



**HAL**  
open science

## Correction and Accuracy of High- and Low-Resolution CTD Data from Animal-Borne Instruments

Lia Siegelman, Fabien Roquet, Vigan Mensah, Pascal Rivière, Etienne  
Pauthenet, Baptiste Picard, Christophe Guinet

► **To cite this version:**

Lia Siegelman, Fabien Roquet, Vigan Mensah, Pascal Rivière, Etienne Pauthenet, et al.. Correction and Accuracy of High- and Low-Resolution CTD Data from Animal-Borne Instruments. *Journal of Atmospheric and Oceanic Technology*, 2019, 36 (5), pp.745-760. 10.1175/JTECH-D-18-0170.1 . hal-02192465

**HAL Id: hal-02192465**

**<https://hal.science/hal-02192465>**

Submitted on 18 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Correction and Accuracy of High- and Low-Resolution CTD Data from Animal-Borne Instruments

LIA SIEGELMAN,<sup>a,b,c</sup> FABIEN ROQUET,<sup>d</sup> VIGAN MENSAH,<sup>e</sup> PASCAL RIVIÈRE,<sup>a</sup>  
ETIENNE PAUTHENET,<sup>f</sup> BAPTISTE PICARD,<sup>g</sup> AND CHRISTOPHE GUINET<sup>g</sup>

<sup>a</sup> Univ. Brest, CNRS, IRD, Ifremer, LEMAR, Plouzané, France

<sup>b</sup> Environmental Science and Engineering, California Institute of Technology, Pasadena, California

<sup>c</sup> NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

<sup>d</sup> Department of Marine Sciences, University of Gothenburg, Gothenburg, Sweden

<sup>e</sup> Institute of Low Temperature Science, Hokkaido University, Sapporo, Japan

<sup>f</sup> Sorbonne Universités, UPMC Univ., Paris 06, UMR 7159, LOCEAN-IPSL, F-75005, Paris, France

<sup>g</sup> Centre d'Études Biologiques de Chizé, CNRS, Villiers-en-Bois, France

(Manuscript received 28 September 2018, in final form 12 February 2019)

### ABSTRACT

Most available CTD Satellite Relay Data Logger (CTD-SRDL) profiles are heavily compressed before satellite transmission. High-resolution profiles recorded at the sampling frequency of 0.5 Hz are, however, available upon physical retrieval of the logger. Between 2014 and 2018, several loggers deployed on elephant seals in the Southern Ocean have been set in continuous recording mode, capturing both the ascent and descent for over 60 profiles per day during several months, opening new horizons for the physical oceanography community. Taking advantage of a new dataset made of seven such loggers, a postprocessing procedure is proposed and validated to improve the quality of all CTD-SRDL data: that is, both high-resolution profiles and compressed low-resolution ones. First, temperature and conductivity are corrected for a thermal mass effect. Then salinity spiking and density inversion are removed by adjusting salinity while leaving temperature unchanged. This method, applied here to more than 50 000 profiles, yields significant and systematic improvements in both temperature and salinity, particularly in regions of rapid temperature variation. The continuous high-resolution dataset is then used to provide updated accuracy estimates of CTD-SRDL data. For high-resolution data, accuracies are estimated to be of  $\pm 0.02^\circ\text{C}$  for temperature and  $\pm 0.03 \text{ g kg}^{-1}$  for salinity. For low-resolution data, transmitted data points have similar accuracies; however, reconstructed temperature profiles have a reduced accuracy associated with the vertical interpolation of  $\pm 0.04^\circ\text{C}$  and a nearly unchanged salinity accuracy of  $\pm 0.03 \text{ g kg}^{-1}$ .

### 1. Introduction

Hydrographic data collected by marine mammals such as southern elephant seals (SES; *Mirounga leonina*), Steller sea lions (*Eumetopias jubatus*), or ribbon seals (*Histiophoca fascia*) equipped with conductivity–temperature–depth Satellite Relay Data Logger (CTD-SRDL) tags (referred to as “tag”) represent a consequent

source of information, especially in the undersampled polar oceans where observations are notoriously hard to acquire (Roquet et al. 2014; Treasure et al. 2017). The temperature and conductivity sensors of these tags, manufactured by Valeport, Ltd., yield high precision ( $\pm 0.005^\circ\text{C}$  for temperature and  $\pm 0.01 \text{ ms cm}^{-1}$  for conductivity; see Boehme et al. 2009) and reasonable accuracies ( $\pm 0.02^\circ\text{C}$  for temperature and  $\pm 0.03 \text{ g kg}^{-1}$  for salinity) after delayed-mode calibration (Roquet et al. 2011). However, to be fitted on marine mammals, the tags are considerably smaller than a traditional CTD and often contain residual noise related to the miniaturized sensors capabilities (Roquet et al. 2017). Recent studies (Nakanowatari et al. 2017; Mensah et al. 2018) demonstrated that the tags are affected by a thermal mass error on their temperature and conductivity cells, particularly pronounced across sharp thermoclines.

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JTECH-D-18-0170.s1>.

Corresponding author: Lia Siegelman, [lsiegelman@caltech.edu](mailto:lsiegelman@caltech.edu)

DOI: 10.1175/JTECH-D-18-0170.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy ([www.ametsoc.org/PUBSReuseLicenses](http://www.ametsoc.org/PUBSReuseLicenses)).

Correction of the thermal mass-induced error have been proposed and shown to work (Nakanowatari et al. 2017; Mensah et al. 2018) but only for salinity and not for temperature. Mensah et al. (2018) demonstrated that the thermal mass error on salinity is of  $O(10^{-2})$   $\text{g kg}^{-1}$  with maxima of  $O(10^{-1})$   $\text{g kg}^{-1}$ . These authors developed a correction, based on Lueck and Picklo (1990), which directly corrects salinity rather than conductivity and yields an error decrease of the salinity measure of up to 50% when using generic parameters values. However, due to the lack of thermal mass correction for temperature, density errors, some of them imputable to temperature, remain.

In addition, some aspects of the accuracy estimates of these data are insufficiently characterized. Indeed, besides the calibration errors discussed in Roquet et al. (2011), errors associated with the dynamical response of the tags and the reconstruction of profiles after compression also exist but have not been estimated.

In the present paper, we take advantage of a newly available continuous high-resolution dataset (described in section 2c) that advantageously contains the ascending and descending phase of dive, allowing to compare upcasts and downcasts for the first time. This unique dataset is used to tackle all the issues mentioned above, that is, the implementation of a thermal lag correction for temperature and conductivity, incorporation of a salinity correction to remove density inversions, and estimation of the dynamic response error of the tags and of the compression error associated with low-resolution data, all of which contribute to improving the post-processing and accuracy estimates of CTD-SRDL data.

Section 2 presents the datasets and method employed in this study. Section 3 introduces the correction procedure, which includes preliminary salinity and temperature adjustments and the thermal mass correction for conductivity and temperature as well as the density inversion removal algorithm. Section 4 describes the parameter estimation for the thermal mass correction. Section 5 presents the correction and accuracy estimates for high-resolution data. Section 6 presents the correction and accuracy estimates for compressed data. A summary and conclusions are proposed in section 7.

## 2. Datasets and method

### a. Satellite-transmitted compressed profile data

Compressed low-resolution profiles are transmitted in near-real time via the Argos satellite system for  $\sim 4$  profiles per day (among the  $\sim 60$  dives per day performed by SES) and are typically made of  $\sim 15$  points chosen via a broken-stick point selection algorithm (Boehme et al. 2009). To date, the majority of available

biologged data is made of low-resolution profiles. They represent 85% of the profiles and 95% of the tags in the Marine Mammals Exploring the Oceans Pole to Pole (MEOP) database, which is the largest database of biologged acquired hydrographic data with over 500 000 profiles collected since 2004 (Treasure et al. 2017). In particular, the heavy subsampling associated with the satellite-transmission method of low-resolution data implies the existence of a compression loss that will be evaluated for the first time in section 6c.

The low-resolution dataset employed in this study is made of 43 tags deployed on SES around the Kerguelen Islands from February 2011 to January 2018 (trajectories in black and red in Fig. 1a), amounting to a total of 14 762 CTD profiles. The environment in which SES dived is marked by temperatures of  $0^{\circ}$ – $4^{\circ}\text{C}$ , salinities of  $34$ – $34.9$   $\text{g kg}^{-1}$ , and densities of  $26.9$ – $27.8$   $\text{kg m}^{-3}$ , characteristic of the areas north of the polar front (PF), in the Antarctic zone (AAZ), which is defined as the area located between the PF and the southern Antarctic Circumpolar Current Front (SACCF) and south of the SACCF (Fig. 1).

### b. High-resolution profile data from retrieved tags

High-resolution profiles are recorded at the sampling frequency of 0.5 Hz for  $\sim 10$  profiles per day and are typically made of  $\sim 1000$  points. However, these data are only available after physical retrieval of the tag due to limited battery capabilities. Thus, compressed profiles are a subsample of high-resolution ones. The high-resolution dataset employed in this study corresponds to the high-resolution version of the low-resolution dataset, that is, the same 14 762 dives (trajectories in black and red in Fig. 1a) but recorded at the sampling frequency of 0.5 Hz. The high-resolution dataset is used to validate the proposed correction scheme for compressed low-resolution data and to derive the compression error (see section 6).

### c. Continuous high-resolution data

Thanks to recent progress in the field of miniaturization and satellite telemetry, tags can now record temperature, salinity, and pressure at the sampling frequency of 0.5 Hz for every dive of the seal's journey ( $>60$  dives per day), giving access, for the first time, to both the ascending and descending phase of a dive. Such tags can last for more than 3 months and cover large oceanic domains at a very high horizontal and vertical resolution (see sections of temperature and salinity in Fig. 2). These state-of-the-art data are opening new horizons for the physical oceanography community and in particular for the study of small-scale processes of  $O(1$ – $50)$  km, also called submesoscales. These submesoscale features,

