

## Realizing the European Vertical Reference System using model-based hydrodynamic leveling data

Afrasteh, Y.; Slobbe, D. C.; Sacher, M.; Verlaan, M.; Klees, R.; Guarneri, H.; Keyzer, L.; Pietrzak, J.; Snellen, M.; More Authors

**DOI**

[10.1007/s00190-023-01778-2](https://doi.org/10.1007/s00190-023-01778-2)

**Publication date**

2023

**Document Version**

Final published version

**Published in**

Journal of Geodesy

**Citation (APA)**

Afrasteh, Y., Slobbe, D. C., Sacher, M., Verlaan, M., Klees, R., Guarneri, H., Keyzer, L., Pietrzak, J., Snellen, M., & More Authors (2023). Realizing the European Vertical Reference System using model-based hydrodynamic leveling data. *Journal of Geodesy*, 97(10), Article 86. <https://doi.org/10.1007/s00190-023-01778-2>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.



# Realizing the European Vertical Reference System using model-based hydrodynamic leveling data

Y. Afrasteh<sup>1</sup> · D. C. Slobbe<sup>1</sup> · M. Sacher<sup>3</sup> · M. Verlaan<sup>1,2</sup> · V. Jahanmard<sup>4</sup> · R. Klees<sup>1</sup> · H. Guarneri<sup>1</sup> · L. Keyzer<sup>1</sup> · J. Pietrzak<sup>1</sup> · M. Snellen<sup>1</sup> · F. Zijl<sup>2</sup>

Received: 9 December 2022 / Accepted: 6 September 2023  
© The Author(s) 2023

## Abstract

All realizations of the European Vertical Reference System (EVRS) computed so far are solely based on geopotential differences obtained by spirit leveling/gravimetry. As such, there are no direct connections between height benchmarks separated by large water bodies. In this study, such connections are added by means of model-based hydrodynamic leveling resulting in a new, yet unofficial realization of the EVRS. The model-derived mean water levels used in computing the hydrodynamic leveling connections were obtained from the Nemo-Nordic (Baltic Sea) and 3D DCSM-FM (northwest European continental shelf) hydrodynamic models. The impact of model-based hydrodynamic leveling on the European Vertical Reference Frame is significant, especially for France and Great Britain. Compared to a solution which only uses spirit leveling/gravimetry, the differences in these countries reach tens to hundreds of kgal mm. We also observed an improved agreement with normal heights obtained by differencing GNSS and the European gravimetric quasi-geoid 2015 (EGG2015) heights. In Great Britain, the south-north slope of  $48 \text{ mm deg}^{-1}$  present in the solution which uses only spirit leveling/gravimetry data reduced to  $2.2 \text{ mm deg}^{-1}$ . In France, the improvement is confined to the southwest. The choice of the period over which water levels are averaged has an impact on the results as it determines, among others, the set of tide gauges available to establish the hydrodynamic leveling connections. When using an averaging period that can be considered as the least preferred choice based on three established criteria, the positive impact for France has gone. For Great Britain, the estimated south-north slope became  $12.6 \text{ mm deg}^{-1}$ . This is larger than the slope obtained using the most preferred averaging period but still substantially lower compared to the slope associated with a solution that uses only spirit leveling/gravimetry.

**Keywords** Hydrodynamic leveling · Height system realization · Tide gauge

## 1 Introduction

All realizations of the European Vertical Reference System (EVRS) (Ihde et al. 2002, 2008) computed so far are solely based on data from the Unified European Leveling Network (UELN). The UELN dataset comprises geopotential differ-

ences between height benchmarks (HBMs) obtained by spirit leveling and gravimetry. The data are provided by all participating countries to the Federal Agency for Cartography and Geodesy (BKG) which is in charge of realizing the EVRS. As discussed in Afrasteh et al. (2021), the exclusive use of spirit leveling/gravimetry imposes limitations on the coverage of the European Vertical Reference Frame (EVRF). Indeed, spirit leveling cannot be used to cross large water bodies. Consequently, the EVRF does neither cover islands not connected to the European mainland leveling network through bridges or tunnels (e.g., Ireland), nor offshore platforms. In addition, the inability of spirit leveling to cross large water bodies reduces the strength of the leveling network in coastal countries. This, in turn, limits the precision and reliability of the computed realization. Apart from this, it should be noted that spirit leveling is prone to systematic errors (Vaníček et al. 1980). Leveling errors accumulate over long distances and

✉ Y. Afrasteh  
y.afrasteh@tudelft.nl

<sup>1</sup> Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands

<sup>2</sup> Deltares, Boussinesqweg 1, 2629 HV Delft, The Netherlands

<sup>3</sup> Federal Agency for Cartography and Geodesy, Richard-Strauss-Allee 11, 60598 Frankfurt am Main, Germany

<sup>4</sup> Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia

may introduce slopes in the height system realizations (e.g., Wang et al. 2012). As for the countries that participate in the UELN project, systematic errors are known to be present in the Belgian (Slobbe et al. 2019), British (Penna et al. 2013), and IGN69 French leveling networks (Sacher and Liebsch 2019).

One way to overcome the limitations in coverage and accuracy of the EVRF is offered by the so-called gravity field approach (Rülke et al. 2012), which has often been solved as a geodetic boundary value problem (e.g., Rummel and Teunissen 1988; Heck and Rummel 1990; Amos and Featherstone 2008; Gerlach and Rummel 2012; Amjadi-parvar et al. 2015; Sánchez and Sideris 2017). Rülke et al. (2012) applied the approach to determine the height reference frame offsets among the European national height systems. In doing so, GNSS heights and different quasi-geoid models were exploited while the EVRF2007 was used as a reference. The offsets were estimated with an accuracy of about 5 cm for most countries. Wu et al. (2018) assessed the potential of using clock networks in height datum unification (Bjerhammar 1985) by simulations using the EUVN (European Unified Vertical Network) (Ihde et al. 2000) as a prior. Although their results are very promising, no clock network is operational yet.

Afrasteh et al. (2021) proposed to overcome the limitations in coverage and accuracy of the EVRF imposed by the exclusive use of spirit leveling/gravimetry by combining the UELN dataset with model-based hydrodynamic leveling data. Model-based hydrodynamic leveling is introduced by Slobbe et al. (2018) as an efficient and flexible alternative method to connect islands and offshore platforms to the height system on land. The method, which is independent from any other method, uses a regional, high-resolution hydrodynamic model to derive mean water level (MWL) differences between tide gauges. By converting the MWL differences to geopotential differences and adding them to the geopotential differences between the HBMs and the MWLs at the tide gauges at both ends of the link, we establish a so-called hydrodynamic leveling connection. Note that the period over which the water level is averaged is arbitrary. Slobbe et al. (2018) showed, however, that higher accuracy can be obtained by choosing a multi-year period and by only averaging over the summer months (the latter avoids the storm surges period). Therefore we use throughout the paper the term summer mean water levels (SMWLs), which refers to the average water level calculated over all May to September months of one or more years.

Assuming the availability of a series of hydrodynamic models each covering a part of the European waters and providing uncorrelated SMWLs at the tide gauges with a uniform variance of  $4.5 \text{ cm}^2$ , Afrasteh et al. (2021) showed that combining UELN data with model-based hydrodynamic leveling data improves the median of the propagated height

standard deviations by 38% compared to a solution that uses only spirit leveling/gravimetry. A variance of  $4.5 \text{ cm}^2$  for the model-derived SMWL corresponds to a standard deviation of 3 cm for each hydrodynamic leveling connection (to establish a connection, we take the difference between two model-derived SMWLs). If the variance is  $12.5 \text{ cm}^2$  (corresponding to a standard deviation of 5 cm for each hydrodynamic leveling connection), the reported improvement is still 29%. They also found increased redundancy numbers for leveling observations close to the coastlines. Afrasteh et al. (2023) reassessed the potential impact of hydrodynamic leveling data on the quality of the EVRF using an empirical noise model for the model-derived SMWLs allowing to include error correlations. Their results suggest an improvement up to 25% if the water levels are averaged over three summer periods.

Both Afrasteh et al. (2021) and Afrasteh et al. (2023) assessed the impact of hydrodynamic leveling data on the quality of the EVRF through geodetic network analyzes, using only (assumed) stochastic information of the model-based hydrodynamic leveling data. This study goes further because it uses *real* model-based hydrodynamic leveling data between tide gauges in the Baltic Sea and the northwest European continental shelf, including the North Sea and Wadden Sea. The main objective is to demonstrate the impact of adding these data on the EVRF. The required model-derived SMWLs are obtained from the Nemo-Nordic hydrodynamic model (Hordoïr et al. 2015, 2019) and the 3D DCSM-FM (Zijl et al. 2020), which cover the Baltic Sea and the northwest European continental shelf, respectively. The water levels are averaged over three successive summer periods that are the same per region. The reason for using three summer periods is that this is the maximum averaging period for which a noise model (needed to build the full noise variance-covariance matrix of the dataset) is available (Afrasteh et al. 2023).

We find it important to mention that the impact is a *provisional* impact and that *we are not presenting a new official realization of the EVRS that replaces the latest release EVRF2019*. The first follows from the fact that our study is subject to the following limitations. First, in establishing the hydrodynamic leveling connections no attempt was made to connect the tide gauges to the UELN by means of real levelings. Instead, we computed the potential differences from adjusted heights obtained during national height system realizations. Second, we made no attempt to reduce the potential differences to the reference epoch adopted for the EVRS (i.e., epoch 2000.0). Doing so requires information about the long-term sea level variability and vertical land motion at the locations of the involved tide gauges. For most tide gauges, these data are not available or there is a mismatch between the period for which the data are available and the period used in this study. Third, we took the used hydrodynamic models as is. That is, we made no attempt to improve their

performance in representing the local MWL at the tide gauge locations (e.g., by improving the (local) spatial resolution of the hydrodynamic models). Finally, no validation is possible due to a lack of accurate control data. However, for those countries showing a large impact we assessed the agreement of the estimated heights with normal heights obtained by differencing GNSS heights and quasi-geoid heights obtained from the European gravimetric quasi-geoid model EGG2015 (Denker 2015).

The paper is organized as follows. Section 2 describes the methodology used to (i) compute the hydrodynamic leveling data, (ii) define the set of connections added to the adjustment, (iii) conduct the adjustment and variance component estimation, and iv) assess the impact. Section 3 introduces the data sets used in this study. In Sect. 4, we present, analyze, and discuss all results. Finally, we conclude by summarizing the main findings of the paper in Sect. 5.

## 2 Methodology

### 2.1 Functional and stochastic model of the hydrodynamic leveling dataset

In a rigorous implementation of model-based hydrodynamic leveling for the realization of the EVRS, the observation equation of the geopotential difference ( $\Delta W$ ) between two UELN height benchmarks (HBMs), one connected to tide gauge  $i$  and the other to tide gauge  $j$ , reads

$$\Delta W_{HBM_i}^{HBM_j} = \Delta W_{HBM_i}^{TGZ_i} + \Delta W_{TGZ_i}^{\overline{SMWL}_i} + \Delta W_{\overline{SMWL}_i}^{\widetilde{SMWL}_j} + \Delta W_{\widetilde{SMWL}_j}^{TGZ_j} + \Delta W_{TGZ_j}^{HBM_j}. \tag{1}$$

Here, TGZ stands for tide gauge zero, and  $\overline{SMWL}$  and  $\widetilde{SMWL}$  are the observation- and model-derived SMWLs, respectively. Note that it is assumed that all data have been reduced to the reference epoch adopted for the EVRS (epoch 2000.0). Also note that the hydrodynamic models assume a constant gravity acceleration, which we also used to compute  $\Delta W_{\overline{SMWL}_i}^{\widetilde{SMWL}_j}$  from the model-derived SMWLs. The adjective ‘rigorous’ refers to an implementation in which the first and last terms in Eq. (1) are determined using spirit leveling and gravimetry. In the context of this project, acquiring these measurements is not realistic and feasible. In addition, none of the available tide gauge records uses TGZ as the vertical reference. The water levels are typically expressed relative to the national height datum (NHD) or chart datum. The metadata to transform the water levels back into water levels with respect to TGZ is missing. Hence, we cannot compute the second and fourth terms in Eq. (1).

