DISSERTATION | DOCTORAL THESIS

Titel | Title

Diagnostic evaluation of coupled Arctic energy and water budgets from observations and climate models

verfasst von | submitted by
Susanna Winkelbauer BSc MSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Doktorin der Naturwissenschaften (Dr.rer.nat.)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 796 605 415
Dissertationsgebiet lt. Studienblatt | Field of study as it appears on the student record sheet:	Meteorologie
Betreut von | Supervisor:	Univ.-Prof. Mag. Dr. Leopold Haimberger

Abstract

This work describes progress towards a better description of the ongoing fundamental changes in the Arctic by (i) creating observation-based, up-to-date estimates of the Arctic water budget, (ii) validating historical energy and water budget simulations of state-of-the-art climate models, and (iii) developing new tools for the precise calculation of oceanic transports on various modeling grids.

The first part of this doctoral thesis deals with the Arctic water budget as represented by observations. Firstly, river discharge into the Arctic Ocean is analyzed and by combining available gauge observations with reanalyses estimates, a new and up-to-date pan-Arctic observation based river discharge estimate is created. Through combination with atmospheric inputs from reanalyses, storage changes from the Gravity Recovery and Climate Experiment (GRACE) and oceanic volume transports from ocean reanalyses, the closure of the non-steric water volume budget is assessed. Finally, using a variational adjustment scheme best estimates for every budget equation term are provided, offering an up-to-date picture of the Arctic water budget and a solid foundation for the second objective of this thesis, the validation of climate models.

The second part of this thesis analyzes relevant components of the Arctic energy and water budgets for climate models from the Coupled Model Intercomparison Project Phase 6 (CMIP6). Simulated long-term averages, trends and seasonal cycles of vertical fluxes at the surface as well as the top-of-atmosphere, lateral transports and storage rates in atmosphere and ocean are validated against our observation-based basis. Large inter-model spreads and systematic biases are found in the representation of long-term averages and annual cycles for several water and energy budget components. Surface freshwater fluxes into the Arctic Ocean tend to be overestimated by most CMIP6 models and about two thirds feature an early timing bias in the seasonal runoff cycle, attributed to an early snow melt bias and the lack of realistic river routing schemes. The net vertical energy flux out of the Arctic Ocean and poleward oceanic heat transports are systematically underestimated by all models. The latter show strong anti-correlation to the mean sea ice cover, atmospheric heat transports, and the long-term ocean warming rate. This strongly suggests that accurate depiction of the mean state is critical for realistic projections of future Arctic warming. This new and comprehensive overview of the models' performance, biases and their ability to reproduce the observed Arctic energy and water budgets should help to evaluate the reliability of the models' future projections.

Another aspect of this doctoral thesis is dedicated to the technical development of new tools for the precise calculation of oceanic transports. Oceanic transports significantly influence the global climate, however their evaluation and validation based on reanalysis and model data is complicated by the distortion of the employed model grids and the diversity of different grid types used. Using two different methods, StraitFlux allows for precise calculations of oceanic fluxes through almost arbitrary sections, enabling straightforward comparisons with observed oceanic volume and energy transports at available sections such as the RAPID array or Arctic gateways. StraitFlux has been in use for transport studies at the Greenland-Scotland Ridge and the Indonesian Troughflow region and was essential in the calculation of Arctic lateral fluxes in this doctoral thesis.

Zusammenfassung

Diese Arbeit beschreibt den Fortschritt hin zu einer besseren Darstellung der laufenden grundlegenden Veränderungen in der Arktis durch (i) die Erstellung von auf Beobachtungen basierten, aktuellen Schätzungen des arktischen Wasserkreislaufs und (ii) die Validierung historischer Simulationen des Energie- und Wasserkreislaufs in Klimamodellen. (iii) Dabei sollen auch neue Methoden für die präzise Berechnung ozeanischer Transporte auf verschiedenen Modellgittern entwickelt werden.

Im ersten Teil dieser Doktorarbeit wird der Abfluss vom Land in den Arktischen Ozean analysiert. Durch die Kombination vorhandener Beobachtungen mit Schätzungen aus Reanalysen wird eine neue und aktuelle, auf Beobachtungen basierende Schätzung des panarktischen Abflusses erstellt. Durch die Kombination mit dem atmosphärischen Eintrag aus Reanalysen, Speicheränderungen aus dem Gravity Recovery and Climate Experiment (GRACE) und ozeanischen Volumenstransporten aus Ozean-Reanalysen wird die Schließung des Wasserhaushaltes bewertet. Unter Verwendung eines variationalen Anpassungsschemas werden die besten Schätzungen für jeden Term der Budgetgleichung erzeugt. Diese auf Beobachtungen basierenden Schätzungen bieten ein aktuelles Bild des arktischen Wasserhaushalts und bilden eine solide Grundlage für das zweite Ziel dieser Arbeit, die Validierung von Klimamodellen.

Der zweite Teil dieser Arbeit gilt der Analyse relevanter Komponenten der arktischen Energie- und Wasserbilanzen für Klimamodelle des Coupled Model Intercomparison Project Phase 6 (CMIP6). Simulierte Langzeitmittelwerte, Trends und saisonale Zyklen der vertikalen Flüsse an der Oberfläche sowie am Oberrand der Atmosphäre, laterale Transporte und Speicherraten in Atmosphäre und Ozean werden mit unserer beobachtungsbasierten Grundlage validiert. Wir finden große Unterschiede zwischen den Modellen und systematische Verzerrungen in der Darstellung von Langzeitmitteln und Jahreszyklen für mehrere Wasser- und Energiebilanzkomponenten. Frischwasserflüsse an der Oberfläche des Arktischen Ozeans werden von den meisten CMIP6-Modellen tendenziell überschätzt und etwa zwei drittel der Modelle zeigen einen zeitlichen Offset im Abflusszyklus, der mit einem frühen Bias in der Schmelze und dem Mangel an realistischen routing in den Modellen in Verbindung steht. Der vertikale Energiefluss aus dem Arktischen Ozean und die polwärts gerichteten ozeanischen Wärmetransporte werden von allen Modellen systematisch unterschätzt. Letztere zeigen eine starke Antikorrelation zur mittleren Meereisbedeckung, atmosphärischen Wärmetransporten und auch der langfristigen Erwärmungsrate des Ozeans. Somit ist für realistische Prognosen zukünftiger Erwärmungen in der Arktis eine genaue Darstellung des mittleren Zustands wichtig. Dieser neue und umfassende Überblick über die Fähigkeiten der Modelle beobachteten Energie- und Wasserhaushalte zu reproduzieren, sollte dazu beitragen, die Zuverlässigkeit der zukünftigen Prognosen der Modelle zu bewerten.

Ein weiterer Aspekt dieser Doktorarbeit war die technische Entwicklung neuer Werkzeuge für die präzise Berechnung ozeanischer Transporte. Ozeanische Transporte formen das globale Klima, jedoch wird ihre Validierung basierend auf Reanalysen- und Modell-Daten durch die Verzerrung der verwendeten Modellgitter und die Vielzahl verschiedener Gittertypen erschwert. Mit Hilfe von zwei verschiedenen Methoden ermöglicht StraitFlux präzise Berechnungen ozeanischer Flüsse durch beliebige ozeanische Abschitte. Dies ermöglicht einen einfachen Vergleich mit beobachteten Transporten an Abschnitten wie etwa dem RAPID-Array oder den arktischen Wasserstraßen. StraitFlux wurde bereits bei Studien an der Greenland-Scotland-Ridge (GSR) und in der Region des Indonesian Throughflow verwendet und war entscheidend bei der Berechnung arktischer lateraler Flüsse in dieser Doktorarbeit.

Contents

1 Introduction

The Arctic is a region of profound climatic and ecological significance, influencing the global climate and sea level, biodiversity and socio-economic factors. Characterized by its extreme conditions and unique environmental dynamics, the Arctic System plays a crucial role in shaping global climate patterns. The Arctic's complex climate system is characterized by vast expanses of sea ice, permafrost-covered landmasses, and close couplings and interactions between atmosphere, ocean and sea ice as well as a tight interplay between energy and water fluxes. The analysis and understanding of this complex system is crucial to address environmental as well as geopolitical challenges, especially in the context of climate warming and its global implications. Notably, the rapid decline of Arctic sea ice opens shorter trans-Arctic trade routes linking the Pacific and Atlantic Oceans, having major economic and potentially environmental implications (Cao et al, 2022; Melia et al, 2017; Yumashev et al, 2017).

Energetically the Earth's climate system is open, absorbing shortwave radiation from the sun and emitting thermal longwave radiation into space at the top of the atmosphere (TOA). In terms of matter, however, the Earth's functions predominantly as a closed system. While it does experience minor losses of volatile gases from the atmosphere, this is significantly mitigated by the Earth's magnetic field maintaining the conditions necessary for habitability (Catling and Zahnle, 2009). This system facilitates the exchange of energy, matter, and momentum across its five components: the atmosphere, hydrosphere, cryosphere, lithosphere, and biosphere, and within them. For instance, the transfer of energy (as latent heat) and matter (in the form of water) between the atmosphere and the hydrosphere during evaporation or the transport of energy and matter within the hydrosphere/atmosphere with ocean currents/wind (Peixoto and Oort, 1992). In the long-term average the climate system gains energy at the top of the atmosphere in the tropics, but loses energy at the poles, meaning the Arctic acts a a "heat sink" in the global energy budget due to a deficit in incoming solar radiation. This latitudinal energy gradient, known as differential heating, is compensated by maintained horizontal energy transports from the equator towards the poles in the atmosphere and ocean (Lorenz, 1967).

The continuous increase in anthropogenic greenhouse gases (GHG) has led to an accumulation of energy in the Earth's climate system, primarily due to the GHGs' ability to absorb and re-emit outgoing longwave radiation back towards the Earth's surface. This results in a positive energy imbalance as the planet receives more incoming solar radiation than it releases thermal radiation, causing rising temperatures in atmosphere and ocean. Consequently, this warming has impacts on the global climate system by intensifying extreme weather events, altering precipitation patterns, and causing sea level rise due to the thermal expansion of ocean water and melting of glaciers and ice sheets. These changes pose significant challenges to ecosystems, human communities and infrastructure across the globe.

Amidst these global changes, the Arctic region experiences particularly rapid warming, approximately at a rate about two (Serreze et al, 2009; Walsh, 2014) to four (Rantanen et al, 2022) times the global average, a phenomenon known as Arctic Amplification (AA). This accelerated warming is driven by a combination of radiative and non-radiative feedback mechanisms, including the ice-albedo feedback, where the melting of ice and snow exposes darker surfaces with lower albedos that subsequently absorb more solar radiation and in turn further increase warming, leading to additional ice and snow melt. Similarly, the lapse rate feedback, which results from a distinct vertical warming pattern in the Arctic troposphere with pronounced warming at the surface and subdued warming in the upper troposphere due to stable stratification and meridional energy transports, significantly contributes to AA (Linke et al, 2023). Additional factors contributing to AA include the reduced rate of heat loss to space compared to the subtropics, increases in total water vapour content in the Arctic atmosphere, northward heat and moisture fluxes, and changes in summer cloud cover (Meredith et al, 2022).

These rapidly increasing temperatures have important implications for the Arctic System itself and also for the global climate system. One of the most visible impacts of Arctic warming is the accelerated melting of sea-ice (Stroeve and Notz, 2018), with the decline in sea ice area and thickness being particularly pronounced in recent decades (Kwok, 2018). As stated in Meredith et al (2022) the strong declines in sea ice extent in September over recent decades (-12.8 ± 2.3 % per decade) are unprecedented in at least 1000 years. Furthermore, the mass loss of terrestrial glaciers and the Greenlandic ice sheet (GIS) (Mouginot et al, 2019) have increased since 1990, together losing almost -500 Gt yr^{-1} (Meredith et al, 2022). Permafrost is another highly affected component of the Arctic (Rowland et al, 2010) and its degradation significantly impacts the climate system by releasing stored carbon dioxide and methane into the atmosphere. Further, permafrost thaw and sea ice loss greatly impact local communities and indigenous populations through altering environments, coastal erosion and damaging infrastructures as well as cultural sites, requiring drastic adaptation strategies (Irrgang et al, 2019; Bartsch et al, 2021). Additionally, Arctic warming significantly impacts the Arctic hydrological cycle, leading to its overall intensification. Snow, ice and permafrost melt lead to increased runoff from terrestrial regions and the GIS and additionally the input of freshwater into the Arctic Ocean is enhanced through increased precipitation.

Effects are also seen outside the Arctic, as the melting of glaciers and the Greenland ice sheet contribute to rapid sea-level rise around the globe (e.g., Box et al, 2022), potentially reshaping coastlines, impacting coastal communities, and posing challenges to global ecosystems (Oppenheimer et al, 2019). The increased input of freshwater into the Arctic Ocean has led to an accumulation of freshwater in the Arctic Ocean, a sudden release of those large amounts of freshwater into the Atlantic Ocean could have major implications for the global oceanic circulation (Haine et al, 2015), influencing the global heat distribution. Furthermore, a reduction in the meridional temperature gradient between the equator and the poles might also affect weather and climate patterns in the mid-latitudes. This idea, however, is the subject of considerable discussion and analysis with conflicting evidence from studies by e.g., Blackport and Screen (2020); Coumou et al (2018); Francis and Vavrus (2012); Screen and Simmonds (2013)

Understanding these pronounced warming trends and resulting effects on the regional and global climate system requires a comprehensive understanding of the underlying physical processes in the Arctic system and the tight couplings between atmosphere, ocean and sea ice. Hence, to improve our knowledge on the complexity of the Arctic system and its far reaching impacts it is crucial to analyse the Arctic energy and water budgets. In general the energy/water budgets describe the balance between energy/water storage in the considered system and energy/water input into as well as output out of the system. Figure 1 shows an overview of the processes involved in the Arctic water (left) and energy (right) cycles. Arrows are pointing in the direction of the fluxes' annual mean values and are scaled proportionally to their long-term averages in the Arctic. Most commonly the freshwater budget of the Arctic Ocean is assessed (e.g., Serreze et al, 2006), however, as Schauer and Losch (2019) stated, any outcomes will be arbitrary and ambiguous due to the assumption of a reference salinity.

As results are dependant on the choice of the reference salinity in a non-linear way, different reference values may lead to very different results. Therefore, as there is no explicit physical constraint for a unique reference salinity (Schauer and Losch, 2019) and to avoid the need of choosing a reference salinity, we assess the water volume budget instead. Generally, we focus on the Arctic Ocean (including sea ice) and the overlying atmosphere, however as the Arctic Ocean is almost entirely surrounded by land masses, with some of the world's largest rivers draining into it, runoff R forms one of the key variables in the Arctic water budget. Therefore, for the water budget, we also consider the surrounding land areas draining into the Arctic Ocean including Greenland and Arctic islands. Freshwater is exchanged between the atmosphere and the underlying surface through precipitation P and evapotranspiration ET, where in the Arctic, in general, precipitation exceeds evapotranspiration leading to a positive net precipitation P-ET. As the Arctic is not a closed system, there is also a lateral exchange of water with lower latitudes in atmosphere (mainly water vapor) and ocean (liquid and solid). However, opposed to energy, there is no exchange of water at the top of the atmosphere. For the energy budget there is a net outgoing energy flux at the top of the atmosphere in the Arctic (RAD_{TOA}). It is primarily composed of three key components dictating the balance of energy entering and leaving the atmospheric system at the top of the atmosphere: incoming shortwave solar radiation, reflected outgoing solar radiation, which represents the portion of incoming solar radiation that is reflected back into space by clouds, aerosols, and the Earth's surface, and outgoing longwave thermal radiation, which is the heat emitted from the Earth back into space. Energy exchanges at the surface F_S additionally include turbulent heat fluxes and downwelling longwave radiation being re-emitted from the atmosphere towards the surface. Turbulent heat fluxes comprise latent fluxes, the energy absorbed or released during phase changes of water such as evaporation, and sensible fluxes, the transfer of energy through thermal conduction resulting from temperature differences between atmosphere and surface. Other forms of energy exchange, like for instance enthalpy fluxes associated with mass transfers (precipitation and evaporation) were also considered, but at the end were neglected in the quantitative evaluations (Mayer et al, 2017, 2024), as described in section 2.2. Additionally, energy is exchanged with lower latitudes via lateral transports in atmosphere and ocean. The water and energy cycles are tightly connected as phase changes of water rely on energy - evaporation requires energy (latent heat), while during condensation energy is being released. Also, various other processes link the water and energy budgets together, for instance, precipitation not only redistributes water but also energy, or the effect that condensation of water vapor leads to cloud formation, which in turn plays a significant role in the energy balance.

Evaluating the Arctic energy and water cycles is of major importance, yet accurately estimating various components of these budgets faces challenges due to limited observational data, particularly for the ocean and sea ice. Satellite observations enable the assessment of surface properties such as sea ice concentration and sea surface temperature and give quite precise estimates of the net radiation at TOA, however in-situ data to asses subsurface properties, like for instance vertically resolved ocean temperatures, are limited. Data paucity also concerns other components of the Arctic energy and water budgets, for instance for river discharge about 20%-40% of the pan-Arctic drainage area is unmonitored (Shiklomanov et al, 2002) and additionally the harsh environmental conditions in the Arctic pose a hindrance to gauge measurements due to river freeze-up and flooding. Similarly, also direct precipitation measurements are very sparse, especially so over the Arctic Ocean (Serreze and Barry, 2014).

To bridge these gaps, reanalyses integrate satellite and in-situ data with numerical models offering spatially homogeneous estimates of climatic and environmental quantities. This integration involves a sophisticated data assimilation process where observations and model outputs are systematically combined to produce comprehensive and coherent estimates of the Earth's climate state. They help to improve our understanding of climate dynamics and variability by providing consistent historical data records that are essential for long-term climate studies, trend analysis, and model validation. They have proven as highly useful tools for climate monitoring and have extensively been used to assess atmospheric, oceanic and land variables on global and regional scales. Also, the usage of ocean reanalyses (ORAs) has proved to be useful to examine past ocean states, ocean variability and long-term climate trends (Palmer et al, 2015; Storto et al, 2019; von Schuckmann et al, 2020; Mayer et al, 2021, 2022). Zuo et al (2011) demonstrate a large positive impact of assimilating hydrographic data in an ocean model on the representation of water mass properties and circulation in the Arctic basin. However, the reliability of ocean reanalyses suffers from observational data paucity, especially in the deep ocean (Palmer et al, 2015) and in regions were observations are generally limited like the Arctic. While trends and temporal variabilities of ocean heat content (OHC) were oftentimes found to be realistically represented in reanalyses (Mayer et al, 2021; Cheng et al, 2022), ocean heat transports feature larger biases when compared to observations (Haines et al, 2013; Mayer et al, 2022). In general, OHC is deemed to be more reliable as it is more strongly constrained in ORAs than oceanic transports, as measurements of currents from moorings exist only at selected locations and are not by default assimilated in reanalyses. As part of the Polar ORA Intercomparison Project Uotila et al (2019) assessed the quality of ten ORAs in the polar regions and found that for the mean state deviations from observations were generally smaller for the ORA ensemble than for individual ORAs, caused mainly due to offsetting biases, indicating that especially the use of an ensemble of multiple ORAs can be useful in examining the physical state of the polar regions.

To achieve comprehensive insights into the interconnected dynamics and variations within the Arctic's climatic system, holistic estimates of the coupled Arctic energy and water budgets are needed. Serreze et al (2009) provide such estimates using mainly observations and reanalyses, however partly outdated data sources and imperfect analytical techniques led to significant discrepancies among the budget components, as indicated by large budget residuals. Generally, it is desirable to eliminate or significantly reduce budget residuals and obtain closed budgets with physical terms only. Mayer et al (2019) combine more recent reanalysis and observational datasets as well as ocean transports from a largely model-independent estimate of OHTs compiled from moored buoys in the main Arctic gateways in a mass-consistent way (Tsubouchi et al, 2018). They find considerably smaller budget residuals than Serreze et al (2009), demonstrating the improvements made in observations, data assimilation techniques and diagnostic methods. To further enhance accuracy and reduce uncertainties, additional observations of the deep Arctic Ocean would be essential to more effectively constrain ocean reanalyses.

To evaluate and understand future changes of the Arctic climate, as well as their global impacts, climate models are employed. Climate models are sophisticated computer codes designed to replicate the Earth's complex climate system, to comprehend the dynamics therein and to enable the assessment of climatic conditions in time and space where no direct observations are available. The Coupled Model Intercomparison Project, a global initiative with its latest generation CMIP6 (Eyring et al, 2016), was developed by the World Climate Research Programme (WCRP) as a collaborative framework to improve our knowledge of climate change. By standardizing the experiment designs, CMIP facilitates a structured exploration

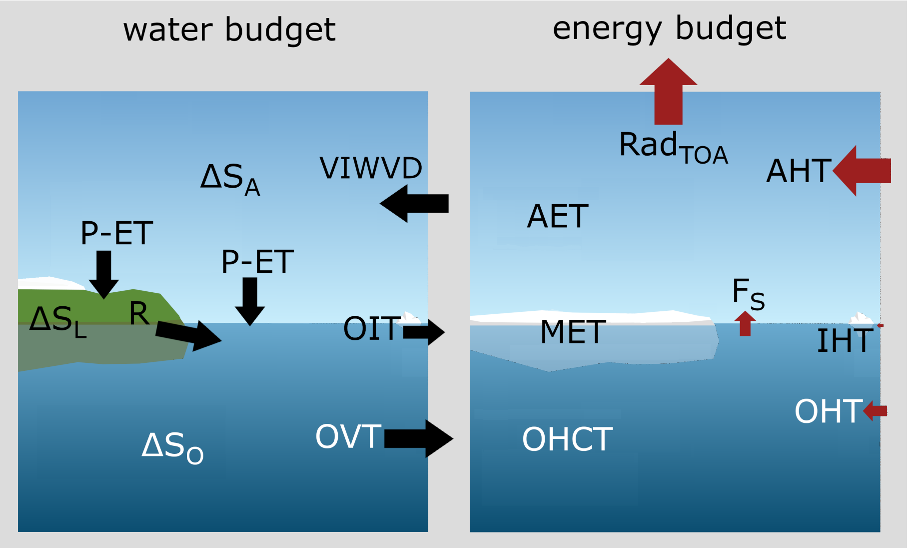

Figure 1: Schematics for the energy and water budget. Arrows are scaled proportionally to the long-term average size of the fluxes in the Arctic. Figures are adapted from (Mayer et al, 2019) and Winkelbauer et al (2022).

of the past and present climate as well as future projections under various greenhouse gas emission scenarios across a wide range of climate models. The project aims not only to refine the accuracy and resolution of climate predictions but also to enhance our understanding of complex climate phenomena and feedback mechanisms. Through these efforts, CMIP contributes to a more robust scientific foundation for assessing the impacts of climate change and informing policymakers and the general public about them. However, the complexity of the Arctic climate system, with the tight couplings between atmosphere, ocean and sea ice, present a major challenge to climate simulations in the Arctic and may introduce biases and large uncertainties (Cai et al, 2021; Knutti, 2008). Therefore, a thorough evaluation of historical climate model simulations against credible observations is needed to identify potential biases and subsequently improve our confidence in future projections of Arctic change.

1.1 Motivation and research goals

Understanding the energy and water budgets in the Arctic and their future changes is of great importance due to their far-reaching implications for the global climate system. Climate models are expected to create reliable predictions of future changes occurring within the Arctic system. However, the complex interplay between atmosphere, ocean and sea ice presents a significant challenge to the accuracy of Arctic climate simulations, introducing large uncertainties and biases even for historical model simulations. This highlights the need for a comprehensive assessment of historical climate model simulations against credible observational data to identify existing deficiencies and enhance the reliability of future Arctic change projections. This thesis is divided into two parts. The first part aims to create reliable observation-based estimates for the Arctic water budget and the second part serves to validate historical simulations of key Arctic water and energy budget terms from CMIP6.

The first objective of this study is to generate up-to-date observationally constrained estimates of the Arctic water budget on annual and seasonal scales. We use a similar approach as Mayer et al (2019) for the energy budget by combining various largely independent, observation-based datasets. Initially, a new observation-based pan-Arctic runoff estimate is generated by combining available gauge observations with credible reanalyses in un-gauged regions. Those are then combined in the water budget equations (see section 2) with atmospheric inputs from various reanalyses, storage rates from satellite observations and oceanic transports from ocean reanalyses to assess the closure of the water budget. Budget residuals are then eliminated in a last step using a variational optimization scheme (see section 2). This motivates the following research question: *How well does the Arctic non-steric water volume budget close, using up-to-date observations and state-of-the art reanalyses and can we reliably remove any residuals to create closed budgets with physical terms only?*

The second and main research goal focuses on the validation of historical climate model runs. Using the observational base created during the first study, as well as (partially updated) energy budget estimates from (Mayer et al, 2019) the following research questions are assessed. *Can climate models accurately reproduce observed seasonalities and trends of the key Arctic energy and water budget components? To what extent do climate models satisfy the energy and water budget equations?*

The third objective of this thesis deals with the technical difficulties of calculating precise oceanic transports of volume and heat on various curvilinear ocean modelling grids as implied for instance in CMIP6. The following research question is formulated. *Can we generate tools to reliably calculate oceanic transports of volume, heat, ice and salt from climate models at any arbitrary section independent of the implied modelling grids?*

This thesis is structured as follows. Section 2 contains information about the employed methods, section 3 presents the three journal articles and section 4 brings together the results of the three research papers and gives an outlook of possible future studies. The employed datasets are described in the respective journal publications.

2 Methods

In this section, all physical and numerical principles relevant for this thesis are described. Sections 2.1 and 2.2 introduce the Arctic water and energy budget equations and section 2.3 describes the variational optimization scheme used to obtain budget closure. Finally, the process of calculating precise oceanic transports of heat, volume and ice is outlined in section 2.4.

2.1 Water Budget

We start by formulating the mass budget for an atmospheric column including water in all states. The field equation for the specific humidity q reads as:

$$\frac{\partial q}{\partial t} + \nabla \cdot q\vec{v} + \frac{\partial q\omega}{\partial p} - S^q = 0 \tag{1}$$

with the horizontal wind vector $\vec{v}$, the horizontal flux of moisture $q\vec{v}$ and the vertical flux $q\omega$. The first term represents the storage of q and the second and third term its flux across the considered volumes horizontal and vertical boundaries. The last term S^q represents the sources of q within the considered volume through phase changes. To obtain the atmospheric water budget equation we integrate vertically from the top of the atmosphere p=0 to the surface of the Earth p=p_s. Following Leibniz rule of integration we can change the order of integration and differentiation for the derivative of q, however as p_s is not constant we have to consider additional marginal flows through changes in surface pressure, which may result through horizontal mass displacements (convergence/divergence) (Chiodo and Haimberger, 2010).

$$\int_0^{p_s} \left(\frac{\partial q}{\partial t} + \nabla \cdot (q\vec{v}) + \frac{\partial q\omega}{\partial p} \right) dp = \frac{\partial}{\partial t} \int_0^{p_s} q dp - q_s \frac{\partial p_s}{\partial t} + \nabla \cdot \int_0^{p_s} (q\vec{v}) dp + q_s\omega_s - \nabla p_s \cdot q_s \vec{v}_s \tag{2}$$

The last term results from changing the order of the divergence operator and the vertical integral. As $\vec{v}_s$ is 0 over land and very small over the ocean this term can be neglected. The second term on the right-hand side of the equation can be omitted for long-time averages, however it has to be considered for shorter time-scales. Furthermore, $\bar{q}\bar{\omega}$ disappears for $p = 0$ and $p = p_s$, however $\overline{q'\omega'}$ at the surface has to be considered as important marginal flow. For the atmospheric water budget this results in:

$$\begin{aligned} \frac{1}{g} \frac{\partial}{\partial t} \int_0^{p_s} q dp &= -\frac{1}{g} \nabla \cdot \int_0^{p_s} q\vec{v} dp - \frac{1}{g} \overline{q'\omega'_{SFC}} + \frac{1}{g} \int_0^{p_s} S^q dp \\ &= -\frac{1}{g} \nabla \cdot \int_0^{p_s} q\vec{v} dp - ET_m - P_m \end{aligned} \tag{3}$$

with the gravitational constant g and surface pressure p_s. The left-hand side of Eq. 4 describes the change of the total mass of water vapour in the atmospheric column (= precipitable water). The first term on the right hand side, the vertically integrated horizontal moisture flux divergence (called VIWVD hereafter), describes the total transport of atmospheric water in the horizontal. The second term describes the vertical flux of water vapor between surface and atmosphere, the evapotranspiration (ET). The last term describes the interaction with condensed water in the atmosphere. It is approximated by the vertical flow of condensed water (= precipitation P). Subscripts ($_m$) for ET and P denote the mass formulation (units of kgm^{-2}s^{-1}).

We will assess volumetric water fluxes in this study, therefore we further divide Eq. 4 by the density of freshwater ρ_w, which is assumed constant at $\rho_w = 1000\ kgm^{-3}$, neglecting dependence on temperature and soluble substances. Further, all terms are integrated over the total considered area to obtain SI units of m^3s^{-1}.

$$\underbrace{\frac{1}{g\rho_w}\frac{\partial}{\partial t}\int_0^{p_s} qdp}_{\Delta S_A} = ET - P - \underbrace{\frac{1}{g\rho_w}\nabla\cdot\int_0^{p_s} q\vec{v}dp}_{VIWVD} \tag{4}$$

The water budget equation for land is derived similarly and reads as:

$$\underbrace{\frac{1}{g\rho_w}\frac{\partial}{\partial t}\left(\underbrace{\int_{p_s}^{p_g} qdp}_{SW} + GS + SD\right)}_{\Delta S_L} = P - ET - \underbrace{\frac{1}{g\rho_w}\nabla\cdot\int_{p_s}^{p_g} q\vec{v}dp}_{R} \tag{5}$$

whereby total land water storage S_L consists of the water stored in the soil (SW) integrated over the soil column (from surface p_s to impenetrable ground p_g), the groundwater storage GS and the snow cover at the soil's surface (snow depth SD). Changes in land water storage are balanced through precipitation P and evapotranspiration ET at the surface as well as runoff R at the surface and within the soil column. Again, all components are integrated over the corresponding areas to yield volumetric tendencies and fluxes.

The mass budget equation for an ocean-sea ice column can be derived similarly. Following Bacon et al (2015) we formulate the oceanic equation including the integration over the considered ocean area. This results in an equation for a closed volume of the integral seawater and sea ice system. The side-wall boundaries are made of ocean or sea ice at defined geographical positions or the coastline. The bottom boundary is assumed to be an impermeable sea bed and the upper boundary is defined by the ocean surface or the surface of sea ice. The mass budget equation reads as:

$$\iiint_V \frac{\partial\rho}{\partial t}dV = \iiint_V \nabla\cdot\rho\vec{v}dV \tag{6}$$

Following Gauss's theorem, the divergence of a vector field in an enclosed volume can be related to the flux of the field through the corresponding closed surface:

$$\iiint_V \vec{\nabla}\cdot\vec{F}dV = \int_\Sigma \vec{F}\cdot\vec{n}d\Sigma \tag{7}$$

with the the closed volume V its complete surface Σ, the vector field $\vec{F}$ and the normal vector of the surface $\vec{n}$. This yields for the right-hand side of Eq. 6:

$$\begin{aligned}\iiint_V \nabla \cdot \rho\vec{v} dV &= \iint_A \rho\vec{v} \cdot \vec{n}\, dA \\ &= P_m - ET_m + R_m + \iint_{A_{sz}} \rho\vec{v} \cdot \vec{n}\, ds\, dz\end{aligned} \tag{8}$$

Where A is the surface surrounding the volume V and $\vec{n}$ is the unit vector normal onto the surface. In the second step the flux is decomposed into its components. Precipitation (P_m), evaporation (ET_m) and runoff from surrounding land areas (R_m) form the vertical fluxes entering the volume through the ocean/sea ice surface. As the seabed is defined to be impermeable, there is no normal flow at the bottom boundary. The last term amounts for the lateral flow of liquid water (OVT) and sea ice (OIT) through the volumes side boundary A_{sz}, which is defined by the horizontal (along-boundary) coordinate s and the vertical coordinate z.

To obtain the volumetric water budget equation we assume ρ to be constant (Boussinesq approximation) and divide Eq. 6 by a reference density ρ_0. This yields an expression where steric effects are ignored and only volume, also considered to be Boussinesq mass ($M_0 = \rho_0 V$), is conserved (Madec et al, 2019; Bacon et al, 2015).

$$\underbrace{\frac{\partial}{\partial t}\iiint_V dV}_{\Delta S_O} = P - ET + R - \underbrace{\iint_{A_{sz}} \vec{v} \cdot \vec{n}\, ds\, dz}_{OVT+OIT} \tag{9}$$

All water budget components, including storage terms and fluxes, are displayed in the left plot of Fig. 1.

2.2 Energy Budget

In the following we will define the energy budget equations for atmosphere, ocean and sea ice individually, as well as for a ocean-sea ice column. In general the energy budget equation describes the balance between vertical energy fluxes, lateral energy transports and the energy storage of the viewed system, a schematic can be seen in the right plot of Fig. 1.

For the atmosphere energy in a closed air parcel can be stored in the forms of internal energy $c_v(T_{air})$, latent heat $L_v(T_{air})q$, potential energy Φ, and kinetic energy k, with their sum being defined as the total energy e. For our budget assessment we follow (Mayer et al, 2019) and utilize a simplified equation for the total energy budget of the atmosphere. To obtain it we start of by formulating the conservation of the total atmospheric energy in an air parcel:

$$\frac{\partial \rho e}{\partial t} + \nabla \cdot (\rho e\vec{v}) = -\nabla \cdot (p\vec{v}) \tag{10}$$

with the total energy

$$e = c_v T_{air} + L_v(T_{air})q + \Phi + k \tag{11}$$

Next we integrate Eq. 10 from the top of the atmosphere (TOA, p=0) to the Earth's surface (p=p_s). We further add source terms in the form of vertical fluxes at the surface (F_S, consisting of the net surface radiation and turbulent heat fluxes) and TOA (F_{TOA}, net radiation at the top of the atmosphere). Further, some assumptions and simplifications are made: all three-dimensional moisture enthalpy fluxes are consistently removed, lateral transports of liquid and frozen water are neglected, all water species are assumed to have the heat capacity of dry air (c_a) and latent heats are taken as constants. While vertical enthalpy fluxes associated with precipitation and evaporation are neglected, the energetic effect of snowfall is kept since it can be significant in high latitudes. For a detailed discussion see Mayer et al (2017, 2024). To retain the notation of the water budget equations we write the storage component on the left side and all relevant fluxes on the right side of the equation:

$$AET = F_{TOA} - F_S - \underbrace{\frac{1}{g}\nabla \cdot \int_0^{p_s} [(1-q)c_a(T_a - T_{00}) + L_v(T_a)q + \phi + k]\vec{v}dp}_{\nabla \cdot F_A} - L_f(T_P)P_{snow} \tag{12}$$

The term on the left-hand-side denotes the tendency of total energy in the atmosphere AET. It is balanced by vertical fluxes at the surface and TOA (both defined positive downward as we follow the direction of integration) as well as the divergence of lateral energy transports $\nabla \cdot F_A$ with atmospheric pressure p, atmospheric water vapor content q, specific heat of dry air at constant pressure c_a=1003 J kg^{-1}K^{-1}, the air temperature T_a and a reference temperature T_{00} (set at the freezing point of seawater T_{00}=-1.8°C; see (Mayer et al, 2019) for detailed discussions), latent heat of vaporization L_v=2.501 × 10^{6}Jkg^{-1}, geopotential ϕ, kinetic energy k and the horizontal wind vector $\vec{v}$. The last term in Eq. 12 is the snowfall term, which accounts for the cooling effect of the surface due to snowfall and the additional latent heat release within the atmospheric column due to net freezing. It is characterized by the latent heat of fusion L$_f$=-0.3337 × 10^6 J kg^{-1} and the snowfall rate P$_{snow}$.

For the ocean we use the vertically integrated sensible heat budget, integrated from the surface to the ocean depth Z, as defined by Mayer et al (2019). Assuming fractional sea ice cover f the ocean energy budget equation reads as:

$$\underbrace{\rho_0 c_p \frac{\partial}{\partial t}\int_0^Z (T_O - T_{00})dz}_{OHCT} = (1-f)F_S + fF_b - \underbrace{\nabla \cdot \int_0^Z \rho_0 c_p (T_O - T_{00})\vec{v}dz}_{\nabla \cdot F_O} + (1-f)L_f(T_p)P_{snow} \tag{13}$$

The term on the left-hand-side denotes the temporal tendency of ocean heat content, with the ocean temperature T_O, the same reference temperature as for the atmosphere above T_{00} and the seawater density ρ_0=1026 kg m^{-3} as well as the specific heat of seawater c_p=3990 kg^{-1}K^{-1}, which are seen as constant. The first two terms on the right-hand side describe the net input of energy at the top of the liquid ocean, which is given by F_S for the ice-free area and by the basal heat flux F_b for the ice-covered area. The third term on the right-hand side denotes the divergence of ocean heat transport with the ocean current vector $\vec{v}$ and the last term again describes the cooling effect of snowfall.

Similarly we formulate the heat budget for sea ice at a grid point with fractional sea ice cover f by comparing storage terms with fluxes and ignoring snow sensible heat:

$$\underbrace{L_f \rho_i \frac{\partial d_i}{\partial t}}_{MET} + \underbrace{\rho_i c_i \frac{\partial [d_i(T_i - T_{00})]}{\partial t}}_{IHCT} = fF_S - fF_b - \underbrace{\nabla \cdot \rho_i d_i L_f \vec{v_i}}_{\nabla \cdot F_I} - L_f \rho_{snow}\frac{\partial d_{snow}}{\partial t} + fL_f(T_p)P_{snow} \tag{14}$$

the first term on the left-hand side describes the required latent heat of fusion for the melting and freezing of ice, known as the sea ice melt tendency MET and the second term is the tendency of sea ice sensible heat. The sea ice density ρ_i is assumed as constant at 928 kgm^{-3}, d_i is the average sea ice thickness per gridpoint, c_i=2106 Jkg^{-1}K^{-1} is the specific heat of sea ice and T_i the temperature of sea ice. The first two terms on the right side describe the vertical fluxes into sea ice. The third term on the right side is the divergence of latent heat transport related to sea ice with the ice drift vector $\vec{v}_i$. The last two terms describe the energy required for snowmelt and the cooling effect of snowfall.

As we will be considering a coupled ocean - sea ice column, we combine Eq. 14 and Eq. 13

$$F_S = OHCT + \nabla \cdot F_O + MET + IHCT + \nabla \cdot F_I - L_f(T_p)P_{snow} + L_f\rho_{snow}\frac{\partial d_{snow}}{\partial t} \tag{15}$$

2.3 Ocean Transport calculations

As described above we may use Gauss's divergence theorem to relate the flux of a vector field through a closed surface to the divergence of the vector field within the region bounded by that surface. Therefore, as we consider the Arctic Ocean bounded by its main gateways as a closed oceanic region, the divergence terms in equations 6, 13, 14 and 15 can be replaced by transports of energy water volume and ice across the lateral boundaries of the considered region, given by the main gateways Fram, Davis and Bering strait as well as the Barents Sea Opening.

Beside the need to know accurate oceanic transports to assess the budget equations and their closure, the assessment of individual oceanic transports and currents is important to understand their crucial role in regulating the Earth's climate. Calculating direct transports instead of divergences gives the added benefit that also transports through individual sections may be calculated. This facilitates the comparison to direct oceanic measurement arrays. However, the calculation of lateral transports for climate models (and reanalyses) is aggravated by the used modelling grids. To avoid the numerical difficulties emerging from the convergence of meridians towards the North Pole in ocean modelling commonly orthogonal curvilinear grids are used where the North Pole is placed over land areas in order to avoid singularities over the ocean (Murray, 1996). Grid-lines are strongly displaced in the northern latitudes when compared to a regular dipolar grid, especially towards the artificial pole(s). Thus, while the numerical problem of the singularity over the ocean is solved, calculations relying on ocean velocities are complicated, as velocities in the direction of the artificial poles do not necessarily point in the direction of the true north. Interpolation onto regular grids is possible but leads to a violation of the model's conservation properties and thus to large errors (see Fig. 9 in Winkelbauer et al, 2024a)). Standard regridding tools, like for instance the python package xESMF (Zhuang et al, 2023), which is built on the ESMF (Earth System Modeling Framework) regridder, assume scalar fields and therefore do not correctly regrid vector fields and can not guarantee conservation of vector properties. Additionally, there is a wide range of different modelling grids, differing in terms of pole number and placement, Arakawa partitions and naming conventions. In the need of a tool to facilitate a consistent comparison of flux estimates from the large set of data sources, a major part of this dissertation was the development of StraitFlux, a new python tool for precise transport calculations on various ocean modelling grids, which is described in the third research paper.

In general oceanic transports through a given section with its starting point x$_s$, its endpoint

x_e and its width x as well as its depth z can be expressed as follows:

$$OT = \int_{x_s}^{x_e} \int_{0}^{z(x)} P\vec{v}_o(x,z) \cdot \vec{n}\, dz\, dx \tag{16}$$

The velocity vector of the oceanic flow is given by $\vec{v}_o$, and the unit vector pointing normal onto the strait by $\vec{n}$. P is the property of the flux, with P=1 for volume transports (OVT) and P=$c_p\rho(\theta(x,z)-\theta_{ref})$ for heat transports (OHT), where θ is the potential temperature, c_p the specific heat and ρ the density of seawater. As true heat transports would actually need closed volume transports, which is generally not the case, a dependence of heat transports on the chosen temperature scale is introduced (Schauer and Losch, 2019). Therefore, each model's heat flux has to be computed relative to a reference temperature θ_{ref}, representing the mean temperature of the assessed flow. This introduces ambiguities as the definition of reference temperatures significantly influences the calculated heat transports (Schauer and Losch, 2019). Such methodological choices can significantly skew results, emphasizing the need for standardized methodologies within the scientific community. In StraitFlux the reference temperature may be chosen to the users needs and it is set universally to 0°C for all models and all Arctic straits in this study to enable easier comparisons between models and to other studies.

The calculation of ice transports proceeds similarly, by integrating the the cross-sectional ice velocity $\vec{v}_i$ multiplied with the unit vector $\vec{n}$ over the grid point average ice depth (d) and along the section width x:

$$OIT = \int_{x_s}^{x_e} d(x)\vec{v}_i(x) \cdot \vec{n}\, dx \tag{17}$$

Latent heat transports associated with sea ice IHT (called $\nabla \cdot F_I$ in Eq. 14) can be estimated by multiplying the ice transports with the sea ice density and the latent heat of fusion.

$$IHT = \int_{x_s}^{x_e} \rho_i L_f d(x)\vec{v}_i(x) \cdot \vec{n}\, dx \tag{18}$$

Additionally, StraitFlux also allows the calculation of salinity transports, which is not relevant for this doctoral thesis and is therefore not discussed further.

In its current implementation StraitFlux works on different curvilinear as well as regular grids. A future advancement could be the implementation of unstructured grids, which is discussed in Section 4.3.1.

2.4 Variational Optimization

To achieve the first research goal of generating up-to-date observation based estimates, we combine a sizeable number of largely independent data sets in the water budget equation. Individually, those estimates represent some of the latest and best estimates of the water budget terms that are currently available. However, when combined in the water budget equations, closure between the various datasets will not be perfect and result in a budget residual. This is driven by data imperfection and simplifications in the used diagnostics. Using the budget equations it is possible to refine the observation based fluxes and storage changes in order to balance the relevant budget equations using physical terms only, while simultaneously remaining within the observational uncertainties of the individual terms. To generate those optimized fluxes we use a variational approach, as already used by (Mayer et al, 2019; Rodell et al, 2015; L'Ecuyer et al, 2015), to enforce budget closure. Therefore, the following cost function J is minimized:

$$J = \sum_i \frac{(F_i - F_i')^2}{\sigma_i'^2} + \lambda \sum_i F_i' \tag{19}$$

With the Lagrange multiplier λ, the adjusted budget terms F_i, the a priori estimates of the budget terms F_i', the uncertainty of the individual budget terms $\sigma_i'^2$ and a budget residual $\sum_i F_i'$.

To obtain best estimates of annual budget components, long-term averages of the individual terms are inserted into Eq. 19. Differentiation in respect to the a priori estimates of the budget terms as well as to λ and solving the system of resulting equations gives us an expression for the adjusted budget terms F_k:

$$F_k = F_k' + \frac{\sigma_k'^2}{\sum_i \sigma_i'^2} \sum_i F_i' \tag{20}$$

So the actual budget residual is divided onto the individual budget terms according to their relative uncertainty. While the a priori uncertainties are derived using the standard deviations of the long-term averaged budget terms, we follow Mayer et al (2018) to calculate the a posteriori uncertainties σ_k^2:

$$\sigma_k^2 = \left(\frac{1}{\sigma_k'^2} + \frac{1}{\sum_i \sigma_i'^2 - \sigma_k'^2} \right) \tag{21}$$

In addition to obtaining average annual estimates, we aim to derive optimal estimates for the mean seasonal cycles of the individual components. It may seem redundant to calculate the annual optimization in the first place, however the observed annual-mean fluxes and their uncertainties are deemed to be more reliable (L'Ecuyer et al, 2015) and budget residuals are generally smaller on an annual scale, leading to different results than soley performing the monthly optimization. Therefore, the monthly optimization process involves two key steps. Firstly, similar to the annual adjustment, fluxes are adjusted for each month separately following Eq. (20). The maximum of the seasonal standard deviations serves as the a priori uncertainty and remains constant across all months. However, after the optimization, the annual means of the resulting monthly fluxes may not align with the annually optimized fluxes. Following the approach outlined by Rodell et al (2015), a second Lagrangian optimization can be applied. In this step, the adjusted monthly fluxes from the first step (F_k) are further adjusted based on their uncertainties, ensuring that their annual mean matches the annually optimized fluxes F_m:

$$FO_k = F_k + \frac{12\sigma_k'^2}{\sum_i \sigma_i'^2} \left(F_m - \frac{1}{12} \sum_i F_i \right) \tag{22}$$

Nevertheless, the second step may again introduce small monthly residuals. Therefore, the entire procedure is iteratively repeated a second time, however using the a posteriori uncertainties obtained from Eq. (21). This yields the desired monthly fluxes, ensuring both consistency with the annually optimized fluxes and a closed budget.

3 Results

This section presents the three peer-reviewed publications of this cumulative thesis, including short overviews, publication details and the main author's contribution to the individual articles.

3.1 Diagnostic evaluation of river discharge into the Arctic Ocean and its impact on oceanic volume transports

3.1.1 Overview

The first publication is an assessment of the Arctic water budget with a special focus on river discharge from Arctic land areas. Initially, available gauge observations are combined with credible reanalyses in un-gauged regions to create a new observation-based pan-Arctic runoff estimate. Then, using largely independent data sources, like reanalyses, ocean reanalyses and satellite as well as in-situ observations and using a non-steric formulation of the Arctic Ocean's volume budget equation, freshwater inputs into the Arctic Ocean via runoff and net precipitation are compared to lateral volume transports over the Arctic's boundaries. To close the budget a variational adjustment procedure (see Sect. 2.3) similar to (Mayer et al, 2019) is used. This yields observation-based best estimates of all water budget terms on annual as well as seasonal scales. This publication addresses the first objective of this thesis of creating observation-based estimates of the Arctic water budget and provides an important basis for the consecutive objective of validating water budget components from climate models.

3.1.2 Publication details

- **Title:** Diagnostic evaluation of river discharge into the Arctic Ocean and its impact on oceanic volume transports
- **Authors:** Susanna Winkelbauer, Michael Mayer, Vanessa Seitner, Ervin Zsoter, Hao Zuo, and Leopold Haimberger
- **Publisher:** Hydrol. Earth Syst. Sci., 26, 279–304
- **Type:** Research article
- **Status:** Published on 24 January 2022, 15 citations on Google Scholar as of March 2024
- **DOI:** `https://doi.org/10.5194/hess-26-279-2022`
- **Own contribution:** My contributions to this work were the acquisition and preparation of the needed data, the analysis and visualization of the results, as well as the preparation of the manuscript under the supervision of MM and LH. The interpretation and discussion of results was done in collaboration with all co-authors. My contributions to this publication are estimated to be at least 75%.

Hydrol. Earth Syst. Sci., 26, 279–304, 2022
https://doi.org/10.5194/hess-26-279-2022

Hydrology and
Earth System
Sciences
Open Access

Diagnostic evaluation of river discharge into the Arctic Ocean and its impact on oceanic volume transports

Susanna Winkelbauer[1], **Michael Mayer**[1,2], **Vanessa Seitner**[1], **Ervin Zsoter**[2], **Hao Zuo**[2], **and Leopold Haimberger**[1]

[1]Department of Meteorology and Geophysics, University of Vienna, Vienna, Austria
[2]European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom

Correspondence: Susanna Winkelbauer (susanna.winkelbauer@univie.ac.at)

Received: 11 June 2021 – Discussion started: 29 June 2021
Revised: 3 December 2021 – Accepted: 14 December 2021 – Published: 24 January 2022

Abstract. This study analyses river discharge into the Arctic Ocean using state-of-the-art reanalyses such as the fifth-generation European Reanalysis (ERA5) and the reanalysis component from the Global Flood Awareness System (GloFAS). GloFAS, in its operational version 2.1, combines the land surface model (Hydrology Tiled European Centre for Medium-Range Weather Forecasts – ECMWF – Scheme for Surface Exchanges over Land, HTESSEL) from ECMWF's ERA5 with a hydrological and channel routing model (LISFLOOD). Furthermore, we analyse GloFAS' most recent version 3.1, which is not coupled to HTESSEL but uses the full configuration of LISFLOOD.

Seasonal cycles as well as annual runoff trends are analysed for the major Arctic watersheds – Yenisei, Ob, Lena, and Mackenzie – where reanalysis-based runoff can be compared to available observed river discharge records. Furthermore, we calculate river discharge over the whole pan-Arctic region and, by combination with atmospheric inputs, storage changes from the Gravity Recovery and Climate Experiment (GRACE) and oceanic volume transports from ocean reanalyses, we assess closure of the non-steric water volume budget. Finally, we provide best estimates for every budget equation term using a variational adjustment scheme.

Runoff from ERA5 and GloFAS v2.1 features pronounced declining trends induced by two temporal inhomogeneities in ERA5's data assimilation system, and seasonal river discharge peaks are underestimated by up to 50 % compared to observations. The new GloFAS v3.1 product exhibits distinct improvements and performs best in terms of seasonality and long-term means; however, in contrast to gauge observations, it also features declining runoff trends. Calculating runoff indirectly through the divergence of moisture flux is the only reanalysis-based estimate that is able to reproduce the river discharge increases measured by gauge observations (pan-Arctic increase of 2 % per decade). In addition, we examine Greenlandic discharge, which contributes about 10 % of the total pan-Arctic discharge and features strong increases mainly due to glacial melting.

The variational adjustment yields reliable estimates of the volume budget terms on an annual scale, requiring only moderate adjustments of less than 3 % for each individual term. Approximately $6583 \pm 84\,\mathrm{km}^3$ of freshwater leaves the Arctic Ocean per year through its boundaries. About two-thirds of this is contributed by runoff from the surrounding land areas to the Arctic Ocean ($4379 \pm 25\,\mathrm{km}^3\,\mathrm{yr}^{-1}$), and about one-third is supplied by the atmosphere. However, on a seasonal scale budget residuals of some calendar months were too large to be eliminated within the a priori spreads of the individual terms. This suggests that systematical errors are present in the reanalyses and ocean reanalysis data sets, which are not considered in our uncertainty estimation.

1 Introduction

Rapid surface warming in the Arctic region has strong impacts on the Arctic water balance and its individual hydrological components, almost certainly leading to an amplification in runoff, evapotranspiration, and precipitation (Rawlins et al., 2010; Collins et al., 2013). Increasing river discharge and precipitation trends and intensified sea ice melt coupled with an increase in freshwater inflow through the Bering Strait lead to an increase in liquid freshwater stored in the Arctic Ocean (Morison et al., 2012; Haine et al., 2015;

Published by Copernicus Publications on behalf of the European Geosciences Union.

Haine, 2020). Ultimately, this could result in enhanced southward exports of low-density waters (Lin et al., 2021) into the Atlantic Ocean, impacting the oceanic circulation also on a global scale. Altogether the hydrological cycle is a complex process with tight coupling between the individual components, having impacts on energy and mass budgets and eventually sea level rise, both regionally (Proshutinsky et al., 2001; Moon et al., 2018) and globally (e.g. Box et al., 2018). Therefore, the quantification of the individual hydrological components and their changes is of great importance.

With the Arctic Ocean being almost entirely surrounded by land masses and some of the world's largest rivers draining into it, the link between ocean and surrounding land is strong. Hence, runoff forms one of the key variables in the Arctic freshwater budget. However, direct quantification of river discharge into the Arctic Ocean is aggravated by the fact that about 30 %–40 % of the pan-Arctic drainage area is now unmonitored (Shiklomanov et al., 2002). Hydrological monitoring suffered a widespread decline from 74 % in 1986 to 67 % by 1999 – in Siberia even 73 % of river gauges were closed between 1986 and 1999 (Shiklomanov et al., 2002; Shiklomanov and Vuglinsky, 2008). In addition, significant portions of the rivers' discharge may bypass the gauging stations through braided channels or as submarine groundwater. Furthermore, climatological conditions pose a hindrance to gauge measurements, as temperatures at the northern latitudes often lead to river freeze-up in late autumn and flooding in spring due to river ice break-up (Syed et al., 2007).

Atmospheric reanalyses produce gridded estimates of atmospheric and land components, providing spatially continuous estimates of variables such as runoff. They represent highly useful tools for climate monitoring. In this study we evaluate runoff from state-of-the-art reanalyses and the Global Flood Awareness System (GloFAS) to provide a best estimate of pan-Arctic river discharge and to incorporate it into the Arctic's freshwater and volume budget. However, data assimilation systems can introduce biases and temporal discontinuities, as changes in the observing system are sometimes inevitable and may lead to inhomogeneities in the time series. One known change is the introduction of the IMS (Interactive Multi-sensor Snow and Ice Mapping System) snow product in ERA5, which led to a negative shift in ERA5's snowmelt and consequently also runoff (Hersbach et al., 2020; Zsótér et al., 2020).

The first complete freshwater budget for the Arctic Ocean is proposed by Aagaard and Carmack (1989) and updated by Serreze et al. (2006) and Dickson et al. (2007). Since then, the amount of available data, in particular of atmospheric and oceanic reanalyses, has opened new possibilities for evaluation of the coupled oceanic and atmospheric energy and hydrological cycles. For example, Mayer et al. (2019) have presented a substantially improved depiction of the Arctic energy cycle compared to earlier assessments. With regard to freshwater or the water volume budget, the data situation has improved as well. New collections of hydrological data of the far north have been published (Shiklomanov et al., 2021b). Tsubouchi et al. (2012, 2018) presented observation-based estimates of volume fluxes through Arctic gateways. This opens the opportunity to overspecify the Arctic volume budgets and thus also to give residual and bias estimates.

This paper is structured as follows. The next section describes the used data and presents the study domain, followed by the methodology. The results are presented in Sect. 4 and are subdivided into seasonal cycles and trends for the four major Arctic watersheds (Sect. 4.1), pan-Arctic seasonalities and trends (Sect. 4.2), and an assessment of budget closure by comparison to oceanic volume fluxes (Sect. 4.3). Section 5 presents conclusions, and in the Appendix a list of acronyms used throughout the text can be found.

2 Data and study domain

Runoff is taken from the European Centre for Medium-Range Weather Forecasts' (ECMWF) fifth-generation global climate reanalysis ERA5 (Hersbach et al., 2020) as well as from its offline simulation ERA5-Land and is downloaded through the Copernicus Climate Change Service (C3S) Climate Data Store (Hersbach et al., 2019; Muñoz Sabater, 2019). Runoff from ERA5 and ERA5-Land are both produced by the land surface model Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land (HTESSEL, Balsamo et al., 2009) of the ECMWF Integrated Forecasting System (IFS). In contrast to ERA5, ERA5-Land is not coupled to the atmospheric model of the IFS, and no direct data assimilation is used; however, observations have an indirect influence, as atmospheric variables from ERA5 are used as atmospheric forcing in ERA5-Land. An advantage of ERA5-Land is the enhanced global resolution of 9 km (31 km for ERA5) (Muñoz Sabater et al., 2021). Runoff data are converted into river discharge (liquid water and ice) by integration over the associated catchment area. We also look into the runoff climatology BT06 (Bourdalle-Badie and Treguier, 2006) that is used in the global ocean ice reanalysis ORAS5.

Furthermore, we consider the GloFAS river discharge reanalysis. GloFAS is developed by ECMWF and the Joint Research Centre (JRC) as part of the Copernicus Emergency Management Service (CEMS). GloFAS is publicly available with data accessible from the Copernicus Climate Change Service Climate Data Store (CDS). Its operational version, GloFAS 2.1 (Harrigan et al., 2019, hereafter denoted $\text{GloFAS}_{\text{E5}}$), combines a simplified version of the hydrological river routing model LISFLOOD (Knijff et al., 2010) to simulate groundwater processes and river routing, with runoff data from HTESSEL, the land surface model used in ERA5 (Harrigan et al., 2020). In addition, we examine an experimental GloFAS version that also uses LISFLOOD's channel routing but forces it with runoff from ERA5-Land – hereafter denoted $\text{GLOFAS}_{\text{E5L}}$ – and GloFAS version 3.1 ($\text{GLOFAS}_{\text{E5}_{\text{new}}}$), which uses the full configura-

Table 1. Positions of gauge observations and GloFAS locations for Ob, Yenisei, Lena, and Mackenzie.

	Gauges	GloFAS
Yenisei	67.48° N; 86.50° E	67.45° N; 86.45° E
Ob	66.57° N; 66.53° E	66.55° N; 66.45° E
Lena	70.70° N; 127.65° E	72.25° N; 126.75° E
Mackenzie	67.45° N; −133.75° E	67.45° N; −133.75° E

tion of the LISFLOOD model and is not coupled to HTESSEL but rather produces its own runoff by using directly precipitation, evaporation, and temperature from ERA5. While $\mathrm{GLOFAS_{E5}}$ and $\mathrm{GLOFAS_{E5_{new}}}$ are available from 1979 to near real time, the experimental version $\mathrm{GLOFAS_{E5L}}$ is only available from 1999 to 2018.

Data from ERA5 and GloFAS are compared to available observed river discharge records. Observing records vary among the different countries and rivers, with the longest time series coming from Russia, where discharge monitoring began in the mid 1930s. In contrast, discharge measurement in North America did not begin until the 1970s (Holmes et al., 2018). The data used in this study come from Roshydromet (Ob, Yenisei, and Lena) and from the Water Survey of Canada (Mackenzie) and were downloaded through the Arctic Great Rivers Observatory (Shiklomanov et al., 2021b). Table 1 shows coordinates of gauge observations and GloFAS sampling locations for Yenisei, Ob, Lena, and Mackenzie. For the pan-Arctic approach, river discharge from an additional 20 rivers was taken. Gauging records for Kolyma, Severna Dvina, Pechora, and Yukon are available for our period of interest 1979–2019 and are also downloaded through the Arctic Great Rivers Observatory (Shiklomanov et al., 2021b), while records for 16 further rivers (Pur, Taz, Khatanga, Anabar, Olenek, Yana, Indigirka, Alazeya, Anadyr, Kobuk, Hayes, Tana, Tuloma, Ponoy, Onega, Mezen) are taken from the Regional Arctic Hydrographic Network data set (R-ArcticNET, Lammers et al., 2001) for the period 1979–1999. We calculated an observation-based pan-Arctic river discharge for the whole period of 1979–2019 by calculating discharge separately for every time step (i.e. every month), using all river discharge measurements available at those time steps. The total pan-Arctic discharge is then obtained by calculating river discharge for the ungauged area at each individual time stamp (using two different calculation methods – see Sect. 4.2) and adding it to the observed discharge. Table 2 lists all runoff and river discharge sources used through this study.

Atmospheric components like precipitation, evaporation, atmospheric storage change, and the divergence of moisture flux (VIWVD) are taken from ERA5, and in Sect. 4.3 we additionally use VIWVD data from the Japanese 55-year Reanalyses JRA55 (Kobayashi et al., 2015) and JRA55-C (Kobayashi et al., 2014), which only assimilate conventional observation data. Land storage is derived from snow depth (given as water equivalent) and soil water changes from ERA5. Groundwater storage is not represented in ERA5 and ERA5-Land, and the representation of frozen land components is also not ideal in HTESSEL, as glaciers are depicted as large amounts of snow which are kept fixed to 10 m of snow water equivalent. When melting conditions are reached, the snow produces a water influx to the soil and consequently contributes to the total runoff. However, the mass balance is not accounted for over glaciers as the snow is restocked to constantly stay at the fixed 10 m level, and hence changes in the glacial storage component cannot be assessed properly. The soil water content includes liquid as well as frozen components and thus also includes permafrost. When the soil temperature reaches melting conditions, the soil water contributes to sub-surface runoff and the soil water storage declines. However, a recent study by Cao et al. (2020) concluded that ERA5-Land soil data are not optimal for permafrost research, due to a warm bias in soil temperature that leads to an overestimation of the active-layer thickness and an underestimation of the near-surface permafrost area. Therefore, we additionally include satellite data from GRACE (Gravity Recovery and Climate Experiment, Tapley et al., 2004) and GRACE Follow-On (Landerer et al., 2020), as land water storage from GRACE includes changes in soil moisture (including permafrost), glaciers, snow, surface water, aquifers, and groundwater. Also, oceanic storage terms are calculated using ocean bottom pressure changes from GRACE. Monthly ocean bottom pressure anomalies and land mass anomalies for the period of April 2002–December 2019 are derived from time-variable gravity observations and are given as equivalent water thickness changes. Due to their limited spatial resolution, GRACE data are prone to signal leakage from land to ocean, and hence we use Mascon solutions (mass concentration blocks, Watkins et al., 2015), which reduce the leakage effect. We examine three Mascon solutions, RL06 v02 from the Center for Space Research at University of Texas, Austin (CSR, Save, 2020; Save et al., 2016), RL06 from the NASA Jet Propulsion Laboratory (JPL, Wiese et al., 2018; Watkins et al., 2015), and RL06 from the Goddard Space Flight Center (GSFC, Loomis et al., 2019) and estimate ocean and land water storage components by taking the mean of those three solutions.

