Annals of Glaciology

Article

Cite this article: Puggaard A, Hansen N, Schlager E, Sørensen LS, Bjerregaard Simonsen S (2025) Filling the GRACE/-FO gap of mass balance observation in central west Greenland by data-driven modelling. *Annals of Glaciology* **66**, e20, 1–11. https://doi.org/10.1017/aog.2025.10019

Received: 19 February 2025
Revised: 2 June 2025
Accepted: 28 July 2025

Keywords:
Greenland Ice Sheet; remote sensing; GRACE; deep learning

Corresponding author: Anna Puggaard;
Email: annpu@space.dtu.dk

cambridge.org/aog

Filling the GRACE/-FO gap of mass balance observation in central west Greenland by data-driven modelling

Anna Puggaard[1,2], Nicolaj Hansen[2], Elke Schlager[2,3], Louise Sandberg Sørensen[1] and Sebastian Bjerregaard Simonsen[1]

[1]Geodesy and Earth Observation, DTU Space, Technical University of Denmark, Kongens Lyngby, Denmark; [2]National Center for Climate Research, Danish Meteorological Institute, Copenhagen, Denmark and [3]Department of Environmental Science, Aarhus University, Roskilde, Denmark

Abstract

Continuous monitoring of the mass balance of the Greenland ice sheet is crucial to assess its contribution to the rise of sea levels. The GRACE and GRACE-FO missions have provided monthly estimates of the Earth's gravity field since 2002, which have been widely used to estimate monthly mass changes of ice sheets. However, there is an 11 month gap between the two missions. Here, we propose a data-driven approach that combines atmospheric variables from the ERA5 reanalysis with GRACE-derived mass anomalies from previous months to predict mass changes. Using an auto-regressive structure, the model is naturally predictive for shorter times without GRACE/-FO observations. The results show a high r2-score (> 0.73) between model predictions and GRACE/-FO observations. Validating the model's ability to reproduce mass anomalies when observations are available builds confidence in estimates used to bridge the GRACE and GRACE/-FO gap. Although GRACE and GRACE-FO are treated equally by the model, we see a decrease in model performance for the period covered by GRACE-FO, indicating that they may not be as well-calibrated as previously assumed. Gap predictions align well with mass change estimates derived from other geodetic methods and remain within the uncertainty envelope of the GRACE-FO observations.

1. Introduction

The Greenland ice sheet has lost mass at increasing rates since the early 1990s, caused by an increased discharge of ice to the ocean through marine-terminating outlet glaciers and a decrease in surface mass balance (SMB); the mass loss is highly variable, with maximum peaks about 2010–12 and 2019 caused by increased melting and runoff (Tedesco and Fettweis, 2020; Otosaka and others, 2023). In this study, we focus on the central west (CW) basin of the Greenland ice sheet as defined by Zwally and others 2012, as this basin holds 130 cm equivalent sea level (Mouginot and others, 2019). The CW basin includes several large and highly dynamic glaciers, including Sermeq Kujalleq (Jakobshavn isbræ) and Kangilliup Sermia (Rink isbræ), which all play an important role in the regional mass balance (MB) (Khazendar and others, 2019; Mouginot and others, 2019; Joughin and others, 2020). Sermeq Kujalleq has naturally received much attention, being the largest outlet of the Greenland ice sheet. Still, the entire CW exhibits considerable variability; Mouginot and others 2019 showed that this basin was in equilibrium from the 1970s until the early 2000s, when it started to accelerate. From 2000 to 2018, the basin lost more than 740 Gt of mass, corresponding to 21% of the entire Greenland mass loss (Mouginot and others, 2019; Shepherd and others, 2020). In 2017 and 2018, the basin experienced a deceleration in mass loss, primarily driven by a deceleration and thickening of Sermeq Kujalleq (Khazendar and others, 2019), followed by a record high mass loss of 2019 (Sasgen and others, 2020; Velicogna and others, 2020).

The MB of the ice sheets is continuously monitored using three geodetic methods: Gravimetry (changes in the gravitational field), altimetry (changes in the surface elevation) and the input–output method (IOM), which estimates MB by comparing SMB (input) with solid ice discharge (output) (Otosaka and others, 2023). The Ice Sheet Mass Balance Intercomparison Exercise (IMBIE) has provided comprehensive assessments of different MB monitoring approaches by systematically comparing and reconciling results from gravimetry, satellite altimetry and the IOM. While IMBIE found overall good agreement between the distinct approaches, the assessment also revealed variability among them, highlighting the use of different geophysical corrections, SMB models and glacial isostatic adjustment (GIA) models (Shepherd and others, 2020; Otosaka and others, 2023). This study focuses on mass changes derived from gravity variations measured by the Gravity Recovery and Climate Experiment (GRACE) satellites and their successor, GRACE Follow-On (GRACE-FO). However, the 1 year

data gap between the conclusion of the GRACE mission and the start of GRACE-FO poses a significant challenge for ensuring continuous MB monitoring. To address this, several attempts have been made using independent methods (Sasgen and others, 2020), including in situ observation from the Greenland GNSS network (Barletta and others, 2024a) and traditional techniques for monitoring mass changes (Otosaka and others, 2023). We aim to provide an alternative to existing methods to bridge the gap between GRACE and GRACE-FO that is consistent with GRACE/-FO observations of mass anomalies.

Recent advances in deep learning techniques have offered various opportunities for ice sheet MB monitoring and modelling. Among these, physics-informed neural networks (PINNs) have proven to be a powerful tool for modelling ice sheet evolution by integrating observational data with physical laws to enhance computational efficiency while maintaining adherence to physical principles (Jouvet and others, 2022; Jouvet and Cordonnier, 2023). In addition to improving computational efficiency, PINNs have proven valuable for inferring basal conditions, such as subglacial topography and basal friction, which are challenging to measure directly (Bolibar and others, 2023; Cheng and others, 2024). For SMB, deep learning methods have been successfully utilized to downscale SMB models (van der Meer and others, 2023). By integrating high-resolution remote sensing data with deep learning techniques, these methods can capture fine-scale spatial variations in SMB processes, such as melt patterns, that are otherwise missed in coarse-resolution models (De Roda Husman and others, 2024). Furthermore, deep learning techniques have proven helpful in automatically detecting supraglacial lake evolution and melt from remote sensing data (Lutz and others, 2023; Zhu and others, 2024).

Additionally, machine learning methods have shown potential for bridging the gap between GRACE and GRACE-FO (Zhang and others, 2022; Shi and others, 2024). Shi and others (2024) use an SVM model based on climate model outputs to reconstruct a gridded GRACE/-FO product, while Zhang and others (2022) apply a nonlinear neural network to integrate both climate model output and ice discharge at the basin scale. Both methods successfully capture the annual MB but struggle to accurately resolve seasonal variations. In this study, we integrate daily atmospheric conditions from the European Centre for Medium-Range Weather Forecasts global atmospheric reanalysis dataset v5 (ERA5) with a history of GRACE/-FO-derived mass anomalies to better capture short-term atmospheric variability and improve the representation of the different drivers of ice mass change. By using daily ERA5 variables instead of monthly aggregates, we aim to retain information of short-term variability, e.g. short-lived but intense precipitation events, which are otherwise smoothed out in the monthly aggregates. The data-driven model provides gridded mass anomalies that are consistent with GRACE/-FO spatiotemporal variability.

In this study, we focus on the CW basin of the Greenland ice sheet. All GRACE/-FO-derived mass anomalies are subject to some degree of spatial leakage effects from neighbouring regions. The northern drainage basins, in particular, are affected not only by leakage from adjacent Greenland drainage basins but also by mass changes in the nearby Canadian Arctic glaciers due to their geographic proximity (Baur and others, 2009; Barletta and others, 2013). Likewise, if we choose a basin further south, the basins are more narrow, which can also cause problems with GRACE/-FO due to the coarse spatial resolution of GRACE/-FO. Therefore, the CW basin is a good and interesting region for this case study. The proposed model addresses gaps in the mass anomaly time series between GRACE and GRACE-FO with an auto-regressive model, enabling it to predict monthly mass anomalies based on past observations and atmospheric drivers. During the gap period, when GRACE observations are unavailable, the model feeds its own predicted mass anomalies from the previous time step into subsequent predictions in place of the missing GRACE data. With the approach, we provide monthly mass anomalies that are consistent with GRACE/-FO observations and exhibit seasonal variability similar to other MB approaches. Therefore, we can provide a 21 year continuous time series of mass anomalies.