Therefore, we establish  $\Delta W_{HBM_i}^{HBM_j}$  using

$$\Delta W_{HBM_i}^{HBM_j} = \left( \gamma_{HBM_i} \widehat{H}_{NHD_i}^{HBM_i} - \gamma_{TG_i} \widehat{H}_{NHD_i}^{\overline{SMWL}_i} \right) + \Delta W_{\overline{SMWL}_i}^{\widetilde{SMWL}_j} + \left( \gamma_{TG_j} \widehat{H}_{NHD_j}^{\overline{SMWL}_j} - \gamma_{HBM_j} \widehat{H}_{NHD_j}^{HBM_j} \right), \tag{2}$$

where  $\gamma$  is the GRS80 normal gravity (Moritz 2000),  $\widehat{H}_{NHD}^{HBM}$  is the adjusted height of the nearest UELN benchmark with respect to the NHD, and  $\widehat{H}_{NHD}^{\overline{SMWL}}$  is the height of the observation-derived SMWL expressed with respect to the NHD. Note that in computing the first and last term, we treat all adjusted heights as normal heights. However, Belgium and The Netherlands use an uncorrected leveled height system, Denmark an orthometric height system, and Great Britain an normal-orthometric height system (Federal Agency for Cartography and Geodesy 2022b). The ‘error’ introduced is believed to be insignificant; the coastal areas in Belgium, The Netherlands, and Denmark have minimal topographic variations, whereas according to Filmer et al. (2010) the difference between normal-orthometric and normal heights does not exceed the 2–3 cm level for Australia where the highest mountain is almost 900 m higher than the one in Great Britain. Apart from that, for a small distance between the HBM and tide gauge the error will partly cancel out in differencing  $\widehat{H}_{NHD}^{HBM}$  and  $\widehat{H}_{NHD}^{\overline{SMWL}}$  as the error is quite systematic. This also applies for any other systematic error present in the adjusted heights. Furthermore, note that  $\widehat{H}_{NHD}^{\overline{SMWL}}$  is in fact the sum of three terms:

$$\widehat{H}_{NHD}^{\overline{SMWL}} = \widehat{H}_{NHD}^{TGBM} + \Delta H_{TGBM}^{TGZ} + H_{TGZ}^{\overline{SMWL}}, \tag{3}$$

where  $\widehat{H}_{NHD}^{TGBM}$  is the adjusted height of the tide gauge benchmark (TGBM) with respect to the NHD,  $\Delta H_{TGBM}^{TGZ}$  is the leveled height difference between the TGBM and TGZ, and  $H_{TGZ}^{\overline{SMWL}}$  the tide gauge derived SMWL expressed relative to TGZ. The TGBM is the HBM to which the TGZ is connected. Most of them are not part of the UELN.

The stochastic model of the hydrodynamic leveling dataset needs to account for the contributions of all three terms in Eq. (2). The contribution of the first and last term is mainly determined by the uncertainty in  $\widehat{H}_{NHD}^{HBM}$  and  $\widehat{H}_{NHD}^{TGBM}$  (Eq. (3)). Indeed, supposing the connections between the TGBMs and TGZs are established by spirit leveling with a precision of 0.5 mm/km (corresponding to the precision of first-order leveling (Bossler 1984)) and that the TGBM is typically close (within a few meters to a few kilometers) to the tide gauges, the contribution of  $\Delta H_{TGBM}^{TGZ}$  is likely below 1 mm in terms of standard deviation. Similarly, the uncertainty of the MWL computed over one month of sea level observations is already < 1 mm based on a 10-minute sampling and assuming white noise with a standard deviation

of 5 cm. As we lack the full variance-covariance matrices of the national height system realizations, we assume that the uncertainty in the first and last term of Eq. (2) is described by the precision of second-order leveling (Bossler 1984), i.e., 1.3 mm/km in terms of standard deviation. The distance between the HBM and TGBM is approximated by the ellipsoidal distance between the nearest UELN benchmark and the tide gauge location.

The full noise variance-covariance matrix associated with the middle term has been obtained using the noise models developed by Afrasteh et al. (2023). For the specific model being used, see Sect. 3.1.

## 2.2 Defining the set of hydrodynamic leveling connections

A key step in compiling the model-based hydrodynamic leveling dataset is to determine between which tide gauges connections should be established. Assuming  $N$  tide gauges are available,  $N - 1$  independent hydrodynamic leveling connections (i.e., connections that do not form any closed circuit) can be established. For this, there are at most  $N^{N-2}$  possibilities (Afrasteh et al. 2021). Similar to Afrasteh et al. (2021, 2023), we use a heuristic search method that identifies the connections one-by-one. Each connection added to the final set is the one that provided the lowest median standard deviation of the adjusted heights. No connections are allowed between tide gauges i) located in different hydrodynamic model domains (i.e., one in the 3D DCSM-FM and the other in the Nemo-Nordic model domain), ii) located within the same country, and iii) located in neighboring countries for which the number of spirit leveling connections between the countries is larger than one. The first criterion stems from the fact that there may be a bias between the water levels obtained from the different hydrodynamic models. The last two criteria are intended to reduce the computational load of the search method. Since we are using real data in this study, there is one additional preparatory step that will be described in the remainder of this section.

As noted in Sect. 1, the SMWL differences will be determined over three consecutive summer periods that are the same per region. The latter reduces potential time-dependent, large-scale errors in the modeled water levels (they cancel out in computing the difference) and simplifies the required noise model (Afrasteh et al. 2023). Model-derived water levels and tide gauge records are available from 1997–2019 and 2017–2021 for 3D DCSM-FM and Nemo-Nordic, respectively. Hence, we need to select which three successive years for each of these time spans will be used. In this choice, the following considerations were taken into account:

- The time difference between the center epoch of the 3-year period and the reference epoch of the EVRS (i.e.,

2000.0)—Ideally, there is no time difference. The greater the difference, the greater the impact of relative differences in long-term sea level variations and vertical land motion at the two involved tide gauge locations.

- Performance of the hydrodynamic model—Afrasteh et al. (2023) showed that for all three hydrodynamic models examined in their study, the performance to represent the SMWLs varies over space and time. As such, it makes sense to select a period in which the model shows a better performance.
- Distribution of the tide gauges—The tide gauges are not homogeneously distributed along the coastline. In addition, the distribution is time-dependent as in many cases the tide gauge records are not full (some tide gauges expired or became available later, and the records may contain gaps). There is no objective way to define the ‘best’ possible distribution. In this study, we use the maximum sea distance (Afrasteh et al. 2023, Sect. 2.1.4) between two adjacent tide gauges. More specifically, we minimize the weighted sum of the maximum ‘sea distance’ between two adjacent tide gauges per country, where the weighting is determined by the length of the country’s coastline.

The criteria will be applied to the sets of tide gauges available per domain to identify what is referred to as the ‘most preferred’ and ‘least preferred’ time spans. These are subsequently used in Experiments I and II to assess the impact of model-based hydrodynamic leveling on the EVRF. The comparison of both experiments allows to evaluate the importance of the time span selection.

## 2.3 Height network adjustment and data weighting

The height network adjustment is conducted using weighted least-squares, see Afrasteh et al. (2021) for the equations. Similar to the approach followed in computing EVRF2019 (Sacher and Liebsch 2019), the datum defect is solved by adding the minimal constraint that for 12 datum points the sum of the height changes is zero. The height of the 12 datum points are obtained from the EVRF2019 adjustment. Hence, the datum of the computed EVRF is the same as the one for the EVRF2019.

After determining which hydrodynamic leveling connections will be added, we conduct the height network adjustment and estimate variance factors for the UELN dataset and each model-based hydrodynamic leveling dataset. In doing so, we used the iterative minimum norm quadratic unbiased estimator (Rao 1971). The iteration was terminated when the relative change of successive variance factors for all observation groups was smaller than  $10^{-4}$ . Note that each of the observation groups that make up the UELN dataset are scaled

with the variance factor estimated by the BKG in computing the EVRF2019 (Sacher and Liebsch 2019, Table 3).

## 2.4 Impact assessment

The impact of hydrodynamic leveling data on the EVRF is assessed by comparing the estimated geopotential numbers to the ones computed using UELN data only. If the differences for a country show a variability larger than 10 kgal mm, we will assess the agreement of the solution with normal heights obtained by differencing GNSS and quasi-geoid heights. In doing so, we interpolate the differences between the adjusted EVRF heights and the heights expressed relative to the NHD to the GNSS data points. By adding the differences to the physical height of the GNSS data points, we obtain their EVRF heights. These are compared with the differences between the GNSS and quasi-geoid heights. Note that the comparison is conducted in the zero-tide system. As the UELN data were provided in the mean-tide system (Sect. 3.3), we apply the following transformation (Federal Agency for Cartography and Geodesy 2022a):

$$C_{\text{zero}} = C_{\text{mean}} - 0.28841 \sin^2(\phi) - 0.00195 \sin^4(\phi) + 0.09722 + 0.08432, \quad (4)$$

where  $C_{\text{zero}}$  and  $C_{\text{mean}}$  are the geopotential numbers in respectively the zero- and mean-tide system, and  $\phi$  is the geodetic latitude.

Apart from reporting statistics of the differences, including the median and standard deviation, we also assess the magnitude of trends in the differences in east–west and/or south–north directions. Here, the standard deviation was estimated as  $1.4826 \times$  the median absolute deviation (Cook and Weisberg 1982; Rousseeuw and Croux 1993). In case we estimate the trend in both directions, the magnitudes are estimated by fitting a plane through the differences. When estimating the south–north slope only, we fit a linear trend and intercept term to the differences as a function of latitude.

## 3 Data

### 3.1 Model-derived water level time series

In this study, two reanalysis products generated by two different hydrodynamic models have been used; one for the northwest European continental shelf generated with the 3D Dutch Continental Shelf Model – Flexible Mesh (3D DCSM-FM) (Zijl et al. 2020), and one for the Baltic Sea generated with the Nemo-Nordic model (Nemo-Nordic NS01) (Hordoir et al. 2019), which is based on version 3.6 of the Nucleus for European Modelling of the Ocean (NEMO) model code

**Table 1** Number of observations and variance factor for each observation group of the UELN dataset

Country	Number of observations	Variance factor
Austria	179	0.832
Belgium	122	1
Bulgaria	118	2.678
Belarus	31	5.389
Croatia	81	1.97
Czech Republic new	185	0.687
Czech Republic old	83	1.41
Denmark	196	0.728
Estonia	418	0.052
Finland	272	0.545
France (IGN69)	344	100
France (NIREF)	1223	1.965
Germany	1112	0.431
Great Britain	61	14.54
Hungary	83	0.306
Italy	203	2.163
Latvia	151	0.722
Lithuania	65	0.566
Macedonia	66	0.905
Netherlands	1373	0.572
Norway new	550	1.805
Norway old	410	2.08
Poland	473	0.743
Portugal	30	4.046
Romania	133	3.044
Russia	176	4.908
Slovakia	196	2.212
Slovenia	89	0.383
Spain	227	5.649
Sweden	4206	0.994
Switzerland	719	0.825
Ukraine	211	2.894

Note that the datasets for Belgium, Bulgaria, and Norway include a number of single connections between a UELN HBM and a GNSS data point. Since no new loops are formed, these connections do not ‘contribute’ to the solution. The variance factors have been estimated by the BKG when computing the EVRF2019. As explained by Sacher and Liebsch (2019), the estimated variance factors for the Belgian and old French datasets (referred to as ‘France’) were not used, but set manually. The variance factor for the British dataset has also been manually set as this observation group was not included in the computation of the official EVRF2019

(Madec et al. 2016) and 3DEnVar data assimilation method (Axell, L. and Liu, Y. 2016). Some alternative products covering our area of interest are publicly available via <https://data.marine.copernicus.eu/products>. Key requirement is that the models used to generate them included all relevant physics contributing to the MWL variability.