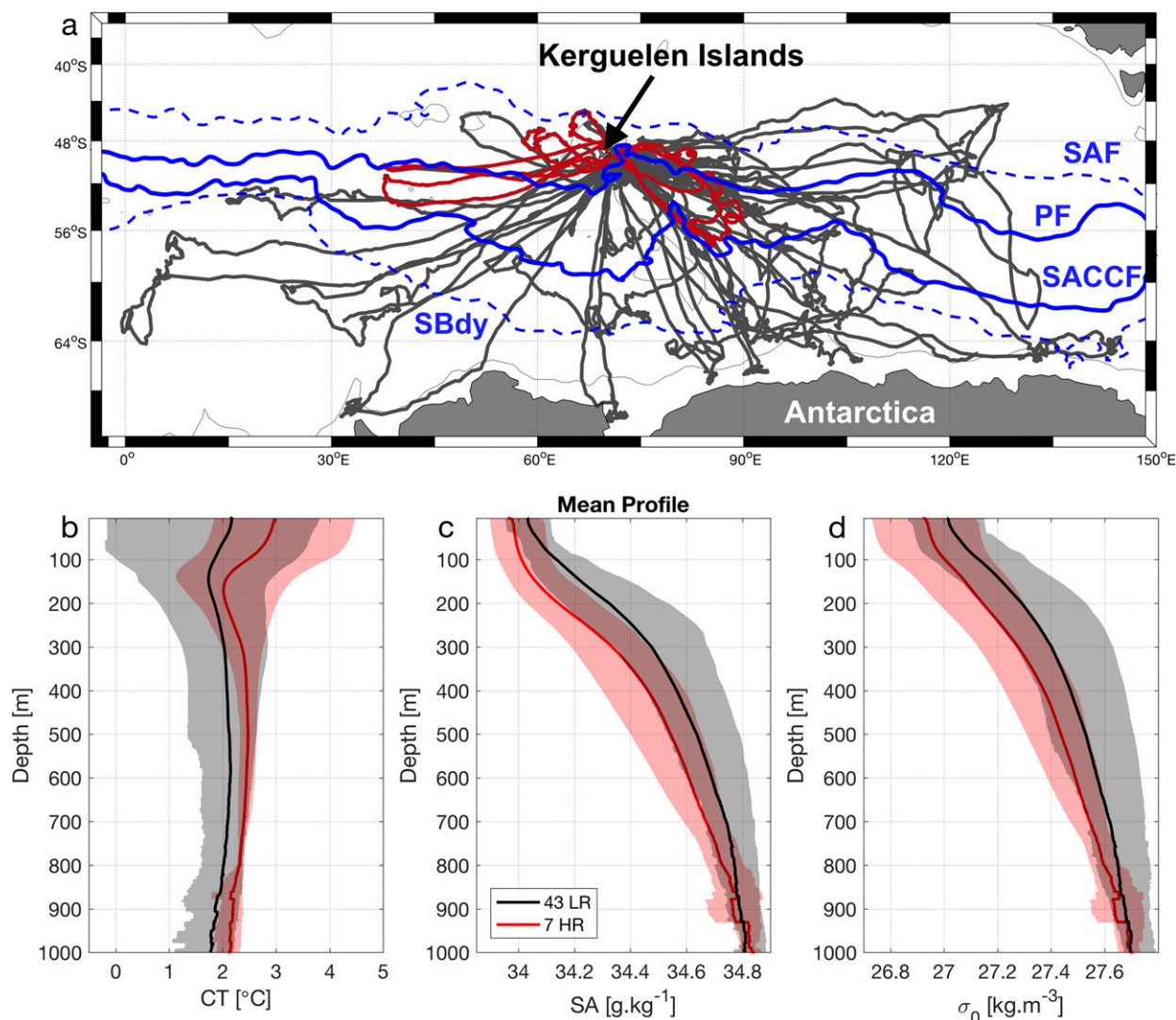


FIG. 1. (a) Spatial distribution of profiles collected by the 43 low-resolution tags (LR; black) among which 7 are also at continuous high resolution (HR; red), deployed from 2011 to 2018 on SES on the Kerguelen Islands. Climatological position of the Subantarctic Front (SAF), Polar Front (PF), Southern Antarctic Circumpolar Current Front (SACCF), and Southern Boundary Front (SBdy) are indicated in blue according to Kim and Orsi (2014). (b) Mean conservative temperature (CT) profile of the 43 LR tags (black) and 7 HR tags (red). (c) As in (b), but for absolute salinity (SA). (d) As in (b), but for potential density  $\sigma_0$ . In (b)–(d) the 20th and 80th percentiles are shaded.

characterized by elongated fronts and filamentary structures, are now captured by such observations, contrary to previous data obtained from animal-borne CTDs that were either of low resolution or only of high vertical (but not horizontal) resolution. To date, such data have only been briefly introduced in Mensah et al. (2018) to validate their salinity-based correction. In this paper, we fully exploit the potential of these new data, referred to as the “continuous high-resolution dataset,” that allows for direct comparison of upcasts and downcasts prior to and after correction.

The continuous high-resolution dataset employed in this study is made of seven tags (39 183 profiles) deployed

on SES around the Kerguelen Islands in the Indian Ocean sector of the Southern Ocean from October 2014 to January 2018 (trajectories in red on Fig. 1a). The environment in which the SES conducted profiles is marked by temperatures of 1°–5°C, salinities of 33.9–34.9  $\text{g kg}^{-1}$ , and densities of 26.7–27.7  $\text{kg m}^{-3}$ , characteristic of the area north of the SACCF (see Figs. 1 and 2).

#### d. General methodology

In addition to the temperature and salinity adjustments introduced in Roquet et al. (2011) and briefly recalled in section 3a, we propose an autonomous two-step

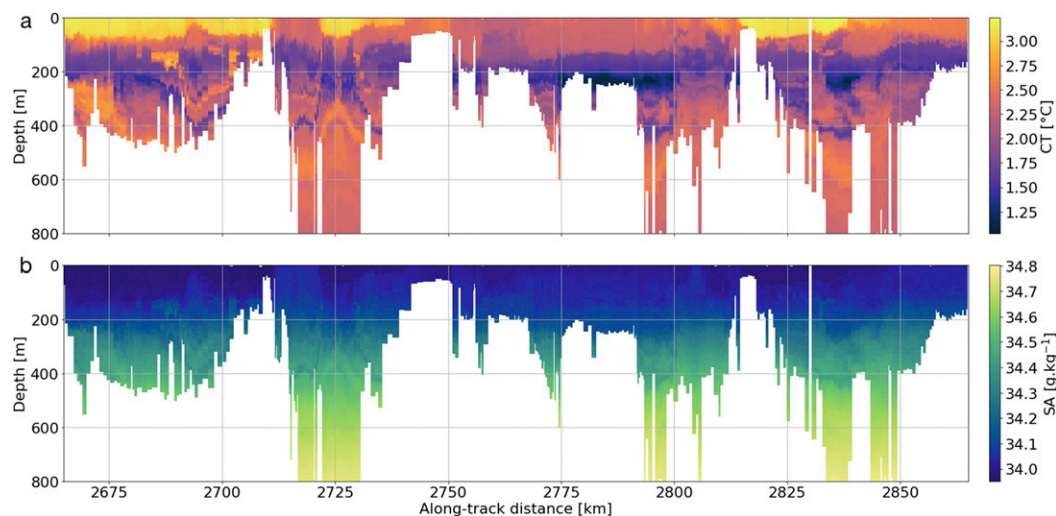


FIG. 2. Example of 200-km-long section of (a) conservative temperature and (b) absolute salinity at high resolution in the Southern Ocean from tag ct112–050–14 (red southwest trajectory from Kerguelen Islands in Fig. 1a). The tag recorded information at the sampling frequency of 0.5 Hz from 28 Oct 2014 to 21 Jan 2015, amounting to 6942 dives, or over 80 dives per day, and for a cumulative distance of 5665 km.

postprocessing procedure for the correction of MEOP's profiles also applicable to other hydrographic data acquired via biologging or with unpumped glider. We first implement a correction scheme for the thermal mass effect where temperature (Morison et al. 1994) and conductivity are simultaneously corrected (Lueck and Picklo 1990) instead of only and directly correcting salinity as is done in Mensah et al. (2018). This was, however, made possible by the theoretical framework delineated in Mensah et al. (2018) (see sections 3b and 4). We then aim at further refining the salinity correction regarding density inversions, which has not been addressed by previous works and is of particular importance for high-resolution profiles. To do so, hydrographic profiles are stabilized within the water column by minimally adjusting absolute salinity (SA) while leaving the values of in situ temperature unchanged with the method developed by Barker and McDougall (2017) (see section 3c).

The procedure is applied to the continuous high-resolution dataset as well as to the low-resolution dataset introduced in section 2. For high-resolution data, the effect of the correction is documented globally and then separately for the areas north and south of the PF (Fig. 1). The correction is then validated without relying on external data. Indeed, we use the continuous high-resolution dataset that advantageously contains both the ascending and descending phase of a dive, to validate the proposed correction by comparing upcast and downcast data before and after correction. Finally, the continuous high-resolution dataset is used to estimate the mean errors associated

with the tag's dynamical response, which are also estimated globally and separately north and south of the PF (see section 5).

For low-resolution data, the effect of the correction is documented for three oceanic zones: north of the PF, in the AAZ, and south of the SACCF (Fig. 1). The correction is then validated by comparing corrected high-resolution profiles to their low-resolution version prior to and after correction. Finally, the compression error associated with low-resolution data is estimated for the first time by comparing corrected low-resolution profiles to their high-resolution counterparts. The compression error is derived globally and then separately for the areas north of the PF, in the AAZ, and south of the SACCF (Fig. 1; see section 6).

Estimation of both the dynamic and compression errors enables us to provide updated accuracy estimates for high- and low-resolution data, which should be of prime importance to the physical oceanography community.

### 3. Correction procedure

#### a. Step 1: Salinity and temperature adjustments

In this section, we briefly recall the delayed-mode calibration designed by Roquet et al. (2011) and applied to the high- and low-resolution datasets before the correction scheme developed in this study (steps 2 and 3; Fig. 3). Temperature and salinity are first corrected for pressure-induced linear biases by comparing CTD-SRDL data to ship-based CTD measurements. An external field effect on the conductivity sensor is also detected, inducing an additional salinity offset.

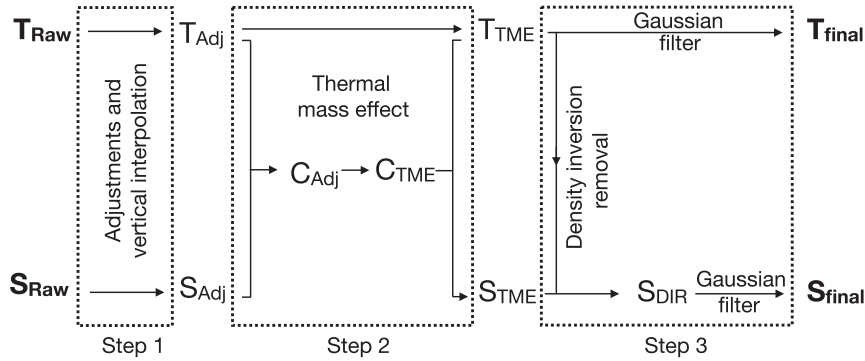


FIG. 3. Schematic summarizing the correction procedure implemented in the MEOP database. From the raw CTD data, temperature  $T_{Raw}$  and salinity  $S_{Raw}$  are first adjusted and vertically interpolated onto a regular grid according to the procedure described in Roquet et al. (2011) (step 1 detailed in section 3a). Then, adjusted conductivity  $C_{Adj}$  and temperature  $T_{Adj}$  are corrected for the thermal mass error (step 2 detailed in section 3b). Density inversions are then removed by minimally adjusting  $S_{TME}$  while leaving  $T_{TME}$  unchanged. Finally, a Gaussian filter with a 1-dbar window is applied to  $T_{TME}$  and  $S_{DIR}$  (step 3 detailed in section 3c).

This salinity offset is corrected with delayed-mode methods. The first method uses the stable salinity maximum characterizing the Lower Circumpolar Deep Water sampled by the seals foraging south of the SACCF and where this approach is not possible, a statistical method of cross comparison of CTD-SRDL measurements is used. For more details on the delayed-mode calibration methods, we refer the reader to Roquet et al. (2011). Finally, temperature and salinity are vertically interpolated onto a regular grid of 1-m resolution.

