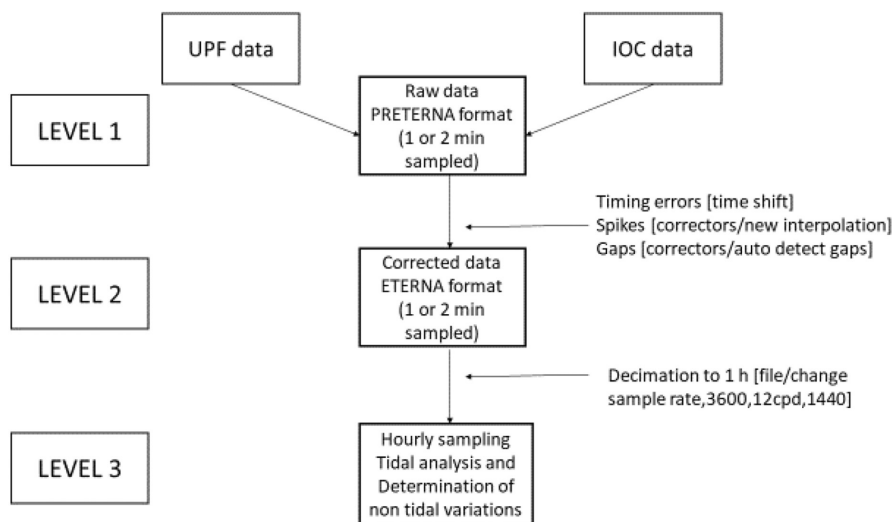|	−23 mn	2 mn	7 mn	32 mn	47 mn	61 mn	62 mn
RANGIROA	0	0	0	0	0	0	555,711
TUBUAI	0	0	877,035	0	0	0	0
MAKEMO	0	7	0	44,760	490,260	0	0
VAIRAO	368,097	0	0	0	0	112,275	0
MANGAREVA	0	0	0	0	0	0	0

FIGURE 3 Preprocessing dataflow. The Tsoft facilities used for corrections are put inside square brackets in the diagram. The PRETERNA and ETERNA formats are described in Wenzel (1995).

4 | DATA VALIDATION AND INTERMEDIATE DATABASE

The raw data cannot be directly inputted in a tidal analysis software. A preprocessing step is necessary, to identify and remove gross errors (incorrect or missing data producing spikes and gaps or timing problems). This preprocessing was done as part as the Tsoft tidal visualization/modelling environment, developed and maintained by the Royal Observatory of Belgium (see TSoft manual, Van Camp & Vauterin, 2005, see http://seismologie.be/en/downloads/tsoft). This software and the tidal analysis software ET34-ANA-V80 V80 (Ducarme & Schueller, 2018; Schueller, 2015; Schueller, 2017) are using specific data presentations called 'PRETERNA and ETERNA formats', described in Wenzel, 1995, 1996. The procedure is divided into three data levels, as shown in Figure 3.

LEVEL 1 data are obtained by rewriting the original records (UPF or IOC data) in the so-called PRETERNA format (Wenzel, 1995). Between LEVEL 1 and LEVEL 2, the data sequence is checked, and preliminary tidal analyses are performed on the less perturbed parts of the data to prepare a preliminary tidal model and to detect timing errors. It is always possible to detect any timing error by considering the coherence of the phase of the main tidal waves with the phase computed from the UPF data, which were archived with the correct *in situ* timing.

In the semidiurnal band, a timing error of 2 min produces a phase difference of one degree at the frequency of 2 cpd (cycles-*per*-day). If necessary, a time shift is applied. Corrections are applied by means of the software Tsoft and a 'remove-restore' procedure: The preliminary tidal model is subtracted from raw data; the corrections are applied on the residuals and the corrected tidal signal is reconstructed by summing up again the removed tidal signal. The outliers, mainly unwanted 'spikes', are identified by eye and removed and small interruptions (gaps no larger than 24 h) are automatically tagged and interpolated with the same procedure.

Between LEVEL 2 and LEVEL 3, the data are averaged down to a 1-hr sampling rate. This averaging process filters out short period noise but maintains all signals with period larger than the lowest Nyquist period considered in our tidal analysis (2-hr period). Besides, this unifies datasets acquired in the same station but with different sampling rates, for example, 1 and 2 min, as it occurred in Mangareva and Tubuai. These LEVEL3 data are convenient for most applications except detection of transient signals such as tsunamis or sea swells, which require a higher sampling rate. The averaging is performed by means of a standard least-squares low-pass filter provided by Tsoft. This filter has a cutoff frequency at 12 cpd and a half-a-length of 1 day, that is, 1,440 or 720 samples for a 1 min or 2 min sampling rate. The LEVEL3 provides the

TABLE 3 Comparison of RADAR and pressure tidal amplitudes A and Greenwich phases α for the main semidiurnal waves

Station	Wave	Pressure		RADAR			
		A_P(mm)	α_P(°)	A_R(mm)	α_R(°)	A_P(mm)/A_R(mm)	α_P(°)-α_R(°)
Rangiroa 2015-06/2019-01	M2	117.3 ±0.5	40.2 ±0.2	117.4 ±0.4	39.6 ±0.2	0.999	0.6
	S2	36.5 ±0.5	13.2 ±0.2	36.4 ±0.4	13.3 ±0.6	1.003	0.1
Makemo 2013-10/2018-11	M2	175.5 ±0.3	69.9 ±0.1	171.9 ±0.3	73.6 ±0.1	1.021	−3.7
	S2	46.8 ±0.3	41.3 ±0.4	45.8 ±0.3	45.1 ±0.4	1.022	−3.8
Vairao 2011-07/2018-11	M2	101.7 ±0.5	−55.4 ±0.3	100.8 ±0.5	−53.9 ±0.3	1.009	−1.5
	S2	86.2 ±0.5	−40.4 ±0.3	85.2 ±0.5	−40.4 ±0.4	1.012	0.0
Tubuai 2010-02/2016-10	M2	152.1 ±0.4	−102.0 ±0.1	152.9 ±0.5	−101.7 ±0.2	0.995	−0.3
	S2	103.0 ±0.2	−52.5 ±0.2	103.5 ±0.5	−52.3 ±0.3	0.995	−0.2
Mangareva 201,203/202101	M2	265.0 ±0.3	−11.3 ±0.1	263.8 ±0.3	−8.8 ±0.1	1.005	−2.5
	S2	92.7 ±0.3	−53.6 ±0.2	92.5 ±0.3	−50.9 ±0.2	1.002	−2.7

Note: Amplitude and phases are rounded up to the last significant digit corresponding to the associated RMS error.

final dataset from which the nontidal sea-level variations are derived from the tidal analysis performed by the software ET34-ANA-V80. The latest version of the software proposes a special option for ocean tides analysis and prediction in agreement with oceanographic conventions (i.e. phases refer to Greenwich and phase lags are counted positive, the changes of sign of the tidal potential with latitude are not taken into account). This program can be downloaded from http://ggp.bkg.bund.de/eterna/.