The 3D DCSM-FM reanalysis is the same as used by Afrasteh et al. (2023). It is obtained using the Delft3D Flexible Mesh software framework that allows for the use of unstructured grids. For this study, we have used software version 2.17.05.72090. The 3D DCSM-FM is a three-dimensional hydrodynamic model with a grid resolution that varies between 0.5 and 4.0 nautical miles. It covers the area between 15° W to 13° E and 43° N to 64° N. The water level time series are generated for the period 1997–2019. For more details about the model setup and forcing data, we refer to Zijl et al. (2020) and (Afrasteh et al. 2023,) [Table 1]. Note that the model is actively maintained and improved. One known issue in the version of the model used in this study is an apparent strong vertical circulation between the bottom and the pycnocline in the deep ocean originating from instabilities close to the open boundaries. This results in a less accurate representation of the MWL in deep ocean waters. Detailed investigations by the model developers showed that the issue has no impact on the shelf. A thorough assessment of 3D DCSM-FM's ability to represent the MWLs computed over one summer period (referred to as the '1-SMWLs') is given by Afrasteh et al. (2023). Their results show that the performance varies over space and time. With regard to the latter, they noticed improved performance in the period 2004–2011, although they had no explanation for it.

The Nemo-Nordic model, developed by the Swedish Meteorological and Hydrological Institutes, is a three-dimensional coupled ocean-sea ice model. The model domain covers both the Baltic and the North Sea (i.e., it ranges from 4.15278° W to 30.1802° E and 48.4917° N to 65.8914° N). In this study, however, we only use the water level time series for the Baltic area. The model is based on the NEMO ocean engine version 3.6. For a detailed description of the model setup and forcing data, as well as a validation of the water level time series, we refer to Jahanmard et al. (2021, 2022). Note that the hourly water levels are exported on grids with a horizontal resolution of 1 nautical mile. The reanalysis period ranges from 2017 to 2021.

The method applied to compute the SMWLs over three successive summer periods, referred to as 3-SMWLs, is described by Afrasteh et al. (2023). For the 3D DCSM-FM dataset, a noise model was computed by Afrasteh et al. (2023), which has been exploited in this study. For the Nemo-Nordic dataset, we used the noise model developed for the Forecasting Ocean Assimilation Model 7 km Atlantic Margin model (AMM7) (Tonani and Ascione 2021). The only motivation we have for using this noise model is that the AMM7 hydrodynamic model, similar to the Nemo-Nordic model, relies on the Nucleus for European Modeling of the Ocean model code (Madec et al. 2016). Unfortunately, we lack the required long-time series of observation- and model-derived water levels to develop an empirical noise model specific to this dataset. The lack of a tailored noise model for the Nemo-

Nordic dataset is among the reasons why variance component estimation is conducted.

### 3.2 Observation-derived water level time series

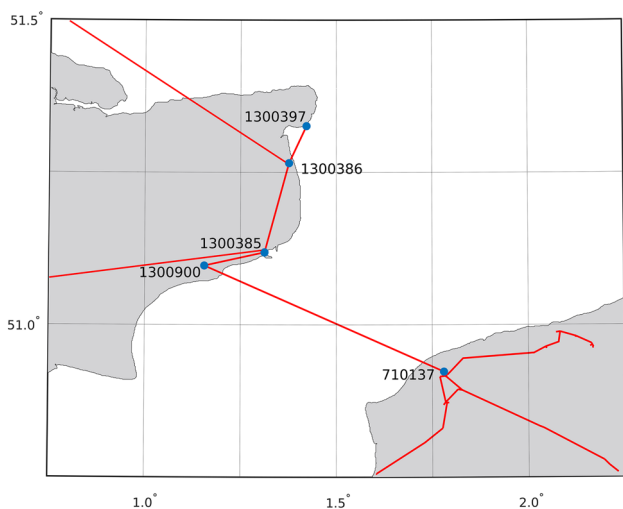
Tide gauge data were obtained from the national authorities in all coastal countries included in the two hydrodynamic model domains, except Spain and Ireland (3D DCSM-FM), and Lithuania and Russia (Nemo-Nordic). Spanish tide gauge data were not included in our database because we did not expect a good performance due to the model issue mentioned in Sect 3.1. For Ireland, we lack spirit leveling data. Regarding Lithuania and Russia, no data were available. The time span of the records is consistent with the time span of the reanalyses; 1997–2019 (3D DCSM-FM) and 2017–2021 (Nemo-Nordic). For the 3D DCSM-FM, the tide gauge dataset is similar to the one used and described by Afrasteh et al. (2023). It includes about 200 records. Note, however, that most of these records do not cover the entire reanalysis time span. For the Nemo-Nordic domain, about 50 tide gauge records were available. The records were used to compute monthly MWL time series by means of a harmonic analysis using UTide (Codiga 2020).

All water levels use the NHD as the vertical reference. For the tide gauges inside the Nemo-Nordic domain, a correction for the vertical land motion induced by glacial isostatic adjustment has been applied. The correction was computed using the regional land uplift model NKG2016LU\_abs (Vestøl et al. 2019). It reduces all water levels to the reference epoch 2000.0. Note that this reference epoch is consistent with the one adopted for the EVRS.

A tide gauge is only considered as a potential candidate for a hydrodynamic leveling connection if it is i) outside the tidal flat areas, ii) connected to the NHD, and iii) within 40 km of the nearest UELN HBM. If multiple tide gauges were connected to the same UELN HBM, we used the one of which the median absolute deviation of the differences between the observation- and model-derived monthly MWLs was lowest. The 3-SMWLs were computed from the monthly MWL time series. Note that all monthly MWLs were excluded for which the value of the difference between the observation- and model-derived monthly MWLs exceeded the median of these differences plus/minus three times the standard deviation (estimated as before).

### 3.3 UELN data

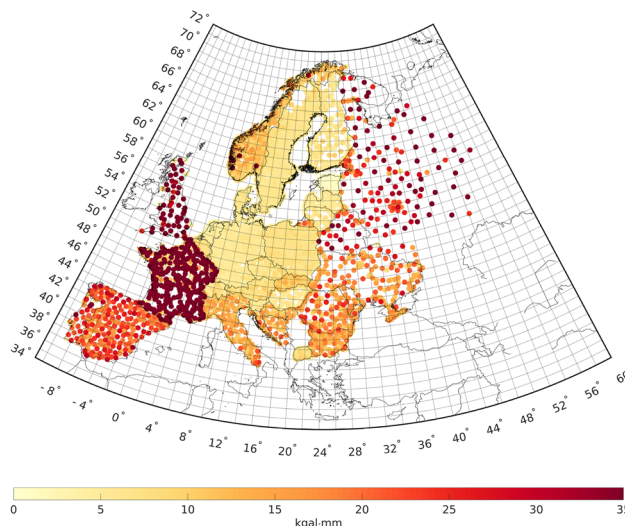
The UELN data were provided by the BKG. The dataset comprises i) the geopotential differences between the UELN HBMs in the mean-tide system, ii) the a-priori variances for the geopotential differences, and iii) a variance factor for each observation group per country estimated by means of variance component estimation (see Table 1).



**Fig. 1** A zoom-in of the UELN, showing the connection between Great Britain and France through the Channel Tunnel

The dataset is, except for the following changes, the same as the one used to compute the EVRF2019 (release of September 2020). First, we included the data of the Third Geodetic Leveling in Great Britain acquired from 1951 to 1959. Second, we added a number of single connections (i.e., connections that do not form new loops) to HBMs for which GNSS data are available in Norway (61), Belgium (9), and Bulgaria (21). These were meant to facilitate a comparison with normal heights obtained by differencing GNSS heights and quasi-geoid heights obtained from the European gravimetric quasi-geoid model EGG2015. Note that the release of the EVRF2019 in September 2020 includes (i) a sign correction in the tidal correction for the Polish dataset and (ii) a number of minor updates to the data for Macedonia (error correction), Latvia, Lithuania, Italy (new nodal points), and Bulgaria (two border connections to Turkey have been included).

In computing the EVRF2019, only the datum shift between the Ordnance Datum Newlyn (ODN) and the EVRF was determined. This was done by adding four leveling connections between the HBMs ‘710137’ and ‘1300900’, ‘1300900’ and ‘1300385’, ‘1300385’ and ‘1300386’, and ‘1300386’ and ‘1300397’ to the French zero order leveling network (NIREF) dataset. With the exception of HBM ‘710137’, all mentioned HBMs are on British territory (see Fig. 1). The reason for determining only the datum shift stems from the fact that previous realizations of the EVRS showed a tilt relative to the ODN reference frame (Sacher and Liebsch 2019). This tilt is caused by systematic errors in the British leveling data (see Penna et al. (2013) and cited references). The ODN has been computed based on an adjustment of data from the Third Geodetic Leveling while fixing the results from the Second Geodetic Leveling (1912–1952). On the other hand, the BKG only had (and still has) access to data



**Fig. 2** Standard deviations of the UELN observations. Note that the variance factors presented in Table 1 have been applied

from the Third Geodetic Leveling. In the dataset we used, the four connections mentioned above are part of the British dataset. The variance factor for this data set has been set equal to 9.867 by the BKG, which means that these four connections are given a lower weight than in the computation of the EVRF2019 (the variance factor for the French NIREF data is 1.965).

Figure 2 shows the standard deviations of the UELN observations. Here, the variance factors presented in Table 1 are already applied.

### 3.4 EUVN\_DA data

The British and French GNSS data used in the impact assessment are part of the dataset assembled in the European Vertical Reference Network—Densification Action (EUVN\_DA) project (Kenyeres et al. 2010). The British dataset comprises 181 data points and the French dataset 164 points. The dataset comprises ellipsoidal coordinates in the European Terrestrial Reference System 1989 (ETRS89) and leveled connections to UELN benchmarks. Kenyeres et al. (2010) did not provide details on when the specific datasets were acquired. It should be somewhere between 2003 and 2008. They also did not specify the length of the sessions over which GNSS data were acquired. They only stated that some countries, including France and Great Britain, “did not observe sessions of 24 hours, but submitted a denser database with a mean site separation close to 50 km.” As such, we expect that the accuracy of the ellipsoidal heights is slightly lower than the 1 cm target accuracy.



### 3.5 The European gravimetric quasi-geoid model EGG2015

The European gravimetric quasi-geoid model EGG2015 (Denker 2015) is the latest of a series of European quasi-geoid models computed by the Institute of Geodesy at the Leibniz University of Hanover. The model was computed from surface gravity data in combination with topographic information, as well as the GOCO05S geopotential model (Mayer-Gürr 2015). Here, the remove-compute-restore technique was exploited. The estimated uncertainty in terms of standard deviation is 1.9 cm (Denker et al. 2018). Further details about the datasets, computational method, and uncertainty are provided by Denker (2013); Denker et al. (2018). Note that EGG2015 is in the zero-tide system. The normal heights obtained by differencing the GNSS and EGG2015 heights are referred to as the ‘GNSS/EGG2015 normal heights’.

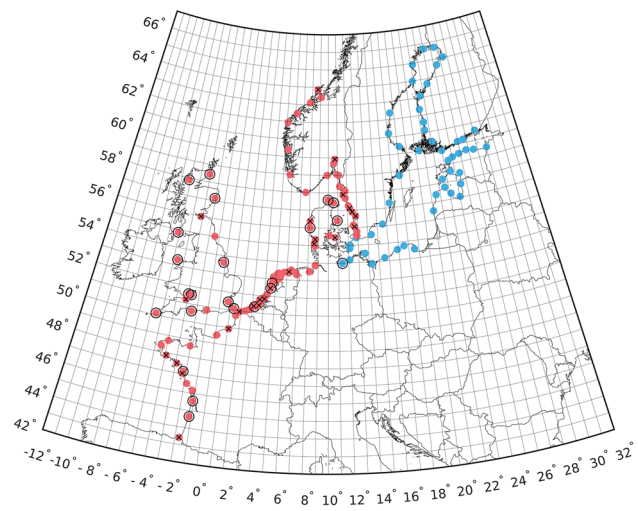
## 4 Results and discussion

### 4.1 Experiment 0: The UELN-only solution

To quantify the impact of model-based hydrodynamic leveling data on the EVRF, we used a solution computed using UELN data only as a reference. This solution is referred to as the ‘UELN-only solution’. As we included the British leveling data as well as a few minor updates to some other observation groups of the UELN dataset (Sect. 3.3), in Experiment 0 we quantify the differences between this solution and the official EVRF2019 (release September 2020, available at <https://evrs.bkg.bund.de/Subsites/EVRS/EN/EVRF2019/evrf2019.html>).