Oceanic volume fluxes through the main Arctic gateways are calculated by integrating the cross-sectional velocity component around the pan-Arctic boundary from the Copernicus Marine Environment Monitoring Service (CMEMS) Global ocean Reanalysis Ensemble Product (GREP, Desportes et al., 2017; Storto et al., 2019), an ensemble of four global ocean reanalyses for the period from 1993 to present. GREP consists of current ocean reanalysis efforts from the Centro Euro-Mediterraneo sui Cambiamenti Climatici (CGLORS, Storto and Masina, 2016), the UK Met Office (FOAM, Blockley et al., 2014), Mercator Ocean (GLORYS, Garric et al., 2017), and ECMWF (ORAS5, Zuo et al.,

Table 2. List of all runoff and river discharge sources.

Product	Description	Variable	Period	
ERA5	Fifth-generation ECMWF reanalysis using IFS (and HTESSEL)	Runoff ($\mathrm{m\,s^{-1}}$)	1979–2019 (back extension to 1950 available)	Hersbach et al. (2020)
ERA5-Land	Offline simulation of ERA5 without DA using HTESSEL	Runoff ($\mathrm{m\,s^{-1}}$)	1981–2019 (back extension to 1950 expected in 2021)	Muñoz Sabater et al. (2021)
$\text{GloFAS}_{\text{E5}}$	ERA5 runoff and simplified LISFLOOD	River discharge ($\mathrm{m^3\,s^{-1}}$)	1979–2019	Harrigan et al. (2020)
$\text{GloFAS}_{\text{E5L}}$	ERA5-Land runoff and simplified LISFLOOD	River discharge ($\mathrm{m^3\,s^{-1}}$)	1999–2018	–
$\text{GloFAS}_{\text{E5new}}$	Full configuration of LISFLOOD	River discharge ($\mathrm{m^3\,s^{-1}}$)	1979–2019	–
BT06	Runoff climatology used in ORCA025	River discharge ($\mathrm{m^3\,s^{-1}}$)	Climatology	Bourdalle-Badie and Treguier (2006)
Observations	Measurements at gauging stations	River discharge ($\mathrm{m^3\,s^{-1}}$)	–	–

2018, 2019). While the GREP ensemble members use the same ocean modelling core and atmospheric forcing (ERA-Interim, Dee et al., 2011), there are differences in used observational data and data assimilation techniques as well as in the reanalysis initial states, NEMO (Nucleus for European Modelling of the Ocean) versions, the sea ice models, physical and numerical parameterizations, and air–sea flux formulations. The data assimilation methods differ in many points, including the deployed assimilation schemes, which range from 3DVAR (three-dimensional variational data assimilation) to SEEK (singular evolutive extended Kalman filter). Furthermore, there are differences in the input observational data set, in surface nudging, in the time windows for assimilation and analysis, as well as in the applied bias correction schemes. All those differences lead to an important dispersion between the reanalysis implementations and add up to the ensemble spread (Storto et al., 2019). For further details, we refer to Storto et al. (2019) as well as the individual data documentations. In addition, volumetric fluxes are derived from moorings within the so-called ARCGATE project (Tsubouchi et al., 2019), covering the period from October 2004 to May 2010. These observation-based estimates of volume fluxes come from a mass-consistent framework that views the Arctic Ocean as a closed box surrounded by land masses and hydrographic observation lines located in the four major Arctic gateways. The hydrographic lines consist of arrays of moored instruments measuring variables temperature, salinity, and velocity, making it possible to calculate fluxes of volume, heat, and freshwater. For more details about the framework, see Tsubouchi et al. (2012, 2018).

Figure 1 presents the study domain. As there is no strict boundary to the south, the definition of the Arctic's geographic extent varies in past studies, and there is no general rule whether to include Greenland and the Hudson Bay or not. We chose our study domain to be consistent with Tsubouchi et al. (2012) as we wanted to compare the oceanic fluxes from ocean reanalysis to the observation-based estimates from the ARCGATE project. The Arctic Ocean is bounded by the position of hydrographic moorings in the main gateways. The Bering Strait forms the only passage to the Pacific Ocean and delivers low-salinity waters into the Arctic, while liquid freshwater and sea ice leave the Arctic Ocean mainly through the Fram and Davis straits. The fourth major strait is the Barents Sea Opening (BSO), where high-salinity waters from the Atlantic Ocean are imported into the Arctic. Furthermore, there are two small passages, Fury and Hecla Strait, that connect the Arctic Ocean with the Hudson Bay.

The terrestrial domain consists of land areas draining into the Arctic Ocean, including the Canadian Arctic Archipelago (CAA) as well as islands along the Eurasian coast. At the Pacific passage the Yukon and Anadyr rivers are considered, as they represent important sources of inflow of low-salinity waters into the Arctic Ocean through the Bering Strait. The total oceanic and terrestrial areas correspond to 11.3×10^6 and $18.2 \times 10^6\,\mathrm{km^2}$ respectively. For volume budget analyses, we further incorporate Greenlandic discharge and storage change north of the Davis and Fram straits, which adds an additional terrestrial catchment area of $0.95 \times 10^6\,\mathrm{km^2}$. The outlines for the individual river catchments were taken from the CEO Water Mandate Interac-

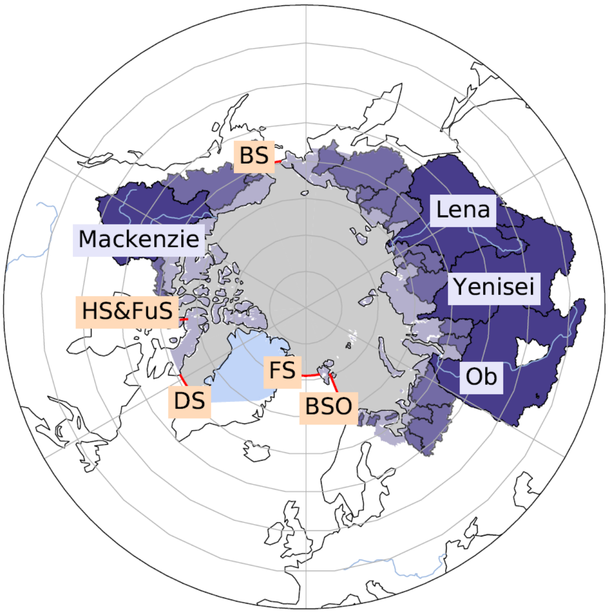

Figure 1. Map of the main study area, consisting of the oceanic area bounded by moorings in the Bering Strait (BS), Davis Strait (DS), Fram Strait (FS), Barents Sea Opening (BSO), as well as Hecla and Fury Strait (HS and FuS) (indicated by grey shading, corresponding to $11.3 \times 10^6\,\mathrm{km}^2$) and the land area draining into it (purple shading, corresponding to $18.2 \times 10^6\,\mathrm{km}^2$ for mainlands and islands and an additional $0.95 \times 10^6\,\mathrm{km}^2$ for Greenland). The colour grading of the land areas indicates the four largest river catchments (Ob, Yenisei, Lena, and Mackenzie; dark purple), the additional 20 catchments with observations, and the smaller catchments and coastal areas where no observations were available (light purple).

tive Database of the World's River Basins (http://riverbasins.wateractionhub.org/, last access: 20 May 2021), and regional outlines like the CAA were taken from the Global Runoff Data Centre (GRDC, 2020).

3 Methods

3.1 Budget equations

A common way to calculate the oceanic freshwater budget is through the assumption of a reference salinity (e.g. Serreze et al., 2006; Dickson et al., 2007; Haine et al., 2015). However, the outcome is dependent on those reference salinities in a non-linear way, so that slight differences in the choice of the reference value lead to very different estimates of freshwater transports, both temporally and spatially. Hence, Schauer and Losch (2019) declared freshwater fractions not useful for the analysis of oceanic regions and rather recommend the usage of salt budgets for salinity assessments. In this paper we do not calculate salt budgets, but we estimate volume budgets and hence also avoid the usage of a reference salinity. Hereinafter, the volumetric budget equations for atmosphere, land, and ocean are formulated.

3.1.1 Atmosphere

The change in water storage in the atmosphere – here expressed as water vapour integrated from the Earth's surface to the top of the atmosphere (i.e. total column water vapour, hereafter denoted as S_A) – denotes the left-hand side of the volumetric budget equation. Atmospheric liquid water and ice are neglected, as they represent only a very small fraction of water in comparison to atmospheric water vapour and lateral moisture fluxes. Generally, atmospheric water in liquid and solid phase are only present in significant amounts in regions with high tropical cumulus clouds and over warm ocean currents (Serreze and Barry, 2014).

The atmospheric storage change (S_A) is balanced by the surface freshwater flux evapotranspiration ET and precipitation P (both in SI units ms^{-1}) and the vertically integrated horizontal moisture flux divergence (last term, hereafter denoted as VIWVD):

$$\frac{1}{g\rho_\mathrm{w}}\frac{\partial}{\partial t}\int_0^{p_\mathrm{s}} q\,\mathrm{d}p = \mathrm{ET} - P - \frac{1}{g\rho_\mathrm{w}}\nabla\cdot\int_0^{p_\mathrm{s}} q\boldsymbol{v}\,\mathrm{d}p, \tag{1}$$

with the gravitational constant g, surface pressure p_s, specific humidity q, density of freshwater ρ_w, and horizontal wind vector $\boldsymbol{v}$. The equation above is probably more familiar to the reader as a mass budget equation, without ρ_w in the denominator. In order to get volumetric water fluxes, we divided by the density of freshwater $\rho_\mathrm{w} = 1000\,\mathrm{kg\,m}^{-3}$, which is assumed constant in this paper, neglecting dependence on temperature and soluble substances. Furthermore, all terms are integrated over the total Arctic area A_total (including the Arctic Ocean and all terrestrial catchments) to obtain SI units of $\mathrm{m}^3\,\mathrm{s}^{-1}$. For presentation of results, we will often use Sverdrup, Sv ($1 \times 10^6\,\mathrm{m}^3\,\mathrm{s}^{-1}$), milli-Sverdrup, mSv ($1 \times 10^3\,\mathrm{m}^3\,\mathrm{s}^{-1}$), or $\mathrm{km}^3\,\mathrm{yr}^{-1}$ as more convenient units.

3.1.2 Land

The change in land water storage (S_L) can be expressed as the sum of changes in volumetric soil water SWVL_n integrated over the corresponding soil depth l_n, changes in snow depth SD – given as snow water equivalent – and changes in groundwater storage GS. Changes in land water storage are balanced through precipitation P_L and evapotranspiration $\mathrm{ET_L}$ over land and runoff R (in SI units ms^{-1}). To obtain volumetric fluxes, we again perform areal integration over the corresponding area (here the land area A_l).

$$\frac{\partial}{\partial t}\left(\mathrm{SD} + \int_{l_n} \mathrm{SWVL}_n \mathrm{d}l_n + \mathrm{GS}\right) = P_\mathrm{L} - \mathrm{ET_L} - R \tag{2}$$

Rearranging Eq. (2), we obtain an estimate of the water discharging into the Arctic Ocean independently of runoff itself. As we prefer analysed quantities, we can further insert Eq. (1) to substitute P_{L} and ET_{L}, which are derived from short-term forecasts from ERA5 through the analysed quantities $\mathrm{VIWVD}_{\mathrm{L}}$ and atmospheric storage change:

$$R = -\frac{\partial S_{\mathrm{L}}}{\partial t} + P_{\mathrm{L}} - ET_{\mathrm{L}} = -\frac{\partial S_{\mathrm{L}}}{\partial t} - \frac{\partial S_{\mathrm{A}}}{\partial t} - \mathrm{VIWVD}. \qquad (3)$$

3.1.3 Ocean

Following Bacon et al. (2015), the oceanic mass budget equation can be expressed as

$$\iiint\limits_{V} \frac{\partial \rho}{\partial t}\mathrm{d}V = F^{\mathrm{surf}} - \iint\limits_{A_{sz}} \rho \boldsymbol{c}\cdot\boldsymbol{n}\,\mathrm{d}s\,\mathrm{d}z. \qquad (4)$$

The left-hand side of Eq. (4) denotes the change in mass over a closed volume, and F^{surf} describes any surface mass input/output in the form of precipitation, evaporation, and runoff. The last term of Eq. (4) denotes ice and ocean side-boundary fluxes from or into the volume, with the horizontal sea water velocity $\boldsymbol{c}$ and integration along the along-boundary coordinate s and depth z. Furthermore, we apply the Boussinesq approximation and assume ρ to be constant. We adopt the reference density used in the Nucleus for European Modelling of the Ocean (NEMO) ocean model of $\rho_{\mathrm{O}} = 1035\,\mathrm{kg\,m^{-3}}$ (Madec et al., 2019) and divide Eq. (4) by ρ_{O}. This yields an expression where steric effects are ignored and only volume, also considered to be Boussinesq mass ($M_{\mathrm{O}} = \rho_{\mathrm{O}} V$), is conserved (Madec et al., 2019; Bacon et al., 2015):

$$\frac{\partial}{\partial t}\iiint\limits_{V} \mathrm{d}V = P_{\mathrm{O}} - \mathrm{ET}_{\mathrm{O}} + R - \iint\limits_{A_{sz}} \boldsymbol{c}\cdot\boldsymbol{n}\,\mathrm{d}s\,\mathrm{d}z. \qquad (5)$$

The change in oceanic volume (derived from bottom pressure changes, hereafter denoted as S_{O}) is balanced by precipitation and evapotranspiration over the oceanic domain (P_{O} and ET_{O}), and runoff from the land domain R and further volume can leave and enter the ocean laterally over its vertical boundaries, described by the last term of the equation (oceanic lateral transport, denoted as F). The liquid portion of F is calculated by integrating the cross-sectional velocity component along the side areas of the Arctic boundary. Additionally, we add ice transports, which are calculated analogously by integrating the cross-sectional ice velocity over the grid-point-average ice depth and integrating it over the Arctic boundary. As volume exchange between liquid ocean and sea ice is conserved in the NEMO model, we additionally remove the liquid water volume that is actually replaced by sea ice, which we call the equivalent liquid water flux. The equivalent liquid water flux at a given grid point is calculated by integrating the liquid volume flux over the grid-point-average ice depth and taking 90 % of the result (as only 90 % of the icebergs are under water). As ice velocities from the public CMEMS data portal are only available from two of the ocean reanalyses (ORAS5 and GLORYS2V4), we calculate the ice flux “correction” term for the GREP ensemble by taking the mean of those two products. However, as the impact of the correction is quite similar for ORAS5 and GLORYS2V4, we believe that the correction is accurate enough for the purpose of this study. Changes in ocean density do not affect the volume as the steric effect is missing due to the Boussinesq approximation.

Table 3. Start and end points of the sections used for lateral flux calculations on the native ORCA grid.

Straits	Latitude/longitude boundaries
Fram	78.80–78.80° N, 20.60° W–11.50° E
Davis	66.60–67.30° N, 61.20–54.00° W
Bering	65.90–65.70° N, 170.00–168.30° W
BSO	77.40–69.70° N, 18.00–20.40° E
Hecla and Fury	69.85–70.00° N, 84.50–84.32° W

In this paper we mostly present monthly means derived by averaging the corresponding fields from reanalyses at their native temporal resolution. Horizontal interpolation and vertical interpolation have been avoided by using all reanalysis products in their native grid representation. Care has also been taken to average over the same area for all products as far as this is possible.

The lateral volume fluxes through the ocean gateways were evaluated along paths on the native ORCA grid, the tripolar grid used in all four GREP ocean reanalyses, that followed the ARCGATE mooring arrays as closely as possible – ORCA coordinates are given in Table 3. This is essential, since the net lateral volume fluxes in and out of the Arctic are very small (∼ 0.2 Sv in the annual mean) compared to the fluxes through individual straits (e.g. ∼ 2.3 Sv for the Davis Strait, Curry et al., 2011).

3.2 Variational approach for budget closure

Given all the various, largely independent data sets we use, closure between the budget terms will not be perfect, resulting in a budget residual. To get rid of any residual and obtain a closed budget with physical terms only, we follow methods by Mayer et al. (2019), L’Ecuyer et al. (2015), and Rodell et al. (2015) and use a variational Lagrange multiplier approach to enforce budget closure on annual and monthly scales. Therefore the following cost function J is minimized:

$$J = \sum_i \frac{\left(F_i - F_i'\right)^2}{\sigma_i'^2} + \lambda \sum_i F_i', \qquad (6)$$

with the Lagrange multiplier λ, the a priori estimates of the budget terms F_i', the adjusted budget terms F_i, the uncertainty of the respective budget term σ'^2_i, and the budget residual $\sum_i F'_i$.

3.2.1 Annual optimization

Inserting the annual means of the individual budget terms into Eq. (6) and differentiating with respect to λ and the a priori estimates of the budget terms yields eight equations with eight unknowns. Solving the system of equations results in an expression for the adjusted budget terms F_k:

$$F_k = F'_k + \frac{\sigma'^2_k}{\sum_i \sigma'^2_i} \sum_i F'_i. \qquad (7)$$

Hence, the budget residual is distributed across the budget terms according to their relative uncertainty. The a priori uncertainties are derived from the standard deviations of the mean annual budget terms. The a posteriori uncertainties σ^2_k are calculated following Mayer et al. (2018):

$$\sigma^2_k = \left(\frac{1}{\sigma'^2_k} + \frac{1}{\sum_i \sigma'^2_i - \sigma'^2_k} \right). \qquad (8)$$

3.2.2 Monthly optimization

Monthly optimization is performed in two steps. First, adjusted fluxes are calculated for each month separately following Eq. (7), whereat the a priori uncertainty is estimated by taking the maximum of the seasonal standard deviations and is kept fixed throughout all months. However, the annual means of the resulting monthly fluxes do not coincide with the annually optimized fluxes. Therefore we follow Rodell et al. (2015) and apply a second Lagrangian optimization, where the adjusted monthly fluxes from the first step (F_k) are adjusted in relation to their uncertainty, so that their annual mean is equal to the annually optimized fluxes F_m:

$$FO_k = F_k + \frac{12\sigma'^2_k}{\sum_i \sigma'^2_i} \left(F_m - \frac{1}{12} \sum_i F_i \right). \qquad (9)$$

However, the second step again generates small monthly residuals. Therefore the whole procedure is performed iteratively a second time, using the a posteriori uncertainties gained through Eq. (8). This results in the desired monthly fluxes that satisfy both a closed budget and consistency with the annually optimized fluxes.

3.3 Trend and relative error calculation

We calculate relative, decadal trends following Zsótér et al. (2020) and Stahl et al. (2012) by applying a linear regression to the annual mean time series:

$$\text{trend} = \frac{10 \cdot \text{slope}}{\text{mean}}. \qquad (10)$$

The "slope" of the time series is the annual change obtained through the linear regression and the "mean" is the long-term annual mean of the time series. The multiplication factor 10 results as we calculate trends over a fixed 10-year period. Hence, a trend of e.g. 0.1 is equal to an increase of 10 % over a decade. All trends are calculated over the common period of the discharge data sets 1981–2019, except for GloFAS$_{\mathrm{E5L}}$, which is calculated over 1999–2018. We do not consider temporal auto-correlation, assuming that subsequent annual means are only weakly correlated, and determine significance using the Wald test with a t distribution, where p values smaller than 0.05 are considered significant.

To compare river discharge estimates from the various reanalyses to river discharge observations, we use Pearson's correlation coefficient r and a normalized root mean square error (NRMSE), which is calculated by dividing the RMSE through the rms of the observed values ($\mathrm{NRMSE}(x) = \mathrm{RMS}(x - \mathrm{obs})/\mathrm{RMS}(\mathrm{obs})$).

4 Results and discussion

We first discuss seasonal cycles and trends for the four major Arctic catchments – Yenisei, Ob, Lena, and Mackenzie. Then we extend our assessments to the pan-Arctic region, where we compare the total terrestrial Arctic runoff with oceanic volume fluxes through the main gateways.

4.1 Analysis of major catchments

4.1.1 Seasonal cycles

Figure 2 shows seasonal cycles of various hydrological components for the Yenisei, Ob, Lena, and Mackenzie catchments. The top panels compare runoff from ERA5 and ERA5-Land to river discharge from GLOFAS$_{\mathrm{E5}}$, GLOFAS$_{\mathrm{E5L}}$, and GLOFAS$_{\mathrm{E5_{new}}}$ as well as to observed discharge values.

Observations show a distinct runoff peak in June due to snowmelt and river ice break-up and weak runoff through winter. The spring flood season of Eurasian rivers depends on the basin size and usually ends by the end of June at small size rivers and by the end of July or beginning of August at large rivers like the Ob, Yenisei, and Lena (Yang et al., 2007). While smaller rivers usually exhibit a low-flow season in the summer–autumn period, discharge from larger rivers mostly shows a slower decrease, also because of summer rainfalls providing additional discharge water. Especially at rivers of eastern Siberia (e.g. Lena), intense rainfalls in the summer–autumn season may occasionally cause rainfall floods (Shiklomanov et al., 2021a). In late winter, with the maximum of river freeze-up, discharge reaches its minimum.

Runoff data from ERA5 and consequently also from GLOFAS$_{\mathrm{E5}}$ underestimate the summer peaks recorded by gauges and reach only about 25 % to 50 % of the observed peak discharge values. In the low-flow season the reanaly-

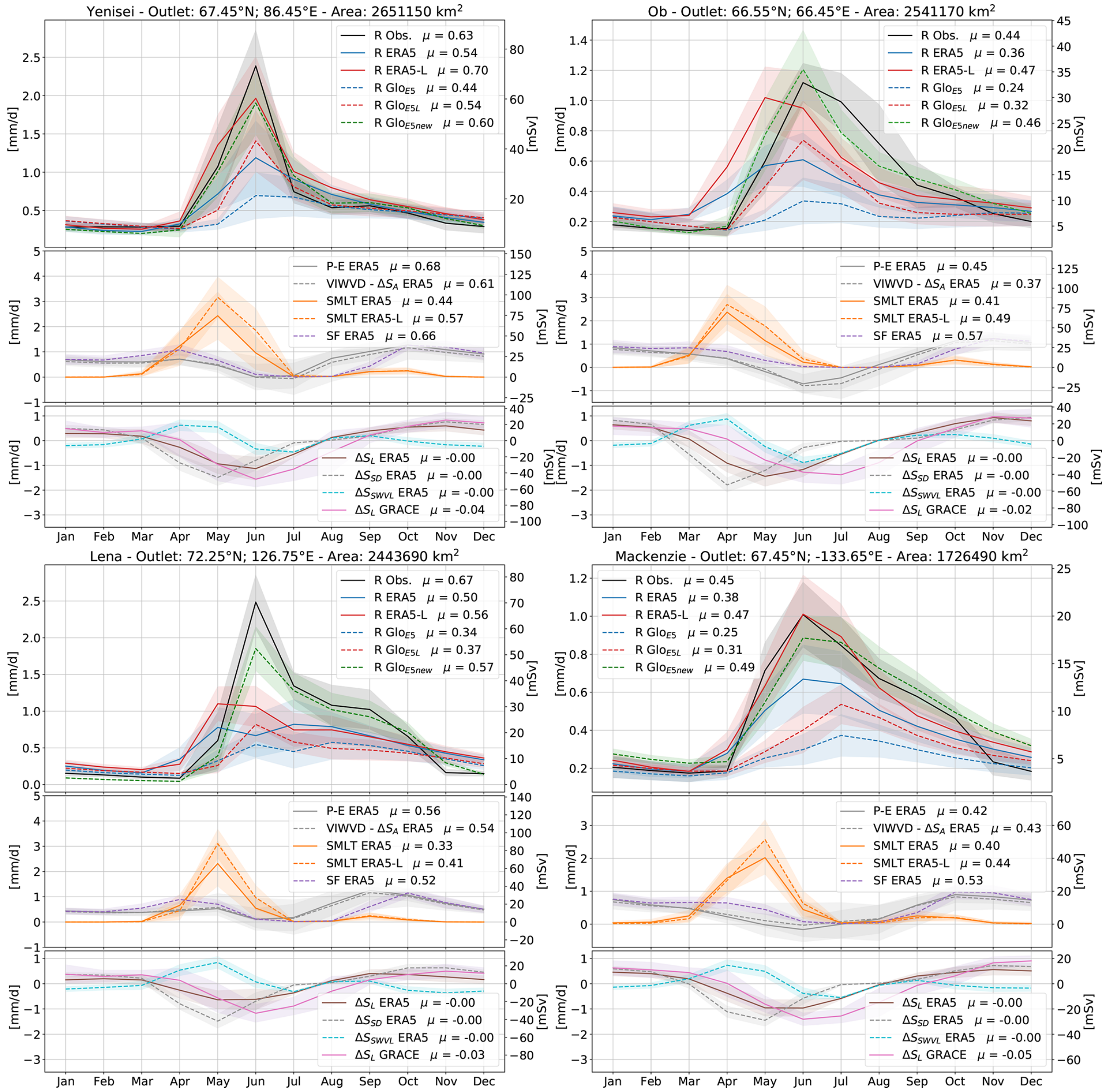

Figure 2. Yenisei, Ob, Lena, and Mackenzie seasonal cycles of runoff and river discharge (R) in the top panel, terrestrial net precipitation ($P-E$), divergence of moisture flux (VIWVD), snowmelt (SMLT), and snowfall (SF) in the middle and land storage change (ΔS_{L}), snow depth change (ΔS_{SD}), and soil water change (ΔS_{SWVL}) in the bottom panel. Seasonal cycles are calculated over the 1981–2019 period (1999–2018 for GloFAS$_{\mathrm{E5L}}$). Shading denotes ±1 SD (1 standard deviation), and μ are the long-term means. Furthermore, geographic coordinates of gauge observations and catchment areas are given. Long-term annual means in mm d^{-1} are given in the legend.

ses slightly exceed observed discharge; however, this does not alter the annual means considerably, resulting in a clear underestimation of discharge by ERA5 and GLOFAS$_{\mathrm{E5}}$ also in annual terms. The difference between ERA5 runoff and GLOFAS$_{\mathrm{E5}}$ discharge is expected to be caused by two sink terms, a groundwater loss component, calibrated in LISFLOOD, that removes water that is lost to deep groundwater systems, and the open water evaporation component, which also removes water through evaporation over water surfaces in LISFLOOD. The negative contribution can reach up to 20 %–40 % of the average flow by any of the two terms in certain regions (described in the LISFLOOD model documentation, Burek et al., 2013). Better accordance in terms of peak height and annual means is achieved by runoff from ERA5-Land – except for the Lena basin – while the sink terms in GloFAS again cause an underestimation by GLOFAS$_{\mathrm{E5L}}$. At the Ob basin the runoff peak in ERA5-Land and also ERA5 occurs in May, i.e. a month earlier than the observed peak.

GloFAS discharge is not in phase with ERA5 runoff but reaches its peak in June, presumably due to the delay by the river routing component. In contrast to $\mathrm{GLOFAS_{E5}}$ and $\mathrm{GLOFAS_{E5L}}$, $\mathrm{GLOFAS_{E5_{new}}}$ reaches values similar to observations and agrees best with the observed values in terms of annual means, peak heights, and seasonality. The middle panels of Fig. 2 show snowfall and snowmelt as well as the atmospheric components net precipitation ($P - E$) and VIWVD. Seasonal cycles of $P - E$ and VIWVD minus storage change agree quite well in terms of peak heights and timing, with low moisture inflow and low net precipitation in summer and higher values in autumn and winter. Annual values of $P - E$ and VIWVD $- \Delta S$ differ by 2–16% depending on the catchment. Seasonal cycles of ERA5 snowmelt show that there is a lag of 1 month between the peak in snowmelt and river discharge. This can partly be explained by the time it takes for the water to reach the river mouth and by water resource management effects but is mostly caused by delayed river ice break-up in the lower parts of the basins. For example, in the upper part of the Ob River, ice breaks up around April to May, while the lower part breaks up between May and June (Yang et al., 2004b). While human impacts through water withdrawals for agricultural use are rather limited compared to rivers at lower latitudes, water resource management via dams and reservoirs can significantly alter the seasonal discharge cycle of the larger Arctic rivers (Shiklomanov et al., 2021a). Especially the Ob and Yenisei basins, but also Lena and Kolyma, are affected by multiple reservoirs using water for hydroelectric power generation and delaying discharge from high-flow periods to the low-flow season (Lammers et al., 2001; Ye et al., 2003; Yang et al., 2004b, a; Shiklomanov et al., 2021a).

The lower panels of Fig. 2 show land storage change from ERA5 and from GRACE. Additionally, ERA5 storage is separated into its components of soil water change ΔS_{SWVL} and snow depth change ΔS_{SD}. In autumn and winter, land water is accumulated through a rise in snow depth, and in summer storage is lost through snowmelt and the associated runoff and to a smaller extent also through evaporation. GRACE shows a significant annual loss of water storage over the past decades, while storage change in ERA5 exhibits only a slight annual decline. In terms of seasonality, GRACE features the largest storage changes in June and July, while ΔS from ERA5 tends to peak 1–2 months earlier. Again, this could be caused by delayed river ice break-up and backwater that is observed by GRACE but not represented in ERA5.

4.1.2 Trends

Figure 3 shows annual river discharge values for the Yenisei, Ob, Lena, and Mackenzie rivers. The corresponding temporal means, standard deviations, decadal trends, normalized RMSEs and correlation coefficients are given in Tables 4 and 5. Observations indicate a slight increase in river discharge over the past decades. The largest changes occurred at the Ob and Lena basins, with a rise of about 4 % per decade. Yenisei shows no significant long-term trend, as, after a rise until the early 2000s, discharge values seem to have slightly decreased over the past decade. In contrast to observations, runoff from ERA5 and discharge from $\mathrm{GLOFAS_{E5}}$ show distinct negative trends of 11 % to 16 % per decade. These strong decreases are a result of changing biases in the snow assimilation in ERA5 and are discussed further in Sect. 3.3.3. The effect of the sink terms removing water in $\mathrm{GLOFAS_{E5}}$ and hence producing a negative shift results in a strong underestimation towards the end of the time series, with values reaching only slightly more than 50 % of the observed values. Runoff from ERA5-Land shows a distinct improvement, with long-term means generally deviating only about 5 % from observations (10 % for Lena) and reasonable NRMSE values. However, the sink terms in $\mathrm{GLOFAS_{E5L}}$ again lead to an underestimation of the observed discharge values. Calculating runoff indirectly through net precipitation minus land storage change and VIWVD minus atmospheric storage change (Eq. 3) yields results mostly within 10 % of the observed discharge values.

Table 4. Mean values and standard deviations of observed river discharge as well as relative decadal trends and their standard errors (see Eq. 10), calculated over the period 1981–2019. Trends that are not significant are marked with a superscript ^{n}.

		Gauge observations	Units
Yenisei	μ	19.2	$(10^3\,\mathrm{m^3\,s^{-1}})$
	σ	1.5	$(10^3\,\mathrm{m^3\,s^{-1}})$
	Trend	0.00 ± 0.01^{n}	(Fraction / decade)
Ob	μ	12.9	$(10^3\,\mathrm{m^3\,s^{-1}})$
	σ	1.8	$(10^3\,\mathrm{m^3\,s^{-1}})$
	Trend	0.04 ± 0.02	(Fraction / decade)
Lena	μ	18.0	$(10^3\,\mathrm{m^3\,s^{-1}})$
	σ	2.5	$(10^3\,\mathrm{m^3\,s^{-1}})$
	Trend	0.04 ± 0.02	(Fraction / decade)
Mackenzie	μ	9.2	$(10^3\,\mathrm{m^3\,s^{-1}})$
	σ	1.0	$(10^3\,\mathrm{m^3\,s^{-1}})$
	Trend	0.02 ± 0.02	(Fraction / decade)

In contrast to $\mathrm{GLOFAS_{E5}}$ and $\mathrm{GLOFAS_{E5L}}$, $\mathrm{GLOFAS_{E5_{new}}}$ exhibits greatly improved river discharge values and, concerning the normalized RMSE values, it provides the best results when compared to observations. Both discharge from $\mathrm{GLOFAS_{E5_{new}}}$ and estimation through $P - E$ also feature slightly negative trends: only calculation through VIWVD exhibits no or even slightly positive trends.

4.2 Pan-Arctic approach

To get a complete estimate of the total amount of freshwater entering the Arctic Ocean via land, which is needed for subsequent budget calculation, river discharge is calculated over

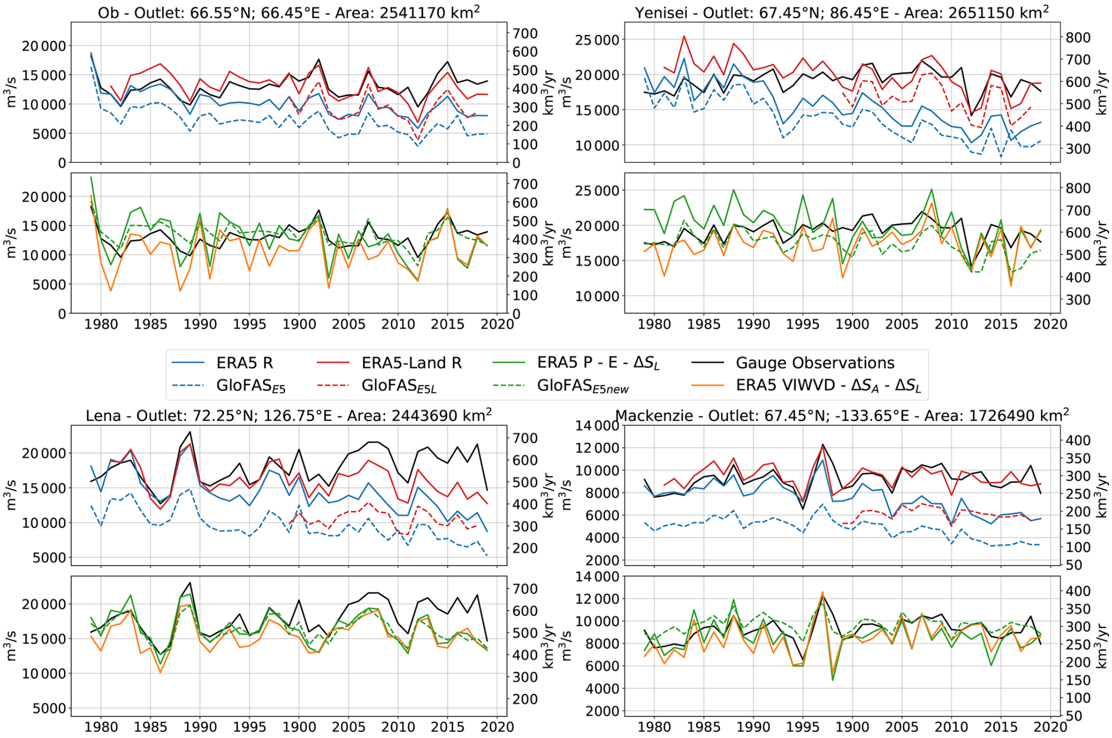

Figure 3. Annual means of observed river discharge, ERA5 runoff, ERA5 net precipitation, ERA5 VIWVD, and ERA5-Land runoff as well as GLOFAS$_{\mathrm{E5}}$, GLOFAS$_{\mathrm{E5L}}$, and GLOFAS$_{\mathrm{E5_{new}}}$ for the period of 1979–2019 (1981–2019 for ERA5-Land and 1999–2018 for GLOFAS$_{\mathrm{E5L}}$).

Table 5. Mean values and standard deviations of river discharge as well as relative decadal trends and their standard errors (see Eq. 10), calculated over the period 1981–2019 (1999–2018 for GloFAS$_{\mathrm{E5L}}$, indicated by the superscript *). Units are the same as in Table 4. Trends that are not significant are marked with a superscript ^{n}. Additionally, normalized RMSE values and correlation coefficients for annual means, calculated in relation to observations (Table 4), are given. Bold values identify the best correlation coefficients and NRMSE values.

		E5 R	E5 $P-E$	E5 VIWVD	E5-L R	GloFAS$_{\mathrm{E5}}$	GloFAS$_{\mathrm{E5L}}$	GloFAS$_{\mathrm{E5_{new}}}$
Yenisei	μ	15.5	19.8	17.4	20.0	13.4	16.4^*	17.3
	σ	3.0	2.9	2.4	2.4	2.9	2.3^*	1.9
	Trend	-0.14 ± 0.02	-0.06 ± 0.02	0.01 ± 0.02^n	-0.06 ± 0.01	-0.16 ± 0.02	$-0.10 \pm 0.05^{*n}$	-0.05 ± 0.01
	NRMSE	0.25	0.15	0.14	0.11	0.34	0.17^*	**0.10**
	r	0.24	0.37	0.49	0.53	0.21	0.81^*	**0.64**
Ob	μ	9.9	12.6	10.6	12.5	6.8	9.3^*	13.4
	σ	1.8	2.9	3.3	2.1	1.7	2.3^*	1.7
	Trend	-0.11 ± 0.02	-0.08 ± 0.03	0.02 ± 0.04^n	-0.06 ± 0.02	-0.15 ± 0.03	$-0.10 \pm 0.99^{*n}$	-0.03 ± 0.02^n
	NRMSE	0.28	0.21	0.25	0.13	0.45	0.32^*	**0.11**
	r	0.39	0.40	0.58	0.66	0.32	0.89^*	**0.68**
Lena	μ	14.2	16.4	15.6	16.1	9.6	10.5^*	16.2
	σ	2.8	2.4	2.2	2.3	2.2	1.3^*	1.8
	Trend	-0.12 ± 0.02	-0.03 ± 0.02^n	0.00 ± 0.02	-0.05 ± 0.02	-0.15 ± 0.02	$-0.03 \pm 0.05^{*n}$	-0.02 ± 0.02^n
	NRMSE	0.27	0.14	0.17	0.15	0.50	0.45^*	**0.14**
	r	0.33	0.65	0.74	0.66	0.27	0.73^*	**0.76**
Mackenzie	μ	7.5	8.6	8.5	9.4	4.8	6.1^*	9.7
	σ	1.3	1.5	1.5	1.0	0.9	0.5^*	0.7
	Trend	-0.11 ± 0.02	-0.01 ± 0.03^n	0.03 ± 0.02	-0.02 ± 0.01^n	-0.13 ± 0.02	$-0.00 \pm 0.04^{*n}$	-0.01 ± 0.01^n
	NRMSE	0.24	0.17	0.16	0.09	0.49	0.36^*	**0.09**
	r	0.29	0.49	0.55	0.65	0.35	0.65^*	**0.73**

the whole pan-Arctic area. First we look at the pan-Arctic seasonal cycle and afterwards at annual means and long-term trends over the whole Arctic drainage area.

Pan-Arctic freshwater discharge estimates vary significantly between different studies. This comes to a big extent from different definitions of the geographic area (Prowse and Flegg, 2000; Shiklomanov and Shiklomanov, 2003) as there is no strict boundary to the south that defines the Arctic, and past studies disagree on whether to include e.g. Greenland and the Hudson Bay or not. Another reason for discrepancies between different studies comes from different approaches to discharge evaluation (Shiklomanov et al., 2021a) – some approaches are based on model simulations and others are derived from river discharge measurements. Using only in situ measurements poses several challenges – especially the handling of the large unmonitored areas, which account for about 30 %–40 % of the total drainage area (Shiklomanov et al., 2002). Most studies adopt the method of hydrological analogy (Shiklomanov and Shiklomanov, 2003), a method where total discharge is calculated by extrapolating gauged runoff to the unmonitored rivers and streams. However, hydrological, climatic, and land cover conditions between gauged and ungauged areas can differ quite a lot, resulting in inaccurate estimates (Shiklomanov et al., 2021a). Hence, we compare two different estimation methods for observed discharge that are displayed in Fig. 4. For that purpose we consider gauging records from the 24 largest Arctic rivers, which account for roughly 70 % of the total drainage area used in this study. Pan-Arctic river discharge for the ungauged areas for each month was estimated as follows: firstly following e.g. Serreze et al. (2006) and applying the method of hydrological analogy for each calendar month by transforming discharge into runoff and, with the assumption that runoff from the ungauged area is the same as runoff from the gauged area, multiplication by the whole drainage area to transform back to river discharge – here denoted as “ae” (area estimate). However, taking runoff from reanalysis (e.g. ERA5-Land) for the same 24 drainage areas and expanding it over the whole area yields different results, with smaller summer peaks, than taking runoff directly from the total drainage area. Hence, monthly correction factors are calculated from our most reliable products ERA5-Land and GloFAS$_{\mathrm{E5}_{\mathrm{new}}}$, and by multiplication by those, a more accurate estimate of observed pan-Arctic river discharge should be possible – denoted as “Ee” (ERA estimate) and “Ge” (GloFAS estimate), leading to more plausible high-flow peaks. Following the results of the previous sections, we trust most in GloFAS$_{\mathrm{E5}_{\mathrm{new}}}$ as it featured the smallest NRMSE values and highest correlation coefficients and hence use the GloFAS estimate Ge for the following pan-Arctic analysis.

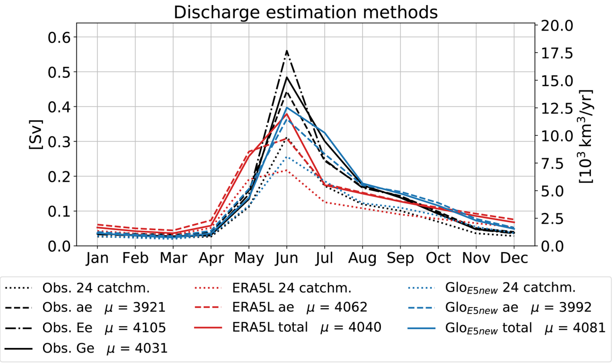

Figure 4. Pan-Arctic discharge estimates using hydrological analogy (ae) and monthly correction factors from ERA5-Land (Ee) and GloFAS$_{\mathrm{E5}_{\mathrm{new}}}$ (Ge) for the period of 1981–1999. Long-term annual means in $\mathrm{km}^3\,\mathrm{yr}^{-1}$ are given in the legend.

4.2.1 Seasonal cycle

Figure 5 shows the mean seasonal cycles of the various pan-Arctic river discharge estimates for the period 1981–2019 (1999–2018 for GLOFAS$_{\mathrm{E5L}}$).

Results are similar to those for the individual catchments, with the largest amounts of water entering the Arctic Ocean in June. Observations (Ge) show summer peaks of about $4.8 \times 10^5\,\mathrm{m}^3\,\mathrm{s}^{-1}$ and annual means of $4031\,\mathrm{km}^3\,\mathrm{yr}^{-1}$. ERA5 runoff and discharge from GLOFAS$_{\mathrm{E5}}$ and GLOFAS$_{\mathrm{E5L}}$ clearly underestimate the summer peaks and annual means; however, they slightly overestimate runoff in the low-flow winter months. ERA5-Land runoff performs quite well in terms of annual means, but the June maximum is still about 25 %–35 % too low. GLOFAS$_{\mathrm{E5}_{\mathrm{new}}}$ performs best in terms of seasonality but still slightly underestimates the summer peak. Indirectly calculated discharge through Eq. (3) does not reach those distinct high June peaks. Annual values from calculations through $P-E$ reach values similar to ERA5-Land runoff, while the VIWVD estimate is about 5 % too low. The runoff climatology Bt06 exhibits a similar seasonal cycle to our observation-based estimate; however, it is roughly 9 % lower than our estimate.

Figure 5b shows land storage change components. The time lag between the observed storage change through GRACE and ERA5 is clearly evident again. Just as for the four major basins, for the sum of all Eurasian and North American catchments (excluding Greenland), GRACE data also show a major decline of land water storage over the past decades, reaching $-131\,\mathrm{km}^3\,\mathrm{yr}^{-1}$ for our area of interest, while land storage from ERA5 shows considerably smaller declines of $-34\,\mathrm{km}^3\,\mathrm{yr}^{-1}$. Additionally, GRACE water storage shows a strong decline of $-115\,\mathrm{km}^3\,\mathrm{yr}^{-1}$ over the north-western part of the Greenlandic ice cap (as specified in Fig. 1), raising the total pan-Arctic storage change to $266\,\mathrm{km}^3\,\mathrm{yr}^{-1}$. The largest changes for GRACE water storage

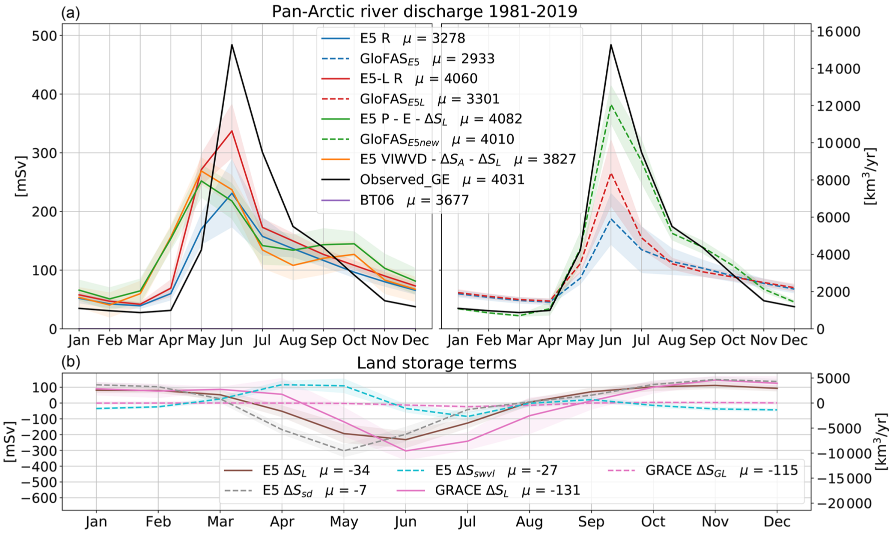

Figure 5. **(a)** Seasonal cycles of pan-Arctic runoff for ERA5, ERA5-Land, and various GloFAS versions as well as indirectly calculated runoff through $P-E$ and VIWVD. Additionally, an estimate of observed river discharge is given. Seasonal cycles are calculated over the period 1981–2019 (1999–2018 for GLOFAS$_{\mathrm{E5L}}$). **(b)** Storage change components from GRACE and ERA5. ERA5 snow depth (dashed brown) and soil water change (dashed cyan) sum up to the total ERA5 storage change (solid brown). Shading denotes variance as ±1 SD (1 standard deviation). Furthermore, long-term annual means in km^3 yr^{-1} are given, and appropriately the right axis is scaled in km^3 yr^{-1}.

occur over the coastal areas of Greenland and the Canadian Arctic Archipelago, and further declining trends can be seen over the Arctic islands Svalbard and Novaya Zemlya as well as over mountainous regions (see Fig. A1 in the Appendix). This suggests a tight linkage to glacial melting.

Wouters et al. (2019) use monthly GRACE Stokes coefficients to examine global glacier mass losses for regions defined in the Randolph Glacier Inventory (RGI), excluding Greenland. Computing the sum over all regions that roughly resemble our area of interest yields a glacial mass change of about −100 Gt (gigatonnes) or roughly −109 km^3 yr^{-1}. Hence, the sum of glacial mass change and storage change from ERA5 resembles the land storage change from GRACE within 10 %.

4.2.2 Trends

Long-term means of annual discharge over the coinciding period of all data sets 1999–2018 (1981–2018 in brackets) and the distribution to Eurasia, North America, and the CAA can be seen in Fig. 6a, and long-term annual means, relative trends, correlation coefficients, and NRMSE values are provided in Table 6. About 75 % of pan-Arctic runoff comes from the Eurasian watersheds, followed by North America, with about 17 %, and the smallest contribution coming from the CAA. Total pan-Arctic runoff values, calculated over 1999–2018, range between 2625 km^3 yr^{-1} for GLOFAS$_{\mathrm{E5}}$ and 3952 km^3 yr^{-1} for indirectly calculated runoff through $P-E-\Delta S_{\mathrm{L}}$ from ERA5. Our observation-based estimates are considerably higher, reaching 4117 km^3 yr^{-1}. Using observations of river discharge only and omitting the CAA, Serreze et al. (2006) obtain values of 3200 km^3 yr^{-1}. Including the CAA but excluding the Yukon River, Haine et al. (2015) combine runoff from ERA-Interim with river discharge observations and report annual values of 3900 ± 390 km^3 yr^{-1} for the period 1980–2000 and 4200 ± 420 km^3 yr^{-1} for the period 2000–2010. Our observation-based estimates, including for the Yukon River, reach values of 3971±31 km^3 yr^{-1} for 1980–2000 and 4155±42 km^3 yr^{-1} for 2000–2010. Subtracting the contribution of the Yukon River (about 200 km^3 yr^{-1}), our estimates are about 5 % lower than those from Haine et al. (2015). Excluding the Yukon River from ERA5-Land runoff and indirectly calculated runoff through ERA5 $P-E$, both estimates are quite close to the estimates made by Haine et al. (2015) for the period 1981–2000; however, they are substantially too low in the 2000–2010 period.

For the whole pan-Arctic region, both ERA5 and GLOFAS$_{\mathrm{E5}}$ feature runoff decreases of 10 %–11 % per decade, and ERA5-Land and indirectly calculated runoff through ERA5 $P-E$ exhibit negative trends of about 2 %–3 % per decade. In contrast, the runoff estimate through ERA5's VIWVD features a slight increase of 2 %, identically to the trend present in our observation-based estimate. GloFAS$_{\mathrm{E5_{new}}}$ comes closest to gauge observations concerning annual means and NRMSE values, while indirectly calculated runoff through VIWVD features the same trend as observations and also the highest correlation concerning annual means. However, the VIWVD estimate generally yields a roughly 5 % lower discharge compared to observations.

Table 6. Mean values and standard deviations of pan-Arctic river discharge in $m^3 s^{-1} \times 10^{-3}$ as well as decadal relative trends, calculated over the period 1981–2019 (1999–2018 for $GLOFAS_{E5L}$, indicated by *). Insignificant trends are indicated by the superscript ^{n}. Bold values identify the best estimates with respect to long-term means, trends, correlation coefficients, and NRMSE values.

	Observed Ge	E5 R	E5 $(P-E)$	E5 VIWVD	E5-L R	$GloFAS_{E5}$	$GloFAS_{E5L}$	$GloFAS_{E5_{new}}$	Units
μ	127.8	104.4	129.4	121.0	128.8	92.2	104.7*	**127.1**	($10^3 m^3 s^{-1}$)
σ	6.1	13.0	7.1	5.9	6.6	11.3	7.0*	5.1	($10^3 m^3 s^{-1}$)
Trend	0.02 ± 0.01	−0.10 ± 0.01	−0.03 ± 0.01	**0.02 ± 0.01**	−0.02 ± 0.01	−0.11 ± 0.01	0.00 ± 0.03*^{n}	−0.01 ± 0.01	(Fraction / decade)
NRMSE	–	0.22	0.06	0.06	0.06	0.30	0.20*	**0.04**	
r	–	−0.14	0.29	**0.78**	0.38	−0.15	0.85*	0.53	

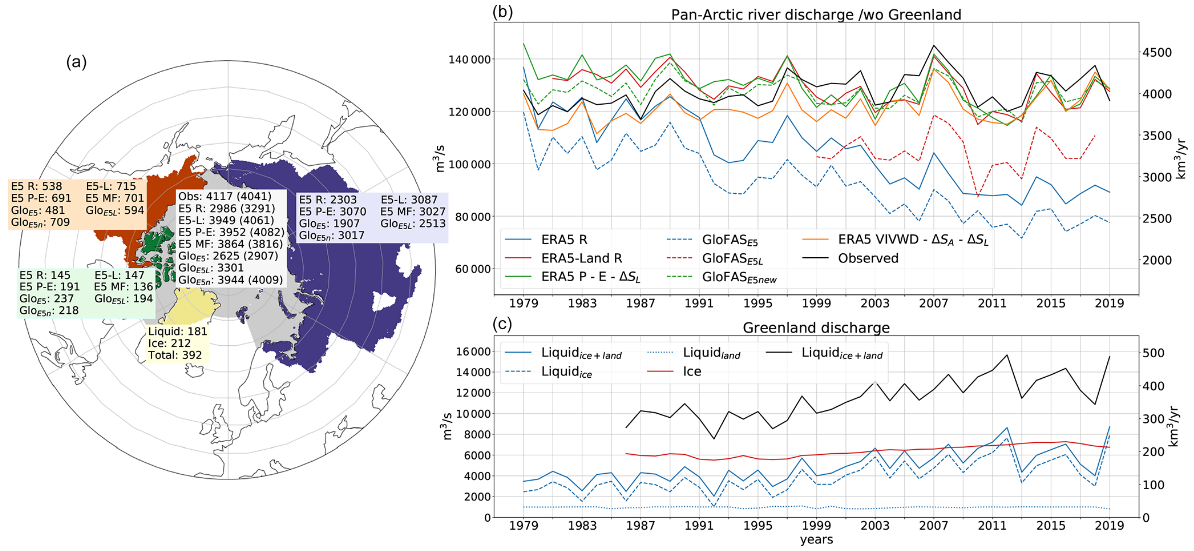

Figure 6. Panel **(a)** shows the drainage area and mean annual discharge values in $km^3 yr^{-1}$ for the whole area as well as for Eurasia, North America, and the CAA individually. Values are calculated over the coinciding period of all data sets 1999–2018 (values in brackets are calculated over the period 1981–2019). The observation-based estimate is calculated using correction factors for the ungauged areas from $GLOFAS_{E5_{new}}$. **(b, c)** Annual means of our observation-based estimate, ERA5, ERA5-Land, $GLOFAS_{E5}$, $GLOFAS_{E5L}$, and $GLOFAS_{E5_{new}}$ runoff as well as indirectly calculated runoff through $P-E$ and VIWVD for the period 1979–2019 (1981–2019 for ERA5-Land and 1999–2018 for $GLOFAS_{E5L}$) are shown. Additionally, Greenlandic liquid (from land and ice; 1979–2019) and solid (1986–2019) discharges are displayed in panel **(c)**.

The strong decreases in ERA5 runoff are reinforced through two discontinuities in the data set, one around 1992 and the second one around 2004, that lead to a significant drop and hence a clear underestimation of runoff over the past decades. As $GLOFAS_{E5}$ takes ERA5 as input, it also adopts those discontinuities. In contrast, ERA5-Land does not exhibit such breaks, suggesting that the error may come from the data assimilation system in ERA5. While the discontinuity around 2004 was traced back to the introduction of the IMS snow product into the assimilation system (Zsótér et al., 2020), the reason for the break around 1992 could not be identified yet and is discussed further in the following section.

In contrast to discharge estimates from reanalyses, observations show opposing trends and a rise in river discharge over the past decades. In NOAA's Arctic Report Card 2018, Holmes et al. (2018) stated that river discharge from Eurasia's largest rivers (Ob, Yenisei, Lena, Kolyma, Pechora, Severnaya Dvina) has increased by 3.3 ± 1.6 % per decade since 1976 and by 2.0±1.8 % per decade for the largest North American rivers (Mackenzie and Yukon). Our observation-based estimates show an increase of 2 ± 1 % for the whole pan-Arctic region.

Runoff from Greenland was left out so far, but contributions of liquid and ice discharge are not negligible. Additionally to Eurasian and North American discharge, Fig. 6 shows discharge from the Greenlandic ice and land areas north of our boundaries in the Davis and Fram straits. Greenlandic liquid discharge from land and ice is taken from Mankoff et al. (2020a), who calculate discharge from daily runoff estimates of the Modèle Atmosphérique Régional (MAR) and the Regional Atmospheric Climate Model (RACMO). We estimate the optimal discharge by taking the mean of MAR and RACMO discharge. Solid ice discharge through calv-

Table 7. Mean values, standard deviations, and trends for liquid discharge from ice and land as well as solid ice discharge for Greenland north of the Davis and Fram straits over the period 1987–2019. Insignificant trends are indicated by the superscript ^{n}.

		$\text{Liquid}_{\text{Ice}}$	$\text{Liquid}_{\text{Land}}$	$\text{Liquid}_{\text{Total}}$	Ice	Total	Units
Greenland	μ	4.2	1.0	5.2	6.7	11.9	($10^3\ m^3\ s^{-1}$)
	σ	1.5	0.1	1.5	0.5	2.0	($10^3\ m^3\ s^{-1}$)
	Trend	0.25 ± 0.05	0.01 ± 0.02^n	0.20 ± 0.04	0.08 ± 0.01	0.13 ± 0.02	(Fraction / decade)

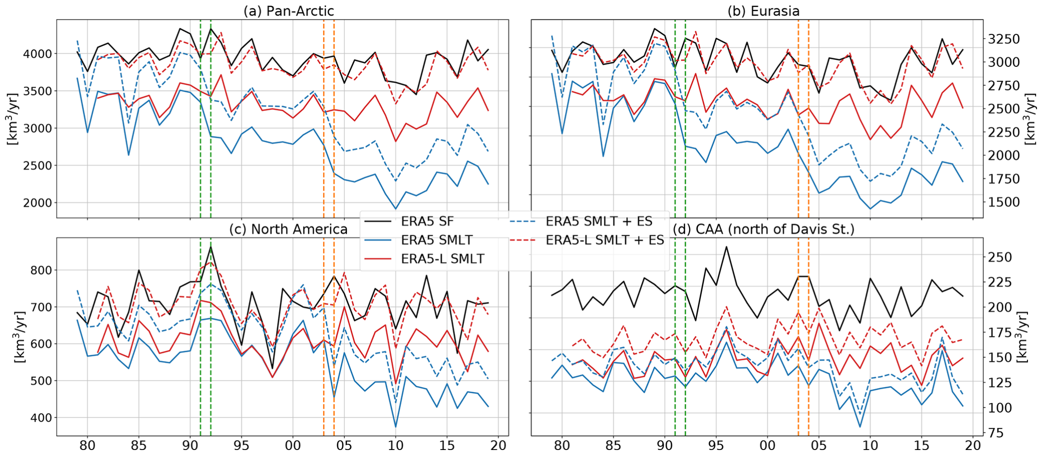

Figure 7. Annual means of ERA5 snowfall (SF, black; 1979–2019), ERA5 SMLT (blue; 1979–2019) and the sum of ERA5 snowmelt and snow evaporation (SMLT + ES, blue dashed; 1979–2019) as well as ERA5-Land snowmelt (red; 1981–2019) and ERA5-Land SMLT + ES (red, dashed; 1981–2019) for the whole pan-Arctic region **(a)**, Eurasia **(b)**, North America **(c)**, and the CAA **(d)**. Vertical lines show the position of the snowmelt discontinuities in 1992 (green) and 2004 (orange).

ing of marine-terminating glaciers and meltwater from ice–ocean boundary melting at submarine glaciers is taken from Mankoff et al. (2020b) over the regions central west (CW), north-west (NW), and north (NO) (see Mankoff et al., 2020b, their Fig. 1) to roughly account for our region of interest. With a total contribution of 392 $km^3\ yr^{-1}$ over the period 1999–2018, discharge from Greenland accounts for roughly 10 % of total pan-Arctic discharge. Most of the freshwater supplied to the Arctic Ocean comes from solid ice discharge, followed by liquid discharge from glaciers and only a small contribution from land runoff. Liquid discharges from land and ice show pronounced seasonalities with peaks in June and July, while solid ice discharge stays relatively stable throughout the year. The bottom figure of the right panel of Fig. 6c shows annual means for the individual Greenlandic discharge components and Table 7 displays mean values and trends. Liquid ice discharge exhibits a vast positive trend of 26 % per decade and solid ice discharge also shows a clear rise of about 8 % per decade. It can be expected that Greenland discharge will further increase in the future (Muntjewerf et al., 2020; Church et al., 2013; Vaughan et al., 2013, e.g.).

Adding Greenlandic discharge to our Ge observation estimate yields a total (pan-Arctic plus Greenland) discharge of 4423 $km^3\ yr^{-1}$. A recent assessment (Shiklomanov and Lammers, 2013; Shiklomanov et al., 2021a) estimates the total discharge to the Arctic Ocean at approximately 4300 $km^3\ yr^{-1}$ for the period 1936–2006. Differences may stem from the use of different data products, from slight variations in discharge area – Shiklomanov and Lammers (2013) use an area of approximately $19 \times 10^6\ km^2$, while our area together with the considered Greenlandic area ($0.95 \times 10^6\ km^2$) sums up to $19.15 \times 10^6\ km^2$ – and from different calculation periods – discharge was very likely smaller in the mid-20th century.

4.2.3 ERA5 runoff discontinuities

As mentioned above, a possible reason for the negative shifts in ERA5 runoff lies in the data assimilation system and its removal of soil moisture. Zsótér et al. (2020) assess the $\text{GLOFAS}_{\text{E5}}$ river discharge reanalysis as well as ERA5 and ERA5-Land runoff and compare them to available river discharge observations. Similarly to our diagnostics for Arctic rivers, Zsótér et al. (2020) find river discharge decreases in $\text{GloFAS}_{\text{E5}}$ for several rivers of the world that are not supported by observations. They find trends in tropical and subtropical areas being driven by changes in precipitation, while changes in snowmelt have a very strong influence on river discharge trends at the northern latitudes. Thus, the runoff decreases in the northern high latitudes are likely linked to snow assimilation and other processes related to snowmelt. Figure 7 shows ERA5 snowfall, snowmelt, and the sum of snowmelt and snow evaporation as well as the corresponding parameters from ERA5-Land for the whole pan-Arctic region, Eurasia, North America, and the CAA. ERA5 shows substantial differences between snow gain and snow loss, and pan-Arctic snowmelt exhibits the same discontinuities around 1992 and 2004 as runoff from ERA5. In contrast, snow gain and snow loss are in balance for ERA5-Land. Zsótér et al. (2020) find similar results and attribute the differences to the land data assimilation, which has impacts on snow and soil moisture. The discontinuity in 2004 was traced back to a change in operational snow analysis through the introduction of the 24 km IMS snow cover information to the snow assimilation system in 2004 (Zsótér et al., 2020). While the discontinuity around 2004 is present in the Eurasian watershed as well as in North America, the discontinuity around 1992 is only present in Eurasia. Figure 8 shows the spatial distribution of the snowmelt discontinuities, calculated as differences between snowmelt climatologies from 1979–1991 and 1992–2003 as well as 1992–2003 and 2004–2019. The latter difference exhibits spurious signals at the coastal mountain range of Alaska, the Rocky Mountains, and mountainous regions in Siberia, and the prior discontinuity spots rather random signals only over Eurasia.

This discontinuity issue is not limited to ERA5 only but rather a general issue in reanalyses, as observation platforms are changing through time, making it practically very hard to make these products perfectly homogeneous in time. Satellite data especially were not available in the early days and were introduced with the development and introduction of new instruments. If the redundancy is large enough, then any discontinuity impact should be less pronounced. However, specifically for snow there is only the IMS product that was introduced in 2004, and hence any inhomogeneity generates a larger impact. Thus, this impact could possibly be reduced when using other data sets on top of the IMS product or instead of it, which ideally go further back in time.