These tidal residuals outputted by this tidal analysis software are a preliminary representation of the variations of the Mean Sea Level (RSL) due, up to the first order, to meteorological phenomena and, up to the second order, to climate related effects.

To perform the different processing steps shown on Figure 3, an intermediate database (see Annex) was set up to collate the raw and validated data as well as the analysis results and the nontidal sea-level fluctuations.

The conception of this intermediate database was derived from the following considerations:

- As already stated, we have two data sources: the University of French Polynesia and 'IOC, Sea Level Monitoring Facility' (www.ioc-sealevelmonitoring. org). As the initial presentation of the data is different, the two sources were treated separately until the final analysis.
- There are two sea-level sensors (and therefore data formats): RADAR and PRESSURE.

The database (annex) was organized in 3 levels corresponding to the 3 levels of data processing described above:

- LEVEL 1: raw data converted in PRETERNA format

  (sampling rate 1 min or 2 min);

- LEVEL 2: data corrected with Tsoft and converted to ETERNA format (Wenzel, 1995).

  (sampling rate 1 min or 2 min);

- LEVEL 3: data averaged to hourly values and tidal analysis residuals.

5 | OCEAN TIDES RESULTS

The tidal regime is described by the form ratio F (van der Stok, 1897):

$$F = \frac{K1 + O1}{M2 + S2}$$

$F > 3$ Diurnal 1 High, 1 Low *per* day;

$0.25 \; F < 3$ Mixed 2 Highs, 2 Lows *per* day, but of different strengths;

$F < 0.25$ Semidiurnal 2 Highs, 2 Lows *per* day, similar strengths.

In French Polynesia, the form ratio F is always lower than 0.2, corresponding to dominating semidiurnal tides. A detailed discussion of the tidal signal can be found in Ducarme et al. (2022). We shall focus here on the internal coherency of the PRESSURE and RADAR records.

We compared the results of the RADAR with those of the PRESSURE gauge (Table 3). The tidal amplitudes are in perfect agreement, except for Makemo and Vairao stations. This calibration difference is confirmed by a linear regression between the associated tidal models (noted Theo):

In Makemo: $\text{Theo}_{PRESSURE} = 1.0158 \pm 0.003 * \text{Theo}_{RADAR}$ (correlation coefficient $r = 0.9980$)

In Vairao: $\text{Theo}_{PRESSURE} = 1.0118 \pm 0.002 * \text{Theo}_{RADAR}$ (correlation coefficient $r = 0.9987$)

In Makemo, a large phase difference of 4° (8 min of time difference) is observed. In Mangareva, also a difference of phase of 2.5° exists between PRESSURE and RADAR records, corresponding to a time difference of 5 min.

We checked also the stability of the tidal amplitudes and phases of the main semidiurnal waves M2 and S2 (Table 4). For that purpose, we cut each dataset in two consecutive blocks in time and compared the results. The differences in phase reach a level of 1° (2 min of time). We can thus expect a precision of 0.5° in phase on the complete dataset. The amplitudes are perfectly stable for the stations of Makemo, Tubuai and Mangareva. For these stations, we expect a precision of 1% for the tidal predictions, that is, better than a centimetre precision. In Rangiroa and Vairao, there is an apparent difference of amplitude for the M2 wave, but S2 is always stable. A change of calibration is thus a very unlikely cause. This phenomenon is perhaps associated with a larger noise level around the M2 frequency compared with the noise level around the S2 frequency, probably linked to seasonal modulations. The precision of the tidal predictions is reduced to a centimetre precision for these stations.

6 | NONTIDAL EFFECTS

The PRESSURE measurements show the same long-term fluctuations as the RADAR measurements, but are affected by jumps and a steady, and sometimes large, drift. For the determination of the mean sea level (RSL), we therefore only considered the RADAR data. Both dataset present long-period fluctuations with seasonal characteristics (Figures 4 and 5). The relative sea level

(RSL), measured with respect to the tide gauge setting and obtained from the RADAR measurements, are given in Table 5 and Figure 6. Their variations (ΔRSL) can be transformed into variations of the mean sea level (ΔMSL) if we can measure the vertical displacement of the crust on the site (Δh), as we have.

$$\Delta MSL = \Delta RSL + \Delta h.$$

Almost all Polynesian islands are subject to slow subsidence (typically at the level of millimetre or sub-millimetre/year), caused by the thermal contraction of their underlying tectonic plates as they move away from mid-ocean ridges, as well as evolution towards isostatic equilibrium with the underlying lithosphere/mantle. Δh values reported in the literature (Becker et al., 2012; Fadil et al., 2011; Martinez-Asensio et al., 2019) are given in Table 6 for Papeete (close to Vairao) and Rikitea. Contradictions are present for Papeete. The treatment of the GPS observations performed at the stations should provide Δh (and then ΔMSL) with respect to the world-wide WGS84 ellipsoid, but at the moment the values that are reported in this work are with respect to each island local datum. This extremely complicated and delicate computer processing of the GPS data, again at the submillimetre level/year, is in progress and will be the subject of a dedicated publication. We refer to Adebisi et al. (2021) for a review on this subject.

At Rangiroa, a fast drift of 80 mm occurred in the beginning of 2014. A one-year interruption in the data acquisition occurred between July 2014 and July 2015, and we found a difference of 50 mm between the computed RSL values of the two blocks of data (before and after the gap). A global adjustment provided the following value: **RSL = 404.0 ± 1.2 mm** with an RMS error on one observation m0 = 47 mm and a variation **ΔRSL = +84.0 ± 3.0 mm** over the time interval 2009–2021. The short period noise reaches a ±40 mm amplitude with long period excursions up to ±100 mm.

At Makemo, the RMS error on one observation is large (m0 = 100 mm). It is associated with large water level rises, up to 75 cm during the winter season, which are associated with sea swell crises (Figure 4). The mean value of the set is **RSL = 655.5 ± 3.4 mm**. Due to the large swell yearly modulation, the amplitude of Sa reaches 66 cm and the long-term trend is not well extracted with **ΔRSL = −7.7 ± 4.9 mm**.

At Vairao, an unusual event happened at the beginning of 2014. There is always a large rise of the sea level starting at the beginning of each year, which was more pronounced at the beginning of that year (Figure 5). Later on, the apparent RSL remains quite stable with **RSL = 530.9 ± 1.3 mm**. The first part (2011/07–2014/01) has a mean value (503.8 mm), 30 mm lower than for the following records. However, the global dataset 2011–2021

TABLE 4 Evolution of tidal amplitudes A and Greenwich phases α for the main semidiurnal waves (RADAR, hourly sampling)