The differences are significant at three HBMs. The largest difference is observed for HBM ‘1706226’ (Poland); it reaches 489.5 kgal mm. A detailed investigation shows that the geopotential number associated with the official EVRF2019 is incorrect. The other two HBMs, i.e., ‘1300386’ and ‘1300397’, are located in Great Britain. In the computation of the EVRF2019, the leveling connections to these HBMs were included in the France NIREF dataset. For both HBMs, the difference equals 6.527 kgal mm.

A quality assessment of the UELN-only/EVRF2019 solutions is out of the scope of this study. The solutions are known to be contaminated by systematic errors in various observation groups, including but not limited to Belgium (Slobbe et al. 2019), France (Duquenne et al. 2015), and Great Britain (Hipkin et al. 2004; Ziebart et al. 2008; Penna et al. 2013). When evaluating the impact of adding hydrodynamic leveling data, though, we included the UELN solution in the comparison with GNSS/EGG2015 normal heights.



**Fig. 3** The locations of the available tide gauges inside the 3D DCSM-FM (red dots) and the Nemo-Nordic (blue dots) model domains. The tide gauges marked with a black circle are those that are only available for the most preferred time span, while those marked with a black cross are only available for the least preferred one. No black circle or cross means the tide gauge is available for both time spans

### 4.2 Experiment I: Adding hydrodynamic leveling data computed over the most preferred time span

In Experiment I, we complemented the UELN dataset with hydrodynamic leveling connections between tide gauges in both the Baltic Sea (Nemo-Nordic dataset) and the northwest European continental shelf (3D DCSM-FM dataset). The 3-SMWLs were computed over the periods 2017–2019 and 2004–2006, respectively. Based on the criteria outlined in Sect. 2.2, these time spans are considered the most preferred ones. The number of available tide gauges is 50 (Nemo-Nordic) and 76 (3D DCSM-FM). We refer to Fig. 3, for a map showing the locations of the tide gauges. In total 124 hydrodynamic leveling connections were added; 49 in the Baltic Sea and 75 in the northwest European continental shelf. For an overview of which connections are being added we refer to Appendix A.

Table 2 provides the estimated variance factors for the three observation groups. Figure 4 shows the differences

**Table 2** Variance factors estimated in Experiments I and II

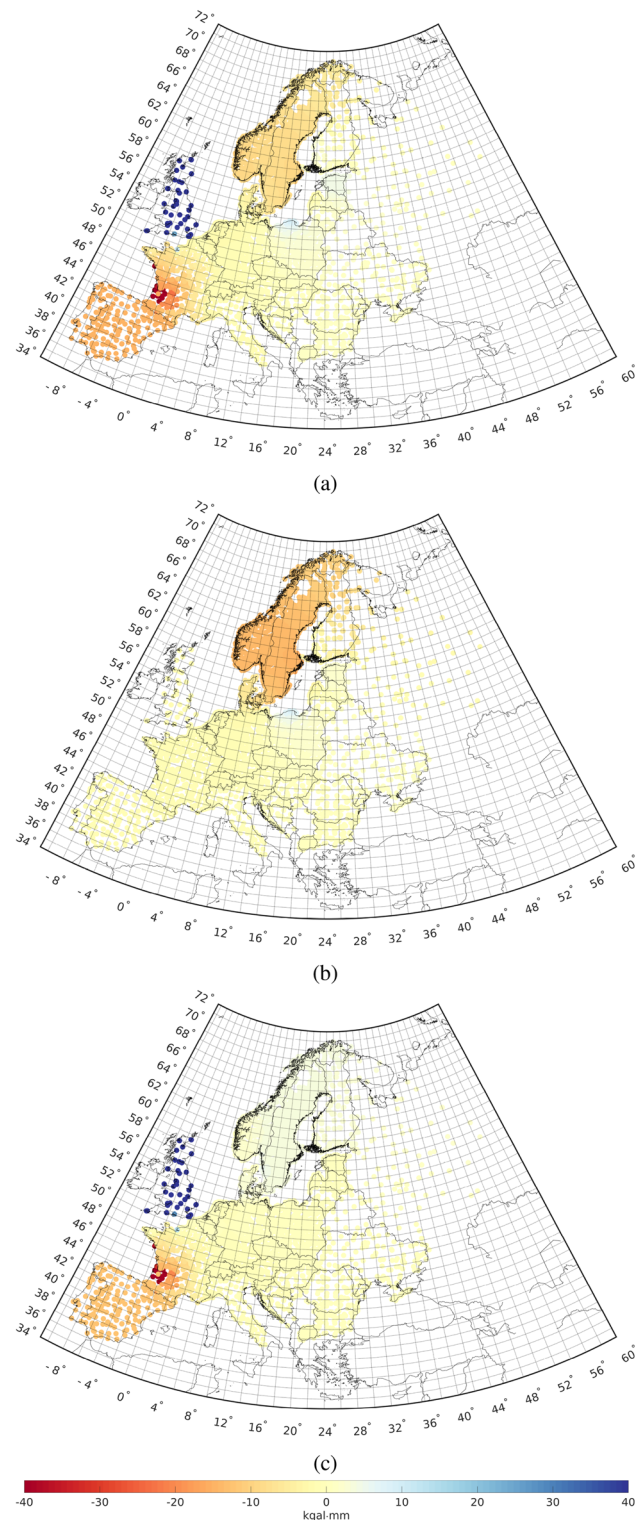
	Experiment I	Experiment II
UELN	0.973	0.964
3D DCSM-FM	2.156	1.121
Nemo-Nordic	2.273	2.174

Note that the UELN dataset is treated as one observation group. The relative weighting factors of all observation groups that make up the UELN dataset are provided in Table 1

between the geopotential numbers of the UELN-only solution and the ones estimated in Experiment I, as well as the contribution to these differences from both hydrodynamic leveling datasets separately. These contributions were obtained by realizing the EVRS based on the UELN dataset and the identified hydrodynamic leveling connections in the specific sea basin. The applied variance factors are the same as those estimated when using both hydrodynamic leveling datasets. The results show the following:

- Both hydrodynamic leveling datasets are downweighted compared to the UELN dataset; for the latter dataset, the estimated variance factor is 0.973, whereas for the 3D DCSM-FM and Nemo-Nordic hydrodynamic leveling datasets the variance factors are 2.156 and 2.273, respectively. This means that the standard deviations of the added hydrodynamic leveling connections increase from 33.8–41.5 kgal mm to 49.6–61.0 kgal mm (3D DCSM-FM) and 38.4–59.9 kgal mm to 58.0–90.3 kgal mm (Nemo-Nordic).
- The differences between the geopotential numbers show a large-scale pattern, of negative (positive) values in the west (east). The median difference is  $-3.1$  kgal mm and the standard deviation (estimated as  $1.4826 \times$  the median absolute deviation (MAD)) is 6.1 kgal mm.
- Great Britain and France stand out in terms of the magnitude of the differences. In Great Britain, the values range between 23.8 and 444.2 kgal mm. In France, the large differences are concentrated in the southwest. The values range between  $-96.7$  and 10.5 kgal mm. For all other countries, the values range between  $-15.2$  (Spain) and 8.9 kgal mm (Poland).
- Only in the countries France and Great Britain the range of the differences exceed 10 kgal mm.
- The contribution of the Nemo-Nordic dataset (see Fig. 4b) is largest in Norway, Sweden, and northern Finland (between  $-16.5$  and  $-10.1$  kgal mm) as well as along the Polish coastline (up to 8.3 kgal mm). Likewise, the contribution of the 3D DCSM-FM dataset (see Fig. 4c) is the largest in Great Britain (between 27.9 and 450.8 kgal mm), France (between  $-91.6$  and 13.8 kgal mm), and Spain (between  $-12.4$  and  $-11.8$  kgal mm).

The observed impact of the hydrodynamic leveling datasets on the EVRF is significant, especially in Great Britain and France where differences reach tens to hundreds of kgal mm. In most other countries the impact is substantially lower; we observe spatially correlated differences of low magnitude. This result is easily explained when considering i) the domain in which model-based hydrodynamic leveling connections have been established, and ii) the weighting (i.e., quality) of



**Fig. 4** Differences between the geopotential numbers associated to the UELN-only solution (Experiment 0) and the ones estimated in Experiment I (a). The panels (b) and (c) show respectively the contributions of the 49 connections added in the Baltic Sea and the 75 in the northwest European continental shelf

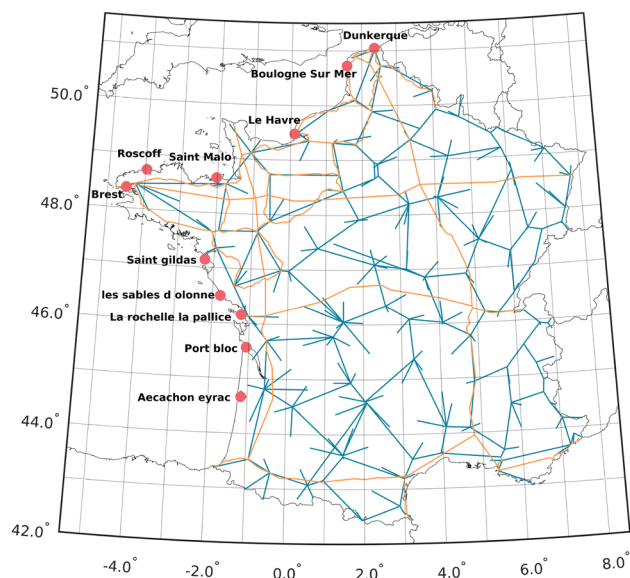
the (individual observation groups within the) UELN dataset relative to the two hydrodynamic leveling datasets.

Indeed, on the former, no major impact was expected for non-coastal countries and the countries around the Mediterranean and the Black Seas because the leveling networks of these countries were not strengthened with model-based hydrodynamic leveling connections. Note that the impact observed in Spain has been propagated from elsewhere; no Spanish tide gauges were used to establish model-based hydrodynamic leveling connections. The difference in impact between Great Britain and France on the one hand and the other North Sea countries on the other hand, as well as the small impact for the Baltic Sea countries, follows directly from the weighting of the different observation groups. Most northwestern European and Baltic Sea countries have high-quality leveling datasets (see Fig. 2). As such, it would have been more useful to include hydrodynamic leveling connections in the Mediterranean and Black Seas. The downweighting of the hydrodynamic leveling datasets relative to the UELN dataset also suppresses the impact. While the downweighting is significant (see the increase in the standard deviations of the added hydrodynamic leveling connections), it is within reasonable bounds. Further discussion of the obtained variance factors follows in the discussion of the results of Experiment II (Sect. 4.3).

The fact that the differences in France are confined to the southwest is probably related to the difference in density of the old IGN69 leveling network versus that of the new NIREF network. As shown in Fig. 5, the two networks only overlap in the northwest. Outside the northwest, the IGN69 network is denser, i.e., contains smaller loops. Given the low weight assigned to this dataset in the adjustment (see Table 1), it is not surprising that the solution along the southwest coast is more impacted by the added hydrodynamic leveling connections.

It is striking to observe that the impact of adding hydrodynamic leveling connections among Baltic Sea tide gauges on the EVRF in Norway and Sweden is larger than when adding leveling connections involving tide gauges along the coasts of Norway and western Sweden (cf. Figure 4b and Fig. 4c). In the latter case, there is hardly any impact. A conclusive explanation for this result is lacking. One explanation might be that the Norwegian and Swedish leveling networks fit well to the central part of the UELN (comprising The Netherlands, Germany, and Denmark) while the Nemo-Nordic dataset introduces some deformation of the network. Alternatively, the Nemo-Nordic dataset may correct some deformation of the network the 3D DCSM-FM dataset is not able to do.