Note that the two-step correction procedure proposed in this paper can be implemented on hydrographic data acquired via biologging and unpumped glider without requiring step 1, which was specifically designed for the tags manufactured by Valeport, Ltd.

### b. Step 2: Thermal mass error

The thermal mass error (TME) is due to the transfer of heat from the sensor's walls to the sample being measured. This well-documented phenomenon (Lueck 1990) affects the conductivity cells of CTDs, which leads to a significant error in the estimation of salinity. This error is a function of the temperature gradient in space and time and has been widely documented for sharp thermocline (Lueck and Picklo 1990; Morison et al. 1994; Mensah et al. 2009; Liu et al. 2015). The tags considered in this study are particularly affected by the TME on their conductivity cells since they are unpumped. In addition, it appears that an error in temperature also related to the TME exists (Mensah et al. 2018). Building on Mensah et al. (2018), who only and directly corrected salinity, we implement a correction scheme for the TME that simultaneously corrects conductivity and temperature.

To correct the conductivity measure, we apply the procedure proposed by Lueck and Picklo (1990). The thermal mass error is modeled as an initial error  $\alpha_C$ , decaying within a relaxation time of  $1/\beta$  (Lueck 1990). Conductivity is then corrected via

$$C_T(n) = \Gamma_C \alpha_C (1 - 0.5\beta\Delta_t)^{-1} T_{hp}(n), \quad (1)$$

with  $C_T$  the conductivity correction added to the conductivity  $n$ th sample,  $T_{hp}(n)$  the high-pass-filtered sample's temperature, using a first-order discrete-time filter with a time constant  $\tau = \beta^{-1} - 0.5\Delta_t$ ,  $n$  the sample's index,  $\Gamma_C = (\partial C/\partial T)|_{S,p}$  the coefficient of sensitivity of conductivity to temperature at fixed salinity, and pressure and  $\Delta_t$  the sampling time interval, with  $\Delta_t = 2$  s for CTD-SRDLs. Note that Eq. (1), expressed as in Mensah et al. (2018), is equivalent to Lueck and Picklo's (1990) original formulation. Salinity is subsequently derived from the corrected conductivity.

Following Morison et al. (1994), temperature is corrected for the TME with a similar scheme according to

$$T_T(n) = \alpha_T (1 - 0.5\beta\Delta_t)^{-1} T_{hp}(n), \quad (2)$$

where the only difference with Eq. (1) is the absence of the sensitivity coefficient  $\Gamma_C$ . This second step is summarized in the schematic on Fig. 3. Adequate values for the three parameters  $\alpha_T$ ,  $\alpha_C$ , and  $\beta$  are estimated in section 4.

### c. Step 3: Density-inversion removal

Density inversions are commonly contained in oceanographic observations and are often due to instrumental noise and phenomena such as salinity spiking. While the

TME correction (step 1) suppresses part of this noise, most of the salinity spikes and density inversions persist after its application. [Barker and McDougall \(2017\)](#) propose a method to stabilize hydrographic profiles that removes part of this residual noise. The procedure minimally adjusts SA while leaving the values of conservative temperature (CT) unchanged. This density inversion removal (DIR) procedure adjusts the profiles such that they never exceed a minimum  $N^2$  threshold ( $N$  is the Brunt–Väisälä frequency), set here, and by default, to  $1 \times 10^{-9} \text{ s}^{-2}$ . Then a Gaussian filter with a 1-dbar window is applied to CT and SA to remove sharp and localized jumps induced by the DIR for SA profiles. These jumps occur at a frequency higher than the original sampling frequency and thus are not the reflection of a physical process. Furthermore, since experimental noise is observed in CT and in SA, which both contribute to the density estimate, smoothing CT in addition to SA ensures potential density  $\sigma_0$  free of density inversion or equivalently maintains a positive  $N^2$ . Indeed,  $N^2$  profiles can exhibit negative values when the filter is only applied to SA but not when it is applied to both SA and CT (not shown). The Gaussian filter also has the advantage to account for the irregular sampling in depth of the raw data that is dependent on the variable vertical speed of the animal. Last, the 1-dbar window does not alter the vertical resolution of the data since the original sampling frequency is of 0.5 Hz and the animal ascending speed is between 1 and  $2 \text{ m s}^{-1}$  ([Hindell and Lea 1998](#)). This third step is summarized in the schematic on [Fig. 3](#).

#### 4. Parameter estimation for the thermal mass error correction

To achieve an equivalent final salinity correction as the one obtained in [Mensah et al. \(2018\)](#), we use their generic parameters values of  $\beta = 0.06 \text{ s}^{-1}$  and  $\alpha_S = 0.04$ . Since [Mensah et al.'s \(2018\)](#) correction is based on the assumption that the correction of salinity is a linear combination of temperature and conductivity corrections, the initial error coefficients  $\alpha_C$ ,  $\alpha_T$ , and  $\alpha_S$  are related by  $\alpha_C = \alpha_T - \alpha_S$ , where  $\alpha_S$  is the initial error for salinity. The relaxation time  $1/\beta$  is assumed to be identical for both temperature and conductivity (thus also for salinity).

To derive an optimal  $\alpha_T$ , we use the continuous high-resolution dataset mentioned in [section 2](#). The temperature correction delineated by [Eq. \(2\)](#) is applied to the ascending and descending phases of every dive ranging from depths of a minimum of 150 m up to 1000 m and occurring within a 10-min interval in order to compare similar water masses, adding up to a total of 32 154 dives.

Differences in temperature between ascent and descent are not only due to the thermal mass effect; changes in the timing and geographic location between both phases actually account for a large part of this difference. However, the temperature difference between ascent and descent caused by natural variability should have no average bias. Consequently, the temperature correction scheme delineated by [Eq. \(2\)](#) is tested through a least squares regression scheme, in which we look for the optimal  $\alpha_T$  that minimizes the mean bias in temperature  $\sum_{z=0}^{400\text{m}} B(\alpha_T, z)$  between the ascending phase of dive  $i$  and the descending phase of dive  $i + 1$ , where

$$B(\alpha_T, z) = \frac{1}{n-1} \sum_{i=1}^{n-1} [T_a^{\alpha_T}(i, z) - T_d^{\alpha_T}(i+1, z)], \quad (3)$$

with  $z \in [0, 400]$  the depth of the measurements (m),  $i$  the  $i$ th dive,  $n$  the number of dives per tag, and  $T_a^{\alpha_T}$  and  $T_d^{\alpha_T}$  the temperature during ascent and descent, respectively, after application of the temperature correction with the coefficient  $\alpha_T$ . Note that below 400 m,  $T_{\text{hp}}$  is close to 0 such that the impact of the correction becomes negligible ([Fig. 4](#)), which justifies the choice of 400 m in  $\sum_{z=0}^{400\text{m}} B(\alpha_T, z)$ . Note also that we do not use the same dive for comparing ascending and descending phases but instead compare the ascending phase of dive  $i$  to the descending phase of dive  $i + 1$  because SES tend to station at the surface for only a few minutes ( $\sim 2$  min) while they swim underwater for longer periods of time ( $\sim 25$  min). It enables us to minimize the natural variability between ascents and descents.

The test is carried out separately for each of the seven tags, with resulting  $\alpha_T$  ranging from 0.07 to 0.10 (see [Table S1](#) in the online supplemental material). An optimal coefficient  $\alpha_T$  equal to 0.09 is then defined as the mean of the seven  $\alpha_T$  obtained for each tag. This optimal  $\alpha_T$  coefficient is used to implement the TME correction for temperature. A value of  $\alpha_C = \alpha_T - \alpha_S = 0.09 - 0.04 = 0.05$  is therefore used in combination with  $\beta = 0.06 \text{ s}^{-1}$ , which produces a salinity correction nearly identical to the one of [Mensah et al. \(2018\)](#). [Figure 4](#) presents the mean  $T_{\text{hp}}$  and temperature bias averaged for the seven tags. Noticeably, the mean temperature bias before correction is inversely proportional to the mean  $T_{\text{hp\_ascent}} - T_{\text{hp\_descent}}$  (black line), which confirms the appropriateness of the TME correction for the temperature field. Indeed, before correction, the mean temperature bias exhibits a significant positive (negative) bias between 0 and 200 m (200 and 400 m) (black line in [Fig. 4b](#)). However, after correction, the mean bias is centered around zero (blue line in [Fig. 4b](#)), indicating the suppression of a systematic bias throughout the water column. More importantly,

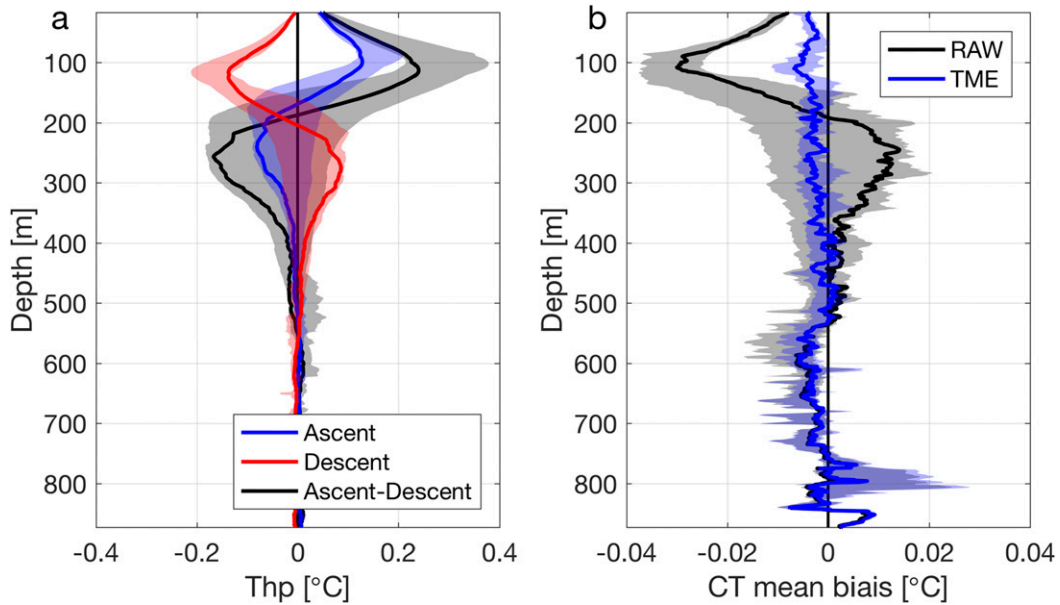


FIG. 4. (a) Mean high-pass filter of temperature  $T_{hp}$  during the ascent (blue line) and descent (red line) and mean  $T_{hp,ascent} - T_{hp,descent}$  (black line) for the 7 HR tags of the continuous high-resolution dataset used to implement the thermal mass error correction. (b) Mean bias prior to correction (black line) and after correction (blue line) for the 7 HR tags of the continuous high-resolution dataset. The 20th and 80th percentiles are shaded.

between 0 and 400 m, where the impact of the correction is the greatest due to the presence of the strongest  $T_{hp}$ , the mean temperature bias is reduced by 75% with an absolute value of  $0.01^{\circ}\text{C}$  before correction and only  $0.003^{\circ}\text{C}$  after correction, which is below the sensor's precision of  $\pm 0.005^{\circ}\text{C}$  (Boehme et al. 2009). Besides the averaged impact of the TME correction for temperature, Fig. 5a presents a typical ascent and descent profile prior to and after correction. The temperature difference between ascent and descent is reduced by  $0.08^{\circ}\text{C}$  at 112 m, a rather considerable improvement. In other words, application of the TME correction to the temperature field successfully removes the mean bias related to the thermal mass effect that is even maintained within the sensor's precision of  $\pm 0.005^{\circ}\text{C}$ .