Station	Wave	Bloc1		Bloc2		Comparison	
		A_1(mm)	α_1(°)	A_2(mm)	α_2(°)	A_1(mm)/A_2(mm)	α_1(°)-α_2(°)
Rangiroa		2009-06/201406 1,570 days		201506/202102 2,074 days			
	M2	119.0±0.5	39.2±0.2	113.5±0.4	40.3±0.2	1.05	−1.1
	S2	36.5±0.5	12.5±0.8	36.0±0.4	13.6±0.7	1.01	−1.1
Makemo		2013-10/201704 1,248 days		201705/202012 1,354 days			
	M2	171.7±0.3	73.6±0.1	172.4±0.5	72.6±0.2	1.00	1.0
	S2	46.1±0.3	44.2±0.5	45.6±0.5	45.3±0.6	1.01	−1.1
Vairao		201107/201603 1,453 days		201604/202101 1,744 days			
	M2	97.8±0.9	−55.2±0.4	101.7±0.5	−53.7±0.3	0.96	−1.5
	S2	84.8±0.9	−41.1±0.5	85.9±0.5	−39.6±0.4	0.99	−1.5
Tubuai		201002/201408 1,633 days		201510/202009 1,457 days			
	M2	152.9±0.5	−101.7±0.2	154.4±0.4	−101.9±0.1	1.00	0.2
	S2	103.5±0.5	−52.3±0.3	104.1±0.4	−51.3±0.2	0.99	−1.0
Mangareva		201203/201612 1,239 days		201712/202012 1,118 days			
	M2	263.5±0.3	−9.1±0.1	264.6±0.4	−7.8±0.1	1.00	−1.3
	S2	92.2±0.3	−51.0±0.2	93.0±0.4	−50.3±0.2	0.99	−0.7

Note: Amplitude and phases are rounded up to the last significant digit corresponding to the associated RMS error. The datasets were divided into two blocks for comparison purposes.
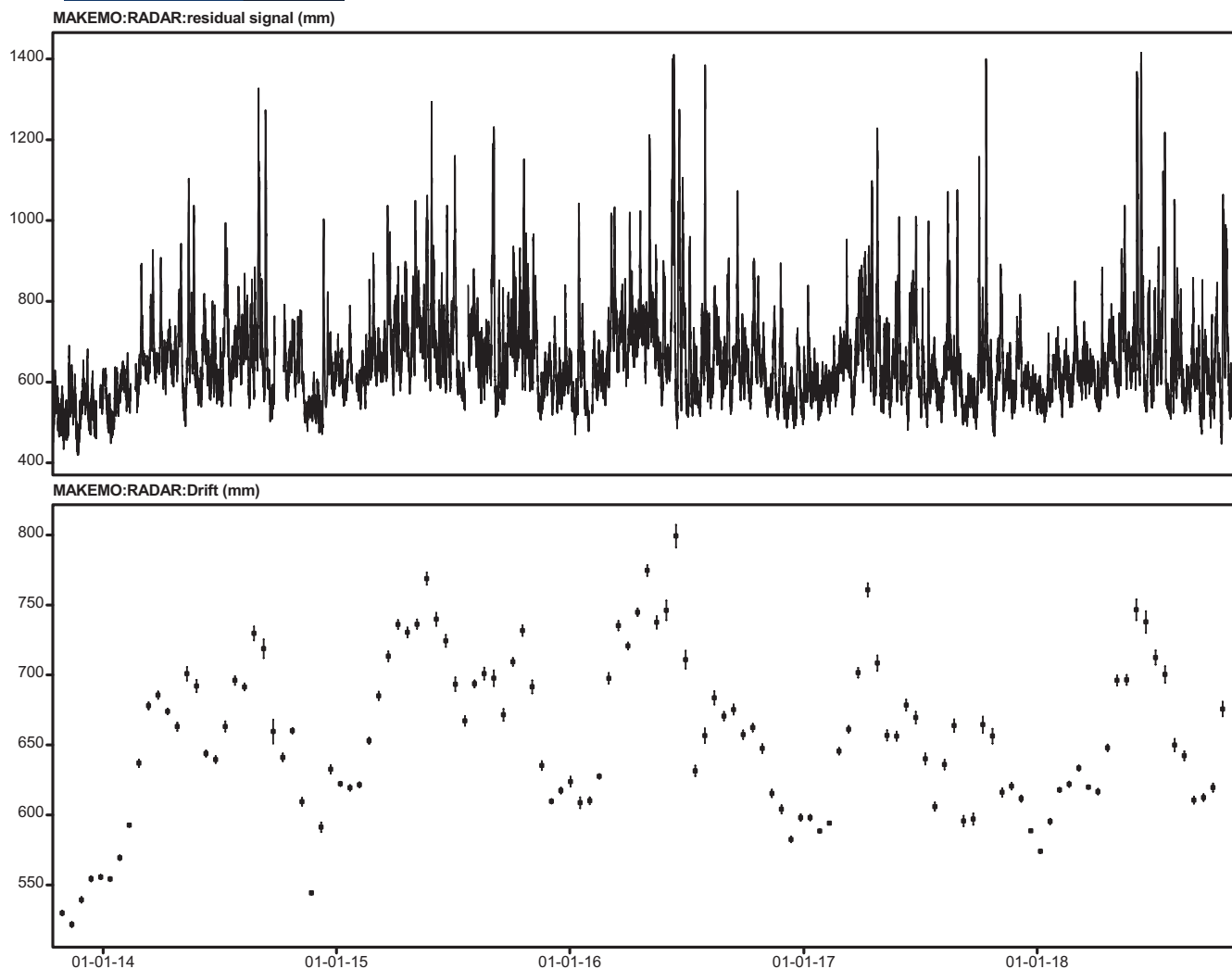
FIGURE 4 Makemo tide station (RADAR sensor). Top: Residual signal showing nontidal sea-level variations. Bottom: Monthly values of the nontidal variations showing the seasonal variations associated with Sa = 5 cm.

provides a very coherent result, **RSL = 518.0 ± 1.2 mm** and **ΔRSL = +36 ± 3 mm**, with m0 = 50 mm.

At Tubuai, the results obtained with a first dataset (2010–2014, sampling rate 2 min) and a second one (2015–2020, sampling rate 1 min), after an interruption of 14 months, are coherent within the associated 2σ error-range for what concerns the tidal signal, but an increase of 40 mm of the apparent RSL (Table 5) is certainly linked to a maintenance intervention at the restart of the station. Here also a fast drift occurred at the beginning of 2014 and this increase is probably an artefact, as a global adjustment provided the value **RSL = 706.0 ± 2.1 mm** with m0 = 74 mm and a variation **ΔRSL = + 54 ± 5 mm** over the time interval 2010–2020.

The large sea-level fluctuation observed in the beginning of 2014 for Rangiroa, Vairao and Tubuai is not affecting the RSL value on the complete dataset but is

hampering the determination of RSL values on smaller blocks (Tables 4 and 5).