Given the small ranges of the differences per country between the solutions obtained in Experiments 0 and I (they are  $< 10$  kgal mm, except for France and Great Britain), a comparison with GNSS/EGG2015 normal heights is only



**Fig. 5** The old French leveling network IGN69 (blue) and the new zero-order leveling network named NIREF (orange). The old network is suspected for a tilt of 23 cm in North–South direction (Sacher and Liebsch 2019)

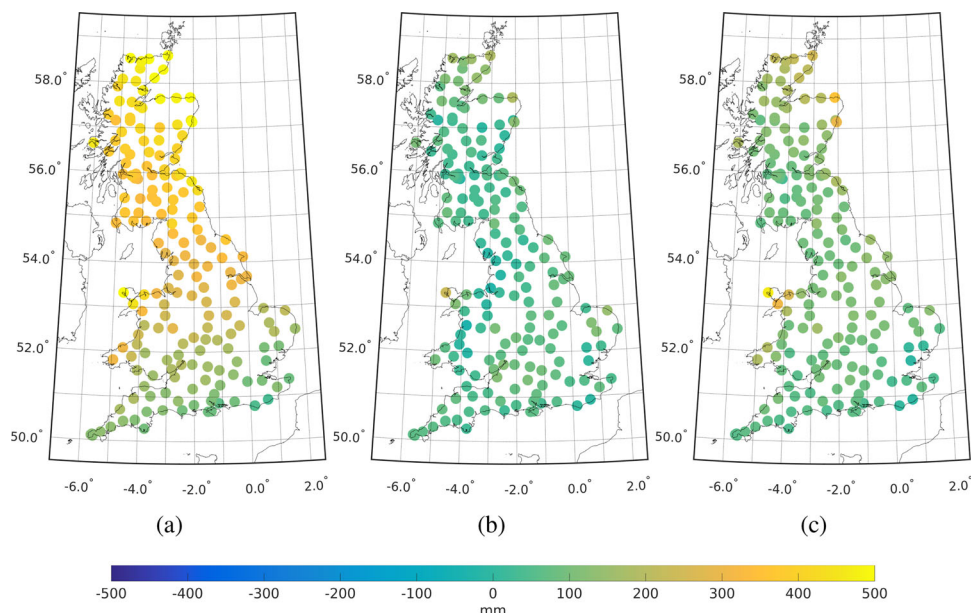
conducted for Great Britain and France. The results are summarized by difference maps (Fig. 6 and top panel of Fig. 8), plots of the differences as a function of latitude (top panel in Fig. 7), plots showing the plane fitted to the differences (Fig. 9) and the histograms of the differences (bottom panels in Figs. 7 and 8).

The comparison shows the following:

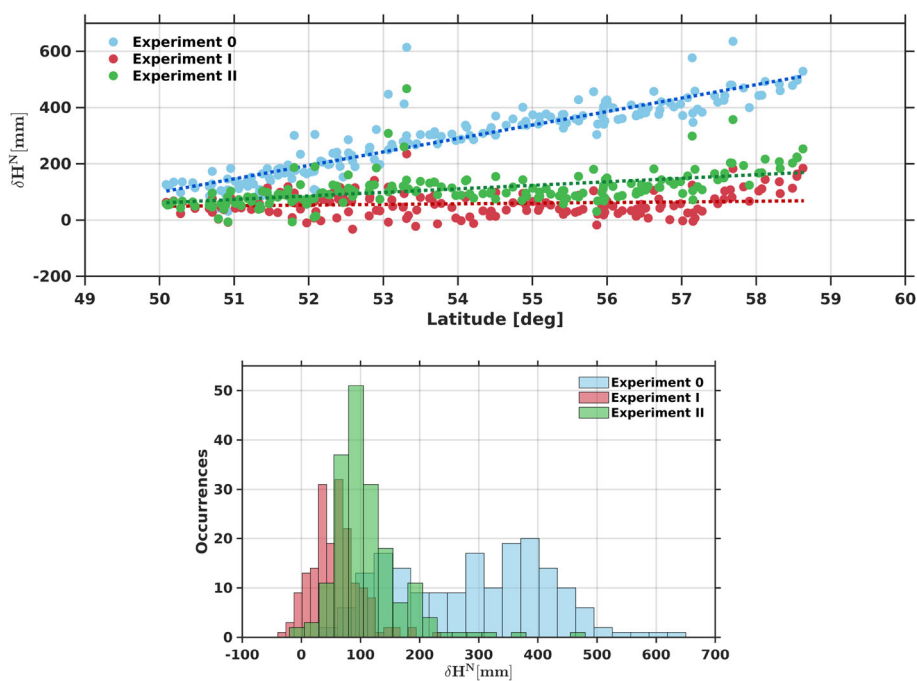
- In Great Britain, the differences for the UELN-only solution show a south-north slope of  $48 \text{ mm deg}^{-1}$ . In the solution including model-based hydrodynamic leveling data (Experiment I) this trend almost disappeared; the estimated slope is  $2.2 \text{ mm deg}^{-1}$ .
- In Great Britain, the median of the differences reduces from 304 mm to 57 mm whereas the standard deviation (estimated as before) reduces from 156 mm to 35 mm.
- In France, the differences show a northwest-southeast slope in both solutions. In the southwest the differences are lower for Experiment I. The south-north slope in Experiment 1 reduces from  $12.6 \text{ mm deg}^{-1}$  to  $10.1 \text{ mm deg}^{-1}$ . However the east–west slope increases from  $5.3 \text{ mm deg}^{-1}$  to  $7.6 \text{ mm deg}^{-1}$ .
- In France, the median difference reduces from  $-31$  mm to  $-21$  mm. The standard deviation (estimated as before) does not change.

The first conclusion to be drawn from the comparison is that the EVRF heights obtained in Experiment I show better agreement with the GNSS/EGG2015 normal heights, both in France and Great Britain. In Great Britain, there is hardly

**Fig. 6** Differences between the EVRF heights estimated in Experiments 0 (a), I (b), and II (c) and the GNSS/EGG2015 normal heights at the EUVN data points in Great Britain



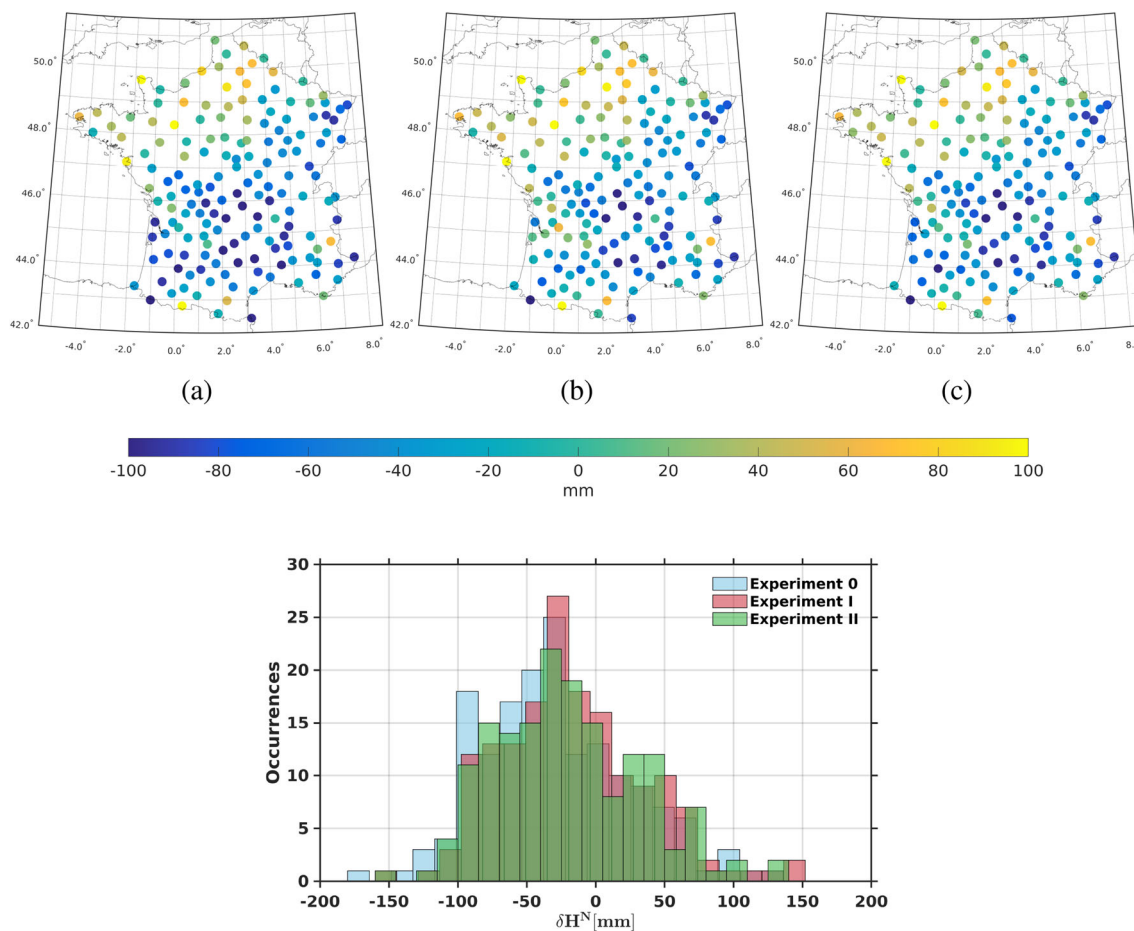
**Fig. 7** The top panel shows the differences between the EVRF heights estimated in the three Experiments and the GNSS/EGG2015 normal heights ( $\delta H^N$ ) at the EUVN data points in Great Britain as a function of latitude as well as the fitted linear model. The bottom panel shows the histograms of the differences



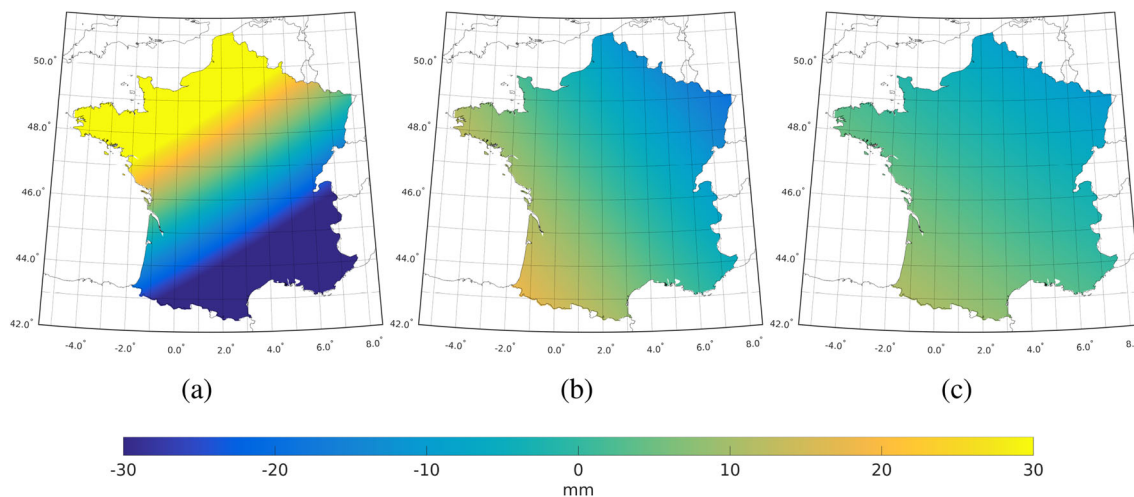
any slope left in the differences associated with Experiment I. In France, the better agreement is concentrated in the southwest. Given the uncertainty of the GNSS data (Sect. 3.4) and the EGG2015 (Sect. 3.5), we can conclude that for these two countries the impact of model-based hydrodynamic leveling data on the quality of the EVRF is positive. The fact that the impact is greater for Great Britain can again be explained by the fact that the border of the British leveling network is completely adjacent to the sea, so that it could be strengthened over the entire perimeter. For France, the strengthening is only on the west side. However, there are possibilities for

further reinforcement; in the southwest, France borders the Mediterranean Sea.