4.3 Comparison to oceanic fluxes

In this section we look at the oceanic volume budget as described in Eq. (5). For this purpose we calculate volume fluxes through the oceanic gateways – BSO, Fram, Davis, Bering, and Hecla and Fury straits – using ocean reanalyses and compare them to our various freshwater input estimates. Oceanic storage change is derived from ocean reanalyses and from GRACE. To close the budget and get rid of any residuals, we use a variational approach (see Sect. 2.1.2).

Figure 9 shows seasonal cycles of lateral net volume fluxes from ORAS5 and from the GREP ensemble for 1993–2018.

The GREP ensemble shows an annual mean lateral transport out of the Arctic region of 203 ± 48 mSv, with ORAS5 being on the higher end of the ensemble with an annual outflow of 232 ± 48 mSv and a summer peak of 578 mSv. The seasonal cycles of the oceanic volume transports resemble the seasonal cycles of the freshwater input and peak in June, as the ocean reacts almost instantaneously to surface freshwater input and generates barotropic waves that lead to mass adjustment within about a week (Bacon et al., 2015).

In addition to the oceanic transports, Fig. 9a shows the forcing terms involved in the generation of the ORAS5 fluxes:

$$F_{\mathrm{ORAS5}} = (P_{\mathrm{O}} - \mathrm{ET}_{\mathrm{O}})_{\mathrm{ERAInterim}} + R_{\mathrm{BT06}} - \frac{\partial S_{\mathrm{ORAS5}}}{\partial t} + \mathrm{damp}_{\mathrm{ORAS5}} + \mathrm{FW}_{\mathrm{adj}}. \quad (11)$$

In addition to net precipitation from ERA-Interim and river discharge from the BT06 climatology, we need to add a surface salinity damping term ($\mathrm{damp}_{\mathrm{ORAS5}}$), which represents an additional non-physical surface freshwater forcing in the ocean reanalyses (last term of Eq. 6.1 in Madec et al., 2019). Additionally, a freshwater adjustment term ($\mathrm{FW}_{\mathrm{adj}}$) should be added due to the assimilation of global mean-sea-level changes. This term is not archived in ORAS5 and hence is not considered in our analysis. However, we deem the freshwater adjustment term to be a rather small source of error and nevertheless would expect good closure concerning Eq. (11). Indeed, with an annual deviation of 11 mSv and an NRMSE value of 0.19, the agreement between F_{ORAS5} and its freshwater input estimate is reasonable. In addition to the FW adjustment term, another cause of the discrepancies between forcing and computed volume fluxes is that the ocean reanalyses compute their own turbulent air–sea fluxes and do not use that from ERA-Interim. However, given the generally low values of sensible and latent heat fluxes, in the Arctic we consider this a moderate source of error.

Figure 9b shows various estimates of freshwater input minus oceanic storage change. To estimate the atmospheric freshwater input over the ocean, we take net precipitation from ERA5 as well as VIWVD from ERA5, ERA-Interim, and the Japanese 55-year Reanalyses JRA55 and JRA55-C. For the reanalysis-based estimates, we use the river discharge reanalyses we have most confidence in, namely $\mathrm{GloFAS}_{\mathrm{E5}_{\mathrm{new}}}$, ERA5-Land runoff, and the indirect estimates through $P - \mathrm{ET}$ and VIWVD minus storage change. Land storage is only derived from GRACE, while oceanic storage is taken from GRACE and from ORAS5. GRACE and ORAS5 show quite similar seasonal cycles, with mass being accumulated by the ocean in summer and released in winter, most likely caused by the seasonal variation of wind stress curl and seasonal changes in Ekman pumping (Bacon et al., 2015). In terms of annual trends, both ORAS5 and GRACE indicate a slight increase in mass over the past decades. The observation-based estimates are calculated using river discharge from our observation estimates (ae, Ee, Ge) and oceanic storage change from GRACE. The atmospheric components VIWVD and atmospheric storage change over the ocean are still taken from reanalyses (ERA5, ERA-Interim, JRA55, JRA55c), as we lack the corresponding observa-

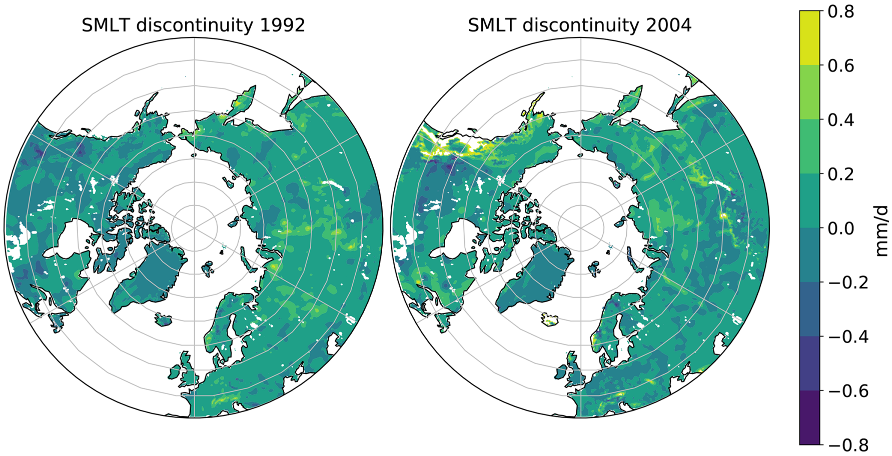

Figure 8. ERA5 snowmelt discontinuities, calculated by subtracting the long-term SMLT means from 1979–1991 and 1992–2003 and respectively 1992–2003 and 2004–2019.

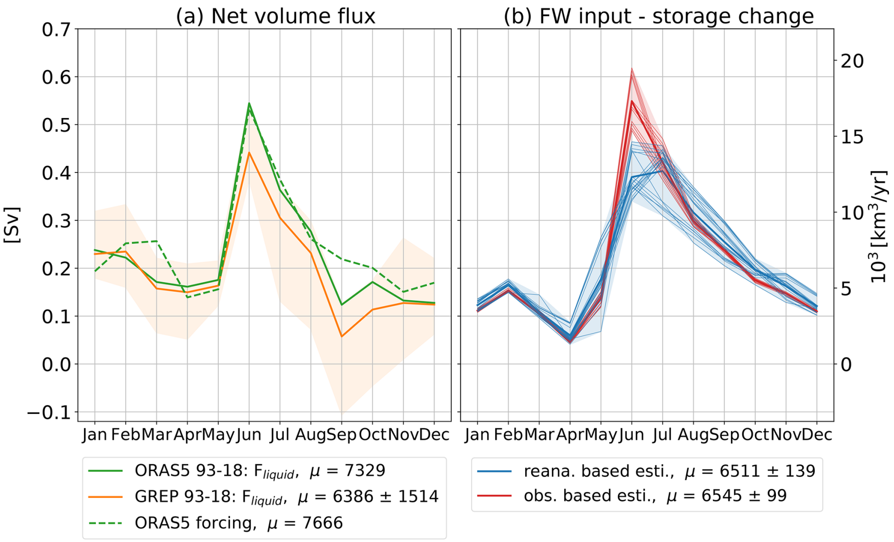

Figure 9. **(a)** Mean seasonal cycles of lateral net volume fluxes (defined as positive out of the Arctic) from ORAS5 (1993–2018) and GREP (1993–2018) as well as the sum of forcing terms generating ORAS5. **(b)** Different realizations of observation- and reanalysis-based estimates of freshwater input minus oceanic storage for the period 1993–2018.

tions. Greenlandic discharge is taken from Mankoff et al. (2020a, b). Calculating the mean over our reanalysis- and observation-based estimates, we obtain annual values of 207 ± 4 and 208 ± 3 mSv respectively. Thus, compared to the GREP volume flux, we obtain an imbalance of 4–5 mSv, representing less than 3 % of the physical fluxes.

Figure 10 shows the budget residuals, calculated as GREP ensemble average volume transport minus the various freshwater estimates. From January to April residuals are generally positive, as the freshwater input from atmosphere and land is low due to river icing and low precipitation values; however, freshwater that entered the Arctic in the past season and moves slowly with the ocean currents still finds its way through the Arctic gateways. In June, the large freshwater peak generates barotropic waves that are also seen in the reanalysis volume transports. Residuals arise from different river discharge estimates and are generally positive for reanalysis- and negative for observation-based estimates. In late summer, river discharge tends to decline and precipitation takes over the main role of delivering freshwater to the ocean. Oceanic reanalyses show a fast decline after the June peak and exhibit a volume transport minimum in September leading to a negative residual peak in all the estimates. We assume that volume that enters the Arctic as freshwater through precipitation is slowly transported with the ocean currents and modified by wind stress, partly freezes on the

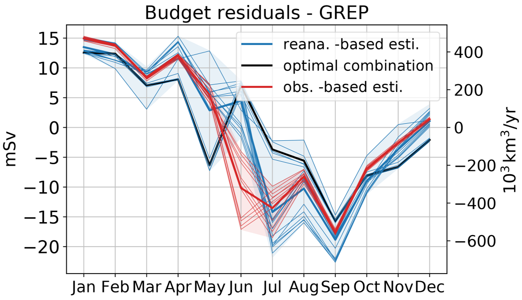

Figure 10. Mean annual cycles of different realizations of the budget residuals (all 1993–2018; reanalysis based: blue lines, observation based: red lines) and the optimal combination yielding the smallest NRMSE value (black line).

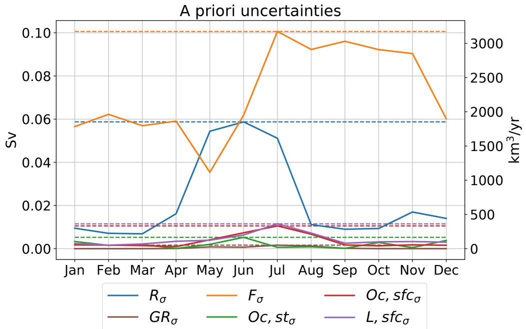

Figure 11. Mean annual cycles (1993–2018) of spread for runoff (R), Greenland runoff (GR), lateral ocean fluxes (F), oceanic storage (Oc, st), and atmospheric input over ocean (Oc, sfc) and land (L, sfc). Spread is calculated as the standard deviation between various estimates (see Table 9) of the individual terms and as an RMSE estimate for GRACE.

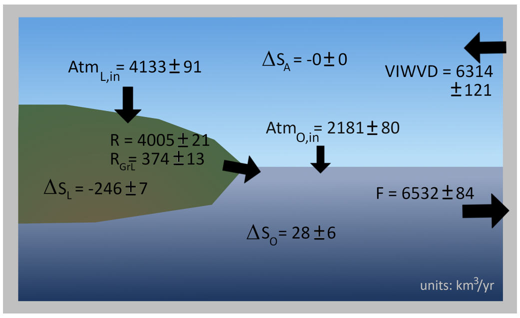

Figure 12. 1993–2018 adjusted long-term means of atmospheric and terrestrial freshwater fluxes and storage terms in the Arctic hydrological cycle as well as oceanic volume transport and storage change. Units are $\mathrm{km^3\,yr^{-1}}$; arrow areas are scaled by the magnitude of the represented terms.

way, and takes weeks or up to months to leave the Arctic area, explaining the elevated winter and spring transports.

Table 8 shows annual means of the various transport estimates, standard deviations, and NRMSE values. The normalized RMSE values are calculated with respect to GREP (^{1}) and ORAS5 (^{2}), indicated by the superscripts. In addition to the estimates from Fig. 9, the observation-based mean volume transport from the ARCGATE project (Tsubouchi et al., 2019, 2012, 2018) is also given. To compare the ARCGATE value to ORAS5 and GREP, mean values and NRMSEs are calculated over the ARCGATE period (October 2004–May 2010, indicated by a *). With an annual mean volume transport of 151 mSv, the ARCGATE estimate is about 25 % lower than the GREP estimate. The high NRMSE values emerge as the observation-based ARCGATE flux does not show any peaks in June but rather stays low throughout the summer (not shown). This is probably a result of the mooring arrays in the gateways being too sparse and hence the velocity field not being measured accurately enough to dissolve the barotropic wave signal. Instead, ARCGATE shows the export of freshwater that travels with the oceanic currents, dilutes on the way, and remains around the continental shelves for some time. Hence, the oceanic reanalyses and the observation-based ARCGATE estimate should only be compared in terms of annual means and not seasonalities.

4.3.1 Volumetric budget closure

To close the volumetric budget and get rid of the small residual, we use a variational adjustment procedure (see Sect. 2.1.2). A priori estimates are calculated by taking the mean over all trustworthy estimates of the individual budget terms – given in Table 9. We perform the adjustment on long-term annual means to close the annual budget and on the monthly climatologies of the individual terms to close the budget on a monthly scale. The uncertainties for the annual optimization are estimated as the standard deviations between the various estimates (see Table 9) of the mean annual budget terms. The uncertainties for the monthly optimization are estimated by taking the annual maximum of the monthly standard deviations, indicated by the dashed lines in Fig. 11. Oceanic lateral volume transport F features the largest uncertainties and hence also experiences the largest adjustments amongst all budget terms.

Figure 12 shows the adjusted long-term annual volume budget over the oceanic and land domains. Oceanic storage indicates a slight increase ($28 \pm 6\,\mathrm{km^3}$) in mass over the past decades. About $6532 \pm 84\,\mathrm{km^3}$ leave the Arctic Ocean through the main gateways on an annual basis. From the water entering the Arctic Ocean through its surface, roughly two-thirds are supplied by runoff ($4379 \pm 25\,\mathrm{km^3}$) and roughly one-third ($2181 \pm 80\,\mathrm{km^3}$) is freshwater delivered by the atmosphere. Oceanic volume transports out of the Arctic domain exceed the atmospheric moisture entering the Arctic ($6314 \pm 121\,\mathrm{km^3}$) by nearly 4 %, indicating an annual loss of water volume of roughly $220\,\mathrm{km^3}$. This loss is mainly driven by terrestrial water mass losses. Even though the representation of frozen land components is not

Table 8. 1993–2018 (*2004–2010) long-term means, NRMSEs, and correlation coefficients of oceanic volume transports from GREP, ORAS5, and ARCGATE as well as estimates derived from the forcing terms used in ORAS5 (as described in Eq. 11). Additionally, our reanalyses and observation-based estimates are given. NRMSE values and correlation coefficients are calculated using monthly means with respect to GREP(^{1}) and ORAS5(^{2}). Units are milli-Sverdrup (mSv).

	GREP	ORAS5	ORAS5 forcing	ARCGATE	Reana.-based estimate	Obs.-based estimate
μ (mSv)	203/200*	232/221*	243	151*	207	208
NRMSE	–	0.22^1	$0.34^1/0.19^2$	$0.79^{1*}/0.83^{2*}$	$0.50^1/0.38^2$	$0.52^1/0.35^2$
r	–	0.97^1	$0.87^1/0.94^2$	$-0.39^{1*}/-0.52^{2*}$	$0.64^1/0.77^2$	$0.74^1/0.87^2$

Table 9. List of employed data sets in the variational adjustment procedure.

F	GREP
R	$\mathrm{Obs_{Ge}}$, $\mathrm{GloFAS_{E5_{new}}}$, ERA5-Land
$R_{\mathrm{Greenland}}$	Liquid and solid from Mankoff et al. (2020a, b)
$\mathrm{Ocean_{Atm,in}}$	VIWVD from ERA5, ERA-Interim, JRA55, JRA55c
$\mathrm{Ocean_{stor}}$	GRACE (CSR, GSFC, JPL)
$\mathrm{Land_{Atm,in}}$	VIWVD from ERA5, ERA-Interim, JRA55, JRA55c
$\mathrm{Land_{stor}}$	GRACE (CSR, GSFC, JPL)

Table 10. 1993–2018 adjusted mean annual cycle of the Arctic Ocean's hydrological cycle; units are km^3 yr^{-1}.

	R	GR	$\mathrm{Atm_{O,in}}$	ΔS_{O}	F
Jan	2264	204	1415	−820	4703
Feb	1770	203	1305	−2361	5640
Mar	1412	203	1385	−830	3830
Apr	2240	207	1487	1252	2682
May	5932	264	1433	2046	5583
Jun	12 775	679	2659	1543	14 570
Jul	7376	1242	3943	301	12 260
Aug	4764	628	3943	543	8792
Sep	3019	234	3265	531	5988
Oct	2826	210	2075	−203	5313
Nov	1973	207	1772	−1098	5050
Dec	1708	205	1491	−572	3976
Mean	4005	374	2181	28	6532

ideal in HTESSEL, the comparison of GRACE mass changes to the sum of ERA5 storage changes (snow and soil water) and glacial changes taken from the literature (e.g. Wouters et al., 2019) agree well. Therefore, we combine land storage changes from ERA5 (excluding glaciers) with storage changes from GRACE to estimate contributions of different terrestrial sources to the diagnosed storage decline in the Arctic. We find that approximately 47 % of the 246 km^3 yr^{-1} decline is generated through liquid and solid discharge from Greenland, while about 45 % comes from Arctic glaciers (excluding Greenland) and the remaining 8 % is the result of a decline in land water storage due to permafrost and snow cover reduction.

The reported increase in oceanic storage is driven by mass increases over the western Arctic Ocean and the coastal areas of Eurasia and North America (see Fig. A1 in the Appendix). For the western Arctic Ocean, various studies indicate an accumulation of freshwater in the Beaufort Gyre due to a combination of favourable wind forcing, redirection of Mackenzie River discharge, inflow of low-salinity waters through the Bering Strait, and sea ice melt (e.g. Proshutinsky et al., 2019). Mass increases along the coastal areas are a result of runoff increases (Ludwigsen et al., 2020; Morison et al., 2012). Mass decreases in the Barents and Kara seas as well as in Baffin Bay (Fig. A1 in the Appendix) counteract those increases, resulting in the reported weak trends for the whole Arctic Ocean. Mass decreases west of Greenland are mainly caused by lowering of the geoid associated with nearby ice mass losses. However, Jeon et al. (2021) also found residual land leakage effects that were not removed in GRACE Mascons. Furthermore, oceanic storage is strongly affected by decadal wind variations (Volkov and Landerer, 2013; Fukumori et al., 2015) and circulation patterns and exhibits strong non-seasonal fluctuations, further complicating the detection of long-term oceanic mass trends. Volkov and Landerer (2013) find that an intensification of the westerly winds over the North Atlantic and over the Russian Arctic continental shelf leads to a decrease in ocean mass in the central Arctic. Furthermore, they found a positive correlation between Arctic Ocean mass fluctuations and northward wind anomalies over the Bering Sea and the north-eastern North Atlantic. They also reveal that cyclonic/anticyclonic anomalies of the large-scale ocean circulation lead to negative/positive Arctic ocean mass anomalies.

Table 10 provides adjusted monthly and annual climatologies, and Fig. 13 shows seasonal cycles of the oceanic budget terms before and after the adjustments. A metric to identify whether the variational adjustment is successful is the comparison of the adjustments of the terms to their respective a priori uncertainty (L'Ecuyer et al., 2015), and hence Fig. 13 also shows the spreads of the a priori estimates (shaded areas). Due to the large differences between oceanic volume transports and freshwater input terms in late winter to early spring and in September, both F and runoff R feature adjustments larger than their a priori spreads, suggesting that the

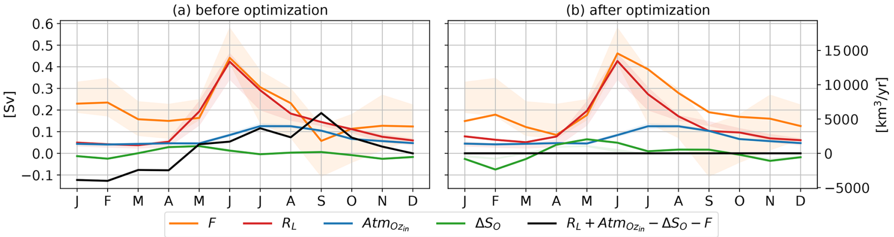

Figure 13. 1993–2018 adjusted mean annual cycles of atmospheric and terrestrial freshwater fluxes in the Arctic hydrological cycle as well as oceanic volume transports and storage changes before **(a)** and after **(b)** the optimization. Shading represents the uncertainties of the a priori estimates (shown in both plots). The black line shows the sum of all budget terms, and by going to zero, it demonstrates the closure of the monthly budgets after the optimization.

a priori uncertainties are larger than estimated. A possible explanation is that systematic biases are not taken into account. The state-of-the-art reanalyses exhibit systematic errors in their runoff seasonalities, as the seasonal runoff peaks in summer are too low in comparison to observations, while winter and spring values are too high. Due to the lack of reliable seasonal observations of the oceanic volume fluxes, it is hard to define systematic biases in the ocean reanalyses. However, all four ocean reanalyses feature quite low September volume fluxes, which are not found in their forcing components (see Fig. 9). Uotila et al. (2019) assess 10 ocean reanalyses, including CGLORS, FOAM, GLORYS, and ORAS5, specifically in the polar regions, and find multiple systematic errors concerning sea ice thickness and extent, temperature profiles, mixed layers, as well as ocean transports. Seasonal cycles of volume transports were not assessed; however, seasonal cycles of sea ice components and heat transports did exhibit systematic errors. Further analyses would be necessary to come up with robust estimates of the bias in seasonal volume transports.

5 Conclusions

We analysed and compared various estimates of runoff into the Arctic Ocean on a seasonal and annual basis and found considerable differences in terms of seasonalities, mean values, and trends. Furthermore, we used a non-steric formulation of the Arctic Ocean's volume budget equation and compare freshwater input into the ocean to lateral volume transports over the Arctic's boundaries. To close the budget and get the best estimates of all the budget terms, we applied a variational adjustment procedure. The main outcomes of this study are the following.

- River discharge observations for the major Arctic catchments show distinct runoff peaks in June due to snowmelt and river ice break-up and low runoff values in winter. Annual trends indicate slight discharge increases over the past decades – the largest increases evident at the Lena and Ob basins with a rise of about 4 ± 2 % per decade. The total pan-Arctic area (excluding Greenland) exhibits an upward trend of 2 ± 1 % per decade. Holmes et al. (2018) show that the six largest Eurasian rivers (Ob, Yenisei,Lena, Kolyma, Pechora, Severnaya Dvina) exhibit an increase of 3.3 ± 1.6 % and the two largest North American rivers (Mackenzie and Yukon) an increase of 2.0 ± 1.8 % per decade.
- We estimate pan-Arctic river discharge from gauge observations using monthly correction factors from $\mathrm{GloFAS}_{\mathrm{E5}_{\mathrm{new}}}$, as the common method of hydrological analogy (e.g. Shiklomanov and Shiklomanov, 2003) tends to underestimate the high-flow summer peaks (see Fig. 4) and obtain a long-term annual flux of 4031 ± 203 km^3 (excluding Greenland). To compare our results to past studies, we adapted the time periods and areal extents accordingly and found reasonable agreement with Haine et al. (2015), who combined runoff from ERA-Interim with river discharge observations and obtained a total discharge about 5 % higher than our estimate. An even better agreement was found with the estimates made by Shiklomanov et al. (2021a), as the total pan-Arctic discharges (including Greenland) agree within 2 %.
- Runoff from ERA5 is underestimated compared to observed discharge, with the largest discrepancies occurring in the high-flow summer months, and features strong declines of 10 ± 1 % per decade on a pan-Arctic basis and even stronger declines at the major catchments. These strong declines are caused by two inhomogeneities (1992 and 2004) in ERA5's snowmelt time series and contradict the discharge increases found in gauge observations. Those inhomogeneities are caused by a loss of snow through changes in the data assimilation system. While the discontinuity around 2004 was traced back to the introduction of the IMS snow cover information to the snow assimilation system, the 1992 inhomogeneity still needs further investigation.

  ERA5-Land does not feature these inhomogeneities and exhibits more moderate runoff declines of 5–6 ± 2 % for the Eurasian major watersheds and 2 ± 1 % for the

Mackenzie River and on the pan-Arctic scale. These declines in ERA5-L runoff are caused by similar declines in $P-E$ from ERA5, as $P-E$ is used as a forcing in ERA5-L (see Fig. 6 and Table 5). As observations agree on an increase in river discharge, these declines are deemed unrealistic. An improvement may possibly be achieved when taking VIWVD as forcing, as VIWVD, which is computed from analysed fields rather than short-term forecasts, features similar trends to discharge observations.

- Calculating runoff through the divergence of moisture flux is the only reanalysis-based estimate that exhibits a slightly positive trend of 2 ± 1 % per decade and thus features the best agreement with observations in terms of trends and the highest correlation coefficient. However, VIWVD tends to underestimate runoff by roughly 5 %.

- River discharge from $\text{GloFAS}_{\text{E5}}$ reflects the variation and the inhomogeneities from ERA5 and shows an additional negative shift due to two sink terms removing water in LISFLOOD that lead to a discharge underestimation of up to 50 % towards the end of the time series. In contrast, $\text{GloFAS}_{\text{E5}_{\text{new}}}$ entails a considerable improvement and reproduces the observed values best in terms of annual means and NRMSE values.

- Liquid and solid discharges from Greenland account for roughly 10 % of total pan-Arctic discharge and hence should not be neglected in assessments of the Arctic freshwater budget. Liquid discharge features a vast increase of 20 ± 4 % per decade and solid discharge an increase of 8 ± 1 %. Due to glacial melting, land storage changes are also particularly strong over the Greenlandic ice sheet, accounting for roughly half of total pan-Arctic land storage change.

- Comparing the estimates of freshwater input into the Arctic Ocean that we have most confidence in after the preceding analysis (listed in Table 7) to oceanic volume transports through the Arctic gateways computed from ocean reanalysis yields only a small imbalance of less than 3% in terms of annual means.

- The variational adjustment procedure provides a best estimate of every budget term for every calendar month – listed in Table 10. About two-thirds of the freshwater provided for oceanic volume transports come from runoff ($4379\pm25\,\text{km}^3\,\text{yr}^{-1}$), and about one-third is provided by the atmosphere. Land areas show a strong decline ($-246\,\text{km}^3\,\text{yr}^{-1}$) of water storage over the past decades, while oceanic storage features only a slight increase. This leads to a surplus of roughly $220\,\text{km}^3\,\text{yr}^{-1}$ of more water leaving the Arctic area than entering.

In summary, we refined past Arctic water budget estimates (e.g. Serreze et al., 2006; Dickson et al., 2007) and their uncertainties by combining some of the most recent reanalyses data sets and observations and by applying a variational optimization scheme. The variational adjustment worked very well on an annual scale and brought reliable estimates of the volume budget terms, requiring only moderate adjustments of less than 3 % for each individual term. Adjustments are considered reliable if budget closure is achievable within the respective terms' error bounds and if the terms are comparable to estimates from past studies. With an annual value of $4379\pm25\,\text{km}^3$ (calculated over 1993–2018), our adjusted runoff estimate is slightly higher than estimates made by Shiklomanov and Lammers (2013) and Shiklomanov et al. (2021a) for the period 1936–2006. However, considering the different calculation periods and assuming a decadal rise of roughly 2 %, the estimates are considered to be in good agreement. On a seasonal scale, however, stronger adjustments were needed to close the budget, and some of the adapted freshwater and volume fluxes fell out of their a priori uncertainty range, suggesting an underestimation of the specified uncertainties. The latter is very likely caused by the presence of systematic errors being present in the data sets or at least in their seasonal cycles, which is not taken into account in our a priori uncertainty estimates. Especially when calculating pan-Arctic runoff, caution is needed. Our results show that seasonal peaks of river discharge are underestimated in almost all of the assessed reanalyses (ERA5, ERA5-Land, $\text{GloFAS}_{\text{E5}}$, $\text{GloFAS}_{\text{E5L}}$). The biggest errors are caused by inhomogeneities in the data assimilation system (ERA5 and $\text{GloFAS}_{\text{E5}}$) which led to a great underestimation of runoff, especially in the latter half of the time series. However, reanalyses without data assimilation (ERA5-Land and $\text{GloFAS}_{\text{E5L}}$) were also not able to reproduce the seasonal cycle of river discharge accurately. On the other hand, we find distinct improvements in the new $\text{GloFAS}_{\text{E5}_{\text{new}}}$ product: especially when investigating seasonal cycles and long-term means, it features considerable enhancements compared to its precursors. However, for inter-annual variability and trend analysis, we recommend the use of the VIWVD estimate, as it reproduces trends from gauge observations more accurately than the other estimates.

When extrapolating observed river discharge to the whole pan-Arctic area, we found that the common method of hydrological analogy tends to underestimate the discharge peaks. We therefore advise using river discharge observations where available and reliable runoff/discharge estimates from reanalyses (e.g. $\text{GloFAS}_{\text{E5}_{\text{new}}}$ or ERA5-Land) to extrapolate discharge to the ungauged areas.

A further possible reason for inconsistencies between runoff and ocean reanalyses is the usage of climatological river discharge data to specify land freshwater input in ocean reanalyses. Our analyses show that the seasonal cycle of the ORAS5 runoff climatology BT06 fits well to our observation-based estimates (see Fig. 5); however, the lack of inter-annual variability in the freshwater input adversely affects i.a. oceanic volume transports. We note that Zuo et al. (2022) work on implementing a time-varying land freshwater input, derived from discharge data from GloFAS, into ORAS5. This should further reduce the inconsistencies between runoff and oceanic volume fluxes from ocean reanalyses.

To further refine the budget estimates, longer time series of all budget terms would be needed. For example, one could repeat the analysis using the back extension of ERA5 which goes back to 1950. There is also a new bias-corrected ERA5 data set (WFDE5, Cucchi et al., 2020) that could be examined in terms of the Arctic water budget. Furthermore, it would help to include a precipitation observation data set, preferably one that combines available satellite-based and gauge-based data sets.

Concerning estimation of biases in ocean reanalyses, one could in principle draw on information from oceanographic data for comparison. A main difficulty with oceanographic data is the generally limited temporal and spatial coverage. Nevertheless, the unique form of the Arctic Ocean (as water leaves and enters only through a handful of gateways) allows relatively easy measurements of the incoming and outgoing fluxes. As an example, measurements from arrays of moored instruments (like e.g. acoustic Doppler current profilers, MicroCAT – CTD Sensors and Seagliders) were taken to estimate transports through the Arctic gateways using a mass-consistent framework (Tsubouchi et al., 2012, 2018). Our results however showed that the moored instruments did not measure the velocity field accurately enough to resolve the barotropic wave signal arising from temporally varying runoff (T. Takamasa Tsubouchi, personal communication, 2018), leading to errors in the seasonality of the net volume flux. A longer measuring period with an even denser monitoring network could help with this aspect.

Appendix A: Additional figures

Figure A1 shows trends of terrestrial and oceanic water storage changes derived from GRACE Mascons.

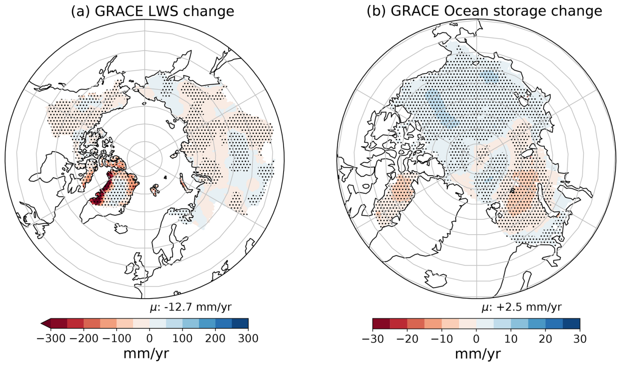

Figure A1. 2002–2019 land water storage trends (left, areal mean of $-12.7\,\mathrm{mm\,yr^{-1}}$) and oceanic storage trends (right, areal mean of $+2.5\,\mathrm{mm\,yr^{-1}}$), calculated as the mean of GRACE JPL, CSR, and GSFC Mascons. Shading denotes significant trends (p values < 0.05).

Appendix B: List of acronyms

ARCGATE	Mooring-derived data of oceanic fluxes through the Arctic gateways
BT06	Runoff climatology used in ORAS5
C3S	Copernicus Climate Change Service
CAA	Canadian Arctic Archipelago
CDS	Climate Data Store
CGLORS	Ocean reanalyses from the Euro-Mediterranean Center on Climate Change
CMEMS	Copernicus Marine Environment Monitoring Service
CEMS	Copernicus Emergency Management Service
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA5	ECMWF's fifth atmospheric reanalysis
ERA5-Land	Offline simulation of ERA5
ERA-Interim	ECMWF interim reanalysis
FOAM	Ocean reanalyses from the UK Met Office
GloFAS	Global Flood Awareness System
GLORYS	Ocean reanalyses from Mercator Ocean
GRACE	Gravity Recovery and Climate Experiment
GREP	Global ocean Reanalysis Ensemble Product
HTESSEL	Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land
JRA55	Japanese 55-year Reanalysis
JRC	Joint Research Centre
LISFLOOD	Model for river-basin-scale water balance and flood simulation
NEMO	Nucleus for European Modelling of the Ocean
NRMSE	Normalized root mean square error
ORAS5	ECMWF's Ocean Reanalysis System 5
ORCA	Tripolar grid used in GREP ocean reanalyses

r	Pearson's correlation coefficient
c	Horizontal seawater velocity
ET	Evapotranspiration
F	Oceanic lateral volume transport
F^{surf}	Surface freshwater fluxes (precipitation, evaporation, runoff)
g	Gravitational constant
P	Precipitation
p_{s}	Surface pressure
q	Specific humidity
R	Runoff or river discharge
ρ_{w}	Density of freshwater
S_{A}	Atmospheric storage
S_{L}	Land storage
S_{O}	Ocean storage
SD	Snow depth
SWVL	Volumetric soil water per layer
$\boldsymbol{v}$	Horizontal wind vector
VIWVD	Vertical integral of divergence of moisture flux

Code and data availability. ERA5 and ERA5-Land data are available in the Copernicus Climate Change Service (C3S) Climate Data Store (Hersbach et al., 2019; Muñoz Sabater, 2019). GloFAS river discharge data are also available in the Copernicus Climate Change Service (C3S) Climate Data Store (Harrigan et al., 2019). GRACE monthly ocean bottom pressure anomalies and land water-equivalent-thickness anomalies are available through the Physical Oceanography Distributed Active Archive Center (Landerer, 2020b, a). River gauge observations are downloadable through the Arctic Great Rivers Observatory (Shiklomanov et al., 2021b) and the Regional Arctic Hydrographic Network data set (Lammers et al., 2001).

Author contributions. SW performed the data analysis, including the production of the figures in the paper, and prepared the manuscript. MM, VS, EZ, HZ, and LH contributed to the interpretation of results and the writing of the manuscript.

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

Acknowledgements. The authors thank Takamasa Tsubouchi for the helpful discussions on ARCGATE data. Susanna Winkelbauer, Vanessa Seitner, and Michael Mayer were supported by Austrian Science Fund project P33177. Vanessa Seitner was additionally supported by Copernicus Marine Environment Service contract 114-R&D-GLO-RAN-CMEMS Lot 8.

Financial support. This research has been supported by the Austrian Science Fund (grant no. P33177) and the Copernicus Marine Environment Service contract (114-R&D-GLO-RAN-CMEMS Lot 8).

Review statement. This paper was edited by Giuliano Di Baldassarre and reviewed by two anonymous referees.

References

Aagaard, K. and Carmack, E. C.: The role of sea ice and other fresh water in the Arctic circulation, J. Geophys. Res.-Oceans, 94, 14485–14498, https://doi.org/10.1029/JC094iC10p14485, 1989.

Bacon, S., Aksenov, Y., Fawcett, S., and Madec, G.: Arctic mass, freshwater and heat fluxes: methods and modelled seasonal variability, Philos. T. Roy. Soc. A, 373, 20140169, https://doi.org/10.1098/rsta.2014.0169, 2015.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., and Betts, A. K.: A Revised Hydrology for the ECMWF Model: Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast System, J. Hydrometeorol., 10, 623–643, https://doi.org/10.1175/2008JHM1068.1, 2009.

Blockley, E. W., Martin, M. J., McLaren, A. J., Ryan, A. G., Waters, J., Lea, D. J., Mirouze, I., Peterson, K. A., Sellar, A., and Storkey, D.: Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new Global FOAM forecasts, Geosci. Model Dev., 7, 2613–2638, https://doi.org/10.5194/gmd-7-2613-2014, 2014.

Bourdalle-Badie, R. and Treguier, A.-M.: A climatology of runoff for the global ocean-ice model ORCA025, available at: https://www.drakkar-ocean.eu/publications/reports/runoff-mercator-06.pdf (last access: 10 January 2021), 2006.

Box, J. E., Colgan, W. T., Wouters, B., Burgess, D. O., O'Neel, S., Thomson, L. I., and Mernild, S. H.: Global sea-level contribution from Arctic land ice: 1971–2017, Environ. Res. Lett., 13, 125012, https://doi.org/10.1088/1748-9326/aaf2ed, 2018.

Burek, P., Van der Knijff, J., and De Roo, A.: LISFLOOD – Distributed Water Balance and Flood Simulation Model – Revised User Manual 2013, https://doi.org/10.2788/24982, 2013.

Cao, B., Gruber, S., Zheng, D., and Li, X.: The ERA5-Land soil temperature bias in permafrost regions, The Cryosphere, 14, 2581–2595, https://doi.org/10.5194/tc-14-2581-2020, 2020.

Church, J., Clark, P., Cazenave, A., Gregory, J., Jevrejeva, S., Levermann, A., Merrifield, M., Milne, G., Nerem, R., Nunn, P., Payne, A., Pfeffer, W., Stammer, D., and Unnikrishnan, A.: Sea Level Change, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, UK and New York, NY, USA, 1535 pp., ISBN 978-1-107-05799-1, 2013.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W. J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A. J., and Wehner, M.: Long-term Climate Change: Projections, Commitments and Irreversibility, in: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, UK and New York, NY, USA, ISBN 978-1-107-05799-1, 2013.

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H., and Buontempo, C.: WFDE5: bias-adjusted ERA5 reanalysis data for impact studies, Earth Syst. Sci. Data, 12, 2097–2120, https://doi.org/10.5194/essd-12-2097-2020, 2020.

Curry, B., Lee, C. M., and Petrie, B.: Volume, Freshwater, and Heat Fluxes through Davis Strait, 2004–05, J. Phys. Oceanogr., 41, 429–436, https://doi.org/10.1175/2010JPO4536.1, 2011.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteorol. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Desportes, C., Garric, G., Régnier, C., Drévillon, M., Parent, L., Drillet, Y., Masina, S., Storto, A., Mirouze, I., Cipollone, A., Zuo, H., Balmaseda, M., Peterson, D., Wood, R., Jackson, L., Mulet, S., and Greiner, E.: CMEMS-GLO-QUID-001-026, EU Copernicus Marine Service Information, available at: https://oceanrep.geomar.de/46419/1/CMEMS-GLO-QUID-001-024.pdf (last access: 2 August 2021), 2017.

Dickson, R., Rudels, B., Dye, S., Karcher, M., Meincke, J., and Yashayaev, I.: Current estimates of freshwater flux through Arctic and subarctic seas, Prog. Oceanogr., 73, 210–230, https://doi.org/10.1016/j.pocean.2006.12.003, 2007.

Fukumori, I., Wang, O., Llovel, W., Fenty, I., and Forget, G.: A near-uniform fluctuation of ocean bottom pressure and sea level across the deep ocean basins of the Arctic Ocean and the Nordic Seas, Prog. Oceanogr., 134, 152–172, https://doi.org/10.1016/j.pocean.2015.01.013, 2015.

Garric, G., Parent, L., Greiner, E., Drévillon, M., Hamon, M., Lellouche, J.-M., Régnier, C., Desportes, C., Le Galloudec, O., Bricaud, C., Drillet, Y., Hernandez, F., and Le Traon, P.-Y.: Performance and quality assessment of the global ocean eddy-permitting physical reanalysis GLORYS2V4, in: EGU General Assembly Conference Abstracts, EGU General Assembly Conference Abstracts, 23–28 April 2017, Vienna, Austria, p. 18776, available at: https://presentations.copernicus.org/EGU2017/EGU2017-18776_presentation.pdf (last access: 15 April 2021), 2017.

GRDC: WMO Basins and Sub-Basins/Global Runoff Data Centre, GRDC, 3rd. rev. ext. Edn., Federal Institute of Hydrology (BfG), Koblenz, Germany, available at: https://www.bafg.de/GRDC/EN/02_srvcs/22_gslrs/223_WMO/wmo_regions_node.html (last access: 4 June 2021), 2020.

Haine, T. W., Curry, B., Gerdes, R., Hansen, E., Karcher, M., Lee, C., Rudels, B., Spreen, G., de Steur, L., Stewart, K. D., and Woodgate, R.: Arctic freshwater export: Status, mechanisms, and prospects, Global Planet. Change, 125, 13–35, https://doi.org/10.1016/j.gloplacha.2014.11.013, 2015.

Haine, T. W. N.: Arctic Ocean Freshening Linked to Anthropogenic Climate Change: All Hands on Deck, Geophys. Res. Lett., 47, e2020GL090678, https://doi.org/10.1029/2020GL090678, 2020.

Harrigan, S., Zsoter, E., Barnard, C. F. W., Salamon, P., and Prudhomme, C.: River discharge and related historical data from the Global Flood Awareness System, v2.1, CDS [data set], https://doi.org/10.24381/cds.a4fdd6b9, 2019.

Harrigan, S., Zsoter, E., Alfieri, L., Prudhomme, C., Salamon, P., Wetterhall, F., Barnard, C., Cloke, H., and Pappenberger, F.: GloFAS-ERA5 operational global river discharge reanalysis 1979–present, Earth Syst. Sci. Data, 12, 2043–2060, https://doi.org/10.5194/essd-12-2043-2020, 2020.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 monthly averaged data on single levels from 1979 to present, CDS [data set], https://doi.org/10.24381/cds.f17050d7, 2019.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. Roy. Meteorol. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Holmes, R., Shiklomanov, A., Suslova, A., Tretiakov, M., McClelland, J., Spencer, R., and Tank, S.: River Discharge, Arctic Report Card: Update for 2018, available at: https://arctic.noaa.gov/Report-Card/Report-Card-2018/ArtMID/7878/ArticleID/786/River-Discharge (last access: 1 June 2021), 2018.

Jeon, T., Seo, K.-W., Kim, B.-H., Kim, J.-S., Chen, J., and Wilson, C. R.: Sea level fingerprints and regional sea level change, Earth Planet. Sc. Lett., 567, 116985, https://doi.org/10.1016/j.epsl.2021.116985, 2021.

Knijff, J. M. V. D., Younis, J., and Roo, A. P. J. D.: LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation, Int. J. Geogr. Inform. Sci., 24, 189–212, https://doi.org/10.1080/13658810802549154, 2010.

Kobayashi, C., Endo, H., Ota, Y., Kobayashi, S., Onoda, H., Harada, Y., Onogi, K., and Kamahori, H.: Preliminary Results of the JRA-55C, an Atmospheric Reanalysis Assimilating Conventional Observations Only, Scient. Online Lett. Atmos., 10, 78–82, https://doi.org/10.2151/sola.2014-016, 2014.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., and Takahashi, K.: The JRA-55 Reanalysis: General Specifications and Basic Characteristics, J. Meteorol. Soc. Jpn. Ser. II, 93, 5–48, https://doi.org/10.2151/jmsj.2015-001, 2015.

Lammers, R. B., Shiklomanov, A. I., Vörösmarty, C. J., Fekete, B. M., and Peterson, B. J.: Assessment of contemporary Arctic river runoff based on observational discharge records, J. Geophys. Res.-Atmos., 106, 3321–3334, https://doi.org/10.1029/2000JD900444, 2001.

Landerer, F.: CSR TELLUS GRACE Level-3 Monthly Ocean Bottom Pressure Anomaly RL06 v03, NASA [data set], https://doi.org/10.5067/TEOCN-3AC63, 2020a.

Landerer, F.: CSR TELLUS GRACE Level-3 Monthly Land Water-Equivalent-Thickness Surface Mass Anomaly RL06 v03, NASA [data set], https://doi.org/10.5067/TELND-3AC63, 2020b.

Landerer, F. W., Flechtner, F. M., Save, H., Webb, F. H., Bandikova, T., Bertiger, W. I., Bettadpur, S. V., Byun, S. H., Dahle, C., Dobslaw, H., Fahnestock, E., Harvey, N., Kang, Z., Kruizinga, G. L. H., Loomis, B. D., McCullough, C., Murböck, M., Nagel, P., Paik, M., Pie, N., Poole, S., Strekalov, D., Tamisiea, M. E., Wang, F., Watkins, M. M., Wen, H.-Y., Wiese, D. N., and Yuan, D.-N.: Extending the Global Mass Change Data Record: GRACE Follow-On Instrument and Science Data Performance, Geophys. Res. Lett., 47, e2020GL088306, https://doi.org/10.1029/2020GL088306, 2020.

L'Ecuyer, T. S., Beaudoing, H. K., Rodell, M., Olson, W., Lin, B., Kato, S., Clayson, C. A., Wood, E., Sheffield, J., Adler, R., Huffman, G., Bosilovich, M., Gu, G., Robertson, F., Houser, P. R., Chambers, D., Famiglietti, J. S., Fetzer, E., Liu, W. T., Gao, X., Schlosser, C. A., Clark, E., Lettenmaier, D. P., and Hilburn, K.: The Observed State of the Energy Budget in the Early Twenty-First Century, J. Climate, 28, 8319–8346, https://doi.org/10.1175/JCLI-D-14-00556.1, 2015.

Lin, X., Massonnet, F., Yang, C., Artale, V., de Toma, V., and Rana, A.: Arctic freshwater cycle and the interaction with the North Atlantic, in: eGU General Assembly 2021, online, 19–30 April 2021, EGU21-436, https://doi.org/10.5194/egusphere-egu21-436, 2021.

Loomis, B. D., Luthcke, S., and Sabaka, T. J.: Regularization and error characterization of GRACE mascons, J. Geod., 93, 1381–1398, https://doi.org/10.1007/s00190-019-01252-y, 2019.

Ludwigsen, C. A., Andersen, O. B., and Rose, S. K.: Assessment of 21 years of Arctic Ocean Absolute Sea Level Trends (1995–2015), Ocean Sci. Discuss. [preprint], https://doi.org/10.5194/os-2020-87, in review, 2020.

Madec, G., Team, N. S., and Team, N. S.: NEMO ocean engine, Zenodo, https://doi.org/10.5281/zenodo.1464816, 2019.

Mankoff, K. D., Noël, B., Fettweis, X., Ahlstrøm, A. P., Colgan, W., Kondo, K., Langley, K., Sugiyama, S., van As, D., and Fausto, R. S.: Greenland liquid water discharge from 1958 through 2019, Earth Syst. Sci. Data, 12, 2811–2841, https://doi.org/10.5194/essd-12-2811-2020, 2020a.

Mankoff, K. D., Solgaard, A., Colgan, W., Ahlstrøm, A. P., Khan, S. A., and Fausto, R. S.: Greenland Ice Sheet solid ice discharge from 1986 through March 2020, Earth Syst. Sci. Data, 12, 1367–1383, https://doi.org/10.5194/essd-12-1367-2020, 2020b.

Mayer, M., Alonso Balmaseda, M., and Haimberger, L.: Unprecedented 2015/2016 Indo-Pacific Heat Transfer Speeds Up Tropical Pacific Heat Recharge, Geophys. Res. Lett., 45, 3274–3284, https://doi.org/10.1002/2018GL077106, 2018.

Mayer, M., Tietsche, S., Haimberger, L., Tsubouchi, T., Mayer, J., and Zuo, H.: An Improved Estimate of the Coupled Arctic Energy Budget, J. Climate, 32, 7915–7934, https://doi.org/10.1175/JCLI-D-19-0233.1, 2019.

Moon, T., Ahlstrom, A., Goelzer, H., Lipscomb, W., and Nowicki, S.: Rising Oceans Guaranteed: Arctic Land Ice Loss and Sea Level Rise, Curr. Clim. Change Rep., 4, 211–222, https://doi.org/10.1007/s40641-018-0107-0, 2018.

Morison, J., Kwok, R., Peralta Ferriz, C., Alkire, M., Rigor, I., Andersen, R., and Steele, M.: Changing Arctic Ocean freshwater pathways, Nature, 481, 66–70, https://doi.org/10.1038/nature10705, 2012.

Muñoz Sabater, J.: ERA5-Land monthly averaged data from 1981 to present, CDS [data set], https://doi.org/10.24381/cds.68d2bb30, 2019.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Muntjewerf, L., Petrini, M., Vizcaino, M., Ernani da Silva, C., Sellevold, R., Scherrenberg, M. D. W., Thayer-Calder, K., Bradley, S. L., Lenaerts, J. T. M., Lipscomb, W. H., and Lofverstrom, M.: Greenland Ice Sheet Contribution to 21st Century Sea Level Rise as Simulated by the Coupled CESM2.1-CISM2.1, Geophys. Res. Lett., 47, e2019GL086836, https://doi.org/10.1029/2019GL086836, 2020.

Proshutinsky, A., Pavlov, V., and Bourke, R. H.: Sea level rise in the Arctic Ocean, Geophys. Res. Lett., 28, 2237–2240, https://doi.org/10.1029/2000GL012760, 2001.

Proshutinsky, A., Krishfield, R., Toole, J. M., Timmermans, M.-L., Williams, W., Zimmermann, S., Yamamoto-Kawai, M., Armitage, T. W. K., Dukhovskoy, D., Golubeva, E., Manucharyan, G. E., Platov, G., Watanabe, E., Kikuchi, T., Nishino, S., Itoh, M., Kang, S.-H., Cho, K.-H., Tateyama, K., and Zhao, J.: Analysis of the Beaufort Gyre Freshwater Content in 2003–2018, J. Geophys. Res.-Oceans, 124, 9658–9689, https://doi.org/10.1029/2019JC015281, 2019.

Prowse, T. D. and Flegg, P. O.: The magnitude of river flow to the Arctic Ocean: dependence on contributing area, Hydrol. Process., 14, 3185–3188, https://doi.org/10.1002/1099-1085(200011/12)14:16/17<3185::AID-HYP170>3.0.CO;2-S, 2000.

Rawlins, M. A., Steele, M., Holland, M. M., Adam, J. C., Cherry, J. E., Francis, J. A., Groisman, P. Y., Hinzman, L. D., Huntington, T. G., Kane, D. L., Kimball, J. S., Kwok, R., Lammers, R. B., Lee, C. M., Lettenmaier, D. P., McDonald, K. C., Podest, E., Pundsack, J. W., Rudels, B., Serreze, M. C., Shiklomanov, A., Øystein, S., Troy, T. J., Vörösmarty, C. J., Wensnahan, M., Wood, E. F., Woodgate, R., Yang, D., Zhang, K., and Zhang, T.: Analysis of the Arctic System for Freshwater Cycle Intensification: Observations and Expectations, J. Climate, 23, 5715–5737, https://doi.org/10.1175/2010JCLI3421.1, 2010.

Rodell, M., Beaudoing, H. K., L'Ecuyer, T. S., Olson, W. S., Famiglietti, J. S., Houser, P. R., Adler, R., Bosilovich, M. G., Clayson, C. A., Chambers, D., Clark, E., Fetzer, E. J., Gao, X., Gu, G., Hilburn, K., Huffman, G. J., Lettenmaier, D. P., Liu, W. T., Robertson, F. R., Schlosser, C. A., Sheffield, J., and Wood, E. F.: The Observed State of the Water Cycle in the Early Twenty-First Century, J. Climate, 28, 8289–8318, https://doi.org/10.1175/JCLI-D-14-00555.1, 2015.

Save, H.: CSR GRACE and GRACE-FO RL06 Mascon Solutions v02, GRACE, https://doi.org/10.15781/cgq9-nh24, 2020.

Save, H., Bettadpur, S., and Tapley, B. D.: High-resolution CSR GRACE RL05 mascons, J. Geophys. Res.-Solid, 121, 7547–7569, https://doi.org/10.1002/2016JB013007, 2016.

Schauer, U. and Losch, M.: "Freshwater" in the Ocean is Not a Useful Parameter in Climate Research, J. Phys. Oceanogr., 49, 2309–2321, https://doi.org/10.1175/JPO-D-19-0102.1, 2019.

Serreze, M. C. and Barry, R. G.: The Arctic Climate System, Cambridge Atmospheric and Space Science Series, 2nd Edn., Cambridge University Press, https://doi.org/10.1017/CBO9781139583817, 2014.

Serreze, M. C., Barrett, A. P., Slater, A. G., Woodgate, R. A., Aagaard, K., Lammers, R. B., Steele, M., Moritz, R., Meredith, M., and Lee, C. M.: The large-scale freshwater cycle of the Arctic, J. Geophys. Res.-Oceans, 111, C11010, https://doi.org/10.1029/2005JC003424, 2006.

Shiklomanov, A. and Lammers, R.: 5.11 – Changing Discharge Patterns of High-Latitude Rivers, in: Climate Vulnerability, edited by: Pielke, R. A., Academic Press, Oxford, 161–175, https://doi.org/10.1016/B978-0-12-384703-4.00526-8, 2013.

Shiklomanov, A., Déry, S., Tretiakov, M., Yang, D., Magritsky, D., Georgiadi, A., and Tang, W.: River Freshwater Flux to the Arctic Ocean, Springer, 703–738, https://doi.org/10.1007/978-3-030-50930-9_24, 2021a.

Shiklomanov, A., Holmes, R., McClelland, J., Tank, S., and Spencer, R. G. M.: Arctic Great Rivers Observatory, Discharge Dataset, Version YYYYMMDD, ArcticGRO [data set], https://www.arcticgreatrivers.org/data (last access: 23 August 2021), 2021b.

Shiklomanov, A. I. and Vuglinsky, V. S.: Hydrological networks, available at: https://www.arcticobserving.org/images/pdf/Workshops/4th_stpetersburg/slideshow/shiklomanov_ppt.pdf (last access: 9 June 2021), 2008.

Shiklomanov, A. I., Lammers, R. B., and Vörösmarty, C. J.: Widespread decline in hydrological monitoring threatens Pan-Arctic Research, Eos Trans. Am. Geophys. Union, 83, 13–17, https://doi.org/10.1029/2002EO000007, 2002.

Shiklomanov, I. and Shiklomanov, A.: Climatic Change and the Dynamics of River Runoff into the Arctic Ocean, Water Resour., 30, 593–601, https://doi.org/10.1023/B:WARE.0000007584.73692.ca, 2003.

Stahl, K., Tallaksen, L. M., Hannaford, J., and van Lanen, H. A. J.: Filling the white space on maps of European runoff trends: estimates from a multi-model ensemble, Hydrol. Earth Syst. Sci., 16, 2035–2047, https://doi.org/10.5194/hess-16-2035-2012, 2012.

Storto, A. and Masina, S.: C-GLORSv5: an improved multipurpose global ocean eddy-permitting physical reanalysis, Earth Syst. Sci. Data, 8, 679–696, https://doi.org/10.5194/essd-8-679-2016, 2016.

Storto, A., Masina, S., Simoncelli, S., Iovino, D., Cipollone, A., Drevillon, M., Drillet, Y., Schuckmann, K., Parent, L., Garric, G., Greiner, E., Desportes, C., Zuo, H., Balmaseda, M., and Peterson, K.: The added value of the multi-system spread information for ocean heat content and steric sea level investigations in the CMEMS GREP ensemble reanalysis product, Clim. Dynam., 53, 287–312, https://doi.org/10.1007/s00382-018-4585-5, 2019.

Syed, T. H., Famiglietti, J. S., Zlotnicki, V., and Rodell, M.: Contemporary estimates of Pan-Arctic freshwater discharge from GRACE and reanalysis, Geophys. Res. Lett., 34, L19404, https://doi.org/10.1029/2007GL031254, 2007.

Tapley, B. D., Bettadpur, S., Watkins, M., and Reigber, C.: The gravity recovery and climate experiment: Mission overview and early results, Geophys. Res. Lett., 31, L09607, https://doi.org/10.1029/2004GL019920, 2004.

Tsubouchi, T., Bacon, S., Naveira Garabato, A. C., Aksenov, Y., Laxon, S. W., Fahrbach, E., Beszczynska-Möller, A., Hansen, E., Lee, C. M., and Ingvaldsen, R. B.: The Arctic Ocean in summer: A quasi-synoptic inverse estimate of boundary fluxes and water mass transformation, J. Geophys. Res.-Oceans, 117, C01024, https://doi.org/10.1029/2011JC007174, 2012.

Tsubouchi, T., Bacon, S., Aksenov, Y., Garabato, A. C. N., Beszczynska-Möller, A., Hansen, E., de Steur, L., Curry, B., and Lee, C. M.: The Arctic Ocean Seasonal Cycles of Heat and Freshwater Fluxes: Observation-Based Inverse Estimates, J. Phys. Oceanogr., 48, 2029–2055, https://doi.org/10.1175/JPO-D-17-0239.1, 2018.

Tsubouchi, T., von Appen, W.-J., Schauer, U., Kanzow, T., Lee, C., Curry, B., de Steur, L., Ingvaldsen, R., and Woodgate, R. A.: The Arctic Ocean volume, heat and fresh water transports time series from October 2004 to May 2010, PANGEA, https://doi.org/10.1594/PANGAEA.909966, 2019.

Uotila, P., Goosse, H., Haines, K., Chevallier, M., Barthelemy, A., Bricaud, C., Carton, J., Fuckar, N., Garric, G., Iovino, D., Kauker, F., Korhonen, M., Lien, V., Marnela, M., Massonnet, F., Mignac, D., Peterson, K. A., Sadikini, R., Shi, L., Tietsche, S., Toyoda, T., Xie, J., and Zhang, Z.: An assessment of ten ocean reanalyses in the polar regions, Clim. Dynam., 52, 1613–1650, https://doi.org/10.1007/s00382-018-4242-z, 2019.

Vaughan, D., Comiso, J., Allison, I., Carrasco, J., Kaser, G., Kwok, R., Mote, P., Murray, T., Paul, F., Ren, J., Rignot, E., Solomina, O., Steffen, K., and Zhang, T.: Observations: Cryosphere, in: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, https://doi.org/10.1017/CBO9781107415324.012, 2013.

Volkov, D. L. and Landerer, F. W.: Nonseasonal fluctuations of the Arctic Ocean mass observed by the GRACE satellites, J. Geophys. Res.-Oceans, 118, 6451–6460, https://doi.org/10.1002/2013JC009341, 2013.

Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., and Landerer, F. W.: Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons, J. Geophys. Res.-Solid, 120, 2648–2671, https://doi.org/10.1002/2014JB011547, 2015.

Wiese, D. N., Yuan, D.-N., Boening, C., Landerer, F. W., and Watkins, M. M.: JPL GRACE Mascon Ocean, Ice, and Hydrology Equivalent Water Height Release 06 Coastal Resolution Improvement (CRI) Filtered Version 1.0, NASA, https://doi.org/10.5067/TEMSC-3MJC6, 2018.

Wouters, B., Gardner, A., and Moholdt, G.: Global Glacier Mass Loss During the GRACE Satellite Mission (2002–2016), Front. Earth Sci., 7, 96, https://doi.org/10.3389/feart.2019.00096, 2019.

Yang, D., Ye, B., and Kane, L. D.: Streamflow changes over Siberian Yenisei River Basin, J. Hydrol., 296, 59–80, https://doi.org/10.1016/j.jhydrol.2004.03.017, 2004a.

Yang, D., Ye, B., and Shiklomanov, A.: Discharge Characteristics and Changes over the Ob River Watershed in Siberia, J. Hydrometeorol., 5, 595–610, https://doi.org/10.1175/1525-7541(2004)005<0595:DCACOT>2.0.CO;2, 2004b.

Yang, D., Zhao, Y., Armstrong, R., Robinson, D., and Brodzik, M.-J.: Streamflow response to seasonal snow cover mass changes over large Siberian watersheds, J. Geophys. Res.-Earth, 112, F02S22, https://doi.org/10.1029/2006JF000518, 2007.

Ye, B., Yang, D., and Kane, D. L.: Changes in Lena River streamflow hydrology: Human impacts versus natural variations, Water Resour. Res., 39, 1200, https://doi.org/10.1029/2003WR001991, 2003.

Zsótér, E., Cloke, H. L., Prudhomme, C., Harrigan, S., de Rosnay, P., Munoz-Sabater, J., and Stephens, E.: Trends in the GloFAS-ERA5 river discharge reanalysis, ECMWF, https://doi.org/10.21957/p9jrh0xp, 2020.

Zuo, H., Alonso-Balmaseda, M., Mogensen, K., and Tietsche, S.: OCEAN5: The ECMWF Ocean Reanalysis System and its Real-Time analysis component, ECMWF, https://doi.org/10.21957/la2v0442, 2018.

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., and Mayer, M.: The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: a description of the system and assessment, Ocean Sci., 15, 779–808, https://doi.org/10.5194/os-15-779-2019, 2019.

Zuo, H., de Boisseson, E., Zsoter, E., Harrigan, S., de Rosnay, P., Wetterhall, F., and Prudhomme, C.: Benefits of dynamically modelled river discharge input for ocean and coupled atmosphere-land-ocean systems, J. Adv. Model. Earth Syst., in preperation, 2022.

3.2 Validation of key Arctic energy and water budget components in CMIP6

3.2.1 Overview

The second publication (Winkelbauer et al, 2024b) pursues the second, and major, research goal and assesses historical simulations of key Arctic energy and water budget components for a large ensemble of Coupled Model Intercomparison Project Phase 6 (CMIP6) models. Observation-based estimates from the first publication (Winkelbauer et al, 2022) and from Mayer et al (2019), as well as independent observations, are used to validate the models' simulated seasonal cycles, long-term averages and trends of vertical fluxes at top-of-atmosphere and the surface as well as lateral transports and storage rates in atmosphere and ocean.

3.2.2 Publication details

- **Title:** Validation of key Arctic energy and water budget components in CMIP6
- **Authors:** Susanna Winkelbauer, Michael Mayer, and Leopold Haimberger
- **Publisher:** Climate Dynamics
- **Type:** Research article
- **Status:** Published on 3 February 2024
- **DOI:** `https://doi.org/10.1007/s00382-024-07105-5`
- **Own contribution:** My contributions involved the acquisition and preparation of the CMIP6 data, computation of all energy and water budget components, the statistical analysis and the visualization of the results, as well as the preparation of the manuscript. The interpretation and discussion of results took place in collaboration with both co-authors. Furthermore, the computation of oceanic transports involved the development of new diagnostic tools, which are described in the final publication. My contributions to this publication are estimated to be about 80%.