2. Data

We use monthly GRACE/-FO mass anomaly observations (Barletta and others, 2013) to provide a historical record of the total MB from 2002 to 2023. Within the same period, we use daily ERA5 variables (Hersbach and others, 2020) to describe the atmospheric/surface components of the total MB observed by GRACE/-FO. Basin-scale solid ice discharge estimates (Mankoff and others, 2020) are included to derive the basin SMB, used exclusively to deep-constrain the model to fit basin SMB during training. The data-driven model estimates are compared to two geodetic datasets derived from different approaches: the IOM Mankoff and others (2021) and altimetry (Khan and others, 2025). All data will focus on the CW basin, defined as basin no. 7 in Zwally and others 2012, and we describe the data in further detail in the following sections.

2.1. ERA5 surface variables

We use daily means of the ERA5 global atmospheric reanalysis (Hersbach and others, 2020). ERA5 has a spatial resolution of 0.25°, corresponding to 30 km over the Greenland ice sheet, with 137 vertical levels. We include only the surface variables that primarily control the SMB: the 2 m temperature, total precipitation, surface pressure and short-wave and long-wave downward surface radiation. We aggregate the hourly ERA5 data into daily values since this temporal resolution should be enough to predict the monthly mass changes on the same grid as GRACE/-FO.

2.2. GRACE gravimetric MB

GRACE and GRACE-FO are dedicated missions to map temporal and spatial variations in Earth's gravity field. GRACE, launched in 2002 and operational until 2017, utilized a pair of satellites to monitor changes in the distance between them caused by variations in Earth's gravitational pull. These changes reflect mass redistributions in the Earth system due to, e.g. ice sheet mass changes (Velicogna and Wahr, 2005; Tapley and others, 2019). GRACE-FO, launched in 2018, continues this legacy, although leaving a gap in the observations between October 2017 and June 2018 between the two missions.

Here, we use a dataset of gravimetric ice sheet mass changes (Barletta and others, 2024b), which are based on GRACE and GRACE-FO monthly solutions (CSR, RL06.2) using the point-mass inversion method described in Barletta and others 2013. In Barletta and others (2024b), the C20 and C30 coefficients are replaced, and degree-1 coefficients have been added according to the release centre technical notes (Swenson and others, 2008; Sun and others, 2016; Landerer, 2024). The dataset offers both gridded estimates of 22 km disks and integrated drainage basin estimates. For the basin mass changes, estimates of the uncertainty are provided, accounting for the propagation of formal errors from the

L2 GRACE/-FO data, uncertainty in degree-one terms, GIA corrections (Caron and others, 2018) and uncertainties in ocean and atmospheric models (Barletta and others, 2013).

2.3. Surface mass balance

Simplified, the MB of ice sheets is the difference between SMB and solid ice discharge (D). Hence, by combining the GRACE/-FO-derived mass anomaly observations over the basin and the solid ice discharge, we can derive the SMB as,

$$SMB = MB + D \tag{1}$$

Here, D is the solid ice discharge. We do not account for the contribution of mass loss from basal processes, e.g. melting at the base, as the MB in the CW basin is mainly driven by SMB and solid ice discharge (Shepherd and others, 2020; Karlsson and others, 2021).

Solid ice discharge estimates between 2002 and 2023 are calculated from the ice velocity and thickness of fast-flowing glaciers (Mankoff and others, 2020). Ice velocity post-2000 is derived from a combination of PROMICE (Solgaard and others, 2021) and MEaSUREs (Howat and Ohio State University, 2017) ice velocity datasets. The ice thickness is derived from surface elevation (Howat and others, 2014) and bedrock elevations (Morlighem and others, 2017), with adjustments over time based on changes in surface elevation (Khan and others, 2016). Flux gates are selected automatically using a threshold of 100 m yr^{-1} since SMB dominates outlet glaciers below this threshold. Finally, the solid ice discharge is calculated per pixel along the flux gate using the density of ice (Mankoff and others, 2020).

2.4. Data for inter-comparison

For inter-comparison purposes, we apply two geodetic MB records: The daily mass changes estimates from the IOM (Mankoff and others, 2021) and the monthly mass changes derived from altimetry (Khan and others, 2022). Mankoff and others (2021) computes the SMB within the GRACE/-FO period using the average of three regional climate models: HIRHAM, RACMO and MAR. The solid ice discharge estimates are from Mankoff and others (2020). To improve the MB estimates of the IOM, Mankoff and others (2021) includes the basal MB term in the MB. We refer to Mankoff and others (2021) for a more detailed description of the IOM dataset. However, it is relevant to note that the solid ice discharge dataset included in the Mankoff and others (2021) dataset is also used for deep supervision during model training in this study. Thus, Mankoff and others (2021) is not a completely independent dataset.

Khan and others 2025 estimate the mass changes of the Greenland Ice Sheet from 2003 to 2023 using a combination of airborne and satellite altimetry data, including measurements from CryoSat-2, Envisat, ICESat, ICESat-2 and Operation IceBridge. Elevation changes are interpolated onto a regular grid using kriging, and elevation changes are converted to mass changes using the density from RACMO RCM (Noël and others, 2019). To correctly convert volume to mass, firn compaction is accounted for using a simple firn model that includes melt and refreezing. As the data product is based on altimetry elevation changes, it should be noted that this method cannot produce rapid changes but rather a temporally "smoothed" data product. Additionally, the regression procedure of Khan and others 2025 imposed a strong inter-annual cycle in MB.

3. Methods

The proposed neural network architecture can handle the different spatiotemporal resolutions of ERA5 and GRACE/-FO. To ensure faster and more stable learning, we scale all data before feeding it to the model. Furthermore, we split data into training and testing sets to evaluate the model's performance on unseen data and prevent overfitting.

3.1. Preprocessing

For each GRACE/-FO solution, a data cube of 30 days of ERA5 daily data leading up to the midpoint of the GRACE/-FO solution is created. Together with the previous GRACE solution at time $T-1$, the data cube of the 30 days of ERA5 data between time $T-1$ and T is used to predict the mass anomaly at time T. It is important to note that the GRACE solutions are irregularly sampled in time. While most have an interval of roughly 30 days, longer gaps occasionally occur. Due to model architecture constraints, we only sample 30 days of ERA5 data for each GRACE solution, even when the GRACE observational interval exceeds 30 days. However, $<7\%$ of the GRACE/-FO solutions have a time interval >30 days.

To ensure fast and stable learning, input data are linearly scaled to a similar order of magnitude. Since the different ERA5 variables and GRACE are on different scales, the larger ranges can disproportionately influence the model's learning rate, leading to biased predictions and poor generalization. We select the scaling strategy based on the distribution of each input variable prior to scaling. For the monthly GRACE/-FO mass anomalies and daily ERA5 temperature and long-wave downward radiation, we apply a z-score normalization, which removes the mean and linearly scales to unit variance. It is appropriate to apply a z-score normalization in this case, as all three variables exhibit approximately normal distributions prior to scaling. The daily ERA5 surface pressure, short-wave downward radiation and precipitation are linearly scaled between zero and one based on the minimum and maximum values. Unlike the z-score normalization, this scaling method preserves the effect of outliers after scaling. Figure S1 in the Supporting Information illustrates the data distributions before and after scaling.

Data are divided into three subsets: training, validation and testing. The testing dataset comprises data from 2009 up to and including 2011, which corresponds to 15% of the dataset. The testing period is intentionally chosen to be in the middle of the GRACE period rather than at the beginning or end, as the GRACE solutions generally have lower uncertainty. In the middle of the GRACE mission period, data are also consistently available monthly, which is not the case towards the end of the mission, where data availability becomes more irregular. Having reliable data for model testing ensures that any differences between GRACE and model predictions stem from the model's performance rather than biases in the GRACE observations. Furthermore, the period also includes the extreme melt year of 2010, leading to a high SMB-driven mass loss (Tedesco and others, 2011). The remaining data are randomly split into training and validation sets in an 80:20 ratio. A random split is used instead of a consecutive split to ensure that most periods are represented in the training dataset to account for the large variability in the mass loss observed over the past two decades (Khazendar and others, 2019). We further explore the ability to bridge the gap between the GRACE and GRACE-FO missions. To do so, we create 30 day timestamps to build ERA5 data cubes. Since the previous GRACE solution is unavailable during the gap period, we use an auto-regressive approach where we take the previously predicted mass anomalies instead.