Note that the magnitude of the estimated slope associated with the UELN-only solution is larger than the  $-(20-25) \text{ mm deg}^{-1}$  reported in (Penna et al. 2013). This difference originates from the use of both the Second and Third Geodetic Leveling datasets in the realization of the British vertical datum (Ordnance Datum Newlyn), whereas we only used the latter one (see also Sect. 3.3). To what extent model-based hydrodynamic leveling data can help to iden-



**Fig. 8** Top row shows the differences between the EVRF heights estimated in Experiments 0 (a), I (b), and II (c) and the GNSS/EGG2015 normal heights at the EUVN data points in France. The bottom panel shows the histogram of the differences



**Fig. 9** Panel a shows the plane fitted to the differences between the EVRF heights estimated in Experiments 0 and the GNSS/EGG2015 normal heights. Panels b and c show the change in the plane fitted in

respectively Experiments I and II with respect to the one for Experiment 0. To enhance visibility, the mean values have been removed. These are 548.4,  $-110.8$ , and  $-88.1$  in (a), (b), and (c), respectively

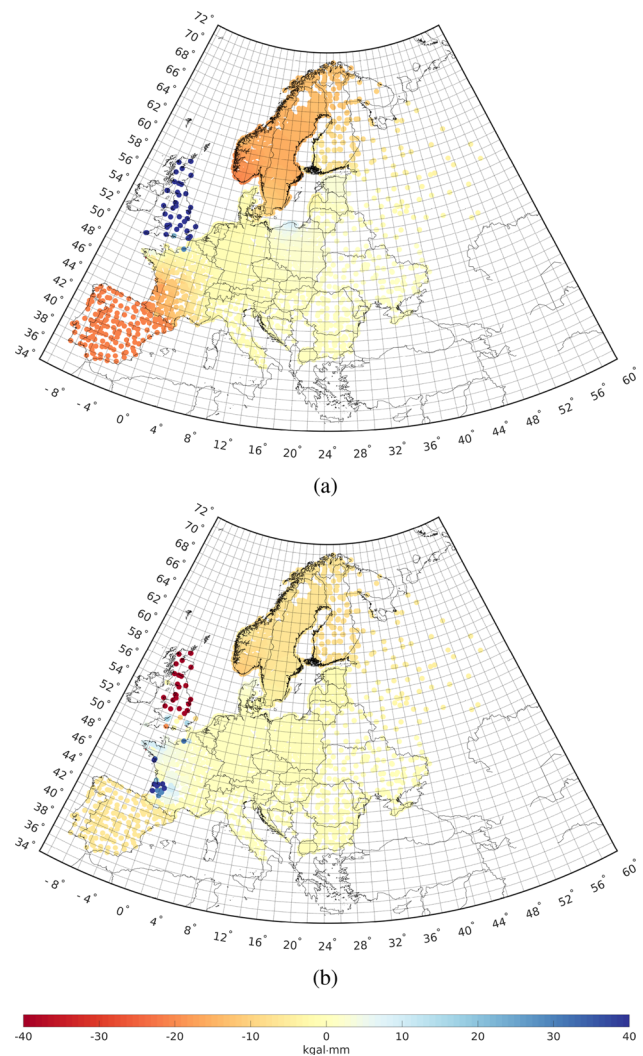
tify the exact source of the systematic leveling errors in the British dataset is outside the scope of this study.

### 4.3 Experiment II: Adding hydrodynamic leveling data computed over the least preferred time span

In Experiment II, we repeat Experiment I while using time spans to compute the 3-SMWLs that can be considered as the least preferred choice based on the criteria outlined in Sect. 2.2. These are 2015–2017 for the 3D DCSM-FM dataset and 2018–2020 for the Nemo-Nordic dataset. Note that a different time span mainly affects the 3D DCSM-FM dataset; the time span over which the reanalysis is conducted is 22 years. For the Nemo-Nordic dataset, we only have 4 years of data available. Hence, there will be an overlap of 2 years. The choice of the time span impacts the quality of the hydrodynamic leveling connections (remember that the performance of the 3D DCSM-FM reanalysis varies over time (Afrasteh et al. 2023)) as well as the set of tide gauges available to establish connections. In Experiment II, we will assess to what extent the choice of the time span changes the impact of model-based hydrodynamic leveling on the EVRF observed in Experiment I.

For the Baltic Sea and northwest European continental shelf, we have 49 and 79 tide gauges available, respectively (see Fig. 3). Table 2 provides the estimated variance factors for the three observation groups. Figure 10 shows the differences between the geopotential numbers of the UELN-only solution and the ones estimated in Experiment II, as well as the differences with respect to the solution obtained in Experiment I. The comparison with the GNSS/EGG2015 normal heights are summarized in Figs 6–8. Moreover, Fig. 9 shows the changes in the plane fitted to the differences with the GNSS/EGG2015 normal heights with respect to the one fitted in Experiment 0. The results show the following:

- The downweighting observed in Experiment I is lower. For the Nemo-Nordic dataset, the variance factor reduced from 2.273 to 2.174. For the 3D DCSM-FM, the value reduced from 2.156 to 1.210.
- The overall pattern in the differences compared to the UELN-only solution is the same.
- Changes compared to the solution obtained in Experiment I reach  $-0.4$  kgal mm and  $2.2$  kgal mm in terms of the median and standard deviation (estimated as before), respectively. They are largest in Great Britain (between  $-235.8$  and  $16.5$  kgal mm), France (between  $-33.0$  and  $79.7$  kgal mm), and the Scandinavian countries (between  $-12.5$  and  $3.8$  kgal mm).
- The agreement with the GNSS/EGG2015 normal heights reduced compared to the solution in Experiment I for both France and Great Britain. For Great Britain, we observe



**Fig. 10** Differences between the geopotential numbers associated to the UELN-only solution (Experiment 0) and the ones estimated in Experiment II (a) and the ones associated to Experiments I and II (b)

an increased offset and slope. The slope in this case increases to  $12.6$  mm deg $^{-1}$ . For France, we observe that the improved agreement in the southwest has gone. The tilt of the fitted plane in latitudinal direction does not show any significant change compared to Experiment I (i.e., it changed from  $-10.1$  mm deg $^{-1}$  to  $-10.6$  mm deg $^{-1}$ ). In longitudinal direction, however, the tilt decreases from  $7.6$  mm deg $^{-1}$  in Experiment I to  $6.4$  mm deg $^{-1}$  in Experiment II.

The impact of the different time spans used to compute the 3-SMWLs is significant and to be expected considering the differences between the sets of tide gauges available to establish the hydrodynamic leveling connections (see Fig. 3). Compared to the set available in Experiment I, in Experiment II we miss 12 tide gauges in Great Britain while only 2

new ones were added. In southwest France, the set in Experiment I includes two more tide gauges. Also, the set of Danish tide gauges is much less homogeneously distributed in Experiment II. At the same time, the number of Swedish tide gauges on the Kattegat side has increased substantially. In fact, given the small number of tide gauges along the British coast in Experiment II (4), it is surprising that there is still such a level of agreement with the GNSS/EGG2015 normal heights compared to what is observed for the UELN-only solution.

A striking difference between the two Experiments is observed in the estimated variance factors for the 3D DCSM-FM dataset. In Experiment II the number is significantly smaller. Compared to the variance factor of the UELN dataset, the downweighting observed in Experiment II is lower. This result can be understood by considering the regional variability in performance of the 3D DCSM-FM reported by Afrasteh et al. (2023); the best performance was obtained in the Kattegat-Skagerrak region. In British waters (i.e., Bristol Channel, Celtic Sea, Irish Sea and St. George's Channel, and the Inner Seas off the West Coast of Scotland), the performance was less. Given the aforementioned differences between the two sets of tide gauges available for establishing hydrodynamic leveling connections, Experiment II added more connections from/to tide gauges in the Kattegat-Skagerrak than to tide gauges in British waters. With regard to the Nemo-Nordic dataset, we would like to remark that a tailored noise model is missing; the length of the available time series is not sufficient to develop such a model. Instead, we relied on a noise model developed for the AMM7 hydrodynamic model covering the northwest European continental shelf (Sect. 3.1). The downweighting of the Nemo-Nordic dataset observed in both experiments suggests that the quality of the Nemo-Nordic model in representing the SMWLs is lower than that of the AMM7 model.

## 5 Summary and conclusion

This study presents the first realization of a regional height reference system, namely the European Vertical Reference System (EVRS), based on the combination of geopotential differences determined with spirit leveling/gravimetry and model-based hydrodynamic leveling. The latter technique was introduced by Slobbe et al. (2018) as an efficient and flexible alternative method to connect islands and offshore platforms to the height system on land. The study built upon two previous studies on exploiting model-based hydrodynamic leveling data. In Afrasteh et al. (2021), we demonstrated the potential of the technique for the realization of the EVRS by means of geodetic network analyzes. That was, we assessed the potential impact on the quality of the European Vertical Reference Frame (EVRF). In Afrasteh et al. (2023), we presented empirical noise models

for three different reanalysis products available for the northwest European continental shelf and reassessed the quality impact.

In the present study, we assessed the impact of model-based hydrodynamic leveling on the EVRF using real data. In doing so, we computed summer mean water levels (SMWLs) over three subsequent years (referred to as the 3-SMWLs) with the Nemo-Nordic and 3D DCSM-FM models covering the Baltic Sea and the northwest European continental shelf, respectively. About 250 coastal tide gauges, each located inside one of the two model domains but outside the tidal flat areas, were connected to the nearest Unified European Leveling Network (UELN) height benchmark (HBM). In establishing these connections, we relied on the adjusted heights obtained in the national height system realizations. Note that all tide gauges for which the distance to the nearest UELN HBM is  $> 40$  km were excluded. Moreover, in case multiple tide gauges were connected to the same UELN HBM, we used the one in which the median absolute deviation of the differences between the observation- and model-derived monthly MWLs was lowest. For both reanalysis products, the 3-SMWLs were calculated over two different time spans. The first is considered the most preferred time span based on three criteria, while the second reflects the least preferred choice. The results obtained using these time spans are presented in Experiments I and II. Note that the choice of the time span is particularly relevant for the 3D DCSM-FM dataset; the reanalysis covers the period 1997–2019. The reanalysis based on the Nemo-Nordic model runs from 2017 to 2021. In determining the connections between the tide gauges, we made use of the heuristic search algorithm developed by Afrasteh et al. (2021). This algorithm identifies the set of hydrodynamic leveling connections that provide the lowest median of the propagated height standard deviations.

The impact of model-based hydrodynamic leveling on the EVRF was determined by comparing the solutions with the so-called UELN-only solution. Except for 3 HBMs, this solution is identical to the EVRF2019 (release September 2020). The differences include a mean difference of 6.527 kgal mm at two British HBMs included in the EVRF2019.

Based on the comparison with the UELN-only solution, the impact of model-based hydrodynamic leveling in Experiment I has a long wave character and is significant for France and Great Britain where the differences range between  $-96.7$  and  $10.5$  kgal mm, and  $23.8$  and  $444.2$  kgal mm, respectively. The range of differences for the other countries is lower than  $10$  kgal mm. In Experiment II, the large-scale pattern is the same. However, the impact for both France and Great Britain is lower.

A comparison with normal heights obtained by differencing GNSS and EGG2015 quasi-geoid heights showed that in Experiment I for both France and Great Britain the sys-

tematic differences between both height data sets decreased. In Great Britain, the south-north slope disappeared almost completely after adding model-based hydrodynamic leveling data. In France, the decrease is mainly visible in the south-west.

Based on the results, we conclude that model-based hydrodynamic leveling has a positive impact on the EVRF. At the same time, and not unexpectedly, this impact does depend on the number and locations of tide gauges available for establishing the connections.

The results presented here show a provisional impact. Indeed, as indicated in Sect. 1, this study is subject to a number of limitations. First, in connecting the tide gauges to the UELN we mostly relied on potential differences computed from adjusted heights rather than real levelings. Second, we made no attempt to reduce the potential differences to the reference epoch adopted for the EVRS (i.e., epoch 2000.0). Third, we made no attempt to improve the performance of the hydrodynamic models in representing the local MWL at the tide gauge locations. In addition, to reduce the computational load no connections were allowed between tide gauges located i) within the same country, or ii) in neighboring countries for which the number of spirit leveling connections between the countries is larger than one. In particular for countries having a long coastline, it could be useful to relax these constraints. Finally, no connections have been added in the Mediterranean and Black Seas. Given the quality of the UELN data in the southern European countries, this would be a valuable addition.