Climate Dynamics
https://doi.org/10.1007/s00382-024-07105-5

ORIGINAL ARTICLE

Validation of key Arctic energy and water budget components in CMIP6

Susanna Winkelbauer[1,2,4] · **Michael Mayer**[1,2,3,4] · **Leopold Haimberger**[1,4]

Received: 10 August 2023 / Accepted: 4 January 2024

Abstract
We investigate historical simulations of relevant components of the Arctic energy and water budgets for 39 Coupled Model Intercomparison Project Phase 6 (CMIP6) models and validate them against observation-based estimates. We look at simulated seasonal cycles, long-term averages and trends of lateral transports and storage rates in atmosphere and ocean as well as vertical fluxes at top-of-atmosphere and the surface. We find large inter-model spreads and systematic biases in the representation of annual cycles and long-term averages. Surface freshwater fluxes associated with precipitation and evaporation as well as runoff from Arctic lands tend to be overestimated by most CMIP6 models and about two thirds of the analysed models feature an early timing bias of one month in the runoff cycle phase, related to an early snow melt bias and the lack of realistic river routing schemes. Further, large biases are found for oceanic volume transports, partly because data required for accurate oceanic transport computations has not been archived. Biases are also present in the simulated energy budget components. The net vertical energy flux out of the ocean at the Arctic surface as well as poleward oceanic heat transports are systematically underestimated by all models. We find strong anti-correlation between average oceanic heat transports and mean sea ice cover, atmospheric heat transports, and also the long-term ocean warming rate. The latter strongly suggests that accurate depiction of the mean state is a prerequisite for realistic projections of future warming of the Arctic. Our diagnostics also provide useful process-based metrics for model selection to constrain projections.

Keywords Arctic · Energy budget · Water budget · CMIP6 · Validation · Bias

1 Introduction

The Arctic has undergone major changes in recent decades due to climate warming, with important implications not only for the Arctic itself, but also for the global climate system. Various feedback mechanisms like the ice-albedo feedback or the Planck feedback (Goosse et al. 2018) lead to a faster warming of the Polar regions compared to the globe. While earlier studies report a warming about twice the global average (Serreze et al. 2009; Walsh 2014), more recent observational datasets suggest an even stronger warming about 4 times the global average (Rantanen et al. 2022). Rising temperatures provoke the degradation of permafrost (Rowland et al. 2010), thawing of the Greenland ice sheet (Mouginot et al. 2019) and sea ice melt (Stroeve and Notz 2018). The decline in sea ice area and thickness has been particularly prominent in recent decades (Kwok 2018) and is caused by both atmospheric and oceanic processes (Docquier and Koenigk 2021). Northward heat transports in both the atmosphere and ocean counterbalance an average net loss of energy to space in the Arctic. Variability and trends in those transports have major impacts on the state and change of the Arctic system, including sea ice, the atmosphere and the ocean (Docquier and Koenigk 2021).

Arctic warming also has a strong impact on the Arctic water balance, leading to an increase of runoff from land areas and the Greenland ice sheet as well as increases in precipitation. The reasons for enhanced Arctic precipitation changes are still under debate. While earlier studies attribute increases in area-integrated evaporation due to increased

---

✉ Susanna Winkelbauer
susanna.winkelbauer@univie.ac.at

[1] Department of Meteorology and Geophysics, University of Vienna, Vienna, Austria

[2] b.geos, Korneuburg, Austria

[3] European Centre for Medium-Range Weather Forecasts, Bonn, Germany

[4] Austrian Polar Research Institute (APRI), Vienna, Austria

open water areas together with enhanced moisture transports from lower latitudes (Bintanja and Selten 2014), more recent studies argue that the changes are consequences of the Planck feedback and therefore energetically driven (Pithan and Jung 2021; Bonan et al. 2023)

The effects of Arctic warming are not only limited to the Arctic—the melting of glaciers and the Greenland ice sheet contribute to rapid sea-level rise around the globe (e.g., Moon et al. 2018; Box et al. 2022) and the release of larger amounts of freshwater to the Atlantic Ocean (Haine et al. 2015) could have major implications for the oceanic circulation at a global scale. Further, there is contrasting evidence regarding the hypothesis that a reduction in the meridional temperature gradient might affect weather and climate in the mid-latitudes (e.g., Blackport and Screen 2020; Coumou et al. 2018; Francis and Vavrus 2012; Screen and Simmonds 2013)

Thus, the Arctic represents a complex system marked by tight couplings between atmosphere, ocean and sea ice, encompassing processes on various spatial and temporal scales. Analyzing the Arctic energy and water budgets is crucial to understand the physical processes of the system as well as the couplings between its components and to comprehend the pronounced warming trend and the resulting impacts on the Arctic system itself and globally. Further, improved process understanding and accurate validation data is needed to develop and enhance climate models and subsequently improve our knowledge of future Arctic change.

The development of the Coupled Model Intercomparison Project, a global collaborative initiative with its latest generation CMIP6 (Eyring et al. 2016), whose data are used to i.a. underpin the 6th Assessment Report of the Intergovernmental Panel on Climate Change (e.g., Fox-Kemper et al. 2021), helps to assess projected future changes under various greenhouse gas emission scenarios and is essential in understanding and quantifying the strength and the effects of climate change. However, the complex interactions between atmosphere, ocean and sea ice pose a major challenge to Arctic climate simulations and introduce large uncertainties and biases (Cai et al. 2021; Knutti 2008). This raises the need for a thorough evaluation of historical climate model simulations against observations in order to detect model biases, find potential shortcomings and improve our confidence in future projections.

However, due to the harsh environmental conditions and sheer remoteness, measurements in the polar regions are relatively sparse (Khosravi et al. 2022), complicating especially ocean and sea ice diagnostics. Satellite observations help in the quantification of surface properties, however in-situ data to assess subsurface properties, like vertically resolved temperatures in the ocean, are limited.

Over the past years, the usage of ocean reanalyses (ORAs) proved to be useful to study past ocean states, long-term climate trends and investigate ocean variability (Storto et al. 2019b; von Schuckmann et al. 2020; Mayer et al. 2021c, 2022). However, as their reliability depends i.a. on the quality and quantity of observational data assimilated into the models, the reanalyses are affected by data paucity in the Arctic. Nevertheless, Mayer et al. (2021c) show that ORAs realistically represents observed trends and temporal variabilities of ocean heat content (OHC) in the Norwegian Sea. Cheng et al. (2022) find that the uncertainty of Arctic OHC is larger than for the other world basins, however they still find consistent trends for Arctic OHC between observations and a reanalysis product. Mayer et al. (2022) find a good agreement between ORAs and observations of the variability of ocean heat transport (OHT) anomalies into the Arctic Mediterranean, but they find OHT to be biased small by about 14%. In general, OHC is more strongly constrained in ORAs than oceanic transports and hence are deemed to be more reliable. A largely observation-based estimate of OHTs is provided by Tsubouchi et al. (2018), who derive transport estimates from moorings in a mass-consistent way, creating a largely model-independent estimate of Arctic OHTs.

Serreze et al. (2009) provide holistic estimates of annual cycles and long-term means of the coupled Arctic energy and water budget. However, their results contained inconsistencies of the various terms as indicated by large budget residuals, which is likely related to inaccurate data and suboptimal diagnostic methods (such as a biased atmospheric budget framework, see Mayer et al. 2017). Therefore, Mayer et al. (2019) combine transports from Tsubouchi et al. (2018) with state-of-the-art reanalyses and other observational products and provide updated and improved, consistent estimates of the coupled Arctic energy budget for the period 2005–2009. Similarly Winkelbauer et al. (2022) provide observationally constrained estimates of the key components of the Arctic water budget using observational datasets as well as reanalyses for 1993–2019.

In this study, we will use the observationally constrained estimates from Mayer et al. (2019) and Winkelbauer et al. (2022) as well as updated estimates from observations and reanalyses to evaluate a large ensemble of CMIP6 models. We aim to analyse the models' ability to accurately simulate some of the key components of the Arctic energy and water budgets and analyse the simulated long-term averages and seasonal cycles of the various energy and water cycle variables and their connections to understand typical model biases.

The paper is structured as follows. Section 2 introduces the main energy and water budget equations and describes the numerical methods used for calculating them, and Sect. 3 describes the data sets analysed and the study area. The results are presented in Sect. 4 and are divided into water (Sect. 4.1) and energy (Sect. 4.2) budget analyses. Conclusions and discussions follow in Sect. 5.

2 Methods

In this section we formulate the vertically integrated energy and water balance equations for the Arctic and describe the analytical methods used.

2.1 Energy and water budgets

For the Arctic energy cycle, we follow Mayer et al. (2019) and define the equation for the total energy budget of the atmosphere as

$$F_S = F_{TOA} - AET - \nabla \cdot F_A - L_f(T_P)P_{snow} \tag{1}$$

with the net (turbulent plus net radiative) vertical energy flux at the surface F_S, the net radiation at the top of the atmosphere F_{TOA}, the atmospheric total energy tendency AET and the divergence of vertically integrated lateral atmospheric energy transports $\nabla \cdot F_A$, which is equal to atmospheric energy transports over the lateral boundaries (AHT). The last term represents the cooling of the surface due to falling snow and consists of the latent heat of fusion L_f ($-$ 0.3337 $\times 10^6$ J kg^{-1}) and the snowfall rate P_{snow}. Vertical fluxes are defined as positive downwards. The energy budget equation for an ocean-sea ice column reads as follows:

$$\begin{aligned} F_S =& OHCT + \nabla \cdot F_O + MET + IHCT + \nabla \cdot F_I \\ & - L_f(T_p)P_{snow} + L_f \rho_{snow} \frac{\partial d_{snow}}{\partial t} \end{aligned} \tag{2}$$

with the temporal tendency of ocean heat content OHCT, the divergence of vertically integrated ocean heat transport $\nabla \cdot F_O$, the sea ice melt energy tendency MET (i.e. the energy absorbed or released during melt and freeze, respectively, computed as the product of monthly sea ice thickness change and L_f), the sea ice sensible heat content tendency IHCT, the divergence of latent heat transport associated with sea ice transports $\nabla \cdot F_I$ and the snowfall term. The last term describes latent heat changes in conjunction with changes in grid-point-averaged snow thickness (d_{snow}).

For the oceanic water budget equation we follow Winkelbauer et al. (2022) and formulate it in its volumetric form:

$$\Delta S_O = P + ET + R - \nabla \cdot F_{vol} \tag{3}$$

with the change of ocean volume denoted as ΔS_O, the surface water fluxes precipitation P and evapotranspiration ET (counted positive downward), runoff from surrounding land areas R and the divergence of lateral oceanic volume fluxes $\nabla \cdot F_{vol}$.

Furthermore, following Gauss's divergence theorem the divergence terms in equations 2 and 3 can be replaced by transports of energy and volume across the lateral boundaries when considering closed oceanic regions.

2.2 Oceanic transports

Oceanic transports of volume (OVT), heat (OHT) and ice (OIT) through a given strait are defined as follows:

$$OVT = \int_{x_s}^{x_e} \int_0^{z(x)} \vec{v}_o(x,z) \cdot \vec{n}\, dz\, dx \tag{4}$$

$$OIT = \int_{x_s}^{x_e} d(x)\vec{v}_i(x) \cdot \vec{n}\, dx \tag{5}$$

$$OHT = c_p\rho \int_{x_s}^{x_e} \int_0^{z(x)} (\theta(x,z) - \theta_{ref})\vec{v}_o(x,z) \cdot \vec{n}\, dz\, dx \tag{6}$$

where $\vec{v}_o$ is the velocity vector of the oceanic flow and $\vec{n}$ is the vector normal to the strait. Furthermore, x defines the width along the strait, with the straits' starting point x_s and the end point x_e. The straits' depth is given by z, where x and z together form the cross sectional area of the strait. Ice transports are calculated by integrating the cross-sectional ice velocity $\vec{v}_i$ over the grid point average ice depth (d) and integrating over the section. Latent heat transports into the study area through ice exports (IHT) are then estimated by multiplying OIT with the sea ice density (assumed constant at 928 kgm^{-3}) and the latent heat of fusion L_f ($-$ 0.3337 $\times$ 10^6 J kg^{-1}). Computation of heat transports requires potential temperature θ, the specific heat of seawater c_p and the density of seawater ρ. Throughout this study, c_p and ρ are kept constant at 3996 Jkg^{-1} K^{-1} and 1026 kgm^{-3}, respectively, because variations in c_p and ρ tend to compensate each other and together lead to only small changes in the computed heat transports (Fasullo and Trenberth 2008) which are neglected in the context of this study.

As discussed by Schauer and Beszczynska-Möller (2009), unambiguous heat transports would actually demand closed volume transports through the examined straits, which is not the case for the single straits considered here, and only approximately satisfied for the total oceanic transport through all straits. As a result, heat transports have to be calculated relative to a reference temperature θ_{ref}, which should represent the mean temperature of the assessed flow. Strictly speaking this reference temperature should vary spatially and temporally according to the investigated flow (Bacon et al. 2015). While changes in the reference temperature have only minor effects on the net Arctic transports (not shown), they are larger for transports through individual straits and may become significant the stronger θ_{ref} changes. However, to simplify the analysis we follow e.g. Tsubouchi et al. (2012), Tsubouchi et al. (2018), Muilwijk et al. (2018), Shu et al. (2022), Heuzé et al. (2023) and calculate all heat transports relative to a $0°C$ reference. Usage of the same reference temperature for all models and straits also allows

Table 1 List of models included in the analysis, their modelling components and links to relevant references

i	Model name	Atmosphere	Land	Ocean	Ice	References
1	ACCESS-CM2	UM10.6 GA7.1	CABLE2.5	MOM5.1	CICE5.1.2	Bi et al. (2020)
2	ACCESS-ESM1-5	UM7.3 GA1	CABLE2.4	MOM5.1	CICE4.1	Ziehn et al. (2020)
3	BCC-CSM2-MR	BCC-AGCM3-MR	BCC AVIM2	MOM4	SIS2	Wu et al. (2019)
4	BCC-ESM1	BCC-AGCM3-Chem	BCC AVIM2	MOM4	SIS2	Wu et al. (2019)
5	CAMS-CSM1-0	ECHAM5-CAMS	CoLM 1.0	MOM4	SIS1	Chen et al. (2019)
6	CanESM5	CanAM5	CLASS3.6-CTEM	NEMO3.4	LIM2	Swart et al. (2019)
7	CAS-ESM2-0	IAP AGCM5.0	CoLM	LICOM2.0	CICE4.0E	Zhang et al. (2020)
8	CESM2	CAM6	CLM5	POP2	CICE5.1	Danabasoglu et al. (2020)
9	CESM2-WACCM	WACCM6	CLM5	POP2	CICE5.1	Danabasoglu et al. (2020)
10	CMCC-CM2-HR4	CAM4	CLM4.5	NEMO3.6	CICE4.0	N/A
11	CMCC-CM2-SR5	CAM5.3	CLM4.5	NEMO3.6	CICE4.0	N/A
12	CNRM-CM6-1	ARPEGE-Clim6.3	Surfex 8.0c	NEMO3.6	Gelato 6.1	Voldoire et al. (2019)
13	CNRM-CM6-1-HR	ARPEGE-Clim6.3	Surfex 8.0c	NEMO3.6	Gelato 6.1	Voldoire et al. (2019)
14	EC-Earth3P-HR	IFS cy36r4	HTESSEL	NEMO3.6	LIM3	Haarsma et al. (2020)
15	EC-Earth3	IFS cy36r4	HTESSEL	NEMO3.6	LIM3	Döscher et al. (2022)
16	FGOALS-f3-L	FAMIL2.2	CLM4.0	LICOM3.0	CICE4.0	He et al. (2019)
17	FGOALS-g3	GAMIL3	CAS-LSM	LICOM3.0	CICE4.0	Li et al. (2020)
18	FIO-ESM2-0	CAM5	CLM4.0	POP2	CICE4	Bao et al. (2020)
19	GFDL-CM4	AM4.0	LM4.0	MOM6	SIS2	Held et al. (2019)
20	GFDL-ESM4	AM4.1	LM4.1	MOM6	SIS2	Dunne et al. (2020)
21	GISS-E2-1-G	GISS-E2.1	GISS LSM	GISS Ocean v1	GISS SI	Kelley et al. (2020)
22	HadGEM3-GC31-LL	UM-HG3-GA7.1	JULES-HG3-GL7.1	NEMO-HG3-GO6	CICE-HG3-GS18	Williams et al. (2018)
23	HadGEM3-GC31-MM	UM-HG3-GA7.1	JULES-HG3-GL7.1	NEMO-HG3-GO6	CICE-HG3-GS18	Williams et al. (2018)
24	INM-CM5-0	INM-AM5-0	INM-LND1	INM-OM5	INM-ICE1	N/A
25	IPSL-CM6A-LR	LMDZ	ORCHIDEE	NEMO-OPA	NEMO-LIM3	Boucher et al. (2020)
26	IPSL-CM6A-LR-INCA	LMDZ	ORCHIDEE	NEMO-OPA	NEMO-LIM3	Boucher et al. (2020)
27	KACE-1-0-G	UM-HG3-GA7.1	JULES-HG3-GL7.1	MOM4p1	CICE-HG3-GSI8	Lee et al. (2020)
28	KIOST-ESM	AM2	LM3.0	MOM5	SIS	Pak et al. (2021)
29	MIROC-ES2L	CCSR-NIES AGCM	MATSIRO6	COCO4.9	COCO4.9	Hajima et al. (2020)
30	MIROC6	CCSR AGCM	MATSIRO6	COCO4.9	COCO4.9	Tatebe et al. (2019)
31	MPI-ESM1-2-HR	ECHAM6.3	JSBACH3.20	MPIOM1.63	MPIOM1.63	Mauritsen et al. (2019)
32	MPI-ESM1-2-LR	ECHAM6.3	JSBACH3.21	MPIOM1.63	MPIOM1.63	Mauritsen et al. (2019)
33	MRI-ESM2-0	MRI-AGCM3.5	HAL 1.0	MRI.COMv4	MRI.COMv4	Yukimoto et al. (2019)
34	NESM3	ECHAM6.3	JSBACH	NEMO3.4	CICE4.1	Cao et al. (2018)
35	NorCPM1	CAM(OSLO4.1)	CLM4	MICOM1.1	CICE4	Bethke et al. (2021)
36	NorESM2-MM	CAM6	CLM5	BLOM	CICE5.1.2	Seland et al. (2020)
37	SAM0-UNICON	CAM5-UNICON	CLM4.0	POP2	CICE4.0	Park et al. (2019)
38	TaiESM1	TaiAM1	CLM4.0	POP2	CICE4	Wang et al. (2021)
39	UKESM1-0-LL	UM-HG3-GA7.1	JULES-ES− 1.0	NEMO-HG3-GO6	CICE-HG3-GS18	Sellar et al. (2019)

for better inter-comparisons with one another (Muilwijk et al. 2018).

Transports must be calculated on the native grids of the models to maintain the conservation properties of the models. However, ocean models often use curvilinear grids where the North Pole is placed over land areas to avoid singularities over the ocean. The number of poles (tri- vs. dipolar), the exact location of the poles, and the Arakawa partition vary between models, resulting in a large number of different grid types, making it difficult to compare models and with observations. We have developed two methods for calculating accurate ocean transports on different CMIP6 model grids, which are described in Winkelbauer et al. (2023) and are available via the Python package StraitFlux (Winkelbauer 2023).

Net Arctic transports are calculated as the sum of transports through Fram Strait, Davis Strait, the Barents Sea

Table 2 List of all CMIP6 variables used through this study, including their units, number of available models n and the indices of missing models

Variable	Description	Unit	n	Missing models
rsus, rsds, rlus, rlds	Surface up-/downward, short-/longwave raditations	Wm^{-2}	39	–
rsut, rsdt, rlut	Toa outgoing short-/longwave and incident shortwave raditations	Wm^{-2}	39	–
hfss, hfls	Surface sensible and latent heat flux	Wm^{-2}	38	28
pr	Precipitation flux	$\mathrm{kg\,m^{-2}\,s^{-1}}$	39	–
evspsbl	Evaporation (incl. sublimation and transpiration)	$\mathrm{kg\,m^{-2}\,s^{-1}}$	38	28
mrro	Runoff flux	$\mathrm{kg\,m^{-2}\,s^{-1}}$	36	5, 34, 35
thetao	Sea water potential temperature	degC	38	27
uo,vo	Sea water x/y velocity	ms^{-1}	36	20,27,35
thkcello	Ocean cell thickness	m	26	7, 8, 9, 16, 17, 18, 21, 24, 27, 28, 33, 34, 35
sithick / sivol	Actual thickness of sea ice / total volume of sea ice divided by grid-cell area (=sithick*siconc)	m	35	17, 21, 24, 27
siconc	Percentage of grid cell covered by sea ice	%	36	20, 21, 27
siu, siv	Sea ice x/y velocity	ms^{-1}	20	7, 8, 9, 13, 16, 17, 18, 20, 21, 24, 27, 28, 29, 30, 33, 34, 35, 36, 38

All variables where downloaded through ESGF

Opening and Bering Strait (see Fig. 1 below for the location of the cross-sections).

2.3 Metrics

To validate CMIP6 output against observations, scalar quantities are regridded to regular grids in the resolution of the available observation-based data (0.25°×0.25° and 1°×1° grids). However, quantities using vector-based components are computed on the respective native grids of the models to avoid any errors associated with the interpolation of vector quantities. Spatial averages are calculated over the Arctic areas as defined in Fig. 1 and long-term average seasonal cycles are determined over the 1993–2014 period.

We calculate decadal trends by applying a linear regression to the monthly anomaly (i.e., deseasonalized) time series. Significance is determined by the Wald test with a t-distribution, with p-values less than 0.05 considered significant. Inter-model correlations are calculated using Pearson's correlation coefficient r and to assess seasonal model performance we use normalised mean errors (nME). The normalisation for each variable is done using the largest error of all models for the variable in question to facilitate inter-model comparisons. For instance, the nME for model j over N years (whereby annual averages are calculated using only the assessed season, e.g. DJF for winter) is calculated as follows:

$$nME_j = \frac{\sum_{i=1}^{N}(data_{j,i} - reference_{j,i})}{MAX_k^K(nME_k)} \tag{7}$$

To determine sampling errors of long-term averages that can arise, e.g., from different states of natural variability modes in the model runs compared to observations, we use a bootstraping approach of random sampling with replacement. Thus, for every model and variable we calculate 1000 long-term averages of the desired period (e.g. 22 years for the 1993–2014 period) out of randomly drawn annual averages within the most recent decades (1980–2014). The sampling error is then estimated as 2-sigma standard deviation from the distribution of the randomly sampled long-term averages.

Confidence ellipses for two-dimensional datasets (see all scatter-diagrams in Sec. 4 and the supplementary material) are calculated using the Pearson correlation coefficient as described at https://carstenschelp.github.io/2018/09/14/Plot_Confidence_Ellipse_001.html. They are determined for the 2-sigma standard deviation and therefore encompass about 95% of all values in the 2D space.

3 Data and study domain

3.1 CMIP6 models

We use monthly output from 39 models that participated in the Climate Model Intercomparison Project Phase 6 [CMIP6, Eyring et al. (2016)]. Table 1 lists all the models used in this study, including their modelling components, and provides links to key references. We use historical model runs for 37 models and the hist-1950 model run for EC-Earth3P-HR and HadGEM3-GC31-MM. We use one ensemble member per model and choose the first available member per model, r1i1p1f2 for CNRM-CM6-1, CNRM-CM6-1-HR, MIROC-ES2L and UKESM1-0-LL, r1i1p1f3 for HadGEM3-GC31-LL and HadGEM3-GC31-MM, r1i1p2f1 for EC-Earth3P-HR and r1i1p1f1 for the remaining 32 models. The models have different horizontal and vertical resolutions (please refer to the individual model documentation listed in Tab. 1) and differ in their modelling components for atmosphere, land, ocean and sea ice. However, the models are not completely independent and often overlap in one or more modelling components. Therefore, when calculating the multi-model mean (MMM), models should ideally be preselected to avoid overlapping components or weighted with respect to their independence and performance (Brunner et al. 2020). Hence, results might differ when compiling a model ensemble that maximizes independence of its members, but this is not the focus of this study.

All data are obtained from the Earth System Grid Federation (ESGF) website (https://esgf-node.llnl.gov/search/cmip6/). We assess different components of the energy and water budgets. Table 2 lists the variables used in this study, not all variables are available for all models, therefore the number of available models (n) and a list of missing models (numbers correspond to indices in Table 1) are also given in Table 2. The variables listed in Table 2 are used to derive the main budget components represented by Eqs. 1 to 3, such as F_s, F_{TOA}, AET, OHCT, MET, OHT, OVT and OIT. F_{TOA} and F_s are calculated directly using all available radiative and turbulent heat flux components (see Table 2) and oceanic transports are calculated using StraitFlux (Winkelbauer et al. 2023). Heat content tendencies in the ocean (OHCT) and sea ice (MET) are calculated from sea water potential temperature and sea ice thicknesse respectively, using a Theil-Sen trend estimator. Atmospheric energy tendencies (AET) are calculated on temperature and humidity levels using central differences of monthly mean values and the atmospheric heat transport AHT, which is equal to the divergence term $\nabla \cdot F_A$, is estimated indirectly using equation 1.

Sea ice extent was calculated similarly to Shu et al. (2020) as the area of all grid cells with sea ice concentration (*siconc*) greater than 15%. For sea ice thickness we either use the variable *sivol* or multiply *sithick* with *siconc*, depending on the availability of the variable through ESGF.

3.2 Observational data

To quantify the representation of the energy and water budget components in CMIP, we compare the modelled seasonal cycles and long-term averages with observation based estimates.

Winkelbauer et al. (2022) provide observationally constrained estimates of the key components of the Arctic water budget using in-situ and satellite observations as well as reanalyses, and enforcing budget closure with a variational

Table 3 List of datasets used to calculate the energy and water budget variables

Variable	Data	Time period
R	Adapted from Winkelbauer et al. (2022)	1993–2014
P-E	Adapted from Winkelbauer et al. (2022)	1993–2014
OVT	GREP	1993–2014
	ArcGate	10/2004–05/2010
OIT	GREP	1993–2014
	ArcGate	10/2004-05/2010
F_S	Mayer et al. (2021b)	1993–2014
	Mayer et al. (2019)	2005–2009
F_{TOA}	DEEPC	1993–2014
	Mayer et al. (2019)	2005–2009
OHCT	GREP	1993–2014
	Mayer et al. (2019)	2005–2009
MET	GREP	1993–2014
	CS2SMOS	October–March 2011–2014
	Mayer et al. (2019)	2005–2009
AET	ERA5	1993–2014
	Mayer et al. (2019)	2005–2009
AHT	ERA5	1993–2014
	Mayer et al. (2019)	2005–2009
OHT	GREP	1993–2014
	ArcGate	10/2004-05/2010

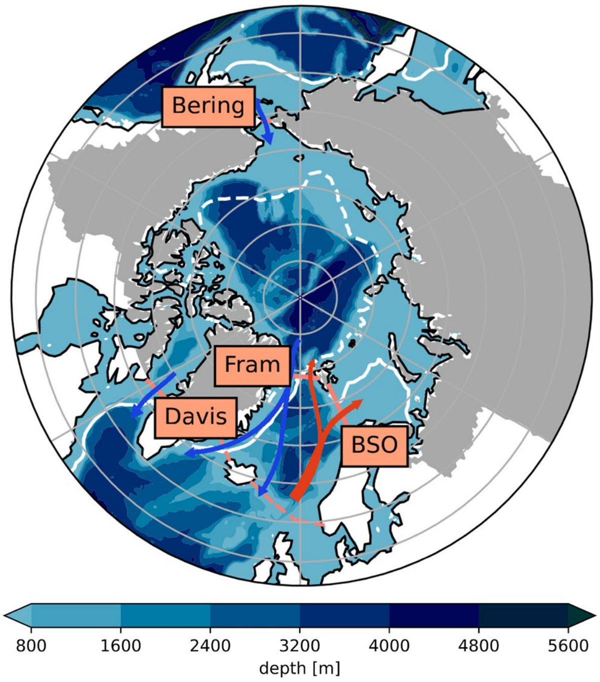

Fig. 1 Map of the main study area, consisting of the oceanic area bounded by the main Arctic gateways (indicated by solid orange lines; corresponds to 11.3×10^6 km^2) and the terrestrial drainage area (grey shading; corresponds to 18.2×10^6 km^2 for mainlands and islands and additional 0.95×10^6 km^2 for Greenland). The orange dashed line indicates the position of the Greenland-Scotland Ridge, which bounds, together with Fram Strait and BSO, the region of the Nordic Seas. Additionally, the main currents flowing in and out of the Arctic (red and blue arrows for warm inflow and cold outflow, respectively) and 1993–2014 mean March (white solid) and September (white dashed) 30% sea ice concentration lines (taken from the GREP reanalyses ensemble) are shown. Shading in the oceanic areas indicates the bathymetry

approach. To avoid use of fluxes based on short-term forecasts from reanalyses, which are known to be biased (Trenberth et al. 2011), the net surface water flux (P-E) was derived from moisture flux divergence, which can be computed from analysed state quantities and thus is more strongly constrained by observations. We adapt results from Winkelbauer et al. (2022) to the 1993–2014 period and use them to validate seasonal cycles and trends of the freshwater input components R and P-E into the Arctic Ocean simulated by the CMIP6 models. As we also want to assess lateral oceanic transports through individual straits and for liquid water and sea ice separately, we additionally calculate oceanic transports directly from the Copernicus Marine Environment Monitoring Service (CMEMS) Global ocean Reanalysis Ensemble Product (GREP, Desportes et al. 2017; Storto et al. 2019a), an ensemble of four global ocean reanalyses: the CMCC Global Ocean Physical Reanalysis System (CGLORS, Storto and Masina 2016), the Forecasting Ocean Assimilation Model (FOAM, MacLachlan et al. 2015), Global Ocean Reanalysis and Simulation Version 4 (GLORYS2V4, Garric et al. 2017) and Ocean Reanalysis System 5 (ORAS5, Zuo et al. 2015). The GREP ensemble members use the NEMO ocean model and are all run at 1/4° horizontal resolution with 75 vertical levels. They all use the same atmospheric forcing (ERA-Interim Dee et al. 2011), however there are differences in the data assimilation methods, used observational products, the reanalysis initial states, NEMO versions, the sea ice models, physical and numerical parameterizations, and air-sea flux formulations. For further details, we refer to the individual data documentations and Storto et al. (2019a). Additionally, we look into an improved version of FOAM (GloRanV14, hereinafter called FOAMv2). Unlike the other reanalyses, FOAMv2 uses a non-linear free surface scheme (NLFS), which introduces some differences when looking into seasonal cycles of volume transports (see Section 4.1.1). Further, we use mooring-derived transports from the so-called ArcGate project (Tsubouchi et al. 2012, 2018), which are available from October 2004 to May 2010.

For the energy budget, we compare the CMIP6 output to results from Mayer et al. (2019), who provide a consistent, closed estimate of the seasonal cycle of the Arctic energy budget for the period 2005–2009 using observations and reanalyses and also a variational optimization approach. They calculate energy budget terms from Eqs. 1 and 2 using satellite observations, various reanalyses and ocean reanalyses as well as oceanic transport derived from moorings. As here we assess longer time periods, i.a. to reduce sampling uncertainties, we additionally calculate the major budget components using observations and reanalyses directly: Net TOA fluxes are compared with the DEEP-C dataset (Liu et al. 2020; Allan et al. 2014,; publicly available at https://doi.org/10.17864/1947.271), a backward extension of the net TOA fluxes from the Clouds and the Earth's Radiant Energy System-Energy Balanced and Filled (CERES-EBAF) satellite product in version 4.1 (Loeb et al. 2018), where fluxes prior to the CERES period have been reconstructed using satellite observations, atmospheric reanalysis and model simulations (Liu et al. 2020). F_s is compared with inferred net surface energy fluxes derived from mass-consistent energy budgets using ERA5 data (Mayer et al. 2021b). The snowfall term in Eqs. 1 and 2 as well as atmospheric transports (via the *vertical integral of divergence of total energy flux*) and the atmospheric tendency term (using central differences on the *vertical integral of total energy*) are additionally calculated using data from the ERA5 reanalyses (Hersbach et al. 2020). Energy tendency components OHCT and MET, as well as latent heat transports associated with sea ice transports (IHT) are estimated using the GREP reanalysis ensemble and oceanic transports of heat are calculated using GREP and mooring-derived transports from ArcGate. Additionally, we use the merged data product from CryoSat2 and

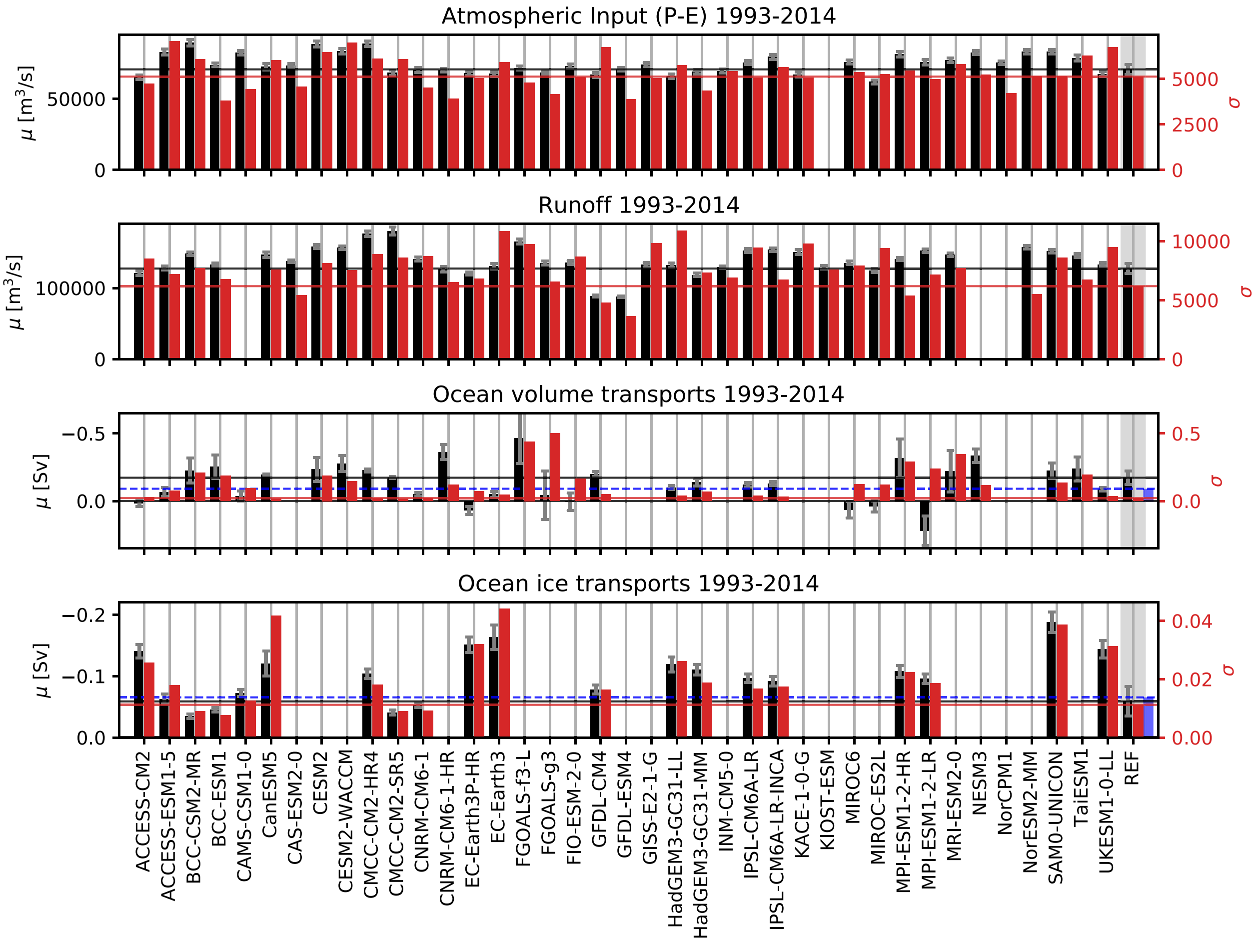

Fig. 2 Averages (black, left axis) and standard deviations of annual averages (red, right axis) for the major Arctic water budget components. Reference values (REF) for P-E and R are taken from Winkelbauer et al. (2022). They are indicated by horizontal lines and shown on the right hand side of the panels. For the oceanic transports REF shows transports from the GREP reanalyses (1993–2014) and additionally also transports from the ArcGate project (2005–2010) are shown (blue bars and dashed lines). Error bars denote sampling errors for CMIP6 and errors calculated from the spread of used observational data for REF

the Soil Moisture and Ocean Salinity satellites (CS2SMOS Ricker et al. 2017), which has not been assimilated in the used ocean reanalyses, to validate sea ice thickness and MET data.

The datasets used to estimate the major energy and water budget components are summed up in Table 3. The calculation of reference uncertainties depends on the used data sources: for P-E and R we use uncertainties provided by Winkelbauer et al. (2022), for oceanic transports as well as OHCT and MET we use the spread of the GREP ensemble and the remaining uncertainties are based on the standard deviations of monthly mean values.

While Mayer et al. (2019) and Winkelbauer et al. (2022) provide closed budgets and therefore consistent estimates of the budget components, the various other, independent data products for some of the budget components (as described above) are not expected to be fully consistent with each other and therefore budget closure for our observational reference estimates is not expected.

3.3 Study area

We consider the Arctic Ocean, which is bounded by hydrographic mooring lines in Fram Strait, Bering Strait, Davis Strait and the Barents Sea Opening (BSO). There are also two small passages, Fury and Hecla Straits, which connect the Arctic Ocean to Hudson Bay through the Canadian Arctic Archipelago (CAA). However, as Tsubouchi et al. (2012) and Bacon et al. (2022) pointed out, volume fluxes through these passages are very small and are not considered in this study. Figure 1 shows the study area, which was chosen to

Table 4 Long-term averages for the major water budget components

Units	P-E	R	R_G	OVT	OIT
	$[10^3\ m^3/s]$ ([frac/dec])			[mSv = $10^3\ m^3/s$] ([frac/dec])	
ACCESS-CM2	65.9 (0.04*)	122.9 (0.06*)	1.3 (0.12)	33 (− 0.80)	− 141 (0.13*)
ACCESS-ESM1-5	85.1 (0.01)	130.8 (0.01*)	7.6 (0.03)	− 66 (0.03)	− 63 (0.19*)
BCC-CSM2-MR	91.8 (0.03*)	149.4 (0.01)	9.4 (0.11*)	− 226 (0.08)	− 35 (0.28*)
BCC-ESM1	75.2 (0.01)	133.8 (0.03*)	9.5 (− 0.03)	− 255 (0.33)	− 45 (0.09)
CAMS-CSM1-0	82.5 (0.00)	–	–	− 37 (− 1.38*)	− 72 (0.0)
CanESM5	75.0 (0.01)	150.1 (0.03*)	8.2 (0.10)	-194 (− 0.03)	− 120 (0.39*)
CAS-ESM2-0	74.7 (0.04*)	139.5 (0.00)	4.6 (0.07)	–	–
CESM2	89.8 (0.04*)	157.7 (0.02*)	4.7 (0.13)	− 234 (0.83*)	–
CESM2-WACCM	84.8 (0.03*)	159.4 (− 0.00)	4.5 (0.07)	− 277 (0.48*)	–
CMCC-CM2-HR4	91.6 (0.04*)	180.6 (0.01)	3.0 (0.13)	− 227 (0.02)	-104 (0.01)
CMCC-CM2-SR5	69.3 (0.05*)	179.3 (0.01*)	4.4 (0.14)	-174 (− 0.01)	− 41 (0.27*)
CNRM-CM6-1	70.9 (0.02)	145.2 (0.01)	9.7 (0.09)	-50 (− 0.02)	− 53 (0.02)
CNRM-CM6-1-HR	71.2 (0.00)	129.4 (0.01*)	11.4 (0.00)	-363 (0.21)	–
EC-Earth3P-HR	68.3 (0.00)	120.0 (0.03*)	1.5 (− 0.01)	70 (0.48)	− 151 (0.11*)
EC-Earth3	68.2 (− 0.00)	135.7 (0.07*)	1.8 (0.12)	− 48 (− 1.17)	− 163 (0.31*)
FGOALS-f3-L	72.8 (0.01)	170.0 (0.00)	2.5 (0.16)	− 465 (0.99*)	–
FGOALS-g3	70.2 (− 0.00)	136.7 (0.04*)	1.5 (0.09)	− 43 (1.31*)	–
FIO-ESM2-0	73.6 (0.02)	138 (0.02*)	1.9 (0.16)	6.6 (1.05)	–
GFDL-CM4	69.4 (0.08*)	89.0 (0.03*)	0.8 (0.03)	− 200 (− 0.15)	− 78 (0.18*)
GFDL-ESM4	72.2 (0.01*)	88.5 (− 0.01*)	0.4 (− 0.06)	–	–
GISS-E2-1-G	75.8 (− 0.02*)	136.0 (− 0.02*)	0.5 (0.01)	–	–
HadGEM3-GC31-LL	66.5 (0.04*)	134.6 (0.08*)	3.8 (0.27)	-99 (− 0.44)	− 119 (0.31*)
HadGEM3-GC31-MM	69.5 (− 0.02*)	119.8 (0.06*)	2.4 (0.18)	-139 (− 0.33)	− 111 (0.16*)
INM-CM5-0	71.1 (0.00*)	131.3 (0.00)	11.9 (0.11)	–	–
IPSL-CM6A-LR	75.9 (0.01*)	156.6 (0.05*)	12.9 (0.05)	− 120 (− 0.13)	− 97 (0.12*)
IPSL-CM6A-LR-INCA	80.9 (0.00)	154.4 (0.03*)	13.2 (− 0.00)	− 128 (− 0.09)	− 91 (0.07)
KACE-1-0-G	67.1 (0.00)	151.5 (0.01)	19.0 (0.03)	–	–
KIOST-ESM	–	132.6 (0.03*)	6.2 (0.03)	–	–
MIROC-ES2L	63.2 (0.07*)	127.5 (0.05*)	8.3 (0.04)	39 (1.76*)	–
MIROC6	76.7 (0.06*)	137.6 (0.03*)	9.9 (0.05)	64 (1.04*)	–
MPI-ESM1-2-HR	82.1 (− 0.02)	141.4 (0.02*)	0.3 (0.02)	− 316 (0.47*)	− 108 (0.22*)
MPI-ESM1-2-LR	76.4 (− 0.00)	153.8 (− 0.01*)	0.4 (0.04)	221 (− 0.53*)	− 96 (0.15*)
MRI-ESM2-0	78.0 (0.01)	149.2 (0.03*)	3.5 (0.30)	− 220 (− 1.26*)	–
NESM3	84.5 (0.02*)	–		− 336 (− 0.03)	–
NorCPM1	76.5 (0.04*)	–		–	–
NorESM2-MM	84.3 (0.07*)	158.8 (0.01*)	9.8 (0.05)	–	–
SAM0-UNICON	84.5 (0.04*)	154.8 (0.05*)	1.4 (− 0.06)	− 224 (− 0.14)	− 188 (0.15*)
TaiESM1	78.6 (− 0.00)	146.9 (0.03*)	8.0 (0.04)	− 238 (0.30*)	–
UKESM1-0-LL	68.4 (0.01)	135.5 (0.04*)	1.9 (0.37)	− 86 (− 0.33)	− 144 (0.27*)
MMM	75.9 (0.02*)	141.1 (0.02*)	5.6 (0.06)	− 145 (0.66)	− 101 (0.18*)
REF	69.2 ± 2.5 (0.02*)	127.0 ± 1.1 (0.02*)	11.9 ± 0.4 (0.13*)	− 151 ± 43 (− 0.00) / − 91^A	− 60 (0.19*) / − 65^A
Winkelbauer et al. (2022)	—"—	—"—	—"—	− 207	

The MMM is calculated using all available models and REF denotes the observation based reference values. Values in brackets show decadal relative trends (* are significant). Reference values for P-E, R and R_G are taken from Winkelbauer et al. (2022), while OVT and OIT are calculated using the GREP reanalyses (1993–2014) and ArcGate (2005–2010, denoted by ^{A}). OVT and OIT are defined as positive northward. Reference uncertainties (±) are taken from Winkelbauer et al. (2022) or calculated from the spread of the GREP ensemble be the same as in Mayer et al. (2019) and Winkelbauer et al. (2022).

To analyse water entering the ocean from the surrounding land areas, we additionally introduce the terrestrial domain, which consists of all land areas draining into the Arctic Ocean, including the CAA as well as islands along the Eurasian coast. We use the catchments as defined by Winkelbauer et al. (2022) and use the same area for all models. The total oceanic and terrestrial areas are 11.3×10^6 km^2 and 18.2×10^6 km^2 respectively, and Greenland provides an additional terrestrial catchment area of 0.95×10^6 km^2.

4 Results

4.1 Water budget

This section looks at the main components of the Arctic water budget. We assess their long-term averages, trends and seasonal cycles.

Figure 2 and Table 4 show long-term averages of surface fresh water flux (P-E) and runoff (R), as well as lateral oceanic fluxes of water volume and ice for the period 1993–2014 and compared with reference values from Winkelbauer et al. (2022). Figure 2 also shows standard deviations and values in brackets in Table 4 show decadal trends.

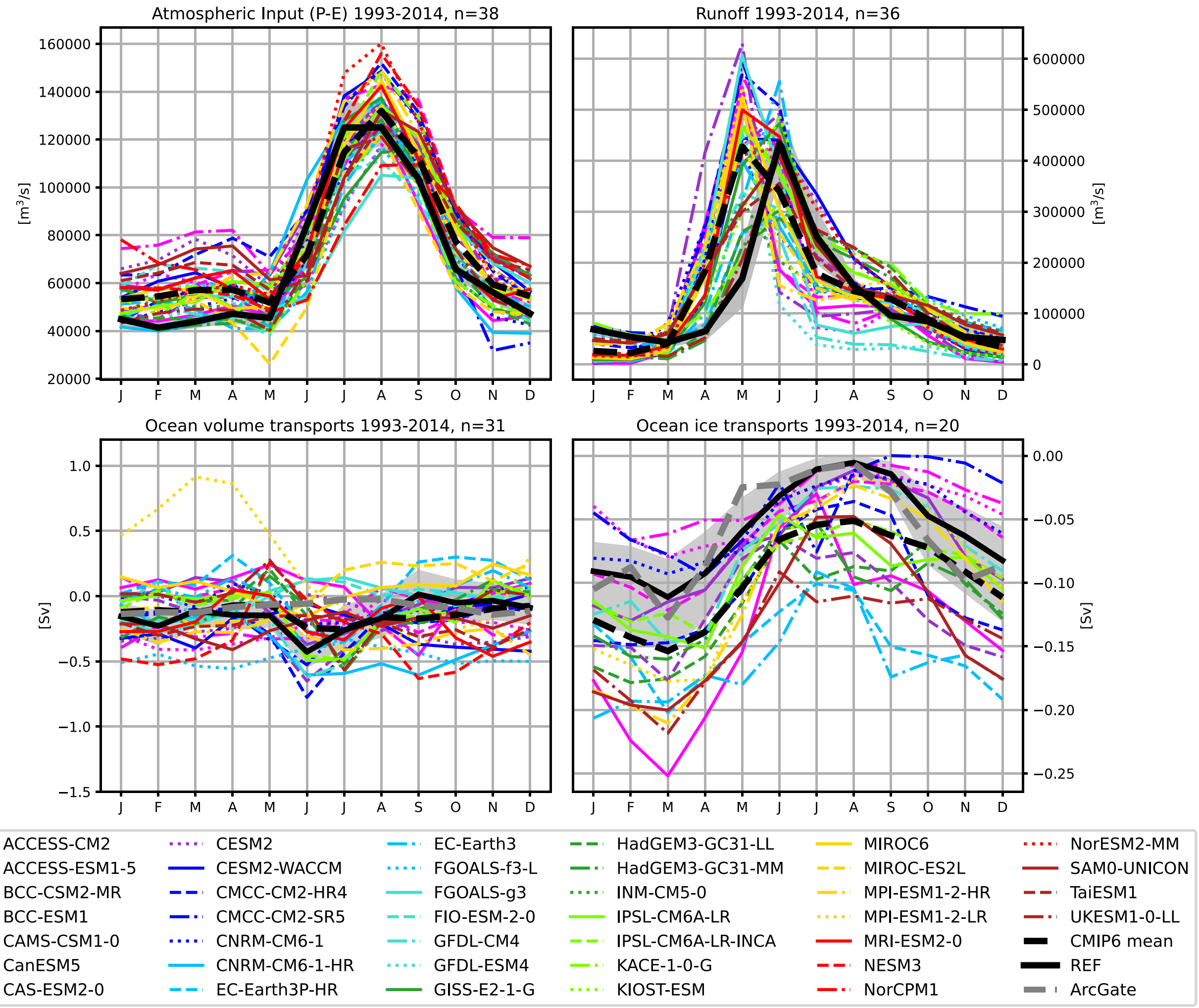

Fig. 3 Mean annual cycles of key terms of the oceanic Arctic water budget: **a** atmospheric freshwater input into the ocean (P-E), **b** runoff from Arctic lands (R), **c** oceanic volume transports and (d) oceanic ice transports across the main Arctic gateways. Ice exports from the ArcGate project are based on the PIOMAS reanalysis and not direct measurements. Shading indicates the uncertainty range of the reference values and is either adopted from Winkelbauer et al. (2022) (top panels) or calculated from the spread (2σ) of the GREP ensemble (bottom panels)

Reference values indicate a long-term mean net freshwater input to the Arctic Ocean from surface fluxes of about 208× 10^3 m^3 s^{-1}. About one-third comes from net precipitation (69.2×10^3 m^3 s^{-1}), two-thirds from runoff from Arctic lands (127.0×10^3 m^3 s^{-1}) and about 5% are melt water and ice discharge from the Greenlandic ice cap R_G (11.9×10^3 m^3 s^{-1}). The MMMs for oceanic P-E (38 models) and R (36 models) are about 10 % higher than our observational references. Net precipitation ranges between 63.2×10^3 m^3 s^{-1} and 91.8× 10^3 m^3 s^{-1}, while runoff ranges between 88.5 and 180.6× 10^3 m^3 s^{-1}, with the lowest values coming from GFDL-CM4 and GFDL-ESM4 and the highest values simulated by CMCC-CM2-SR5 and CMCC-CM2-HR4. Greenlandic runoff from CMIP6 models varies between 0.3 and 19.0× 10^3 m^3 s^{-1} and is underestimated by most models, with the MMM being about 50% smaller than the reference value. In contrast, the CMIP6 MMM of P-E over Greenland is about 10% higher than the reference estimate and there is a clear offset between runoff and P-E for most models. As soil moisture content and the surface snow amount do not change considerably (not shown) the mass balance over Greenland does not seem to be closed for the affected models. Possible reasons may include that the catchment area used, which is assumed to be the same for all models, might omit high runoff regions, that discharge coming directly from the ice sheet and/or solid discharge is underestimated or missing or, to a lesser degree, that the models feature conservation issues over Greenland. Further analyses would be needed to get to the origin of these discrepancies, which were not in the scope of this study.

Most models agree on an increase in freshwater input to the Arctic Ocean for the period 1993–2014: 25 models show a significant increase in R (only 3 show a significant decrease) and while all models agree on increasing precipitations and evaporations, trends in precipitation prevail in most models leading to significant positive trends in oceanic P-E for 19 models (only 2 show a significant decrease). The MMMs show an increase in oceanic P-E of 2% per decade and an increase in R of 2% per decade, which is in fairly good agreement with trends in the reference data. These increases in oceanic P-E and R contribute to an increase in liquid freshwater stored in the Arctic Ocean, which has been observed (Rabe et al. 2011; Proshutinsky et al. 2009; McPhee et al. 2009) and simulated by CMIP6 models (Zanowski et al. 2021; Wang et al. 2022), and further may lead to increased oceanic freshwater exports out of the Arctic system.

Reanalyses indicate a net outflow of liquid volume from the Arctic Ocean of − 151 ± 43 mSv, while estimates derived from observation in the ArcGate project reach -91 mSv. Most CMIP6 models agree on an outflow of liquid volume out of the Arctic and the CMIP6 MMM stays within the reference estimates with − 145 mSv, however the inter-model variability is large. Some models significantly overestimate the net outflows (e.g., FGOALS-f3-L, MPI-ESM1-2-HR), while others indicate net inflows into the Arctic of up to 221 mSv (MPI-ESM1-2-LR). However, it has to be noted that diagnosed volume transports are very sensitive to the exact ocean bathymetry, where slight changes may lead to large deviations of multiple Sverdrups. As net Arctic volume transports are comparatively small values resulting from the sum of large in- and outflowing branches, small errors may lead to significant inconsistencies.

Ice volume transports have only been calculated for 20 models, with all models agreeing on an export of ice to the Atlantic. Ice transports vary between − 188 and − 35 mSv, with a MMM about 30% higher than our observational estimates (− 60 ± 19 mSv for GREP and − 65 mSv for ArcGate).

Using those precisely calculated liquid and solid transports and taking into account all volume budget terms, we are still not able to close the simulated volume budgets for the individual models. Possible reasons for those shortcomings are discussed in Sect. 4.3.

While most models simulate an increase in liquid volume exports through the Fram Strait, an increase in imports through the Barents Sea opening and a decrease in exports through the Davis Strait (not shown, see e.g. Wang et al. 2022), the trends in net volume transports for the whole Arctic vary widely between models. For ice transports, the majority of models agree on a decrease in ice exports over the considered 22-year period, with significant trends between 11 and 39% per decade and a MMM trend of 18%. Long-term averages for volume transports and trends through the individual straits are shown in Table 6.

4.1.1 Long term mean seasonal cycles

Figure 3 shows the seasonal cycles of the main components of the Arctic water budget. Reference values (Winkelbauer et al. 2022) indicate a peak in net atmospheric freshwater input to the Arctic Ocean (P-E, Fig. 3a) from July to September and input minima during the cold season. The CMIP6 ensemble shows a large spread throughout the year. Most models are able to simulate the timing of P-E peaks and minima correctly, but tend to overestimate net P-E for most of the year (see also Table 1).

The annual cycles of terrestrial runoff are summarized in Fig. 3b). Observations show a strong runoff peak in June, mainly due to snowmelt and river ice break-up, and weak runoff during winter. CMIP6 models disagree on the timing of the runoff peak, with about two-thirds placing the runoff maximum in May. However, while observations are derived from gauge measurements at river mouths, the discharge estimates for CMIP6 are determined by calculating area integrals of runoff at each individual grid point over the whole Arctic catchment. As we do not use any kind of river routing this may introduce an error in the runoff phase - especially for large catchments, routing can lead to delays of several months (Gosling and Arnell 2011). Hou et al. (2023) feed daily runoff outputs from 12 CMIP6 models into a state-of-the-art global river routing model to obtain discharge estimates at river gauges and compare the results with streamflow observations. In general, they find that models tend to perform better in non-cold regions than in cold environments. They find an early bias in the timing of the simulated maximum discharge for cold regions in most of the CMIP6 models evaluated. Therefore, in addition to differences in river routing, differences in runoff phase are most likely related to the ability of models to accurately simulate cryospheric hydrological processes such as snow and permafrost. Gosling and Arnell (2011) find that especially for catchments where the peak flow is strongly influenced by seasonal snowmelt (e.g. Ob and Mackenzie), models tend to overestimate the magnitude of the peak flow and show an early bias of the seasonal peak flows. Kouki et al. (2022) analyse the seasonal snow cover for 33 CMIP6 models and find that the models generally overestimate the spring snowmelt rate, leading to early snowmelt. In addition, they found that the snow water equivalent is generally overestimated in winter, driven by precipitation biases and that, while temperature and precipitation can partly explain the biases in snowmelt, there may also be other contributing factors like inaccuracies in model parameterizations related to snow and the surface energy budget. The shift in the runoff phase also has implications for the seasonal cycles of terrestrial water storage. Wu et al. (2021) assess the annual cycles of terrestrial water storage for 25 CMIP6 models and find a shift in the phase of water storage for the four largest Arctic river basins compared to GRACE satellite data, with an earlier end of the recharge period and an earlier start of the discharge period, which is consistent with our results. Nevertheless, some models appear to get the timing of the runoff peak right (Fig. 3b), however whether this is caused by

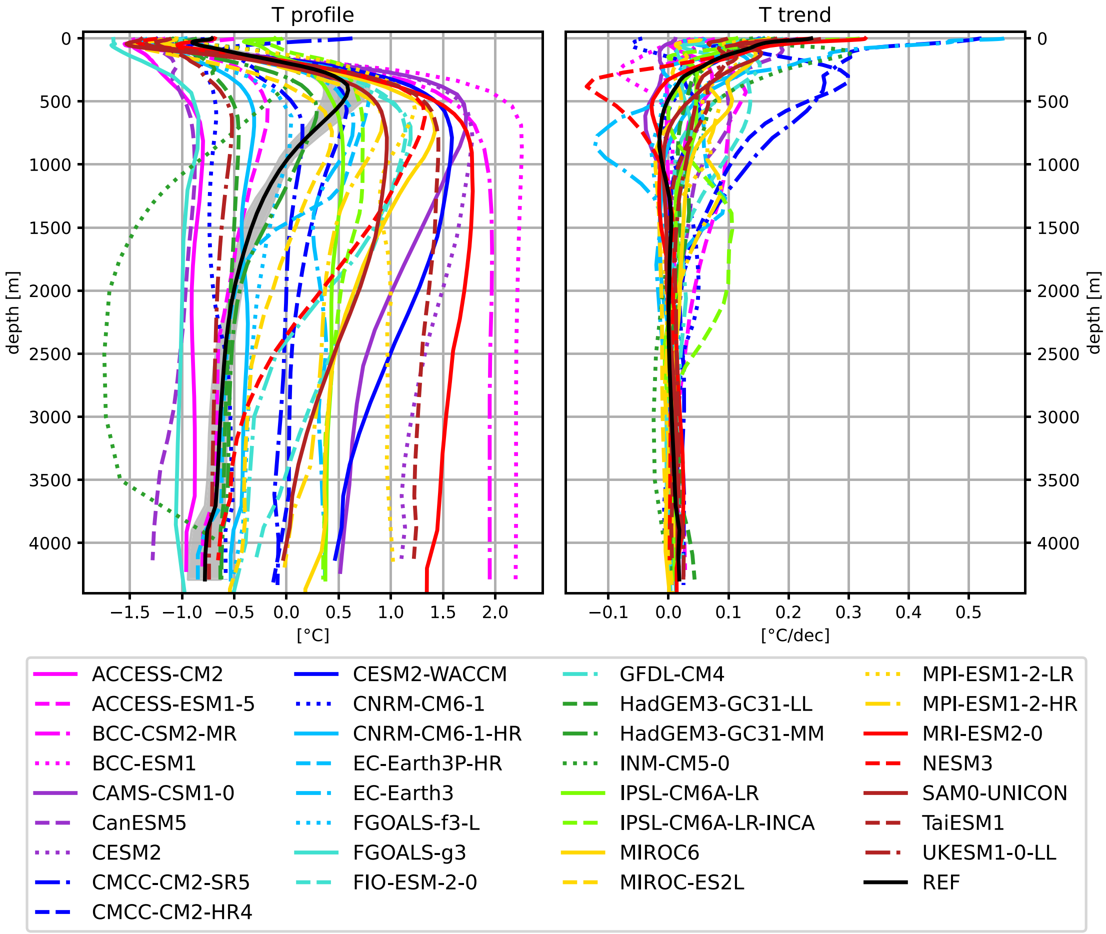

Fig. 4 Vetical profiles of area averaged Arctic temperatures and temperature trends over the 1993–2014 period. Reference values are taken from the GREP reanalyses. Shading indicates the spread (2σ) of the GREP ensemble

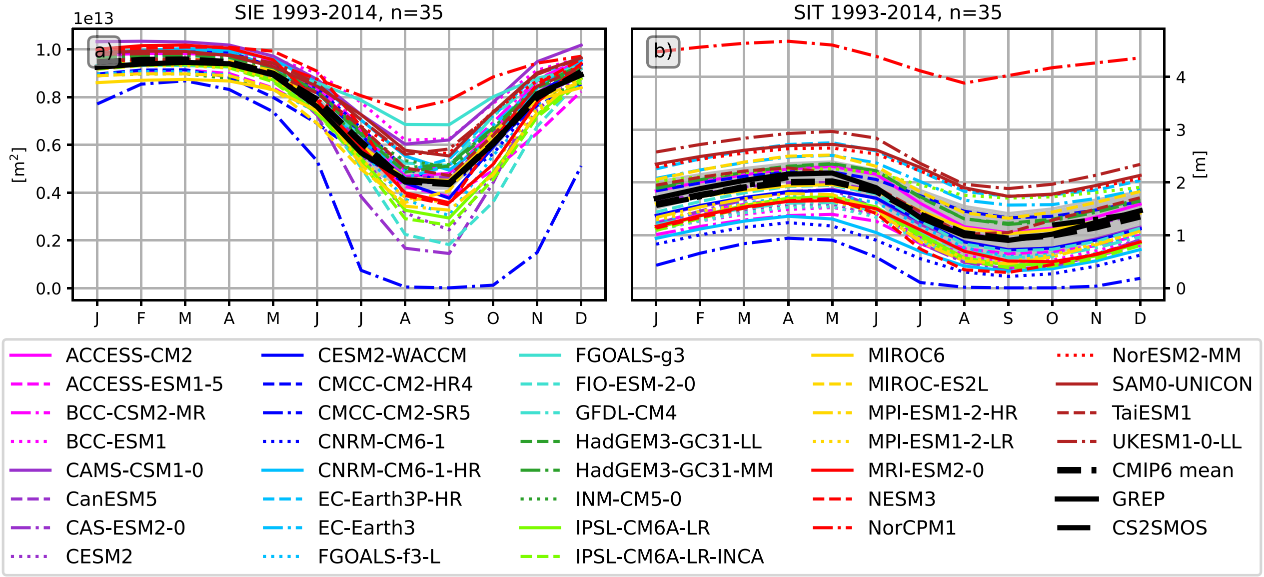

Fig. 5 Mean annual cycles of **a** the mean Arctic sea ice extent (SIE) and **b** the mean sea ice thickness for CMIP6 models and the GREP reanalysis. Additionally SIT estimates from CS2SMOS (10-2002–12-2014) are shown. Shading indicates the spread (2σ) of the GREP ensemble

an actual better representation of the cryospheric processes therein or whether they get the phase right for the wrong reason is not clear and would need further examination.