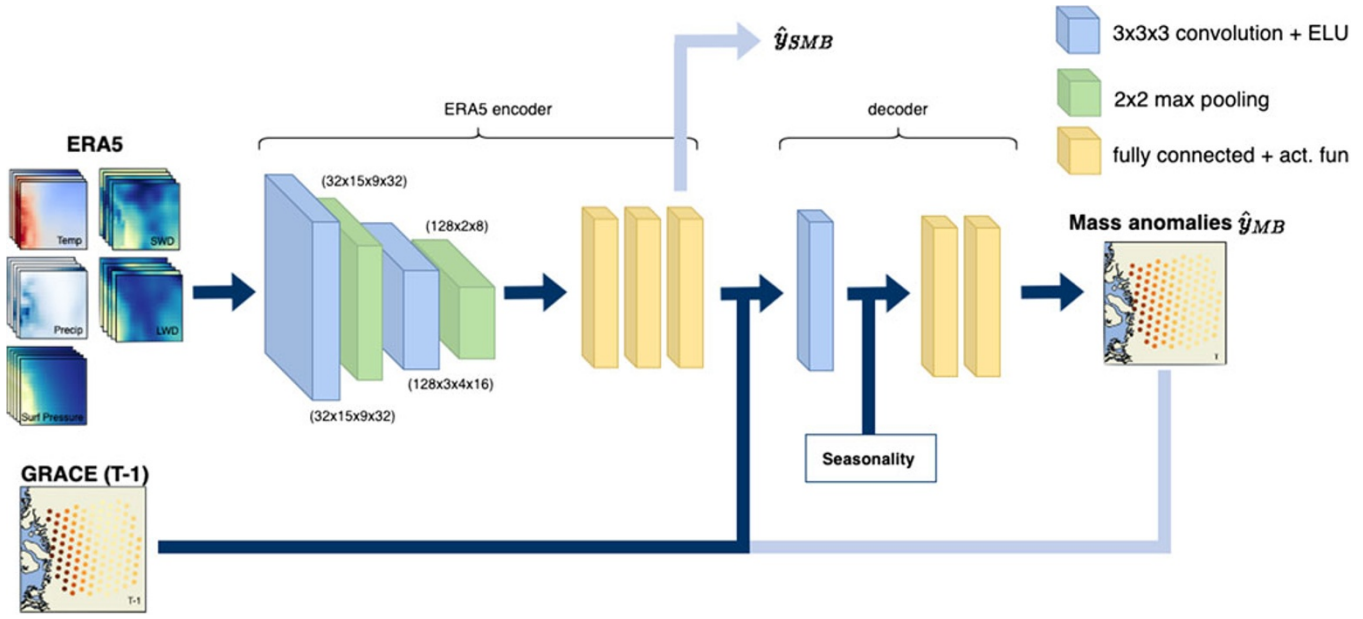

Figure 1. Proposed model architecture consisting of an encoder and decoder. The encoder handles the daily ERA5 data, while the decoder combines the encoded ERA5 data with the previous GRACE/-FO observation and seasonal information to predict the mass anomalies.

3.2. Deep learning architecture

The proposed neural network consists of an ERA5 encoder and a decoder to handle the different spatiotemporal resolutions of ERA5 and GRACE/-FO; see Fig. 1. The ERA5 encoder consists of a CNN with two convolutional layers, each with a kernel size of $3 \times 3 \times 3$ with a 2-pixel stride and the exponential linear unit (elu) as an activation function. The elu activation function was primarily chosen to allow small negative outputs. Between each convolutional layer, we apply a three-dimensional MaxPooling layer and dropout ($p = 0.2$). Then, we apply three fully connected (fc) layers, where the last fc layer reduces the spatial dimension to match GRACE/-FO. The outcome of this encoding step produces feature maps representing the SMB of the basin at time T, being constrained by the SMB derived from GRACE/-FO and the solid discharge. We note that the feature maps outputted from the ERA5 encoder are only used to constrain the model during training and are not regarded as an output once the model is trained. The next step is to merge the GRACE/-FO solution at $T-1$ with the encoded ERA5 data using a convolutional layer before concatenating the seasonal cycle and applying two fc layers to estimate the mass anomaly of the next time step. We represent the seasonality using a sinusoidal transformation. By representing the day of the year as the sine and cosine of the day, it ensures that the cyclical nature of time is preserved. Again, the elu activation function is applied for the convolutional layer, but a hyperbolic tangent function is applied to the first fc layer as it yielded the best performance, and no activation function is applied to the last fc layer to allow negative values smaller than −1.

3.2.1. Training

During training, the Adaptive Moment Estimation (Adam) (Kingma and Ba, 2015) is employed, starting with a learning rate of 10^{-3}. Adam uses an adaptive learning rate by scaling updates for each parameter based on the mean and variance of past gradients. Models are trained with a batch size of 20 for 150 epochs, saving the ten models with the lowest validation loss. We experimented with alternative hyperparameters, including learning rate, loss function and optimiser, but the presented configuration yielded the best results.

A multipart loss function constrains the model:

$$\mathcal{L} = \mathcal{L}_{SMB} + \mathcal{L}_{MB}\,, \tag{2}$$

where $\mathcal{L}_{SMB}$ ensures that the SMB representation from the ERA5 decoder follows the SMB trends on basin scale:

$$\mathcal{L}_{SMB} = \frac{1}{n_1}\sqrt{\sum_{i=1}^{t}\frac{((\sum_{j=1}^{n_1}\hat{y}_{SMB,ij}) - Y_{SMB,i})^2}{t}} \tag{3}$$

where $\hat{y}_{SMB,ij}$ is the output of the ERA5 decoder before merging with GRACE, see Fig. 1, and have the same spatial resolution of ERA5. n_1 is the number of grid points of $\hat{y}_{SMB,ij}$ and $Y_{SMB,i}$ is the basin SMB derived from GRACE/-FO mass anomalies and solid ice discharge using eq. (1). Note that the SMB includes contributions from basal MB, as we do not remove the signal from basal processes in the GRACE/-FO solutions. $\mathcal{L}_{MB}$ is computed on the mass anomalies output of the model:

$$\mathcal{L}_{MB} = \sum_{i=1}^{t}\sum_{j=1}^{n_2}\frac{(\hat{y}_{MB,ij} - y_{MB,ij})^2}{n_2} + \frac{1}{n_2}\sum_{i=1}^{t}\frac{((\sum_{j=1}^{n_1}\hat{y}_{MB,ij}) - Y_{MB,i})^2}{2\sigma^2} \tag{4}$$

where $y_{MB,ij}$ is the GRACE/-FO-derived mass anomalies, $\hat{y}_{MB,ij}$ is the predicted mass anomalies on the same grid as GRACE/-FO and $Y_{MB,i}$ is the basin mass anomalies measured by GRACE/-FO. n_2 is the number of grid points of GRACE/-FO grid. The first term of $\mathcal{L}_{MB}$ computes the MSE between predicted mass anomalies and mass anomalies observed by GRACE/-FO. For basin-scale mass anomalies, we have the associated uncertainty for the GRACE observation; see Fig. 4. Thus, we can incorporate the uncertainty into the loss function for the second term. This means that when the uncertainty is high, the errors between predictions and observations are given less weight.