At Mangareva, the records at the UPF station are missing between 2017 and 2018. An unusual event happened at the beginning of 2018, but the RSL is the same for the two blocks (before and after the gap, see Tables 4 and 5) and we get for the complete interval 2012–2020 a value **RSL = 645.2 ± 2.0 mm**. We tested the possibility to determine a temporal variation of the RSL. We got a variation **ΔRSL = +7.4 ± 5.2 mm**. This result is not significant as the associated RMS error is very large. Fortunately, a sea-level monitoring station of the University of Hawaii (UH), labelled Rikitea in the IOC catalogue, is installed a few 100 m apart from the UPF station (see Figure 7). We downloaded the data relative to this station from the IOC database and a comparison is presented in the next section.
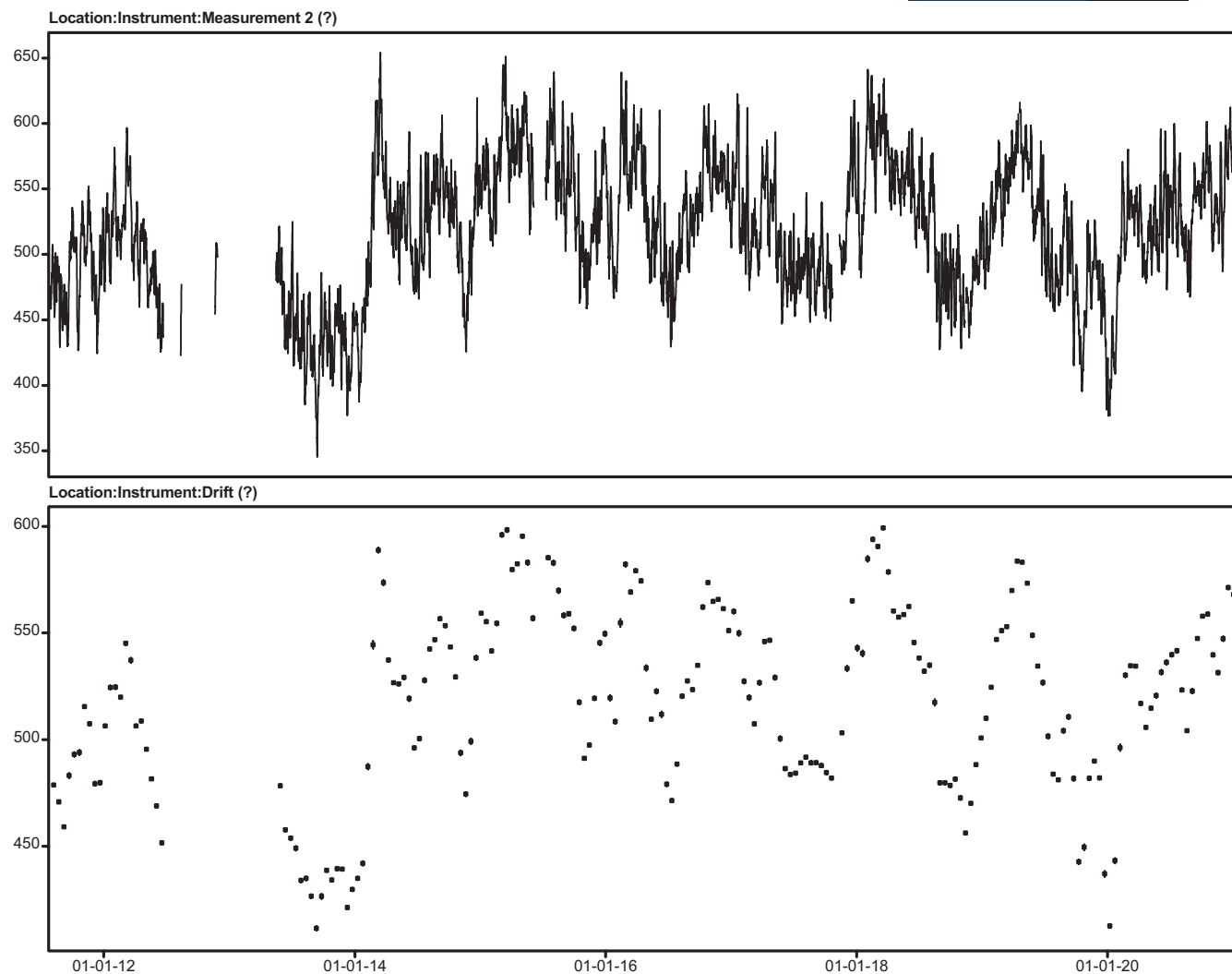
Location:Instrument:Measurement 2 (?)

Location:Instrument:Drift (?)

FIGURE 5 Nontidal sea-level variations at the Vairao tide station (RADAR sensor) showing up the annual signal and the January 2014 unusual event. Top: Original sampling, bottom: Sliding monthly average.

7 | COMPARISONS BETWEEN THE RECORDS FROM THE TIDE GAUGE DATA MAINTAINED BY HAWAII UNIVERSITY AND THE UPF TIDE GAUGE IN MANGAREVA

Currently, the tide gauge of the University of Hawaii hosts three different instruments: A FLOAT sensor with an encoder, a PRESSURE gauge and a RADAR altimeter. The dataset covers the period 2008/07–2021/04. The sampling rate is 5 min for the FLOAT sensor and 1 min for the two other sensors. We found that the PRESSURE sensor was badly perturbed between 2012 and 2019, so we ignored these measurements. The results of the FLOAT sensor and the RADAR are in very good agreement. There is no difference in amplitude with the results of the UPF RADAR but only a slight phase difference of 1° in the semidiurnal band. The nontidal variations are also in perfect agreement (Figure 8), as indicated also by

the differences between sensors shown in Figure 9. The unusual event at the beginning of 2018 disappears on this graph. This clearly indicates that it was a real signal. The difference between the sensor measurements is of the order of 30 mm peak to peak at short periods. There is also a slight differential drift especially at the beginning of the float sensor record.

A linear regression between the different datasets gives:

$\text{RADAR}_{UH} = 1.0025 \pm 0.0002 * \text{RADAR}_{UPF}$ (correlation coefficient $r = 0.999$).

$\text{RADAR}_{UH} = 0.9999 \pm 0.0002 * \text{FLOAT}_{UH}$ (correlation coefficient $r = 0.999$).

There is only a 0.3% calibration difference between the UPF and the UH RADAR measurements. The UH FLOAT sensor calibration agrees with the UH RADAR calibration.

Thanks to the long period of nearly 13 years covered by the University of Hawaii tide gauges in Mangareva, it was possible to determine a secular trend, as the drift of the RADAR and of the FLOAT sensor are very similar.

TABLE 5 Relative sea level (RSL) and RSL variations (ΔRSL) with RADAR (hourly sampling), m0: RMS error on one observation

Rangiroa		Makemo		Vairao		Tubuai		Mangareva	
2009/06–2021/02		2013/10–2020/12		2011/07–2021/01		2010/02–2021/01		2012/05–2021/01	
3,644 days		2,565 days		3,197 days		3,195 days		2009 days	
m0 = 47 mm		m0 = 104 mm		m0 = 50 mm		m0 = 74 mm		m0 = 56 mm	
F = 0.13		F = 0.11		F = 0.15		F = 0.17		F = 0.12	
RSL (mm)	Epoch	RSL (mm)	Epoch	RSL (mm)	Epoch	RSL (mm)	Epoch	RSL (mm)	Epoch
374.4±1.6	2009–2014	653.64±2.8	2013–2020	503.1±2.4	2011–2014	686.5±2.6	2010–2014	644.2±2.6	2012–2016
426.9±1.6	2014–2018			530.9±1.4	2014–2020	725.1±3.2	2015–2020	646.6±2.7	2018–2020
429.0±2.4	2014–2021			518.0±1.2	2011–2021	706.0±2.1	2010–2021	645.2±2.0	2012–2020
404.0±1.2	2009–2021								
ΔRSL (mm)									
+84.0±3.0				+36.0±3.0		+54.0±5.0		+7.0±5.0	

The variation of the RSL between 2008/07 and 2021/04 is +25.2±3.23 mm from the RADAR data and 16.4±3.1 mm for the FLOAT sensor data. The discrepancy is due to the anomalous drift at the beginning of the FLOAT sensor records (Figure 9).