Figure 3c shows the seasonal cycles of oceanic volume transports through the main Arctic gateways. The GREP ocean reanalysis mean resembles the seasonal cycles of freshwater input to the ocean surface and shows an export maximum of 430 mSv in June, an almost instantaneous response of the ocean to surface freshwater input, as the ocean achieves mass adjustment within about a week through the generation of barotropic waves (Bacon et al. 2015). The observational ArcGate estimate does not show this peak in June, most likely because the mooring arrays are too sparse and the velocity field is not measured accurately enough (both in space and time) to resolve barotropic waves. Models using a non-linear free surface scheme (NLFS), where freshwater from sea ice melt is physically dumped into the ocean resulting in barotropic waves (Madec 2016; Roullet and Madec 2000), were corrected by subtracting the seasonal change in sea ice volume. Volume transports without the sea ice volume correction are shown in the Supplementary material (Fig. S1). The FOAMv2 reanalysis, which in contrast to the GREP ensemble also uses NLFS, as well as CMIP6 models with NLFS show much stronger amplitudes and summer peaks up to one order of magnitude larger than the other models and reanalyses. The effect of ice formation and growth on volume transport can be seen in the cold season, as freshwater is removed from the ocean, leading to a net import of water into the Arctic. However, this behaviour is physically not realistic, as in reality sea ice melt and formation should not affect volume transports in and out of the Arctic. While the correction for net Arctic volume transports appears to be relatively straightforward, the correction for individual Arctic straits and heat transports is not as straightforward and is beyond the scope of this study. Unsurprisingly, the effect of the model mass adjustment of 2-3 Sv in summer is also visible to some extent in the heat transports and will be discussed in Sect. 4.2.1. Meanwhile, linear free surface models show smoother cycles with smaller amplitudes as salt is removed from the ocean during ice melt and added during ice growth to simulate brine rejection. After the correction (Fig. 3c) the seasonal cycles of the NLFS models appear smoother and the CMIP6 MMM stays within uncertainty bounds of our reference estimates during 9 out of 12 months. However, the inter-model spread is still large and there are also models showing some spurious patterns. For example, the MPI-ESM1-2-LR model features volume inflows of up to 900 mSv in spring. As mentioned above, due to the sensitivity of volume transports some errors may be a result of inaccuracies in the calculation process related to the ocean bathymetry. We discuss this further in Sect. 5 and Winkelbauer et al. (2023).

Annual cycles of oceanic ice transports are shown in Fig. 3d. Reanalyses (REF) and ArcGate estimates agree on the annual phase of ice export, with a maximum of ice export in March and a minimum from July to September. Of the 20 CMIP6 models used in this study, which provide all the necessary parameters to calculate ice transports, most agree with the observational estimates in terms of the timing of ice discharge, but differ widely in terms of magnitude. Ice

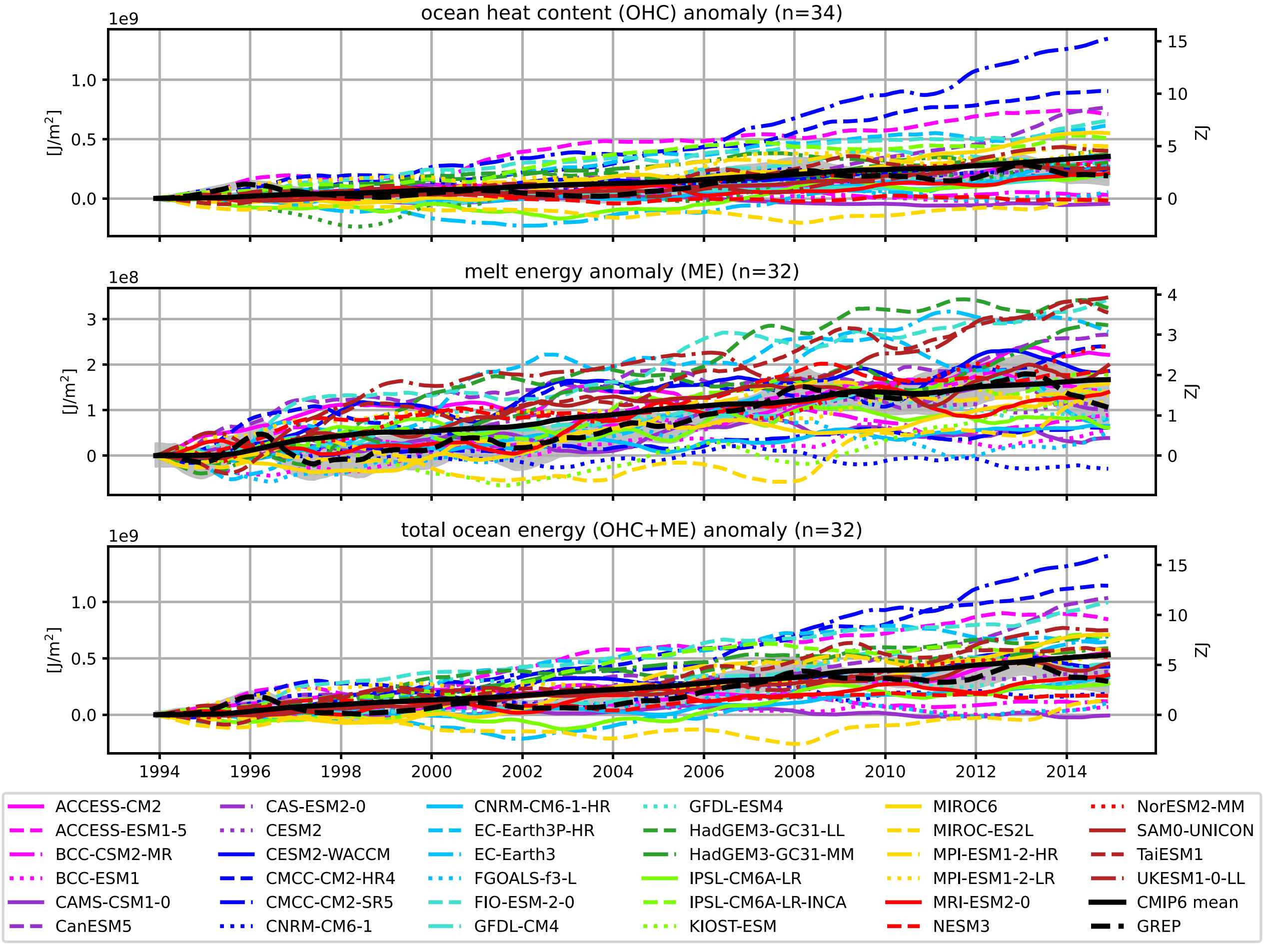

Fig. 6 Full-depth anomalous OHC (top) and ME (middle) accumulation as well as their sum (bottom) in the Arctic Ocean since 1993. Number of available CMIP6 models is given in the titles (n). Left axis shows area-averaged changes in J/m^2 and the right axis shows area integrated changes in ZJ using a conversion factor of 11.3×10^{12} m^2. Shading indicates the spread (2σ) of the GREP ensemble

export maxima in March range from -250 mSv (ACCESS-CM2) to -70 mSv (BCC-CSM2-MR), with most models overestimating total sea ice export throughout the year.

4.2 Energy budget

In this section we will assess the key components of the coupled energy budget of the Arctic. We will start by looking at the mean state of the Arctic system and the accumulation of energy (enthalpy) in the Arctic ocean-sea ice system (heat content of the Artcic Ocean and the enthalpy due to sea ice melt).

Figure 4a shows depth profiles of average Arctic ocean temperatures integrated over our whole study area. The observed halocline lies in the uppermost 250 m, with the warmer and saltier Atlantic Water layer lying underneath. Vertical profiles from the GREP reanalyses (REF) are quite consistent with observed profiles (see e.g. Khosravi et al. 2022) and show an Atlantic water core temperature of about 0.6− 0.8 °C and a core depth of about 450 m. Temperature profiles from the CMIP6 models feature substantial biases, especially so in depths below 500 m. Consistent with Khosravi et al. (2022) and Heuzé et al. (2023), we find that CMIP6 models simulate the Atlantic layer too deep and too thick. Further, CMIP6 models feature a large inter-model spread of more than 3 °C for layers underneath the halocline.

Figure 5 shows the annual cycles of sea ice extent (SIE) and sea ice thickness (SIT). While the CMIP6 models again feature a large inter-model spread with obvious biases for several models, the MMM actually stays within the uncertainty bounds of our reference estimates for both SIE and SIT.

Figure 6 shows the heat accumulation in the Arctic since 1993. The starting year 1993 was chosen because of the availability of our observational reference values. The top panel shows the increase in heat contained in the Arctic Ocean at full depth (OHC), as defined in Fig. 1. Ocean reanalyses show an increase of 0.2 GJ/m^2 (area integrated values may be calculated using the Arctic Ocean area of 11.3×10^{12} m^2 and are provided on the right axes of Fig. 6). While most CMIP6 models agree on an increase in oceanic heat over the period 1993–2014, the amount of heat accumulation varies widely between models. Most models overestimate the heat accumulation, with CMCC-CM2-SR5 being the most extreme with an accumulation of 1.3 GJ/m^2. The MMM of all 34 models is almost twice as high as the observational estimate, reaching 0.35 GJ/m^2 for the 22-year period. Three models (CAMS-CSM1-0, NESM3, MIROC-ES2L) show a slight decrease in oceanic heat storage over the 20-year period and another three models (BCC-ESM1, BCC-CSM2-MR, FGOALS-g3) show insignificantly small heat accumulations, about an order of magnitude smaller than the reference values.

The middle panel shows the accumulation of energy going into sea ice melt (ME). Ocean reanalyses show a heat accumulation of 0.1 GJ/m^2 over the 22-year period, about 57% less than the accumulated OHC change. All CMIP6 models (except one CNRM-CM6-1) agree on an increase of the ME over the last decades, but they again show a huge inter-model spread and range from a total accumulation of 0.0 to 0.4 GJ/m^2. The MMM of all 32 models (0.2 GJ/m^2) is about 30% higher than indicated by the reanalyses, but remains within the uncertainty of the reanalysis ensemble. The total ocean energy accumulation (OHC+ME) is mostly dominated by ocean heat content and is shown in the bottom panel of 6.

For a deeper understanding of the OHC changes we assess the trends of Arctic Ocean temperatures with depth (Fig. 4b). Reanalyses reveal a strong increase of temperatures of about 0.25 °C per decade at the surface. Trends become weaker with depth and beneath about 500 m temperature changes become very small. CMIP6 models show quite diverse trends. While most models agree on an temperature increase at the surface, the strength of the trend ranges from close to zero up to an increase more than twice as high as shown by reanalyses. For the layers below the halocline models differ in terms of sign of the trend and trend strength. For the deep ocean all models agree on comparably small temperature changes. However, it has to be noted, that temperature trends are calculated over a 22-year period, a time-frame short enough that variabilities in inflowing Atlantic Waters may be of importance. Muilwijk et al. (2018) found that variabilities in northward ocean heat transports may impact temperature changes in the deeper Arctic Ocean, with prominent variability on perennial and decadal time scales as well as indicators of variability on multidecadal scales.

Nevertheless, for the 1993–2014 period the large OHC changes simulated by the CMCC-CM2-SR5 model are a result of strong temperature increases from the surface down to about 2000 m depth, while for instance the strong OHC changes for CMCC-CM2-HR4 are mainly driven by temperature changes in the depth of the Atlantic layer core. The NESM3 model, which simulates a slight decrease of heat accumulation, features plausible temperature trends at the surface, however those are compensated by a strong temperature decline around the Atlantic water core depth. For the other five models simulating either insignificantly small heat accumulations or even slight decreases,

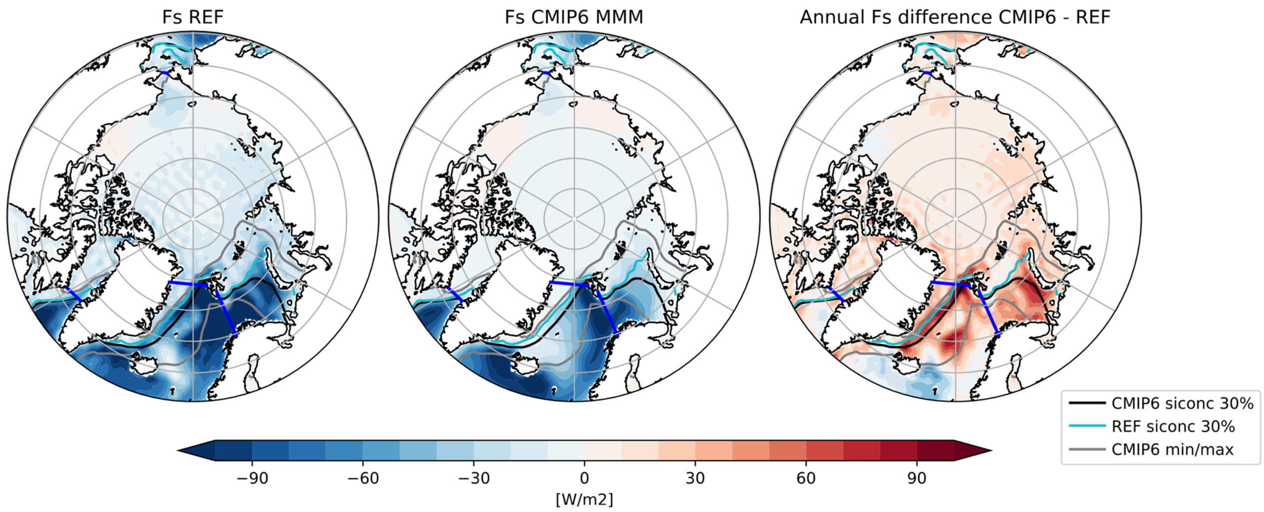

Fig. 7 1993–2014 average F_S for the observation based estimate (REF, left) and the CMIP6 MMM (middle) as well as their difference (right). 30% sea ice concentration lines are indicated in black, cyan and grey, borders of the study area are marked in blue

Table 5 Long-term averages for major energy budget components. The MMM is calculated using all available models and REF denotes the observation based reference values. Averaging periods are 1993–2014 for the MMM and REF and 2005–2009 for the Mayer et al. (2019) estimate. Reference values for OHT are taken from reanalyses and ArcGate (2005–2010, denoted by ^{A}). Oceanic transports are given in TW and may be converted to Wm^{-2} using an integration area of 11.3 × 10^{12} m^2. Reference uncertainties (±) are based on the standard deviations of monthly mean values (F_S, F_{TOA}, AET) or calculated from the spread of the GREP ensemble (OHCT, MET, OHT)

Units	F_S [Wm^{-2}]	F_{TOA}	AET	OHCT	MET	OHT [TW]
ACCESS-CM2	−7.1	−109.3	0.1	0.4	0.3	87.2
ACCESS-ESM1-5	−9.7	−111.7	0.0	1.0	0.2	134.8
BCC-CSM2-MR	−2.8	−114.0	0.0	0.0	0.1	27.6
BCC-ESM1	−3.5	−112.5	−0.0	−0.0	0.1	23.6
CAMS-CSM1-0	−3.5	−110.0	0.0	−0.1	0.0	35.6
CanESM5	−8.0	−109.8	0.0	1.1	0.4	92.2
CAS-ESM2-0	0.1	−103.6	−0.0	-	0.2	–
CESM2	−6.3	−115.6	−0.1	0.4	0.2	68.0
CESM2-WACCM	−5.9	−118.4	0.0	0.3	0.3	60.4
CMCC-CM2-HR4	−14.6	−114.7	0.0	1.3	0.3	189.8
CMCC-CM2-SR5	−9.8	−98.0	0.0	1.9	0.1	144.6
CNRM-CM6-1	−8.1	−115.6	0.0	0.3	−0.0	100.7
CNRM-CM6-1-HR	−7.8	−117.1	−0.0	0.3	0.0	113.5
EC-Earth3P-HR	−9.5	−112.0	0.0	0.8	0.2	115.5
EC-Earth3	−9.8	−112.8	−0.0	0.6	0.3	103.2
FGOALS-f3-L	−5.3	−111.9	−0.0	0.1	0.1	64.4
FGOALS-g3	−5.4	−111.7	0.2	0.3	-	31.8
FIO-ESM2-0	−7.4	−104.7	−0.0	0.9	0.6	106.0
GFDL-CM4	−8.7	−112.9	−0.1	0.4	0.2	121.9
GFDL-ESM4	−7.7	−114.0	0.1	0.3	0.1	–
GISS-E2-1-G	−2.6	−113.3	−0.0	–	–	–
HadGEM3-GC31-LL	−9.9	−108.8	0.0	0.6	0.4	115.3
HadGEM3-GC31-MM	−9.8	−110.9	−0.0	0.4	0.4	122.5
INM-CM5-0	−10.1	−103.6	0.0	0.6	-	-
IPSL-CM6A-LR	−8.7	−114.9	0.0	0.4	0.2	110.6
IPSL-CM6A-LR-INCA	−8.0	−115.5	0.0	0.8	0.3	109.5
KACE-1-0-G	−12.8	−107.9	0.0	–	–	–
KIOST-ESM	-	−115.2	-	-	0.2	-
MIROC-ES2L	−6.7	−100.4	0.1	−0.0	0.2	94.1
MIROC6	−8.0	−109.6	0.0	0.8	0.3	119.3
MPI-ESM1-2-HR	−8.4	−115.7	0.1	0.6	0.2	102.4
MPI-ESM1-2-LR	−10.9	−115.3	0.0	0.6	0.2	124.3
MRI-ESM2-0	−6.1	−115.2	0.1	0.4	0.3	83.9
NESM3	−4.7	−112.6	0.0	−0.1	0.3	40.0
NorCPM1	−8.3	−111.9	−0.1	-	0.2	-
NorESM2-MM	−8.6	−116.3	0.0	0.4	0.4	-
SAM0-UNICON	−7.0	−111.8	−0.1	0.3	0.3	61.8
TaiESM1	−7.9	−113.0	0.1	0.3	0.5	78.8
UKESM1-0-LL	−8.3	−109.7	−0.0	0.5	0.6	93.0
MMM	−7.6	−111.6	0.0	0.5	0.2	92.9
REF	−18.0 ± 3.9	−116.7 ± 0.5	0.1 ± 0.9	0.3 ± 0.4	0.3 ± 0.1	126.7 ± 3.7 / 151.4^A
Mayer et al. (2019)	−16.2	−115.8	−0.1	0.3	0.4	175.2

temperature trends are rather small and partly negative already from the surface down. There are also some other spurious signals to be seen, for instance the EC-Earth3 model simulates net heat accumulations similar to the observed values, however temperature trends at the surface are about twice as high as indicated by our observational reference and in return it features strongly negative temperature trends around 800 m depth. To assess the spurious

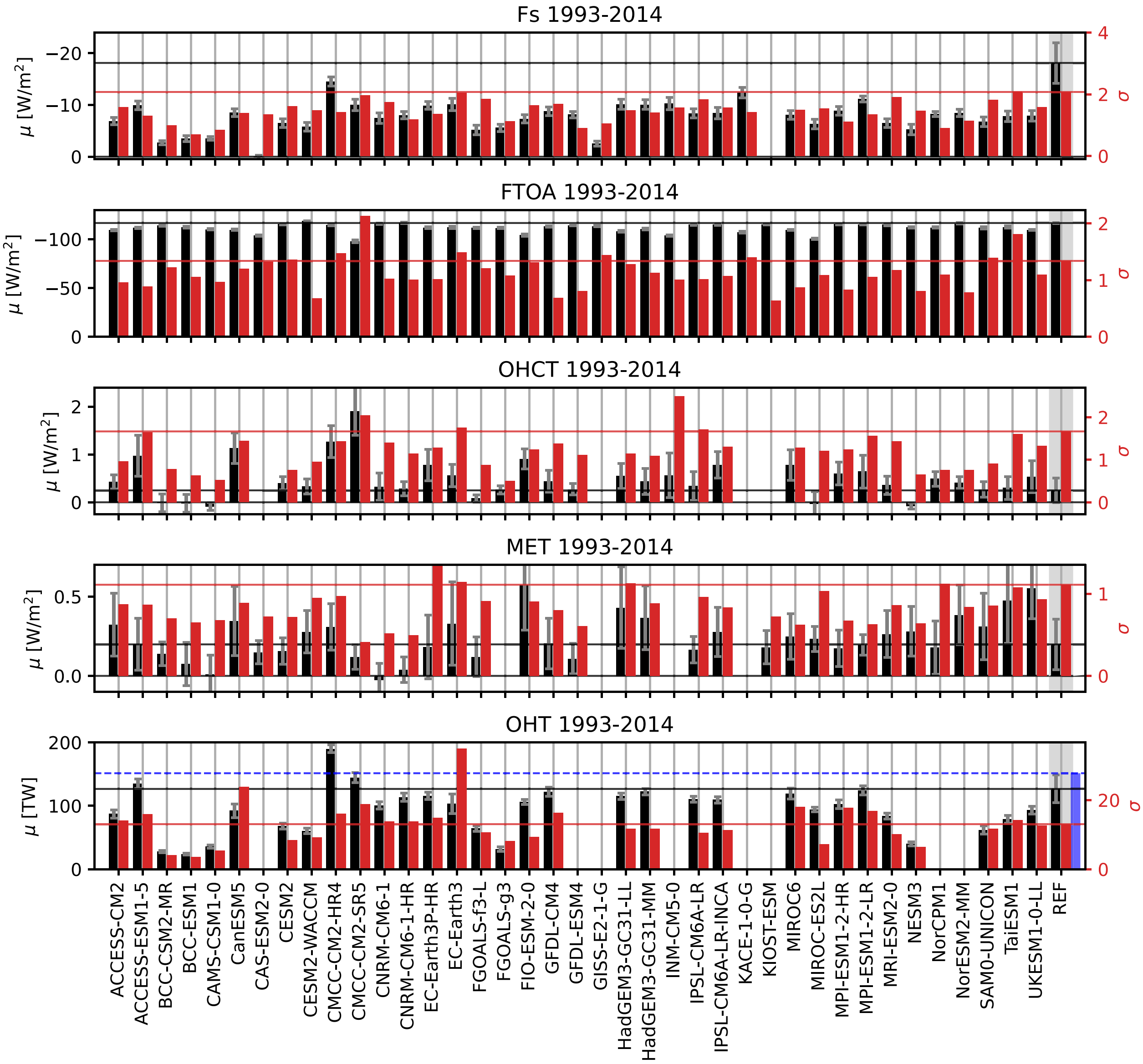

Fig. 8 Averages (black, left axis) and standard deviations of annual averages (red, right axis) for the major Arctic energy budget components. Reference values (REF) are indicated by horizontal lines and shown on the right hand side of the panels. They are taken from Mayer et al. (2021a), DEEPC, ERA5 and the GREP reanalyses (1993–2014). Additionally also transports from the ArcGate project (2005–2010) are shown (blue bars and dashed lines). Error bars denote sampling errors for CMIP6 and errors calculated from the spread of used observational data for REF

trends more closely we looked at longer time periods and found some dubious jumps in the models' OHC time series and partly even changes in the sign of temperature trends when viewing other 22 year periods (not shown). This may indicate that equilibrium was not yet reached by the models and longer spin-up times may be required. We found no clear connection between temperature biases and the strength of temperature trends. We additionally calculated some of the water and energy budget variables for a selection of intra-ensembles containing multiple members of the same models. While intra-model spreads are small for most variables, we found rather large ranges for OHC anomalies, with additionally strong variation from model to model, indicating model-dependence of simulated

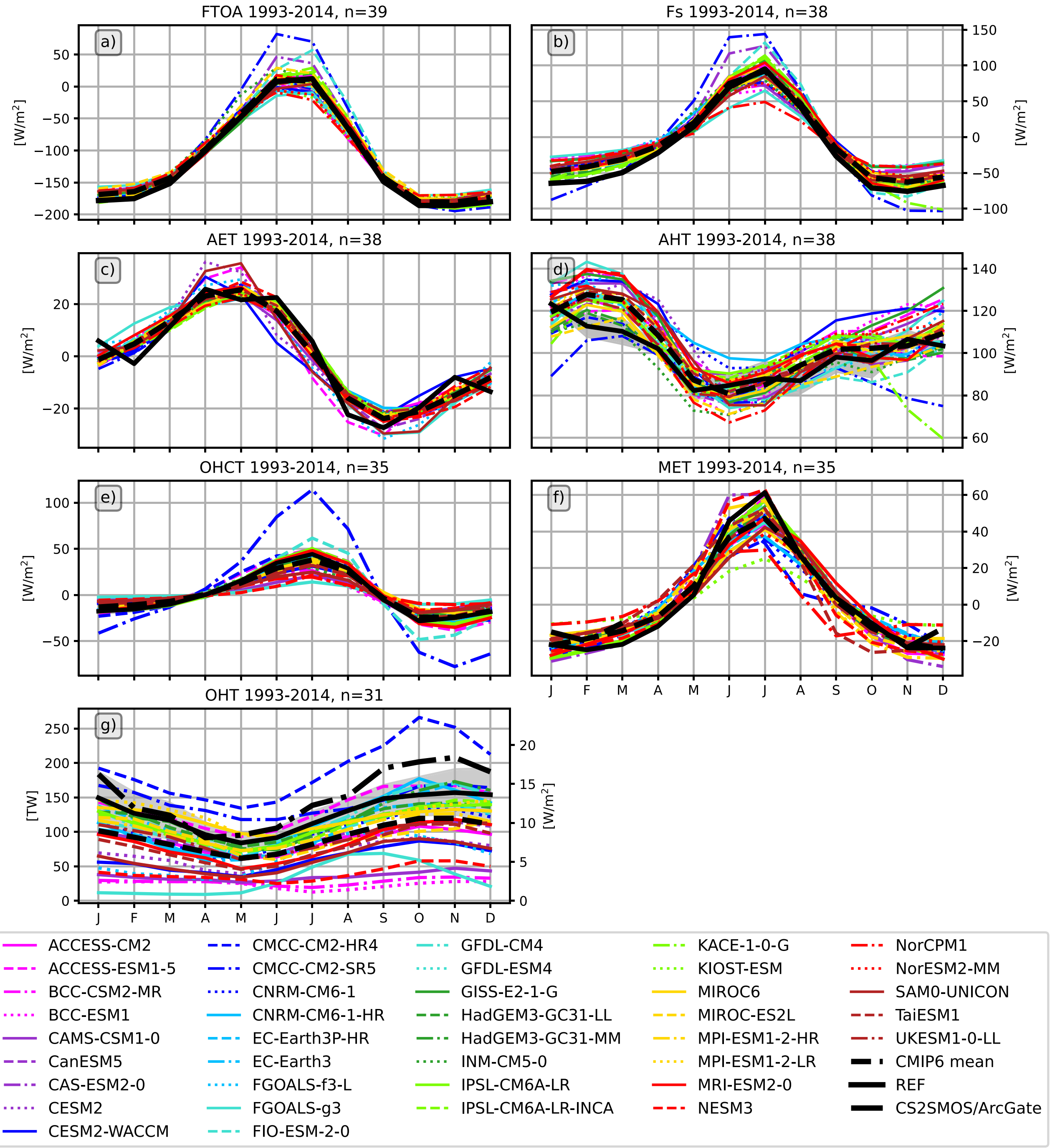

Fig. 9 Mean annual cycles of the key terms of the coupled Arctic energy budget: **a** net radiation at the top of the atmosphere F_{TOA}, **b** net vertical energy flux at the surface F_S, **c** atmospheric energy tendency AET, **d** atmospheric heat transport AHT, **e** full-depth ocean heat content tendency, **f** melt energy tendency (MET), and **g** the oceanic heat transport across the main Arctic gateways. Shading indicates the uncertainty range of the reference values and is either based on the 2σ standard deviations of monthly mean values (F_S, F_{TOA}, AET, AHT) or calculated from the spread of the GREP ensemble (OHCT, MET, OHT)

internal variability. For instance, OHC anomalies at the end of 2014 for an ensemble of 11 CMCC-CM2-SR5 models range between 0.59 and 1.31 GJm^{-2} and for 11 CESM2 models only between 0.19 and 0.29 GJm^{-2}. The larger intra-model errors could again be a sign of internal variability or possible spin-up effects. Nevertheless, as errors estimated via our bootstrapping approach are of a similar or even higher value than those estimated from the model ensembles, we believe our uncertainty estimation to be valid.

The OHC and ME accumulations are converted into tendencies following Mayer et al. (2019) using the Theil-Sen trend estimator. Mean rates for 1993–2014 are given in Table 5. Reanalyses indicate a total ocean warming rate (OHCT+MET) of 0.4 Wm^{-2} for 1993–2014, of which about 40% is due to sea ice melting. The CMIP6 MMM shows a total warming rate of 0.7 Wm^{-2}, of which about one third is due to MET. The atmospheric warming rate (AET) is more than one order of magnitude smaller than OHCT and MET. The CMIP6 models range between − 0.1 and 0.1 Wm^{-2} (the exception being FGOALS-g3), with an MMM of 0.0 Wm^{-2}. Our reference estimate (ERA5) reaches 0.1±0.9 Wm^{-2}, while the estimate from Mayer et al. (2019) suggests − 0.1 Wm^{-2}. As the latter was calculated only over the 2005–2009 period it may be affected by natural variability on various time scales, as AET is assumed to be positive but close to zero on longer time scales (von Schuckmann et al. 2020).

Table 5 also shows long-term averages of vertical and lateral energy fluxes into the Arctic and results suggest strong biases in several energy budget components. Satellite observations show a net radiation at TOA of − 116.7±1.2 Wm^{-2} for the period 1993–2014 and for the area of interest. Most CMIP6 models show smaller fluxes, the whole ensemble ranging from − 118.4 to − 98.0 Wm^{-2}, with a MMM of − 111.6 Wm^{-2}.

The net vertical energy flux at the ocean surface (F_S) from Mayer et al. (2021a) is − 18.0±2.1 Wm^{-2} for the

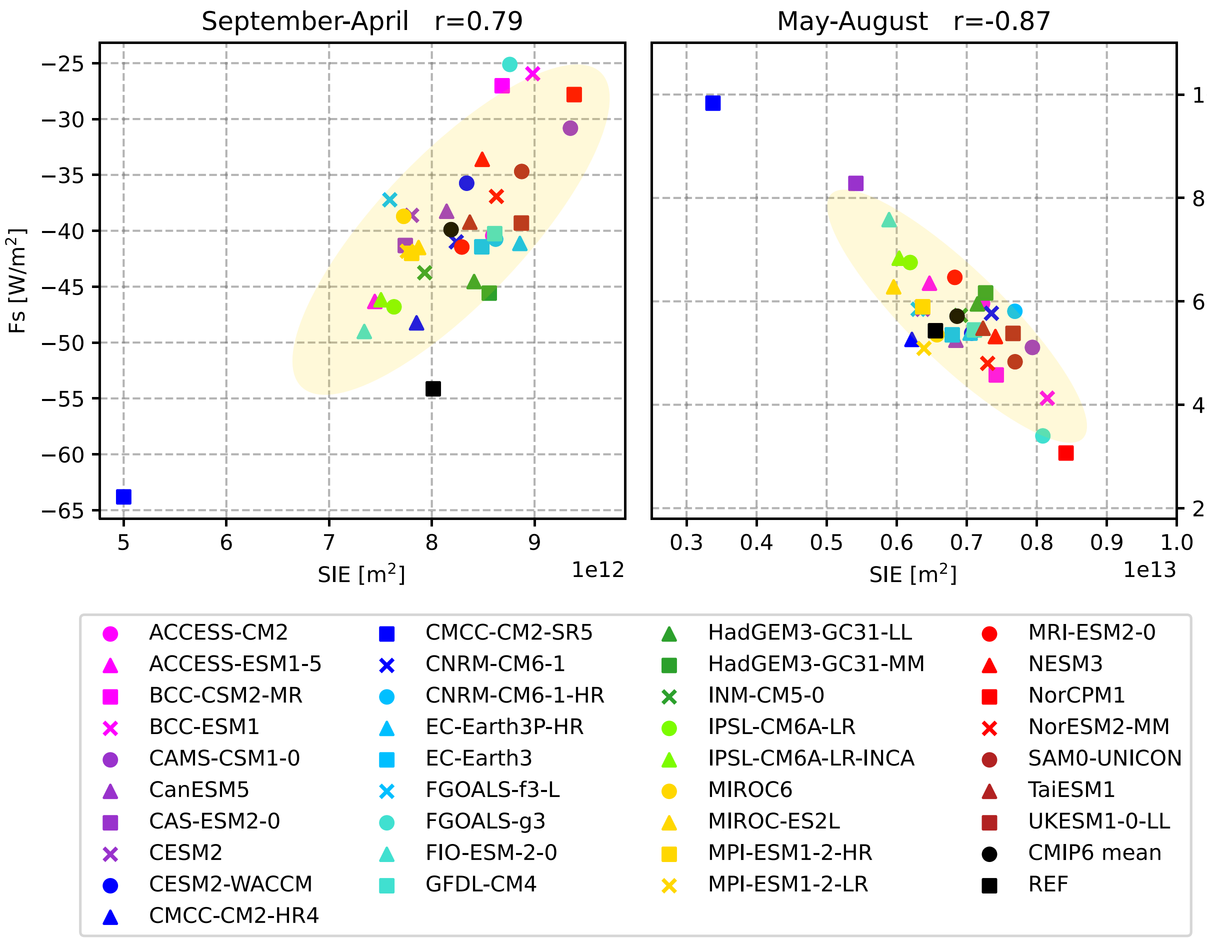

Fig. 10 Scatter plots between the net surface flux F_S and mean sea ice extent SIE averaged over 1993–2014. Left panel: September-April correlation, right panel: May–August correlation. Yellow ellipses show the 2-sigma confidence ellipses for the CMIP6 models

period 1993–2014, while Mayer et al. (2019) estimate a flux of − 16.2 Wm^{-2} for 2005–2009. All CMIP6 models strongly underestimate the outgoing energy fluxes at the surface, ranging from − 14.6 to − 2.6 Wm^{-2}, with one model (CAS-ESM2-0) even showing a slightly positive annual F$_S$ of 0.1 Wm^{-2}. Geographical maps of F$_S$ for the individual models are not shown, but it should be noted that all models are able to simulate reasonable large-scale patterns, with low F$_S$ values over sea ice and high values from the ocean in the Nordic Seas. However, net F$_S$ in the Nordic Seas shows even larger biases with the CMIP6 MMM being about 30% lower than indicated by our reference (not shown). Figure 7 shows the long-term averaged F$_S$ for our observational estimate (left panel), the CMIP6 MMM (middle panel) and their difference (right panel). Furthermore, 30 % sea ice concentration isolines are shown. Differences in F$_S$ over the central, sea ice covered Arctic are small, with slightly higher values around the Kara, Laptev and Chukchi Seas. The largest differences in F$_S$ occur near the sea ice edge between Greenland and Svalbard, in the Barents Sea and in the Norwegian Sea in proximity of the Lofoten Basin, with differences of up to 80 Wm^{-2}. The exact position of the sea ice edge in the Nordic Seas varies considerably between CMIP6 models (indicated by the grey lines in Fig. 7), with the MMM sea ice edge being positioned further south than the reference. Thus, most CMIP6 models simulate too little open water, resulting in smaller net outgoing energy fluxes. For the Labrador Sea and the Bering Sea, the sea ice concentration lines between our reference and CMIP6 are in good agreement and the differences in F$_S$ are comparatively small. Apart from sea ice, the sea surface temperature has major effects on F$_S$. For example, F$_S$ biases in Lofoten Basin are mainly caused by regional cold biases in the simulated sea surface temperatures (not shown).

The loss of energy to space over the Arctic is balanced by northward heat transports in atmosphere and ocean. The CMIP6 models show an enormous range of simulated oceanic heat transports, ranging from 20.30 to 189.82 TW (corresponding to a convergence of 1.80 to 18.80 Wm^{-2}), with a MMM of 93.3 TW (8.26 Wm^{-2}). For the same period, reanalyses indicate a long-term average heat flux of 126.7 TW for 1993–2014. Observational estimates (Tsubouchi et al. 2012) are only available for the period 10/2004-05/2010, but they show an even higher heat transport of 151.4 TW (13.40 Wm^{-2}). For the same period, the reanalysis is 136.2 TW (12.05 Wm^{-2}) and the CMIP6 MMM is 98.1 TW (8.68 Wm^{-2}), clearly underestimating the lateral energy input. Table 5 shows that while most models underestimate the reference value, there are 6 models in particular (BCC-CSM2-MR, BCC-ESM1, CAMS-CSM1-0, FGOALS-g3, FGOALS-f3-L and NESM3) that have exceptionally low transports, with values more than 50% lower than our reference estimates. Some of these use the same ocean model component, BCC-CSM2-MR, BCC-ESM1 and CAMS-CSM1-0 use MOM4, while FGOALS-g3 and FGOALS-f3-L use LICOM3.0. Therefore, it would be a useful step to scale the models in terms of their independence through appropriate weighting algorithms to obtain reliable MMM.

Figure 8 shows long-term averages of the major energy fluxes and tendencies for all models and the reference-based estimates, whereby especially the models' biases in Fs and OHT stand out. Additionally, standard deviations of annual averages and sampling errors are shown. The large sampling errors for OHCT and MET highlight the high temporal variabilities in those variables, which, as discussed above, may indicate possible residual spin-up effects.

4.2.1 Long term mean seasonal cycles

Figure 9 shows the mean annual cycles of the main energy budget terms in Eqs. 1 and 2. Averaging periods depend on the availability of reference data and are indicated in the figure titles. In general, most models are able to simulate the general shape of the annual cycles accurately, but there are also some obvious biases and differences, which are discussed in more detail below.

The net radiation at TOA is shown in Fig. 9 a. It is strongly negative for most of the year and only slightly positive in June and July. This strong seasonal cycle is mainly driven by solar radiation. The spread (max-min) between CMIP6 models is relatively small in winter and the transition seasons, reaching a maximum of 95 Wm^{-2} in summer. A few models reach unrealistically high values during summer, in particular the CMCC-CM2-SR5 model shows a maximum of more than 80 Wm^{-2} (compared to 12 Wm^{-2} from observations), mainly due to strongly underestimated reflected shortwave radiation as a consequence of low sea ice biases (Fig. S2 in Supplementary material). About 20% of the models (8 out of 39) simulate negative F$_{TOA}$ throughout the year. Inter-model spread is higher during summer, nevertheless the CMIP6 MMM is in quite good agreement with the observational estimate (DEEPC) and stays within the observational uncertainty bounds during those months. In winter, the inter-model spread is smaller, but most models underestimate the strong winter minima and the MMM is up to 10 Wm^{-2} lower than indicated by observations. The net radiation at TOA is an important driver of the annual cycle of the surface energy flux F$_S$, so F$_S$ shows a similarly strong annual cycle. F$_S$ remains negative (outgoing) during winter, and with the maximum of incoming shortwave radiation in May, F$_S$ becomes positive and reaches its maximum in summer as sea ice melt progresses. Similar to F$_{TOA}$, some models

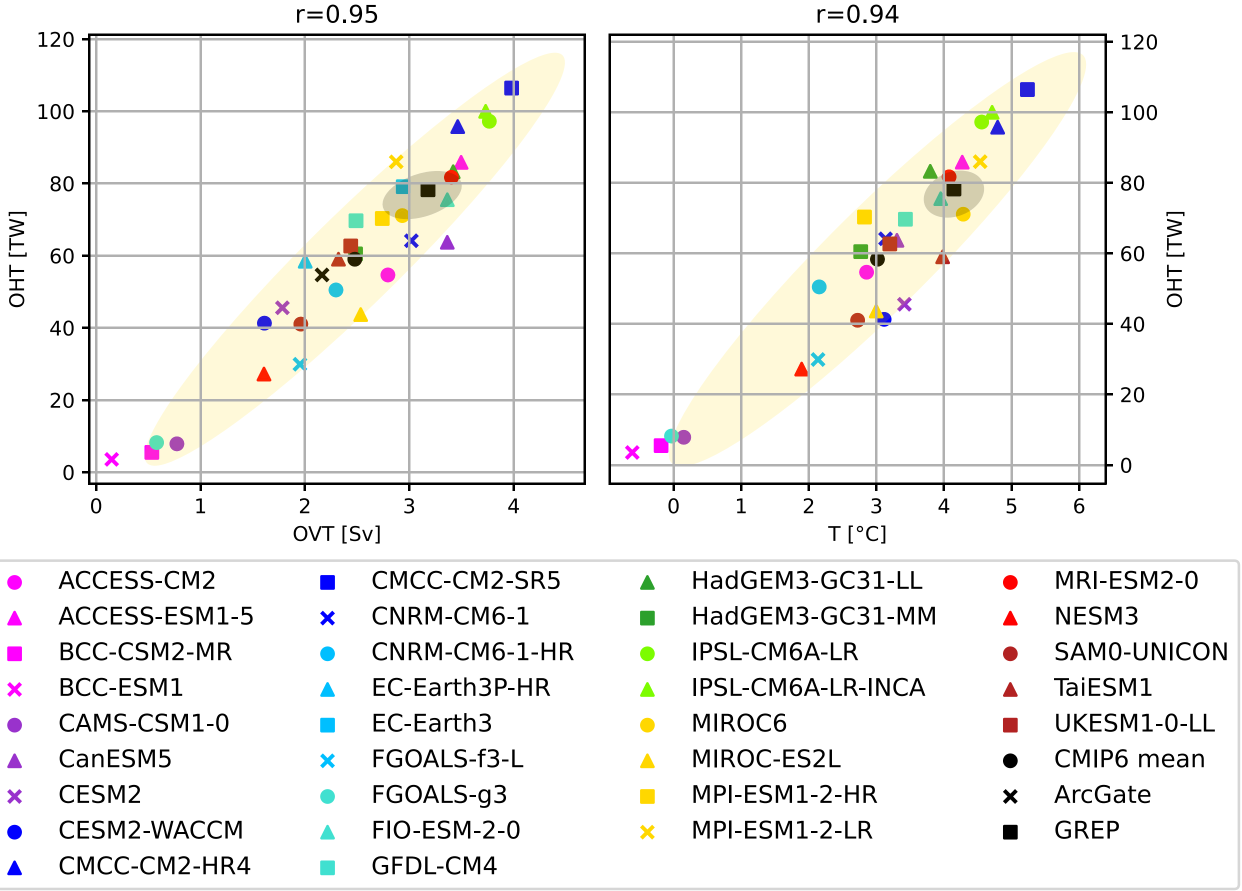

Fig. 11 Barents Sea Opening correlations of long-term annual averaged ocean heat transports (OHT) and ocean volume transports (OVT, left panel) as well as BSO average temperatures (right panel) for various CMIP6 models (1993–2014), the GREP reanalyses mean (1993–2014) and ArcGate observations (2005–2010). Positive values denote transports into the Arctic. Yellow ellipses show the 2-sigma confidence ellipses for the CMIP6 models and grey ellipses for the GREP reanalyses

have unrealistically high summer maxima (CMCC-CM2-SR5, CAS-ESM2-0, FIO-ESM-2-0), caused by underestimated reflected shortwave radiation due to too little sea ice (not shown). The CMCC-CM2-SR5 model strongly underestimates both the total extent of sea ice in the area of interest and the thickness of the sea ice cover (Fig. 5), and in the summer months CMCC-CM2-SR5 even simulates an ice-free Arctic. CAS-ESM2-0 and FIO-ESM-2-0 are also at the lower end of the SIE ensemble, while models with high SIE during the summer months (e.g. NorCPM1, FGOALS-g3) also simulate low F_S during the summer months. In winter, most models simulate lower net upward F_S than our reference estimate (inferred F_S, Mayer et al. 2021a). However, CMCC-CM2-SR5 overestimates the winter minima because the SIE is quite small in winter, leading to unrealistically strong outgoing longwave radiation and latent heat fluxes (not shown). Figure 10 shows scatter plots between long-term average F_S and SIE. The correlations are divided into the season with negative net F_S (September - April) and the season with positive net F_S (May - August). The correlations are high throughout the year. In summer, when incoming solar radiation is high, models with little sea ice simulate higher incoming net radiations, mostly caused by reduced reflected shortwave radiations (not shown). In autumn and winter, when the incoming solar radiation is low to non-existent, models with less sea ice simulate higher outgoing longwave radiations (not shown) and therefore lead to higher negative net radiations.

Figure 9c shows that the annual cycle of the atmospheric energy storage component AET is moderate compared to the other atmospheric components, and that CMIP6 models reproduce the observed cycles (ERA5) quite well. Atmospheric energy transport (AHT) for CMIP6 is estimated as residual using Eq. 1. Inter-model spread is relatively high throughout the year with most CMIP6 models simulating higher transports than indicated by our observational reference (ERA5). Biases are strongest from late autumn to early spring and are connected to biases in surface energy fluxes and therefore biases in the position of the sea ice edge as well as sea surface temperatures. However, in summer, where biases in F_S and F_{TOA} are at their peaks in some models (e.g., CMCC-CM2-SR5, CAS-ESM2-0 and NorCPM1), compensating effects lead to smaller biases in AHT. Meanwhile, biases in AET play a less prominent role and adjust the total AHT biases with smaller reinforcing and compensating effects.

Figure 9e shows the annual cycles of the oceanic storage component OHCT. The models agree on ocean warming in the summer months and ocean cooling in the winter, but the amplitude of the cycles varies considerably. The model scatter is large for most of the year, with 95% of the models within 12–43 Wm^{-2} in summer and -24 and -4 Wm^{-2} in winter. The most obvious exception is CMCC-CM2-SR5, which has summer maxima about three times higher than the MMM and winter minima about three times lower. These large variations are again closely related to the underestimation of sea ice in CMCC-CM2-SR5. An amplification of the OHCT seasonal cycle with declining sea ice is expected, and in fact has already been observed over recent decades (Mayer et al. 2016), but of course to a much lesser degree compared to CMCC-CM2-SR5. The annual cycles of the melt tendency (MET) are shown in figure 9f. The reference values are calculated from the GREP ensemble. The majority of models simulate the phase of the annual MET cycle correctly, but the inter-model spread is large throughout the year. Winter values range from -31 to -10 Wm^{-2} and summer peaks are between 16 and 52 Wm^{-2}. The MMM amplitude is generally lower than the reference estimate, with weaker freezing in late winter and early spring and weaker melting during the summer months.

The annual cycles of the net oceanic heat transports are shown in Fig. 9g. The large inter-model variability is evident throughout the year, but all models agree on an inflow of heat to the Arctic in all calendar months. Most models are able to simulate the timing of the inflow extremes correctly, with a minimum in May and a maximum in late autumn and early winter. Reference values (REF) are derived from the GREP ocean reanalysis ensemble, and the observational

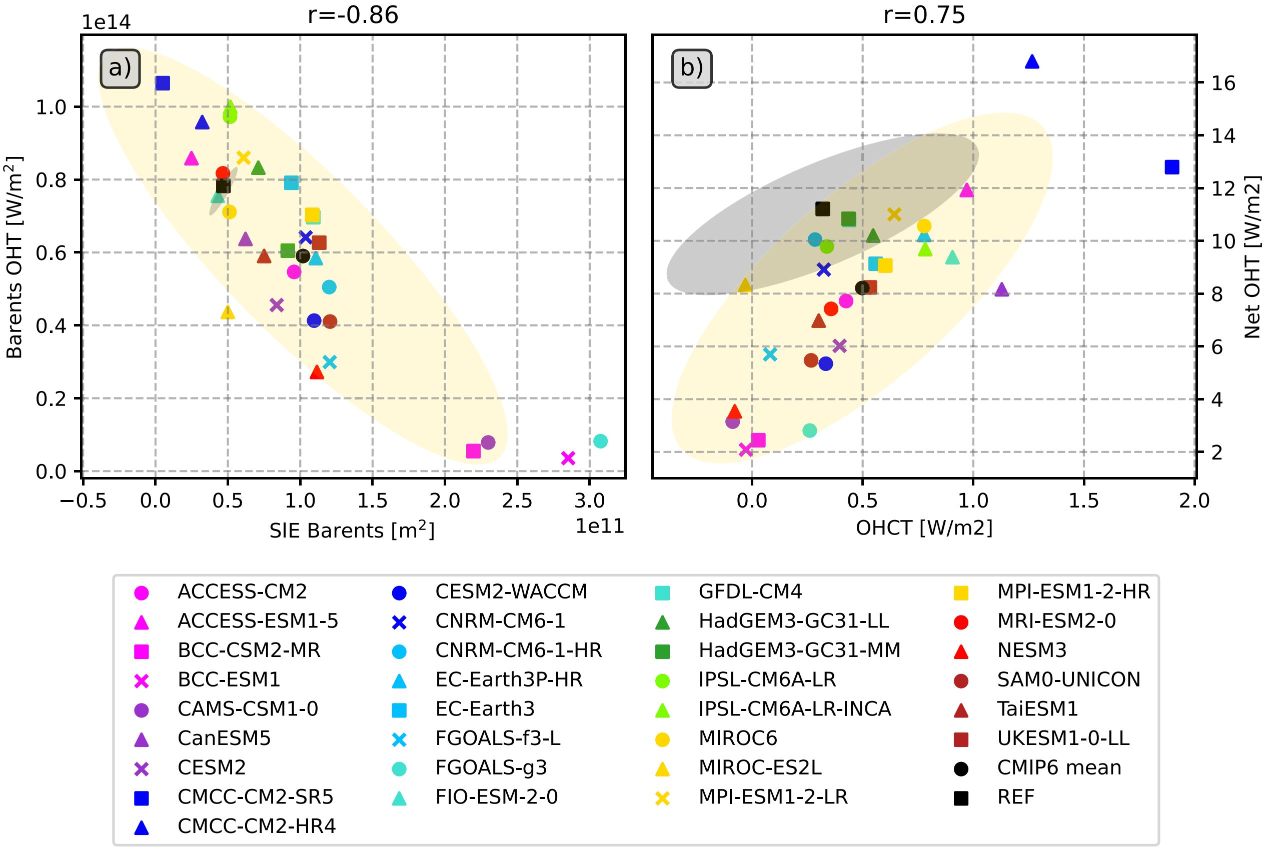

Fig. 12 Scatter plots of the effect of ocean heat transports on sea ice and the ocean warming rate. a) correlations between oceanic heat transports through the Barents Sea Opening and the mean sea ice extent in the Barents Sea, b) correlations between net Arctic oceanic heat transports and the oceanic heat content tendency. All values are long-term annual averages over the 1993–2014 period. Yellow ellipses show the 2-sigma confidence ellipses for the CMIP6 models and grey ellipses for the GREP reanalyses

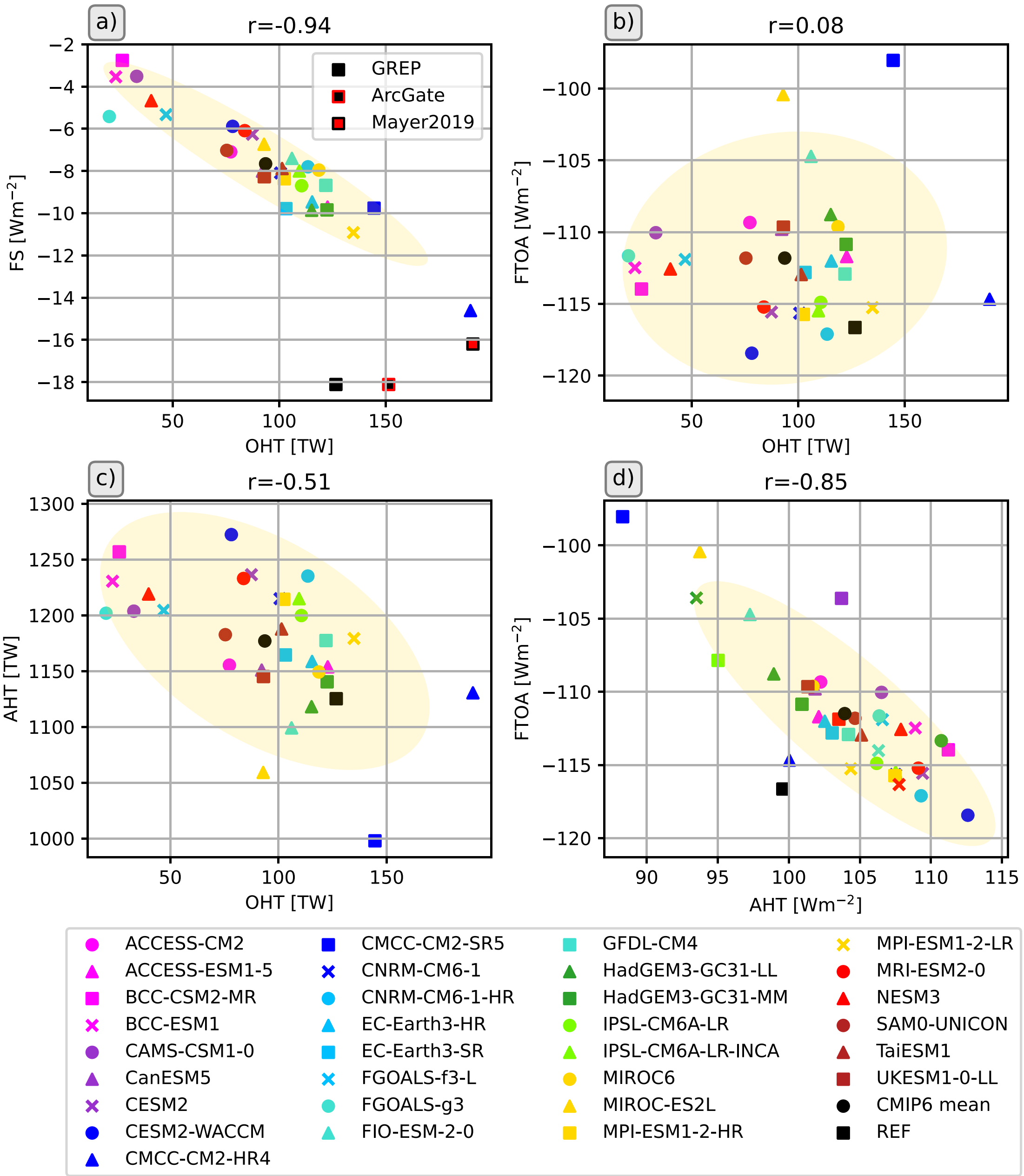

Fig. 13 Scatter plots of long-term annual averages of oceanic heat transports and atmospheric energy budget components. Correlations between **a** OHT and the net surface energy flux F_s, **b** OHT and the net energy flux at the top of the atmosphere F_{TOA}, **c** OHT and atmospheric heat transports AHT, **d** AHT and F_{TOA}. Yellow ellipses show the 2-sigma confidence ellipses for the CMIP6 models

annual cycle from ArcGate is also shown. Almost all models underestimate the heat influx compared to REF and ArcGate. Only CMCC-CM2-HR4 simulates larger heat transports throughout the year than indicated by observations, with an October peak about 25% higher than the ArcGate estimate of about 200 TW. In addition, CMCC-CM2-SR5, MPI-ESM1-2-HR and MPI-ESM1-2-LR exceed observations in spring. The heat transports for BCC-CSM2-MR, BCC-ESM1, CAMS-CSM1-0, FGOALS-g3, FGOALS-f3-L and NESM3 are too low in all calendar months and are mainly caused by biases in the inflow of Atlantic waters through the Barents Sea opening. Heat transports through the individual Arctic straits are shown in the Supplementary material (Fig. S3). In general, while most models are able to simulate the shape of the annual transport cycles to some extent, the inter-model spread is large for all Arctic straits. Seasonal cycles for the

BSO feature similar spreads and biases as the net Arctic heat inflow, reflecting the leading role of BSO in determining the amount of oceanic heat entering the central Arctic. Figure 11 shows correlations between BSO heat transports with BSO volume transports and BSO average ocean temperatures. Correlations are high both for volume transports and temperatures indicating biases in the simulated temperatures and currents. It is worth noting that volume transports and strait average temperatures are not independent of each other and feature moderate to high correlations (not shown). The models with exceptionally low OHT values show mean temperatures around 0 degrees Celsius or even slightly negative values. Additionally, they simulate Norwegian Coastal Currents (NCs) that are generally too weak, slowed down too far south, or even negative, while high OHT values are driven by high volume transports due to strong NCs and higher temperatures. Figure 4 revealed that some of the models (including BCC-CSM2-MR, BCC-ESM1 and CAMS-CSM1-0) feature large positive temperature biases in and underneath the Atlantic water layer. However, in the upper most layers some of those models show negative biases and underestimate the actual temperatures in the surface and halocline layers. Figure S4 shows temperature profiles averaged along the individual straits. The BSO profile reveals that while the largest temperature biases for BCC-CSM2-MR, BCC-ESM1, CAMS-CSM1-0 and FGOALS-g3 are found at the surface, where three of the models even simulate temperatures below 0°C, negative temperature biases are present in all layers of the rather shallow BSO. These low temperatures near the surface are tightly coupled to the overlying sea ice cover. However, the sea ice cover does not only affect heat transport, but the link between oceanic transports and sea ice goes both ways, as increased heat transports also lead to less sea ice (Årthun et al. 2019).

Heat transports through Fram Strait (Fig. S3a) are too small in the majority of models, but it is possible that models with the NLFS scheme are affected by the mass adjustment due to sea ice melt, as the differences in volume transports between models with non-linear and linear surfaces are largest for Fram Strait (not shown), leading to a simulated maximum volume outflow in summer for NLFS models and a minimum about 2 Sv smaller for those without NLFS. Volume transports through Fram Strait are generally biased low in CMIP6 (Heuzé et al. 2023). However, while net transports show a large spread similar to the BSO (Tab. 6), the MMM actually stays well within the uncertainty range of the reference values. Temperature profiles (Fig. S4) at Fram strait show that virtually all models feature a positive temperature bias below 500 m and above that the majority of models features a negative temperature biases. Therefore, the low biases in OHT are mainly caused by warm biased deep waters flowing out of the Arctic through the East Greenlandic Current (EGC) and cold biased waters flowing into the Arctic via the more shallow West Spitsbergen Current (WSC).

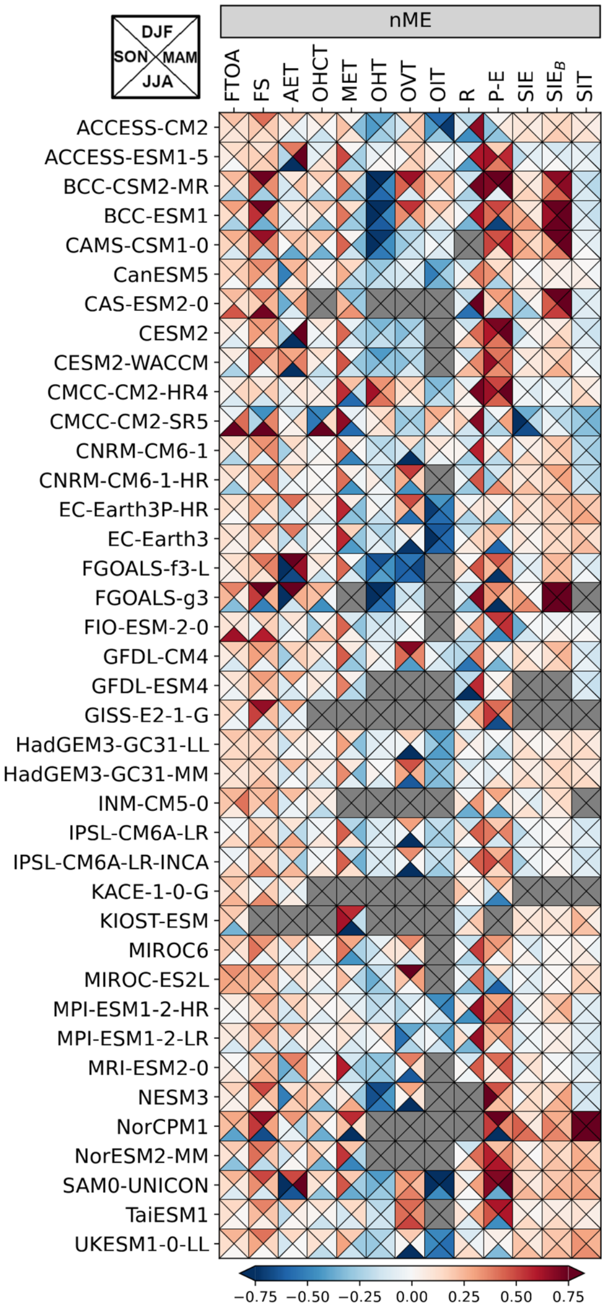

Fig. 14 "Portrait" diagram of seasonal normalized mean errors (nME) for various water and energy budget components for 1993–2014. Triangles indicate the respective seasons DJF (upper triangle), MAM (right triangle), JJA (bottom triangle), and SON (left triangle). SIE$_B$ denotes the sea ice extent in the BSO

The effect of model spatial resolution can be seen for heat transports through Davis Strait (Fig. S3c). All CMIP6 models with a horizontal resolution of 1/4 degree simulate an OHT peak in autumn, similar to the observational ArcGate estimate, while coarser resolution models do not show such a peak. Somewhat surprisingly, the reanalysis-based estimates, which also feature a horizontal resolution of 1/4 degree, do not simulate such a peak, however they are known to have a cold bias in the West Greenland Current (Pietschnig et al. 2017). The high resolution CMIP6 models however feature stronger and warmer West Greenlandic Currents and stronger, but similarly tempered, Baffin Island currents during autumn (see Fig. S5).

The strength of OHT has important implications for the state of the Arctic Ocean and sea ice. Figure 12 shows scatter plots of OHT, sea ice extent and the ocean warming rate OHCT. As mentioned above there is a tight coupling between sea ice and heat transports. The left panel shows this correlation for the BSO, as models with higher/lower heat transports simulate smaller/larger sea ice areas. This leaves two possibilities: either reduced OHTs allow more sea ice to form, or a larger sea ice cover slows down currents, cools the ocean and therefore leads to lower heat transports. While the effect of OHT on Arctic sea ice has been discussed in various observational (e.g., Årthun et al. 2012; Onarheim and Årthun 2017) and modelling (e.g., Årthun et al. 2019; Dörr et al. 2021) studies, the influences of changes in Arctic sea ice on oceanic circulations, temperatures and therefore heat transports have been less investigated and still pose many unknowns (Docquier and Koenigk 2021). More thorough analysis and model experiments would be required to clarify this possible bidirectional effect, but, this is beyond the scope of this study.