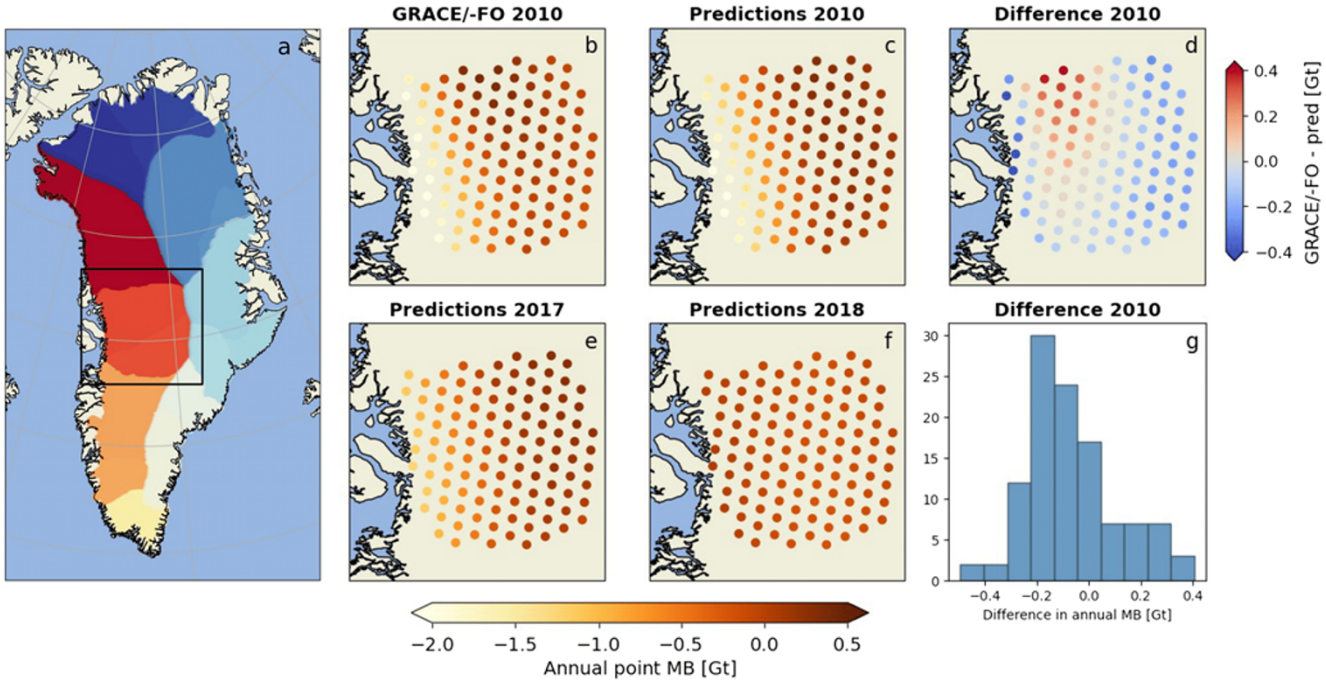

Figure 2. Overview of basin definitions by Zwally and others 2012 (a) with a focus on CW basin and annual point MB estimates. For 2010, the GRACE observations from Barletta and others 2013 (b) and the model predictions (c) are compared in (d) with distributions of the differences shown in (g). (e) and (f) show the MB of 2017 and 2018. 2010 corresponds to one of the years in the test dataset, and both 2017 and 2018 correspond to the GRACE/-FO gap. Scatter point MB estimates by the model are on the same resolution as GRACE/-FO with a disk radius of 22 km.

We tested different strategies for weighting $\mathcal{L}_{SMB}$ and $\mathcal{L}_{MB}$ in eq. (3), including assigning different weights as well as adaptive weighting using SoftAdapt (Heydari and others, 2019) and GradNorm (Chen and others, 2018). However, equal weights but ensuring equal units of $\mathcal{L}_{SMB}$ and $\mathcal{L}_{MB}$ yielded the best results.

3.2.2. Auto-regression

We only train the data-driven model when GRACE/-FO is available. GRACE solutions are provided as input into the model until 10 June 2017, as indicated by the dark blue arrow in Fig. 1. When the GRACE mission ends, and then the model operates in an auto-regressive manner, shown by the light blue arrow in Fig. 1. This means the mass anomaly output of the previous time step is fed into the model as input for the next prediction step instead of the observed GRACE solutions. This allows for the MB reconstructions during the observational gap between GRACE and GRACE-FO. To make a realistic evaluation of the reliability of the auto-regressive predictions, we also apply the auto-regressive model during the test period, where GRACE observations are available for comparison.

4. Results

Due to the auto-regressive nature of the model, we can provide monthly mass anomaly estimates both when GRACE/-FO solutions are available and when they are not. In Fig. 2, we compute the annual MB of three different years (2010, 2017 and 2018) based on monthly estimates by the model. The annual MB in 2010 (Fig. 2c) is part of the testing dataset, whereas both 2017 (Fig. 2e) and 2018 (Fig. 2f) are partially within the GRACE/-FO gap. For 2010, the reported MB by Barletta and others 2013 is illustrated in Fig. 2b, and the difference between the model predictions and GRACE/-FO observations is shown in Fig. 2d. Fig. 2d shows a slightly negative bias in 2012, meaning that the model predicts a slightly greater MB (0–0.2 Gt). Only in the northern part of the does GRACE observe a higher MB in 2010 (Fig. 2d). The GRACE solutions are provided as input into the model until 10 June 2017, when the GRACE mission ends, and then the model operates in an auto-regressive manner. Both 2010 and 2018 show a greater mass loss at the margin of the ice sheet compared to 2017, corresponding with the reported decline of discharge of the marine-terminating glaciers of the CW basin during 2017 (Mankoff and others, 2021). For the CW basin, 2017 was also a low melt year (Tedesco and others, 2017), followed by an increase in melt in 2018 compared to the previous year (Tedesco and others, 2018).

Figure 3 shows the observed GRACE/-FO mass anomalies and predicted mass anomalies by the deep learning model. The plot showcases both the testing dataset (2009–11), marked as unfilled points, and the training dataset, marked as filled points. For the training data, we further differentiate between the GRACE (dark blue) and GRACE-FO missions (light blue). The training data have RMSE of 15.04 Gt and r^2 of 0.83, while the test dataset have RMSE of 14.61 Gt and r^2 of 0.87. It is important to note that the testing data only includes data from the GRACE as input, whereas the training data includes both GRACE and GRACE-FO data. For the GRACE period only, the model predictions have a RMSE of 14.54 Gt and r^2 of 0.99, while the model performs worse in the GRACE-FO period with a RMSE of 19.70 Gt and r^2 of 0.73.

Figure 4 presents the GRACE/-FO mass anomalies with associated uncertainties and the model-predicted mass anomalies. Both training and testing periods are included, with the testing periods highlighted as the vertical grey areas. The model predictions generally fall within the uncertainty range of the GRACE/-FO observations for both training and testing periods. Figure 4 also

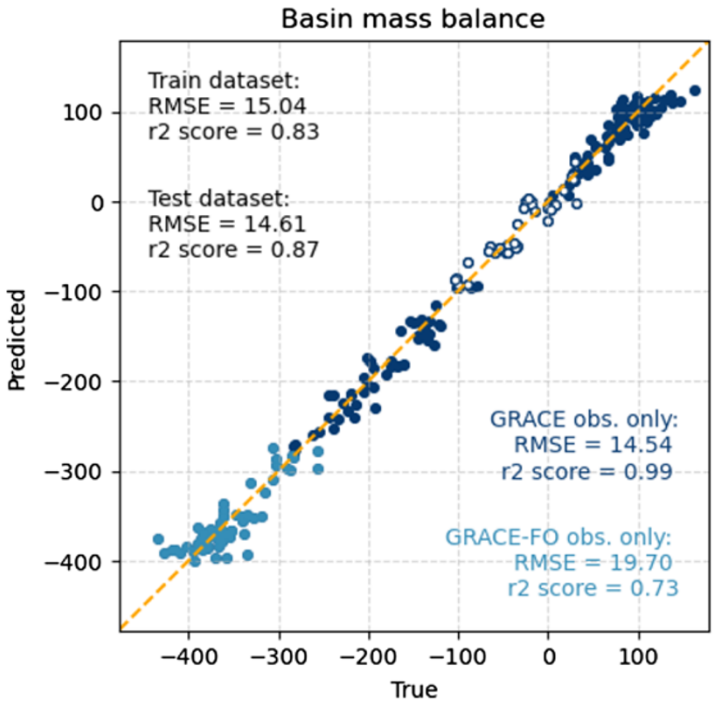

Figure 3. Predicted mass anomalies versus GRACE/-FO-derived mass anomalies with RMSE and r^2-score. Dark blue is within the GRACE period and light blue is within the GRACE-FO period. Unfilled points indicate the testing period, which is included in the GRACE period. RMSE and r^2-scores are shown for both training data and testing data. RMSE and r^2-score for the training period are also divided into GRACE and GRACE-FO period.