8 | CONCLUSIONS

This study provides cleaned, calibrated and with a verified time-tagging of sea-level data for the five tide gauge geodetic stations installed in French Polynesia (see Figure 1) for the period 2009–2020. For preprocessing and analysis of the data, a three-level operational database was set up:

- LEVEL 1 raw data with original sampling rate,
- LEVEL 2 corrected data with original sampling rate,
- LEVEL 3 hourly averaged data and tidal analysis residues.

The database, entitled 'SEA LEVEL collected from TIDE STATIONS in South Pacific Ocean from 2009-06-13 to 2021-01-28', can be accessed on the NOAA National Centers for Environmental Information (NCEI) data servers as 'NCEI Accession 0244182' and contains two subsets: The first one is relative to the original sampling rate (LEVEL1 and LEVEL2) and the second one (LEVEL3) is relative to the hourly re-sampling. The harmonic tide model coefficients derived from the analysis of the hourly data subset are also provided for each tide gauge station.

The RADAR and PRESSURE gauge measurements of the five UPF/OGT tide stations have been both validated for the precise measurement of the tides and short-term phenomena such as tsunami waves. The RADAR dataset is the only one suitable for long-term analysis of Mean Sea Level (MSL) variations (see Figure 6). Of course, as who can do more can do less, RADAR measurements are also suitable for the study of short-term phenomena.

A metrological modelling of the absolute variations of the Mean Sea Level (ΔMSL) requires the establishment of a common datum for the five tide gauges with respect to a reference ellipsoid (through GNSS positioning), as well as the modelling of long-term fluctuations affecting measurements, whether of instrumental or geophysical origin (meteorological, subsidence of the islands Δh). Nearly all atolls and islands in the South Pacific are of volcanic origin and are subject to subsidence during their lifetime, with vertical velocities ranging from zero to a few mm/year (Fadil et al., 2011). This delicate task has still to be done. We nevertheless performed preliminary separated adjustments of the relative sea-level trends (ΔRSL) for each station, with the following conclusions:

FIGURE 6 Nontidal variations shown at the same scale (500 mm full scale). From top to bottom: Rangiroa, Makemo, Vairao, Tubuai and Mangareva tide stations. The red arrows indicate an unusual fast sea-level rise at the beginning of 2014.

TABLE 6 Previous results found in the literature

Station	Papeete		Rikitea	
	ΔRSLy (mm/year)	Δh (mm/year)	ΔRSLy (mm/year)	Δh (mm/year)
Becker et al., 2012	34 years 2.9 ± 0.5	11.4 years −0.4 ± 0.1	40 years 2.1 ± 0.4	
Martinez-Asensio et al., 2019	1969–2014 3.2 ± 0.4	2004–2013 −1.9 ± 0.2	1968–2015 1.7 ± 0.3	2000–2016 −1.0 ± 0.4
Fadil et al., 2011		1998–2009 −0.5 ± 0.1		

ΔRSLy, Annual relative sea-level trend; Δh, Vertical motion rate from GPS observations.

1. In Rangiroa, there is an apparent upward drift of 7.3 ± 0.3 mm *per* year, for the period 2009–2021, but the time series presents a large interruption. This value is large compared with the global sea-level rate of change derived from satellite altimetry (3.4 ± 0.4 mm from GFSC, 2021), but there are perhaps local factors

to take into account, like a local subsidence of the tide gauge;

2. In Makemo, the sea level remains stable, but the RMS error m0 is very large. It is associated with large water level rises, up to 75 cm during the winter season, which seems to correspond to sea swell crises. The series
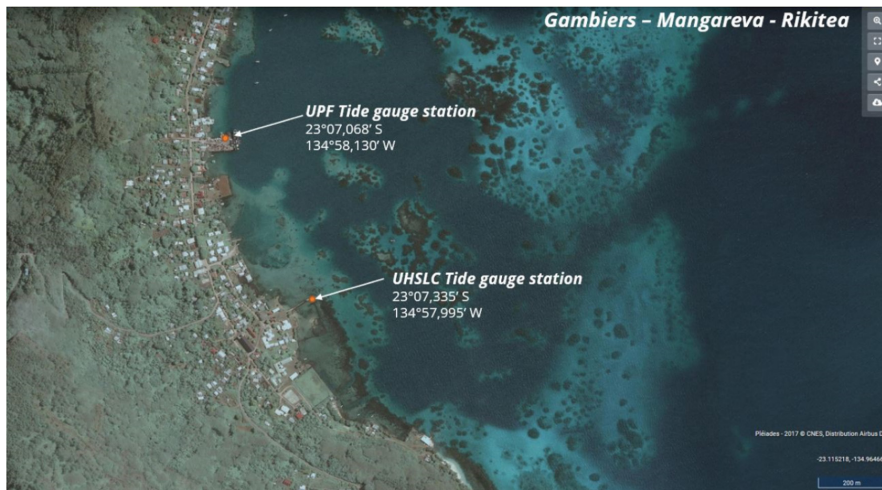
FIGURE 7 Locations of the tide gauges maintained by the University of Hawaii (UH) and the University of French Polynesia (UPF) in the island of Mangareva (Gambier archipelago). They are separated by less than 1 km. The first one (UH) is in a very calm fishpond, totally buffered from the lagoon, and the second one (UPF) is in the main harbour of the island, close to the boat terminal going to the airport.
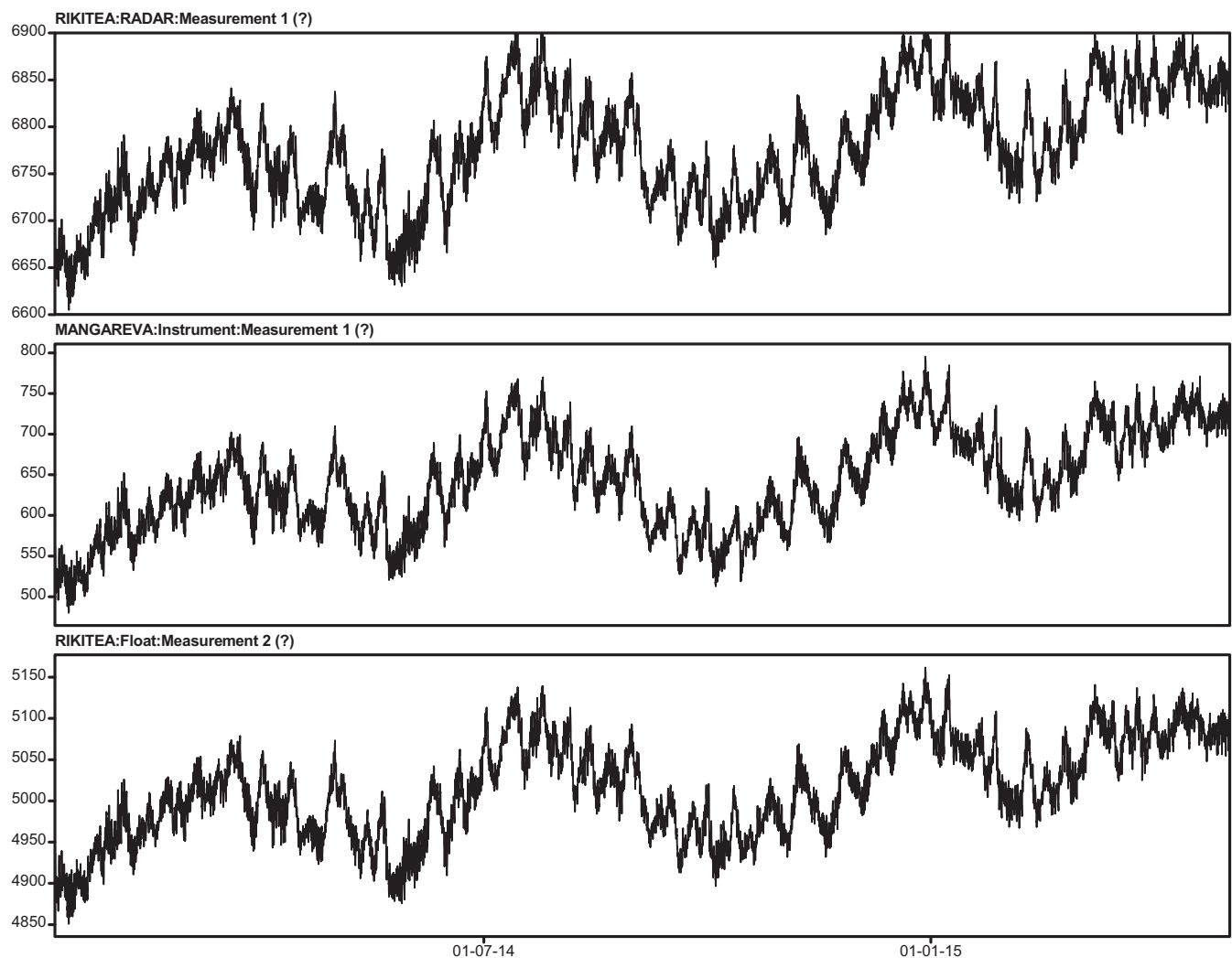
FIGURE 8 Residual signal in mm after subtraction of the tidal model at Mangareva/Rikitea for different sensors. Top: RADAR UH, middle: RADAR UPF, bottom; FLOAT sensor UH.