All these aspects have to be addressed to achieve a rigorous implementation of model-based hydrodynamic leveling. However, this requires collaboration with experts who know which processes contribute to the (summer) mean water level variability at the tide gauge locations, and who have access to adequate models and datasets to model this variability. But also experts who can assess any vertical movement of the tide gauges and who can connect the tide gauges to the UELN. Undoubtedly, some things will not be realized overnight. Many data are not (publicly) available. A key dataset is leveling connections between the TGBMs and the nearest UELN HBMs. Likely, these connections have to be established which takes time and costs. At the same time, the prospect of having a technique that can connect islands and offshore platforms to the mainland height system and suppresses systematic errors in leveling networks makes the effort worthwhile.

Future research could also focus on using GNSS/leveling to establish (additional) connections between HBMs (separated by large water bodies). In addition to the required GNSS data, doing so requires the *full* noise variance-covariance matrix of the quasi-geoid model used. Obtaining the latter is probably the biggest challenge. Even if both are available,

the question remains how to validate the results, as there is no dataset left to validate the solution.

**Acknowledgements** This study was performed in the framework of the Versatile Hydrodynamics project, funded by the Netherlands Organization for Scientific Research (NWO). This support is gratefully acknowledged. The author VJ is funded by the Estonian Science Council grant PRG1785. Tide gauge data were kindly provided by the Agentschap Maritieme Dienstverlening en Kust, Belgium; Danish Coastal Authority; Danish Meteorological Institute; Danish Maritime Safety Administration; Service Hydrographique et Océanographique de la Marine, France; Die Wasserstraßen- und Schifffahrtsverwaltung des Bundes, Germany; Marine Institute, Ireland; Rijkswaterstaat, the Netherlands; Norwegian Hydrographic Service; Swedish Meteorological and Hydrological Institute; and U.K. National Tidal and Sea Level Facility (NTSLF). Part of this study has been conducted using E.U. Copernicus Marine Service Information; <https://doi.org/10.48670/moi-00059>. We thank the editor and three anonymous reviewers for their constructive remarks and corrections, which helped us to improve the quality of the manuscript.

**Author Contributions** CS proposed the initial idea. YA and CS designed the algorithm and experiments. YA, MS, VJ, and CS compiled and pre-processed the data. YA and CS developed the software and wrote the manuscript. All analyzed the results of the study and reviewed the manuscript.

**Data Availability** The hydrodynamic models 3D DCSM-FM is developed by Deltares. The Nemo-Nordic model is developed by the Swedish Meteorological and Hydrological Institutes. Neither of the hydrodynamic models is publicly available. The AMM7 model output is available at the Copernicus Marine Service (<https://doi.org/10.48670/moi-00059>). Tide gauge data are acquired from publicly accessible data services hosted by the national authorities. German tide gauge records were obtained upon request. EGG2015 model is publicly available ([https://www.isgeoid.polimi.it/Geoid/Europe/europe2015\\_g.html](https://www.isgeoid.polimi.it/Geoid/Europe/europe2015_g.html)). UELN and EUVN data are provided by the Federal Agency for Cartography and Geodesy (BKG). These data are not publicly available.

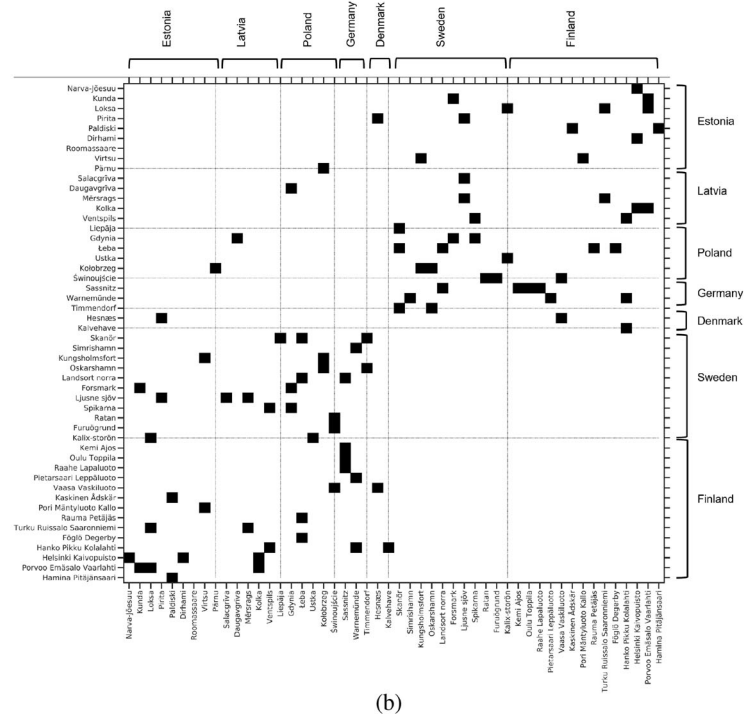
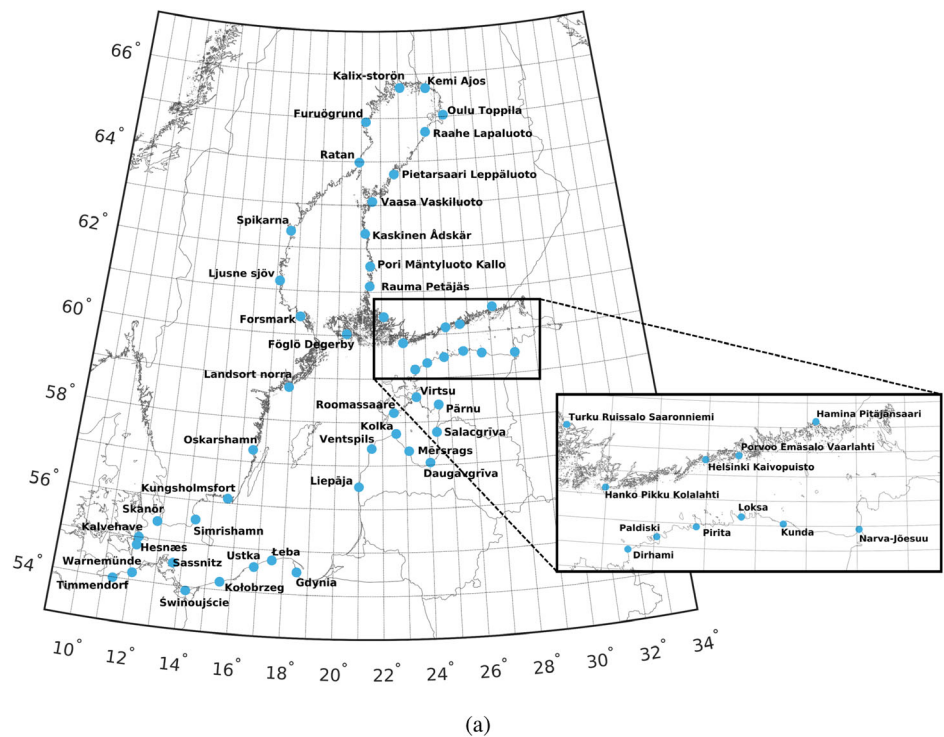
## Declarations

**Conflict of interest** All the authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.



**Fig. 11** Location of the tide gauges (a), and the connectivity matrix showing which model-based hydrodynamic leveling connections are included in the Nemo-Nordic dataset (b)

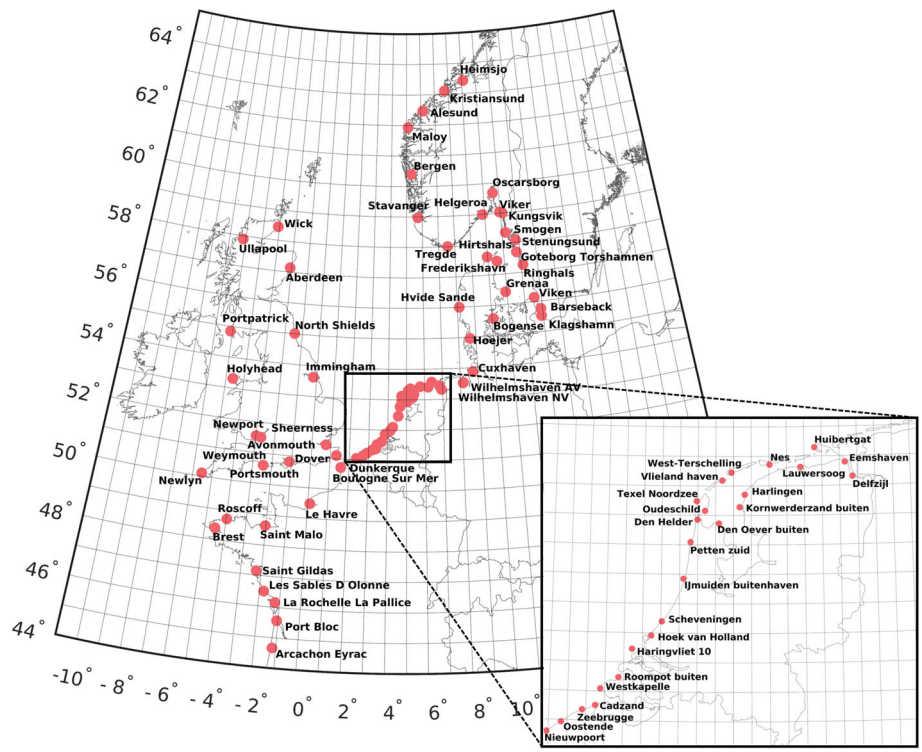


### Appendix A The hydrodynamic leveling connections being added

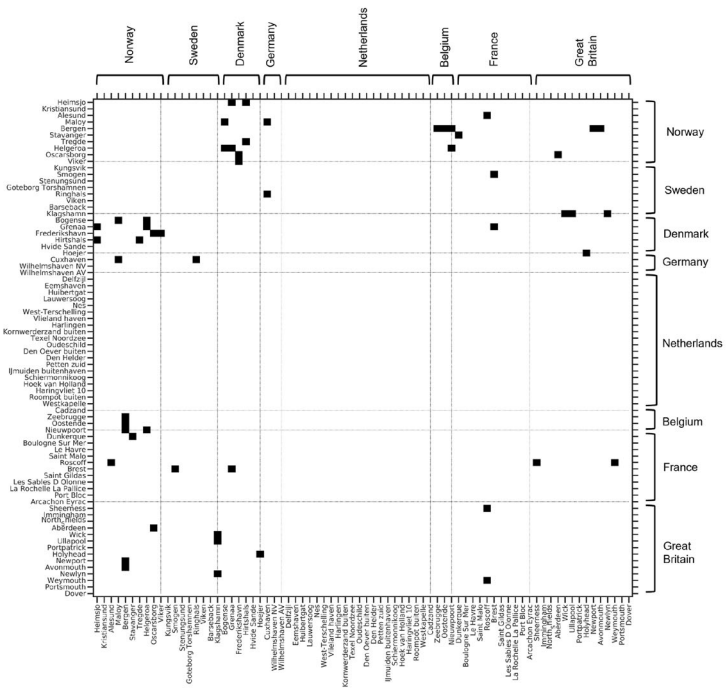
Figures 11b and 12b show the model-based hydrodynamic leveling connections included in, respectively, the Nemo-Nordic and 3D DCSM-FM hydrodynamic leveling datasets.

These connections are the ones being identified by our heuristic search algorithm described in Sect. 2.2. The maps showing the corresponding locations of the tide gauges are included in Figs. 11a and 12a.