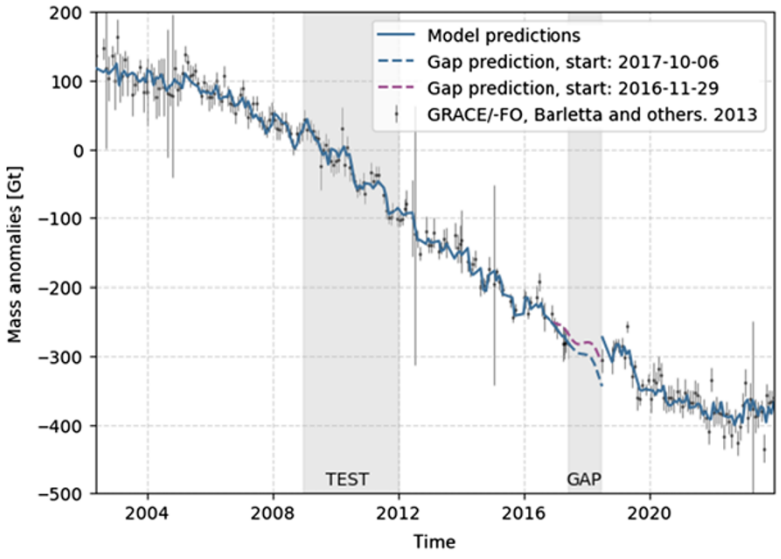

Figure 4. Mass anomalies for the CW basin. The black dots are the GRACE observation with uncertainties (1 σ) as lines. The first GRACE-FO solution is excluded due to a short baseline. The vertical grey areas illustrate the two testing cases: one where data are not included in training the model and the GRACE/-FO gap. The blue is the mass anomalies estimated by the model. The two dotted lines show the result of the auto-regression in the gap between GRACE and GRACE-FO, both including (blue) and excluding (green) GRACE data from early 2017. Figure S2 in the Supporting Information shows the auto-regression over the gap but excludes 2, 4 and 6 GRACE solutions before the end of the mission.

shows the auto-regressive model applied to the GRACE/-FO gap twice: Once where the auto-regression starts on 29 November 2016 and once where it starts on 10 June 2017, the last available GRACE solution. 29 November 2016 is chosen based on Figure S2 in the Supporting Information, where we explore different starting times for the auto-regression. When the auto-regression begins on 10 June 2017, the predicted mass anomalies are about 30 Gt lower than what GRACE-FO measures at the start of its record. For the gap prediction initiated on 29 November 2016, the model predicts mass anomalies in accordance with the GRACE-FO uncertainty envelope.

We compare the annual MB of the CW basin in Fig. 5. Here we use a hydrological year, running from October to September. The data-driven annual MB shows good agreement with GRACE/-FO observations (correlation coefficient = 0.79). The correlation with altimetry-derived MB is lowest (correlation coefficient = 0.68), and the correlation with the IOM MB is the highest (correlation coefficient = 0.85). However, the data-driven estimates are consistently higher annual MB compared to the IOM, showing a high offset between the two. The GRACE/-FO in Fig. 5d and e also shows a lower correlation between the annual MB from IOM and altimetry. The correlation between altimetry and IOM altimetry in Fig. 5 is high but also shows a similar high offset between the two methods.

Figure 6 shows the computed mass changes over the basin compared with IOM and altimetry. Since the altimetry-derived mass changes natively is a smoothened dataset, we apply a 4 month running mean to the GRACE/-FO solutions, model estimates and the IOM data for the records to be comparable. Figure 6a shows mass change for the full period, Fig. 6b within the testing period, and Fig. 6c within the gap between GRACE and GRACE-FO. The

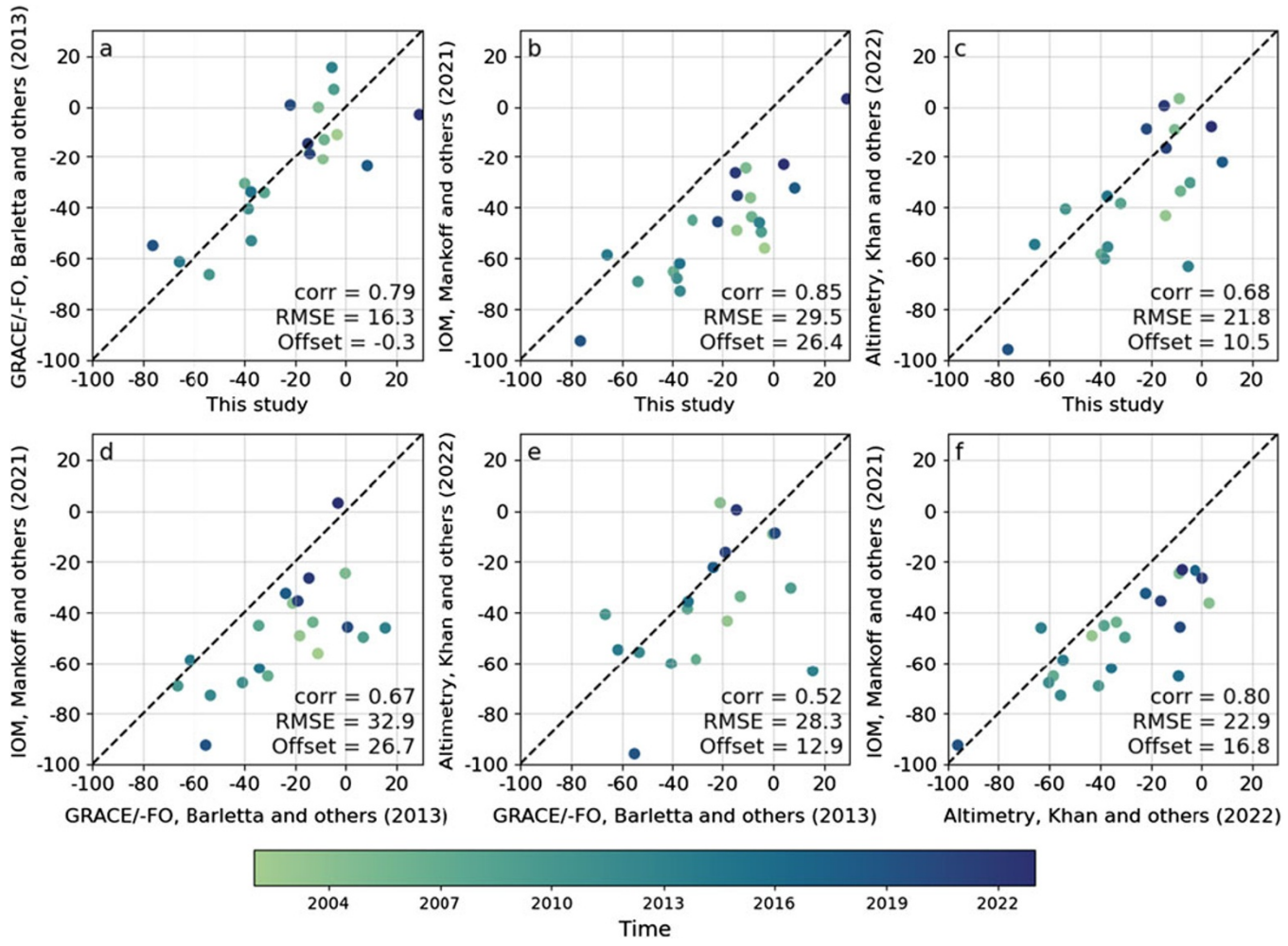

Figure 5. Comparison of annual MB for the CW basin for hydrological years (Oct–Sept) between the data-driven model estimates, Barletta and others (2013), Mankoff and others (2021), and Khan and others (2022). All axis units are in Gt/year and offset is calculated as $1/n\sum(x_i - y_i)$, units in Gt/yr. Due to the gap between GRACE and GRACE-FO, incomplete hydrological years are excluded in the data-driven model estimates and Barletta and others (2013) plots (a-e). Only (f) includes all years between 2002 and 2023.

mass changes for the testing period compare well with both inter-comparison datasets. Figure 6b also includes the mass changes from the auto-regressive model starting in January 2011. The auto-regression performs similarly for the first year but shows less variability and less mass loss in the summer of 2011. In the following winter, the auto-regressive model predicts similar MBs as the inter-comparison datasets but deviates in the 2012 melt season. Figure 6c also includes the mass changes from the auto-regressive model that were initialized at different times: 29 November 2016 and 10 June 2017. There are only small differences between mass changes predicted by the two auto-regressive predictions within the gap between GRACE and GRACE-FO. When compared with the IOM and altimetry-derived mass changes, both comparison datasets show more negative mass change in the summer of 2017. The following winter and summer the auto-regressive predictions show similar mass changes when compared with IOM and altimetry. When looking at the mass changes leading up to the GRACE/-FO gap, the GRACE data appear to be out of phase compared with IOM and altimetry, which is partially also seen in the model predictions. When the auto-regression starts before 2017, the model can capture the increase in mass change and hence fits better the comparison datasets.