is also too short for a correct determination of RSL variations;

3. In Vairao, the series is homogeneous with an apparent upward drift of 3.8 ± 0.3 mm *per* year, for the period 2011–2021. At the Fare Ute Point tide stations of the University of Hawaii, situated at about 40 km from Vairao in the main harbour of Papeete, Becker et al. (2012) and Martinez-Asensio et al. (2019) report,
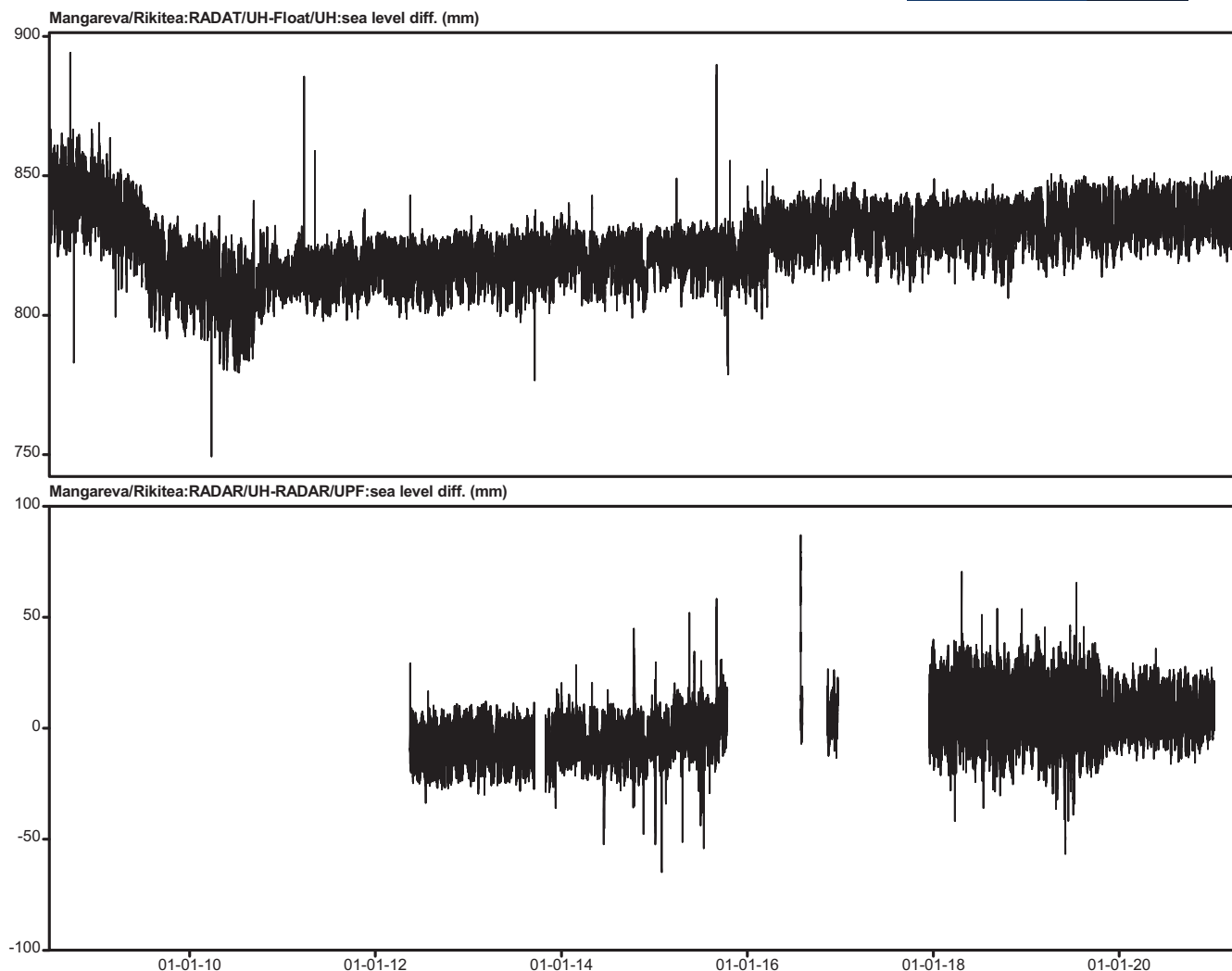
FIGURE 9 Comparison of the different sensors between Mangareva (UPF) and Rikitea (UH) tide stations. Top: UH RADAR–UH FLOAT, bottom: UH RADAR–UPF RADAR. The short-term noise is of the order of 30 mm peak to peak.