Ocean heat transports also affect the change in oceanic temperature. Figure 12 b) shows the correlation of heat transports and the change in ocean heat content: models with larger/smaller OHT show a faster/slower warming of the Arctic Ocean.

The consequences of biases in the oceanic components may also pass over to the Arctic atmosphere, potential effects are shown in Fig. 13. There are strong correlations between simulated long-term averaged OHT and F_s (Fig. 13a), as OHT driven changes of sea ice and ocean temperature strongly affect the reflected shortwave radiation during summer and outgoing longwave radiations as well as turbulent energy fluxes. However, there are no significant correlations

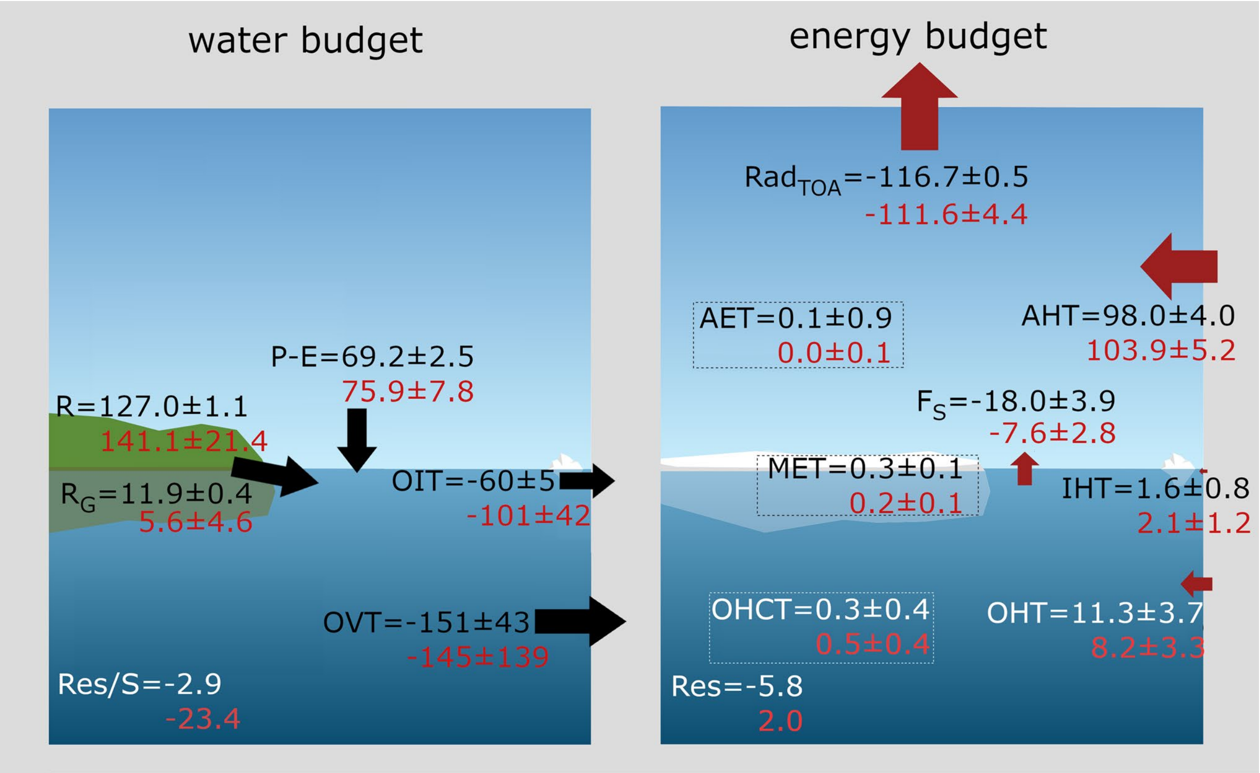

Fig. 15 Water (left, in 10^3 m^3/s) and energy (right, in Wm^{-2}) fluxes and storage rates for the reference estimates (black) and the CMIP6 MMM (red) for 1993–2014. Additional estimates from ArcGate and Winkelbauer et al. (2022) as well as Mayer et al. (2019) are given in Table 7. The graphic designs of the schematics are adapted from Winkelbauer et al. (2022) and Mayer et al. (2019)

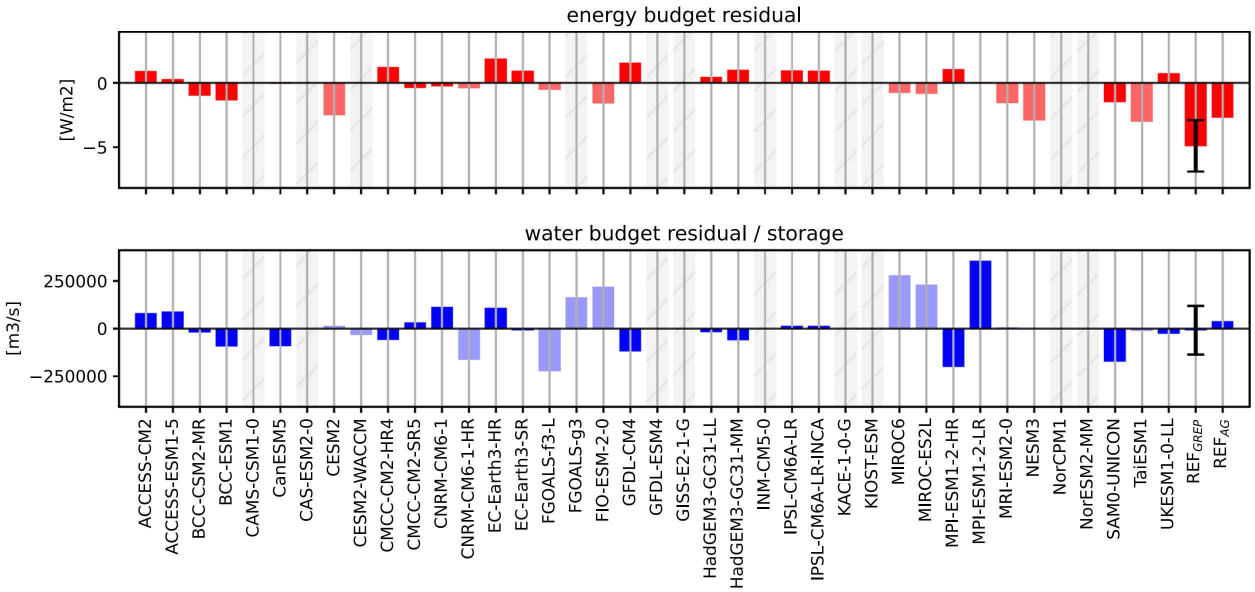

Fig. 16 Budget residuals for the energy (top) and water (bottom) budget of the Arctic Ocean

between OHT and the net radiation at the top of the atmosphere (Fig. 13b) and biases do not seem to reach up to the top of the atmosphere. In contrast, Fig. 13d shows high correlations between long-term averaged atmospheric heat transports AHT and F_{TOA}, as models with weaker outgoing F_{TOA} also feature weaker AHT. OHT and AHT feature moderate anti-correlation (Fig. 13c) and the atmosphere compensates for variances in OHT to the extent that its biases are not seen at the TOA. Models with stronger OHT feature stronger F_s and therefore weaken the atmospheric gradients and subsequently AHT. Note the deviation of the reference values from the model based reference ellipse in 13a. This is caused by inconsistency in our reference based estimates for F_s and OHT and will be discussed futher in the next section.

The impact of OHT on other components of the Arctic system highlights the importance of detecting the exact source of any possible biases therein. To check whether the biases in OHT are also present further south, where less sea ice is present, we calculated transports through the Greenland-Scotland Ridge (GSR, dashed orange line in Fig. 1). Figure S6 shows heat transports at the GSR and figure S7a shows scatter plots between heat transports though the GSR and the sum of transports through Fram Strait and the BSO. They show a high correlation with biases of heat transports being also present further south in the Nordic Seas, tightly coupled to biases in the GSR across strait temperatures (Fig. S7b). Figure S7a shows a group of models slightly to the left of the reference estimates, simulating realistic GSR transports and lower Fram and BSO transports. This shift is actually caused by too much sea ice in the models, which forces the heat out of the ocean in the Nordic Seas through higher outgoing surface energy fluxes.

Figure 14 summarises the seasonal performance of the models and shows normalised mean errors for all models and variables for the energy and water budgets. Seasons are subdivided by triangles, as indicated in the top left-hand corner of the figure. Mean errors for each variable have been normalised by the largest error of the concerning variable to allow for better inter-model comparisons. The closer the values are to 0, the smaller the model bias and the better the model performance. For instance, the net surface energy flux F_s is biased positive from autumn to spring and biased negative in summer for most models, meaning that there is less outgoing energy during the colder seasons and less net incoming energy during summer for those models (see Fig. 9). In contrast, the CMCC-CM2-SR5 model shows a positive bias during summer (more net incoming energy) and a negative bias during winter, caused mainly by its large negative sea ice bias. Further, the connection of biases in sea ice and oceanic heat transports is evident, as models with positive biases in sea ice extent have a negative OHT bias, while biases in sea ice thickness seem to be less relevant with regard to OHT. Seasonal biases in OVT for some models are caused by the models NFL scheme. The affected models are biased negative in summer (stronger outgoing flux due to the sea ice melt effect) and biased positive in winter (effect of ice formation and growth). However, also models without the NFL scheme tend to feature some spurious features. Biases in runoff are mostly due to the one-month shift in the simulated annual cycle, with summer runoff being biased small and spring runoff being biased large. MET biases are largest in the transitional seasons (spring and autumn), while OHCT biases are largest in winter and summer, indicating a weakened amplitude of the annual cycle for most models and an enhanced amplitude for the CMCC-CM2-SR5 model.

4.3 Budget closure

Non-closure of global budgets may contribute to unforced long-term changes/trends in state variables, the so-called model "drift". This may distort the estimate of forced changes in coupled climate simulations and lead to false interpretations. On a global scale Irving et al. (2021) find non-negligible drift trends in time-integrated ocean heat and freshwater fluxes, F_{TOA} and moisture flux into the atmosphere (evaporation minus precipitation), suggesting a considerable leakage of mass and energy in the simulated climate system. To our knowledge, budget closure on a more regional scale for the Arctic area has not been assessed yet for CMIP6. We use all terms from equations 2 and 3 and ignore changes in the oceanic volume storage, which are considered small (Winkelbauer et al. 2022), to assess the energy and water budget closure for the Arctic Ocean:

$$Res_{energy} = F_S - OHCT - MET - \nabla \cdot F_O - \nabla \cdot F_I + L_f(T_p)P_{snow} - L_f \rho_{snow} \frac{\partial d_{snow}}{\partial t} \tag{8}$$

$$Res_{water} = P + ET + R - \nabla \cdot F_{vol} \tag{9}$$

Annual mean fluxes and storage components for the water and energy budgets as simulated by the CMIP6 MMM (red values) and our reference estimates (black values) are shown in Fig. 15. Note that the reference estimates were taken from multiple independent data sources and are not consistent and therefore the observational budget estimates are not closed but rather feature budget residuals. Comparisons with estimates from Mayer et al. (2019) and Winkelbauer et al. (2022) are given in Table 7.

Figure 16 shows residuals for the energy (top) and water (bottom) budgets. With a snowfall term of 1 Wm^{-2} (ERA5) energy budget residuals for the reference estimates using oceanic transports from the GREP ensemble (REF_{GREP}) are at − 4.8 Wm^{-2}. Residuals using oceanic transports from ArcGate (REF_{AG}) are smaller at − 2.6 Wm^{-2}. As already seen in Fig. 13a, the largest inconsistencies are found between net surface energy fluxes derived from a combintation of CERES-EBAF TOA fluxes and atmospheric energy budget quantities provided by Mayer et al. (2021a) and oceanic lateral heat transports from the GREP ensemble. Surface energy fluxes as seen by the ocean reanalyses are actually about 3 Wm^{-2} smaller (not shown), explaining the observation based budget residuals. While Mayer et al. (2021a) use ERA5 data, the ocean reanalyses in GREP are not coupled and use atmospheric forcing from ERA-Interim, which already features significantly smaller surface energy fluxes than ERA5 (not shown). Further, the GREP reanalyses calculate their own upwelling fluxes influenced by their own ice thicknesses and skin temperatures, while ERA5 sees constant sea ice thickness of 1.5m and is known to have a warm temperature biases over sea ice (Wang et al. 2019). For the water budget reference based residuals are at − 2.9×10^3 m^3 (REF_{GREP}) and − 52.1×10^3 m^3 (REF_{AG}).

Figure 16 further shows budget residuals for the individual CMIP6 models. Residuals for the energy budget are comparatively small with values between − 2.5 and 2 Wm^{-2}. Residuals for the water volume budget are mostly smaller than ± 100×10^3 m^3. Some models feature larger residuals (e.g. MPI-ESM1-2-LR), however for those models, as discussed above, our volume transports calculations may not be accurate enough.

There are multiple potential reasons for non-closure, some of them are listed below:

- Even though we are confident in our methods of calculation, we still can not preclude problems with our technical analyses. Especially the calculation of Arctic ocean volume transports is very sensitive to the ocean bathymetry and many large fluxes of opposing sign sum up to a relatively small net transport. Therefore, small inaccuracies in the methods of calculation may lead to major errors in net integrated transports. It also has to be noted, that the needed information to calculate exact oceanic volume transports, like exact ocean depths, is not readily available for all models. This will be discussed further in the conclusions section.
- We consider the budget equations as complete as possible, however there is still the possibility that we are missing some smaller budget terms. While small themselves, they still could have effects when trying to close the budgets. For example, oceanic transports are calculated as the sum of the four major gateways, but we neglect transports through the smaller channels of Hecla and Fury Strait. Further, we ignore the small fluxes associated with the change in sensible heat content of ice (IHCT in Eq. 2) and also the temporal (sub-monthly) eddy component of oceanic transports. Also, it's possible that not all components are provided in the CMIP6 model output, e.g. small terms like numerical diffusion and mass leak increments.
- Imbalances may also arise from deficiencies in the models itself, including model coupling, numerical schemes and/or physical processes. While it is desirable for

regional budgets in climate models to be closed, achieving a perfect closure can be challenging. The closure of regional budgets depends on the accuracy and representation of processes within the model, the spatial and temporal resolution of the model, and the quality of parameterizations. However, due to the complexity of Earth's climate system, including interactions between different components (atmosphere, ocean, land, and ice), achieving complete closure at a regional scale is challenging (Lauritzen et al. 2022).

5 Summary and discussion

This study analyses the performance of 39 CMIP6 models in simulating the energy and water budgets of the Arctic. We find systematic biases in several energy and water budget components and large inter-model spreads in most evaluated parameters when compared to the uncertainty of the observationally constrained estimates.

We assessed model performance by comparing historical long-term averages and seasonal cycles of key energy and water cycle components with observational reference data. The main results of this study are summarised below.

Long-term averaged surface freshwater fluxes tend to be overestimated by most of the models analysed, and apart from large model spreads in their seasonal cycles, we also found an early timing bias of one month in the runoff cycle phase, most likely related to the models' disability to correctly simulate the timing of snow melt and permafrost degradation.

The introduction of the StraitFlux tools Winkelbauer et al. (2023) allowed the calculation of oceanic transports consistent with the discretization schemes of the respective models, allowing a fair comparison and avoiding spurious artifacts that would be caused by interpolation. However, the results of oceanic volume and ice transports show strong biases. Inter-model spread is large and the majority of models fail to simulate the annual cycles of the net Arctic volume transports correctly. The largest errors and some spurious peaks in summer are introduced by the use of the NLFS scheme for sea ice meltwater. Seasonal cycles of volume transports corrected for sea ice volume still show a large spread with some suspicious-looking models, but the MMM of the corrected fluxes is in better agreement with the reference cycles and is within the uncertainty range of the reanalyses during 9 of 12 months. The calculation of volume transports is very sensitive to the exact ocean bathymetry. We found that for the individual Arctic straits biases due to inaccurate handling of the bathymetry are comperatively small and mostly amount to less than 10 %. However, as the individual fluxes sum up to a rather small net Arctic volume flux, those biases may cause some significant errors for the net transports. As discussed in Winkelbauer et al. (2023) caution is advised especially when calculating volume transports for shallow or bathymetrically more complicated straits, where currents are intensified in the proximity of the ocean ground or coast. While the calculation of heat and salinity transports is not as sensitive, it is still not neglectable in most cases. To improve the calculation of transports we would need the exact cell thicknesses either at the positions where the oceanic temperatures and velocities are defined or, if only thicknesses at the middle of the grid cell are supplied, provide the transformation equations to transform the thicknesses to the cell faces (Arakawa-C) or edges (Arakawa-B) for all models. Unfortunately, these data are not available for all CMIP6 models, and we hope the situation will improve in future CMIPs.

Surface energy fluxes in CMIP6 are generally strongly underestimated compared to the observationally constrained reference estimates, with the largest biases occurring in autumn, winter and spring. Radiative fluxes at TOA are closer to observations, but some models still show biases, especially in summer. Errors in F_s and F_{TOA} are closely related to the extent of the simulated sea ice area. Therefore, models with particularly large biases in their simulation of sea ice (CMCC-CM2-SR5, CAS-ESM2-0, FIO-ESM2-0, NorCPM1, FGOALS-g3) also have the largest errors in F_s and F_{TOA}. Also problems in the models' energy conservation may lead to errors in the net global energy budget at TOA and at the surface (Wild 2020), however those errors should be comparatively small.

As with the water budget, also for the energy budget the largest uncertainties and biases seem to be generated in the ocean. While most models are able to correctly simulate the timing of the oceanic lateral heat inflows, the inter-model spread is exceptionally large and most models show a systematic underestimation of the heat transports. Six models (BCC-CSM2-MR, BCC-ESM1, CAMS-CSM1-0, FGOALS-g3, NESM3) simulate particularly small heat transports, mostly due to temperature biases but also because of too weak simulated Barents Sea volume transports. We find strong relationships between lateral oceanic heat transports and the mean state of the Arctic. Furthermore, oceanic transports have strong effects on sea ice cover and ocean warming rates, demonstrating the importance of the mean state on projected trends.

Biases in Arctic deep waters were shown to be caused by the lack of ventilation through shelf overflows and inaccurate oceanic transports (Heuzé et al. 2023). In addition, longer spin-up times may be required as deep waters may take longer to equilibrate to the initial conditions. A more detailed assessment of oceanic transports would be necessary to determine the exact source of these biases.

Despite the use of more accurate oceanic transport estimates and the assessment of more complete budgets, it was still not possible to close the energy and water budgets for the individual models completely. Nevertheless, energy budget residuals are smaller than 2 Wm^{-2} for most models, which is still small when compared to the inter-model spreads in most energy budget components. Small residuals could be due to both technical issues on our side and deficiencies in the models, including model coupling, physical processes and numerical schemes. More extensive evaluations of these imbalances could help to further identify and address biases and limitations, leading to improved representations of regional processes and more balanced budgets.

Furthermore, it must be reiterated that all multi-model averages were computed using all available models without any kind of model weighting, which should be applied to mitigate biases, uncertainties and discrepancies between models and provide a more balanced representation of the overall model ensemble. The results of this study can nevertheless help us to understand typical model biases in the Arctic, and using these results it may be possible to generate physically based metrics to detect outliers from the model ensemble. These metrics may prove may be useful in reducing the spread of future projections of Arctic change.

Large model spreads can be exacerbated by several sources of error. First and foremost, we used only one realisation per model, which is known to introduce a sampling error as each different realisation simulates a different possible outcome of the chaotic climate system (Wang et al. 2022). However, past studies suggested intra-model biases to be quite small compared to inter-model biases (e.g., Zanowski et al. 2021; Khosravi et al. 2022; Wang et al. 2022). We used a bootstraping approach to estimate those sampling errors and found this to be true for for most variables in our study. Also, observations similarly account for only one realisation and therefore the sampling error should be of the same value for our observational estimates. So, in most cases, biases between models and observations when looking at long-term means are very likely to be true systematic biases inherent in the model. However, for variables with larger sampling errors, like e.g. OHCT and MET, and also when looking at trends of the relatively short period of 22-years, variabilities on different time scales may introduce sampling uncertainty. In those cases, the best solution would be to look at longer time scales, whereby this oftentimes is problematic due to the length of available observations and spinup effects during the earlier part of the model simulations.

In addition, errors may be introduced by missing processes or different treatment of processes in the models. For example, as we saw in Fig. 3, the inclusion of a non-linear free surface scheme leads to biased seasonal cycles of oceanic volume transports, at least in the current generation of climate models. Errors in the calculation of energy and water budget variables have been minimised by using the native grid files of all variables where interpolation can corrupt the result.

In conclusion, the biases we find in some of the Arctics' energy and water budgets of the evaluated models have substantial effects on the simulated mean state and changes within the system and therefore possibly also on projections of future warming of the Arctic. To obtain more realistic simulations of the Arctic and processes therein more observations would be needed to constrain the models, as well as higher resolution and improved parametrizations, as already discussed by e.g., Heuzé et al. (2023). Nevertheless, the diagnostics framework presented here can be applied to measure progress made with upcoming new versions of coupled model runs, performed, e.g., within CMIP7. The presented diagnostics may also be used to generate more process-based metrics compared to earlier studies (e.g., Brunner et al. 2020) that focused on state quantities to detect outliers from the model ensemble and therefore reduce the spread of future projections of Arctic change.

Appendix

See Tables 6, 7, 8.

Table 6 Averaged volume and heat transports through the main Arctic gateways for the 1993–2014 period

	Fram		Barents		Davis		Bering	
	Volume	Heat	Volume	Heat	Volume	Heat	Volume	Heat
ACCESS-CM2	− 3.3 (0.01)	27.6 (− 0.02)	2.8 (− 0.09*)	54.7 (− 0.04*)	− 0.2 (1.26*)	3.9 (0.21*)	0.7 (− 0.00)	1.1 (0.05)
ACCESS-ESM1-5	− 4.2 (− 0.09*)	35.6 (0.02)	3.5 (− 0.02*)	85.9 (− 0.01)	− 0.7 (0.64*)	6.3 (− 0.21*)	1.3 (0.02*)	7.0 (0.17*)
BCC-CSM2-MR	− 0.7 (− 0.14*)	21.8 (0.03*)	0.5 (0.21*)	5.5 (0.29*)	− 0.0 (4.65)	0.3 (0.04)	− 0.0 (0.04*)	0.1 (− 0.00)
BCC-ESM1	− 0.3 (− 0.02)	20.1 (0.08*)	0.1 (− 0.27*)	3.6 (− 0.02)	− 0.1 (1.24*)	− 0.2 (0.57)	− 0.0 (− 0.03*)	0.1 (− 0.14*)
CAMS-CSM1-0	− 0.9 (0.14*)	20.7 (− 0.02*)	0.8 (− 0.16*)	7.9 (− 0.29*)	− 0.5 (0.00)	7.1 (0.07*)	0.5 (− 0.08*)	− 0.1 (0.04)
CanESM5	− 3.7 (− 0.22*)	23.4 (0.26*)	3.4 (0.18*)	63.7 (0.35*)	− 0.6 (0.38*)	3.2 (− 0.16*)	0.8 (0.02)	3.0 (0.52*)
CAS-ESM2-0	-	-	-	-	-	-	-	-
CESM2	− 1.2 (− 0.12*)	10.2 (0.12*)	1.8 (0.07*)	45.6 (0.12*)	− 1.9 (0.09*)	6.2 (− 0.11*)	1.1 (0.02*)	6.0 (0.38*)
CESM2-WACCM	− 1.2 (− 0.17*)	9.0 (0.08*)	1.6 (0.12*)	41.3 (0.10*)	− 1.8 (0.07*)	6.4 (− 0.01)	1.1 (0.00)	3.7 (0.44*)
CMCC-CM2-HR4	− 2.5 (0.07*)	58.8 (0.15*)	3.5 (− 0.01)	95.8 (0.00)	− 2.7 (− 0.05*)	25.9 (0.09*)	1.5 (0.01)	8.7 (0.19*)
CMCC-CM2-SR5	− 3.3 (0.01)	17.1 (0.37*)	4.3 (0.01*)	106.4 (0.11*)	− 2.6 (− 0.05*)	8.2 (− 0.03)	1.3 (0.03*)	11.9 (0.34*)
CNRM-CM6-1	− 3.4 (− 0.23*)	24.0 (− 0.06*)	3.4 (− 0.14*)	64.1 (− 0.15*)	− 1.6 (0.76*)	8.8 (− 0.31*)	1.5 (0.01*)	7.0 (0.05)
CNRM-CM6-1-HR	− 1.0 (− 0.19*)	29.8 (0.30*)	1.9 (0.05*)	50.5 (0.18*)	− 2.2 (0.09*)	27.2 (− 0.04*)	0.9 (− 0.06*)	6.6 (0.02)
EC-Earth3P-HR	− 2.3 (− 0.14*)	34.9 (0.25*)	2.1 (0.08*)	58.4 (0.09*)	− 0.9 (0.15*)	16.2 (− 0.10*)	1.1 (0.01)	6.0 (− 0.08*)
EC-Earth3	− 2.6 (− 0.32*)	3.8 (1.80*)	3.2 (0.33*)	79.1 (0.45*)	− 2.4 (− 0.16*)	7.7 (− 0.13*)	1.8 (0.06*)	12.6 (0.37*)
FGOALS-f3-L	− 3.6 (0.11*)	6.1 (0.48*)	2.0 (− 0.04*)	29.9 (− 0.03*)	− 0.9 (0.16*)	14.4 (− 0.21*)	2.1 (− 0.02*)	14.0 (0.06*)
FGOALS-g3	− 1.1 (0.40*)	9.4 (0.29*)	0.6 (− 0.29*)	8.2 (0.50*)	− 1.4 (0.16*)	8.2 (− 0.17*)	1.9 (0.01*)	6.0 (0.43*)
FIO-ESM2-0	− 1.4 (− 0.03*)	11.1 (− 0.08*)	2.4 (− 0.02)	75.5 (− 0.04*)	− 1.9 (0.06*)	10.5 (− 0.10*)	0.9 (0.04*)	6.0 (0.43*)
GFDL-CM4	− 1.7 (− 0.29*)	26.1 (0.08*)	2.4 (0.14*)	69.7 (0.15*)	− 1.8 (0.07*)	24.8 (0.02)	0.9 (0.02*)	1.9 (0.37*)
GFDL-ESM4	–	–	–	–	–	–	–	–
GISS-E2-1-G	–	–	–	–	–	–	–	–
HadGEM3-GC31-LL	− 3.4 (− 0.04*)	23.5 (0.21*)	3.8 (0.00)	83.3 (0.05*)	− 1.8 (0.03*)	7.5 (− 0.11*)	1.3 (0.03*)	2.4 (0.62*)
HadGEM3-GC31-MM	− 2.3 (− 0.17*)	41.2 (0.03)	2.5 (0.06*)	60.5 (0.09*)	− 1.4 (0.07*)	17.9 (0.17*)	1.0 (0.09*)	2.9 (0.34*)
INM-CM5-0	-	-	-	-	-	-	-	-
IPSL-CM6A-LR	− 4.1 (− 0.13*)	2.3 (0.87*)	4.1 (− 0.02*)	97.2 (0.00)	− 1.3 (0.49*)	4.6 (− 0.08*)	1.1 (− 0.01)	2.9 (0.34*)
IPSL-CM6A-LR-INCA	− 3.6 (− 0.26*)	− 1.1 (− 1.59*)	4.0 (0.10*)	100.0 (0.08*)	− 1.7 (0.24*)	3.2 (0.14*)	1.1 (0.12*)	6.6 (0.20*)
KACE-1-0-G	–	–	–	–	–	–	–	–
KIOST-ESM	–	–	–	–	–	–	–	–
MIROC-ES2L	− 3.2 (0.02)	42.7 (0.10*)	2.5 (− 0.07*)	71.1 (0.13*)	0.2 (0.74*)	4.7 (0.03)	0.5 (− 0.03)	3.1 (0.26*)
MIROC6	− 2.8 (0.02*)	40.5 (0.16*)	2.9 (0.06*)	43.7 (− 0.07*)	− 1.1 (− 0.18*)	4.7 (0.04)	1.0 (− 0.00)	3.0 (0.36*)
MPI-ESM1-2-HR	− 2.8 (− 0.11*)	17.5 (− 0.06)	2.7 (0.18*)	70.3 (0.22*)	− 0.8 (− 0.10*)	8.6 (0.16*)	0.7 (0.07*)	3.3 (0.42*)
MPI-ESM1-2-LR	− 2.1 (− 0.32*)	26.0 (− 0.22*)	2.9 (0.12*)	86.0 (0.13*)	− 1.1 (0.23*)	10.7 (− 0.21*)	0.6 (0.04*)	1.6 (0.52*)
MRI-ESM2-0	− 2.6 (− 0.25*)	9.3 (− 0.29*)	_2.5 (0.15*)_	61.1 (0.22*)	− 1.7 (0.00)	6.8 (0.23*)	0.8 (− 0.05*)	6.6
NESM3	− 0.8 (− 0.29*)	10.2 (− 0.11*)	1.7 (0.11*)	27.2 (0.34*)	− 1.5 (0.01)	4.8 (− 0.16*)	0.3 (0.05*)	− 1.9 (0.21*)
NorCPM1	–	–	–	–	–	–	–	–
NorESM2-MM	–	-	–	–	–	–	–	–
SAM0-UNICON	− 1.2 (− 0.21*)	10.0 (− 0.08*)	2.0 (0.24*)	41.0 (0.28*)	− 2.0 (− 0.11*)	10.0 (0.15*)	1.0 (− 0.04*)	0.7 (1.37*)
TaiESM1	− 1.6 (− 0.12*)	8.0 (− 0.30*)	2.3 (0.08*)	59.0 (0.13*)	− 2.0 (0.01*)	9.8 (0.04*)	1.0 (0.04*)	2.0 (0.78*)
UKESM1-0-LL	− 2.0 (− 0.11*)	19.4 (0.07*)	2.8 (− 0.02*)	62.7 (0.06*)	− 2.2 (0.13*)	9.9 (− 0.30*)	1.3 (− 0.03*)	2.2 (0.59*)
MMM	− 2.3 (− 0.11)	20.2 (0.09)	2.4 (0.05)	57.4 (0.11)	− 1.3 (0.11)	9.1 (− 0.03)	0.8 (0.02)	3.6 (0.31)
REF	− 2.2 (− 0.10)	34.0 (0.03)	3.2 (0.07)	78.3 (0.12)	− 2.4 (− 0.00)	9.5 (− 0.16)	1.2 (− 0.02)	4.9 (0.12)
ArcGate	− 1.4	67.6	2.2	54.7	− 1.9	31.8	1.0	− 2.6

Volume Transports are given in SV and heat transports in TW. REF values are calculated using the GREP reanalyses ensemble. Values in brackets show Trends in [fraction/decade]

Table 7 Water and energy fluxes and storage rates for the reference estimates (Ref) and the CMIP6 MMM for 1993–2014

[10^3 m^3/s]	Ref	CMIP6	Winkelbauer et al. (2022)	ArcGate
P-E	69.2 ± 2.5	75.9 ± 7.8	Same as Ref	–
R	127.0 ± 1.1	141.1 ± 21.4	Same as Ref	–
R$_G$	11.9 ± 0.4	5.6 ± 4.6	Same as Ref	–
OVT	− 151 ± 43	− 145 ± 139	− 207	− 91
OIT	− 60 ± 5	− 101 ± 42		− 65
Res/S	− 2.9	− 23.4	− 1.1	− 52.1
[Wm^{-2}]	Ref	CMIP6	Mayer et al. (2019)	ArcGate
Rad$_{TOA}$	− 116.7 ± 0.5	− 111.6 ± 4.4	− 115.8	–
F$_S$	− 18.0 ± 3.9	− 7.6 ± 2.8	− 16.2	–
AET	0.1 ± 0.9	0.0 ± 0.1	− 0.1	–
OHCT	0.3 ± 0.4	0.5 ± 0.4	0.3	–
MET	0.3 ± 0.1	0.2 ± 0.1	0.4	–
AHT	98.0 ± 4.0	103.9 ± 5.2	99.6	–
OHT	11.3 ± 3.7	8.2 ± 3.3	15.5	13.4
IHT	1.6 ± 0.8	2.1 ± 1.2	1.4	–
Res	− 5.8	2.0	0.0	− 3.6

Additionally ArcGate estimates and estimates from Winkelbauer et al. (2022) and Mayer et al. (2019) are given

Table 8 List of Acronyms

ArcGate	Mooring-derived data of oceanic fluxes through the Arctic gateways
BSO	Barents Sea opening
C2SMOS	Sea ice product merged from Cryosat-2 and Soil Moisture and Ocean Salinity satellites
CAA	Canadian Arctic Archipelago
CERES-EBAF	Clouds and the Earth's Radiant Energy System-Energy Balanced and Filled
CGLORS	Ocean reanalyses from the Euro-Mediterranean Center on Climate Change
CMEMS	Copernicus Marine Environment Monitoring Service
CMIP6	Coupled Model Intercomparison Project Phase 6
EGC	East Greenlandic Current
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA5	ECMWF's fifth atmospheric reanalysis
ERA-Interim	ECMWF interim reanalysis
ESGF	Earth System Grid Federation
FOAM	Ocean reanalyses from the UK Met Office
GLORYS	Ocean reanalyses from Mercator Ocean
GRACE	Gravity Recovery and Climate Experiment
GREP	Global ocean Reanalysis Ensemble Product
MMM	Multi-model mean
NC	Norwegian Coastal Current
NEMO	Nucleus for European Modelling of the Ocean
NLFS	Non-linear free surface
NRMSE	Normalized root mean square error
ORAS5	ECMWF's Ocean Reanalysis System 5
OSR	Ocean Reanalysis
REF	Reference data
WSC	West Spitsbergen Current
AE(T)	Atmospheric energy (tendency)
AHT	Atmospheric energy/heat transport
c_p	Specific heat of seawater
d_{snow}	Snow thickness
$\nabla \cdot F_A$	Divergence of lateral atmospheric energy transports
$\nabla \cdot F_I$	Divergence of latent heat transport associated with sea ice transports
$\nabla \cdot F_O$	Divergence of ocean heat transports
$\nabla \cdot F_{vol}$	Divergence of lateral oceanic volume fluxes
ΔS_O	Change of ocean volume
ET	Evapotranspiration
F_s	Surface energy flux
F_{TOA}	Net radiation at the top of the atmosphere (TOA)
IHCT	Sea ice sensible heat content tendency
L_f	Latent heat of fusion
ME(T)	Sea ice melt energy (tendency)
$\vec{n}$	Vector normal to strait
nME	Normalised mean error
OHC(T)	Ocean heat content (tendency)
OHT	Oceanic transports of heat
OIT	Oceanic transports of ice
OVT	Oceanic transports of volume
P	Precipitation
P_{snow}	Snowfall rate
R	Runoff
ρ	Density of sea water
ρ_{snow}	Snow density

Table 8 (continued)

SIE	Sea ice extent
SIE_B	Sea ice extent in the Barents Sea
SIT	Sea ice thickness
θ	Potential temperature
θ_{REF}	Reference temperature
VIWVD	Vertical integral of divergence of moisture flux
$\vec{v}_o$	Ocean velocity vector
$\vec{v}_i$	Ice velocity vector
x	Width along the oceanic strait
x_s, x_e	Start and end point of oceanic strait
z	Ocean depth

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00382-024-07105-5.

Acknowledgements Susanna Winkelbauer and Michael Mayer were supported by the Austrian Science Fund project P33177 and the Copernicus Marine Environment Service contract 21003-COP-GLORAN Lot 7. The authors sincerely thank Richard Renshaw from Met Office for the helpful discussions on GloRanV14 (FOAMv2).

Author contributions SW, MM and LH conceptualized the study. SW performed the data analysis, including the production of the figures in the paper, and prepared the manuscript. MM and LH contributed to the interpretation of results and the writing of the manuscript. All authors have read and agreed to the publication of the present version of the manuscript.

Funding Open access funding provided by University of Vienna. This work was supported by the Austrian Science Fund project P33177 and the Copernicus Marine Environment Service contract 21003-COP-GLORAN Lot 7.

Data availability CMIP6 data is available through the Earth System Grid Federation (ESGF) website https://esgf-node.llnl.gov/search/cmip6/. ERA5 data (Hersbach et al. 2019) as well as mass-consistent energy fluxes derived from ERA5 (Mayer et al. 2021b) are available in the Copernicus Climate Change Service (C3S) Climate Data Store and DEEP-C fluxes are publicly available at https://doi.org/10.17864/1947.271.

Declarations

Conflict of interest The authors declare that they have no competing interests.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

Allan RP, Liu C, Loeb NG et al (2014) Changes in global net radiative imbalance 1985–2012. Geophys Res Lett 41(15):5588–5597. https://doi.org/10.1002/2014GL060962

Årthun M, Eldevik T, Smedsrud LH et al (2012) Quantifying the influence of Atlantic heat on Barents sea ice variability and retreat. J Clim 25(13):4736–4743. https://doi.org/10.1175/JCLI-D-11-00466.1

Årthun M, Eldevik T, Smedsrud L (2019) The role of Atlantic heat transport in future arctic winter sea ice loss. J Clim 32(11):3327–3341. https://doi.org/10.1175/JCLI-D-18-0750.1

Bacon S, Aksenov Y, Fawcett S et al (2015) Arctic mass, freshwater and heat fluxes: methods and modelled seasonal variability. Phil Trans R Soc A. https://doi.org/10.1098/rsta.2014.0169

Bacon S, Garabato A, Aksenov Y et al (2022) Arctic ocean boundary exchanges: a review. Oceanography. https://doi.org/10.5670/oceanog.2022.133

Bao Y, Song Z, Qiao F (2020) Fio-esm version 2.0: model description and evaluation. J Geophys ResOceans. https://doi.org/10.1029/2019JC016036

Bethke I, Wang Y, Counillon F et al (2021) Norcpm1 and its contribution to cmip6 dcpp. Geosci Model Dev 14(11):7073–7116. https://doi.org/10.5194/gmd-14-7073-2021

Bi D, Dix M, Marsland S et al (2020) Configuration and spin-up of access-cm2, the new generation Australian community climate and earth system simulator coupled model. J South Hemisphere Earth Syst Sci. https://doi.org/10.1071/ES19040

Bintanja R, Selten FM (2014) Future increases in arctic precipitation linked to local evaporation and sea-ice retreat. Nature 509:479–482. https://doi.org/10.1038/nature13259

Blackport R, Screen JA (2020) Weakened evidence for mid-latitude impacts of arctic warming. Nat Clim Chang 10(12):1065–1066. https://doi.org/10.1038/s41558-020-00954-y

Bonan DB, Feldl N, Zelinka MD et al (2023) Contributions to regional precipitation change and its polar-amplified pattern under warming. Environ Res Clim 2(3):035010. https://doi.org/10.1088/2752-5295/ace27a

Boucher O, Servonnat J, Albright AL et al (2020) Presentation and evaluation of the ipsl-cm6a-lr climate model. J Adv Model Earth Syst. https://doi.org/10.1029/2019MS002010

Box J, Hubbard A, Bahr D et al (2022) Greenland ice sheet climate disequilibrium and committed sea-level rise. Nat Clim Change. https://doi.org/10.1038/s41558-022-01441-2

Brunner L, Pendergrass AG, Lehner F et al (2020) Reduced global warming from cmip6 projections when weighting models by performance and independence. Earth Syst Dyn 11(4):995–1012. https://doi.org/10.5194/esd-11-995-2020

Cai Z, You Q, Wu F et al (2021) Arctic warming revealed by multiple cmip6 models: evaluation of historical simulations and quantification of future projection uncertainties. J Clim 34:4871–4892. https://doi.org/10.1175/JCLI-D-20-0791.1

Cao J, Wang B, Yang YM et al (2018) The nuist earth system model (nesm) version 3: description and preliminary evaluation. Geosci Model Dev 11(7):2975–2993. https://doi.org/10.5194/gmd-11-2975-2018

Chen HM, Li J, Su JZ et al (2019) Introduction of cams-csm model and its participation in cmip6. Adv Clim Chang Res 15(5):540. https://doi.org/10.12006/j.issn.1673-1719.2019.186

Cheng L, von Schuckmann K, Abraham JP et al (2022) Past and future ocean warming. Nat Rev Earth Environ 3(11):776–794. https://doi.org/10.1038/s43017-022-00345-1

Coumou D, Di Capua G, Vavrus S et al (2018) The influence of arctic amplification on mid-latitude summer circulation. Nat Commun. https://doi.org/10.1038/s41467-018-05256-8

Danabasoglu G, Lamarque JF, Bacmeister J et al (2020) The community earth system model version 2 (cesm2). J Adv Model Earth Syst. https://doi.org/10.1029/2019MS001916

Dee DP, Uppala SM, Simmons AJ et al (2011) The era-interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137(656):553–597. https://doi.org/10.1002/qj.828

Desportes C, Garric G, Régnier C et al (2017) CMEMS-GLO-QUID-001-026, E.U. Copernicus Marine Service Information

Docquier D, Koenigk T (2021) A review of interactions between ocean heat transport and arctic sea ice. Environ Res Lett 16(12):123002. https://doi.org/10.1088/1748-9326/ac30be

Dörr J, Årthun M, Eldevik T et al (2021) Mechanisms of regional winter sea-ice variability in a warming arctic. J Clim 34:1–56. https://doi.org/10.1175/JCLI-D-21-0149.1

Döscher R, Acosta M, Alessandri A et al (2022) The ec-earth3 earth system model for the coupled model intercomparison project 6. Geosci Model Dev 15(7):2973–3020. https://doi.org/10.5194/gmd-15-2973-2022

Dunne JP, Horowitz LW, Adcroft AJ et al (2020) The gfdl earth system model version 4.1 (gfdl-esm 4.1): overall coupled model description and simulation characteristics. J Adv Model Earth Syst. https://doi.org/10.1029/2019MS002015

Eyring V, Bony S, Meehl GA et al (2016) Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. Geosci Model Dev 9(5):1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Fasullo JT, Trenberth KE (2008) The annual cycle of the energy budget. Part I: global mean and land-ocean exchanges. J Clim 21(10):2297–2312. https://doi.org/10.1175/2007JCLI1935.1

Fox-Kemper B, Hewitt H, Xiao C et al (2021) Ocean, cryosphere and sea level change. Cambridge University Press, Cambridge and New York, pp 1211–1362. https://doi.org/10.1017/9781009157896.011

Francis JA, Vavrus SJ (2012) Evidence linking arctic amplification to extreme weather in mid-latitudes. Geophys Res Lett. https://doi.org/10.1029/2012GL051000

Garric G, Parent L, Greiner E, et al (2017) Performance and quality assessment of the global ocean eddy-permitting physical reanalysis GLORYS2V4. In: EGU General Assembly Conference Abstracts, EGU General Assembly Conference Abstracts, p 18776

Goosse H, Kay JE, Armour KC et al (2018) Quantifying climate feedbacks in polar regions. Nat Commun. https://doi.org/10.1038/s41467-018-04173-0

Gosling SN, Arnell NW (2011) Simulating current global river runoff with a global hydrological model: model revisions, validation, and sensitivity analysis. Hydrol Process 25(7):1129–1145. https://doi.org/10.1002/hyp.7727

Haarsma R, Acosta M, Bakhshi R et al (2020) Highresmip versions of ec-earth: Ec-earth3p and ec-earth3p-hr - description, model computational performance and basic validation. Geosci Model Dev 13(8):3507–3527. https://doi.org/10.5194/gmd-13-3507-2020

Haine TW, Curry B, Gerdes R et al (2015) Arctic freshwater export: status, mechanisms, and prospects. Global Planet Change 125:13–35. https://doi.org/10.1016/j.gloplacha.2014.11.013

Hajima T, Watanabe M, Yamamoto A et al (2020) Development of the miroc-es2l earth system model and the evaluation of biogeochemical processes and feedbacks. Geosci Model Dev 13(5):2197–2244. https://doi.org/10.5194/gmd-13-2197-2020

He B, Bao Q, Wang X et al (2019) Cas fgoals-f3-l model datasets for cmip6 historical atmospheric model intercomparison project simulation. Adv Atmos Sci 36(8):771–778. https://doi.org/10.1007/s00376-019-9027-8

Held IM, Guo H, Adcroft A et al (2019) Structure and performance of gfdl's cm4.0 climate model. J Adv Model Earth Syst 11(11):3691–3727. https://doi.org/10.1029/2019MS001829

Hersbach H, Bell B, Berrisford P et al (2020) The era5 global reanalysis. Q J R Meteorol Soc 146(730):1999–2049. https://doi.org/10.1002/qj.3803

Hersbach H, Bell B, Berrisford P, et al (2019) Era5 monthly averaged data on single levels from 1979 to present. https://doi.org/10.24381/cds.f17050d7

Heuzé C, Zanowski H, Karam S et al (2023) The deep arctic ocean and fram strait in cmip6 models. J Clim 36(8):2551–2584. https://doi.org/10.1175/JCLI-D-22-0194.1

Hou Y, Guo H, Yang Y et al (2023) Global evaluation of runoff simulation from climate, hydrological and land surface models. Water Resour Res 59(1):e2021WR031817. https://doi.org/10.1029/2021WR031817

Irving D, Hobbs W, Church J et al (2021) A mass and energy conservation analysis of drift in the cmip6 ensemble. J Clim 34(8):3157–3170. https://doi.org/10.1175/JCLI-D-20-0281.1

Kelley M, Schmidt GA, Nazarenko LS et al (2020) Giss-e2.1: configurations and climatology. J Adv Model Earth Syst. https://doi.org/10.1029/2019MS002025

Khosravi N, Wang Q, Koldunov N et al (2022) The arctic ocean in cmip6 models: biases and projected changes in temperature and salinity. Earth's Future 10(2):e2021EF002282. https://doi.org/10.1029/2021EF002282

Knutti R (2008) Should we believe model predictions of future climate change? Philos Trans Royal Soc A Math Phys Eng Sci 366(1885):4647–4664. https://doi.org/10.1098/rsta.2008.0169

Kouki K, Räisänen P, Luojus K et al (2022) Evaluation of northern hemisphere snow water equivalent in cmip6 models during 1982–2014. Cryosphere 16(3):1007–1030. https://doi.org/10.5194/tc-16-1007-2022

Kwok R (2018) Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). Environ Res Lett 13(10):105005. https://doi.org/10.1088/1748-9326/aae3ec

Lauritzen P, Kevlahan N, Toniazzo T et al (2022) Reconciling and improving formulations for thermodynamics and conservation principles in earth system models (esms). J Adv Model Earth Syst. https://doi.org/10.1029/2022MS003117

Lee J, Kim J, Sun MA et al (2020) Evaluation of the Korea meteorological administration advanced community earth-system model (k-ace). Asia Pac J Atmos Sci. https://doi.org/10.1007/s13143-019-00144-7

Li L, Yu Y, Tang Y et al (2020) The flexible global ocean-atmosphere-land system model grid-point version 3 (fgoals-g3): description and evaluation. J Adv Model Earth Syst 12(9):e2019MS002012. https://doi.org/10.1029/2019MS002012

Liu C, Allan RP, Mayer M et al (2020) Variability in the global energy budget and transports 1985–2017. Clim Dyn 55:3381–3396. https://doi.org/10.1007/s00382-020-05451-8

Loeb NG, Doelling DR, Wang H et al (2018) Clouds and the earth's radiant energy system (ceres) energy balanced and filled (ebaf) top-of-atmosphere (toa) edition-4.0 data product. J Clim 31(2):895–918. https://doi.org/10.1175/JCLI-D-17-0208.1

MacLachlan C, Arribas A, Peterson KA et al (2015) Global seasonal forecast system version 5 (glosea5): a high-resolution seasonal forecast system. Q J R Meteorol Soc 141(689):1072–1084. https://doi.org/10.1002/qj.2396

Madec G (2016) Nemo ocean engine—version 3.6. Note du Pole de modélisation, Institut Pierre-Simon Laplace (IPSL) 27

Mauritsen T, Bader J, Becker T et al (2019) Developments in the mpi-m earth system model version 1.2 (mpi-esm1.2) and its response to increasing co2. J Adv Model Earth Syst 11(4):998–1038. https://doi.org/10.1029/2018MS001400

Mayer M, Haimberger L, Pietschnig M et al (2016) Facets of arctic energy accumulation based on observations and reanalyses 2000–2015. Geophys Res Lett 43(19):10,420-10,429. https://doi.org/10.1002/2016GL070557

Mayer M, Haimberger L, Edwards JM et al (2017) Toward consistent diagnostics of the coupled atmosphere and ocean energy budgets. J Clim 30(22):9225–9246. https://doi.org/10.1175/JCLI-D-17-0137.1

Mayer M, Tietsche S, Haimberger L et al (2019) An improved estimate of the coupled arctic energy budget. J Clim 32(22):7915–7934. https://doi.org/10.1175/JCLI-D-19-0233.1

Mayer J, Mayer M, Haimberger L (2021) Consistency and homogeneity of atmospheric energy, moisture, and mass budgets in era5. J Clim 34(10):3955–3974. https://doi.org/10.1175/JCLI-D-20-0676.1

Mayer M, Vidar SL, Kjell AM et al (2021) Ocean heat content in the highnorth. J Oper Oceanogr 14(Suppl. 1):S17–S23. https://doi.org/10.1080/1755876X.2021.1946240

Mayer M, Tsubouchi T, von Schuckmann K et al (2022) Atmospheric and oceanic contributions to observed nordic seas and arctic ocean heat content variations 1993–2020. J Oper Oceanogr 15(Suppl. 1):S20–S28. https://doi.org/10.1080/1755876X.2022.2095169

Mayer J, Mayer M, Haimberger L (2021b) Mass-consistent atmospheric energy and moisture budget monthly data from 1979 to present derived from era5 reanalysis. https://doi.org/10.24381/cds.c2451f6b

McPhee MG, Proshutinsky A, Morison JH et al (2009) Rapid change in freshwater content of the arctic ocean. Geophys Res Lett. https://doi.org/10.1029/2009GL037525

Moon T, Ahlstrøm A, Goelzer H et al (2018) Rising oceans guaranteed: arctic land ice loss and sea level rise. Curr Clim Change Rep. https://doi.org/10.1007/s40641-018-0107-0

Mouginot J, Rignot E, Bjørk A et al (2019) Forty-six years of Greenland ice sheet mass balance from 1972 to 2018. Earth Atmos Planet Sci 116(19):9239–9244. https://doi.org/10.1073/pnas.1904242116

Muilwijk M, Smedsrud LH, Ilicak M et al (2018) Atlantic water heat transport variability in the 20th century arctic ocean from a global ocean model and observations. J Geophys Res Oceans 123(11):8159–8179. https://doi.org/10.1029/2018JC014327

Onarheim IH, Årthun M (2017) Toward an ice-free barents sea. Geophys Res Lett 44(16):8387–8395. https://doi.org/10.1002/2017GL074304

Pak G, Noh Y, Lee MI et al (2021) Korea institute of ocean science and technology earth system model and its simulation characteristics. Ocean Sci J. https://doi.org/10.1007/s12601-021-00001-7

Park S, Shin J, Kim S et al (2019) Global climate simulated by the Seoul national university atmosphere model version 0 with a unified convection scheme (sam0-unicon). J Clim 32(10):2917–2949. https://doi.org/10.1175/JCLI-D-18-0796.1

Pietschnig M, Mayer M, Tsubouchi T et al (2017) Volume and temperature transports through the main arctic gateways: a comparative study between an ocean reanalysis and mooring-derived data. Ocean Sci Discussions 2017:1–32. https://doi.org/10.5194/os-2017-98

Pithan F, Jung T (2021) Arctic amplification of precipitation changes-the energy hypothesis. Geophys Res Lett 48(21):e2021GL094977. https://doi.org/10.1029/2021GL094977

Proshutinsky A, Krishfield R, Timmermans ML et al (2009) Beaufort gyre freshwater reservoir: state and variability from observations. J Geophys Res Oceans. https://doi.org/10.1029/2008JC005104

Rabe B, Karcher M, Schauer U et al (2011) An assessment of arctic ocean freshwater content changes from the 1990s to the 2006–2008 period. Deep Sea Res Part I Oceanogr Res Papers 58(2):173–185. https://doi.org/10.1016/j.dsr.2010.12.002

Rantanen M, Karpechko AY, Lipponen A et al (2022) The arctic has warmed nearly four times faster than the globe since 1979. Commun Earth Environ. https://doi.org/10.1038/s43247-022-00498-3

Ricker R, Hendricks S, Kaleschke L et al (2017) A weekly arctic sea-ice thickness data record from merged cryosat-2 and smos satellite data. Cryosphere 11(4):1607–1623. https://doi.org/10.5194/tc-11-1607-2017

Roullet G, Madec G (2000) Salt conservation, free surface, and varying levels: a new formulation for ocean general circulation models. J Geophys Res Oceans 105(C10):23927–23942. https://doi.org/10.1029/2000JC900089

Rowland JC, Jones CE, Altmann G et al (2010) Arctic landscapes in transition: responses to thawing permafrost. EOS Trans Am Geophys Union 91(26):229–230. https://doi.org/10.1029/2010EO260001

Schauer U, Beszczynska-Möller A (2009) Problems with estimation and interpretation of oceanic heat transport—conceptual remarks for the case of fram strait in the arctic ocean. Ocean Sci 5(4):487–494. https://doi.org/10.5194/os-5-487-2009

Screen JA, Simmonds I (2013) Exploring links between arctic amplification and mid-latitude weather. Geophys Res Lett 40(5):959–964. https://doi.org/10.1002/grl.50174

Seland Ø, Bentsen M, Olivié D et al (2020) Overview of the Norwegian earth system model (noresm2) and key climate response of cmip6 deck, historical, and scenario simulations. Geosci Model Dev 13(12):6165–6200. https://doi.org/10.5194/gmd-13-6165-2020

Sellar AA, Jones CG, Mulcahy JP et al (2019) Ukesm1: description and evaluation of the u.k. earth system model. J Adv Model Earth Syst 11(12):4513–4558. https://doi.org/10.1029/2019MS001739

Serreze MC, Barrett AP, Stroeve JC et al (2009) The emergence of surface-based arctic amplification. Cryosphere 3:11–19

Shu Q, Wang Q, Song Z et al (2020) Assessment of sea ice extent in cmip6 with comparison to observations and cmip5. Geophys

Res Lett 47(9):e2020GL087965. https://doi.org/10.1029/2020GL087965

Shu Q, Wang Q, Årthun M et al (2022) Arctic ocean amplification in a warming climate in cmip6 models. Sci Adv 8(30):eabn9755. https://doi.org/10.1126/sciadv.abn9755

Storto A, Masina S (2016) C-glorsv5: an improved multipurpose global ocean eddy-permitting physical reanalysis. Earth Syst m Sci Data 8(2):679–696. https://doi.org/10.5194/essd-8-679-2016

Storto A, Alvera-Azcárate A, Balmaseda MA et al (2019) Ocean reanalyses: recent advances and unsolved challenges. Front Marine Sci. https://doi.org/10.3389/fmars.2019.00418

Storto A, Masina S, Simoncelli S et al (2019) The added value of the multi-system spread information for ocean heat content and steric sea level investigations in the cmems grep ensemble reanalysis product. Clim Dyn. https://doi.org/10.1007/s00382-018-4585-5

Stroeve J, Notz D (2018) Changing state of arctic sea ice across all seasons. Environ Res Lett 13(10):103001. https://doi.org/10.1088/1748-9326/aade56

Swart NC, Cole JNS, Kharin VV et al (2019) The Canadian earth system model version 5 (canesm5.0.3). Geosci Model Dev 12(11):4823–4873. https://doi.org/10.5194/gmd-12-4823-2019

Tatebe H, Ogura T, Nitta T et al (2019) Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in miroc6. Geosci Model Dev 12(7):2727–2765. https://doi.org/10.5194/gmd-12-2727-2019

Trenberth KE, Fasullo JT, Mackaro J (2011) Atmospheric moisture transports from ocean to land and global energy flows in reanalyses. J Clim 24(18):4907–4924. https://doi.org/10.1175/2011JCLI4171.1

Tsubouchi T, Bacon S, Naveira Garabato AC et al (2012) The arctic ocean in summer: a quasi-synoptic inverse estimate of boundary fluxes and water mass transformation. J Geophys Res Oceans. https://doi.org/10.1029/2011JC007174

Tsubouchi T, Bacon S, Aksenov Y et al (2018) The arctic ocean seasonal cycles of heat and freshwater fluxes: observation-based inverse estimates. J Phys Oceanogr 48(9):2029–2055. https://doi.org/10.1175/JPO-D-17-0239.1

Voldoire A, Saint-Martin D, Sénési S et al (2019) Evaluation of cmip6 deck experiments with cnrm-cm6-1. J Adv Model Earth Syst 11(7):2177–2213. https://doi.org/10.1029/2019MS001683

von Schuckmann K, Cheng L, Palmer MD et al (2020) Heat stored in the earth system: where does the energy go? Earth Syst Sci Data 12(3):2013–2041. https://doi.org/10.5194/essd-12-2013-2020

Walsh JE (2014) Intensified warming of the arctic: causes and impacts on middle latitudes. Global Planet Change 117:52–63. https://doi.org/10.1016/j.gloplacha.2014.03.003

Wang C, Graham RM, Wang K et al (2019) Comparison of era5 and era-interim near-surface air temperature, snowfall and precipitation over arctic sea ice: effects on sea ice thermodynamics and evolution. Cryosphere 13(6):1661–1679. https://doi.org/10.5194/tc-13-1661-2019

Wang YC, Hsu HH, Chen CA et al (2021) Performance of the Taiwan earth system model in simulating climate variability compared with observations and cmip6 model simulations. J Adv Model Earth Syst 13(7):e2020MS002353. https://doi.org/10.1029/2020MS002353

Wang S, Wang Q, Wang M et al (2022) Arctic ocean freshwater in cmip6 coupled models. Earth's Future 10(9):e2022EF002878. https://doi.org/10.1029/2022EF002878

Wild M (2020) The global energy balance as represented in cmip6 climate models. Clim Dyn. https://doi.org/10.1007/s00382-020-05282-7

Williams KD, Copsey D, Blockley EW et al (2018) The met office global coupled model 3.0 and 3.1 (gc3.0 and gc3.1) configurations. J Adv Model Earth Syst 10(2):357–380. https://doi.org/10.1002/2017MS001115

Winkelbauer S (2023) Straitflux. https://doi.org/10.5281/zenodo.10053555

Winkelbauer S, Mayer M, Seitner V et al (2022) Diagnostic evaluation of river discharge into the arctic ocean and its impact on oceanic volume transports. Hydrol Earth Syst Sci 26(2):279–304. https://doi.org/10.5194/hess-26-279-2022

Winkelbauer S, Mayer M, Haimberger L (2023) Straitflux—precise water strait fluxes on various modelling grids. [Manuscript in preparation]

Wu T, Lu Y, Fang Y et al (2019) The Beijing climate center climate system model (bcc-csm): the main progress from cmip5 to cmip6. Geosci Model Dev 12(4):1573–1600. https://doi.org/10.5194/gmd-12-1573-2019

Wu RJ, Lo MH, Scanlon BR (2021) The annual cycle of terrestrial water storage anomalies in cmip6 models evaluated against grace data. J Clim 34(20):8205–8217. https://doi.org/10.1175/JCLI-D-21-0021.1

Yukimoto S, Kawai H, Koshiro T et al (2019) The meteorological research institute earth system model version 2.0, mri-esm2.0: description and basic evaluation of the physical component. J Meteorol Soc Jpn Ser II 97(5):931–965. https://doi.org/10.2151/jmsj.2019-051

Zanowski H, Jahn A, Holland MM (2021) Arctic ocean freshwater in cmip6 ensembles: declining sea ice, increasing ocean storage and export. J Geophys Res Oceans 126(4):e2020JC016930. https://doi.org/10.1029/2020JC016930

Zhang H, Zhang M, Jin J et al (2020) Description and climate simulation performance of cas-esm version 2. J Adv Model Earth Syst 12(12):e2020MS002210. https://doi.org/10.1029/2020MS002210

Ziehn T, Chamberlain MA, Law RM et al (2020) The Australian earth system model: Access-esm1.5. J South Hemisphere Earth Syst Sci. https://doi.org/10.1071/ES19035

Zuo H, Balmaseda M, Mogensen K (2015) The new eddy-permitting orap5 ocean reanalysis: description, evaluation and uncertainties in climate signals. Clim Dyn 49(3):791–811. https://doi.org/10.1007/s00382-015-2675-1

3.3 StraitFlux - Precise computations of Water Strait fluxes on various Modelling Grids

3.3.1 Overview

The third publication describes new tools that were developed to tackle the challenges in the precise calculation of oceanic transports on models' native grids due to the nature and distortion of the curvilinear ocean model grids used.

StraitFlux works on different modelling grids, independent of the grid pole placement, orientation, Arakawa partition and naming conventions. It is available via zenodo (Winkelbauer, 2023) and github (`https://github.com/susannawinkelbauer/StraitFlux`) and downloadable via pypi.

3.3.2 Publication details

- **Title:** StraitFlux - Precise computations of Water Strait fluxes on various M odelling Grids
- **Authors:** Susanna Winkelbauer, Michael Mayer, and Leopold Haimberger
- **Publisher:** Geoscientific Model Development
- **Type:** Research article
- **Status:** Accepted on 15.04.2024, waiting for publication
- **DOI:** `https://doi.org/10.5194/egusphere-2023-2883`, preprint
- **Own contribution:** The idea and concept of the methods was first created for one specific grid type (ORCA0.25) and was developed in collaboration with all co-authors. My contributions further involved the advancement of the methods to work on various modelling grids, the implementation of the methods in a versatile Python package and the preparation of the manuscript, including the preparation of the data and the visualization of the methods and results. All co-authors took part in the interpretation and discussion of results. My contributions to this publication are estimated to be at least 80%.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

StraitFlux - Precise computations of Water Strait fluxes on various Modelling Grids

Susanna Winkelbauer[1,2], Michael Mayer[1,2,3], and Leopold Haimberger[1,2]
[1]Department of Meteorology and Geophysics, University of Vienna, Vienna Austria
[2]b.geos, Korneuburg, Austria
[3]European Centre for Medium-Range Weather Forecasts, Bonn, Germany
Correspondence: Susanna Winkelbauer (susanna.winkelbauer@univie.ac.at)

Abstract. Oceanic transports shape the global climate, but the evaluation and validation of this key quantity based on reanalysis and model data is complicated by the distortion of the used curvilinear ocean model grids towards their displaced north poles. Combined with the large number of different grid types, this has made the exact calculation of oceanic transports a challenging and time-consuming task. Use of data on standard latitude/longitude grids is not an option since transports computed from those are not mass consistent. We present two methods for transport calculations on grids with variously shifted north poles, different orientations, and different Arakawa partitions. The first method calculates net transports through arbitrary sections using line integrals, while the second method generates cross-sections of the vertical-horizontal planes of these sections using vector projection algorithms. Apart from the input data on the original model grids the user only needs to specify the start and end points of the required section to get the net transports (with both methods) and their integrand (for the second method). This allows to calculate oceanic fluxes through almost arbitrary sections, to compare them with observed oceanic volume and energy transports at available sections such as the RAPID array or at Fram strait and other Arctic gateways, or to compare them amongst reanalyses and to model integrations from the Coupled Model Intercomparison Projects (CMIP).