5. Discussion

Our findings confirm that a deep learning model can effectively combine atmospheric conditions with prior GRACE/-FO-derived mass anomaly measurements to emulate mass anomalies. This capability highlights the potential of the data-driven model when filling the gap of observations in the gravitational record of ice mass loss. As seen in Figs. 3 and 4, the model demonstrates the ability to capture the complex interactions between atmospheric drivers and the observed mass anomalies on the entire CW basin, showing low RMSE (< 20Gt) and high r^2-score (> 0.73) between model predictions and observations. The model has slightly better test statistics than training, indicating that the model is generalizing well to unseen data. The testing period was intentionally chosen to be in the middle of the GRACE period rather than at the beginning or end, as the GRACE solutions consistently showcase lower errors, thus making them more reliable for testing the model's performance. Still, the training data do include all data independent of data quality, which explains the slightly better test performance. Furthermore, GRACE-FO was not included in the test dataset but only in the training data. Within the GRACE-FO period, the model shows slightly lower performance compared

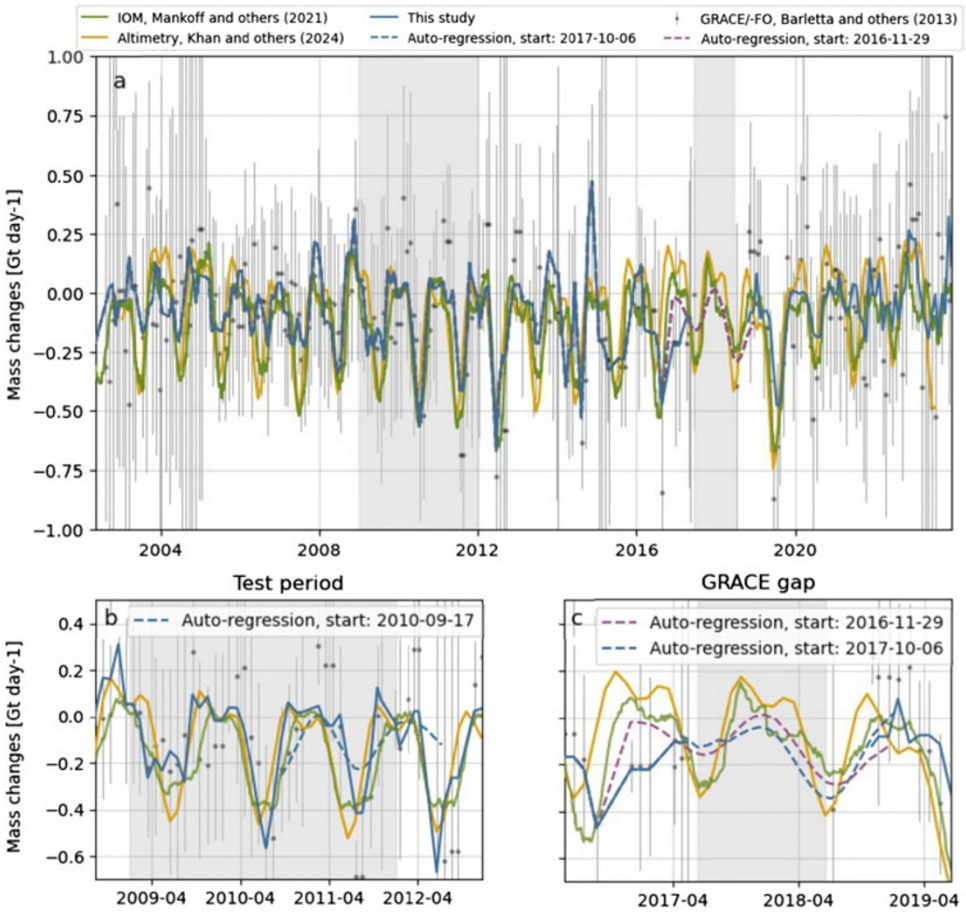

Figure 6. Mass changes estimated by the deep learning model (blue), GRACE/-FO observations (black), the IOM (green) and derived from altimetry (yellow). Since the altimetry-derived mass changes are temporally smoothed, we apply a 4 month running mean to the mass changes by data-driven model. (This study), GRACE/-FO observations and the IOM (a) show the full GRACE/-FO period, whereas (b) and (c) include the testing periods. In both (b) and (c), the dotted lines show the auto-regressive predictions.

with the GRACE period. This suggests that, despite the assumption that GRACE and GRACE-FO data are equivalent due to their identical systems and underlying physics, the model struggles to capture MB variability during the GRACE-FO period accurately. This may point to potential differences within the CW basin between data from the two missions. Due to the accelerometer data degradation on GRACE-D, replaced with transplant data (Behzadpour and others, 2021), the GRACE-FO measurement uncertainty is generally larger. These increased uncertainties propagate into the GRACE/-FO mass anomaly estimates, which is evident in Figs. 4 and 6. We incorporate the uncertainty of the mass anomalies into the model training, constraining the model to stay within the uncertainty bounds of the GRACE/-FO observations. Thus, in periods with larger uncertainty, the model can deviate more from the observations, which might explain the slight drop in model performance during the GRACE-FO period. As the operational record of GRACE-FO increases in the future, it would be interesting to see if these differences persist.

The mass anomalies for the basin in Fig. 4 show that the model-predicted mass anomalies generally fall within the GRACE/-FO uncertainty envelope. Evident in Fig. 6, the model captures the general trend of the mass changes but fails to capture the full magnitude of these mass changes. GRACE/-FO mass change observations show several high-amplitude mass changes both within the training and testing period. Such rapid increases are typically driven by short-lived, large precipitation events, extreme melt events or brief periods of extremely high calving rates, each of which occurs only rarely. As a result, these events are also rare in the data, making it difficult for the model to effectively learn and capture them. This is also evident with the peak in late 2014, where we see a strong high mass change in both the GRACE solution and the data-driven model, but not in the IOM and altimetry. The performance of the data-driven model is limited by the data it is trained with. This challenge is not specific to the method presented in this study but represents a broader limitation for data-driven models (Goodfellow and others, 2016). Furthermore, we also note that these high-amplitude mass changes are not present in either IOM or altimetry mass changes. IOM relies on RCMs to describe the MB and, thus, inherits any biases present in the RCMs. Moreover, the ice velocity data used to calculate solid ice discharge vary in temporal coverage, which introduces additional 'smoothing' to the dataset. Altimetry-derived mass changes are natively temporally 'smoothed'. Consequently, neither IOM nor altimetry can capture these high-amplitude mass changes. On the other hand, GRACE/-FO can observe these high-amplitude changes, since the GRACE/-FO directly measures the changes in Earth's gravity field and does not rely on RCMs for surface processes or smoothed ice velocity data. However, on a shorter time scale, the measurements can be susceptible to noise, mainly due to aliasing of short-term mass

variations in the atmosphere and ocean, which are not removed by the background models and, therefore, are more prominent at short time scales (Wahr and others, 2006).

As shown in the annual validation plot (Fig. 5a), the data-driven model aligns well with the GRACE/-FO solutions, demonstrating high correlation and minimal offset. This indicates that the data-driven model captures the annual variability in MB observed by GRACE/-FO. In contrast, there is a high RMSE (> 21 Gt/yr) and substantial offset (> 10 Gt/yr) between the data-driven model and both the IOM (Fig. 5b) and altimetry-derived (Fig. 5c) MB estimates. Furthermore, the correlation between the data-driven and the altimetry-derived MB is low, but a lower correlation is seen between altimetry and GRACE/-FO MB in Fig. 5e. This suggests that it is not a limitation of the data-driven model but rather a discrepancy between the altimetry-derived MB and other estimates. The IOM method shows a high offset with the data-driven, meaning that on average, the IOM estimates 26.4 Gt lower MB, compared with the data-driven model. Similarly, the IOM also shows positive offsets with GRACE/-FO solutions and altimetry, again suggesting that the discrepancy is not a limitation of the data-driven model, but rather a systematic offset in the IOM. While both the IOM and GRACE/-FO observations do not include the outer glaciers, GRACE/-FO observations are affected by signal leakage from outer glaciers, especially from Disko Island. However, signal leakage cannot explain more than 20 Gt/yr offset between IOM and GRACE/-FO observations in the CW basin.

While Fig. 5 revealed interannual variability and offsets between the MB approaches, the model-estimated mass changes generally show the same seasonal variability as the IOM and altimetry-derived mass changes in Fig. 6. However, in 2017, leading up to the GRACE/-FO gap, the GRACE solutions drift away from the mass change estimated by both the IOM and altimetry. Since the applied loss function accounts for the uncertainties in GRACE/-FO data by assigning less weight to observations with high uncertainty, it is unsurprising that the model performs poorly during the first years of the GRACE mission, where the GRACE uncertainties are generally high. However, in 2017, the uncertainties associated with GRACE do not reflect any decrease in data quality near the end of the mission, even though the solutions fail to capture the seasonal trends observed in the IOM and altimetry mass changes. The model similarly fails to capture the seasonal trend, inheriting this limitation from the GRACE data. Since data-driven models, in general, learn the relationship between input and output data, the model is naturally sensitive to the quality of the GRACE/-FO solutions. The model will reproduce systematic biases and uncertainties in the GRACE/-FO data. It is, therefore, important to carefully consider data quality when applying data-driven models.