respectively, $+2.9 \pm 0.5$ mm/year and $+3.2 \pm 0.4$ mm/year (Table 6). Δh values range between -0.4 mm/year and -1.9 mm/year. These results are compatible with an absolute trend in sea level $\Delta MSL = 2.75 \pm 1.46$ mm/year, between 1993 and 2019, estimated using observations from satellite altimeters (GFSC, 2021);

4. In Tubuai, there is an apparent upward drift of 4.9 ± 0.5 mm *per* year, for the period 2010–2021, but the time series presents a large interruption;

5. In Mangareva, the 2017 data records are missing in both the UPF and IOC databases. The two data sections (before and after the gap) do not give different RSL values. It was possible to determine the variation of the RSL between 2008/07 and 2021/04 by using the uninterrupted data of the University of Hawaii at $+2.0 \pm 0.4$ mm *per* year with the RADAR. Becker et al. (2012) and Martinez-Asensio et al. (2019) report, respectively, $+2.1 \pm 0.4$ mm/year and $+1.7 \pm 0.3$ mm/year. (Table 6). Δh value reported by Martinez-Asensio et al.

is -1.0 ± 0.4 mm/year. The absolute trend in sea level ΔMSL from 1993 to 2019 using observations from satellite altimeters for this site is estimated at 2.76 ± 1.18 mm *per* year (GFSC, 2021).

In conclusion, there is a reasonable agreement between the sea-level rate of change determined with satellite altimetry and our result in Vairao and in Mangareva. Our values are larger in Tubuai and particularly in Rangiroa, but these datasets are less reliable as they present large interruptions.

AUTHOR CONTRIBUTIONS

visualization (equal); writing – original draft (equal). **Bernard Ducarme:** Data curation (equal); formal analysis (equal); methodology (equal); software (equal); writing – original draft (equal); writing – review and editing (equal). **Guy Wöppelmann:** Resources (equal); writing – original draft (equal); writing – review and editing (equal). **Gaël André:** Resources (equal); writing – original draft (equal); writing – review and editing (equal). **Alban Gabillon:** Data curation (equal); software (equal); writing – review and editing (equal).

OPEN RESEARCH BADGES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at [[https://www.ncei.noaa.gov/archive/accession/0244182]]. Learn more about the Open Practices badges from the Center for Open Science: https://osf.io/tvyxz/wiki

ORCID

Jean-Pierre Barriot https://orcid.org/0000-0002-2112-9685

REFERENCES

Adebisi, N., Balogun, A.L., Min, T.H. & Tella, A. (2021) Advances in estimating sea level rise: a review of tide gauge, satellite altimetry and spatial data science approaches. *Ocean and Coastal Management*, 208, 105632. https://doi.org/10.1016/j.oceco aman.2021.105632

Barriot, J.P., Serafini, J., Sichoix, L., Reymond, D. & Hyvernaud, O. (2011) The tsunami of march 11, 2011 as observed by the network of tide gauges of French Polynesia. *Journal of Marine Science and Technology*, 20(6), 639–646. https://doi.org/10.6119/JMST-012-0430-1

Ducarme, B., Barriot, J.-P. & Zhang, F. (2022) Zhang, combination of Tsoft and ET34-ANA-V80 software for the preprocessing and analysis of tide gauge data in French Polynesia. *Geodesy and Geodynamics*. https://doi.org/10.1016/j.geog.2022.05.002

Becker, M., Meyssignac, B., Letetrel, C., Llovel, W., Cazenave, A. & Delcroix, T. (2012) Sea level variations at tropical Pacific islands since 1950. *Global and Planetary Change*, 80-81, 85–98.

Cazenave, A., Palanisamy, H. & Ablain, M. (2018) Contemporary Sea level changes from satellite altimetry: what have we learned? What are the new challenges? *Advances in Space Research*, 62(7), 1639–1653. https://doi.org/10.1016/j.asr.2018.07.017, https://www.sciencedirect.com/science/article/pii/S0273 117718305799

Ducarme, B. & Schueller, K. (2018) Canonical wave grouping as the key to optimal tidal analysis. *Bulletin d'Information des Marées Terrestres*, 150, 12131–12245. http://www.bim-icet.org/

Fadil, A., Sichoix, L., Barriot, J.-P., Ortéga, P. & Willis, P. (2011) Evidence for a slow subsidence of the Tahiti Island from GPS, DORIS, and combined satellite altimetry and tide gauge sea level records. *Comptes Rendus Geoscience*, 343(5), 331–341. https://doi.org/10.1016/j.crte.2011.02.002

GSFC, (2021) Global Mean Sea Level Trend from Integrated Multi-Mission Ocean Altimeters TOPEX/Poseidon, Jason-1, OSTM/Jason-2, and Jason-3 Version 5.1. Ver. 5.1 PO.DAAC, CA, USA. https://doi.org/10.5067/GRSLM-TJ151. https://climate.nasa.gov/vital-signs/sea-level/

Hague, B.S., Jones, D.A., Trewin, B., Jakob, D., Murphy, B.F., Martin, D.J. et al. (2021) ANCHORS: A multi-decadal tide gauge dataset to monitor Australian relative sea level changes. *Geoscience Data Journal*, 1–17. https://doi.org/10.1002/gdj3.136

Intergovernmental Oceanographic Commission Manuals and Guides, (2020) Quality control of in situ sea level observations: a review and progress towards automated quality control. volume 1, 71 pp. IOC/2020/MG/83Vol.1

Marcos, M., Puyol, B., Amores, A., Pérez Gómez, B., Fraile, M.Á. & Talke, S.A. (2021) Historical tide gauge sea-level observations in Alicante and Santander (Spain) since the 19th century. *Geoscience Data Journal*, 8, 144–153. https://doi.org/10.1002/gdj3.112

Martinez-Asensio, A., Wöppelmann, G., Ballu, V., Becker, M., Testut, L., Magnan, A.K. et al. (2019) Relative Sea-level rise and the influence of vertical land motion at tropical Pacific Islands. *Global and Planetary Change*, 176, 132–143. https://www.sciencedirect.com/science/article/pii/S0921818118306751

Miguez, B.M., Testut, L. & Wöppelmann, G. (2008) The Van de Casteele test revisited: an efficient approach to tide gauge error characterization. *Journal of Atmospheric and Oceanic Technology*, 25, 1238–1244. https://doi.org/10.1175/2007JTECHO554.1

Schueller, K. (2015) Theoretical basis for earth tide analysis with the new ETERNA34-ANA-V4.0 program. *Bulletin d'Information des Marées Terrestres*, 149, 12024–12061. http://www.bim-icet.org/

Schueller, K.. (2017) Theoretical basis for Earth tide analysis and prediction, Documentation ET34-X-Vmn. http://ggp.bkg.bund.de/eterna/.

Van Camp, M. & Vauterin, P. (2005) Tsoft: graphical and interactive software for the analysis of time series and earth tides. *Computers and Geosciences*, 31, 631–640. https://doi.org/10.1016/j.cageo.2004.11.015

Van der Stok, J. P. (1897) Wind and water, currents, tides and tidal streams in the east Indian archipelago, ISBN-13: OCLC:458366268.