**Fig. 12** Location of the tide gauges (a), and the connectivity matrix showing which model-based hydrodynamic leveling connections are included in the 3D DCSM-FM dataset (b)



(a)



(b)

## References

- Afrasteh Y, Slobbe DC, Verlaan M, Sacher M, Klees R, Guarneri H, Keyzer L, Pietrzak J, Snellen M, Zijl F (2021) The potential impact of hydrodynamic leveling on the quality of the European vertical reference frame. *J Geodesy*. <https://doi.org/10.1007/s00190-021-01543-3>
- Afrasteh Y, Slobbe DC, Verlaan M, Sacher M, Klees R, Guarneri H, Keyzer L, Pietrzak J, Snellen M, Zijl F (2023) An empirical noise model for the benefit of model-based hydrodynamic leveling. *J Geodesy* 97(1):1–21
- Amjadiparvar B, Rangelova E, Sideris MG (2015) The GBVP approach for vertical datum unification: recent results in North America. *J Geodesy* 90(1):45–63. <https://doi.org/10.1007/s00190-015-0855-8>
- Amos MJ, Featherstone WE (2008) Unification of New Zealand's local vertical datums: iterative gravimetric quasigeoid computations. *J Geodesy* 83(1):57–68. <https://doi.org/10.1007/s00190-008-0232-y>
- Axell L, Liu Y (2016) Application of 3-D ensemble variational data assimilation to a Baltic Sea reanalysis 1989–2013. *Tellus A: Dyn Meteorol Oceanog* 68(1):24,220. <https://doi.org/10.3402/tellusa.v68.24220>
- Bjerhammar A (1985) On a relativistic geodesy. *Bull géodésique* 59(3):207–220. <https://doi.org/10.1007/BF02520327>
- Bossler JD (1984) Standards and specifications for geodetic control networks. National Geodetic Information Branch, NOAA, FGCC
- Codiga DL (2020) UTide unified tidal analysis and prediction functions. <https://www.mathworks.com/matlabcentral/fileexchange/46523-utide-unified-tidal-analysis-and-prediction-functions>
- Cook RD, Weisberg S (1982) Residuals and influence in regression. Chapman and Hall, New York
- Denker H (2013) Regional gravity field modeling: theory and practical results. Springer, Berlin, Heidelberg, pp 185–291. [https://doi.org/10.1007/978-3-642-28000-9\\_5](https://doi.org/10.1007/978-3-642-28000-9_5)
- Denker H (2015) A new European Gravimetric (Quasi)Geoid EGG2015. 26th IUGG General Assembly, June 22–July 2, Prague, Czech Republic
- Denker H, Timmen L, Voigt C, Weyers S, Peik E, Margolis HS, Delva P, Wolf P, Petit G (2018) Geodetic methods to determine the relativistic redshift at the level of  $10^{-18}$  in the context of international timescales: a review and practical results. *J Geodesy* 92(5):487–516
- Duquenne F, Coulomb A, LÉcu F (2015) French approach to modernize its vertical reference system, FIG Working Week 2015, Sofia, Bulgaria, 17–21 May 2015. [http://www.fig.net/resources/proceedings/fig\\_proceedings/fig2015/papers/ts06g/TS06G\\_duquenne\\_coulomb\\_et\\_al\\_7676.pdf](http://www.fig.net/resources/proceedings/fig_proceedings/fig2015/papers/ts06g/TS06G_duquenne_coulomb_et_al_7676.pdf)
- Federal Agency for Cartography and Geodesy (2022a) EVRF2019. <https://evrs.bkg.bund.de/Subsites/EVRS/EN/EVRF2019/evrf2019.html>. Accessed 16 Nov 2022
- Federal Agency for Cartography and Geodesy (2022b) Height datum relations. <https://evrs.bkg.bund.de/Subsites/EVRS/EN/Projects/HeightDatumRel/height-datum-rel.html>. Accessed 16 November 2022
- Filmer MS, Featherstone WE, Kuhn M (2010) The effect of egm2008-based normal, normal-orthometric and Helmert orthometric height systems on the Australian levelling network. *J Geodesy* 84(8):501–513. <https://doi.org/10.1007/s00190-010-0388-0>
- Gerlach C, Rummel R (2012) Global height system unification with GOCE: a simulation study on the indirect bias term in the GBVP approach. *J Geodesy* 87(1):57–67. <https://doi.org/10.1007/s00190-012-0579-y>
- Heck B, Rummel R (1990) Strategies for solving the vertical datum problem using terrestrial and Satellite Geodetic Data. In: Sea surface topography and the geoid. Springer, New York, pp 116–128. [https://doi.org/10.1007/978-1-4684-7098-7\\_14](https://doi.org/10.1007/978-1-4684-7098-7_14)
- Hipkin R, Haines K, Beggan C, Bingley R, Hernandez F, Holt J, Baker T (2004) The geoid EDIN2000 and mean sea surface topography around the British Isles. *Geophys J Int* 157:565–577. <https://doi.org/10.1111/j.1365-246X.2004.01989.x>
- Hordoïr R, Axell L, Löptien U, Dietze H, Kuznetsov I (2015) Influence of sea level rise on the dynamics of salt inflows in the Baltic Sea. *J Geophys Res: Oceans* 120(10):6653–6668. <https://doi.org/10.1002/2014JC010642>
- Hordoïr R, Axell L, Höglund A, Dieterich C, Fransner F, Gröger M, Liu Y, Pemberton P, Schimanke S, Andersson H et al (2019) Nemo-Nordic 1.0: a NEMO-based ocean model for the Baltic and North seas—research and operational applications. *Geosci Model Develop* 12(1):363–386. <https://doi.org/10.5194/gmd-12-363-2019>
- Ihde J, Adam J, Gurtner W, Harsson BG, Sacher M, Schlüter W, Wöppelmann G (2000) The height solution of the European vertical reference network (EUVN). *Veröff Bayer Komm für die Internat Erdmessung, Astronom Geod Arb* 61:132–145
- Ihde J, Augath W, Sacher M (2002) The Vertical Reference System for Europe. In: International Association of Geodesy Symposia. Springer, Berlin, Heidelberg, pp 345–350. [https://doi.org/10.1007/978-3-662-04683-8\\_64](https://doi.org/10.1007/978-3-662-04683-8_64)
- Ihde J, Mäkinen J, Sacher M (2008) Conventions for the definition and realization of a European Vertical Reference System (EVRS)—EVRS Conventions 2007. *EVRS Conv*, pp 1–10
- Jahanmard V, Delpeche-Ellmann N, Ellmann A (2021) Realistic dynamic topography through coupling geoid and hydrodynamic models of the Baltic Sea. *Cont Shelf Res* 222(104):421. <https://doi.org/10.1016/j.csr.2021.104421>
- Jahanmard V, Delpeche-Ellmann N, Ellmann A (2022) Towards realistic dynamic topography from coast to offshore by incorporating hydrodynamic and geoid models. *Ocean Model* 180(102):124. <https://doi.org/10.1016/j.ocemod.2022.102124>
- Kenyeres A, Sacher M, Ihde J, Denker H, Marti U (2010) EUVN densification action—final report. Tech. rep., EUVN\_DA Working Group, [https://evrs.bkg.bund.de/SharedDocs/Downloads/EVRS/EN/Publications/EUVN\\_DA\\_FinalReport.pdf?\\_\\_blob=publicationFile&v=1](https://evrs.bkg.bund.de/SharedDocs/Downloads/EVRS/EN/Publications/EUVN_DA_FinalReport.pdf?__blob=publicationFile&v=1). Accessed 6 Oct 2022
- Madec G, Bourdallé-Badie R, Bouttier PA, Bruciau C, Bruciaferri D, Calvert D, Chanut J, Clementi E, Coward A, Delrosso D, et al (2016) NEMO ocean engine. Tech. rep., Notes du Pôle de modélisation de l'Institut Pierre-Simon Laplace (IPSL), (27). ISSN 1288-1619
- Mayer-Gürr T (2015) The combined satellite gravity field model GOCO05s. In: EGU general assembly conference abstracts, p 12364
- Moritz H (2000) Geodetic reference system 1980. *J Geod* 74(1):128–162. <https://doi.org/10.1007/s001900050278>
- Penna NT, Featherstone WE, Gazeaux J, Bingham RJ (2013) The apparent British sea slope is caused by systematic errors in the levelling-based vertical datum. *Geophys J Int* 194(2):772–786. <https://doi.org/10.1093/gji/ggt161>
- Rao C (1971) Estimation of variance and covariance components—minique theory. *J Multiv Anal* 1(3):257–275. [https://doi.org/10.1016/0047-259X\(71\)90001-7](https://doi.org/10.1016/0047-259X(71)90001-7)
- Rousseeuw PJ, Croux C (1993) Alternatives to the median absolute deviation. *J Am Stat Assoc* 88(424):1273–1283. <https://doi.org/10.1080/01621459.1993.10476408>
- Rülke A, Liebsch G, Sacher M, Schäfer U, Schirmer U, Ihde J (2012) Unification of European height system realizations. *J Geod Sci* 2(4):343–354. <https://doi.org/10.2478/v10156-011-0048-1>
- Rummel R, Teunissen P (1988) Height datum definition, height datum connection and the role of the geodetic boundary value problem. *Bulletin Géodésique* 62(4):477–498. <https://doi.org/10.1007/bf02520239>

- Sacher M, Liebsch G (2019) EVRF2019 as new realization of EVRS. [https://evrs.bkg.bund.de/SharedDocs/Downloads/EVRS/EN/Publications/EVRF2019\\_FinalReport.pdf?\\_\\_blob=publicationFile&v=2](https://evrs.bkg.bund.de/SharedDocs/Downloads/EVRS/EN/Publications/EVRF2019_FinalReport.pdf?__blob=publicationFile&v=2)
- Sánchez L, Sideris MG (2017) Vertical datum unification for the international height reference system (IHRs). *Geophys J Int* 209(2):570–586
- Slobbe DC, Klees R, Verlaan M, Zijl F, Alberts B, Farahani HH (2018) Height system connection between island and mainland using a hydrodynamic model: a case study connecting the Dutch Wadden islands to the Amsterdam ordnance datum (NAP). *J Geod* 92(12):1439–1456. <https://doi.org/10.1007/s00190-018-1133-3>
- Slobbe DC, Klees R, Farahani HH, Huisman L, Alberts B, Voet P, De Doncker F (2019) The impact of noise in a GRACE/GOCE global gravity model on a local quasi-geoid. *J Geophys Res: Solid Earth* 124(3):3219–3237. <https://doi.org/10.1029/2018jb016470>
- Tonani M, Ascione I (2021) Product user manual: ocean physical and biogeochemical reanalysis NWSHELF\_MULTIYEAR\_PHY\_004\_009 NWSHELF\_MULTIYEAR\_BGC\_004\_011. <https://catalogue.marine.copernicus.eu/documents/PUM/CMEMS-NWS-PUM-004-009-011.pdf>
- Vaniček P, Castle RO, Balazs EI (1980) Geodetic leveling and its applications. *Rev Geophys* 18(2):505–524. <https://doi.org/10.1029/RG018i002p00505>
- Vestøl O, Ågren J, Steffen H, Kierulf H, Tarasov L (2019) NKG2016LU: a new land uplift model for Fennoscandia and the Baltic Region. *J Geodesy* 93(9):1759–1779. <https://doi.org/10.1007/s00190-019-01280-8>
- Wang YM, Saleh J, Li X, Roman DR (2012) The US Gravimetric Geoid of 2009 (USGG2009): model development and evaluation. *J Geodesy* 86(3):165–180. <https://doi.org/10.1007/s00190-011-0506-7>
- Wu H, Müller J, Lämmerzahl C (2018) Clock networks for height system unification: a simulation study. *Geophys J Int* 216(3):1594–1607. <https://doi.org/10.1093/gji/ggy508>
- Ziebart MK, Illiffe JC, Forsberg R, Strykowski G (2008) Convergence of the UK OSGM05 GRACE-based geoid and the UK fundamental benchmark network. *J Geophys: Res Solid Earth*. <https://doi.org/10.1029/2007JB004959>
- Zijl F, Laan S, Groenenboom J (2020) Development of a 3D model for the NW European Shelf (3D DCSM-FM). <https://www.deltares.nl/app/uploads/2020/12/Development-of-a-3D-model-for-the-NW-European-Shelf-3D-DCSM-FM.pdf>