## 5. Correction and accuracy of high-resolution profiles

### a. Effect of the correction scheme

Here we only consider the ascending phase of a dive since it is the one retained in the final MEOP database. We use the same continuous high-resolution dataset as in section 4 but with no restrictive criteria for depth or time. As a result, the two-step postprocessing correction is now implemented for the 39 183 high-resolution profiles of the continuous high-resolution dataset.

A typical pre- and postcorrection profile at high resolution is presented on the top panels of Fig. 6. The

TME (blue curve) first adjusts CT, SA, and  $\sigma_0$  on a low-frequency scale of  $O(100)$  m. Then salinity spikes and subsequent density inversions occurring at a scale of  $O(1)$  m are removed by the DIR (red curve).

Differences between raw and corrected data are computed at both steps and RMS values are presented on Fig. 7. Overall, the largest corrections take place between the surface and 300 m, that is, where stronger gradients of temperature are located (Fig. 4). As expected, the TME has a much larger contribution to the CT correction, whereas both TME and DIR contribute significantly to the correction of SA. Indeed, during the DIR the only effective correction applied to CT is the Gaussian filter. For CT, the root-mean-square (RMS) between raw and corrected data is  $5.9 \times 10^{-3}^{\circ}\text{C}$  after application of the TME and only  $6.1 \times 10^{-3}^{\circ}\text{C}$  after application of the additional DIR step (Fig. 7a). On the other hand, for SA, the RMS between raw and corrected data is of  $2.8 \times 10^{-3} \text{ g kg}^{-1}$  after the TME step and  $5.3 \times 10^{-3} \text{ g kg}^{-1}$  after the DIR. Likewise, for  $\sigma_0$ , the RMS is of 2.7 and  $4.5 \times 10^{-3} \text{ kg m}^{-3}$  after the TME and DIR, respectively. These results highlight the equally important impact of each step for the correction of SA and  $\sigma_0$  (Figs. 6 and 7). Furthermore, the order of the steps is important and applying the TME prior to the DIR considerably reduces the amplitude of the DIR (not shown here). In other words, the DIR has a lesser impact on the profiles when applied after the TME than directly on raw data, which points out the contribution



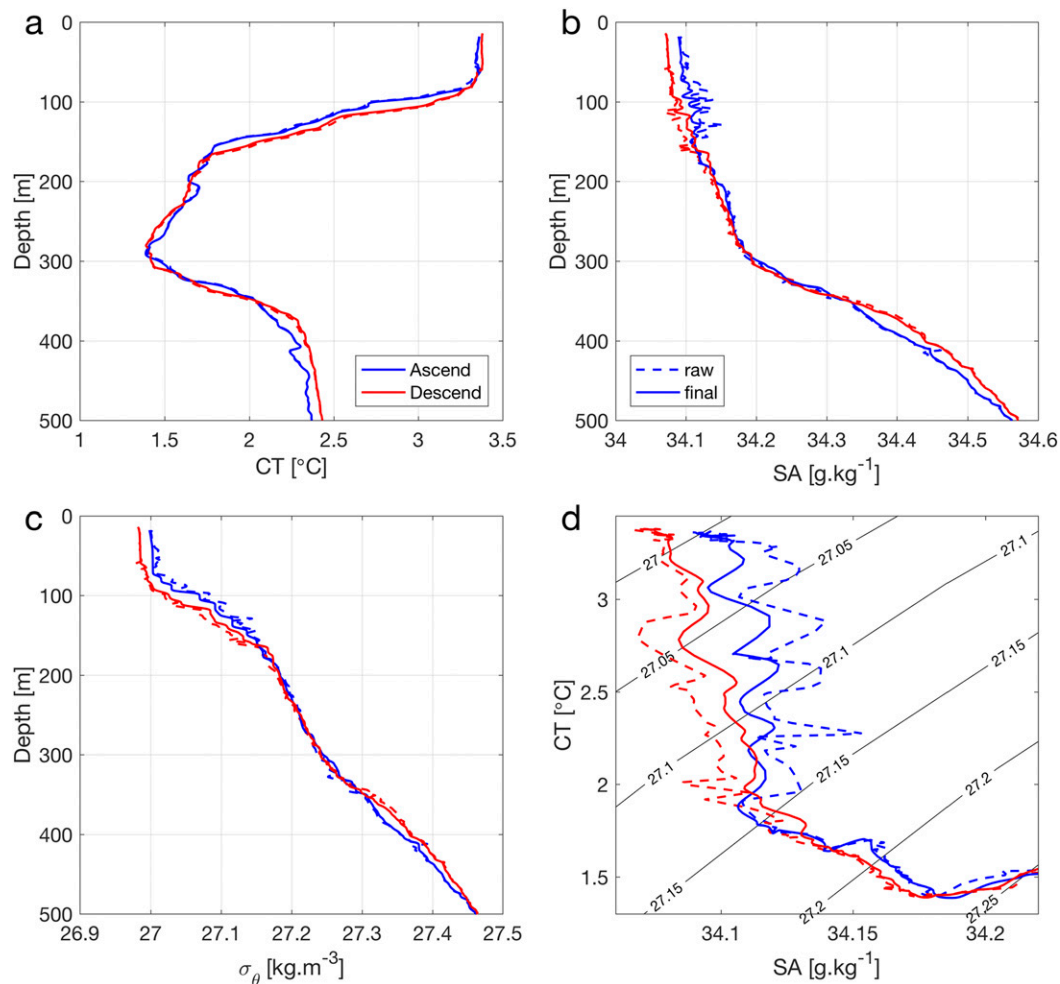


FIG. 5. Typical ascent (blue line) and descent (red line) profiles prior to correction (dashed line) and after correction (solid line), zoomed from 0 to 500 m. (a) Conservative temperature. (b) Absolute salinity. (c) Potential density. (d) CT–SA diagram.

of the thermal mass effect correction for reducing density inversions. The mean RMS between raw and corrected data is  $1.35 \times 10^{-2} \text{ }^\circ\text{C}$  for CT and  $1.17 \times 10^{-2} \text{ g kg}^{-1}$  for SA within the upper 300 m. These values are largely above the sensor's precision, and underscore the relative importance of the correction.

The impact of the correction for high-resolution profiles is now documented for two zones of the Southern Ocean: north and south of the PF, following the climatological definition of Kim and Orsi (2014). Unfortunately, the continuous high-resolution dataset used in this study does not contain enough profiles south of the SACCF (191 profiles or <0.05% of the dataset) to compute separate statistics for the AAZ and the area south of the SACCF. North of the PF, the profiles are marked by a mean temperature and salinity of  $2.50^\circ\text{C}$  and of  $34.42 \text{ g kg}^{-1}$ , while south of the PF, the means are of  $2.01^\circ\text{C}$  for temperature and  $34.57 \text{ g kg}^{-1}$  for salinity

(Table 1). RMSs between raw and corrected data are presented per area and depth range in Table 1. While below 300 m, the impact of correction is equivalent north and south of the PF, it differs between 0 and 300 m, where the correction is stronger by  $\sim 20\%$  south of the PF for both CT, SA, and  $\sigma_0$ . For CT, the mean RMS is  $1.4 \times 10^{-2} \text{ }^\circ\text{C}$  north of the PF and  $1.6 \times 10^{-2} \text{ }^\circ\text{C}$  south of the PF. For SA, the mean RMS is  $1.1 \times 10^{-2} \text{ g kg}^{-1}$  north of the PF and  $1.3 \times 10^{-2} \text{ g kg}^{-1}$  south of the PF. For  $\sigma_0$ , it is  $9 \times 10^{-3} \text{ kg m}^{-3}$  north of the PF and  $1.1 \times 10^{-2} \text{ kg m}^{-3}$  south of it (Table 1).

#### b. Validation of the correction scheme for high-resolution profiles

Following the same methodology as in section 4, the two-step correction is applied to the ascending and descending phases of the 39 183 dives that are deeper than 150 m and occur within a 10-min interval, which

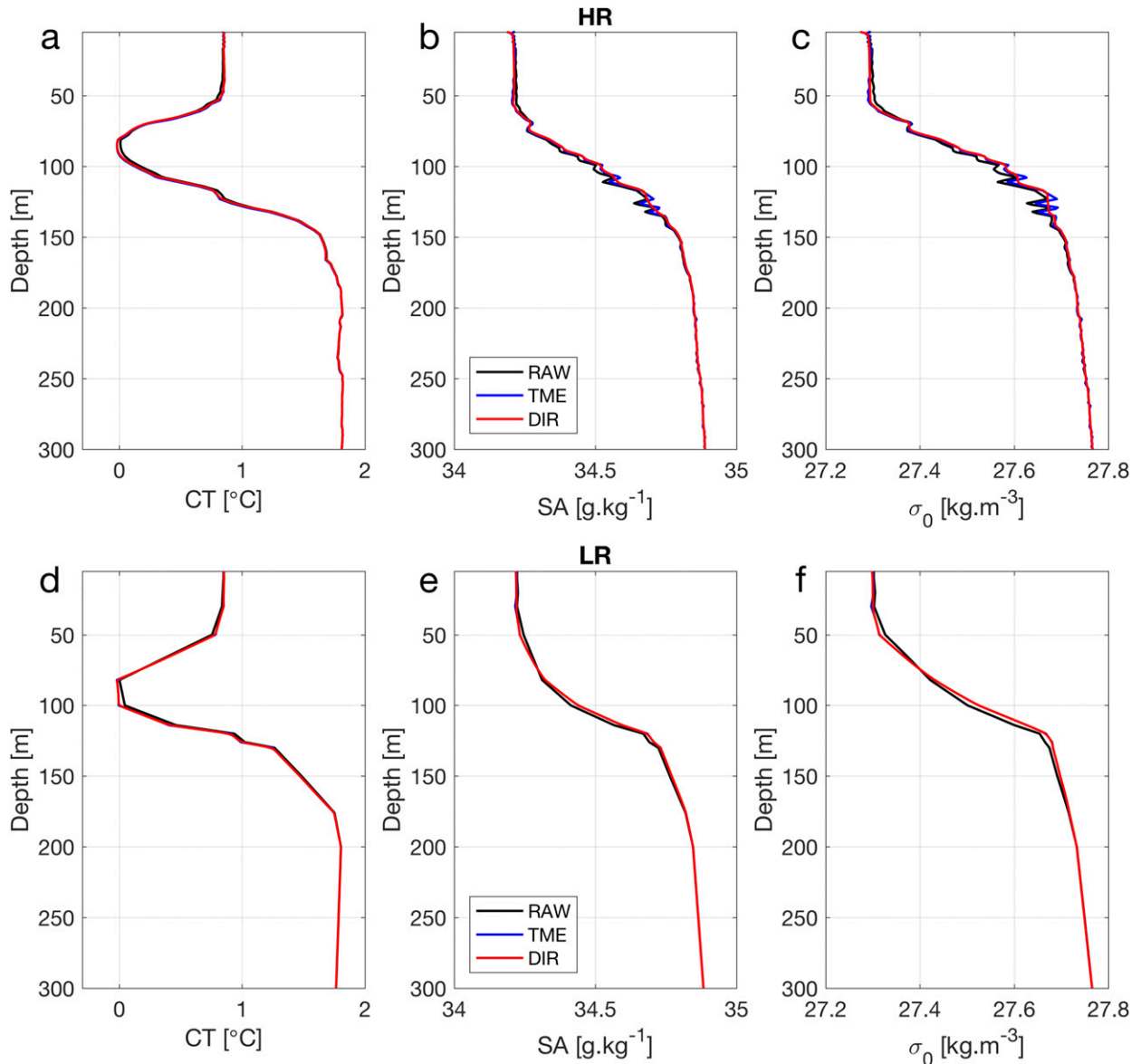


FIG. 6. Typical profile undergoing the correction scheme. High-resolution profiles of (a)–(c) conservative temperature, absolute salinity, and potential density, respectively. (d)–(f) As in (a)–(c), respectively, but showing low-resolution profiles.