To test the performance of the auto-regressive nature of the model, we perform an auto-regression during the testing period where GRACE solutions are available (Fig. 6b). Here, we can evaluate how well the auto-regressive model performs by comparing the auto-regressive predictions to GRACE observations and the regular model (seeing the entire GRACE observational period). The auto-regression correctly captures the general trend but does not exhibit the same variability as the GRACE solution or regular model. With the auto-regressive model, we can capture the general climatic trends of the MB, but mass changes induced by weather variability are not fully captured. Although the auto-regressive model cannot capture the full variability of the MB, it can reproduce the long-term trends (climatic variability) and thus still be useful for bridging the gap between GRACE and GRACE-FO in Fig. 6c. Since GRACE fails to capture the seasonal trends in early 2017, we initiate the auto-regression twice: first on 29 November 2016 to exclude GRACE solutions from 2017, and again at the end of the final GRACE solution on 10 June 2017. Compared to IOM and altimetry, neither auto-regressive predictions capture as large a mass loss during the 2017 melt season. The subsequent mass increase during the winter of 2017 and mass loss of the 2018 melt season compared well with IOM and altimetry, indicating that the auto-regressive model successfully captures mass changes during this part of the gap. Leading up to the end of the GRACE mission, we can compare the regular and auto-regressive models. While the auto-regressive model can only capture the climatic trends, it still compares better with IOM and altimetry than the regular model, as this model inherits the bias from the GRACE observations. Furthermore, this also explains the 30 Gt difference between the two auto-regressive model predictions. Thus, the auto-regressive model, which begins before 2017, produces the best estimates. While integrating the actual GRACE/-FO observation into the model is preferred, 2017 illustrates a case where the autoregressive model outperforms the regular model due to a bias in the GRACE observation.

Capturing the evolution of mass anomalies over time requires a deep learning architecture that can effectively model both short-term variability and long-term trends. For future work, we suggest improving the model by including more advanced temporal modelling approaches in the network architecture for future work to enhance the model's auto-regressive capabilities. Recurrent Neural Networks such as Long Short-Term Memory networks or Gated Recurrent Units are well-suited to capturing temporal correlations. They can aid the model to better understand the time evolution of the system's dynamics (Goodfellow and others, 2016). Additionally, incorporating attention mechanisms within the temporal framework could further enhance the network's ability to focus on key time steps or critical transitions, such as periods of rapid ice loss or extreme melt seasons affecting the properties of the firn in the subsequent years. By incorporating these components into the architecture, the generative performance of the network could be improved, leading to more accurate reconstructions and predictions of mass anomalies within the GRACE/-FO gap.

6. Conclusion

This study demonstrates that a data-driven model can successfully emulate mass anomalies by combining atmospheric conditions with a history of GRACE/-FO-derived mass anomalies. The model captures basin-scale mass anomalies well, showing low RMSE and high r^2 scores for the GRACE period but slightly lower performance in the GRACE-FO period, suggesting differences between the two missions. This difference is likely due to the higher uncertainty associated with the GRACE-FO observations, which is accounted for during model training. Overall, the model-predicted mass anomalies generally fall within the GRACE/-FO uncertainty envelope, successfully capturing the broader mass change trends but missing high-amplitude changes driven by extreme precipitation, melt and calving events. These limitations stem from the data-driven approach, as these extreme events are rare occurrences in the input data, making it difficult for the data-driven model to learn and capture them effectively. Annually, the MB estimated by the data-driven model aligns well with GRACE/-FO observations. In contrast, both the data-driven model and the GRACE/-FO observations reveal an offset in the IOM and low correlation with the altimetry-derived MB, suggesting discrepancies are between the traditional methods rather than in the data-driven

approach. Using the auto-regressive nature of the model, we fill the gap between GRACE and GRACE-FO by replacing the previous GRACE mass anomaly observation with the previously predicted mass anomaly. Results show that the auto-regressive model successfully predicts climatic variability but struggles to capture weather-imposed mass changes. Leading up to the end of the GRACE mission in 2017, the seasonal trend is not captured in GRACE-derived mass changes, whereas it remains evident in both IOM- and altimetry-derived mass changes. The data-driven model, which sees the entire GRACE observational period, inherits this trend from the GRACE observations, whereas the auto-regressive predictions do not. Results show that by excluding 2017 GRACE solutions in the auto-regressive predictions, we can better bridge the gap between GRACE and GRACE-FO while remaining within the GRACE-FO uncertainty envelope. Thus, we can create a 21 year continuous time series of mass anomalies that are consistent with GRACE/-FO observations using a data-driven model, providing an alternative to traditional methods such as IOM and altimetry for estimating ice sheet MB during periods with limited or missing GRACE/-FO data.

Supplementary material. The supplementary material for this article can be found https://doi.org/10.1017/aog.2025.10019.

Acknowledgements. Both AP and NH received support for this study from the National Centre for Climate Research (NCKF). Furthermore, NH is supported by the Novo Nordisk Foundation project PRECISE (NNF23OC0081251). The authors acknowledge the ESA Climate change initiative for the Greenland ice sheet funded via ESA-ESRIN contract number 4000104815/11/I-NB. Furthermore, the authors acknowledge Valentina Barletta for providing the GRACE/-FO-derived mass anomaly dataset.

References

Barletta VR, Bordoni A and Khan SA (2024a) GNEt derived mass balance and glacial isostatic adjustment constraints for greenland. *Geophysical Research Letters* **51**, e2023GL106891. doi:10.1029/2023GL106891

Barletta VR, Sørensen LS and Forsberg R (2013) Scatter of mass changes estimates at basin scale for Greenland and Antarctica. *The Cryosphere* **7**, 1411–1432. doi:10.5194/tc-7-1411-2013

Barletta VR, Sørensen LS, Simonsen SB and Forsberg R (2024b) Gravitational mass balance of Greenland 2003 to present v2.4. doi:10.11583/DTU.12866579.V4

Baur O, Kuhn M and Featherstone WE (2009) GRACE-derived ice-mass variations over Greenland by accounting for leakage effects. *Journal of Geophysical Research: Solid Earth* **114**, 2008JB006239. doi:10.1029/2008JB006239

Behzadpour S, Mayer-Gürr T and Krauss S (2021) GRACE follow-on accelerometer data recovery. *Journal of Geophysical Research: Solid Earth* **126**, e2020JB021297. doi:10.1029/2020JB021297

Bolibar J, Sapienza F, Maussion F, Lguensat R, Wouters B and Pérez F (2023) Universal differential equations for glacier ice flow modelling. *Geoscientific Model Development* **16**, 6671–6687. doi:10.5194/gmd-16-6671-2023

Caron L, Ivins ER, Larour E, Adhikari S, Nilsson J and Blewitt G (2018) GIA model statistics for GRACE hydrology, cryosphere, and ocean science. *Geophysical Research Letters* **45**, 2203–2212. doi:10.1002/2017GL076644

Chen Z, Badrinarayanan V, Lee CY and Rabinovich A (2018) Gradnorm: gradient normalization for adaptive loss balancing in deep multitask networks. In *Proceedings of the 35 th International Conference on Machine Learning*.

Cheng G, Morlighem M and Francis S (2024) Forward and inverse modeling of ice sheet flow using physics-informed neural networks: application to Helheim Glacier, Greenland. *Journal of Geophysical Research: Machine Learning and Computation* **1**, e2024JH000169. doi:10.1029/2024JH000169

De Roda Husman S and 7 others (2024) Physically-informed super-resolution downscaling of antarctic surface melt. *Journal of Advances in Modeling Earth Systems* **16**, e2023MS004212. doi:10.1029/2023MS004212

Goodfellow I, Bengio Y and Courville A (2016) *Deep Learning*. Cambridge, Massachusetts, USA: MIT Press.

Hersbach H and 43 others (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **146**, 1999–2049. doi:10.1002/qj.3803

Heydari AA, Thompson CA and Mehmood A (2019) Softadapt: techniques for adaptive loss weighting of neural networks with multi-part loss Functions doi:10.48550/arXiv.1912.12355. [cs, math, stat].