Wenzel, H.-G. (1995) Format and structure for the exchange of high precision tidal data. *Bulletin d'Information des Marées Terrestres*, 121, 9097–9101. http://www.bim-icet.org/

Wenzel, H.-G. (1996) The nanogal software: earth tide data processing package ETERNA 3.30. *Bulletin d'Information des Marées Terrestres*, 124, 9425–9439. http://www.bim-icet.org/

APPENDIX A

A.1 | Database for sea-level monitoring

- There are two data sources: the University of French Polynesia and 'IOC, Sea Level Monitoring Facility' (www.ioc-sealevelmonitoring.org). The two sources were treated separately until the final analysis.
- There are two sensors: RADAR and pressure gauge. The measured values are different.
- There are 3 levels of data processing (Figure A.1):

- Level 1: raw data converted in PRETERNA format (Figure A.2)

(sampling rate 1 min or 2 min);
- Level 2: data corrected using Tsoft in ETERNA format (Figure A.3)

(sampling rate 1 min or 2 min);
- Level 3: data averaged to hourly values and analysis results (Figure A.4)
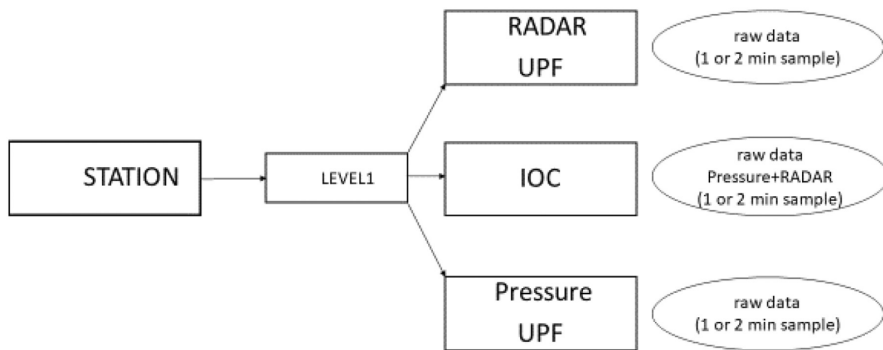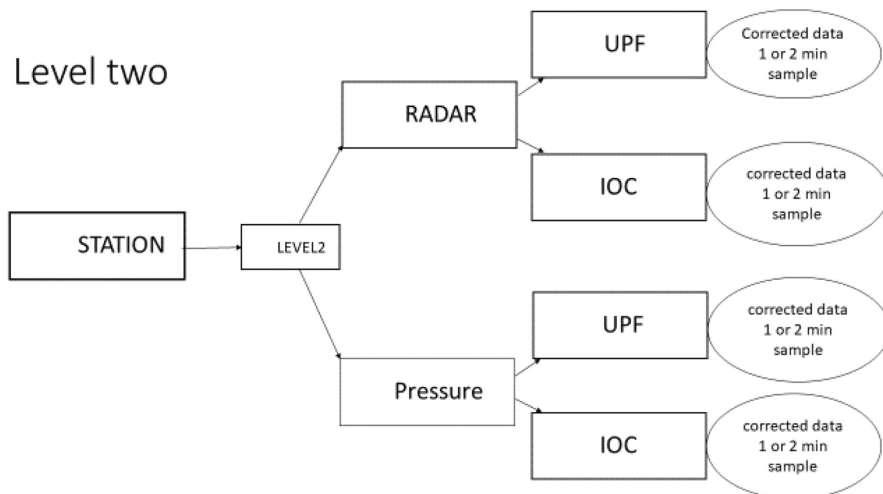
FIGURE A.1
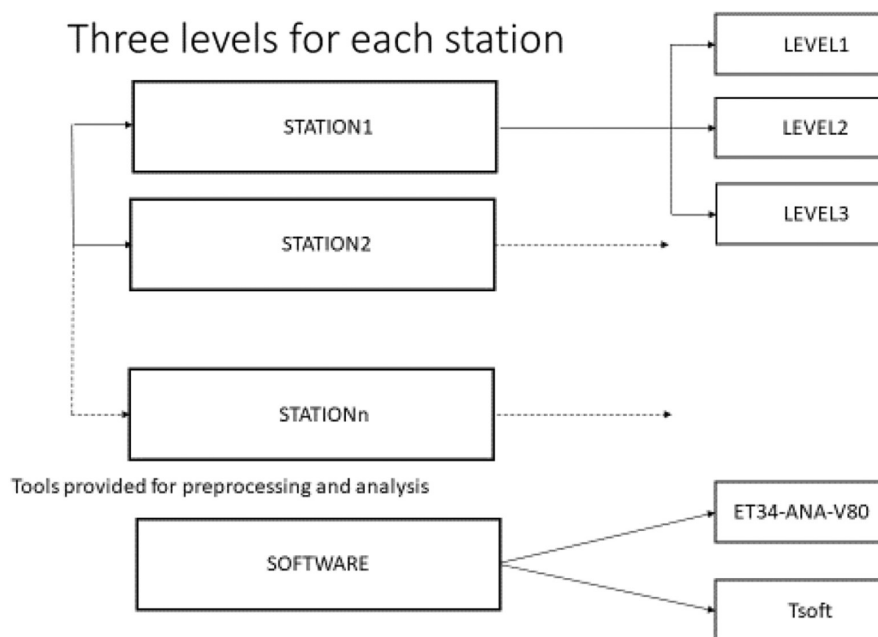
Level one

FIGURE A.2

Level two
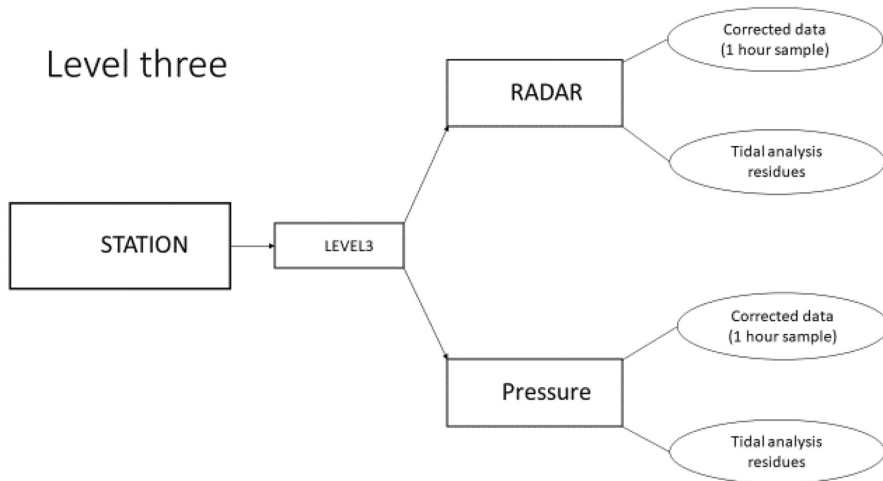

FIGURE A.3

Level three

FIGURE A.4