amounts to 32 154 dives. Comparison between the ascending phase of dive  $i$  and the descending phase of dive  $i + 1$  is done before correction as well as at both steps of the procedure. For each tag, the RMS error of ascend versus descend is calculated at each depth  $z$  according to

$$\text{RMS}(z) = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n-1} [X_a(z, i) - X_d(z, i+1)]^2}, \quad (4)$$

with  $n$  the total number of dives per tag,  $X_a$  and  $X_d$  the field (CT, SA, or  $\rho_0$ ) during the ascent and descent,

respectively. This RMS is calculated for the uncorrected ( $\text{RMS}_{\text{raw}}$ ) and corrected ( $\text{RMS}_{\text{cor}}$ ) CT, SA, and  $\rho_0$  at the TME and DIR steps. The performance of the correction is then evaluated through the difference  $D_{\text{cor}} = \text{RMS}_{\text{raw}} - \text{RMS}_{\text{cor}}$  at both steps, with positive values indicating a decrease of the RMS error due to a convergence of ascending and descending profiles post correction. All seven tags see a continuous and significant improvement at each step of the procedure for all three fields. The difference between the ascending and descending phases for both CT, SA, and  $\sigma_0$  is reduced after correction, as can be inferred from

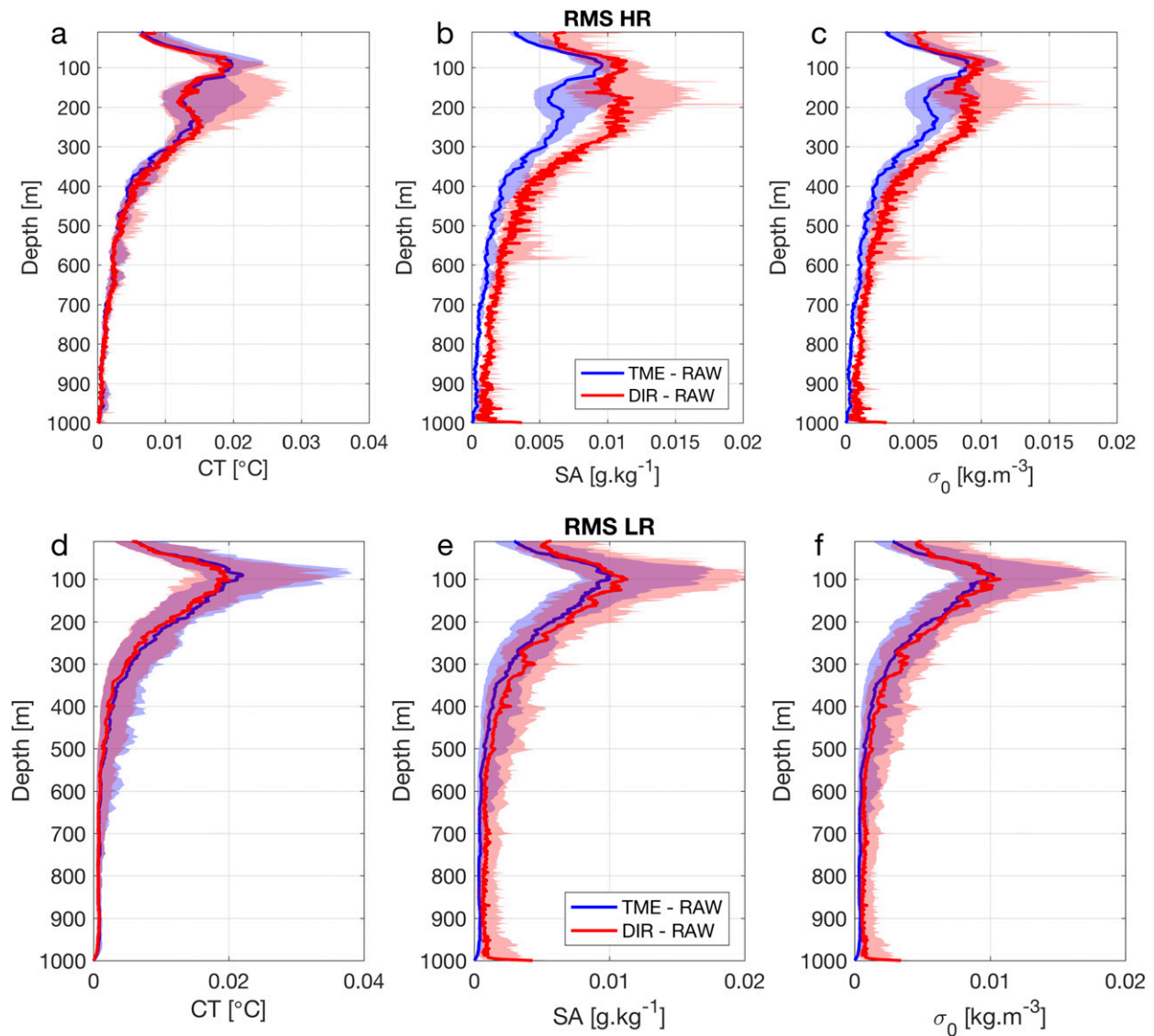


FIG. 7. Mean RMS between raw and corrected data for (top) the 7 HR tags for (a) CT, (b) SA, and (c)  $\sigma_0$  and (bottom) the 43 LR tags for (d) CT, (e) SA, and (f)  $\sigma_0$ . The 20th and 80th percentiles are shaded.

the positive  $D_{\text{cor}}$  presented on the upper panels of Fig. 8, even though between 200 and 400 m the difference is slightly negative for CT (but inferior to the sensor's precision). The TME reduces most the difference between ascents and descents for CT with mean  $|D_{\text{cor}}|$  over the water column of  $1.3 \times 10^{-3} \text{ } ^\circ\text{C}$  at both steps. For SA and  $\sigma_0$ , the two steps incrementally contribute to reduce the difference between ascents and descents with a mean  $|D_{\text{cor}}|$  of 1.5 and  $3.0 \times 10^{-3} \text{ g kg}^{-1}$  for SA at step TME and DIR, respectively, and 1.4 and  $2.9 \times 10^{-3} \text{ kg m}^{-3}$  for  $\sigma_0$  at step TME and DIR, respectively.

Overall, the error difference  $D_{\text{cor}}$  is most reduced in areas of strong temperature gradients, that is, between the surface and 300-m depth (Figs. 4 and 8), which itself

coincides with the strongest impact of the correction procedure (Fig. 7), emphasizing the procedure's efficiency. Besides the effect of the correction scheme presented on Fig. 8 over the entire water column, a typical pre- and postcorrection profile, between the surface and 300 m, is shown on Fig. 5. The profile, which occurs in a region presenting sharp temperature variations and thus particularly prone to thermal mass and salinity spiking effects, sees its temperature difference between ascent and descent reduced by  $8 \times 10^{-2} \text{ } ^\circ\text{C}$  at 112 m and its salinity by  $4 \times 10^{-2} \text{ g kg}^{-1}$  at 130 m, two rather considerable improvements which are also reflected on the density profile (Fig. 5c) and in the TS plan (Fig. 5d).

TABLE 1. Statistics relative to high-resolution (HR) data per oceanic zone and depth range averaged for the seven tags of the continuous HR dataset used in this study. Mean CT, SA, and RMS between raw and final profiles are presented, as is the mean error bias defined by Eq. (5), for which 14 904 profiles are used north of the PF and 17 250 south of the PF, adding to a total of 32 154 profiles. The location of the PF is defined according to Kim and Orsi (2014).

| Zone        | No. of HR profiles | Depth (m) | CT (°C) | SA (g kg <sup>-1</sup> ) | RMS of CT (°C) | RMS of SA (g kg <sup>-1</sup> ) | RMS of $\sigma_0$ (kg m <sup>-3</sup> ) | CT mean bias (°C) | SA mean bias (g kg <sup>-1</sup> ) | $\sigma_0$ mean bias (kg m <sup>-3</sup> ) |
|-------------|--------------------|-----------|---------|--------------------------|----------------|---------------------------------|-----------------------------------------|-------------------|------------------------------------|--------------------------------------------|
| North of PF | 17 592             | 0–300     | 2.64    | 34.04                    | 0.014          | 0.011                           | 0.009                                   | 0.004             | 0.005                              | 0.005                                      |
|             |                    | 300–1000  | 2.45    | 34.58                    | 0.003          | 0.003                           | 0.002                                   | 0.003             | 0.004                              | 0.004                                      |
|             |                    | 0–1000    | 2.50    | 34.42                    | 0.006          | 0.006                           | 0.004                                   | 0.003             | 0.005                              | 0.004                                      |
| South of PF | 21 591             | 0–300     | 1.66    | 34.20                    | 0.016          | 0.013                           | 0.011                                   | 0.003             | 0.004                              | 0.004                                      |
|             |                    | 300–1000  | 2.17    | 34.73                    | 0.003          | 0.002                           | 0.002                                   | 0.001             | 0.002                              | 0.002                                      |
|             |                    | 0–1000    | 2.01    | 34.57                    | 0.007          | 0.007                           | 0.005                                   | 0.002             | 0.003                              | 0.003                                      |
| All         | 39 183             | 0–300     | 2.38    | 34.10                    | 0.014          | 0.012                           | 0.010                                   | 0.003             | 0.005                              | 0.004                                      |
|             |                    | 300–1000  | 2.36    | 34.63                    | 0.003          | 0.003                           | 0.002                                   | 0.002             | 0.004                              | 0.003                                      |
|             |                    | 0–1000    | 2.36    | 34.47                    | 0.006          | 0.006                           | 0.004                                   | 0.003             | 0.004                              | 0.003                                      |

### c. Accuracy estimates

Besides calibration errors discussed in Roquet et al. (2011) and estimated to be of  $\pm 0.02^\circ\text{C}$  for temperature and  $\pm 0.03 \text{ g kg}^{-1}$  for salinity after delayed-mode calibration, the tags are also subject to a thermal lag response as documented in this study as well as in previous works (Nakanowatari et al. 2017; Mensah et al. 2018). Even after correction, a residual error imputable to thermal lag effects remains. In this section, we quantify this dynamical bias present in CTD-SRDL data, and also detail it north and south of the PF.