Howat I (2017) *MEaSURES Greenland Ice Velocity: Selected Glacier Site Velocity Maps From Optical Images*, Version 2, Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. doi:10.5067/VM5DZ20MYF5C

Howat IM, Negrete A and Smith BE (2014) The Greenland Ice Mapping Project (GIMP) land classification and surface elevation data sets. *The Cryosphere* **8**, 1509–1518. doi:10.5194/tc-8-1509-2014

Joughin I, Shean DE, Smith BE and Floricioiu D (2020) A decade of variability on Jakobshavn Isbræ: ocean temperatures pace speed through influence on mélange rigidity. *The Cryosphere* **14**, 211–227. doi:10.5194/tc-14-211-2020

Jouvet G and Cordonnier G (2023) Ice-flow model emulator based on physics-informed deep learning. *Journal of Glaciology* **69**(278), 1941–1955. doi:10.1017/jog.2023.73

Jouvet G, Cordonnier G, Kim B, Lüthi M, Vieli A and Aschwanden A (2022) Deep learning speeds up ice flow modelling by several orders of magnitude. *Journal of Glaciology* **68**(270), 651–664. doi:10.1017/jog.2021.120

Karlsson NB, Solgaard AM, Mankoff KD, Gillet-Chaulet F, MacGregor JA and Box JE (2021) A first constraint on basal melt-water production of the Greenland ice sheet. *Nature Communications* **12**, 3461. doi:10.1038/s41467—23739-z

Khan SA and 24 Others (2025) Smoothed Monthly Greenland ice Sheet Elevation Changes During 2003–2023. *Earth System Science Data* **17**(6), 3047-3071. doi:10.5194/essd-17-3047-2025

Khan SA, Bamber JL, Rignot E, Helm V, Aschwanden A and Holland DM (2022) Greenland mass trends from airborne and satellite altimetry during 2011–2020. *Journal of Geophysical Research: Earth Surface* **127**, e2021JF006505. doi:10.1029/2021JF006505

Khan SA, Sasgen I, Bevis M, Van Dam T, Bamber JL and Wahr J (2016) Geodetic measurements reveal similarities between post–last glacial maximum and present-day mass loss from the Greenland ice sheet. *Science Advances* **2**, e1600931. doi:10.1126/sciadv.1600931

Khazendar A and 13 others (2019) Interruption of two decades of jakobshavn isbrae acceleration and thinning as regional ocean cools. *Nature Geoscience* **12**, 277–283. doi:10.1038/s41561-019-0329-3

Kingma DP and Ba JL (2015) Adam: a method for stochastic optimization 3rd international conference on learning Representations. International Conference for Learning Representations, San Diego, California, USA. In *ICLR 2015-Conference Track Proceedings* 1.

Landerer F (2024) Monthly estimates of degree-1 (geocenter) gravity coefficients, generated from GRACE (04-2002 - 06/2017) and GRACE-FO (06/2018 onward) RL0602 solutions. Technical report, NASA Jet Propulsion Laboratory.

Lutz K, Bahrami Z and Braun M (2023) Supraglacial lake evolution over Northeast Greenland using deep learning methods. *Remote Sensing* **15**, 4360. doi:10.3390/rs15174360

Mankoff KD and 13 others (2021) Greenland ice sheet mass balance from 1840 through next week. *Earth System Science Data* **13**, 5001–5025. doi:10.5194/essd-13-5001-2021

Mankoff KD, Solgaard A, Colgan W, Ahlstrøm AP, Khan SA and Fausto RS (2020) Greenland Ice Sheet solid ice discharge from 1986 through March 2020. *Earth System Science Data* **12**, 1367–1383. doi:10.5194/essd-12-1367-2020

Morlighem M and 35 others (2017) BedMachine v3: complete bed topography and ocean bathymetry mapping of Greenland from multibeam echo sounding combined with mass conservation. *Geophysical Research Letters* **44**, 11,051–11,061. doi:10.1002/2017GL074954.

Mouginot J and 8 others (2019) Forty-six years of Greenland ice sheet mass balance from 1972 to 2018. *Proceedings of the National Academy of Sciences* **116**, 9239–9244. doi:10.1073/pnas.1904242116

Noël B, van de Berg WJ, Lhermitte S and van den Broeke MR (2019) Rapid ablation zone expansion amplifies north Greenland mass loss. *Science Advances* **5**, eaaw0123. doi:10.1126/sciadv.aaw0123

Otosaka IN and 83 others (2023) Mass balance of the greenland and antarctic ice sheets from 1992 to 2020. *Earth System Science Data* **15**, 1597–1616. doi:10.5194/essd-15-1597-2023

Sasgen I and 8 others (2020) Return to rapid ice loss in greenland and record loss in 2019 detected by the grace-fo satellites. *Communications Earth & Environment* **1**, 1–8. doi:10.1038/s43247-020-0010-1

Shepherd A and 106 others (2020) Mass balance of the Greenland ice sheet from 1992 to 2018. *Nature* **579**, 233–239. doi:10.1038/s41586-019-1855-2

Shi Z, Wang Z, Zhang B, Geng H, An J, Wu S, Liu M and Wu Y and Wu, H (2024) Bridging the spatiotemporal ice sheet mass change data gap between GRACE and GRACE-FO in Greenland using machine learning method. *Journal of Hydrology* **629**, 130622. *Journal of Hydrology* doi: 10.1016/j.jhydrol.2024.130622

Solgaard A and 13 others (2021) Greenland ice velocity maps from the promice project. *Earth System Science Data* **13**, 3491–3512. doi:10.5194/essd-13-3491-2021

Sun Y, Riva R and Ditmar P (2016) Optimizing estimates of annual variations and trends in geocenter motion and J_2 from a combination of GRACE data and geophysical models. *Journal of Geophysical Research: Solid Earth* **121**, 8352–8370. doi:10.1002/2016JB013073

Swenson S, Chambers D and Wahr J (2008) Estimating geocenter variations from a combination of GRACE and ocean model output. *Journal of Geophysical Research: Solid Earth* **113**, 2007JB005338. doi:10.1029/2007JB005338

Tapley BD and 21 others (2019) Contributions of grace to understanding climate change. *Nature Climate Change* **9**, 358–369. doi:10.1038/s41558-019-0456-2

Tedesco M and 7 others (2011) The role of albedo and accumulation in the 2010 melting record in Greenland. *Environmental Research Letters* **6**, 014005. doi:10.1088/1748-9326/6/1/014005

Tedesco M and 10 others (2017) Arctic report card 2017: Greenland ice sheet. In *Arctic Report Card 2017*, NOAA.

Tedesco M and 9 others (2018) Arctic report card 2018: Greenland Ice Sheet. In *Arctic Report Card 2018*, NOAA.

Tedesco M and Fettweis X (2020) Unprecedented atmospheric conditions (1948–2019) drive the 2019 exceptional melting season over the Greenland ice sheet. *The Cryosphere* **14**, 1209–1223. doi:10.5194/tc-14-1209-2020

van der Meer M, De Roda Husman S and Lhermitte S (2023) Deep learning regional climate model emulators: a comparison of two downscaling training frameworks. *Journal of Advances in Modeling Earth Systems* **15**, e2022MS003593. doi:10.1029/2022MS003593

Velicogna I and 11 others (2020) Continuity of ice sheet mass loss in Greenland and Antarctica from the GRACE and GRACE Follow-On missions. *Geophysical Research Letters* **47**, e2020GL087291. doi:10.1029/2020GL087291

Velicogna I and Wahr J (2005) Greenland mass balance from GRACE. *Geophysical Research Letters* **32**, 2005GL023955. doi:10.1029/2005GL023955

Wahr J, Swenson S and Velicogna I (2006) Accuracy of GRACE mass estimates. *Geophysical Research Letters* **33**, 2005GL025305. doi:10.1029/2005GL025305

Zhang B, Yao Y and He Y (2022) Bridging the data gap between GRACE and GRACE-FO using artificial neural network in Greenland. *Journal of Hydrology*, **608**, 127614. doi: 10.1016/j.jhydrol.2022.127614

Zhu Q and 6 others (2024) Automated surface melt detection over the Antarctic from Sentinel-1 imagery using deep learning. *International Journal of Applied Earth Observation and Geoinformation* **130**, 103895. doi:10.1016/j.jag.2024.103895

Zwally JH, Giovinetto MB, Beckley MA and Saba JA (2012) *Antarctic and Greenland drainage systems*. GSFC cryospheric sciences laboratory. http://icesat4.gsfc.nasa.gov/cryo_data/ant_grn_drainage_systems.php.