We implemented our methods in a Python package called StraitFlux. This paper represents its scienfic documentation and demonstrates its application on outputs of multiple CMIP6 models and several ocean reanalyses. We also analyse the robustness and computational performance of the tools as well as the uncertainties of the results. The package is available on github and zenodo and can be installed using pypi.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

1 Introduction

Oceanic transports of heat, volume, and salt are integral components of the Earth's energy and mass budgets, playing a key role in regulating the Earth's climate. For instance, the Atlantic Meridional Overturning Circulation (AMOC) plays a crucial role heating the North Atlantic by transporting warm surface waters from the tropics to the North Atlantic via the Gulf Stream and cold, dense waters southward at depth. It influences the weather and climate of eastern North America and western Europe (Jackson et al., 2015), and subsequently also affects the Arctic climate and sea ice (Liu and Fedorov, 2022; Mahajan et al., 2011). A weakening of the AMOC has been reported (Caesar et al., 2018; Mayer et al., 2023a; Rahmstorf et al., 2015) over recent decades and a potential future collapse of the AMOC (Ditlevsen and Ditlevsen, 2023) would have major effects on the North Atlantic region and beyond. Monitoring oceanic currents is therefore particularly important in today's rapidly changing climate.

There exist several mooring arrays and other measurement devices capable of recording deep water velocities and other sea state variables in the oceans. For example, there are mooring lines in the Arctic gateways (Tsubouchi et al., 2012, 2018) or the Rapid Climate Change–Meridional Overturning Circulation and Heatflux Array (RAPID-MOCHA, e.g., Rayner et al., 2011) and the Overturning in the Subpolar North Atlantic Program (OSNAP, Lozier et al., 2017) for measuring the AMOC. It is desirable to compare transports calculated from those instruments with ocean reanalyses and climate models. This is challenging because the moorings are not aligned with the model grids, and the grids of ocean models, particularly in the Arctic, are complicated. A tool that facilitates a consistent comparison of flux estimates from this growing set of sources is therefore needed.

The convergence of meridians towards the North Pole pose challenges in ocean modelling. Murray (1996) proposed several global orthogonal curvilinear grids where the North Pole is placed over land areas in order to avoid singularities over the ocean. Those ideas were picked up by many modelling centres and are now commonly used in the world of ocean modelling. Fig. 1 shows examples of the two most common grid types. Many ocean models use tripolar grids, where two mesh north poles are placed over North America and Eurasia, whereat the exact location of those two North Poles varies between models. Ellipses around those poles and their normals create the new grid-lines (Madec and Imbard, 1996), which are strongly displaced in the northern latitudes when compared to a regular dipolar grid. The second main grid type is the displaced dipolar grid, where the North Pole is displaced to somewhere over land areas, usually Greenland. Hereby especially the gridlines in the proximity of the artifical pole feature a strong distortion. While solving the numerical problem of a singularity over the ocean, those curvilinear grids complicate the calculation of oceanic transports, especially in the proximity of the poles, as velocities in the direction of the artificial poles do not point in the direction of the true north. The exact position of the poles and the number and angle of the grid-lines varies between different models, forming a vast amount of different grid types that complicate inter-comparison between different models and to observations.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

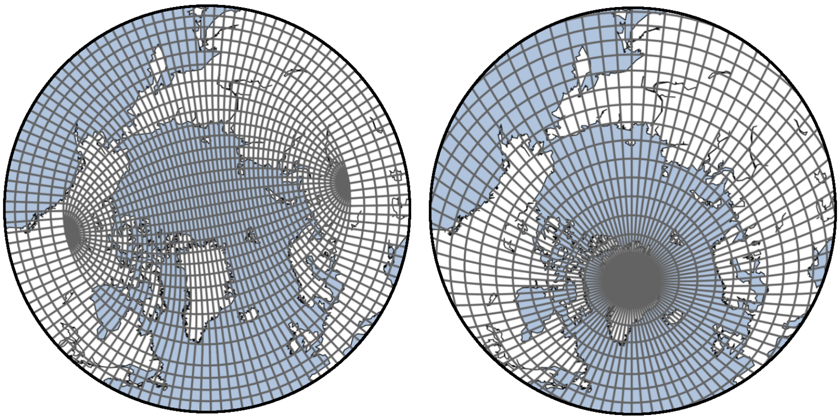

Figure 1. Examples of two curvilinear grids typically used for ocean modelling. Left: a tripolar grid with 2 northern poles (1 over Eurasia, 1 over North America); Right: a displaced dipolar grid with one northern pole displaced over Greenland.

For optimal accuracy and consistency the transports have to be calculated on the native grids. Horizontal interpolation of vector quantities (u,v) onto more convenient regular grids prior to the transport calculation compromises the conservation properties of the respective models, potentially leading to spurious effects and misleading results (an example is provided in Sect. 3). Finding the nearest points of the model grid has so far been done mostly manually for selected straits (e.g., Heuzé et al., 2023). This is time consuming and becomes increasingly impractical when dealing with multiple models and multiple straits and increasing model resolution.

We have developed two methods for calculating oceanic transports on arbitrary oceanic sections, independent of grid pole placement, orientation and Arakawa partition. The first method, using line integration, yields net transports of volume, heat, salt, and ice across defined straits. The second method employs vector projection algorithms to generate cross-sections of currents, temperature, and salinity in the vertical plane. We will refer to the methods as Line Integration Method (LM) and Vector Projection Method (VPM). For both methods the tedious point selection process is fully automatized. We tested our methods on various tri- and dipolar grids from the Climate Model Intercomparison Project Phase 6 (CMIP6, Eyring et al. (2016)) and show some exemplary results in Sect. 3.1.

This paper is structured as follows: In Sect. 2, we describe the fundamental concepts of both calculation methods and their mathematical foundation. Furthermore, in Sect. 2.2, we describe the implementation of the methods in the open-source Python package StraitFlux. Sect. 3 assesses the robustness of the tools, examines their accuracy, provides application examples, and analyzes their computational performance. The final section outlines the strengths and weaknesses of StraitFlux and draws conclusions regarding its utility.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

2 Methods

2.1 Mathematics/General idea

The general idea is to calculate oceanic volume, heat, salinity and ice transports across any chosen vertical section, typically a straight section between two land masses. We define the oceanic transports of volume (OVT), heat (OHT), salinity (OST) and ice (OIT) through a given strait as follows:

$$OVT = \int_{x_1}^{x_2} \int_{0}^{z_b(x)} \boldsymbol{v}_o(x,z) \cdot \boldsymbol{n} \, dz \, dx \tag{1}$$

$$OST = \int_{x_1}^{x_2} \int_{0}^{z_b(x)} S\boldsymbol{v}_o(x,z) \cdot \boldsymbol{n} \, dz \, dx \tag{2}$$

$$OHT = c_p \rho \int_{x_1}^{x_2} \int_{0}^{z_b(x)} (\theta(x,z) - \theta_{ref}) \boldsymbol{v}_o(x,z) \cdot \boldsymbol{n} \, dz \, dx \tag{3}$$

$$OIT = \int_{x_1}^{x_2} d(x) \boldsymbol{v}_i(x) \cdot \boldsymbol{n} \, dx \tag{4}$$

where $\boldsymbol{v}_o$ ($\boldsymbol{v}_i$) represents the velocity vector of liquid water (sea ice) and $\boldsymbol{n}$ is the vector normal to the strait - therefore their product gives velocities normal to the considered coast-to-coast section. Further, x defines the along-strait extent and z its depth. The boundaries z_b, x_1 , x_2 have to be chosen such that no water can "escape" the desired coast-to-coast section. This can be ensured if x_e and x_w are land points and the auxiliary fields describing model ocean depths are used appropriately. S is the sea water salinity, c_p the specific heat of seawater, ρ the density of seawater and θ the potential temperature. Throughout this study, we will use θ and T both synonymously for the potential temperature of seawater. For the validation in Sect. 3, c_p and ρ are set to 3996 $\mathrm{Jkg^{-1}K^{-1}}$ and 1026 $\mathrm{kgm^{-3}}$, respectively, as default values. However it is easy to adapt those values to individual model needs. Schauer and Losch (2019) correctly point out that true heat transports would actually demand closed volume transports through the examined straits, which is generally not the case. However, the "heat fluxes" as defined above are commonly used to ensure comparability with transports derived from observations. Therefore, we will further refer to heat transports when calculating "heat fluxes" as defined here. Additionally each model's heat flux should be computed relative to a reference temperature θ_{ref}, representing the mean temperature of the assessed flow. The validation part of this paper focuses on Arctic straits, therefore, we follow e.g. Heuzé et al. (2023); Muilwijk et al. (2018) and choose a universal reference temperature

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

of 0°C. Generally, in StraitFlux, reference temperatures are set per default to 0°C, but may be changed to individual needs. Furthermore, following the approach of Schauer and Losch (2019) and Heuzé et al. (2023), salinity fluxes are calculated instead of freshwater fluxes to avoid the need for a reference salinity, which would vary for each model. This simplification facilitates model comparisons.

2.1.1 Line Integration Method (LM)

The basic principle of LM is outlined in Fig. 2a. A closed line is generated following the strait (red line) as closely as possible by connecting the faces of the individual grid cells (blue line). For an Arakawa-C grid (see inset of 2a for definition), the u and v components are positioned exactly at the faces of the grid cells. This positioning allows for a straightforward integration of the u/v components along the meridional/zonal width of the grid cell and its depth. The sum of all grid cells vertically and horizontally then provides the net transport through the strait. In the case of an Arakawa-B grid, however, the u and v components are defined at the edges of the grid cells and must be transformed to the faces of the tracer grid cells. To obtain u and v velocities at the faces of the tracer cells, we adapt equations 6.46 and 6.50 from Griffies et al. (2004) (in accordance we also use the same multi-letter variable names). The u velocities at eastern faces of the tracer cells (uet) are calculated as:

$$uet_{i,j,k} = \frac{dyu_j\, dhu_j\, u_j + dyu_{j-1}\, dhu_{j-1}\, u_{j-1}}{2\, dyet_{i,j}} \tag{5}$$

with dyu_j the meridional width of the u/v cell, dhu_j the depth of the u/v cell and $dyet_{i,j}$ the meridional width of the tracer cell's east side. The v velocities at northern faces of the tracer cells (vnt) are calculated as:

$$vnt_{i,j,k} = \frac{dxu_i\, dhu_i\, v_i + dxu_{i-1}\, dhu_{i-1}\, v_{i-1}}{2\, dxnt_{i,j}} \tag{6}$$

with dxu_j the zonal width of the u/v cell, dhv_j the depth of the u/v cell and $dxnt_{i,j}$ the zonal width of the tracer cell's north side. After transformation, transports are calculated identically as for the Arakawa-C grid. For the case of an Arakawa-A cell, where T, u and v are placed in the middle of the grid cell, we implement a similar method as for Arakawa-B and move the u and v components onto the cell faces. Note that while volume transports are calculated without any further use of interpolation, for heat and salinity transports, the scalar quantities of T and S have to be interpolated to the faces of the tracer grid cells. This is done using linear interpolation (similar to Madec, 2016, Sect. 12.3.1).

2.1.2 Vector Projection Method (VPM)

The second method uses simple vector projection algorithms to obtain the share of the u and v components that passes orthogonally through the strait. Fig. 2b shows a schematic of the VPM. For every grid cell touching the strait we calculate direction vectors of the u and v components (blue and green arrows), and normal vectors pointing from the tracer grid cell in the direction of the strait (yellow arrows). Then, using equations 7, the u and v vectors are projected onto the normal vector:

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

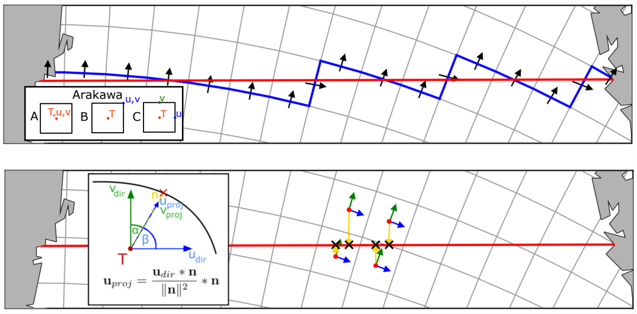

Figure 2. a) Schematic of the Line Integration Method, with the reference line (bold red line), the closed line generated on the native model grid (blue line) and used u and v components (black arrows). b) Schematic of the Vector Projection Method, the direct u and v vectors (blue and green) are projected onto the normal vectors (yellow) to find the portions of u/v that pass orthogonally through the strait (red line).

2

$$\boldsymbol{u}_{proj} = \frac{\boldsymbol{u}_{dir} \cdot \boldsymbol{n}}{\|\boldsymbol{n}\|^2} \cdot \boldsymbol{n}; \qquad \boldsymbol{v}_{proj} = \frac{\boldsymbol{v}_{dir} \cdot \boldsymbol{n}}{\|\boldsymbol{n}\|^2} \cdot \boldsymbol{n} \tag{7}$$

where $\boldsymbol{u}_{dir}$ and $\boldsymbol{v}_{dir}$ represent the direction vectors of the u and v components, $\boldsymbol{n}$ the normal vector between tracer grid cell and strait and $\boldsymbol{u}_{proj}$ and $\boldsymbol{v}_{proj}$ the projection vectors of $\boldsymbol{u}_{dir}$ and $\boldsymbol{v}_{dir}$, passing orthogonally through the strait. The closer the angle between direction vector and normal vector (α and β in Fig. 2b) to 0 (90) or 180 (270), the larger (smaller) the amount that actually passes through the strait. The projection vectors' magnitudes are then used to compute orthogonal transports at all grid cells touching the strait. In the final step, these transports are interpolated bilinearly onto the closest points on the reference line (black crosses in Fig. 2b, called T_{proj} henceforth) and divided by the respective cell thicknesses on the reference line to obtain velocities. This results in velocity cross-sections of the vertical plane which are spaced irregularly along the along-strait distance (x) in accordance with the distribution of T_{proj} points. The interpolation onto evenly distributed points on the section, to e.g. enable the calculation of differences with other models/reanalyses, is initially left to the user and eventually will be included in a future version of StraitFlux. Integrating the cross-sections along the along-strait distance (x) and depth (z) provides net transports, which agree very well with the values obtained by the LM (see Sect. 3). To obtain heat and salinity transports the velocity cross-sections simply have to be multiplied with the T and S cross-sections, which are obtained using bilinear interpolation, before integration.

2.2 Software Implementation

In this section, we describe the implementation of the transport calculation tools into the open-source python package StraitFlux. We provide an overview of the code structure and its usage.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

2.2.1 Preprocessing

The script is designed to work with ocean reanalyses and various CMIP6 models, which may differ in terms of grid orientation, partitioning, coordinate names, units, and dimension names. Therefore, prior to the transport calculations, data preprocessing is conducted to standardize these attributes. This includes ensuring consistency in dimension names (x, y, lev, bnds), coordinate names (lon, lat), units (SI), and the shape of longitude and latitude coordinates (2D). For CMIP6 models, preprocessing is carried out using selected tools from the open-source python package `xMIP` (https://cmip6-preprocessing.readthedocs.io/en/latest/). Ocean reanalyses are treated in a similar fashion by adapting some of the `xMIP` tools. There is no regridding involved.

To integrate transports along cell faces and depths accurately, precise information about the cell extents is required. As horizontal grid metrics are not always provided by CMIP6 models, the script automatically determines the zonal and meridional extents using the `calc_dxdy` function. The vertical extent has to be specified prior to the calculation process. For instance, for CMIP6 the variable "cell thickness" (*thkcello*), which is available through the Earth System Grid Federation (ESGF) website (https://esgf-node.llnl.gov/search/cmip6/) for most models, may be used. Since the cell thickness is not available for all models, the function `calc_dz` enables the calculation of cell thicknesses by supplying the variable "total ocean depth" (*deptho*, also available via ESGF), and the vertical level boundaries (*lev_bnds* for CMIP6, contained in every three dimensional ocean variable). If *deptho* is not available it is also possible to supply it from another model, preferably one with a similar grid, and the variable will automatically be interpolated to the needed grid. However, especially volume transport calculations are very sensitive to the exact ocean bathymetry used by the model. Therefore, if possible, it is advised to supply the exact fields (e.g. *thkcello* for CMIP6) and not necessarily to rely on `calc_dz`.

2.2.2 Index extraction

The determination of section positions for transport calculations is accomplished in the `def_indices` function. Users can specify start and end points, and also points in between if the section is "kinked". The function generates a reference line (ref_{line}) consisting of equally spaced latitude-longitude pairs.

In order to calculate the net transports via line integration, a polyline along the edges of grid boxes has to be generated following the ref_{line} as closely as possible (red line in Fig. 2). The function `check_availability_indices` determines the indices of the u and v components along the closed integration line:

The first point of the integration line is found by selecting the nearest point on the native grid (center of a cell) within a selection window with the size of 2 degrees around the first point on the reference curve - the size of the selection window was chosen to work properly on grids with a resolution of 1 degree and higher, if needed the window size can be adapted to coarser resolutions. The subsequent points on the reference line are then found iteratively using the `select_points` function. The

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

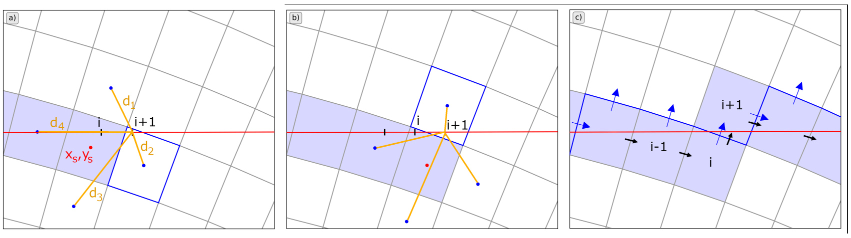

Figure 3. Illustration of the indices selection process using `select_points`. a) and b) Determination of consecutive grid-cells by comparing distances (orange lines) of 4 surrounding grid-cells. c) Specification whether a u or v component should be be taken.

basic principle of `select_points` is outlined in Fig. 3. The starting point at any iteration i is defined as coordinate pair [x_s, y_s]. The next point is found by comparing the distances of the four surrounding points ([x_s, y_{s+1}],[x_{s+1}, y_s],[x_s, y_{s-1}],[x_{s-1}, y_s], marked in blue), to the next point $i+1$ on the ref_{line} (red line). The point closest to the reference curve becomes the next point on the model grid that will be included in the integration line (outlined in blue, Fig. 3a) and the new starting point [x_s, y_s] (Fig. 3b) to find the subsequent point. This is done for all points on the reference line and results in a closed sequence of grid cells along the reference line (filled blue cells). To prevent water from "escaping", the first and last point should be land grid points. Should this not be the case, a warning will be given to the user to recheck the given lat/lon coordinates and the computed transports should be treated with caution. The user is provided with figures of the selected line and the model land-sea mask, which can be used to check the position and length of the desired strait.

To determine whether the u or v component of each grid cell has to be taken, we look at the positioning relative to the previous cell (see Fig. 3c). When coming from left or right to the new grid cell, the v component of the cell is taken. When coming from above/below to the new cell, the u component of the previous/current cell is taken. For instance, to get to cell i in Fig. 3c, we came from left, hence the v component of cell i is taken. In order to get to cell $i+1$ we come from below, therefore the u component of cell $i+1$ is taken.

Depending on the orientation of the reference curve relative to the distorted grid lines, the sign of the u component may vary - this is illustrated in Fig. 4: u is counted positive when cell $i+1$ is below cell i (=coming from above, i.e. from a higher index in y direction) and negative when $i+1$ is above i (=coming from below, i.e. from a lower index in y direction). In some cases also the sign of the v component may change to negative in positive y direction, especially due to the strong bend of dipolar grids in the proximity of Greenland. Fig. 5 shows the change of the v sign for Fram strait on the dipolar grid as used by the POP2 ocean model (Smith et al., 2010) in the CMIP6 model SAM0-UNICON (Park et al., 2019). If the new cell is reached by coming from the left, the v component if positive (lower part of the strait in Fig. 5), and when coming from the right the whole grid is oriented upside-down and v is negative (upper part of the strait in Fig. 5). Straits may be defined from west to east as well as from east to west (also north to south and south to north) in the `def_indices` function and integrated transports are

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

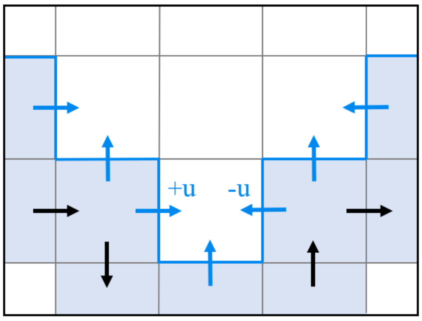

Figure 4. Illustration of the specification of the sign for the u component. Black arrows show the direction of integration along the reference line (blue line). Blue arrows show the direction of the used u and v components. When coming from above (i.e. from a higher index in y) u is counted positive and negative when coming from below.

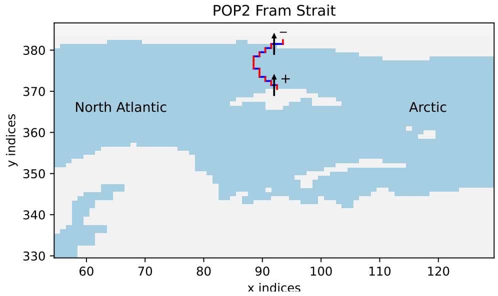

Figure 5. Change in sign of the v component at Fram strait in the CMIP6 model SAM0-UNICON (Park et al., 2019), using a displaced dipolar POP2 ocean model. Blue lines indicate cells where v components are used and red lines where u components are used. In the eastern (lower in y-direction) part of the strait positive v components are tilted from the strait toward the true north and are therefore counted positive, in the upper part of the strait v components are tilted from the strait toward the true south and are therefore counted negative.

defined to be positive pointing to the left of the direction of the strait. Therefore, transports for straits defined from north-west to south-east will be positive towards the north-east and negative towards the south-west.

190 2.2.3 North fold boundary and periodic cells

Due to the periodicity of the Earth, the model domain boundaries pose a further challenge for the transport computations. Concerning the eastern and western boundaries, most ocean models have various options for dealing with periodicity. The most common options are cyclic east-west boundaries and closed boundaries. For the cyclic boundaries the values of the last 1-2 columns are set to the values of the first 1-2 columns - therefore, whatever flows out of the western end of the basin enters the

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

eastern end and vice-versa (Madec, 2016). For the closed boundary conditions solid walls are enforced at all model boundaries and the first and last columns are set to zero. Same is true for the north-south boundary conditions. For our application the displaced dipolar model grids are uncomplicated concerning the north-south conditions. As both poles are placed over land the southern and northern most grid-cells consist of land areas only and no oceanic transports pass the northern boundary - the grids are topologically equivalent to a cylinder (periodic in x but not in y, Smith et al. (2010)). The tripolar grids however require additional consideration along the northern boundary, as the ocean is divided by the line between the two northern grid poles. For instance, the three-polar ORCA grid (used in multiple CMIP6 models) uses a North fold boundary with a T-point pivot (see Sect. 8.2.2 in Madec (2016)), where the upper three grid-cells are duplicated and pivoted around the line connecting the two north poles. For more information on the model specific boundary conditions see e.g., Madec (2016); Griffies et al. (2004); Smith et al. (2010).

These conditions have to be handled with care, as especially the volume transport calculation is very sensitive and can yield useless results when there is a gap in the integration line or if any grid-cells are counted twice.

2.2.4 Line Integration Method (LM)

The actual oceanic transports are calculated through the `transports` function by integrating the velocities at the chosen u and v indices over the cells zonal or meridional extents and their actual vertical extent. See Sect. 2.2.1 for the computation of the needed mesh files.

For heat transports the indexed cells are additionally multiplied by the ocean's potential temperature prior to integration and for salinity transports with the cells salinity. These are defined at grid cell midpoints for Arakawa B and C grids. Therefore, the fields of T and S have to be interpolated to the faces of the grid cells first, which is done using linear interpolation (similar to Madec, 2016, Sect. 12.3.1). Furthermore heat/salinity transports have to be multiplied by specific heat and density (see equation 3). Those are set as constant to 3996 Jkg^{-1}K^{-1} and 1026 kgm^{-3} per default, however it is possible to adapt them to individual needs.

Transports may be calculated for longer time periods at once, where the time period may be set by the 'time_start' and 'time_end' arguments in `transports`. The final result of `transports` gives net integrated transports through the strait with the coordinate time, which are saved per default as netCDF file. The used indices and values through the cell faces can be saved on request.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

2.2.5 Velocity Projections

The LM has good conservation properties but since the faces of the polyline can point into very different directions, it is difficult to plot cross-sections. Therefore a second method for calculating the cross-strait transports at points on the reference line has been developed. The direction of the normal vectors thus changes smoothly and allows the calculation of horizontally and vertically resolved contributions to the total transport through a respective strait.

Similar to LM, again the first step is to find the closest points on the native grids to the reference line. However, herein the closest 4 points for every point on the reference line are found, in order to be able to interpolate the calculated transports onto the reference line in the last step of the calculation.

For the projection of the u and v velocities onto the strait, direct vectors and normal vectors for every grid-cell are determined using the functions `calc_dir_vec` and `calc_normvec`. Direct vectors are assumed to point from one grid-cell to the neighboring ones and are simply calculated by taking the difference between the $u_{x,y}$ and $u_{x+1,y}$ for $\boldsymbol{u}_{dir}$ and the difference between $v_{x,y}$ and $v_{x,y+1}$ for $\boldsymbol{v}_{dir}$.

Normal vectors are calculated using three consecutive cross products. The basic principle is illustrated in Fig. 6. For each point T, u and v, we find the two closest points r1 and r2 on the reference line. Transforming them into Cartesian coordinates, we can take their cross-product and get the vector $\boldsymbol{rn}$, standing orthogonally onto the surface spanned by r1 and r2 (blue surface). The cross-product of $\boldsymbol{rn}$ with T, yields the vector $\boldsymbol{Tr}$ (orthogonal to the green surface). Finally taking the cross-product of $\boldsymbol{rn}$ with $\boldsymbol{Tr}$ and normalizing the resulting vector with the Earth's radius, yields the point T_{proj} on the reference line. The normal vector $\boldsymbol{n}$ for every grid-cell is given as the vector pointing from point T to T_{proj}. The projection points T_{proj} are later used as interpolation points on the reference line.

The function `proj_vec` uses the direct and normal vectors and equations 7 to calculate the projection vectors at every grid-cell. The actual calculation is done using `vel_projection`. Using the norm of the projection vectors, the u and v components of every vertical layer are then scaled according to their position with respect to the reference line of the section and multiplied by the actual cell thicknesses of the cells. The signs of the u and v components are determined by comparing the angles between direct and normal vector - the general idea is outlined in Fig. 7. Then, in a final step, the scaled fluxes for every vertical layer are interpolated onto the reference line and divided by the respective layer thickness at the reference line.

The final result of `vel_projection` gives cross-sections of the velocities passing through the strait, with coordinates time, depth and x - the along strait distance. Exemplary cross-sections are shown in Sect. 3.1. To simplify the integration of the cross-sections to net transports, the horizontal (dx_int) and vertical extents (dz_int) of every point on the reference line are output as well. Therefore, the net volume transport - similar to the end product of the LM - can be calculated by multiplying the section with dx_int and dz_int and summing up over x and depth.

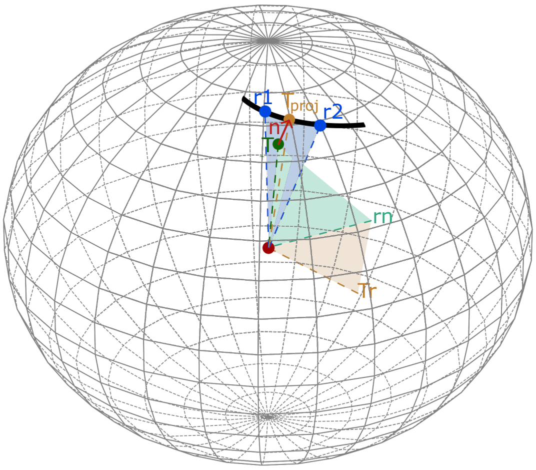

Figure 6. Illustration of the normal vector computation process using three consecutive cross products: r1 and r2 feature the two closest points to T on the reference line, $\boldsymbol{rn}$ is the normal vector on the surface spanned by r1 and r2, $\boldsymbol{Tr}$ is the normal vector of $\boldsymbol{rn}$ with T, T_{proj} is the projection point on the reference line and results by normalizing the cross-product of $\boldsymbol{rn}$ with $\boldsymbol{Tr}$ with the Earth's radius. T and T_{proj} then provide the normal vector $\boldsymbol{n}$.

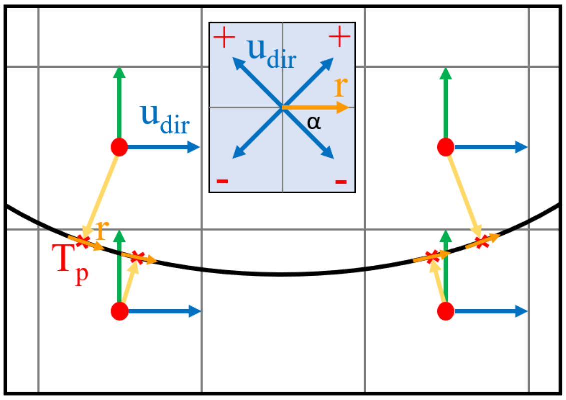

Figure 7. Determination of the u component signs using the angle α between the direct vector of u and the direction vector of the reference line r at the corresponding projection point. When $\alpha < 180$ the u component is counted negative, when $\alpha > 180$ the u component is counted positive. The v signs are calculated similarly using $\boldsymbol{v_{dir}}$.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

Cross-sections of temperature and salinity are simply calculated by interpolating the scalar quantities of T and S onto the strait defined by the T_{proj} points. To obtain the vertical profiles of heat and salinity transports, the T and S cross-sections have to be multiplied with the velocity cross-sections.

The velocity projection method has originally been developed for visualization purposes, however, as is shown in the next section, almost everywhere it provides nearly as accurate estimates of total fluxes through a strait as the LM.

3 Validation

In this section we will assess the robustness of the tools and their accuracy with multiple approaches. Firstly, the results of our computations are compared with naive calculation of the fluxes from interpolated velocity fields. Secondly, we specify simple u,v and T fields where the transports can be calculated analytically. These fields are then transformed exactly to the respective ocean model grids using the analytical mapping functions. The transports are then calculated using our LM and compared to the analytic solutions. Thirdly, we show the consistency between VPM and LM and then we check the correspondence between area integrated divergence fields and the transports through the array boundary. Lastly, we compare our results to results taken from an independent study where transports through Fram Strait were calculated by picking indices and signs for all grid points by hand.

For verification we use harmonic functions to specify simple two dimensional u,v and T fields where the transports can be calculated analytically:

$$u(\lambda,\varphi) = 0; \qquad v(\lambda,\varphi) = v_r(\varphi) + v_0 \cos\varphi \sin k\lambda; \qquad T = T_r + T_0 \cos\varphi \sin(k(\lambda+\psi)) \tag{8}$$

with longitude λ, the latitude φ, wavenumber k and phase shift ψ given in radiant. With the Earth radius a, we get for the transport:

$$F(\varphi) = a\cos\varphi \int_0^{2\pi} T v \, d\lambda \quad = v_r(\varphi) T_r \cos\varphi + \frac{v_0 T_0 \cos^3\varphi}{2} \cos\psi \tag{9}$$

The second term on the right-hand-side of (9) vanishes for phase shift $\pi/2$. The defined v and T fields are transformed to different ocean model grids (CMIP6) - the four used modelling grids are shown in Fig. 8 top and were chosen to be as different as possible in terms of horizontal resolution, number and location of poles, strength and extent of the distortion and used Arakawa partition. Transports are then calculated for full parallels at different latitudes (φ = -70°N to 85°N) and for different wavenumbers (k = 1 to 100) using our LM, which leads to the generation of nontrivial polylines depending on the curvature of the respective grid. Solutions are then compared to the analytic solutions of the transport integral (eq. 9). Differences remain low (mostly below $\pm 1\%$) for all four grid types over all assessed latitudes and wavenumbers as shown in Fig. 8 bottom. The biggest errors occur for the lower resolution grids at higher latitudes and higher wavenumbers. This is most likely caused by the

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

coarse resolution and discretization of the models, which are not able to resolve the smaller generated waves, and less by the curvature of the grid alone, as the higher resolution model features very low errors up to k=100. At latitudes with regular grid lines errors due to discretization are deemed to be small. We calculated transports using the defined spherical harmonics for the tripolar CanESM5 grid at 20°N for various horizontal resolutions and found a difference of just about 0.015% between a 1° and a 10° resolution. Other small differences are mainly caused due to inaccuracies in the latitude selection. While analytic solutions are calculated at full latitudes, the position of the polylines may be shifted north or south due to the grids resolution. For instance, the latitudinal shift of the 20°N line in the CanESM5 model leads to an error of 0.2%, explaining practically the total recorded error. Further, differences in the Earth radius (we assume a = 6371km) may lead to minor discrepancies.

The grey areas in the lower left panels of Fig. 8 indicate missing values and are due to the nature of the two dipolar grids, as they skip areas over the Greenlandic ice cap around the artificial North Poles. This is problematic for this application, as we define complete parallels for our analytic solutions. However, it does not affect actual transport calculations as this complication does not occur over oceanic areas.

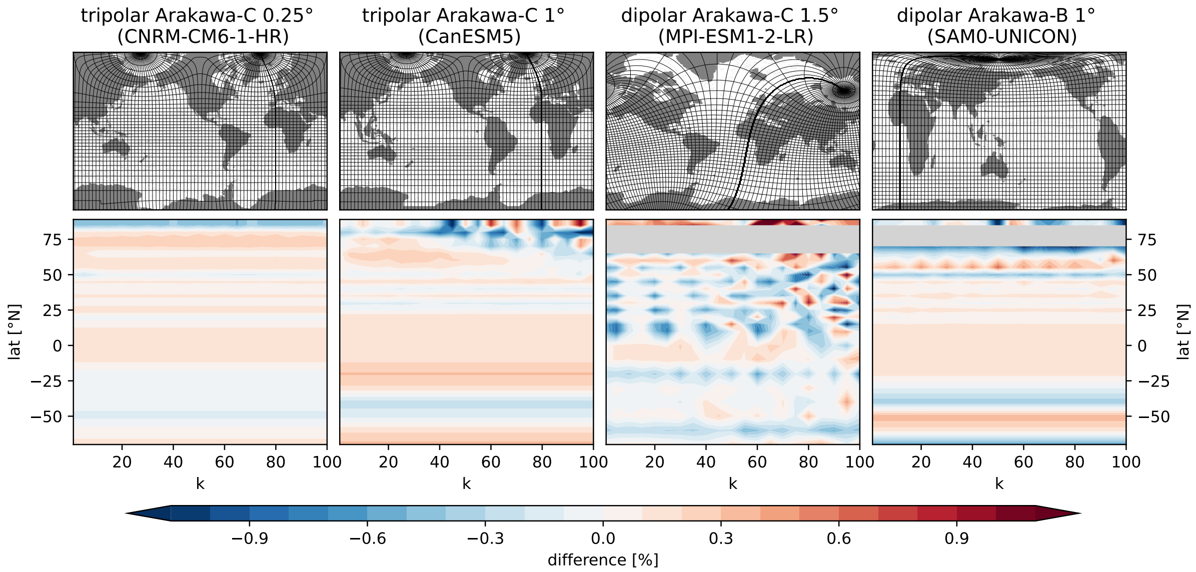

Figure 8. Top: grids used for the generation of polylines for the transport calculation via LM. The gridlines show the position of the regular gridlines on the distorted modelling grids (resolution of 5°). Bottom: differences between the LM and analytic solutions of transports using spherical harmonics as fields. Grey areas in the bottom plots indicate areas where the calculation was not performed due to the absence of grid points in the dipolar grids over parts of Greenland.

In summary, results from the LM correspond very well with analytical results.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

Fig. 9 compares transports calculated from interpolated u and v values on a regular grid with those derived from u and v on the native grid. We compare both, bilinear and conservative interpolation as defined in ESMFgrid, and calculate volume and heat transports through the sum of Arctic main gateways (Fram Strait, Barents Sea Opening, Davis Strait and Bering Strait) for the CanESM5 model. Volume transports for the net Arctic should be close to zero (about -0.2 Sv according to observation-based estimates, Winkelbauer et al., 2022). Transports derived from our two methods fit well to those estimates, however transports derived from the interpolated fields feature severe errors. For heat transports errors are considerably smaller in relative terms (albeit still significant) when considering the whole Arctic, however errors for single straits are still major (not shown). It is thus clear that the use of interpolated vector components is inappropriate for all kinds of transport calculations.

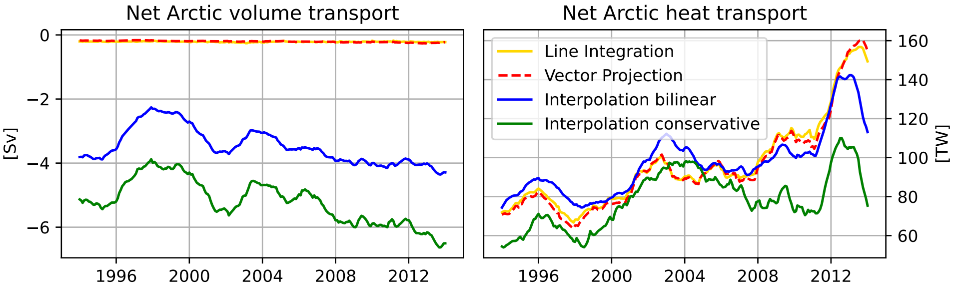

Figure 9. Display of the interpolation error for volume and heat transports for the Net Arctic (Fram, Davis, Bering Strait and Barents Sea Opening) from the CanESM5 model (1° resolution).

While we have not compared the cross-section method with the analytical solution as we did for the LM, we show the credibility of the VPM by comparing volume transports obtained through the LM and the VPM (Fig. 10). We choose volume transports as those are more sensitive than heat and salinity transports and therefore should better indicate potential shortcomings. Ideally, both methods should provide the same results, however due to differences in the calculation process small differences are expected. We choose a strait in the Arctic region - Fram strait - in order to come close to the strongest distortion of the curvlinear grids and show volume transports for three models that use different grid types and Arakawa partitions. The CMCC-CM2-SR5 model uses a tripolar grid with an Arakawa-C partition, IPSL-CM6A-LR (Boucher et al., 2020) uses a tripolar grid with an Arakawa-B partition and SAM0-UNICON uses a displaced dipolar grid with an Arakawa-B partition. Depending on the model, transports obtained through the different calculation methods match within a couple of percent of their total value. We consider this a very good result, given that this method was more designed for plotting purposes than for maximum accuracy of the integrated result.

Another test to validate our transport calculation tools is the comparison of the transports across a whole latitudinal circle to the divergence of transports north of that latitude (i.e., validation of the sentence of Gauss). This is done for the ORAS5 ocean reanalysis (Zuo et al., 2019) and shown in Fig. 11. While the VPN differs from the values obtained by the LM and the divergence integral, the differences are still very small compared to those found in Fig. 9. Those differences may be caused by

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

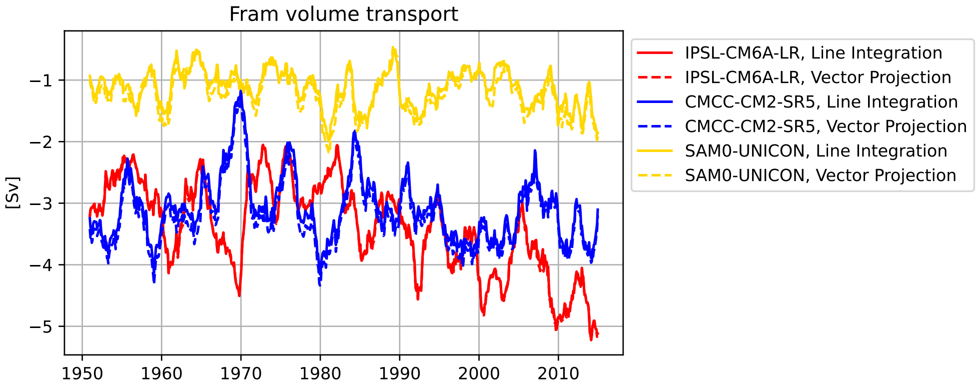

Figure 10. Comparison of volume transports obtained through the LM (solid) and through integration of cross-sections obtained through the VPM (dashed). The selected models use different grid types (see text) all with a horizontal resolution of about 1°.

the increasing difference in integration area between the two methods with stronger grid curvature further north or also by an inaccuracy in the treatment of the North Fold boundary points. This needs to be further investigated and may be resolved in a later version of the software.

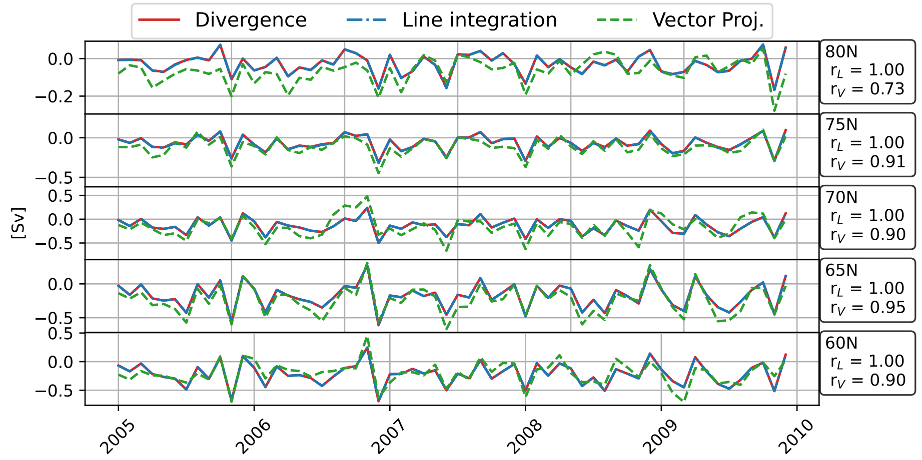

Figure 11. Integrated volume fluxes across different circles of latitude derived from ORAS5. Transports computed using the LM and VPM as well as through the integration of the divergence of transports north of the section in question.

Lastly, we compared our methods to transports obtained by Heuzé et al. (2023) available via PANGAEA (Zanowski et al., 2023), who calculate transports of salinity, heat and volume through Fram strait for various CMIP6 models by choosing the coordinates for each model by hand. Fig. 12 shows the comparison of our transports to those obtained by Heuzé et al. (2023)

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

for 10 selected models. For most models the results match within an expected range of uncertainty - differences may arise from differences in the exact positioning of the straits and differences in the definitions of ρ and c_p.

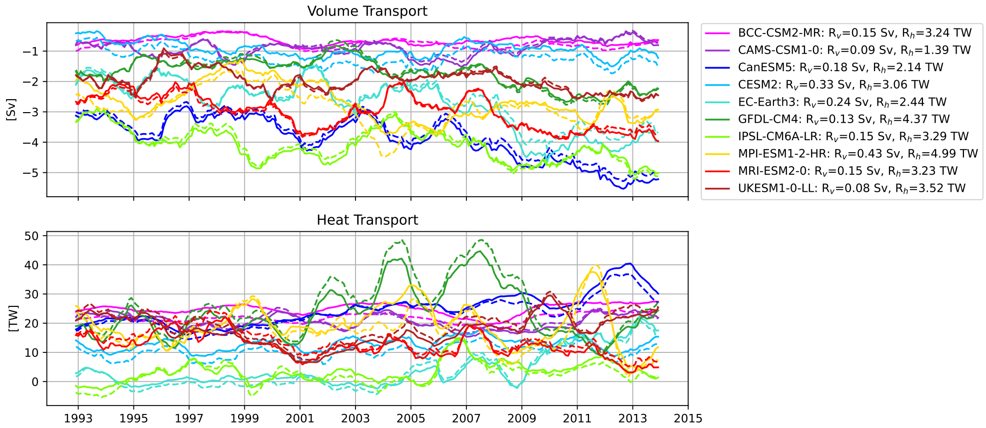

Figure 12. Volume and heat transport time series for ten selected CMIP6 models from Heuzé et al. (2023) (solid) and our LM estimates (dashed). RMS-differences between Heuze's estimates and ours for volume (R_v) and heat (R_h) transports over the given time range are added in the line labels.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

3.1 Application examples

To illustrate the abilities of StraitFlux we present some sample results and refer to studies where the tools have already been successfully used.

Results from the LM and VPM have already been shown in the validation section. Additionally to the net integrated transports the VPM also provides cross-sections of the vertical plane. Fig. 13 shows exemplary cross-sections of currents, temperature and salinity for the Greenland-Scotland Ridge (GSR) for two CMIP6 models with different horizontal resolutions - note the big difference in bottom topography and also the depiction of individual currents between the models.

Mayer et al. (2023b) use StraitFlux to compare oceanic transports across the GSR from ocean reanalyses against largely independent observations. They use the results from StraitFlux to partition the water masses into Atlantic, Overflow and Polar water, enabling a more in-depth analysis. They find that ocean reanalyses underestimate the observed Atlantic Water inflow by up to 15%, causing a low bias in oceanic heat transports (OHT) of 5%–22%. Further, they attribute a pronounced anomaly in OHT during the two-year period around 2018 to a reduction in Atlantic Water inflow through the Faroe–Shetland branch in combination with anomalously cool temperatures of Atlantic Water arriving at the GSR due to a recent strengthening of the North Atlantic subpolar gyre. Winkelbauer et al. (2023) use StraitFlux to calculate net transports of volume and heat passing into and out of the Arctic through Fram Strait, Davis Strait, Bering Strait and the BSO. They assess the transports' seasonal cycles and find clear correlations between oceanic transports and the Arctic's mean state. Fritz et al. (2023) use StraitFlux to assess transports in the ITF region and find reasonable agreements between reanalysis-based transports and observations in terms of means, seasonal cycles, and variability. Furthermore, transports have been calculated at the RAPID and OSNAP sections.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

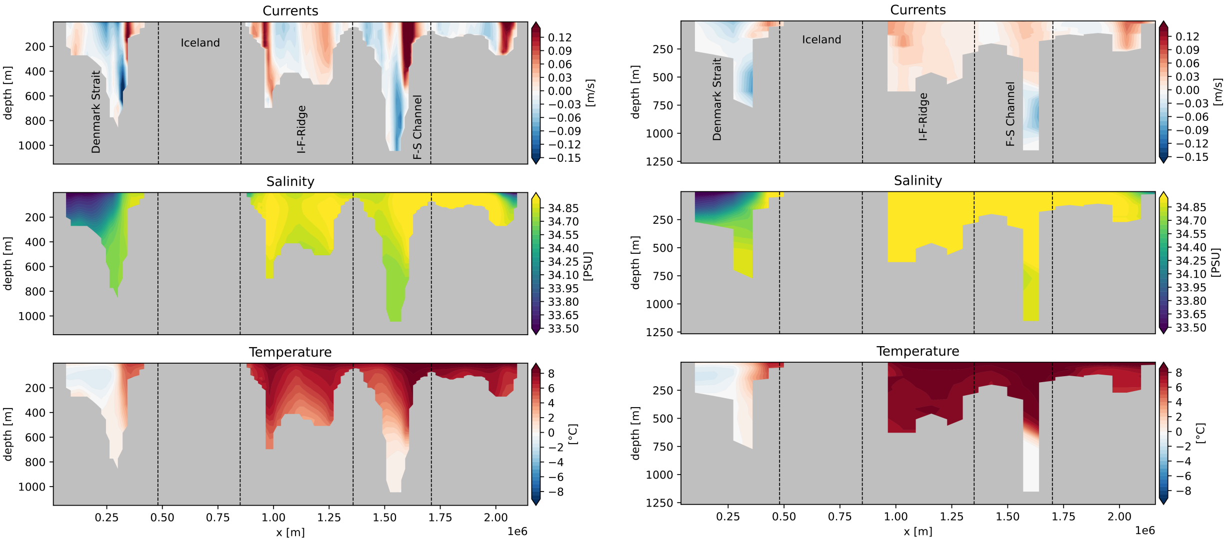

Figure 13. Sample Cross-sections of currents (top), salinity (middle) and temperature (bottom) for the Greenland Scotland Ridge for the CNRM-CM6-1-HR (0.25 deg horizontal resolution; Voldoire et al., 2019) and CNRM-CM6-1 (1 deg horizontal resolution; Voldoire et al., 2019) CMIP6 models.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

3.2 Computational performance

The transport calculations usually need to involve only a small fraction of the 3D field values stored in the CMIP or reanalysis archives. As the current archives do not support extraction of subareas, the global fields need to be downloaded and consequently a fair amount of the total computational time is spent in reading and preprocessing the files.

In order to calculate e.g. the temperature flux, the following fields from a typical 1/4 degree CMIP6 model are needed:

Field	approx. Dimensions	approx. Size [MB]
Sea Water X Velocity u	1400x1000x75	300
Sea Water Y Velocity v	1400x1000x75	300
Sea Water Potential Temperature T	1400x1000x75	300
Ocean Model Cell Thickness thkcello	1400x1000x75	300

Table 1. Approximate dimensions and sizes of variables for typical 1/4 degree models needed for the transport calculations of one time step.

To detect the indices of the section, calculate horizontal meshes and determine the Arakawa partition the software reads one vertical layer of the data files. Once the section is known, the software chooses the subregion, so that not the whole files need to be read. While reading the files takes up the majority of the calculation time, the calculation itself is performed relatively fast. Approximate times for the major calculation steps for a Xeon Gold 6148 CPU for a 1 degree model (CanESM5) and a 1/4 degree model (EC-Earth3P-HR) at Fram Strait are given in the table below. The calculation of the mesh files and detection of the Arakawa partition has to be performed only once per model, and the calculation of indices and parameters for the VPM (normal and direct vectors, signs of velocity components etc.) once per model and strait, when the functions' parameter `saving` is set to `True` (default). This speeds up subsequent calculations, e.g. for different months or straits, considerably.

With monthly time resolution, for 1/4 degree models it is possible to calculate transports directly for multiple years (e.g. the calculation of the 65 year period takes about 60 s), for higher resolution models we advise to loop the calculation (e.g. over 12 months) to avoid high memory consumption. For faster performance calculations may of course be done in parallel. Also the flux calculation for other ensemble members can be done in parallel as well.

3.3 Availability

StraitFlux is available as open source python package at github and zenodo and can be installed from pypi. The gitlab repository also contains an example script and some example datasets as well as a requirements file, to simplify the installation and usage of StraitFlux.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

horizontal resolution [°]	1°x1°, 75 l, 12 m	0.25°x0.25°, 75 l, 12 m
read files for mesh and index calculation **	1.5 s	3.5 s
calculate indices **	0.9 s	1.4 s
determine Arakawa partition *	3 ms	3 ms
read files (subselected)	2.2 s	3.5 s
calculate dz at cell faces **	25 ms	0.15 s
Line Integration Method (LM)		
calculate mesh files *	10 s	44 s
calculate transports	50 ms	80 ms
Total	15 s / 2.3 s	55 s / 3.8 s
Vector Projection Method (VPM)		
calculate projection vectors and constants **	8.2 s	31 s
calculate transports	0.2 s	3.3 s
regrid to section	0.1 s	0.2 s
Total	12.5 s / 2.5 s	43 s / 7.5 s

Table 2. Approximate calculation times for 12 months of monthly data for two exemplary models with 1 degree and 1/4 degree resolution and 75 vertical layers at Fram Strait. Functions are divided into "used by both" (top), LM (middle) and VPM (bottom). The asterix * indicates functions that only have to be calculated once per model and ** functions that need to be calculated once per model and strait. Total times are given for the first calculation performed (left values) and every consecutive calculation (right values), when the functions' parameter 'saving' is set to 'True' (default).

StraitFlux is a free software and can be redistributed and/or modified under the terms of the GNU General Public License version 3 as published by the Free Software Foundation.

4 Conclusions

In this study, we have introduced StraitFlux, an open-source Python package designed to facilitate the calculation and analysis of oceanic transports through arbitrary oceanic straits and sections. We give a comprehensive overview of StraitFlux, including its underlying principles, software implementation, validation, and application examples. StraitFlux facilitates scientific studies to validate models and gain valuable insights into ocean circulation, heat transports, and water mass exchanges, making it a useful tool for climate scientists, oceanographers and modelers.

StraitFlux works on various curvilinear ocean modeling grids and is written so flexible that it is expected to work for future versions (e.g. CMIP7) as well. Unstructured grids are not included in this release. However, the methods have already been successfully adapted and tested for the FESOM2 ocean model (Danilov et al., 2017), the successor of FESOM which is i.a. used in the AWI Climate Model (AWI-CM) in CMIP6, and are planned to be included in future versions of StraitFlux as well.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

The tools include two methods for calculating oceanic transports: the Line Integration Method (LM) and the Vector Projection Method (VPM). The LM creates a closed polyline along grid cell faces to compute net integrated transports, while the second method employs vector projection algorithms to estimate the share of u and v components passing orthogonally through the strait and generates cross-sections of velocities, temperatures and salinities in the vertical plane.

Both methods have been thoroughly validated and produce reliable results across various ocean models and grids. Our validation efforts have demonstrated that StraitFlux consistently matches analytical solutions, even in complex grid configurations and regions with strong distortion. Both methods deliver net transports that match within a couple % of their total value, even at the most distorted sections. The tool's accuracy is further affirmed by comparisons with the divergence of transports and independent transport calculations. One problem remains at the northern most latitudes for the VPM, which we hope to resolve soon.

The applications of StraitFlux extend to a wide range of research areas. Researchers can use the package to analyze seasonal cycles, mean states, and variability in oceanic transports. Furthermore, the ability to generate cross-sections of currents, temperature, and salinity provides a detailed view of the ocean's vertical structure and flow patterns.

In summary, the user-friendly implementation and broad applicability make it a valuable tool for studying the Earth's climate system and its dynamics. The simplified comparison to observational data highlights its suitability for model validation and assessment. We hope that StraitFlux empowers researchers to explore and understand oceanic transports more thoroughly, given their importance in the climate system and changes therein.

Code availability. The Python implementation of StraitFlux is available at https://github.com/susannawinkelbauer/StraitFlux (last access: 30 November 2023) and can be installed from pypi. The github repository additionally contains the notebook 'Examples.ipynb' with some easy examples to get started with the transport calculations. Data files used in the notebook may be downloaded via ESGF (https://esgf-node.llnl.gov/search/cmip6/). Version 1.0.1 of StraitFlux, which is described and used in the paper, is long-term archived at zenodo (Winkelbauer, 2023).

Data availability. CMIP6 data used in the validation section of this paper may be downlaoded via ESGF (https://esgf-node.llnl.gov/search/cmip6/). Reanalyses data is available via the Copernicus Marine Service (https://marine.copernicus.eu/).

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

Author contributions. SW, MM and LH conceptualized the study. SW developed the Python implementation of StraitFlux with the help of MM and LH. SW performed the data analysis, including the production of the figures in the paper, and prepared the manuscript. All authors wrote, edited and reviewed the manuscript and agreed to the publication of the present version of the manuscript.

Competing interests. The authors declare that they have no competing interests.

Acknowledgements. Susanna Winkelbauer and Michael Mayer were supported by the Austrian Science Fund project P33177 and the Copernicus Marine Service contract 21003-COP-GLORAN Lot 7. The authors sincerely thank Céline Heuzé from the University of Gothenburg for the provision of manually calculated transport data to use as validation and helpful discussions on the transport calculation methodology. We acknowledge work by Vanessa Seitner who developed an earlier version of Straitflux.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

References

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, Lionel, E., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS002010, 2020.

Caesar, L., Rahmstorf, S., Robinson, A., Feulner, G., and Saba, V.: Observed fingerprint of a weakening Atlantic Ocean overturning circulation, Nature, 556, 191–196, https://doi.org/10.1038/s41586-018-0006-5, 2018.

Danilov, S., Sidorenko, D., Wang, Q., and Jung, T.: The Finite-volumE Sea ice–Ocean Model (FESOM2), Geoscientific Model Development, 10, 765–789, https://doi.org/10.5194/gmd-10-765-2017, 2017.

Ditlevsen, P. and Ditlevsen, S.: Warning of a forthcoming collapse of the Atlantic meridional overturning circulation, Nature Communications, 14, https://doi.org/10.1038/s41467-023-39810-w, 2023.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fritz, M., Mayer, M., Haimberger, L., and Winkelbauer, S.: Assessment of Indonesian Throughflow transports from ocean reanalyses with mooring-based observations, Ocean Science, 19, 1203–1223, https://doi.org/10.5194/os-19-1203-2023, 2023.

Griffies, S., Harrison, M., Pacanowski, R., Rosati, A., Liang, Z., Schmidt, M., Simmons, H., and Slater, R.: A Technical Guide to MOM4, 2004.

Heuzé, C., Zanowski, H., Karam, S., and Muilwijk, M.: The Deep Arctic Ocean and Fram Strait in CMIP6 Models, Journal of Climate, pp. 1 – 68, https://doi.org/10.1175/JCLI-D-22-0194.1, 2023.

Jackson, L. C., Kahana, R., Graham, T., Ringer, M. A., Woollings, T., Mecking, J. V., and Wood, R. A.: Global and European climate impacts of a slowdown of the AMOC in a high resolution GCM, Climate Dynamics, 45, https://doi.org/10.1007/s00382-015-2540-2, 2015.

Liu, W. and Fedorov, A.: Interaction between Arctic sea ice and the Atlantic meridional overturning circulation in a warming climate, Climate Dynamics, 58, https://doi.org/10.1007/s00382-021-05993-5, 2022.

Lozier, M. S., Bacon, S., Bower, A. S., Cunningham, S. A., de Jong, M. F., de Steur, L., deYoung, B., Fischer, J., Gary, S. F., Greenan, B. J. W., Heimbach, P., Holliday, N. P., Houpert, L., Inall, M. E., Johns, W. E., Johnson, H. L., Karstensen, J., Li, F., Lin, X., Mackay, N., Marshall, D. P., Mercier, H., Myers, P. G., Pickart, R. S., Pillar, H. R., Straneo, F., Thierry, V., Weller, R. A., Williams, R. G., Wilson, C., Yang, J., Zhao, J., and Zika, J. D.: Overturning in the Subpolar North Atlantic Program: A New International Ocean Observing System, Bulletin of the American Meteorological Society, 98, 737 – 752, https://doi.org/https://doi.org/10.1175/BAMS-D-16-0057.1, 2017.

Madec, G.: Nemo Ocean Engine version 3.6, Note du Pole de modélisation, Institut Pierre-Simon Laplace (IPSL), 27, 2016.

Madec, G. and Imbard, M.: A global ocean mesh to overcome the North Pole singularity, Climate Dynamics, 12, 381–388, https://doi.org/10.1007/BF00211684, 1996.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

Mahajan, S., Zhang, R., and Delworth, T. L.: Impact of the Atlantic Meridional Overturning Circulation (AMOC) on Arctic Surface Air Temperature and Sea Ice Variability, Journal of Climate, 24, 6573 – 6581, https://doi.org/10.1175/2011JCLI4002.1, 2011.

Mayer, J., Haimberger, L., and Mayer, M.: A quantitative assessment of air–sea heat flux trends from ERA5 since 1950 in the North Atlantic basin, Earth System Dynamics, 14, 1085–1105, https://doi.org/10.5194/esd-14-1085-2023, 2023a.

Mayer, M., Tsubouchi, T., Winkelbauer, S., Larsen, K. M., Berx, B., Macrander, A., Iovino, D., Jónsson, S., and Renshaw, R.: Recent variations in oceanic transports across the Greenland-Scotland-Ridge, t.b.a., 2023b.

Muilwijk, M., Smedsrud, L. H., Ilicak, M., and Drange, H.: Atlantic Water Heat Transport Variability in the 20th Century Arctic Ocean From a Global Ocean Model and Observations, Journal of Geophysical Research: Oceans, 123, 8159–8179, https://doi.org/10.1029/2018JC014327, 2018.

Murray, R. J.: Explicit Generation of Orthogonal Grids for Ocean Models, Journal of Computational Physics, 126, 251–273, https://doi.org/doi.org/10.1006/jcph.1996.0136, 1996.

Park, S., Shin, J., Kim, S., Oh, E., and Kim, Y.: Global Climate Simulated by the Seoul National University Atmosphere Model Version 0 with a Unified Convection Scheme (SAM0-UNICON), Journal of Climate, 32, 2917 – 2949, https://doi.org/10.1175/JCLI-D-18-0796.1, 2019.

Rahmstorf, S., Box, J. E., Feulner, G., Mann, M. E., Robinson, A., Rutherford, S., and Schaffernicht, E. J.: Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation, Nature Climate Change, 5, 475–480, https://doi.org/10.1038/nclimate2554, 2015.

Rayner, D., Hirschi, J. J.-M., Kanzow, T., Johns, W. E., Wright, P. G., Frajka-Williams, E., Bryden, H. L., Meinen, C. S., Baringer, M. O., Marotzke, J., Beal, L. M., and Cunningham, S. A.: Monitoring the Atlantic meridional overturning circulation, Deep Sea Research Part II: Topical Studies in Oceanography, 58, 1744–1753, https://doi.org/10.1016/j.dsr2.2010.10.056, climate and the Atlantic Meridional Overturning Circulation, 2011.

Schauer, U. and Losch, M.: "Freshwater" in the Ocean is Not a Useful Parameter in Climate Research, Journal of Physical Oceanography, 49, 2309 – 2321, https://doi.org/10.1175/JPO-D-19-0102.1, 2019.