The mean dynamical error is estimated from the continuous high-resolution dataset through the calculation of  $B(\alpha_T, z)$  defined by Eq. (3) after application of the two-step correction. Results are presented for the entire water column on the top panels of Fig. 9 and summary statistics per zone can be found in Table 1, in which the mean dynamical error over the water column is defined by

$$M_{z_{\max}} = \frac{1}{z_{\max}} \sum_{z=0}^{z_{\max}} |B(\alpha_T, z)|, \quad (5)$$

where  $z_{\max}$  is the maximal depth, here, 1000 m. For both CT and SA, the mean dynamical bias  $M_{1000}$  is of  $2.7 \times 10^{-3}^\circ\text{C}$  and  $4.1 \times 10^{-3} \text{ g kg}^{-1}$ , respectively, which is well below the sensor's accuracies of  $\pm 0.005^\circ\text{C}$  for temperature and  $\pm 0.02 \text{ g kg}^{-1}$  for salinity. The  $\sigma_0$  also has a weak mean bias with a value of  $3.5 \times 10^{-3} \text{ kg m}^{-3}$ . For all three quantities  $M_{1000}$  is larger north than south of the PF by 75% for CT and 50% for SA and  $\sigma_0$ . Furthermore, across all areas, the mean dynamical bias is larger in the first 300 m by 45% for CT and 25% for SA and  $\sigma_0$ .

In conclusion, application of the TME and DIR considerably improves the final accuracy of CTD-SRDL data, with mean dynamical errors estimated to be below

the sensor's precision of  $\pm 0.005^\circ\text{C}$  for temperature and  $\pm 0.02 \text{ g kg}^{-1}$  for salinity. This means that after application of the correction procedure, dynamical errors are negligible compared to calibration errors such that the accuracies of  $\pm 0.02^\circ\text{C}$  for temperature and  $\pm 0.03 \text{ g kg}^{-1}$  for salinity presented in Roquet et al. (2011) are still valid for high-resolution data.

## 6. Correction and accuracy of compressed profiles

### a. Effect of the correction scheme

The low-resolution version of the high-resolution profile shown on Fig. 6a is presented on the bottom panels of the same figure. Overall, the correction procedure behaves similarly but with a few noticeable differences. Similarly to the high-resolution profile, CT is primarily affected by the TME. However, for SA and  $\sigma_0$ , only the effect of the TME is large while that of DIR has become insignificant. As mentioned before, the TME acts on scales of  $O(100)$  m while the DIR tends to act on scales of  $O(1)$  m. The strong (weak) impact of the TME (DIR) is thus consistent with the low resolution of the profile that does not have enough data points to generate salinity spiking. In addition, the Gaussian filter has a negligible impact for all three quantities (not shown), which is also coherent with the low resolution of the profile.

These observations can be generalized to the entire dataset and the application of the correction procedure to the 14 762 low-resolution profiles yields similar results as for high-resolution ones (Fig. 7, lower panels). The shape of the correction's impact within the water column is preserved, with the strongest corrections occurring once again between the surface and 300-m depth. However, as mentioned before, the TME now has the main impact for CT, SA, and  $\sigma_0$  with a mean

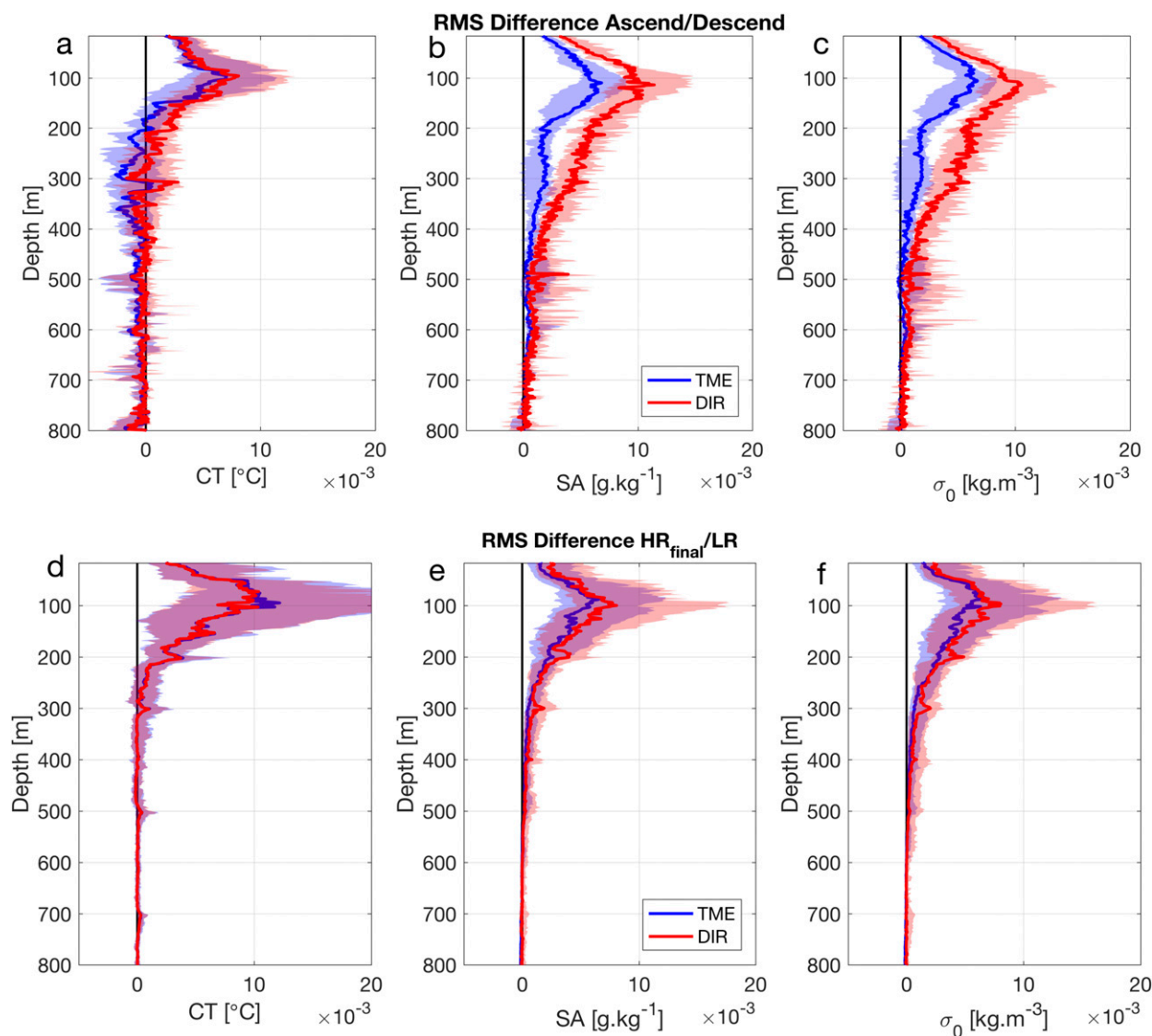


FIG. 8. Mean RMS difference between ascending and descending phases of a dive at each step of the correction procedure for (top) the 7 HR tags for (a) CT, (b) SA, and (c)  $\sigma_0$  and (bottom) the 43 LR tags for (d) CT, (e) SA, and (f)  $\sigma_0$ . The 20th and 80th percentiles are shaded.

RMS between raw and corrected data of  $4.9$  and  $4.5 \times 10^{-3} \text{C}$  for CT at step TME and DIR, respectively,  $2.3$  and  $3.0 \times 10^{-3} \text{g kg}^{-1}$  for SA, and  $2.3$  and  $2.7 \times 10^{-3} \text{kg m}^{-3}$  for  $\sigma_0$ . Once again, the correction is noticeably stronger between 0 and 300 m with average RMS between raw and final data of  $11.9 \times 10^{-3} \text{C}$  for CT,  $7.1 \times 10^{-3} \text{g kg}^{-1}$  for SA, and  $6.6 \times 10^{-3} \text{kg m}^{-3}$  for  $\sigma_0$ .

As for the high-resolution profiles, a regional analysis is conducted on the low-resolution dataset. Since a large amount of profiles are located south of the SACCF (5438 profiles or 37% of the dataset), three zones are now analyzed: the area north of the PF, the AAZ, and the area south of the SACCF, where the fronts are also defined according to Kim and Orsi (2014). Mean temperatures

are  $2.77^\circ$ ,  $2.17^\circ$ , and  $0.84^\circ \text{C}$  and mean salinities  $34.43$ ,  $34.52$ , and  $34.62 \text{g kg}^{-1}$  for the areas north of the PF, in the AAZ, and south of the SACCF, respectively. The temperature correction is equivalent in all three zones regardless of the depth range. For SA and  $\sigma_0$ , the correction is equivalent in the three zones below 300 m but in the upper 300 m it is more important south of the PF (in both the AAZ and the area south of the SACCF), as for to the correction impact on high-resolution data. For SA, the mean RMS is  $0.005 \text{g kg}^{-1}$  north of the PF and  $0.007$  and  $0.008 \text{g kg}^{-1}$  in the AAZ and south of the SACCF, respectively. For  $\sigma_0$ , it is  $0.005 \text{kg m}^{-3}$  north of the PF and  $0.007 \text{kg m}^{-3}$  south of it (Table 2).

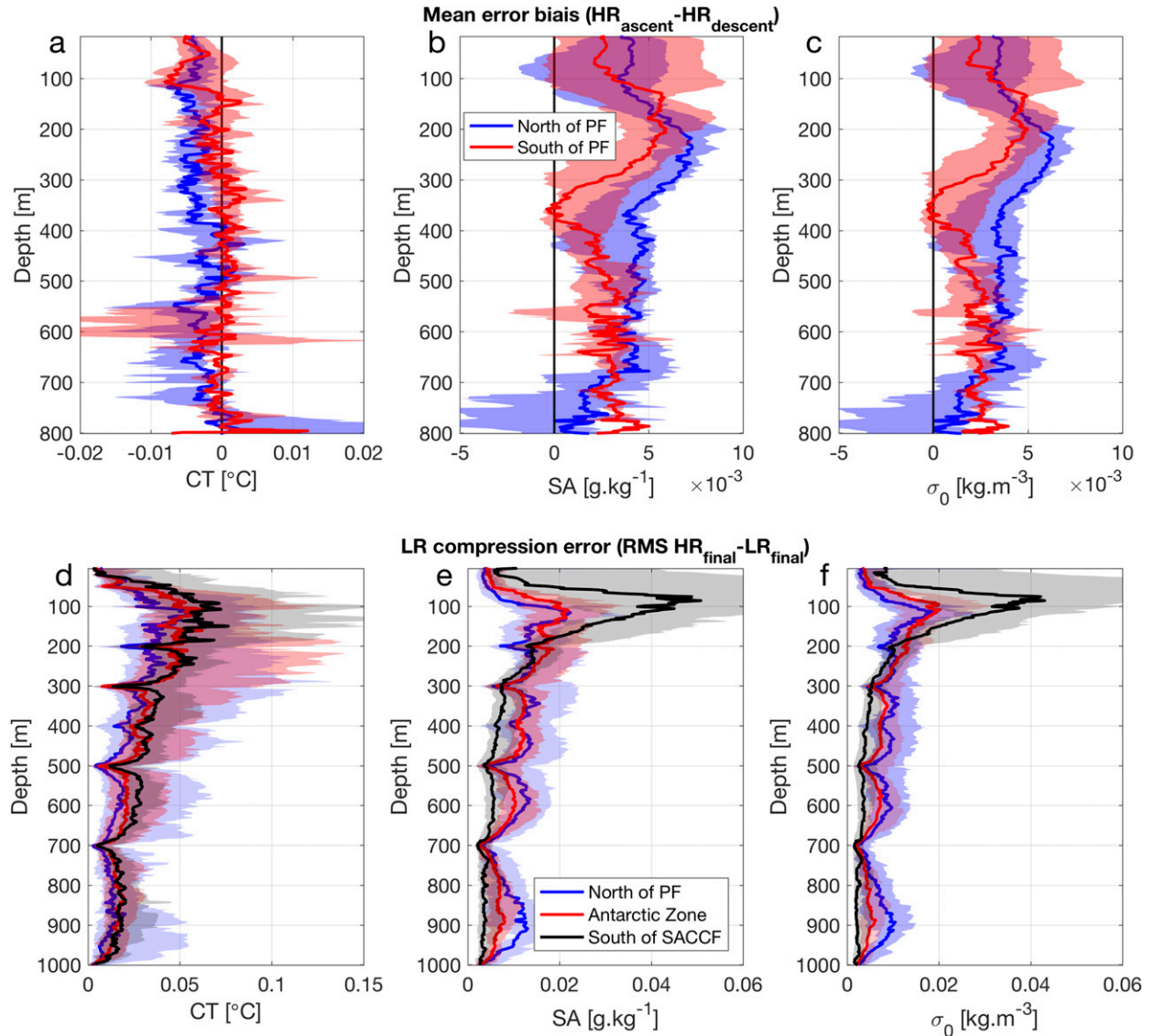


FIG. 9. (top) Mean error bias of high-resolution profiles related to the dynamical response of the tag. The mean error bias is calculated as the mean difference between the corrected ascents and descents [see Eq. (3)] for (a) CT, (b) SA, and (c)  $\rho_0$  averaged for the 7 HR tags of the continuous high-resolution dataset employed in this study and separated by frontal zones according to Kim and Orsi (2014). (bottom) Compression error associated with the low-resolution dataset. The compression error is calculated as the mean RMS difference between the corrected high- and low-resolution profiles for (d) CT, (e) SA, and (f)  $\rho_0$  averaged for the 43 tags and separated by oceanic zones according to Kim and Orsi (2014). The 20th and 80th percentiles are shaded.

### b. Validation of the correction scheme for compressed profiles

To validate the correction for low-resolution profiles, we undertake a comparison between the 14 762 low-resolution profiles at each step of the correction procedure (raw, TME, DIR) and their available corrected high-resolution version. Similar to the validation of high-resolution data, the RMS error of high versus low resolution is calculated, for each tag, at each depth  $z$  according to

$$\text{RMS}(z) = \sqrt{\frac{1}{n} \sum_{i=1}^n [X_{\text{HR}_{\text{final}}}(z, i) - X_{\text{LR}}(z, i)]^2}, \quad (6)$$

with  $n$  the total number of dives,  $X_{\text{HR}_{\text{final}}}$  the high-resolution corrected field (CT, SA, or  $\rho_0$ ), and  $X_{\text{LR}}$  the low-resolution field (CT, SA, or  $\rho_0$ ). Note that in Eq. (6), the high-resolution profiles are those obtained after application of the correction scheme and referred to as  $X_{\text{HR}_{\text{final}}}$ . RMS is calculated for the uncorrected ( $\text{RMS}_{\text{raw}}$ ) and corrected ( $\text{RMS}_{\text{cor}}$ ) low-resolution profiles of CT,

TABLE 2. Statistics relative to low-resolution (LR) data per oceanic zone and depth range averaged for the 43 tags at low resolution considered in this study. Mean CT, SA, and RMS between raw and final profiles are presented, as is mean LR compression error defined by Eq. (6). The frontal zones are defined according to Kim and Orsi (2014).

| Zone           | No. of LR profiles | Depth (m) | CT (°C) | SA (g kg <sup>-1</sup> ) | RMS of CT (°C) | RMS of SA (g kg <sup>-1</sup> ) | RMS of $\sigma_0$ (kg m <sup>-3</sup> ) | CT compression error (°C) | SA compression error (g kg <sup>-1</sup> ) | $\sigma_0$ compression error (kg m <sup>-3</sup> ) |
|----------------|--------------------|-----------|---------|--------------------------|----------------|---------------------------------|-----------------------------------------|---------------------------|--------------------------------------------|----------------------------------------------------|
| North of PF    | 3333               | 0–300     | 3.05    | 34.09                    | 0.012          | 0.005                           | 0.005                                   | 0.031                     | 0.011                                      | 0.009                                              |
|                |                    | 300–1000  | 2.65    | 34.58                    | 0.001          | 0.001                           | 0.001                                   | 0.015                     | 0.010                                      | 0.007                                              |
|                |                    | 0–1000    | 2.77    | 34.43                    | 0.003          | 0.002                           | 0.002                                   | 0.020                     | 0.010                                      | 0.008                                              |
| Antarctic Zone | 5991               | 0–300     | 2.04    | 34.14                    | 0.013          | 0.007                           | 0.007                                   | 0.040                     | 0.014                                      | 0.011                                              |
|                |                    | 300–1000  | 2.22    | 34.68                    | 0.001          | 0.001                           | 0.001                                   | 0.019                     | 0.008                                      | 0.005                                              |
|                |                    | 0–1000    | 2.17    | 34.52                    | 0.004          | 0.003                           | 0.002                                   | 0.025                     | 0.010                                      | 0.007                                              |
| South of SACCF | 5438               | 0–300     | 0.19    | 34.36                    | 0.012          | 0.008                           | 0.007                                   | 0.046                     | 0.021                                      | 0.017                                              |
|                |                    | 300–1000  | 1.12    | 34.73                    | 0.001          | 0.001                           | 0.001                                   | 0.022                     | 0.005                                      | 0.003                                              |
|                |                    | 0–1000    | 0.84    | 34.62                    | 0.004          | 0.003                           | 0.003                                   | 0.029                     | 0.010                                      | 0.007                                              |
| All            | 14 762             | 0–300     | 1.91    | 34.21                    | 0.012          | 0.007                           | 0.007                                   | 0.048                     | 0.015                                      | 0.012                                              |
|                |                    | 300–1000  | 2.06    | 34.66                    | 0.001          | 0.001                           | 0.001                                   | 0.025                     | 0.009                                      | 0.006                                              |
|                |                    | 0–1000    | 2.02    | 34.53                    | 0.004          | 0.003                           | 0.003                                   | 0.032                     | 0.011                                      | 0.008                                              |

SA, and  $\rho_0$  at both steps of the procedure. Similar to section 5b, the performance of the correction is then evaluated through the difference  $D_{\text{cor}} = \text{RMS}_{\text{raw}} - \text{RMS}_{\text{cor}}$  after application of the first step (TME) and second step (DIR). Positive  $D_{\text{cor}}$  indicates a decrease of the RMS error due to the convergence between the reference corrected high-resolution profile  $X_{\text{HR}_{\text{final}}}$  and the low-resolution profile undergoing the correction scheme. The results are presented on the lower panels of Fig. 8 for the averaged 43 tags and RMSs are of  $O(10^{-3})^\circ\text{C}$  for CT,  $O(10^{-3})\text{g kg}^{-1}$  for SA, and  $O(10^{-3})\text{kg m}^{-3}$  for  $\rho_0$ . Similarly to high-resolution profiles, there is a continuous decrease of the RMS error at each step of the procedure, validating the correction for low-resolution profiles. Once again, the strongest error decrease occurs in the first upper hundreds of meters of the water column. In contrast with high-resolution profiles, for which the first two steps have an equal impact, and in accordance with the correction impact on low-resolution profiles (see section 5a), it is now the TME that reduces most the RMS error. Indeed, for CT, the mean  $|D_{\text{cor}}|$  is  $1.0 \times 10^{-3}^\circ\text{C}$  after TME and  $1.3 \times 10^{-3}^\circ\text{C}$  after DIR. For SA, the mean  $|D_{\text{cor}}|$  is 0.9 and  $1.3 \times 10^{-3}\text{g kg}^{-1}$  after TME and DIR, respectively and for  $\sigma_0$ , it is 1.4 and  $1.3 \times 10^{-3}\text{kg m}^{-3}$  after TME and DIR, respectively.

### c. Compression error on low-resolution data

In this section, we estimate the compression error associated with the linear interpolation between transmitted data and quantify it for three major zones of the Southern Ocean: the area north of the PF, the AAZ, and the area south of the SACCF. The compression error is derived through the calculation of the RMS error between corrected high- and low-resolution profiles [see Eq. (6)]. The results are presented for the three zones on the lower panels of Fig. 9 and in Table 2. For CT, the compression error is important and has a mean value of  $\pm 0.032^\circ\text{C}$ , which is larger than the calibration error value of  $\pm 0.02^\circ\text{C}$  obtained after delayed-mode calibration (Roquet et al. 2011). This error increases from north to south with values of  $\pm 0.02^\circ\text{C}$  north of the PF,  $\pm 0.025^\circ\text{C}$  in the AAZ, and  $\pm 0.029^\circ\text{C}$  south of the SACCF. For SA, the compression error is overall relatively small, with a value of  $\pm 1.1 \times 10^{-2}\text{g kg}^{-1}$  across all areas. However, south of the SACCF, the compression error in salinity is important in the first 200 m, with a maximal error of  $\pm 5 \times 10^{-2}\text{g kg}^{-1}$  at 87 m. Furthermore, in the upper 300 m this compression error also increases from north to south with mean values of  $\pm 1.1 \times 10^{-2}\text{g kg}^{-1}$  north of the PF,  $\pm 1.4 \times 10^{-2}\text{g kg}^{-1}$  in the AZ, and  $\pm 2.1 \times 10^{-2}\text{g kg}^{-1}$  south of the SACCF. However, from 300 to 1000 m this tendency is reversed with mean compression errors of  $\pm 1 \times 10^{-2}\text{g kg}^{-1}$  north

of the PF,  $\pm 8 \times 10^{-3} \text{ g kg}^{-1}$  in the AZ, and  $\pm 5 \times 10^{-3} \text{ g kg}^{-1}$  south of the SACCF. Similar patterns are obtained for  $\sigma_0$ , with a mean compression error of  $\pm 0.008 \text{ kg m}^{-3}$  across all areas and a maximal error south of the SACCF of  $\pm 4 \times 10^{-2} \text{ kg m}^{-3}$  at 87 m. Note that the compression error profiles (Fig. 9) present lower values at fixed depths of 200, 300, 400, 500, 700, and 1000 m, which is due to the compression algorithm that picks these depth levels more often.