Smith, R., Jones, P., Briegleb, B., Bryan, F., Danabasoglu, G., Dennis, J., Dukowicz, J., Eden, C., Fox-Kemper, B., Gent, P., hecht, M., Jayne, S., Jochum, M., Large, W., Lindsay, K., Maltrud, M., Norton, N., Peacock, S., Vertenstein, M., and Yeager, S.: The Parallel Ocean Program (POP) Reference Manual, 2010.

Tsubouchi, T., Bacon, S., Naveira Garabato, A. C., Aksenov, Y., Laxon, S. W., Fahrbach, E., Beszczynska-Möller, A., Hansen, E., Lee, C. M., and Ingvaldsen, R. B.: The Arctic Ocean in summer: A quasi-synoptic inverse estimate of boundary fluxes and water mass transformation, Journal of Geophysical Research: Oceans, 117, https://doi.org/doi.org/10.1029/2011JC007174, 2012.

Tsubouchi, T., Bacon, S., Aksenov, Y., Garabato, A. C. N., Beszczynska-Möller, A., Hansen, E., de Steur, L., Curry, B., and Lee, C. M.: The Arctic Ocean Seasonal Cycles of Heat and Freshwater Fluxes: Observation-Based Inverse Estimates, Journal of Physical Oceanography, 48, 2029 – 2055, https://doi.org/10.1175/JPO-D-17-0239.1, 2018.

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., Colin, J., Guérémy, J.-F., Michou, M., Moine, M.-P., Nabat, P., Roehrig, R., Salas y Mélia, D., Séférian, R., Valcke, S., Beau, I., Belamari, S., Berthet, S., Cassou, C., Cattiaux, J., Deshayes, J., Douville, H., Ethé, C., Franchistéguy, L., Geoffroy, O., Lévy, C., Madec, G., Meurdesoif, Y., Msadek, R., Ribes, A., Sanchez-Gomez, E., Terray, L., and Waldman, R.: Evaluation of CMIP6 DECK Experiments With CNRM-CM6-1, Journal of Advances in Modeling Earth Systems, 11, 2177–2213, https://doi.org/10.1029/2019MS001683, 2019.

Winkelbauer, S.: StraitFlux, https://doi.org/10.5281/zenodo.10053554, 2023.

https://doi.org/10.5194/egusphere-2023-2883
Preprint. Discussion started: 2 January 2024

Winkelbauer, S., Mayer, M., Seitner, V., Zsoter, E., Zuo, H., and Haimberger, L.: Diagnostic evaluation of river discharge into the Arctic Ocean and its impact on oceanic volume transports, Hydrology and Earth System Sciences, 26, 279–304, https://doi.org/10.5194/hess-26-279-2022, 2022.

490 Winkelbauer, S., Mayer, M., and Haimberger, L.: Validation of key Arctic energy and water budget components in CMIP6, PREPRINT (Version 1), Research Square, https://doi.org/10.21203/rs.3.rs-3245244/v1, 2023.

Zanowski, H., Heuzé, C., Muilwijk, M., and Karam, S.: Historical (1985-2015) Fram Strait volume, heat, and salt transports in CMIP6 models, https://doi.org/10.1594/PANGAEA.949182, 2023.

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., and Mayer, M.: The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: a description of the system and assessment, Ocean Science, 15, 779–808, https://doi.org/10.5194/os-15-779-2019, 2019.

4 Conclusions

4.1 Summary and Discussion

Three main research goals were defined in the introduction of the present thesis: (i) create up-to-date estimates of the Arctic water budget using the latest observation-based data, (ii) validate historical CMIP6 model simulations of the Arctic energy and water budgets and (iii) develop new tools for the precise calculation of oceanic transports on various modelling grids.

For the first goal of this thesis the Arctic non-steric water volume budget as well as its closure are assessed and observation-based, up-to-date estimates of the Arctic water budget components are generated. In a first step runoff into the Arctic Ocean is evaluated from different datasets as well as available observations. Runoff forms a key water budget component, especially so in the Arctic Ocean where the volume of river water entering the ocean is especially high as the Arctic Ocean contains about 1% of the global oceanic water, but receives about 11% of the global river discharge (Shiklomanov and Shiklomanov, 2003). However, reliable quantitative estimates of pan-Arctic river runoff estimates are hampered and oftentimes associated with rather large uncertainties as only about 60-80% of the Arctic drainage area are monitored (Shiklomanov et al, 2002). There are different approaches how to deal with the unmonitered area, while some ignore the ungauged areas completely, others rely on model simulations. A very common approach is the method of hydrological analogy, were the available runoff data is expanded to unmonitored regions with similar hydrological, climatic and land cover conditions (Shiklomanov et al, 2021). Serreze et al (2006) for instance, assume that runoff from the unmonitored area is identical to runoff from the gauged area and therefore estimate pan-Arctic river discharge simply by multiplying the gauged runoff with the total drainage area (monitored + unmonitored). While there is a good consistency for pan-Arctic long-term mean runoff estimated with different methods (Shiklomanov et al, 2021), we find seasonal differences, with the method of hydrological analogy tending to underestimate the high-flow summer peaks. Therefore, we combine gauge observations with monthly correction factors calculated from the Global Flood Awareness System (GloFAS) v3.1 (Zsótér et al, 2021), the reanalyses that we find to be most reliable when assessing individual catchment areas, to obtain complete estimates of pan-Arctic river discharge. We find an long-term average river discharge of 4031 ± 203 km^{3}yr^{-1} for the total pan-Arctic area with an upward trend of 2 ± 1% per decade. Adapting our discharge estimates to different catchment areas we find good agreement with past studies (e.g., Shiklomanov et al, 2021). The assessed reanalyses, however, all feature major biases and unrealistic negative runoff trends. Biases were found to be connected to discrepancies in the used data assimilation systems (snow assimilation in ERA5, see also Zsótér et al, 2020), problems with the hydrological LISFLOOD model (GloFAS 2.1) or the propagation of negative trends from the used forcing field components P-ET (ERA5-Land). The latter might possibly be improved when replacing the short-term forecasts of P-ET by VIWVD as forcing, which is computed from analyzed fields and features similar trends to discharge observations.

In a next step, the observation-based discharge estimate is combined with atmospheric freshwater input (P-ET) from atmospheric reanalyses, oceanic lateral fluxes from ocean reanalyses and storage tendencies from the GRACE satellite product in the non-steric formulation of the Arctic Ocean's volume budget equation. On an annual scale this yields only a small imbalance of less than 3%. To obtain a closed budget we use a variational adjustment scheme on annual and monthly values. The adjustment is considered as reliable if budget closure is achieved while remaining in the respective terms' uncertainty estimate. On an annual scale adjustments are considerably small and yield reliable estimates for all budget terms. However, to close the budget on a monthly scale stronger adjustments are needed. This causes some of the adapted fluxes to fall out of the respective a priori uncertainty estimates, which suggests that the uncertainties are underestimated in the first place as possible systematic errors in the data sets are not considered in the uncertainty estimation. Possible reasons for inconsistencies might involve the general underestimation of the seasonal river discharge peak by all reanalyses, as well as the usage of climatological river discharge data in ocean reanalyses. Efforts are being made in integrating time-varying river discharge data into ocean reanalyses, as done e.g., in the BRONCO (Benefits of dynamically modelled River discharge input for OceaN and COupled atmosphere-land-ocean systems) project (Zuo, 2020). This may further reduce inconsistencies between the oceanic volume fluxes from ocean reanalyses and runoff estimates. To evaluate biases within ocean reanalyses ideally oceanographic data would be used. The Arctic Ocean's unique geography, with water entering/leaving the Arctic only through a handful of oceanic gateways, allows for measurements of in- and outflows with considerable but manageable efforts. Mooring arrays on the various straits exist since 1997 for Fram Strait and the Barents Sea Opening, and since 2004 for Davis and Bering Strait and are maintained by six research institutes located worldwide (Tsubouchi, 2014). With careful post-processing it is possible to estimate the overall energy influx into the Arctic through the gateways. For instance, Tsubouchi et al (2018) use a mass-consistent framework on the arrays of moored instrument data and estimate net transports through the Arctic's main gateways for the period 10/2004 to 05/2010. However, using their data, our results show errors in the transports' seasonality as the velocity fields are not measured accurately enough to resolve barotropic wave signals. This again confirms the challenges and limitations induced due to generally limited temporal and spatial coverage of oceanographic data. Enhancing the density of the monitoring network, both spatially and temporally, might help to mitigate this issue and also help in constraining ocean reanalyses.

Nevertheless, past Arctic water budget estimates (e.g., Serreze et al, 2006), as well as uncertainty estimates for the individual budget terms have been refined using observations and the most recent reanalyses data sets and by applying a variational optimization scheme. On an annual scale only small imbalance of less than 3% are found, which is considerably smaller than the residual found by Serreze et al (2006) (about 8%) for the oceanic freshwater budget. Those improvements demonstrate the progress made in observations and data assimilation techniques as well as diagnostic methods and build a solid foundation for the second part of this study, the validation of climate model simulations.

Before discussing the second goal of this study we will briefly consider the third goal, which is the generation of new tools for the precise calculation of oceanic transports through arbitrary oceanic straits and sections. The evaluation of transports based on reanalyses and model data is complicated due to the employed modelling grids as the distortion of curvilinear grids towards their displaced north poles causes that velocities do not point in the direction of the true north/east anymore. Interpolation of the vector components onto regular latitude-longitude grids would require the exact angles between the models' native grid lines and lines of constant latitude and longitude in order to rotate the velocities into the directions of a regular latitude-longitude grid. As long as those grid angles are not standard output for models and reanalyses, an accurate interpolation and consecutive transport calculation is not possible. However, even if the grid angles and other model configurations are archived correctly, ensuring the conservation characteristics of the interpolated fluxes is still challenging, stressing the importance of calculating transports only on the models' native grids. The unfeasability of interpolation and the large number of different grid specifics in existing models and reanalyses makes the exact calculation of transports for a large ensemble of data a challenging and time consuming task. Given the importance of oceanic transports in the climate system and changes therein, this raised the need for a versatile tool to calculate precise transports on various ocean modelling grids. This gap was closed through the development of the open-source Python package StraitFlux. The third paper gives an overview of principles underlying the transport calculation, it documents its software implementation and also provides a thorough validation of the employed methods as well as some application examples. It is shown that both employed methods, the Line Integration Method (LM) and the Vector Projection Method (VPM), produce reliable results across various ocean models and grid types even for complex grid configurations and at regions with strong distortions. StraitFlux considerably facilitates the analysis of ocean currents, net oceanic transports of heat and salinity as well as exchanges of water masses.

The flexibility to calculate transports at arbitrary sections allows for direct comparisons to observations and simplifies the validation of climate model simulations. It is possible to analyze mean states, variabilities as well as seasonal cycles of net oceanic transports and the ocean's vertical structure and flow patterns can be assessed through cross-sections of currents, temperature and salinity. StraitFlux should prove useful to study the Earth's climate system and better understand its dynamics and physical processes, in the past as well as for future projections. Adhering to the FAIR (Findability, Accessibility, Interoperability, Reusability) principles (Wilkinson et al, 2016) make StraitFlux a very user-friendly implementation and its broad applicability makes it a useful tool for climate scientists, oceanographers and modelers and, apart from the usage for this thesis, its usefulness is shown for reanalyses at the Greenland-Scotland Ridge (Mayer et al, 2023) and for the Indonesian Throughflow Region (Fritz et al, 2023). StraitFlux allows the calculation of oceanic transports consistent with the discretization schemes of the respective models avoiding spurious artifacts that would be caused by interpolation and therefore allowing for a fair comparison between models as well as to observations and enabling holistic budget assessments where particularly precise transports are needed.

StraitFlux was an important asset for the second goal of this thesis, the assessment ad validation of historical simulations of 39 CMIP6 models and their performance in simulating the energy and water budgets of the Arctic. Long-term averages and seasonal cycles are compared against the observation-based basis gained through the first study (Winkelbauer et al, 2022) as well as observational estimates for the energy budget. Substantial biases are found for components of the mean state of the Arctic system as well as changes therein. The main results are summarized below.

Most models overestimate the long-term averaged surface freshwater fluxes into the Arctic Ocean, with the MMM being about 10% higher than our observational reference for both net precipitation and runoff. Trends of oceanic P-E and R are generally in fairly good agreement with observations and indicate an increase of freshwater input into the Arctic Ocean of 2%. While the timing of freshwater input through P-E concords with observations, the seasonal runoff cycle shows an early timing bias of one month. Those deviations are most likely related to biases in the timing of the models' snow melt and permafrost degradation. Kouki et al (2022) find that biases in snowmelt in turn are related to biases in precipitation and temperature and especially during the snow melt season in spring also to other factors like possible inaccuracies in model parametrizations related to hydrology and the surface energy budget. Additionally, as we compare observation-based river discharge at the river's mouth with grid-point averaged runoff from models, deviations in the runoff cycle might stem from the omission of river routing in some climate models, which may introduce delays, especially in larger catchments and for catchments with strong seasonal snowmelt (Gosling and Arnell, 2011).

While StraitFlux has facilitated more precise calculations of oceanic volume and ice transports, the output still reveals significant inter-model spreads and biases, with considerable inter-model variations concerning the long-term averages and annual cycles of the net transports. Some models feature spurious volume transport peaks during the summer months, which can be traced back to the implementation of non-linear free surface schemes, where sea ice melt water is physically dumped into the ocean and generates barotropic waves (Madec et al, 2019) during the melt season while freshwater is removed from the ocean during the cold season leading to net imports of water into the Arctic. This non-physical behaviour can be corrected using the models' sea ice volume changes. While this leads to smoother seasonal cycles for the net Arctic volume transports, inter-model spreads still remain large with some models featuring spurious signals. Volume transports are very sensitive to the exact horizontal and vertical cell extents in the models, therefore errors might be related to inaccuracies in the used ocean bathymetries. Especially for shallow or bathymetically more complicated straits, where currents tend to be accelerated in the proximity of the coast or ocean ground, inaccuracies in the bathymetries might cause larger errors. To calculate transports as accurately as possible cell thicknesses would be either needed both for the positions of temperatures and velocities, or if only thicknesses at cell centers are provided, the transformation equations to transform the thicknesses to the needed positions would be required. Unfortunately the needed data is not readily available for all CMIP6 models. This will be discussed further below.

The analysis of the energy budget further underscores the substantial biases and variabilities in the representation of the Arctic climate system by CMIP6 models. Surface energy fluxes are generally underestimated compared to observationally constrained reference estimates. Biases are smaller for fluxes at TOA, however still significant for some models. Those errors in the vertical energy fluxes are closely related to biases in the simulated sea ice extents. In summer, models with little sea ice simulate increased net radiation due to lower reflection of shortwave radiation, while in autumn and winter, they simulate heightened outgoing longwave radiation. Additionally, small errors in the global energy budget at the surface and TOA might be introduced due to inconsistencies in the models' energy conservation (Wild, 2020). This will be explored in more detail at the discussion of budget closure below.

Similar to the water budget, also for the energy budget the largest biases and uncertainties are found in the ocean. Concerning the state of the Arctic Ocean, past studies (e.g., Heuzé et al, 2023) have already shown that biases in Arctic deep waters are caused by the combination of the lack of shelf overflows and therefore ventilation as well as erroneous oceanic transports. Using our ensemble of CMIP6 models we confirm substantial biases and high variabilities in the Arctic Ocean as well as for sea-ice. Generally, models tend to simulate the Atlantic water layer too deep and too thick and tend to overestimate the accumulation of energy within the ocean and sea ice over the past decades. Major inter-model and intra-model differences in OHC trends as well as spurious signals in the depth profiles of temperature trends further could be a sign that equilibrium was not yet reached by the models. Especially deep waters may take longer to equilibrate to the initial conditions, suggesting that longer spin-up times may be required.

Concerning lateral energy transports all models agree on an inflow of heat into the Arctic and most models correctly simulate the timing of inflow peaks, however inter-model spreads again are substantial and most models systematically underestimate the magnitude of the observed heat influx. Particularly a subset of six CMIP6 models (all using the same ocean model component) feature unrealistically low transports, mainly caused by heat flux biases through the Barents Sea Opening (BSO) due to temperature biases and insufficiently simulated Norwegian Coastal Currents. The strength of OHT impacts oceanic temperatures and trends therein as well as sea ice extent, underscoring the robust relationship between sea ice cover and oceanic conditions. Further impacts are also seen on atmospheric energy fluxes, emphasizing the critical impact of oceanic transports on the broader Arctic climate and changes therein. Additionally, the spatial resolution of models affects the accuracy of simulated heat transports with higher resolution models aligning closer to observational estimates, especially in the Davis Strait.

Thus, our analyses revealed large inter-model spreads in multiple energy and water budget components. Model variability can stem from several factors, including the use of a single realization per model. This is known to introduce sampling errors due to the climate system's chaotic nature. However, comparisons to intra-model spreads suggest that differences within models are generally minor relative to those between different models, implying that discrepancies for most components when examining long-term averages are likely representing true systematic biases inherent in the models. However, when considering trends the short time period of 22 years makes the estimates prone to sampling errors through variabilities. The time period is, however, limited by the availability of observational time series and the need to avoid spin-up effects in model simulations. Additionally, errors might be introduced by missing processes or different treatment of processes in the models, as seen for instance for seasonal cycles of volume transports when including a non-linear free surface scheme (NLFS) scheme. Additional uncertainties due to methodological choices, such as interpolation errors, have been reduced to a minimum by using native grid files for calculating the energy and water variables.

We further assessed to what extent climate models satisfy the energy and water budget equations. Non-closure of global budgets may contribute to spurious long-term changes/trends in state variables, the so-called model "drift". This may distort the estimate of forced changes in coupled climate simulations and lead to false interpretations. On a global scale previous studies (Irving et al, 2021; Hobbs et al, 2016) showed good conservation after drift removal concerning energy budgets, however bad conservation concerning ocean mass and atmospheric water budgets. On more regional scales we assessed budget closure over the Arctic domain, also including atmospheric and oceanic lateral transports of energy and water. Despite adopting more precise oceanic transport estimates, complete closure of the energy and water budgets for individual models was not possible. While especially energy budget residuals remained relatively small when compared to the inter-model variation and observational references, budget residuals for the water volume budget were significantly higher relative to individual fluxes for some models. Possible reasons for those discrepancies include technical inaccuracies in our analyses. Especially when calculating the net Arctic oceanic volume transports small inaccuracies might potentially cause major errors as large fluxes through the individual straits sum up to a relatively small net flux and due to the sensitivity of volume transports towards the exact ocean bathymetry. Additionally, methodological choices like the choice of reference temperatures for oceanic heat transports, can significantly skew results. This emphasizes the need for standardized methodologies recognized within the scientific community to better compare, validate, and interpret the vast array of climate data generated by different models. Furthermore, while we consider the used budget equations as complete as possible, some smaller terms are still missing, like for instance the tendency of sea ice sensible heat, the oceanic transports' temporal (sub-monthly) eddy component or generally transports through the narrow passages of Hecla and Fury Strait. Also, deficiencies within the models themselves, such as issues with model coupling, physical processes, and numerical schemes, might lead to non-closure and it is possible that small terms like for instance mass leak increments or numerical diffusion are not provided by the CMIP6 model output. Addressing those imbalances through comprehensive evaluations could help in identifying and reducing biases and therefore lead to more accurate regional process representations and more balanced budgets.

Furthermore, refined analytical strategies, like for instance an adjustment towards a weighted or constrained averaging approach could help in ensuring more robust assessments of climate projections and simulations and better bridge the gap between model outputs and observations. In our study all multi-model averages were computed using an arithmetic mean over all available models without any kind of model weighting. However, as models vary greatly in performance and are oftentimes not entirely independent of one another, which might lead to skewed and biased results (Knutti et al, 2017; Brunner et al, 2020). Therefore, to mitigate biases, uncertainties and discrepancies between models and provide a more balanced representation of the overall model ensemble a more sophisticated method, such as model weighting or constraining should be used. This will be discussed further in the next section.

A significant challenge encountered was the incomplete availability of data from CMIP6 models. In our study this particularly refers to the oceanic cell thicknesses, which partially hampered accurate transport calculations and deeper analyses of oceanic processes. Generally, thicknesses are needed at the positions where the oceanic temperatures and also the velocities are discretisized, or at least, transformation equations should be provided to convert the thicknesses to the required locations. While most models provide thicknesses at least at one position in the grid cell, some models do not provide the required variable at all. Additionally, transformation of cell thicknesses for many models was complicated as the documentations of the models in the CMIP archive where in part insufficient, oftentimes challenging to locate, or in some instances even entirely absent. This lack of comprehensive documentation and data accessibility hinder our ability to fully leverage the CMIP6 datasets and therefore complicated our research process. Enhanced transparency and accessibility of both data and model documentation are a critical area of improvement in future climate modeling efforts. Fully adhering to the FAIR, as done for instance for the Sixth Assessment Report (AR6) of the Intergovernmental Panel on Climate Change (IPCC) (Iturbide et al, 2022), should be an important target for future climate model intercomparison efforts such as CMIP7. This would also contribute to the FAIRness of Straitflux.

4.2 Outlook

The results presented and the insights gained throughout this thesis motivate further ideas for future analyses and improvements. Two of them are presented here.

4.2.1 Model weighting

Generally, climate projections, and therefore our ability to predict climate change, contain uncertainties stemming from three main sources: (1) uncertainties arising from different scenarios of technological and societal development, (2) internal variability of the climate system and (3) model uncertainty due to model imperfections (Knutti et al, 2017). (1) can be accounted for by creating different possible projections dependent on different emission scenarios as done e.g. by the Shared Socioeconomic Pathways (SSPs) in CMIP6. (2) originates from the chaotic behavior of the climate system and can generally not be reduced on time scales longer than 5-10 years (Deser et al, 2012). Thus, to reduce uncertainties of future projections it is best to start with (3) which arises due to our incomplete comprehension of natural processes and their depiction in models. Multimodel ensembles are essential for assessing the range of uncertainty in climate simulations and projections, yet the traditional approach of using an arithmetic multimodel mean and assigning equal weights to each model's output, as it was done in this work using all available CMIP6 models, might not be ideal. Models oftentimes share code and model components and not all models perform equally well in terms of representing the observed climate and trends (e.g., Knutti et al, 2013). Using some sort of pre-selection, constraining approach or weighting algorithm instead of the concept of "model democracy" might help to mitigate uncertainties, biases and discrepancies between models and provide a more balanced representation of the overall model ensemble.

There are numerous different ways to sub-select, constrain or weight models. A method that incorporates weighting concerning the models' performance and independence was introduced by Knutti et al (2017); Lorenz et al (2018); Brunner et al (2020) and implemented as ClimWIP (`https://github.com/lukasbrunner/ClimWIP`). For each model of the considered ensemble a weight w_i is estimated considering model dependence (denominator) and model performance (numerator):

$$w_i = \frac{e^{-\frac{D_i^2}{\sigma_d^2}}}{1 + \sum_{j \neq i}^{M} e^{-\frac{S_{ij}^2}{\sigma_s^2}}} \tag{23}$$

where D_i is the distance of model i to observations, S_{ij} the distance between models i an j and M the total number of used model runs. The two parameters σ_D and σ_S determine how strongly model performance and model similarity are weighted. Models that are further away from the observed state get less weight (larger D_i) and models similar to one another get less weight (larger S_{ij}). Model performance (D_i) and model similarity (S_{ij}) can for instance be measured using skill scores like the root-mean-square error and need some choice of used diagnostic(s) (e.g. climatological means or trends) and variables(s) relevant for the projected target variable.

ClimWIP has already been used for different targets (e.g. sea ice decline or global mean temperature increase) using different sets of metrics. Continuing on the work from this dissertation we would like to test whether the usage of more process based quality metrics in the model weighting scheme may lead to further improvements in constraining projection uncertainties in the Arctic. Due to the high correlation between oceanic heat transports and

Arctic sea ice as well as ocean heat content tendencies and Arctic sea ice, we are currently testing whether the inclusion of those oceanic variables as diagnostics can further decrease the uncertainty of future projections of Arctic sea ice loss, without being overconfident. These diagnostics will be usable also for other target variables and periods when sea ice will have disappeared. Results will be directly compared to the work of Knutti et al (2017), who restricted their diagnostics to sea ice and surface temperature variables as well as Zhao et al (2022), who additionally use summer sea level pressure.

4.2.2 Advancement of StraitFlux

Unstructured grids: While StraitFlux works on curvilinear and regular grids, unstructured grids are not implemented yet. However, numerous oceanic models employ unstructured meshes, such as the Finite volumE Sea ice – Ocean Model (FESOM, Danilov et al, 2017; Wang et al, 2014) or the Icosahedral Nonhydrostatic Weather and Climate Model (ICON-O, Korn et al, 2022). Given the increasing adoption of unstructured meshes due to their ability to enhance resolution in specific geographical areas—thereby potentially capturing dynamics that would otherwise require higher global resolutions or nesting in structured grids—we plan to include transport calculations for various unstructured meshes in a future version of StraitFlux.

Same as for the curvilinear grids, also for unstructured grids vector components should not be interpolated as the models' satisfaction of continuity is violated and significant errors may be introduced. However, at least for models with vertically aligned meshes, the implementation into StraitFlux should be possible with limited effort. The general idea is similar to the second method presented by Sidorenko et al (2020) for the calculation of meridional overturning and barotropic streamfunctions on unstructured meshes and is outlined below and sketched in Fig. 2 for the FESOM2 (Danilov et al, 2017) grid. The discretization used by FESOM2 can be seen in Figure 1 in Sidorenko et al (2020). Horizontal velocities are placed on the centroids of the triangles and scalars at the vertices. Horizontal fluxes are calculated between scalar control volumes (=hexagons around the cell vertices), thus along the mesh edges (blue arrows in Fig. 2). Following equation 1 in Sidorenko et al (2020) volume transports F through the faces of the scalar control volumes (= yellow lines in Fig. 2) for a vertical layer k and the mesh edge e can be calculated as:

$$F_e = \left[-(\vec{e}_z \times \vec{d}_{ec_1}) \cdot \vec{u}_{kc_1} h_{kc_1} + (\vec{e}_z \times \vec{d}_{ec_2}) \cdot \vec{u}_{kc_2} h_{kc_2}\right] \tag{24}$$

where $\vec{e}_z$ is the vertical unit vector at the edge centeroids, d_{ec_1} and d_{ec_2} are the vectors pointing from edge centeroids to the cell centeroids (yellow lines), u_{kc_1} and u_{kc_2} are the horizontal velocities at the corresponding cell centeroids (green dots) and h_{kc_1} and h_{kc_2} are the layer thicknesses at the respective velocity points (cell centeroids). Transports are calculated for every edge centeroid along the yellow line in Fig. 2 and summed up to obtain total transports. Sidorenko et al (2020) view transports through lines of constant latitude and obtain the direction/sign of the transports by comparing the latitudes of the vertices with the latitudes of the reference line. However, for lines that do not follow a specific latitude the determination of the signs gets a bit more complicated and is still work in progress.

Depending on the discretization used by the models, the location of the velocity vectors differs and Eq. 24 has to be adapted accordingly. For instance for the ICON-O model normal velocities are placed at triangle edges making the transport calculations considerably simpler.

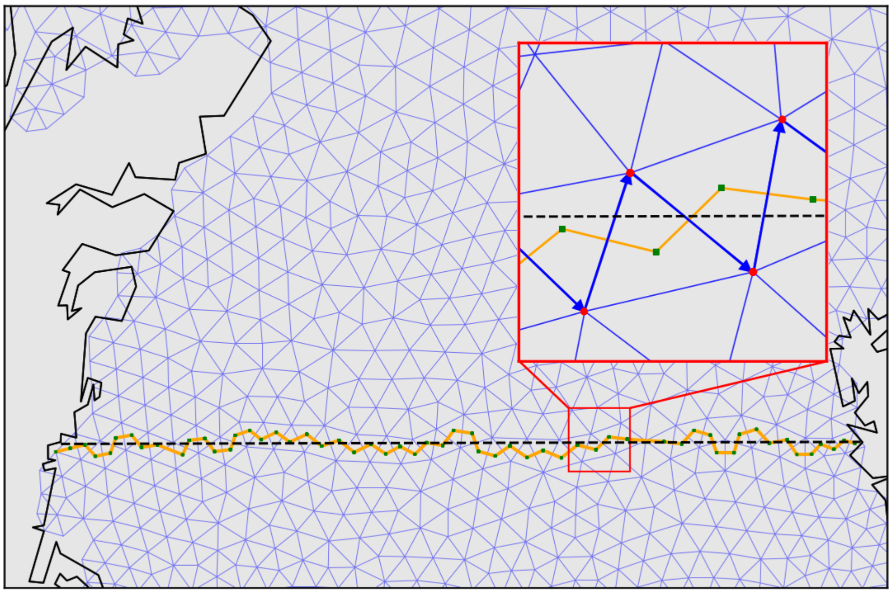

Figure 2: StraitFlux's Line Integration Method at Fram Strait on the unstructured FESOM2 grid. Thin blue lines show triangular grid cells, black dashed line shows the position of Fram Strait. Cell centers are marked by green dots and cell vertices by red dots. Horizontal velocities are defined at triangle centroids. The broken line (yellow) formed by connecting cell centers to edge centeroids forms the closest line to the reference strait where exact fluxes are available in FESOM2. It will be used as intergation line for the Line Integration Method. Blue arrows show the fluxes associated with each line segment (see Sidorenko et al (2020), Equation 1 for the exact calculation).

While the theoretical framework is basically set, the advancement of the StraitFlux tools for unstructured grids has still to be done.

Helmholtz decomposition: Furthermore, in a future version of StraitFlux a third method for the transport calculations might be included based on the fundamentals of Helmholtz decomposition. For every vector we can formulate the following Poisson-equation:

$$\vec{v} = \vec{\nabla}^2 \vec{P} \quad \text{with} \quad \vec{\nabla}^2 \vec{P} = \vec{\nabla}(\vec{\nabla} \cdot \vec{P}) - \vec{\nabla} \times (\vec{\nabla} \times \vec{P}) \tag{25}$$

with $\phi = \vec{\nabla} \cdot \vec{P}$ and $\vec{\psi} = \vec{\nabla} \times \vec{P}$ we get:

$$\vec{v}_\phi = \nabla\phi + (-\nabla \times \vec{\psi}) \tag{26}$$

thus we can decompose every vector field into a divergent (curl-free) component given by the scalar potential ϕ and a rotational (divergence-free) component given by the vector potential $\vec{\psi}$. For a 2-dimensional (horizontal) flow, the curl-free component can be written as:

$$-\nabla \times \vec{\psi} = \vec{k} \times \nabla\psi \tag{27}$$

with ψ being the scalar streamfunction. The scalar fields of vorticity and streamfunction can then be remapped onto a regular grid without the violation of conversation properties. Eastward and northward velocity components could then be recovered using gradients.

AMOC: The Atlantic Meridional Overturning Circulation (AMOC) plays a key role in the regulation of the climate system by redistributing heat northward. Climate models predict a strong weakening of the AMOC (Weijer et al, 2020), which would have major impacts on the global climate. The driving mechanisms for this weakening are still under debate and are believed to be linked to changes in buoyancy forcing, wind forcing or a combination of the two (Madan et al, 2024). Similarly to net transports, the calculation of the AMOC for large model ensembles is also complicated due to the used modelling grids. In the present version of StraitFlux the meridional overturning streamfunction has to be calculated manually from the velocity cross-sections generated by the vector projection method. In the next version of StraitFlux, however, an additional function will automatically provide streamfunctions through the specified sections, thereby simplifying and accelerating the calculation of the meridional overturning at different positions in the ocean. This enhancement will facilitate further systematic research of the AMOC and changes therein, and might potentially help in the quantification of AMOC weakening.

ESMValTool: The integration of StraitFlux into ESMValTool (Righi et al, 2020) would additionally bring many benefits. ESMValTool is a community driven diagnostic and performance metrics tool that facilitates reproducible scientific analysis and assessment of Earth system models. As ESMValTool is widely used by researchers and is affiliated with large model intercomparison projects like CMIP, the integrating of StraitFlux into ESMValTool could greatly enhance its visibility and application across the climate modeling community. Additionally, the exposure to a community-driven development process, with collective insights and expertise from a large community of scientists and developers, could lead to significant enhancements in the functionality and reliability of the tools.

Straitflux was so far implemented only for ocean model output. It should be relatively straightforward to implement it also for atmospheric model output, which can also be on irregular grids. This would help to align atmospheric and oceanic transports as closely as possible.

4.3 Extracurricular activities

As seen, budget evaluations put high demands on the quality of model output and, as mentioned in the conclusion section, adhering to the FAIR principles is crucial for fostering transparency, efficiency, and innovation in climate modeling. It is important to engage in activities such as Fresh Eyes on CMIP, a new working group integrating the voices of early career researchers (ECRs) into CMIP, to put forward those demands on a high level so that they are met by all CMIP7 contributors. Together with a group of motivated ECRs I was able to join the documentation subgroup, where we mainly address challenges related to the experiment and model documentations and try to facilitate the communication of relevant information to the CMIP users. In the process we get the opportunity to collaborate with the broader CMIP community and communicate ideas directly to the CMIP7 task teams.

I am also a member of the Austrian Polar Reasearch Institute. Together with a voluntary group of ECRs we are in the process of expanding and better connecting young scientists working on polar science in Austria. In this context we want to integrate APRI ECR into the Association of Polar Early Career Scientists (APECS), as already done by many other nations. This would enable more interdisciplinary and also international collaborations, help young scientists find national and international career and education opportunities and promote better outreach to the public and possible future generations of polar researchers.

References

Bacon S, Aksenov Y, Fawcett S, Madec G (2015) Arctic mass, freshwater and heat fluxes: methods and modelled seasonal variability. Phil Trans R Soc A 373. `https://doi.org/10.1098/rsta.2014.0169`

Bartsch A, Pointner G, Nitze I, Efimova A, Jakober D, Ley S, Högström E, Grosse G, Schweitzer P (2021) Expanding infrastructure and growing anthropogenic impacts along arctic coasts. Environmental Research Letters 16(11):115013. `https://doi.org/10.1088/1748-9326/ac3176`

Blackport R, Screen JA (2020) Weakened evidence for mid-latitude impacts of arctic warming. Nature Climate Change 10(12):1065–1066. `https://doi.org/10.1038/s41558-020-00954-y`

Box J, Hubbard A, Bahr D, Colgan W, Fettweis X, Mankoff K, Wehrlé A, Noël B, Van den Broeke M, Wouters B, Bjørk A, Fausto R (2022) Greenland ice sheet climate disequilibrium and committed sea-level rise. Nature Climate Change 12. `https://doi.org/10.1038/s41558-022-01441-2`

Brunner L, Pendergrass AG, Lehner F, Merrifield AL, Lorenz R, Knutti R (2020) Reduced global warming from cmip6 projections when weighting models by performance and independence. Earth System Dynamics 11(4):995–1012. `https://doi.org/10.5194/esd-11-995-2020`

Cai Z, You Q, Wu F, Chen H, Chen D, Cohen J (2021) Arctic warming revealed by multiple cmip6 models: Evaluation of historical simulations and quantification of future projection uncertainties. Journal of Climate 34:4871–4892. `https://doi.org/10.1175/JCLI-D-20-0791.1`

Cao Y, Liang S, Sun L, Liu J, Cheng X, Wang D, Chen Y, Yu M, Feng K (2022) Trans-arctic shipping routes expanding faster than the model projections. Global Environmental Change 73:102488. `https://doi.org/10.1016/j.gloenvcha.2022.102488`

Catling D, Zahnle K (2009) The planetary air leak. Scientific American 300:36–43. `https://doi.org/10.1038/scientificamerican0509-36`

Cheng L, von Schuckmann K, Abraham JP, Trenberth KE, Mann ME, Zanna L, England MH, Zika JD, Fasullo JT, Yu Y, Pan Y, Zhu J, Newsom ER, Bronselaer B, Lin X (2022) Past and future ocean warming. Nature Reviews Earth & Environment 3(11):776–794. `https://doi.org/10.1038/s43017-022-00345-1`

Chiodo G, Haimberger L (2010) Interannual changes in mass consistent energy budgets from era-interim and satellite data. Journal of Geophysical Research: Atmospheres 115(D2). `https://doi.org/10.1029/2009JD012049`

Coumou D, Di Capua G, Vavrus S, Wang L, Wang SY (2018) The influence of arctic amplification on mid-latitude summer circulation. Nature Communications 9. `https://doi.org/10.1038/s41467-018-05256-8`

Danilov S, Sidorenko D, Wang Q, Jung T (2017) The finite-volume sea ice–ocean model (fesom2). Geoscientific Model Development 10(2):765–789. `https://doi.org/10.5194/gmd-10-765-2017`

Deser C, Phillips A, Bourdette V, Teng H (2012) Uncertainty in climate change projections: the role of internal variability. Climate Dynamics 38(3):527–546. `https://doi.org/10.1007/s00382-010-0977-x`

Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, Taylor KE (2016) Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. Geoscientific Model Development 9(5):1937–1958. `https://doi.org/10.5194/gmd-9-1937-2016`

Francis JA, Vavrus SJ (2012) Evidence linking arctic amplification to extreme weather in midlatitudes. Geophysical Research Letters 39(6). `https://doi.org/10.1029/2012GL051000`

Fritz M, Mayer M, Haimberger L, Winkelbauer S (2023) Assessment of indonesian throughflow transports from ocean reanalyses with mooring-based observations. Ocean Science 19(4):1203–1223. `https://doi.org/10.5194/os-19-1203-2023`

Gosling SN, Arnell NW (2011) Simulating current global river runoff with a global hydrological model: model revisions, validation, and sensitivity analysis. Hydrological Processes 25(7):1129–1145. `https://doi.org/10.1002/hyp.7727`

Haine TW, Curry B, Gerdes R, Hansen E, Karcher M, Lee C, Rudels B, Spreen G, de Steur L, Stewart KD, Woodgate R (2015) Arctic freshwater export: Status, mechanisms, and prospects. Global and Planetary Change 125:13–35. `https://doi.org/10.1016/j.gloplacha.2014.11.013`

Haines K, Stepanov VN, Valdivieso M, Zuo H (2013) Atlantic meridional heat transports in two ocean reanalyses evaluated against the rapid array. Geophysical Research Letters 40(2):343–348. `https://doi.org/10.1029/2012GL054581`

Heuzé C, Zanowski H, Karam S, Muilwijk M (2023) The deep arctic ocean and fram strait in cmip6 models. Journal of Climate 36(8):2551 – 2584. `https://doi.org/10.1175/JCLI-D-22-0194.1`

Hobbs W, Palmer MD, Monselesan D (2016) An energy conservation analysis of ocean drift in the cmip5 global coupled models. Journal of Climate 29(5):1639 – 1653. `https://doi.org/10.1175/JCLI-D-15-0477.1`

Irrgang AM, Lantuit H, Gordon RR, Piskor A, Manson GK (2019) Impacts of past and future coastal changes on the yukon coast — threats for cultural sites, infrastructure, and travel routes. Arctic Science 5(2):107–126. `https://doi.org/10.1139/as-2017-0041`

Irving D, Hobbs W, Church J, Zika J (2021) A mass and energy conservation analysis of drift in the cmip6 ensemble. Journal of Climate 34(8):3157 – 3170. `https://doi.org/10.1175/JCLI-D-20-0281.1`

Iturbide M, Fernández J, Gutiérrez J, Pirani A, Huard D, Al Khourdajie A, Medina J, Bedia J, Casanueva A, Cimadevilla E, Cofiño A, De Felice M, Diez-Sierra J, García-Díez M, Goldie J, Herrera D, Herrera García S, Manzanas R, Milovac J, Yelekçi 2022) Implementation of fair principles in the ipcc: The wgi ar6 atlas repository. Scientific Data 9(1). `https://doi.org/10.1038/s41597-022-01739-y`

Knutti R (2008) Should we believe model predictions of future climate change? Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 366(1885):4647–4664. `https://doi.org/10.1098/rsta.2008.0169`

Knutti R, Masson D, Gettelman A (2013) Climate model genealogy: Generation cmip5 and how we got there. Geophysical Research Letters 40(6):1194–1199. `https://doi.org/10.1002/grl.50256`

Knutti R, Sedláček J, Sanderson BM, Lorenz R, Fischer EM, Eyring V (2017) A climate model projection weighting scheme accounting for performance and interdependence. Geophysical Research Letters 44(4):1909–1918. `https://doi.org/10.1002/2016GL072012`

Korn P, Brüggemann N, Jungclaus JH, Lorenz SJ, Gutjahr O, Haak H, Linardakis L, Mehlmann C, Mikolajewicz U, Notz D, Putrasahan DA, Singh V, von Storch JS, Zhu X, Marotzke J (2022) Icon-o: The ocean component of the icon earth system model—global simulation characteristics and local telescoping capability. Journal of Advances in Modeling Earth Systems 14(10):e2021MS002952. `https://doi.org/10.1029/2021MS002952`, e2021MS002952 2021MS002952

Kouki K, Räisänen P, Luojus K, Luomaranta A, Riihelä A (2022) Evaluation of northern hemisphere snow water equivalent in cmip6 models during 1982–2014. The Cryosphere 16(3):1007–1030. `https://doi.org/10.5194/tc-16-1007-2022`

Kwok R (2018) Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). Environmental Research Letters 13(10):105005. `https://doi.org/10.1088/1748-9326/aae3ec`

Linke O, Quaas J, Baumer F, Becker S, Chylik J, Dahlke S, Ehrlich A, Handorf D, Jacobi C, Kalesse-Los H, Lelli L, Mehrdad S, Neggers RAJ, Riebold J, Saavedra Garfias P, Schnierstein N, Shupe MD, Smith C, Spreen G, Verneuil B, Vinjamuri KS, Vountas M, Wendisch M (2023) Constraints on simulated past arctic amplification and lapse rate feedback from observations. Atmospheric Chemistry and Physics 23(17):9963–9992. `https://doi.org/10.5194/acp-23-9963-2023`

Lorenz E (1967) Nature and theory of the general circulation of the atmosphere. WMO 161pp

Lorenz R, Herger N, Sedláček J, Eyring V, Fischer EM, Knutti R (2018) Prospects and caveats of weighting climate models for summer maximum temperature projections over north america. Journal of Geophysical Research: Atmospheres 123(9):4509–4526. `https://doi.org/10.1029/2017JD027992`

L'Ecuyer TS, Beaudoing HK, Rodell M, Olson W, Lin B, Kato S, Clayson CA, Wood E, Sheffield J, Adler R, Huffman G, Bosilovich M, Gu G, Robertson F, Houser PR, Chambers D, Famiglietti JS, Fetzer E, Liu WT, Gao X, Schlosser CA, Clark E, Lettenmaier DP, Hilburn K (2015) The observed state of the energy budget in the early twenty-first century. Journal of Climate 28(21):8319 – 8346. `https://doi.org/10.1175/JCLI-D-14-00556.1`

Madan G, Gjermundsen A, Iversen SC, LaCasce JH (2024) The weakening amoc under extreme climate change. Climate Dynamics 62(2):1291–1309. `https://doi.org/10.1007/s00382-023-06957-7`

Madec G, Team NS, Team NS (2019) NEMO ocean engine. `https://doi.org/10.5281/zenodo.1464816`

Mayer M, Haimberger L, Edwards JM, Hyder P (2017) Toward consistent diagnostics of the coupled atmosphere and ocean energy budgets. Journal of Climate 30(22):9225 – 9246. `https://doi.org/10.1175/JCLI-D-17-0137.1`

Mayer M, Alonso Balmaseda M, Haimberger L (2018) Unprecedented 2015/2016 indo-pacific heat transfer speeds up tropical pacific heat recharge. Geophysical Research Letters 45(7):3274–3284. `https://doi.org/10.1002/2018GL077106`

Mayer M, Tietsche S, Haimberger L, Tsubouchi T, Mayer J, Zuo H (2019) An improved estimate of the coupled arctic energy budget. Journal of Climate 32(22):7915 – 7934. `https://doi.org/10.1175/JCLI-D-19-0233.1`

Mayer M, Vidar SL, Kjell AM, von Schuckmann K, Monier M, Greiner E (2021) Ocean heat content in the highnorth. Journal of Operational Oceanography 14(Suppl.1):S17–S23. `https://doi.org/10.1080/1755876X.2021.1946240`

Mayer M, Tsubouchi T, von Schuckmann K, Seitner V, Winkelbauer S, Haimberger L (2022) Atmospheric and oceanic contributions to observed nordic seas and arctic ocean heat content variations 1993-2020. Journal of Operational Oceanography 15(Suppl.1):S20–S28. `https://doi.org/10.1080/1755876X.2022.2095169`

Mayer M, Tsubouchi T, Winkelbauer S, Larsen KMH, Berx B, Macrander A, Iovino D, Jónsson S, Renshaw R (2023) Recent variations in oceanic transports across the greenland–scotland ridge. State of the Planet 1-osr7:14. `https://doi.org/10.5194/sp-1-osr7-14-2023`

Mayer M, Kato S, Bosilovich M, Bechtold P, Mayer J, Schröder M, Behrangi A, Wild M, Kobayashi S, Li Z, L'Ecuyer T (2024) Assessment of atmospheric and surface energy budgets using observation-based data products. Surveys in Geophysics 1573-0956. `https://doi.org/10.1007/s10712-024-09827-x`

Melia N, Haines K, Hawkins E (2017) Implications from opening arctic sea routes. Report, University of Reading, Gov.uk

Meredith M, Sommerkorn M, Cassotta S, Derksen C, Ekaykin A, Hollowed A, Kofinas G, Mackintosh A, Melbourne-Thomas J, Muelbert M, Ottersen G, Pritchard H, Schuur E (2022) Polar regions. In: The Ocean and Cryosphere in a Changing Climate: Special Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, p 203–320, `https://doi.org/10.1017/9781009157964.005`

Mouginot J, Rignot E, Bjørk A, van den Broeke M, Millan R, Morlighem M, Noël B, Scheuchl B, Wood M (2019) Forty-six years of greenland ice sheet mass balance from 1972 to 2018. EARTH, ATMOSPHERIC, AND PLANETARY SCIENCES 116(19):9239–9244. `https://doi.org/10.1073/pnas.1904242116`

Murray RJ (1996) Explicit generation of orthogonal grids for ocean models. Journal of Computational Physics 126(2):251–273. `https://doi.org/doi.org/10.1006/jcph.1996.0136`

Oppenheimer M, Glavovic B, Hinkel J, Roderik v, Magnan A, Abd-Elgawad A, Rongshu C, Cifuentes M, Robert D, Ghosh T, Hay J, Ben M, Meyssignac B, Sebesvari Z, A.J. S, Dangendorf S, Frederikse T (2019) Sea Level Rise and Implications for Low Lying Islands, Coasts and Communities, pp 321–445. `https://doi.org/10.1017/9781009157964.006.`

Palmer M, Roberts C, Balmaseda M, Chang YS, Chepurin G, Ferry N, Yosuke F, Good S, Guinehut S, Haines K, Hernandez F, Köhl A, Lee T, Martin M, Masina S, Masuda S, Peterson K, Storto A, Toyoda T, Xue Y (2015) Ocean heat content variability and change in an ensemble of ocean reanalyses. Climate Dynamics 49. `https://doi.org/10.1007/s00382-015-2801-0`

Peixoto JP, Oort AH (1992) Physics of climate. Springer 520pp

Rantanen M, Karpechko AY, Lipponen A, Nordling K, Hyvärinen O, Ruosteenoja K, Vihma T, Laaksonen A (2022) The arctic has warmed nearly four times faster than the globe since 1979. Communications Earth & Environment 3(1). `https://doi.org/10.1038/s43247-022-00498-3`

Righi M, Andela B, Eyring V, Lauer A, Predoi V, Schlund M, Vegas-Regidor J, Bock L, Brötz B, de Mora L, Diblen F, Dreyer L, Drost N, Earnshaw P, Hassler B, Koldunov N, Little B, Loosveldt Tomas S, Zimmermann K (2020) Earth system model evaluation tool (esmvaltool) v2.0 – technical overview. Geoscientific Model Development 13(3):1179–1199. `https://doi.org/10.5194/gmd-13-1179-2020`

Rodell M, Beaudoing HK, L'Ecuyer TS, Olson WS, Famiglietti JS, Houser PR, Adler R, Bosilovich MG, Clayson CA, Chambers D, Clark E, Fetzer EJ, Gao X, Gu G, Hilburn K, Huffman GJ, Lettenmaier DP, Liu WT, Robertson FR, Schlosser CA, Sheffield J, Wood EF (2015) The observed state of the water cycle in the early twenty-first century. Journal of Climate 28(21):8289 – 8318. `https://doi.org/10.1175/JCLI-D-14-00555.1`

Rowland JC, Jones CE, Altmann G, Bryan R, Crosby BT, Hinzman LD, Kane DL, Lawrence DM, Mancino A, Marsh P, McNamara JP, Romanvosky VE, Toniolo H, Travis BJ, Trochim E, Wilson CJ, Geernaert GL (2010) Arctic landscapes in transition: Responses to thawing permafrost. Eos, Transactions American Geophysical Union 91(26):229–230. `https://doi.org/10.1029/2010EO260001`

Schauer U, Losch M (2019) "freshwater" in the ocean is not a useful parameter in climate research. Journal of Physical Oceanography 49(9):2309 – 2321. `https://doi.org/10.1175/JPO-D-19-0102.1`

von Schuckmann K, Cheng L, Palmer MD, Hansen J, Tassone C, Aich V, Adusumilli S, Beltrami H, Boyer T, Cuesta-Valero FJ, Desbruyères D, Domingues C, Garcia-Garcia A, Gentine P, Gilson J, Gorfer M, Haimberger L, Ishii M, Johnson GC, Killick R, King BA, Kirchengast G, Kolodziejczyk N, Lyman J, Marzeion B, Mayer M, Monier M, Monselesan DP, Purkey S, Roemmich D, Schweiger A, Seneviratne SI, Shepherd A, Slater DA, Steiner AK, Straneo F, Timmermans ML, Wijffels SE (2020) Heat stored in the earth system: where does the energy go? Earth System Science Data 12(3):2013–2041. `https://doi.org/10.5194/essd-12-2013-2020`

Screen JA, Simmonds I (2013) Exploring links between arctic amplification and mid-latitude weather. Geophysical Research Letters 40(5):959–964. `https://doi.org/10.1002/grl.50174`

Serreze MC, Barry RG (2014) The Arctic Climate System, 2nd edn. Cambridge Atmospheric and Space Science Series, Cambridge University Press, `https://doi.org/10.1017/CBO9781139583817`

Serreze MC, Barrett AP, Slater AG, Woodgate RA, Aagaard K, Lammers RB, Steele M, Moritz R, Meredith M, Lee CM (2006) The large-scale freshwater cycle of the arctic. Journal of Geophysical Research: Oceans 111(C11). `https://doi.org/10.1029/2005JC003424`

Serreze MC, Barrett AP, Stroeve JC, Kindig DN, Holland MM (2009) The emergence of surface-based arctic amplification. Cryosphere 3:11–19

Shiklomanov A, Déry S, Tretiakov M, Yang D, Magritsky D, Georgiadi A, Tang W (2021) River Freshwater Flux to the Arctic Ocean, pp 703–738. `https://doi.org/10.1007/978-3-030-50930-9_24`

Shiklomanov AI, Lammers RB, Vörösmarty CJ (2002) Widespread decline in hydrological monitoring threatens pan-arctic research. Eos, Transactions American Geophysical Union 83(2):13–17. `https://doi.org/10.1029/2002EO000007`

Shiklomanov I, Shiklomanov A (2003) Climatic change and the dynamics of river runoff into the arctic ocean. Water Resources 30:593–601. `https://doi.org/10.1023/B:WARE.0000007584.73692.ca`

Sidorenko D, Danilov S, Koldunov N, Scholz P, Wang Q (2020) Simple algorithms to compute meridional overturning and barotropic streamfunctions on unstructured meshes. Geoscientific Model Development 13(7):3337–3345. `https://doi.org/10.5194/gmd-13-3337-2020`

Storto A, Alvera-Azcárate A, Balmaseda MA, Barth A, Chevallier M, Counillon F, Domingues CM, Drevillon M, Drillet Y, Forget G, Garric G, Haines K, Hernandez F, Iovino D, Jackson LC, Lellouche JM, Masina S, Mayer M, Oke PR, Penny SG, Peterson KA, Yang C, Zuo H (2019) Ocean reanalyses: Recent advances and unsolved challenges. Frontiers in Marine Science 6. `https://doi.org/10.3389/fmars.2019.00418`

Stroeve J, Notz D (2018) Changing state of arctic sea ice across all seasons. Environmental Research Letters 13(10):103001. `https://doi.org/10.1088/1748-9326/aade56`

Tsubouchi T (2014) Assembling an arctic ocean boundary monitoring array. Data Science Journal 13:PDA72–PDA78. `https://doi.org/10.2481/dsj.IFPDA-12`

Tsubouchi T, Bacon S, Aksenov Y, Garabato ACN, Beszczynska-Möller A, Hansen E, de Steur L, Curry B, Lee CM (2018) The arctic ocean seasonal cycles of heat and freshwater fluxes: Observation-based inverse estimates. Journal of Physical Oceanography 48(9):2029 – 2055. `https://doi.org/10.1175/JPO-D-17-0239.1`

Uotila P, Goosse H, Haines K, Chevallier M, Barthélemy A, Bricaud C, Carton J, Fuckar N, Garric G, Iovino D, Kauker F, Korhonen M, Lien V, Marnela M, Massonnet F, Mignac D, Peterson K, Sadikni R, Shi L, Zhang Z (2019) An assessment of ten ocean reanalyses in the polar regions. Climate Dynamics 52. `https://doi.org/10.1007/s00382-018-4242-z`

Walsh JE (2014) Intensified warming of the arctic: Causes and impacts on middle latitudes. Global and Planetary Change 117:52–63. `https://doi.org/10.1016/j.gloplacha.2014.03.003`

Wang Q, Danilov S, Sidorenko D, Timmermann R, Wekerle C, Wang X, Jung T, Schröter J (2014) The finite element sea ice-ocean model (fesom) v.1.4: formulation of an ocean general circulation model. Geoscientific Model Development 7(2):663–693. `https://doi.org/10.5194/gmd-7-663-2014`

Weijer W, Cheng W, Garuba OA, Hu A, Nadiga BT (2020) Cmip6 models predict significant 21st century decline of the atlantic meridional overturning circulation. Geophysical Research Letters 47(12):e2019GL086075. `https://doi.org/10.1029/2019GL086075`, e2019GL086075 10.1029/2019GL086075

Wild M (2020) The global energy balance as represented in cmip6 climate models. Climate Dynamics 55:553–577. `https://doi.org/10.1007/s00382-020-05282-7`

Wilkinson M, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, Bonino da Silva Santos LO, Bourne P, Bouwman J, Brookes A, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo C, Finkers R, Mons B (2016) The fair guiding principles for scientific data management and stewardship. Scientific Data 3. `https://doi.org/10.1038/sdata.2016.18`

Winkelbauer S (2023) Straitflux. `https://doi.org//10.5281/zenodo.10053555`

Winkelbauer S, Mayer M, Seitner V, Zsótér E, Zuo H, Haimberger L (2022) Diagnostic evaluation of river discharge into the arctic ocean and its impact on oceanic volume transports. Hydrology and Earth System Sciences 26(2):279–304. `https://doi.org/10.5194/hess-26-279-2022`

Winkelbauer S, Mayer M, Haimberger L (2024a) Straitflux – precise computations of water strait fluxes on various modelling grids. EGUsphere 2024:1–26. `https://doi.org/10.5194/egusphere-2023-2883`

Winkelbauer S, Mayer M, Haimberger L (2024b) Validation of key arctic energy and water budget components in cmip6. Climate Dynamics (1432-0894). `https://doi.org/10.1007/s00382-024-07105-5`

Yumashev D, van Hussen K, Gille J, Whiteman G (2017) Towards a balanced view of arctic shipping: estimating economic impacts of emissions from increased traffic on the northern sea route. Climatic Change 143(1):143–155. `https://doi.org/10.1007/s10584-017-1980-6`

Zhao J, He S, Wang H, Li F (2022) Constraining cmip6 projections of an ice-free arctic using a weighting scheme. Earth's Future 10(10):e2022EF002708. `https://doi.org/10.1029/2022EF002708`, e2022EF002708 2022EF002708

Zhuang J, raphael dussin, Huard D, Bourgault P, Banihirwe A, Raynaud S, Malevich B, Schupfner M, Filipe, Levang S, Gauthier C, Jüling A, Almansi M, RichardScottOZ, RondeauG, Rasp S, Smith TJ, Stachelek J, Plough M, Pierre, Bell R, Caneill R, Li X (2023) pangeo-data/xesmf: v0.8.2. `https://doi.org/10.5281/zenodo.8356796`

Zsótér E, Cloke HL, Prudhomme C, Harrigan S, de Rosnay P, Munoz-Sabater J, Stephens E (2020) Trends in the glofas-era5 river discharge reanalysis (871). `https://doi.org/10.21957/p9jrh0xp`

Zsótér E, Harrigan S, Barnard C, Wetterhall F, Ferrario I, Mazzetti C, Alfieri L, Salamon P, Prudhomme C (2021) River discharge and related historical data from the global flood awareness system. v3.1. copernicus climate change service (c3s) climate data store (cds). URL `https://cds.climate.copernicus.eu/cdsapp#!/dataset/cems-glofas-historical`

Zuo H (2020) Bronco: Benefits of dynamically modelled river discharge input for ocean and coupled atmosphere-land-ocean systems. `https://marine.copernicus.eu/about/research-development-projects/2018-2020/bronco`, accessed: 2024-04-30

Zuo H, Mugford RI, Haines K, Smith GC (2011) Assimilation impacts on arctic ocean circulation, heat and freshwater budgets. Ocean Modelling 40(2):147–163. `https://doi.org/10.1016/j.ocemod.2011.08.008`

A Appendix

A.1 Acknowledgment of tools for writing

Online services like dictionaries and translators based on AI tools have been used to correct spelling and grammatical mistakes and to get suggestions to improve the clarity and flow of the text: ”deepl.com” was used for translations and improvements of the writing style, ”ChatGPT (3.5)” was used to improve the writing style and flow of the text and ”grammarly.com” was used to find grammar, spelling and punctuation mistakes.

A.2 Acknowledgments

First of all I would like to thank my colleagues at the Department of Meteorology and Geophysics, University of Vienna. Special thanks go to my supervisor Leo Haimberger and colleague Michael Mayer for their comprehensive support, valuable discussions and patience over the years. They helped me to design and develop the study presented here and, moreover, gave me the opportunity to pursue this doctorate by employing me within their research projects. Thanks go also to all colleagues and friends that I was able to make within APRI and at conferences and summer schools over the years.

I would also like to thank my friends and family for their unconditional moral support and for highlighting other things in life than work. Your support has sustained me through the highs and lows of this endeavor. Special thanks go to Simone, for her unwavering love, encouragement, and understanding throughout the challenging journey of completing this thesis. Thanks to my parents, Małgorzata and Stefan, for their encouragement and belief in me, and for making my studies possible financially. Thank you to my brother, Jan, for always supporting his little sister and all his ideas and help concerning the mathematical foundations of StraitFlux and thank you to Katja for our great lunch meetings and chats about our doctoral studies. And, of course, special thanks to my two cats, Sammy and Balu, for their 'distracting' support during my home office days.

A.3 List of abbreviations

AA Arctic Amplification

AET Atmospheric Energy Tendency

AHT Atmospheric Heat Transport

AMOC Atlantic Meridional Overturning Circulation

APECS ssociation of Polar Early Career Scientists

APRI Austrian Polar Reasearch Institute

ARCGATE Mooring-derived data of oceanic fluxes through the Arctic gateways

C3S Copernicus Climate Change Service

CAA Canadian Arctic Archipelago

CDS Climate Data Store

CGLORS ocean reanalyses from the Euro-Mediterranean Center on Climate Change

CMEMS Copernicus Marine Service

CMIP Coupled Model Intercomparison Project

ECMWF European Centre for Medium-Range Weather Forecasts

ECR Early Career Researcher

ERA5 ECMWF's fifth atmospheric reanalysis

ERA5-Land offline simulation of ERA5

ERA-Interim ECMWF interim reanalysis

ET evapotranspiration

F oceanic lateral volume transport

F^{surf} surface freshwater fluxes (precipitation, evaporation, runoff)

FESOM Finite volumE Sea ice – Ocean Model

FOAM ocean reanalyses from the UK Met Office

GHG greenhouse gases

GIS Greenlandic Ice Sheet

GloFAS Global Flood Awareness System

GLORYS ocean reanalyses from Mercator Ocean

GRACE Gravity Recovery and Climate Experiment

GREP Global ocean Reanalysis Ensemble Product

GS groundwater storage

GSR Greenland Scotland Ridge

HTESSEL Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land

ICON-O Icosahedral Nonhydrostatic Weather and Climate Model

IHT Ice Heat Transport

JRA55 Japanese 55-year Reanalysis

JRC Joint Research Centre

LISFLOOD model for river basin scale water balance and flood simulation

MET Melt ice Tendency

NEMO Nucleus for European Modelling of the Ocean

NLFS non-linear free surface scheme

NRMSE normalized root mean square error

OHC Ocean Heat Content

OHCT Ocean Heat Content Tendency

OHT Oceanic Heat Transport

OIT Oceanic Ice Transport

OVT Oceanic Volume Transport

ORAs Ocean Reanalyses

ORAS5 ECMWF's Ocean Reanalysis System 5

P precipitation

r Pearson's correlation coefficient

R runoff or river discharge

$\mathbf{Rad}_{TOA}$ energy flux at TOA

$\mathbf{S}_A$ atmospheric water storage

$\mathbf{S}_L$ land water storage

$\mathbf{S}_O$ ocean water storage

SD snow depth

SSP Shared Socioeconomic Pathway

SW soil water

SWVL Volumetric soil water per layer

TOA top of the atmosphere

VIWVD Vertical integral of divergence of moisture flux

WCRP World Climate Research Programme

A.4 List of variables and constants

$\vec{c}$ horizontal sea water velocity

$\mathbf{c}_a$ heat capacity of dry air

$\mathbf{c}_i$ specific heat of sea ice

$\mathbf{c}_p$ specific heat of seawater

$\mathbf{c}_v$ ..

d sea ice thickness

e total energy

g gravitational constant

k kinetic energy

L_f latent heat of fusion

L_v latent heat of vaporization

$\vec{n}$ normal vector

ϕ geopotential

Φ potential energy

p_s surface pressure

q specific humidity

ρ_w sea water density

ρ_w sea ice density

ρ_w density of freshwater

T_a air temperature

T_i temperature of sea ice

v horizontal wind vector

ω vertical wind component