Overall, the compression error is important for CT but negligible for SA and  $\sigma_0$  with mean values of  $\pm 0.032^\circ\text{C}$ ,  $\pm 0.011 \text{ g kg}^{-1}$ , and  $\pm 0.008 \text{ kg m}^{-3}$ , respectively. Since the compression error can be assumed to be independent from the calibration error, new accuracy estimates for low-resolution data can be derived as  $\pm \sqrt{\text{Error}_{\text{compression}}^2 + \text{Error}_{\text{calibration}}^2}$ , which is equal to  $\pm 0.04^\circ\text{C}$  for temperature and  $\pm 0.03 \text{ g kg}^{-1}$  for salinity. These accuracy estimates can now be used to quantify the error associated with low-resolution data of the MEOP database ( $\sim 85\%$  of the profiles). However, if more precision is needed relative to a specific depth or oceanic zone, we encourage the users to refer to the compression error profiles of Fig. 9 as well as to Table 2.

## 7. Summary and conclusions

In this paper, we took advantage of the new generation of data available at the sampling frequency of 0.5 Hz to improve the quality of CTD-SDRL data. We provided evidence that the temperature sensors are affected by a thermal mass error. A correction scheme was therefore developed to improve temperature estimates by simultaneously correcting conductivity and temperature. This correction builds upon Mensah et al. (2018), who only and directly corrected salinity. In a second step, we refined salinity estimates by removing salinity spiking and density inversions according to Barker and McDougall (2017). This two-step postprocessing correction was implemented for two datasets collected by Kerguelen SES: a continuous high-resolution dataset made of 7 tags (39 183 profiles) deployed from 2014 to 2018 and a low-resolution dataset made of 43 tags (14 762 profiles) deployed from 2011 to 2018. The continuous high-resolution dataset was used to test the robustness and efficiency of the procedure for high-resolution data by comparing upcast and downcast data, which significantly converged after application of the correction. The correction scheme was then validated for low-resolution data by comparing corrected high- and low-resolution profiles, which showed similar improvements. To date, correction and validation of the profiles is mainly done by comparison with regional atlases of historical data and

ship-based CTDs (Roquet et al. 2011), thus relying on external observations acquired at a different date and location, and containing their own measure bias. A major advantage of this method is that it does not rely on external data and can be independently implemented. In addition to improving the tags' data quality, updated accuracy estimates were derived. Dynamical biases were found to be negligible in front of calibration errors such that the accuracy of high-resolution data is estimated to be  $\pm 0.02^\circ\text{C}$  for temperature and  $\pm 0.03 \text{ g kg}^{-1}$  for salinity. For the first time, the compression error associated with the interpolation of low-resolution data was quantified and found to be  $\pm 3.2 \times 10^{-2}^\circ\text{C}$  for temperature and  $\pm 1.1 \times 10^{-2} \text{ g kg}^{-1}$  for salinity (Table 2). This implies that the accuracies of low-resolution data are  $\pm 0.04^\circ\text{C}$  for temperature and  $\pm 0.03 \text{ g kg}^{-1}$  for salinity (section 6c). This analysis is of prime importance for users of the MEOP database, as it will enable them to carefully quantify the error associated with the MEOP data and ultimately refine their analyses.

Both datasets used in this study are located in the Southern Ocean. This suggests that the TME coefficients derived in section 3 may not be optimal for datasets located in other areas, in particular in regions marked by stronger thermoclines. Hence, we encourage users to determine their own optimal coefficient in the areas of deployment following the same methodology as the one presented in section 3 whenever possible. Nevertheless, our results have shown that the TME parameters provide satisfactory results for the Southern Ocean, which is where most of the MEOP data are located. In addition, this method should yield similar results for the polar areas of the Northern Hemisphere that are characterized by similar temperature gradients and where most of the remaining MEOP data are located.

Another potential drawback of this method is that it removes all density inversions within the water column yet they may occur on scales of order 1–5 m as the result of physical instabilities such as gravitational or Kelvin–Helmholtz's ones (Kundu and Cohen 1990). This is particularly the case during winter when these instabilities are known to be more frequent (Haine and Marshall 1998) and will become a growing concern with the advent of continuous high-resolution tags capable of detecting such density inversions. To refine the correction, an option would be to use, when available, accelerometry sensors deployed on marine mammals in conjunction with CTD tags to detect the animal's head movements, speed, and direction variations. These motions recorded by the accelerometer, and often linked to the apex predator's foraging effort (Viviant et al. 2010),



may cause the water to flow in a reverse direction within the unpumped cell of the tag. This may in turn generate thermal mass and salinity spiking effects. Being able to discriminate density inversions resulting from the animal's behavior and physical processes is key to perfect the correction procedure. Incorporation of such biological data in the correction of biologged profiles in conjunction with sensor accuracy improvements appears as the next step forward to refine the quality and accuracy of animal tag oceanographic data.

In conclusion, we have presented a new postprocessing method that yields conclusive results for the improvement of CTD-SDRL data. This technique may also be useful for unpumped glider data. Given the recent technological advances in biologging, continuous high-resolution dataset are expected to become increasingly available such that the approaches proposed here should have increasing utility.

*Acknowledgments.* The present study is a contribution to the Observing System—Mammals as Samplers of the Ocean Environment (SO-MEMO) with funding and logistic support from CNES-TOSCA, IPEV, and CORIOLIS in situ ocean observation program. The authors wish to thank Paul Barker and Trevor McDougall for the use of their correction scheme. L.S. is supported by a CNES and LabexMer doctoral fellowship.

#### REFERENCES

- Barker, P. M., and T. J. McDougall, 2017: Stabilizing hydrographic profiles with minimal change to the water masses. *J. Atmos. Oceanic Technol.*, **34**, 1935–1945, <https://doi.org/10.1175/JTECH-D-16-0111.1>.
- Boehme, L., P. Lovell, M. Biuw, F. Roquet, J. Nicholson, S. E. Thorpe, M. P. Meredith, and M. Fedak, 2009: Animal-borne CTD-satellite relay data loggers for real-time oceanographic data collection. *Ocean Sci.*, **5**, 685–695, <https://doi.org/10.5194/os-5-685-2009>.
- Haine, T. W., and J. Marshall, 1998: Gravitational, symmetric, and baroclinic instability of the ocean mixed layer. *J. Phys. Oceanogr.*, **28**, 634–658, [https://doi.org/10.1175/1520-0485\(1998\)028<0634:GSABIO>2.0.CO;2](https://doi.org/10.1175/1520-0485(1998)028<0634:GSABIO>2.0.CO;2).
- Hindell, M. A., and M.-A. Lea, 1998: Heart rate, swimming speed, and estimated oxygen consumption of a free-ranging southern elephant seal. *Physiol. Biochem. Zool.*, **71**, 74–84, <https://doi.org/10.1086/515890>.
- Kim, Y. S., and A. H. Orsi, 2014: On the variability of Antarctic Circumpolar Current fronts inferred from 1992–2011 altimetry. *J. Phys. Oceanogr.*, **44**, 3054–3071, <https://doi.org/10.1175/JPO-D-13-0217.1>.
- Kundu, P., and L. Cohen, 1990: *Fluid Mechanics*. Academic Press, 638 pp.
- Liu, Y., R. H. Weisberg, and C. Lembke, 2015: Glider salinity correction for unpumped CTD sensors across a sharp thermocline. *Coastal Ocean Observing Systems*, Elsevier, 305–325.
- Lueck, R. G., 1990: Thermal inertia of conductivity cells: Theory. *J. Atmos. Oceanic Technol.*, **7**, 741–755, [https://doi.org/10.1175/1520-0426\(1990\)007<0741:TIOCCT>2.0.CO;2](https://doi.org/10.1175/1520-0426(1990)007<0741:TIOCCT>2.0.CO;2).
- , and J. J. Picklo, 1990: Thermal inertia of conductivity cells: Observations with a sea-bird cell. *J. Atmos. Oceanic Technol.*, **7**, 756–768, [https://doi.org/10.1175/1520-0426\(1990\)007<0756:TIOCCO>2.0.CO;2](https://doi.org/10.1175/1520-0426(1990)007<0756:TIOCCO>2.0.CO;2).
- Mensah, V., M. Le Menn, and Y. Morel, 2009: Thermal mass correction for the evaluation of salinity. *J. Atmos. Oceanic Technol.*, **26**, 665–672, <https://doi.org/10.1175/2008JTECHO612.1>.
- , F. Roquet, L. Siegelman-Charbit, B. Picard, E. Pauthenet, and C. Guinet, 2018: A correction for the thermal mass-induced errors of CTD tags mounted on marine mammals. *J. Atmos. Oceanic Technol.*, **35**, 1237–1252, <https://doi.org/10.1175/JTECH-D-17-0141.1>.
- Morison, J., R. Andersen, N. Larson, E. D'Asaro, and T. Boyd, 1994: The correction for thermal-lag effects in sea-bird CTD data. *J. Atmos. Oceanic Technol.*, **11**, 1151–1164, [https://doi.org/10.1175/1520-0426\(1994\)011<1151:TCFTLE>2.0.CO;2](https://doi.org/10.1175/1520-0426(1994)011<1151:TCFTLE>2.0.CO;2).
- Nakanowatari, T., and Coauthors, 2017: Hydrographic observations by instrumented marine mammals in the Sea of Okhotsk. *Polar Sci.*, **13**, 56–65, <https://doi.org/10.1016/j.polar.2017.06.001>.
- Roquet, F., J.-B. Charrassin, S. Marchand, L. Boehme, M. Fedak, G. Reverdin, and C. Guinet, 2011: Delayed-mode calibration of hydrographic data obtained from animal-borne satellite relay data loggers. *J. Atmos. Oceanic Technol.*, **28**, 787–801, <https://doi.org/10.1175/2010JTECHO801.1>.
- , and Coauthors, 2014: A southern Indian Ocean database of hydrographic profiles obtained with instrumented elephant seals. *Sci. Data*, **1**, 140028, <https://doi.org/10.1038/sdata.2014.28>.
- , and Coauthors, 2017: Ocean observations using tagged animals. *Oceanography*, **30** (2), 139, <https://doi.org/10.5670/oceanog.2017.235>.
- Treasure, A. M., and Coauthors, 2017: Marine mammals exploring the oceans pole to pole: A review of the MEOP consortium. *Oceanography*, **30** (2), 132–138, <https://doi.org/10.5670/oceanog.2017.234>.
- Viviant, M., A. W. Trites, D. A. Rosen, P. Monestiez, and C. Guinet, 2010: Prey capture attempts can be detected in Steller sea lions and other marine predators using accelerometers. *Polar Biol.*, **33**, 713–719, <https://doi.org/10.1007/s00300-009-0750-y>.